Construction and phenotypic characterization of *Mycobacterium smegmatis* mutants deficient in the MutY DNA glycosylase

Farzanah Hassim

Dissertation submitted to the Faculty of Health Science, University of the Witwatersrand, Johannesburg, in fulfillment of the requirements for the degree of Master of Science in Medicine.

Johannesburg 2013

“Success is simply the outward evidence of the inward development”

~Author Unknown~

Declaration

I declare that this dissertation is my own, unaided work. It is being submitted for the degree of Master of Science in Medicine at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.

28 May 2013

Farzanah Hassim — Date

Presentations

Farzanah Hassim and Bhavna Gordhan
Construction and phenotypic characterization of *M. smegmatis* mutants deficient in the MutY DNA glycosylase
SASBMB/FASBMB Congress 2012
The South African Society of Biochemistry and Molecular Biology
29 January – 1 February 2012
Champagne Sports Resort, Drakensberg, KwaZulu-Natal
Poster Presentation

Farzanah Hassim and Bhavna Gordhan
Construction and phenotypic characterization of *M. smegmatis* mutants deficient in the MutY DNA glycosylase
Wits Faculty of Health Sciences Research Day and PG Expo 2012
19 September 2012
University of Witwatersrand, Medical School, Johannesburg
Oral Presentation

Farzanah Hassim and Bhavna Gordhan
Construction and phenotypic characterization of *M. smegmatis* mutants deficient in the MutY DNA glycosylase
MRC Early Career Scientists Conference 2012
24 – 25 October 2012
South African Medical Research Council Centre, Cape Town
Oral Presentation

Abstract

Mycobacterium tuberculosis is a facultative pathogen that causes tuberculosis and it accounted for 1.4 million deaths worldwide in 2011. During infection reactive species are released by macrophages as part of the hosts' immune response, causing damage to the pathogen's DNA. Since mycobacterial genomes are G+C rich the guanine base is more susceptible to oxidative damage which results in the formation of 7,8-dihydro-8-oxoguanine (8-oxoG) lesions which are subsequently repaired by the Fpg/Nei family of DNA glycosylases of the base excision repair (BER) pathway. Mycobacteria possess four copies of the Fpg/Nei glycosylases with the Fpg homologue displaying an overlapping role with MutM in the GO system. Loss of the Fpg/Nei DNA glycosylases leads to mispaired bases during replication which are subsequently repaired by the MutY DNA glycosylase. Previously, in our laboratory we showed that a deficiency of two Fpg/Nei glycosylases displayed no differences in survival of *M. smegmatis* mutants under oxidative stress and a 2-3 log difference was observed only when three or all four of the DNA glycosylases were inactivated. Surprisingly, there was no corresponding increase in mutator phenotype for all the combinatorial Fpg/Nei deficient mutants compared to the parental strain. A recent study further showed that a deficiency in *M. smegmatis* MutY glycosylase also displayed no notable susceptibility to oxidative stress or increase in mutagenesis. Since in *E.coli* a double *mutM* and *mutY* mutant displayed an increased mutator phenotype compared to the individual mutants, it was plausible to investigate the role of MutY in combination with the Fpg/Nei family of DNA glycosylases in mycobacteria to understand whether these DNA glycosylases display overlapping and/or compensatory functions in dealing with oxidative damage. Using homologous recombination the *mutY* gene was deleted in the parental and in the Fpg/Nei deficient mutant strains. Deletion of *mutY* together with the Fpg/Nei family of DNA glycosylases

displayed similar *in vitro* growth kinetics as the parental strain under normal culture conditions. However, under *in vitro* oxidative stress conditions the *mutY* deficiency especially in the absence of all four Fpg/Nei DNA glycosylases results in a greater reduction in survival of the mutants with a general increase in mutation rates. Consistent with these data was the significant increase in C → A and A → C mutations with the progressive loss of the DNA glycosylases as assessed by the spectral analysis of rifampicin resistant mutants. Taken together these data indicate that the mycobacterial MutY DNA glycosylase has antimutator properties and possibly has a more significant role in mycobacterial genome maintenance compared to the Fpg/Nei family of DNA glycosylases.

Acknowledgements

I would like to sincerely thank God the Almighty for all his favours bestowed upon me.

It has been an honor and privilege to be awarded funding from the National Research Foundation (NRF) and the DST-NRF Centre of Excellence for Biomedical TB Research. The following funding is acknowledged to Dr. Bhavna Gordhan: The South African Tuberculosis and AIDS Training (SATBAT) program (1U2RTW007370/3), NHLS Research Trust Grant and the NRF incentive funding.

To Dr. Bavesh Kana, thank you for your constant encouragement and guidance. Your intelligence and leadership is invaluable to this department.

I would like to extend my sincere gratitude to my mentor and supervisor Dr. Bhavna Gordhan for the opportunity and unique experience. Being a part of this institution has led me on a path of self discovery and personal growth which will hopefully put me in good stead of becoming a great scientist one day. I appreciate your patient guidance, good judgment and useful critiques in this research project. Thank you for your valuable time and constructive suggestions.

Heartfelt thanks to all my colleagues (past and present) and support staff at the Centre of Excellence for Biomedical TB Research for your assistance and kindness. It has been a motivational and a rewarding journey. A special thanks to my good "friend" and colleague Nabiela Moolla, I will always treasure your friendship.

Lastly I would like to express my appreciation to my family and friends for their duaa's, support and encouragement.

Table of Contents

List of Figures

List of Tables

Abbreviations

A	Adenine
Amp	Ampicillin
amp^R	Ampicillin resistant
AP	Apurinic/Apyrimidinic
aph	Gene encoding kanamycin resistance
BER	Base excision repair
bp	Base pair
C	Cytosine
cfu	Colony forming units
CMR	Central microbial resources
CTAB	Cetyltrimethylammonium bromide
DCO	Double cross over
DIG	Dioxygen
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
DOTS	Directly observed therapy – short course
DS	Downstream
DTT	Dithiothreitol
dUTP	Deoxyuridine triphosphate
EDTA	Ethylenediaminetetra acetic acid
Fpg	Formamidopyrimidine DNA glycosylase
G	Guanine
G+C	Guanine and cytosine
H_2O	Water
HCl	Hydrochloric acid
HIV	Human immuno-deficiency virus
hr	Hour
Hyg	Hygromycin
hyg^R	Hygromycin resistance
kbp	Kilobase pair
Km	Kanamycin
km^R	Kanamycin resistant
LA	Luria-Bertani agar
lacZ	Gene encoding β-galactosidase activity
LB	Luria-Bertani broth
MDR-TB	Multi-drug resistant tuberculosis
$MgCl_2$	Magnesium chloride
min	Minute
ml	Milliliter
MMR	Mismatch repair
MOPS	3-(N-morpholino) propanesulfonic acid
MutY	Adenine DNA glycosylase
m-value	Number of mutational events

NaCl	Sodium chloride
NaOH	Sodium hydroxide
Nei	Endonuclease VIII
NER	Nucleotide excision repair
N_t	Final size of the population
Nth	Endonuclease III
OD_{600nm}	Optical density at 600 nm
Ori	Origin of replication
*ori*E	*E. coli* origin of replication
PCR	Polymerase chain reaction
rif^R	Rifampicin resistant
RNA	Ribonucleic acid
RNase	Ribonuclease
RNS	Reactive nitrogen species
ROS	Reactive oxygen species
rpm	Revolutions per minute
rpoB	Codes for the β subunit of RNA polymerase
RRDR	Rifampicin resistance determining region
RT-PCR	Reverse transcriptase-polymerase chain reaction
sacB	Gene encoding levansucrase
SCO	Single cross over
SDS	Sodium dodecyl sulphate
sec	Seconds
SOD	Superoxide dismutase
T	Thymine
TB	Tuberculosis
TBE	Tris, boric acid and EDTA
TDR-TB	Totally drug resistant tuberculosis
TE	Tris-EDTA
TIGR	The Institute for Genomic Research
Tris	Tris (hydroxymethyl) aminomethane
Tween	Polyoxyethylenesorbitanmonooleate
US	Upstream
WHO	World Health Organization
XDR-TB	Extensively drug resistant tuberculosis
X-gal	5-bromo-4-chloro-3-indolyl-β-galactoside

1 Introduction

1.1 Tuberculosis, treatment and drug resistance

Tuberculosis (TB) is an airborne infection caused by the facultative intracellular pathogen, *Mycobacterium tuberculosis* (MTB) and is a major public health problem worldwide despite the availability of several antibiotics. The World Health Organization (WHO) reported 1.4 million TB deaths in 2011 and a quarter of all deaths were people co-infected with human immunodeficiency virus (HIV) (WHO, 2011). DOTS (Directly Observed Therapy Short course) is an internationally developed program by the WHO for the control of TB. The currently accepted and most effective treatment is the six months regimen involving: two months of isoniazid, rifampicin, ethambutol and pyrazinamide (daily intake) followed by four months of isoniazid and rifampicin (daily or thrice a week) with cure rates of up to 95%. This treatment program is among the most cost effective however, the disadvantage is that due to the long treatment time there is patient non-adherence which leads to drug resistance; MDR (multidrug-resistant)-TB and XDR (extensively drug-resistant)-TB (WHO, 2011) thus making the disease difficult to treat.

MDR-TB strains are resistant to at least two first line drugs such as isoniazid and rifampicin, therefore patients are treated for 18-24 months with a combination of second line drugs which are more costly, less effective and more toxic to the patients thus having much lower cure rates of 50-70% (Koul *et al*., 2011). Second line drugs are made up of six classes: aminoglycosides (amikacin and kanamycin), polypeptides (capreomycin, viomycin and enviomycin), fluoroquinolones (ciprofloxacin, levofloxacin and moxifloxacin), thioamides (ethionamide and prothionamide), cycloserine (closerin) and terizidone thus all of these drugs may not be readily available due to their toxic properties. While XDR-TB strains are resistant to

isoniazid, rifampicin, any of the fluoroquinolones and one of the second-line injectable drugs (amikacin, kanamycin or capreomycin) resulting in high mortality rates (Meya and McAdam, 2007). Drug resistance surveys in Southern Africa have reported increased cases of MDR-TB. In Kwa-Zulu Natal a 98% mortality rate was reported for patients with XDR-TB, 50% of the cases were patients not previously diagnosed with TB, due to limited diagnostic coverage in rural areas (Gandhi *et al.,* 2006). There has also been a report on totally drug resistant-TB (TDR) in Iran. Of the 146 MDR-TB patients tested 15 strains were identified as TDR isolates and they were found to be resistant to first and second-line drugs (Velayati *et al.*, 2009). The TB/HIV co-infection further complicates the treatment due to the interaction between antiretrovirals and anti-TB drugs, therefore the need for shorter and simpler drug regimens that are safe and effective against both drug resistant strains and TB/HIV individuals. It is the development of drug resistance, HIV co-infection and TB/HIV co-treatment that has made the disease difficult to diagnose and treat.

The development of 11 promising compounds (new or repurposed) are currently in phase I, II and III clinical trials with the hope to develop better drugs with different modes of action that may shorten the treatment time, thus alleviating the current drug resistant problem (Lienhardt *et al.*, 2012). However, more is needed thus it is important to understand how MTB survives the harsh environment of the host as this knowledge will help in discovering potential drug targets leading to the development of new novel drugs and/or redesigning the current drugs.

1.2 Infection and the host immune response

TB is a disease of the lung which is transmitted by the inhalation of aerosols that contain *M. tuberculosis* cells from an infected individual. Upon entry in the lung alveoli the adaptive immune response is stimulated. Alveolar macrophages in the vicinity phagocytose the bacteria,

resulting in either the clearance of the initial infection or having a 5% risk of developing into active disease. While the remaining 95% of infected individuals develop latent TB infections which may reactivate at a later stage in immunocompromised individuals such as HIV co-infected individuals. The infected macrophages may lyse causing other macrophages, T lymphocytes and immune cells (such as natural killer and B cells) to come to the site of infection. The accumulation of these cells results in inflammation causing the formation of granulomas (Emile *et al.*, 1997). The central region of the granuloma may undergo necrosis and/or apoptosis thereby becoming hypoxic, while further necrosis of cells would lead to the dissemination of bacilli into the lung. The infected lung maybe exposed to exogenous (generated by the hosts immune system) and endogenous (such as the incomplete reduction of oxygen) oxidants resulting in a redox imbalance ultimately damaging the pathogen's DNA which may trigger mechanisms in the pathogen to enter a persistent or dormant state (Kumar *et al.*, 2011).

The virulence of *M. tuberculosis* is dependent on the antimicrobial activity of reactive species that are produced by macrophages during infection. Reactive oxygen species (ROS) (such as hydrogen peroxide, superoxide, oxygen and hydroxyl radicals) and reactive nitrogen species (RNS) (such as nitric oxide) produced by macrophages causes oxidative stress and plays a major role in the host's antimicrobial defense to control the infection by causing damage to the pathogen's DNA which induces mutations and eventually death if the damage is not repaired (Ehrt and Schnappinger, 2009).

Mycobacteria possess catalases, peroxidases and superoxide dismutases which are able to detoxify ROS and RNS generated by macrophages (Bartos *et al.*, 2004). Among the resistance mechanism is *katG*, a catalase-peroxidase that possesses dual activity as it has the ability to decompose hydrogen peroxide (into water and oxygen) and utilizes hydrogen peroxide to oxidize

various substrates (Bartos *et al*., 2004). While an alkyl hydroperoxide reductase (AphC) is able to detoxify organic peroxides (such as cumene peroxide) thus protecting the organism from oxidative damage (Springer *et al.*, 2001). Superoxide dismutases (SODs) are also resistant to the oxidative burst which is generated by macrophages by converting superoxide radicals to hydrogen peroxide and water thereby removing the toxic superoxide and preventing higher levels of hydrogen peroxide to be generated by other reactions (Teixeira *et al.*, 1998). Studies have shown that *M. tuberculosis* exports active form of superoxide dismutase which may protect the external cell wall structures of the pathogen by neutralizing oxygen molecules (Harth and Horwitz, 1999). Despite these anti-oxidative systems, mycobacteria do encounter damage to their DNA under these oxidative conditions which is repaired by several efficient DNA repair pathways to maintain genome stability.

1.3 DNA repair mechanisms and DNA glycosylases

The common types of oxidative base damage are pyrimidine dimers, single or double stranded breaks, mismatch base pairs and apurinic and apyrimidinic (AP) sites. Mycobacteria possess elaborate mechanisms that can detoxify ROS and remove oxidative base damage. Most organisms possess genes for the nucleotide excision repair (NER), the base excision repair (BER), the SOS response and the mismatch repair (MMR) (Dos Vultos *et al.*, 2009; Kurthkoti and Varshney, 2012). Bioinformatic analyses of mycobacterial genomes have identified genes for the NER, BER and SOS response but homologues for genes involved in MMR were absent in *M. leprae*, *M. tuberculosis* and *M. smegmatis* (Mizrahi and Andersen, 1998; Cole *et al.*, 1998; Cole *et al.*, 2001; Garnier *et al.*, 2003; The Institute for Genome Research [http://www.tigr.org]). In other organisms the MMR pathway plays a critically important role in mutation avoidance and

genome stabilization. Despite the absence of this repair pathway, hypermutability in mycobacteria was not observed *in vitro* under normal culture conditions, suggesting that mycobacteria relies on the NER and BER pathways for genome maintenance (Kurthkoti and Varshney, 2011; Kurthkoti and Varshney, 2012).

NER is an efficient, non-specific repair pathway that is able to repair a wide range of lesions that cause major distortions specifically induced by UV light (Dos Vultos *et al.*, 2009). It involves a multi-enzyme complex consisting of UvrA, UvrB and UvrC endonucleases which act in concert to recognize and excise damaged lesions by removing several nucleotides including the lesion and the repair is subsequently completed by polymerases and ligases. NER is however, unable to recognize and repair minor distortions in the DNA such as mismatched bases, methylated bases, oxidized guanines (8-oxoG) or base analogs.

These damaged bases are dealt with by the BER pathway, which is a multistep, multi-enzyme repair pathway in which DNA glycosylases are the primary enzymes which recognize and remove oxidatively damaged bases (Figure 1.1). DNA glycosylases cleave the N-glycosyl bond thus removing the damaged base and creates an AP site in which polymerases fill in a new base and ligases seal the gap. These glycosylases form part of the Fpg/Nei family and the Nth superfamily (endonulease III (*nth*)) of DNA glycosylases. The Fpg/Nei family is made up of Fpg/MutM/fapy and endonulease VIII (Nei) which are structurally related but differ in their primary substrate specificity (Krokan *et al*., 1997; Wallace *et al*., 2003). Fpg (also called MutM) glycosylases recognize oxidized purines (7,8-dihydro-8-oxoguanine (8-oxoG); 4,6-diamino-5-formamidopyrimidine (FapyA); 2,6-diamino-4-hydroxy-5-formamidopyrimidine (FapyG)), while Nei and Nth have the same substrate specificity for oxidized pyrimidines and FapyA (Krokan *et al*., 1997; Mizrahi and Andersen, 1998; Wallace *et al*., 2003).

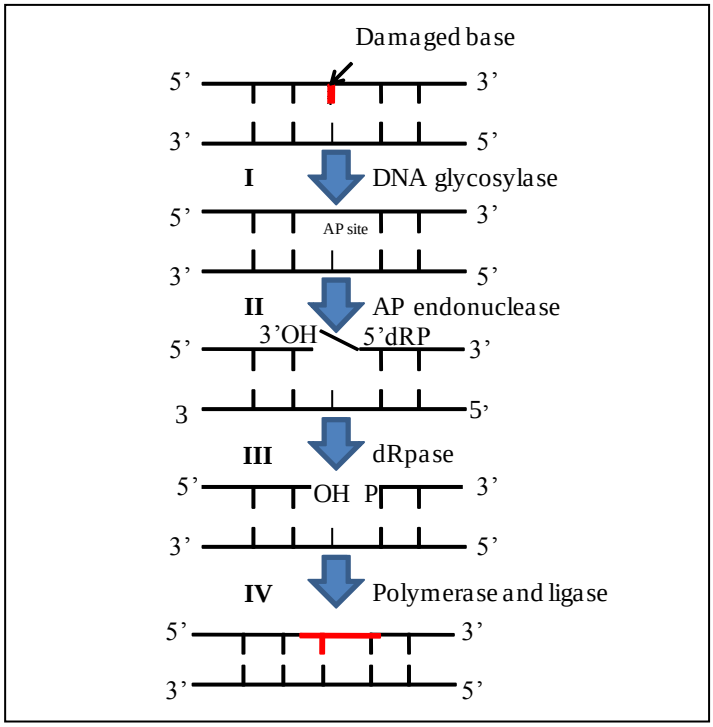


Figure 1.1: The base excision repair (BER) pathway. An efficient multistep multi-enzyme pathway that repairs minor distortions in the DNA by several DNA glycosylases. **I.** DNA glycosylase recognizes and excises the damaged base; **II.** The AP endonuclease removes the sugar-phosphate backbone; **III.** dRpase (deoxyribophosphodiesterase) converts the 3’terminal to a hydroxyl group; **IV.** DNA polymerase inserts a new nucleotide and DNA ligase seals the gap. Modified from Friedberg *et al*., 1995.

Bioinformatic analysis of the genome sequences of *M. tuberculosis* and *M. smegmatis* showed a unique duplication of these Fpg/Nei homologues. In both organisms two Fpg (I and II) and two Nei (I and II) DNA glycosylases were identified which have maintained the necessary domains for Fpg and Nei protein function with the exception of *fpg II* in *M. tuberculosis* which lacks the N-terminal domain, important for DNA binding and DNA glycosylase activity (Mizrahi and Anderson 1998; Cole *et al.*, 1998; Guo *et al.*, 2010). Therefore, *fpgII* in *M. tuberculosis* has been classified as a pseudogene. The mycobacterial Fpg DNA glycosylases also showed strong

resemblances to each other as well as to previously characterized Fpg family of glycosylases (Krokan *et al*., 1997; Wallace *et al*., 2003; Guo *et al*., 2010).

Recent work in our laboratory showed that the Fpg/Nei glycosylases are dispensable for growth since *M. smegmatis* mutants lacking either one, two, three or all four homologues of the Fpg/Nei family of DNA glycosylases revealed no significant differences in growth kinetics under normal culture conditions with no considerable increase in mutator phenotype for the various mutant strains (Goosens, 2008; Goosens *et al*., unpublished). However, preliminary data showed a 2-3 log difference in survival under oxidative damage as generated by hydrogen peroxide for mutants lacking three or more homologues of the Fpg/Nei family of DNA glycosylases (Goosens *et al.*, unpublished).

Mycobacterial genomes are G+C rich therefore, the guanine base is chemically more susceptible to 'guanine specific stress' caused by the interaction of the guanine bases with the reactive species resulting in the formation of 8-oxoG (GO lesions) (O'Sullivan *et al*., 2005). This damage is generally identified by Fpg/Nei glycosylases which initiates the BER pathway and efficiently repairs the damage (Krokan *et al*., 1997; Kurthkoti and Varshney, 2011). However, during replication the 8-oxoG lesion frequently mispairs with an adenine base and if not repaired and undergoes further replication it will induce G → T transversions in the genome. MutY another DNA glycosylase which forms part of the GO DNA repair system has the ability to reduce these mutagenic effects by removing the mispaired base and ensures that the correct base is inserted before replication (Michaels and Miller, 1992).

1.4 The GO system

The GO system is fully characterized in *Escherichia coli* and is suggested to be one of the more important error avoidance systems, composed of three proteins MutT, MutM (Fpg) and MutY (Michaels and Miller, 1992; Fowler *et al.*, 2003; Krokan *et al.*, 1997). Figure 1.2 illustrates the GO repair system.

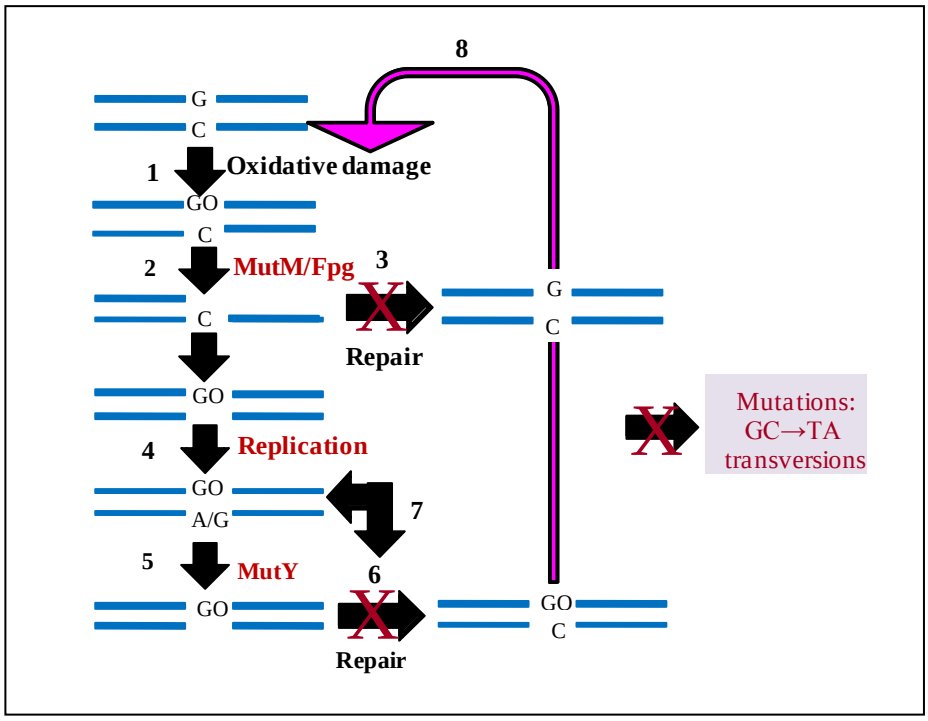


Figure 1.2: The GO system is composed of three proteins *mutT*, *mutM* and *mutY*. Oxidative damage can lead to the formation of 7, 8-dihydro-8-oxoguanine (referred to as 8-oxoG or GO lesions) in the DNA which is repaired by MutM and MutY, while MutT ensures that the nucleotide pool is free from these lesions that occur due to error prone DNA polymerases. **1.** Oxidative damage produces an oxidized base 8-oxoG; **2.** MutM/Fpg initiates the BER and removes the 8-oxoG; **3.** Repair of the lesion and the correct base is inserted; **4.** Replication occurs prior to repair by MutM/Fpg in which A or G is misincorporated; **5.** MutY recognizes and removes the mispaired base; **6.** Repair of the lesion and the correct base is inserted; **7.** If unrepaired by MutY, A or G is misincorporated again, **5.** MutY removes the mispaired base followed by **6.** repair; **8.** The cycle is repeated. If repair does not take place at step **3** and **6** it results in mutations. Modified from Kurthkoti *et al.*, 2010.

MutT is a hydrolase that coverts the 8-oxoGTP to 8-oxoGMP and pyrophosphate thus ensuring that the guanines in the nucleotide pool remains undamaged and prevents the

misincorporation of 8-oxoGTP in the DNA, in the absence of this gene AT → CG transversions results (Fowler *et al*., 2003; Dos Vultos *et al*., 2006).

Cloning and sequence analysis has shown that the *E. coli mutM* gene is identical to the *fpg* gene (Michaels *et al.*, 1991) thus having perhaps overlapping roles in both the BER pathway and the GO system. MutM functions as a formamidopyrimidine-DNA glycosylase that recognizes and excises purines and 8-oxoG lesions that are generated by oxidative damage and disruption of this gene leads to increased levels of GC → TA transversions (Cabrera *et al*., 1988; Michaels *et al.*, 1991).

The MutY DNA glycosylase helps to correct errors that result during DNA replication. This glycosylase recognizes and excises adenines or guanines paired with 8-oxoG or guanine. Deletion of *mutY* leads to an increase in GC → TA transversions (Nghiem *et al.*, 1988).

Strong biochemical and genetic evidence suggests that both *mutM* and *mutY* are involved in the same repair system (Michael and Miller, 1992). A subsequent study showed that a *mutM mutY* double deletion resulted in a higher mutation rate (GC → TA transversions) than for either individual deletion (Michaels *et al.*, 1992). This study further showed that when the MutM protein was overexpressed in a *mutY* deficient strain the mutation rate was reduced to wild type levels, suggesting that these glycosylases may have overlapping and/or compensatory role(s) under oxidative conditions. Further studies have shown that in a triple deletion mutant (*mutT mutM mutY*), *mutT* did not contribute in enhancing the effect of GC → TA transversions as compared to a double deletion mutant (*mutM mutY*) (Fowler *et al.*, 2003). Since in the absence of both these genes the organisms ability to protect itself from oxidative generated lesions is compromised it is suggestive that *mutM* and *mutY* are co-dependent. Nevertheless all three

proteins play an important role in error avoidance as well as display antimutator roles against GO lesions in the chromosomal DNA and in the nucleotide pool.

Unlike *E. coli* that has single copies of MutT, MutM and MutY, mycobacteria harbors four putative Nudix hydrolases which are annotated as MutT proteins (Dos Vultos *et al*., 2006), four copies of MutM (Fpg/Nei glycosylases) (Guo *et al.*, 2010; Jain *et al*., 2007; Olsen *et al*., 2009) and a single MutY homologue (Kurthkoti *et al*., 2010). A recent study by Kurthkoti *et al*., (2010) reported no significant increase in mutation rate for a *M. smegmatis mutY* deficient mutant strain indicating that there are other enzymes in the DNA repair pathway that may be compensating for this deficiency. Interestingly, *M. leprae* has reduced its genome size by retaining selective genes and has still managed to remain a major human pathogen. It has lost all MutT homologues, as well as the genes involved in the Fpg/Nei family of glycosylases (i.e. *fpgII*, *neiI* and *neiII*), however it has retained the *fpgI*, *nth* and *mutY* genes (Young and Robertson, 2001; Cole *et al*., 2001), suggesting that these genes are essential for DNA repair and genome maintenance and the other homologues were lost due to their redundant functions.

1.5 Structural and functional characteristics of MutY

The structural and functional characteristics of mycobacterial MutY are not well understood and much of the investigations rely on information available mainly from the *E. coli* counterpart as well as from studies in eukaryotes (humans and mice) and prokaryotes (*B. subtilis*, *D. radiodurans*, *H. pylori*, *P. aeruginosa*, *N. meningitides* and *N. gonorrhoeae*). Biochemical and protein studies in these organisms have revealed that MutY has similar substrates and binding preferences as the *E. coli* counterpart thus confirming the role of the MutY DNA

glycosylase in DNA repair (Slupska *et al.*, 1996; Debora *et al.*, 2011; Li and Lu, 2001; Eutsey *et al*., 2007; Oliver *et al*., 2002; Davidsen *et al.*, 2005).

Protein assays in *E. coli* have revealed the following: MutY is a 39kda adenine glycosylase that has N-glycosylase activity which functions to remove undamaged adenines that are mispaired with guanine, cytosine or 8oxoG and a 3' AP endonuclease activity that acts on the first phosphodiester bond on the 3' end of the AP site (Manuel *et al*., 1996). It consists of two domains namely the large NH_2 (N) - terminal domain (p26) and the COOH (C) - terminal domain (p13) (Manuel *et al*., 1996). The crystal structure of MutY is illustrated in Figure 1.3 below, the N - terminal domain is the catalytic domain containing the critical residues for catalysis which removes the mispaired adenine base in the DNA. While the C - terminal domain is responsible for recognizing the 8-oxoG and discriminates between A·G and A·8-oxoG mispairs.

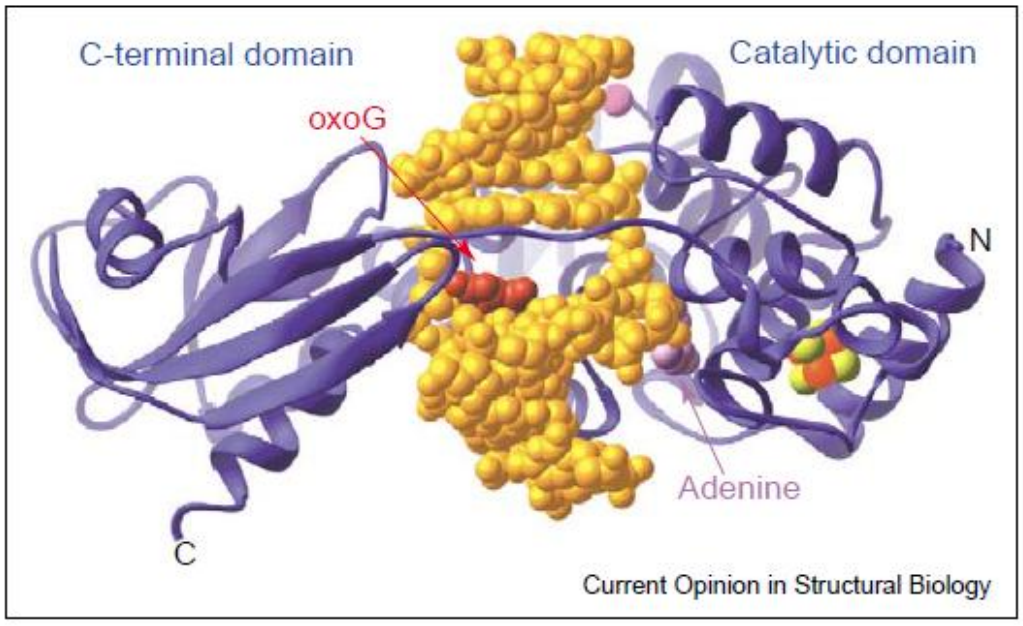


Figure 1.3: Crystal structure of MutY. Depicts MutY bound to DNA containing an adenine mispaired to 8-oxoG. The NH_2 (p26) and COOH (p13) terminal domains are depicted by N and C respectively. Gold – DNA, Red – 8-oxoG base, Purple – adenine base. (Taken from Fromme *et al*., 2004)

Extensive biochemical analysis was carried out by Guo *et al*., (2010) to characterize *M. tuberculosis* DNA glycosylases which revealed that mycobacterial DNA glycosylases have overlapping substrate preferences but are distinct compared to those of *E. coli* even though they share high sequence similarity. The *E. coli* Fpg and *M. tuberculosis* FpgI both recognize the 8-oxoG when paired with cytosine, thymine and guanine with the major difference being that the *M. tuberculosis* FpgI does not have strong specificity for 8-oxoG that is paired to an adenine. The reason for this is unknown even though this gene shares high sequence similarity in the two organisms (Guo *et al.*, 2010). The study further showed strong biochemical evidence that mycobacterial DNA glycosylases have substrate overlap with the *E. coli* counterparts (Guo *et al.*, 2010), thus suggesting the importance of these genes in the BER pathway and the GO system. There is evidence in *M. tuberculosis* and *M. smegmatis* that the MutY DNA glycosylase can recognize adenine, guanine and thymine but not cytosine that is base paired to 8-oxoG (possible due to it not being recognized as a mispair); while the *E. coli* MutY homologue only recognizes an 8-oxoG adenine mispair (Kurthkoti *et al.*, 2010). Collectively the Fpg/Nei and MutY glycosylases work in concert to remove the 8-oxoG bases as well as the bases mispaired with 8-oxoG hence preventing mutations and genome instability.

The availability of several Fpg/Nei deficient *M. smegmatis* mutants in our laboratory provided an opportunity to investigate the combinatorial role of Fpg/Nei and MutY DNA glycosylases to better understand the repair function of these glycosylases in the BER pathway and the GO system.

1.6 Aim

M. smegmatis is an important non-pathogenic model organism that was used as a surrogate host in this study due to its rapid growth, close structural and metabolic similarities to *M. tuberculosis.* In order to improve our understanding on the individual and collective role of MutY in the presence and absence of the Fpg/Nei family of DNA glycosylases in its ability to maintain the genome stability of mycobacteria, the following objectives were carried out:

1. Construction of a *mutY* deficient mutant in the parental *M. smegmatis* mc^2155 strain and in selected mutant strains lacking the Fpg/Nei family of DNA glycosylases.
2. Genotypic verification of the *mutY* deficient mutant strains.
3. Phenotypic characterization of the parental and the various MutY and Fpg/Nei mutant strains under normal and oxidative stress conditions.
4. Evaluation of mutation rates of the parental and the various MutY and Fpg/Nei mutant strains using the fluctuation assay.
5. Spectral analysis of rifampicin resistant mutants isolated from the parental and the various MutY and Fpg/Nei mutant strains to assess for increases in GC → TA transversions.

2 Materials and Methods

2.1 Bacterial strains, plasmids and their maintenance

E. coli strains were grown overnight with shaking (350 rpm, Horizontal Orbital Shaker Incubator, Lasec) at 37°C in Luria-Bertani broth (LB) [5g yeast extract, 10g tryptone and 10g NaCl] containing the appropriate antibiotics (100μg/ml ampicillin, 50μg/ml kanamycin and/or 50μg/ml hygromycin). While strains with large plasmids carrying knockout constructs (>8kb) were grown with shaking (100 rpm, Innova^{TM} 4000 Incubator Shaker, New Brunswick Scientific) at 30°C to prevent plasmid rearrangements. *E. coli* cells were streaked or spread on Luria-Bertani agar (LA) [5g yeast extract, 10g tryptone, 10g NaCl and 15g agar] plates supplemented with the appropriate antibiotics and/or 40μg/ml 5-bromo-4-chloro-3-indolyl-galactoside (X-gal) and/or 5% sucrose and grown at 37°C or at 30°C (in the case of large plasmids).

M. smegmatis strains were grown in Middelbrook 7H9 media (Difco) supplemented with 0.085% NaCl, 0.2% glucose, 0.2% glycerol, and 0.05% Tween 80 or on Middelbrook 7H10 media supplemented with 0.085% NaCl, 0.2% glucose and 0.5% glycerol. Where appropriate, 25 μg/ml kanamycin and/or 50μg/ml hygromycin and/or 200μg/ml rifampicin and/or 40μg/ml X-gal and/or 2-5% sucrose was included in the culture medium.

All bacterial strains and plasmids used are listed in Table 2.1 and 2.2 respectively. *E. coli* strains were stored in 66% glycerol (v/v) and *M. smegmatis* strains were stored in growth media at -80ºC. All purified plasmids were stored at -20ºC.

Table 2.1: Bacterial strains used in this study

Name	Characteristics	Source/Reference
***E. coli* DH5α**	*sup*E44 Δ*lacU169 hsd*R17 recA1 *end*A1 *gyr*A96 *thi*-1 *rel*A1	Hanahan, 1983
WT (*Mycobacterium smegmatis* mc²155)	*ept*-1, efficient plasmid transformation mutant of mc²6	Snapper *et al.*, 1990
Q1 (Δ*fpgI*Δ*fpgII*Δ*neiI*Δ*neiII*)	Derivative of mc²155 containing *fpgI fpgII neiI neiII* mutant alleles MSMEG_2419; MSMEG_5545; MSMEG_1756 and MSMEG_4683 respectively	Goosens *et al.*, unpublished
Q47 (Δ*neiI*Δ*neiII*Δ*fpgI*Δ*fpgII*)	Derivative of mc²155 containing *neiI neiII fpgI fpgII* mutant alleles MSMEG_1756; MSMEG_4683; MSMEG_2419 and MSMEG_5545 respectively	Goosens *et al.*, unpublished
Δ*Y* (Δ*mutY*)	Derivative of mc²155 containing a *mutY* mutant allele	This study
Q1Δ*Y* (Δ*fpgI*Δ*fpgII*Δ*neiI*Δ*neiII*Δ*mutY*)	Derivative of mc²155 containing *fpgI fpgII neiI neiII mutY* mutant alleles MSMEG_2419; MSMEG_5545; MSMEG_1756; MSMEG_4683 and MSMEG_6636 respectively	This study
Q47Δ*Y* (Δ*neiI*Δ*neiII*Δ*fpgI*Δ*fpgII*Δ*mutY*)	Derivative of mc²155 containing *neiI neiII fpgI fpgII mutY* mutant alleles MSMEG_1756; MSMEG_4683; MSMEG_2419; MSMEG_5545 and MSMEG_6636 respectively	This study
Δ*Y:Y* (Δ*mutY:mutY*)	Δ*Y* carrying a functional *mutY* integrated at the *attB* locus	This study
Q1Δ*Y:Y* (Δ*fpgI*Δ*fpgII*Δ*neiI*Δ*neiII*Δ*mutY:mutY*)	Q1Δ*Y* carrying a functional *mutY* integrated at the *attB* locus	This study
Q47Δ*Y:Y* (Δ*neiI*Δ*neiII*Δ*fpgI*Δ*fpgII*Δ*mutY*)	Q47Δ*Y* carrying a functional *mutY* integrated at the *attB* locus	This study

2.2 DNA isolation

2.2.1 Plasmid isolation from *E. coli*

2.2.1.1 Small-scale isolation (Miniprep)

Single colonies grown overnight to stationary phase (1.5 ml) were centrifuged at 13 000 rpm for 5 min, cells were re-suspended in 100 µl solution I (50mM glucose, 25mM Tris-HCl pH 8.0, 10mM EDTA) followed by addition of 200 µl of solution II (0.2M NaOH, 1% SDS, distilled H_2O). The tubes were inverted to mix and left at room temperature for 5 min. Next, 150 µl of solution III (3M potassium acetate, 11.5% glacial acetic acid, distilled H_20) was added to the

cells, shaken vigorously and cooled on ice for 5 min before centrifugation at 13 000 rpm for 5 min. The supernatant was transferred to a fresh Eppendorf tube containing 10mg/ml RNAse A and allowed to digest for 30 min at 37°C. The DNA was precipitated with 350 µl of isopropanol and centrifuged at 13 000 rpm for 10 min after which the pellet was washed with 70% ethanol, vacuum dried (SpeedVac, Savant) at 65°C for 10-15 min and re-suspended in 10 µl distilled H_2O.

2.2.1.2 Large-scale isolation (Maxiprep)

For large scale DNA extractions, *E. coli* cultures were grown in 50 ml LB overnight with shaking and the cells were harvested in a centrifuge (Allegra™ X-22R, Beckman Coulter) at 4500 rpm for 10 min. The extraction method was the same as described above with the exception that the solution volumes were increased by a factor of 10. The DNA was re-suspended in 300 µl distilled H_2O and purified by the addition of equal volumes of phenol: chloroform. The aqueous phase of the solution was added to an equal volume of chloroform and centrifuged at 13 000 rpm for 2 min. The plasmid DNA was re-precipitated with 1/10 volume 3M sodium acetate (pH 5.5), twice the volume of cold 100% ethanol and centrifugation at 13 000 rpm for 20 min. The DNA pellet was vacuum dried at 65°C for 10-15 min and re-suspended in 50-100 µl distilled H_2O.

For cloning purposes DNA was extracted using the Nucleobond plasmid DNA purification kit (Macherey-Nagel) according to the manufacturer's instructions to obtain good quality and higher yields of DNA.

Table 2.2: Cloning vectors used in this study

Name	Characteristics	Source/Reference
pGEM3 Zf(+)	*E. coli* cloning vector; amp^R; *lacZ*-alpha; *ori*E	Promega
p2NIL	*E. coli* cloning vector and mycobacterial suicide plasmid; km^R; *ori*E	Parish and Stoker, 2000
pGOAL17	Plasmid carrying *lacZ* and *sacB* genes as a *Pac*I cassette; amp^R; *ori*E	Parish and Stoker, 2000
pIJ963	Plasmid carrying *hyg* as a *Bgl*II cassette; amp^R; hyg^R; *ori*E	Blondelet-Rouault *et al*., 1997
pTWEETY	Integration vector carrying an integrase gene; km^R	Pham *et al*., 2007
pGΔYUS	pGEM cloning vector carrying the deleted *mutY* upstream region; amp^R; *lacZ*-alpha; *ori*E	This study
pGΔYDS	pGEM cloning vector carrying the deleted *mutY* downstream region; amp^R; *lacZ*-alpha; *ori*E	This study
p2ΔY	p2NIL suicide vector carrying the deleted *mutY* upstream and downstream region; km^R; *ori*E	This study
p2ΔY::hyg	p2ΔY carrying a *hyg* cassette from pIJ963; hyg^R; km^R; *ori*E	This study
p2ΔY::hyg::p17	p2ΔY::hyg carrying selectable and counter selectable markers from pGOAL17; hyg^R; km^R; *lac*Z; *sac*B; *ori*E	This study
p2ΔY::p17	p2ΔY::hyg::p17 removal of the *hyg* cassette to generate an unmarked knockout vector; km^R; *lac*Z; *sac*B; *ori*E	This study
pTWmutY	pTWEETY integrating vector carrying *mutY*; km^R	This study

amp^R – ampicilin resistance; *ori*E– origin of replication in *E. coli*; km^R- kanamycin resistance; hyg^R-hygromycin resistance.

2.2.2 Chromosomal DNA isolation from *M. smegmatis*

2.2.2.1 Small-scale chromosomal isolation

Single *M. smegmatis* colonies were re-suspended in 50 µl of distilled H_2O and boiled for 15 min at 65˚C after which 50 µl of chloroform was added and the contents mixed by inversion. The aqueous phase containing the DNA was isolated by centrifugation at 13 000 rpm for 5 min and 5-10 µl was used for PCR amplification.

2.2.2.2 Large-scale chromosomal isolation

Log-phase cells (5 ml) were harvested at 4500 rpm for 10 min and the pellets re-suspended in 500 µl TE buffer (10mM Tris-HCl pH 8.0, and 1mM EDTA). The suspension was heat killed at 65°C for 35 min then treated with 50 µl of 10mg/ml lysozyme for 1 hr at 37°C after which 70 µl of 10% SDS and 6 µl of 10mg/ml proteinase K was added and the tubes incubated at

65°C for a further hour. Next 100 µl of 5M NaCl was added and the contents mixed to which 80 µl of CTAB/NaCl solution (10% CTAB made in 0.7M NaCl) was added and incubated at 65ºC for 10 min. An equal volume of chloroform:isoamyl alcohol (24:1 v/v) was added and centrifuged at 13 000 rpm for 10 min. The aqueous phase was placed into a new Eppendorf tube and the DNA precipitated with 450 µl isopropanol, followed by centrifugation at 13 000 rpm for 20 min at 4°C. The pellet was washed with ice cold 70% ethanol and vacuum dried at 65°C for 10-15 min. The pellets were re-suspended in 20-40 µl of distilled H_2O.

2.3 DNA modifications

2.3.1 Restriction enzyme digestions (restriction analysis)

Restriction enzymes and their buffers were obtained from New England Biolabs, Inc., Roche Diagnostics or Fermentas and were used according to the manufacturer's instructions. For plasmid DNA 1-2µg of DNA was digested in a 10-20 µl total volume for 1 hr and 2-3µg of chromosomal DNA was digested in a 20 µl total volume overnight. The concentration of the DNA used was quantified using the UV spectroscopy (section 2.3.6). All digests were incubated at 37°C, heat inactivated at 65°C for 15 min and the DNA fragments were separated by agarose gel electrophoresis (section 2.3.4).

2.3.2 Phosphorylation of DNA

Blunt ended PCR (section 2.3.10) products were phosphorylated using polynucleotide kinase (Roche Diagnostics) which inserts phosphates to the 5' end, for successful ligation to a dephosphorylated vector. The reactions were set-up according to the manufacturer's instructions; the reactions were incubated at 37˚C for 15 min and the enzyme inactivated at 65˚C for 5 min.

2.3.3 Dephosphorylation of plasmid DNA

In order to prevent vector re-ligation and decrease the vector background during cloning, the 5'-phosphates were removed by adding antarctic alkaline phosphatase (New England Biolabs) to the linearized vector DNA and the reaction was carried out according to the manufacturer's instructions. The reaction was incubated at 37°C for 1 hr, before inactivation of the enzyme at 65°C for 20 min. The vector fragment was separated on an agarose gel and purified using a commercial DNA binding column (section 2.3.5).

2.3.4 Agarose gel electrophoresis

Agarose gels (0.8% - 2 %) were prepared in TAE buffer (1mM EDTA, 40mM Tris-acetic acid pH8.5) containing 2.5 µl of ethidium bromide (10mg/ml) in a total volume of 50 ml. Appropriate DNA molecular weight markers (Roche Diagnostics) shown in Figure 5.2.8 (Appendix B) were used to determine DNA fragment sizes; gels were electrophoresed at 80V and visualized using the gel fluorescence imaging system (BIORAD).

2.3.5 DNA fragment extraction and purification

The required DNA fragments were excised from agarose gels exposed to UV light and purified using the Nucleospin gel extraction kit (Macherey-Nagel) according to the manufacturer's instructions. Briefly, the appropriate sized fragments were excised and melted in the buffer provided and then loaded onto the column provided. The DNA was eluted with the elution buffer (provided) and the DNA quantified as described below (section 2.3.6).

2.3.6 DNA quantification

Plasmid and chromosomal DNA after extraction or gel purification was quantified on the NanoDrop ND-1000 Spectrophotometer and on agarose gels to estimate the concentration of the DNA in comparison to the DNA molecular weight marker III (Figure 5.2.8 Appendix B).

2.3.7 Ligation reactions

DNA ligations were performed using the T4 DNA Ligase kit (Roche Diagnostics) according to the manufacturer's instructions. Ligation reactions were incubated at room temperature for 1 hr – overnight and the reactions were inactivated at 65˚C for 20 min. The equation below was used to calculate the volume of insert required for 50ng of vector DNA for optimal transformation efficiency. The ligation reactions were subsequently transformed into competent *E. coli* DH5α cells (section 2.3.8.1).

$$\text{[insert DNA] (ng)} = \frac{\text{insert size (bp) X concentration of vector DNA}}{\text{vector size (bp)}}$$

2.3.8 Transformation of bacteria

2.3.8.1 *E. coli* DH5α chemically competent cells

Rubidium chloride competent cells were prepared as follows: 1 ml of overnight stationary phase culture was inoculated in 100 ml LB, grown to an OD_{600nm} of 0.48 - 0.55 and the cells were placed on ice for 15 min followed by harvesting at 4500 rpm for 5 min at 4°C. The pellets were re-suspended in 20 ml TfbI solution (30mM potassium acetate, 100mM rubidium chloride, 10mM calcium chloride, 50mM manganese chloride, and 15% v/v glycerol - pH 5.8) and placed on ice for a further 15 min. The cells were re-harvested at 4500 rpm at 4°C for 5 min, re-suspended in 2 ml TfbII solution (10mM MOPS, 75mM calcium chloride, 10mM rubidium chloride and 15% v/v glycerol-pH 6.5) and used immediately or were frozen in an ethanol bath and stored at -80°C until further use.

For transformations, 10 µl of a ligation reaction was added to 100 µl of thawed competent *E. coli* DH5α cells and incubated on ice for 20 min. The cells were heat-shocked at 42°C for 90 sec and returned to the ice for 2 min. The transformed cells were recovered with 800 µl LB at

37°C for 1 hr followed by centrifugation at 13 000 rpm for 5 min. The pellets were re-suspended in 100 µl LB and spread onto LA plates containing the appropriate antibiotics. Plates were scored after overnight incubation at 37°C.

2.3.8.2 *M. smegmatis* competent cells

A pre-culture was grown from freezer stocks in 7H9 media at 37°C to stationary phase which was then used to inoculate 100 ml of fresh media. The cells were grown at 37°C to an OD_{600nm} of 0.8. Cells were harvested at 4500 rpm for 10 min and the pellet re-suspended in 30 ml ice cold 10% glycerol. This process of centrifugation and washing of cell was repeated 3 times to ensure all salts were removed. The final suspension was used immediately to electroporate the various suicide plasmids.

400 µl of competent *M. smegmatis* cells were electroporated with 1µg of plasmid DNA in a 2mm electroporation cuvette and pulsed using the Gene Pulser Xcell (BIORAD) with the following conditions: 2500V, 25µF and 1000Ω. Electroporated cells were recovered immediately with 800 µl 2xTY followed by incubation at 37°C for 3 hrs. The cells were centrifuged at 13 000 rpm for 5 min, the supernatant was discarded and the pellet re-suspended in 100 µl of fresh media before spreading on media supplemented with the appropriate antibiotics. The plates were incubated at 37˚C for 3-5 days before scoring.

2.3.9 Sequencing

Cloning vectors and PCR products were sequenced either at Inqaba Biotech Ltd or the DNA Sequencing Facility at Stellenbosch University. EditSeq and SeqMan programs available with the DNASTAR software were used to analyze the sequenced data. The primers used for sequencing are listed in Table 5.1.2 (Appendix A).

2.3.10 Polymerase Chain Reaction (PCR)

PCR reactions were performed using FastStart Taq DNA polymerase (Roche Diagnostics) in a 25-50 µl total volume according to the manufacturer's instructions with the following cycling conditions: denaturation at 94°C for 5 min; 30 cycles of denaturation at 94°C for 60 sec; annealing for 30 sec at the specified annealing temperature; extension at 72°C for 60 sec and a final extension at 72°C for 7 min. Once the PCR conditions were optimized with Taq DNA polymerase, for cloning purposes the PCR was repeated with a high-fidelity DNA polymerase, Phusion (Finnzymes and Roche Diagnostics). The PCR was performed according to the manufacturer's instructions with the following cycling conditions: denaturation at 98°C for 30 sec; followed by 25 - 35 cycles of denaturation at 98°C for 10 sec; annealing for 30 sec at a higher or same temperature as that used with Taq DNA polymerase, extension at 72°C for 30-90 sec; and a final extension at 72°C for 7 min. The ThermoCycler (BIORAD) was used for all PCR reactions and the primers used are detailed in Table 5.1.1 (Appendix A).

2.3.11 Southern blot analyses

2.3.11.1 Electroblotting

2-3µg of genomic DNA from *M. smegmatis* wild type and mutant strains was digested with the appropriate enzymes and incubated at 37˚C overnight. The fragments were separated by gel electrophoresis, treated with depurination solution I (0.2M HCl) with continuous shaking for 15 min, after which the gels were rinsed twice with distilled H_2O followed by treatment with denaturing solution II (0.5M NaOH and 1.5M NaCl) with continuous shaking for 15 min.

The DNA from the agarose gels was transferred to nitrocellulose membranes (Hybond-N nitrocellulose membrane, Amersham) using the cassette clamps to create a sandwich that consisted of sponge, filter paper, gel, membrane, filter paper and sponge in the respective order.

The cassette was then placed in the OmniPAGE Electroblotting Unit (Cleaver Scientific Ltd CS-300V) equilibrated with 1X TBE buffer (89mM Tris-Borate, 2mM EDTA, pH8.3) and the DNA transferred at 600A, 130V at 4°C for 2 hrs. Thereafter, the membranes were removed from the cassette clamps and crosslinked at 120 000 μJ/cm^2 for 2 min using the UV crosslinker (UVC 500, Amersham).

2.3.11.2 Chemiluminescence probe

The probe was generated using the PCR DIG probe synthesis kit (Roche Diagnostics) according to the manufacturer's instructions. Plasmid and/or genomic DNA was used as the template to generate labeled probes that amplified the upstream (using the USF and USR primer set) and downstream (DSF and DSR primer set) homologous fragments (Figure 2.1 and Table 5.1.1 Appendix A) by PCR to incorporate digoxigenin (DIG)-dUTP.

2.3.11.3 Hybridization and detection

The membranes were placed in roller bottles containing 12 ml of DIG Easy Hyb solution (Roche Diagnostics) and pre-hybridized at 54°C (calculated according to the manufacturer's instructions) for 30 min in the Hybaid oven (Thermo Scientific). The labeled probe was denatured at 95°C for 5 min and added to the membrane and hybridized overnight at 54˚C. The next morning the membranes were washed twice with 25 ml solution I (2× SSC, 0.1% SDS) at room temperature for 5 min followed by two washes with solution II (0.5 × SSC, 0.1% SDS) at 65°C for 15 min each. The DNA was detected with the addition of DIG alkaline phosphatase-conjugated antibody which allows the DNA to be visualized on X-ray film after 30 min – 3 hrs of exposure.

2.4 Bioinformatic Analysis

Various bioinformatic tools were used to identify and analyze the MutY DNA glycosylase in *M. smegmatis* and *M. tuberculosis*. Genome sequence data for *M. tuberculosis* was obtained from the Tuberculosis database TubercuList (http://genolist.pasteur.fr/TubercuList/) and for *M. smegmatis* from the Smegmalist database (http://cmr.tigr.org). *M. smegmatis* sequences TIGR-CMR (http://cmr.TIGR.org/TIGR-scripts/CMR/CmrHomePage.cgi) was used to make genome comparisons of the *mutY* DNA glycosylase gene in *M. smegmatis* and *M. tuberculosis* to establish the degree of sequence similarity and the chromosomal context of the gene in both these organisms.

2.5 Construction of mycobacterial *mutY* knockout vectors

The strategy used to construct the knockout vector is shown in Figure 2.1; the cloning vectors and strategy are illustrated in Appendix B while all other subsequent construct vector maps are shown in Section 3 and Appendix B with detailed characteristics of all vectors in Table 2.2.

2.5.1 Cloning of homologous regions into pGEM3Zf(+)

A *mutY* deficient mutant strain of *M. smegmatis* (ΔY) was generated by homologous recombination (Gordhan and Parish, 2001). Smegmalist (http://mycobrowser.epfl.ch/smegmalist.html) was used to obtain the *M. smegmatis mutY* gene sequence and the Primer3 program (http://frodo.wi.mit.edu/primer3) was used to design primers with unique restriction sites to facilitate cloning. These primers were used to amplify approximately 2kbp of the homologous flanking region upstream and downstream including 100bp on either side of the 882bp *mutY* gene (illustrated with orange arrows in Figure 2.1).

These PCR fragments were ligated separately into pGEM3Zf(+) at the *Sma* I site (Figure 5.2.1 Appendix B). The resulting construct was transformed into competent *E.coli* DH5α cells and transformants were selected on LA plates containing ampicillin. Positive clones were identified by restriction analysis of plasmid extracted from transformants. The correct clones were sequenced to ensure that no mutations were introduced during PCR amplification.

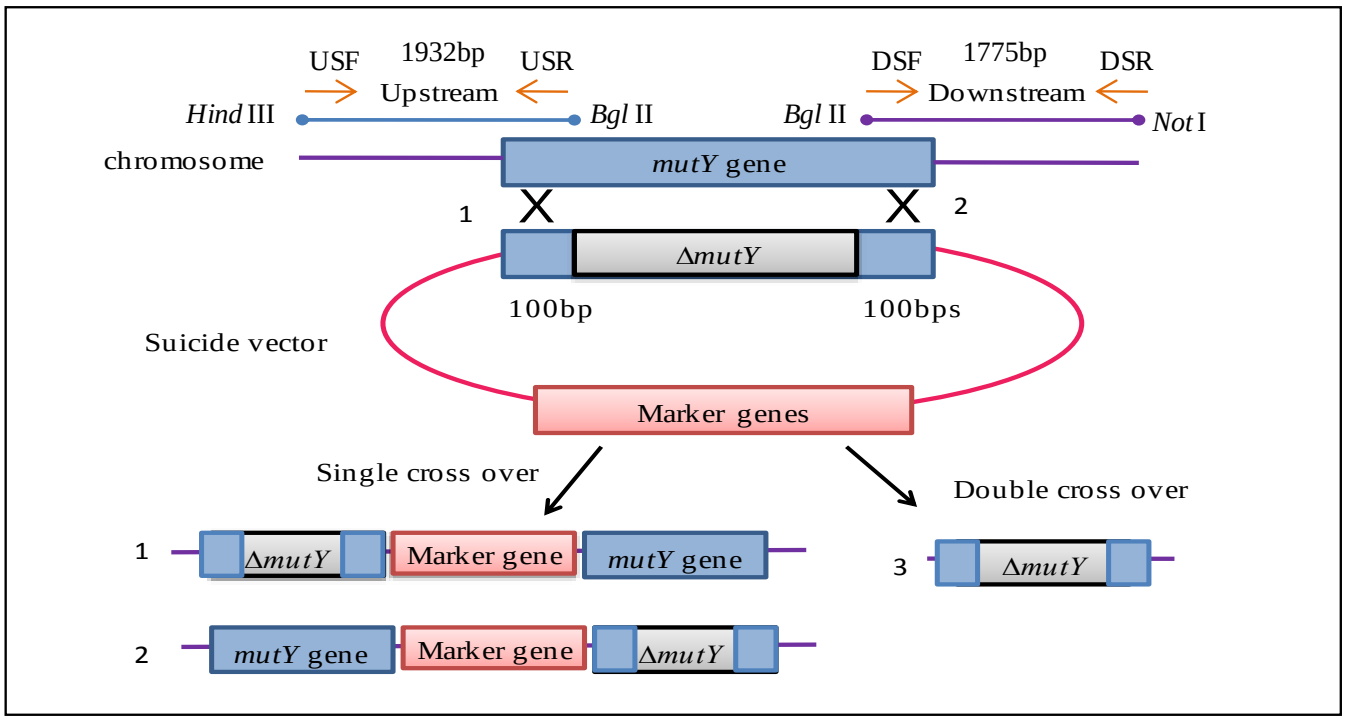


Figure 2.1: Construction of a *ΔmutY* deletion mutant. A deleted *mutY* allele was generated by amplifying the upstream and downstream homologous regions followed by cloning into the suicide vector (p2NIL) by homologous recombination. The suicide vector carrying a disrupted copy of the *mutY* gene was electroporated into *M. smegmatis* to generate single cross over mutants which retains the intact and the disrupted copy of the *mutY* gene (1 and 2). The single cross over mutant undergoes a second selection process using the selectable and counter selectable markers to generate a double cross over mutant in which the functional copy of the *mutY* gene is replaced with the disrupted copy (3). (Adapted from [Gordhan and Parish, 2001]) orange arrows represent the primer pairs used to amplify the upstream and downstream homologous regions with their respective sizes. USF-mutYUSF; USR-mutYUSR; DSF-mutYDSF and DSR-mutYDSR

2.5.2 Three-way cloning into p2NIL

Once the *mutY* upstream and downstream deleted fragments were verified to have no mutations, the upstream PCR fragment containing the *Hind* III and *Bgl* II restriction sites and the downstream PCR fragment containing the *Bgl* II and *Not* I restriction sites were ligated directly into the p2NIL vector at the *Hind* III/*Not* I sites to generate the p2ΔY suicide vector (Figure 5.2.2 Appendix B). The transformations were plated on LA plates containing kanamycin and plasmid DNA extracted from the positive clones was confirmed by restriction enzyme analysis.

2.5.3 Cloning of the *Hyg* and *Pac* I cassette

To generate a marked deletion allele the hygromycin cassette from pIJ963 was inserted between the upstream and downstream regions at the *Bgl* II site shown in the p2ΔY suicide vector map (Figure 5.2.3 Appendix B) to generate the p2ΔY::hyg vector. The construct was transformed into competent *E. coli* cells and transformants were selected on LA plates containing kanamycin and hygromycin. DNA from positive clones was extracted and confirmed by restriction enzyme analysis. In addition, the *Pac* I cassette [counter-selectable markers (*lac*Z and *sacB)*] from pGOAL17 was ligated at the *Pac* I site in p2ΔY::hyg construct (Figure 5.2.4 Appendix B) to generate a marked suicide vector (p2ΔY::hyg::p17) which was transformed into competent *E. coli* cells. The transformations were plated on LA containing hygromycin, kanamycin and X-gal. Blue colonies represented positive clones which were selected for DNA extraction, followed by restriction analysis.

To generate an unmarked suicide vector (p2ΔY::p17), the marked suicide vector (p2ΔY::hyg::p17) was digested with *Bgl* II to remove the hygromycin cassette and re-ligated (Figure 5.2.6 Appendix B). Transformants were selected on LA containing kanamycin and X-gal and the DNA was extracted from blue colonies which were confirmed to be positive clones by

restriction enzyme analysis. A positive clone from the marked and unmarked suicide vector was tested for sucrose sensitivity, 5 µl of serially diluted (10^1-10^6) log phase cultures were spotted on LA plates containing kanamycin, hygromycin and X-gal supplemented with and without 5% sucrose. The unmarked suicide vector was used for all subsequent experiments.

2.5.4 Identification of Single Cross Overs (SCOs)

The unmarked suicide vector carrying the deleted allele of the *mutY* gene was electroporated into the parental strains [mc^2155 (WT)] and the Fpg/Nei deficient mutants [Q1 and Q47] (Table 2.1). Transformants were selected on 7H10 plates containing kanamycin and X-gal after 3-5 days. Single cross over (SCO) transformants carrying the functional and deleted allele of the *mutY* gene were identified as blue colonies due to the hydrolysis of the X-gal by β-galactosidase, encoded by the *lacZ* gene on the vector. Chromosomal DNA was extracted from one SCO mutant for genotypic analysis by Southern blot analysis and if correct was used to generate the *mutY* deletion mutant.

2.5.5 Isolation of *mutY* deletion mutants (DCOs)

To identify *mutY* deletion mutants (Δ*mutY*), the single cross over mutants were forced to undergo a second cross over event in the presence of sucrose to generate a double cross over (DCO) mutant. Colonies that retain the *sacB* cassette will not be viable in the presence of sucrose, since *sacB* expresses levansucrase which converts sucrose to toxic levan which causes bacterial cell death. Hence, white sucrose resistant colonies must lose all the marker genes together with the vector sequence and are indicative of successful double cross over (DCO) mutants or wild type revertants depending on where the second cross over occurred as shown in Figure 2.1. Possible DCOs were further tested on 7H10 media with and without kanamycin to

confirm that the suicide vector together with the counter-selectable markers (*lacZ* and *sacB*) were lost (Gordhan and Parish, 2001).

Chromosomal DNA was extracted from potential DCOs and screened initially by PCR to distinguish between parental and DCOs using primer sets (Table 5.1.1 Appendix A) in the region of the *mutY* deletion. Confirmed positive deletion clones were further assessed by Southern blot analysis to ensure that the genes upstream and downstream of the deletion were genotypically correct.

2.5.6 Construction of complementary *mutY* strain

Primers were designed (Table 5.1.1 Appendix A) to amplify 200 bp upstream of the start codon of the *mutY* gene to include the promoter regions. The amplicon was cloned into an integrating vector pTWEETY at the *Sca* I site shown in Figure 5.2.7 (Appendix B). Transformants were screened by restriction analysis and the correct clone was sent for sequencing to ensure that no mutations were introduced during PCR amplification of the *mutY* fragment. The complemented vector was electroporated into the various *mutY* deficient mutant strains (Table 2.1) and positive clones identified by PCR were assessed for *mutY* expression by Real Time PCR (section 2.5.7).

2.5.7 Real Time PCR of the various complemented strains

RNA from the various strains was isolated according to the manufacturer's instructions with some modifications (NulceoSpin RNA II kit, Machery-Nagel). Mid-log phase cultures (OD_{600nm} ~ 0.67) were harvested at 4500g for 10 min followed by re-suspension of the pellets in TE buffer containing lysozyme (10mg/ml) and incubated at 37°C for 30 min. The treated cells were transferred to lysing matrix B tubes (MP Biomedicals) to which lysis buffer and β-mercaptoethanol was added and the cells lysed in the ribolyzer (Savant Fastprep FP120) at speed

6 for 45 sec. This was repeated three times with cooling of the cells on ice between each lysis step. The lysed cells were passed through several spin columns (provided in the kit) in order to reduce the viscosity of the lysate and to remove salts to make the DNase treatment more effective. The DNase was inactivated with several washes using the wash buffers provided and the RNA was eluted with RNase-free water followed by quantification on a nanodrop.

1 µg of RNA was treated with turbo DNAse as per the manufacturer's instructions (Ambicon). Briefly, 1 µl of Turbo DNAse enzyme and 1µl 10X buffer was added to the RNA in a 20 µl reaction volume and incubated at 37°C for 30 min after which the enzyme was inactivated by the addition of the inactivation reagent. The mixture was centrifuged at 13 000 rpm for 3 min and the supernatant was used for cDNA conversion as described by the manufacturer (Invitrogen). The reactions were performed as follows: to 6.25 µl of supernatant (RNA), 0.5 µl of 10µm dNTP and 1 µl random hexamer were added and incubated at 65°C for 5 min followed by the addition of 1 µl of 5X RT buffer, 1 µl of 0.1M DTT, 2 µl of 25mM $MgCl_2$ and 0.25 µl of 200U/µl superscript III. For the control reaction no superscript was added.

Conditions for the cDNA conversion was as follows: 25°C for 10 min, 50°C for 50 min and 85°C for 5 min. FastStart Taq DNA polymerase was used to amplify 5% of the cDNA using the Real Time primers in Table 5.1.2 (Appendix A) with an annealing temperature of 60°C.

2.6 Phenotypic characterization of the parental and mutant strains

The parental and the mutant strains were characterized for differences in growth kinetics under normal culture conditions, followed by sensitivity to oxidative stress (induced by hydrogen peroxide and cumene peroxide).

2.6.1 Growth kinetics under normal culture conditions

Freezer stocks of the parental, Fpg/Nei deficient mutants and *mutY* deficient mutant strains were thawed at room temperature, inoculated in 5 ml of 7H9 media and grown to log phase OD_{600nm} = 1.0. The pre-culture was then used to inoculate 50 ml of 7H9 media with a starting OD_{600nm} = 0.02. The cultures were incubated at 37˚C in a shaking incubator and the growth of each strain was monitored at 3 hr intervals over a 39 hr period by optical density measurements. The experiment was conducted twice for statistical significance.

2.6.2 Determination of the optimal concentration of the oxidative agents

The optimum concentration for hydrogen peroxide to give an approximate 3 log kill in *M. smegmatis* was previously optimized as 2.5 mM (Goosens, 2008). The optimum concentration of cumene peroxide, menadione and tert-butyl hydroperoxide was determined by spotting 5 µl of early log phase mc^2155 culture OD_{600nm} = 0.2 on 7H10 media containing different concentrations of the agents and incubated at 37˚C for 3 days. The concentration that displayed at least a 2-4 log decrease in growth was used to test survival of mc^2155 in liquid culture. Once the appropriate concentration was determined the parental, Fpg/Nei deficient mutants, the *mutY* deficient mutants and complemented strains were assessed under these conditions.

2.6.3 Sensitivity to oxidative stress induced by hydrogen peroxide and cumene peroxide

The parental, Fpg/Nei deficient mutants, the *mutY* deficient mutants and complemented strains were tested for their sensitivity to oxidative stress as generated by hydrogen peroxide which is a light sensitive compound, therefore demands high level of technical uniformity to achieve consistent data (Goosens *et al*., unpublished). Log phase cultures were grown in 7H9 media (OD_{600nm} = 0.35) and 2.5 mM hydrogen peroxide was carefully added to the 30 ml culture and survival was assessed by sampling at 0, 2, 4 and 6 hr intervals. Appropriate dilutions were

spread on 7H10 plates which were incubated at 37°C and the viable cell counts (cfu) were assessed after 3 days. A second oxidizing agent, cumene peroxide was tested as it is a lot more stable and less light sensitive than hydrogen peroxide. Early log phase cultures (OD_{600nm} = 0.20) were grown in 7H9 media to which 3.8 mM cumene peroxide was added and 100 µl of the appropriate serial dilutions (10^0-10^{-5}) were plated at 0, 3 and 6 hr intervals. Viable cell counts were assessed after 3 days.

2.6.4 Increase in mutation rates as measured by the fluctuation assay

Mutation rates measure the likelihood of a cell to gain a mutation during its lifetime. The Luria-Delbrück fluctuation analysis was used to accurately measure the increase in mutation rates for the various mutant and complemented strains. Since this assay considers the variations between parallel cultures, the possibility of mutations arising early or late during growth and the number of mutational events can be predicted (Rosche and Foster, 2000). Briefly, single colonies were grown in 5 ml of 7H9 media to late log phase (10^8 cells/ml) and a dilution series (10^0-10^{-7}) of the culture was plated onto 7H10 media to measure the initial inoculum which should have no pre-existing mutants. 1 ml of the culture was also plated on 200µg/ml 7H10 rifampicin plates to confirm the absence of any pre-existing mutants. The cell density was reduced from 10^8 cells/ml to 10^2 cells/ml and 2 ml aliquots were dispensed into 30 tubes and incubated at 37°C with shaking for 5-7 days. A dilution series from 5 representative tubes out of the 30 tubes were assessed for population size as cfu/ml on 7H10 media, while the culture from the remaining 25 tubes was plated onto 25 individual 7H10 plates containing 200µg/ml rifampicin to identify spontaneous rifampicin resistant mutants (Rosche and Foster, 2000 and Machowski *et al*., 2007). The mutation rates were calculated using the equations in Figure 3.25 as well as the computer software developed at the CBTBR (Machowski *et al.,* 2007).

2.6.5 Mutation spectrum of spontaneous rifampicin resistant mutants

DNA from between 16-20 rifampicin resistant colonies isolated in the fluctuation assay for the parental, Fpg/Nei deficient mutants, the *mutY* deficient mutants and complemented strains were extracted as described in section 2.2.2.1 and the 184bp fragment of the Rifampicin Resistance Determining Region (RRDR) within the *rpoB* gene was PCR amplified using the primer set within the RRDR region (Table 5.1.2, Appendix A). The PCR products were sequenced and the data was aligned to the Smegmalist sequence using the Seqman program to identify the mutations responsible for the resulting rifampicin resistance.

3 Results

3.1 Bioinformatic analysis of the *mutY* gene

Biochemical activity assays revealed that the mycobacterial MutY (*M. smegmatis* and *M. tuberculosis*) has similar substrate specificities to *E. coli* MutY in recognizing adenines mispaired with 8-oxoG, hence a single copy of the *mutY* gene is annotated as a probable DNA glycosylase (Jain *et al*., 2007; Cole *et al.*, 1998). The TIGR-CMR research tool showed that the *mutY* gene is in the same genomic context in all the organisms analyzed and most of the genes upstream and downstream of *mutY* are also conserved in a similar context (Figure 3.1).

MutY is not only conserved in mycobacteria but homologues are present in most other bacteria such as *Bacillus subtilis* and *Helicobacter pylori* (Figure 3.1) thus indicating a protective role of this DNA repair enzyme in genome maintenance under oxidative stress conditions (Debora *et al*., 2011, Eutsey *et al*., 2007).

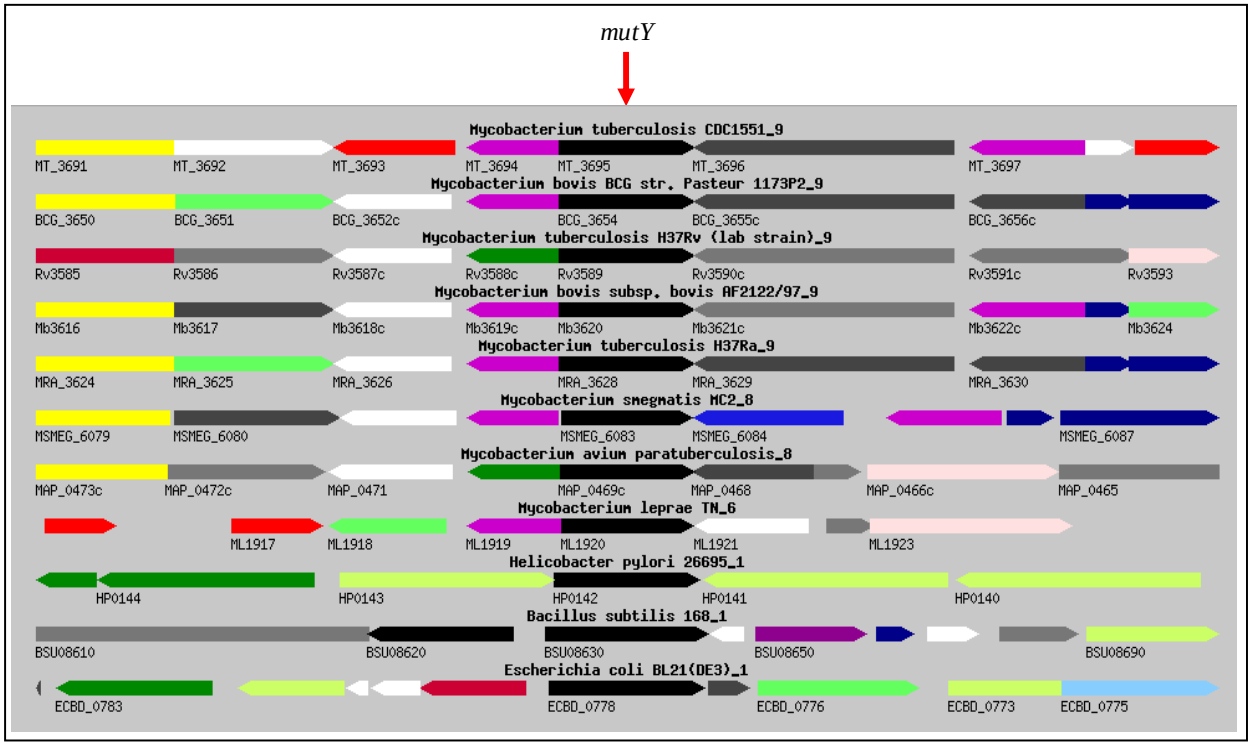


Figure 3.1: Comparative bioinformatic analysis of *mutY* in sequenced genomes. TIGR-CMR (http://cmr.TIGR.org/TIGR-scripts/CMR/CmrHomePage.cgi) was used to analyze the genomic organization of *mutY* (indicated by a red arrow) in which *M. smegmatis* was compared to several other mycobacterial species, *E. coli* and other pathogenic organisms. The same colour arrows indicate homologous genes in the different organism.

While the crystal structure, mechanism and function of MutY is well characterized in *E. coli* (Fromme *et al.*, 2004; Michaels and Miller, 1992) not much is known about this enzyme in mycobacteria except it has been shown biochemically that MutY was able to excise adenines that were base paired to 8-oxoG (Kurthkoti *et al.*, 2010; Guo *et al.*, 2010). Hence, in this study an attempt was made to better understand the physiological role of the mycobacterial MutY in DNA repair in relation to other DNA glycosylases. These investigations were carried out using a close relative of *M. tuberculosis*, *M. smegmatis* as it is non-pathogenic, faster growing, and the *mutY* gene is in the same chromosomal context with 70% sequence identity in both the pathogenic and non-pathogenic counterparts thus, making *M. smegmatis* an ideal model organism for this study.

In *E. coli* it has been shown that MutY plays an important role in DNA repair, specifically in preventing misincorporation of adenine opposite the 8-oxoG mutations (Nghiem *et al*., 1988). The MutM protein removes 8-oxoG lesions from DNA and the deletion of both *mutM* and *mutY* results in a 25-75-fold higher mutation rate than either mutator alone, suggesting that a combinatorial loss of these two repair enzymes severely affects genome stability (Michaels *et al*., 1992). Thus, to better understand the physiological role of the mycobacterial MutY in genome maintenance, its function was investigated in context of the Fpg/Nei family of DNA glycosylases. Homologous recombination (Gordhan and Parish, 2001) was used to inactivate the *mutY* gene in the parental *M. smegmatis* strain as well as in selected combinatorial mutants deficient in the Fpg/Nei family of DNA glycosylases (Goosens *et al*., unpublished).

3.2 Construction of mycobacterial *mutY* knockout vectors

3.2.1 Cloning of homologous regions into pGEM3Zf(+)

The availability of the *M. smegmatis* genome sequence (http://mycobrowser.epfl.ch/smegmalist.html) allowed for the design of a suitable cloning strategy to generate a *mutY* deletion vector using PCR. The *mutY* gene sequence that included 2kbp of upstream and downstream sequence was retrieved from the Smegmalist database. Primers were designed to include unique restriction sites and to amplify the upstream and downstream homologous regions that include only ~100bp on either side of the *mutY* gene as indicated in Figure 3.2. The primers used to amplify the upstream (1932bp) and the downstream (1775bp) regions are denoted in Table 5.1.1 (Appendix A) as (MutYUSF and MutYUSR) and (MutYDSF and MutYDSR) respectively.

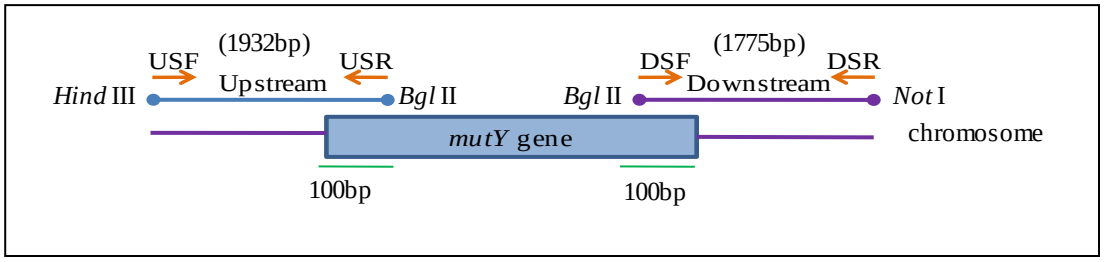


Figure 3.2: PCR amplification strategy for the generation of the upstream and downstream homologous fragments. The orange arrows represent the forward and reverse primers with the unique restriction sites indicated to amplify the upstream and downstream homologous fragments for Δ*mutY* (indicated by the blue and purple lines) that includes a 100bp on either side of the gene (indicated in green).

The PCR amplicons were each cloned separately into the *Sma* I site of pGEM3Zf(+) as mentioned in Section 2.5.1 and the cloning strategy is shown in Figure 5.2.1 (Appendix B), to generate the pGΔYUS and pGΔYDS vectors respectively. The vectors with the cloned upstream and downstream homologous fragments were transformed in competent *E. coli* DH5α; transformants were selected in the presence of ampicillin and the positive clones were screened by restriction enzyme analysis. The expected fragment sizes were obtained with all the restriction enzymes for both pGΔYUS and pGΔYDS vectors (Figures 3.3 and 3.4) except for pGΔYDS (Figure 3.4) an extra band was observed with *Sca* I and *Bgl* II/*Not* I digests in addition to the expected bands. The DNA was partially digested with these restriction enzymes and therefore accounts for the extra band in each case. Even though a high fidelity polymerase was used to PCR amplify the US and DS amplicons, both pGΔYUS and pGΔYDS were sequenced using the universal M13 primers as well as the upstream and downstream internally designed sequencing primers, US1, US2, US3 and DS1, DS2, DS3 respectively (Table 5.1.2 Appendix A), to confirm that no mutations were introduced during the PCR amplification. The sequence alignments showed a total match of the amplified sequences to the database sequence (data not shown) thus the cloning process to generate a suicide vector carrying a *mutY* deletion allele was continued.

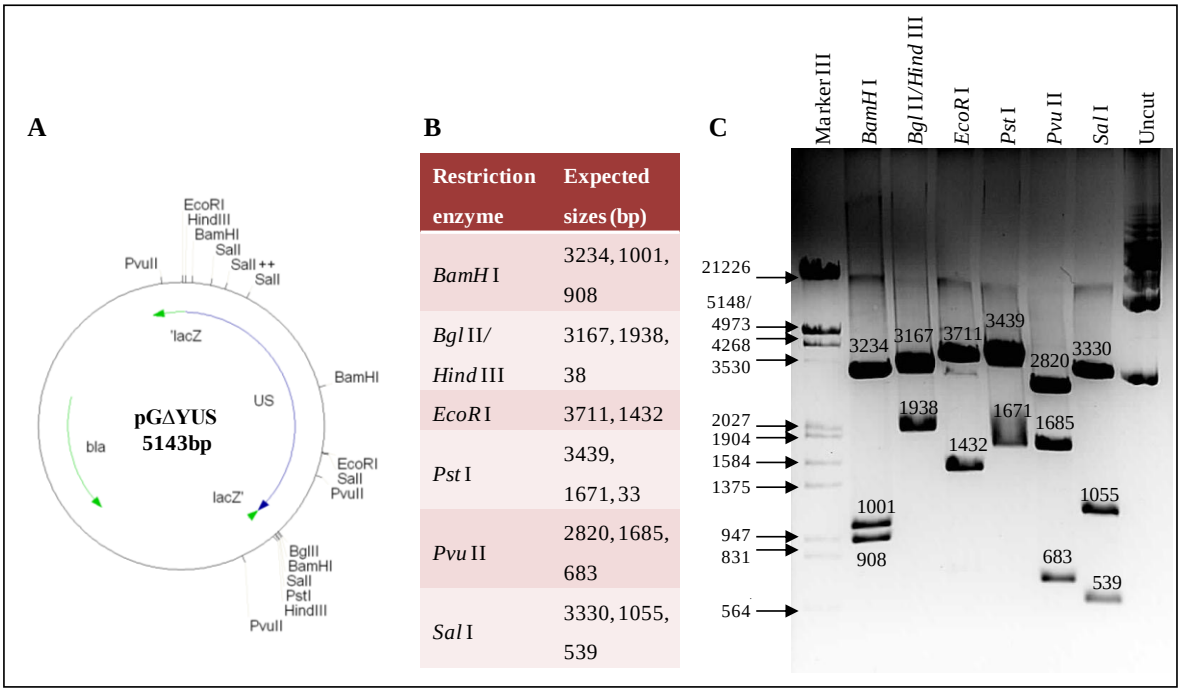


Figure 3.3: Restriction analysis of the upstream (US) PCR fragment in pGEM3Zf(+). A. Restriction map of pGΔYUS vector with the upstream fragment ligated at the *Sma* I site of pGEM3Zf(+), upstream fragment (US) indicated in blue and the marker genes in green. **B**. Expected sizes of fragments for a selection of restriction enzymes used for mapping. **C**. Restriction digest of the pGΔYUS vector in which all the digests correspond to the expected sizes.

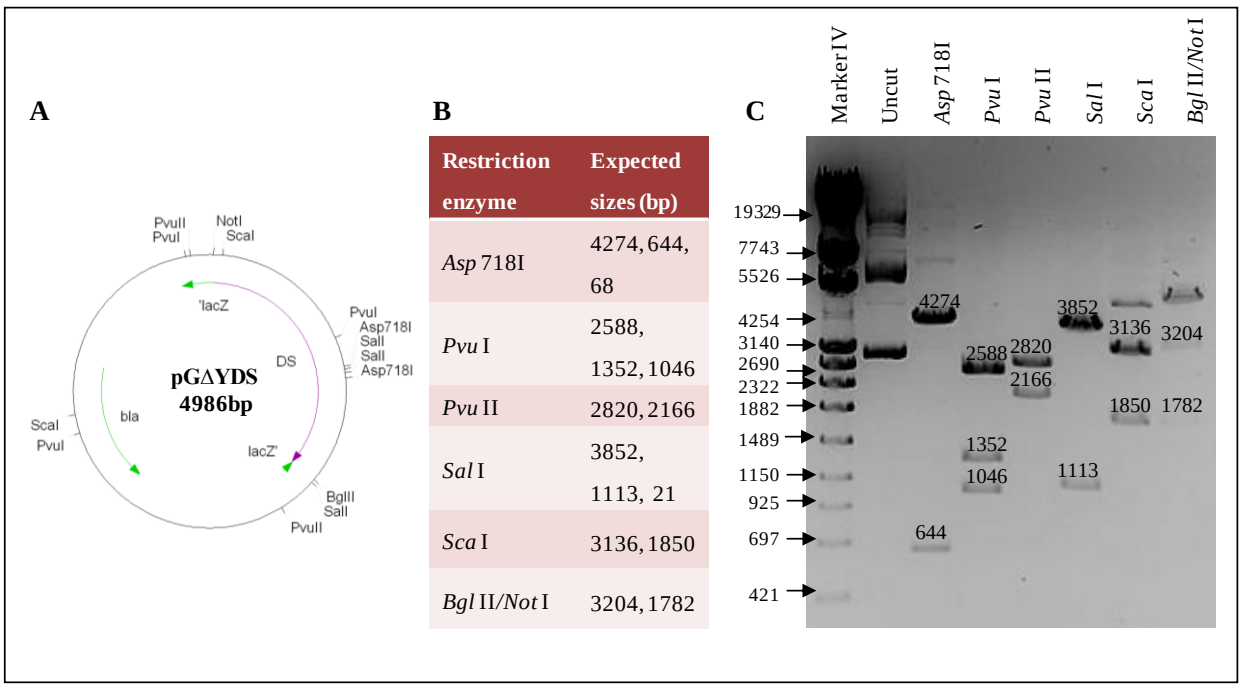


Figure 3.4: Restriction analysis of the downstream (DS) PCR fragment in pGEM3Zf(+). **A**. restriction map of pGΔYDS vector in which the downstream (DS) fragment is represented in purple was ligated at the *Sma* I site of pGEM3Zf(+). **B**. Expected sizes of fragments for a selection of restriction enzymes used for mapping. **C**. Restriction digest of the pGΔYDS vector in which all the digests correspond to the expected sizes with partial digests for *Sca* I and *Bgl* II.

3.2.2 Three-way cloning into p2NIL

To create a *mutY* deletion suicide construct the cloning was carried out as mentioned in Section 2.5.2. The upstream (US) fragment containing the compatible restriction sites *Hind* III and *Bgl* II was excised from pGΔYUS while the downstream (DS) fragment was excised at the *Bgl* II and *Not* I restriction sites from pGΔYDS. These fragments were cloned simultaneously into the *Hind* III/*Not* I compatible sites in the p2NIL vector to generate the suicide vector p2ΔY carrying a deleted allele of *mutY* fused at the *Bgl* II site (Figure 5.2.2 Appendix B). Positive clones were selected and screened by restriction map analysis which confirmed that the clones were stable as the expected fragments were obtained for all the restriction digests (Figure 3.5). As a final cloning step, the selectable and counter-selectable markers were cloned into p2ΔY.

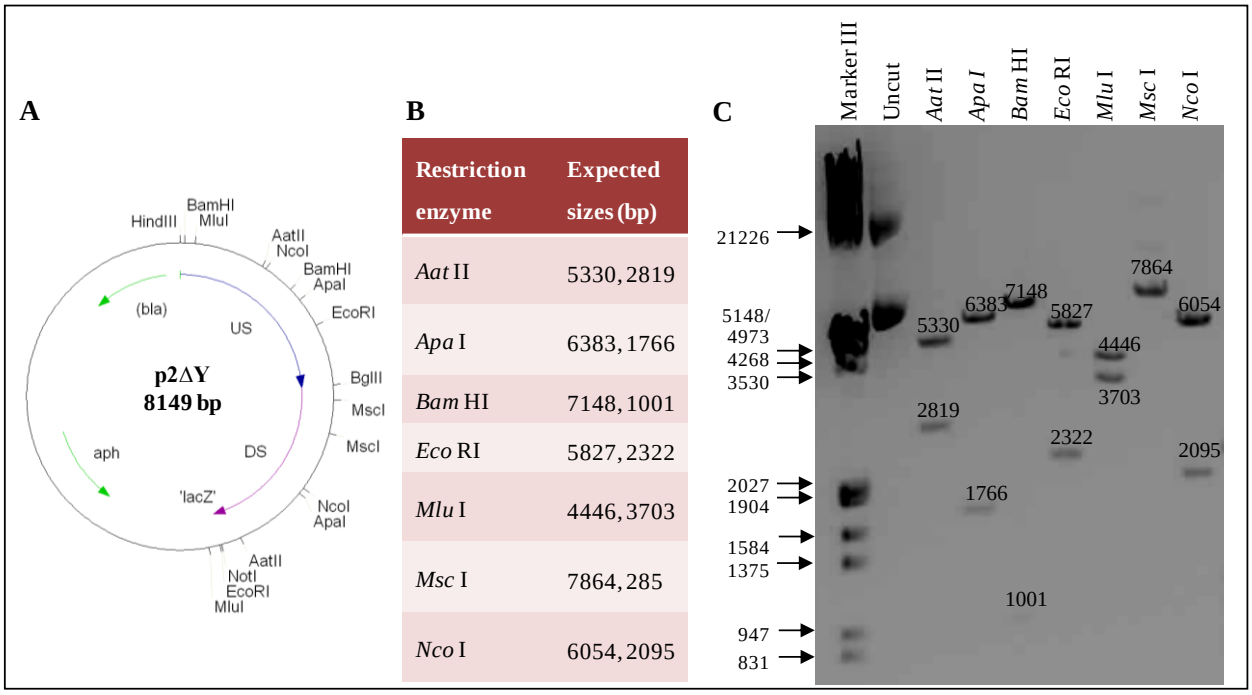
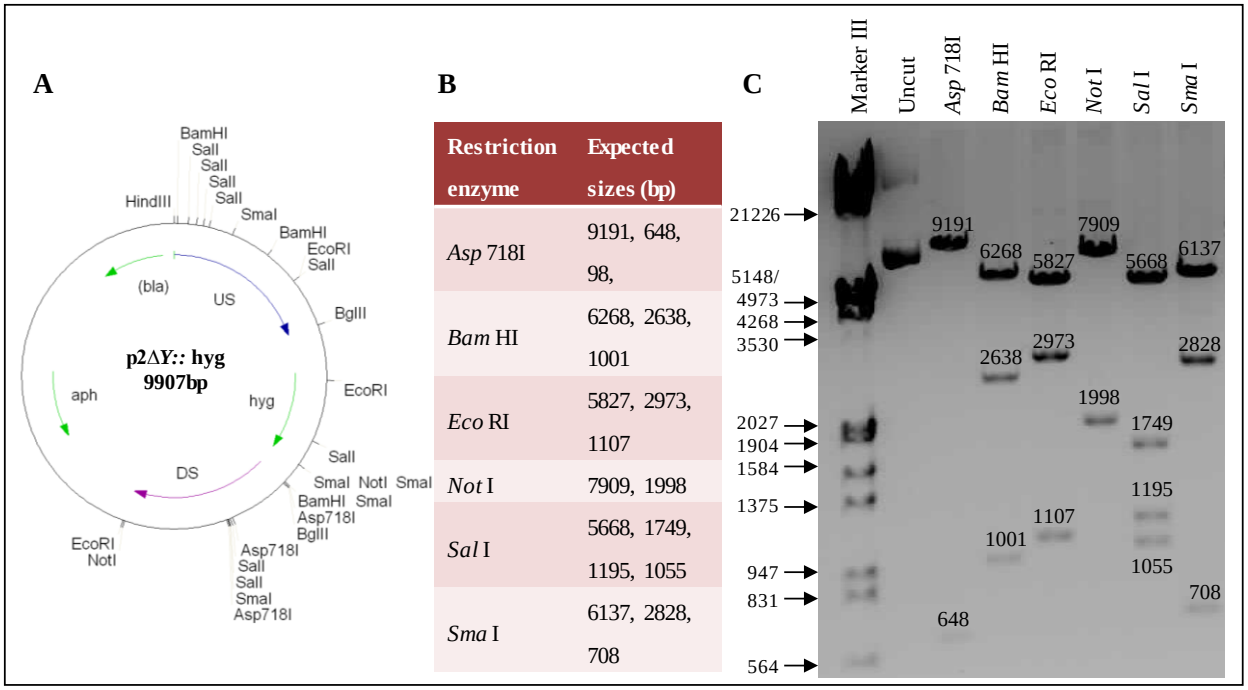


Figure 3.5: Generation of a *mutY* deletion suicide vector. A. p2ΔY restriction map in which the US – upstream fragment is depicted in blue and the DS - downstream fragment in purple. **B.** Expected sizes of fragments for a selection of restriction enzymes used for mapping. **C.** Restriction analysis of the p2ΔY suicide vector, all digests correspond to the expected sizes.

3.2.3 Cloning of selectable and counter-selectable markers

The 1758bp hygromycin resistant antibiotic cassette was isolated as a *Bgl* II fragment from the vector pIJ963 and inserted into p2ΔY at the unique *Bgl* II site joining the US and DS fragments illustrated in Figure 5.2.3 (Appendix B) to generate a marked deletion allele of *mutY*. An antibiotic marker ensures guranteed isolation of positive mutants thereby reducing the downstream screening to identify a positive clone which can be labour intensive. However, marked mutants have limitations and are not very useful for *in vivo* screening as the antibiotic resistance marker could interfere with the phenotyopic outcome. After cloning, a single white colony was selected and screened by several restriction digests to confirm that the clone was correct (Figure 3.6).



Restriction enzyme	Expected sizes (bp)
Asp 718I	9191, 648, 98,
Bam HI	6268, 2638, 1001
Eco RI	5827, 2973, 1107
Not I	7909, 1998
Sal I	5668, 1749, 1195, 1055
Sma I	6137, 2828, 708

Figure 3.6: Generation of a marked suicide vector. A. Restriction map of the marked suicide vector with the hygromycin cassette (indicated in green) ligated in between the upstream (US) and downstream (DS) fragments. **B.** Expected sizes of fragments for a selection of restriction enzymes used for mapping. **C.** Restriction analysis of the p2ΔY::hyg vector in which all the digests corresponded to the expected sizes.

Lastly the counter-selectable marker genes *lacZ* and *sacB* were excised as a *Pac* I cassette (6359bp) from the pGOAL 17 vector, depicted in Figure 5.2.4 (Appendix B) and cloned into the p2ΔY::hyg marked suicide vector to facilitate the generation of a *M. smegmatis mutY* deficient mutant. The agar plates were incubated at 30°C to prevent rearrangements during replication of the vector which had significantly increased in size (16266bp). Restriction map analysis resulted in the expected sizes for all the enzymes tested confirming that the clone was stable (Figure 3.7).

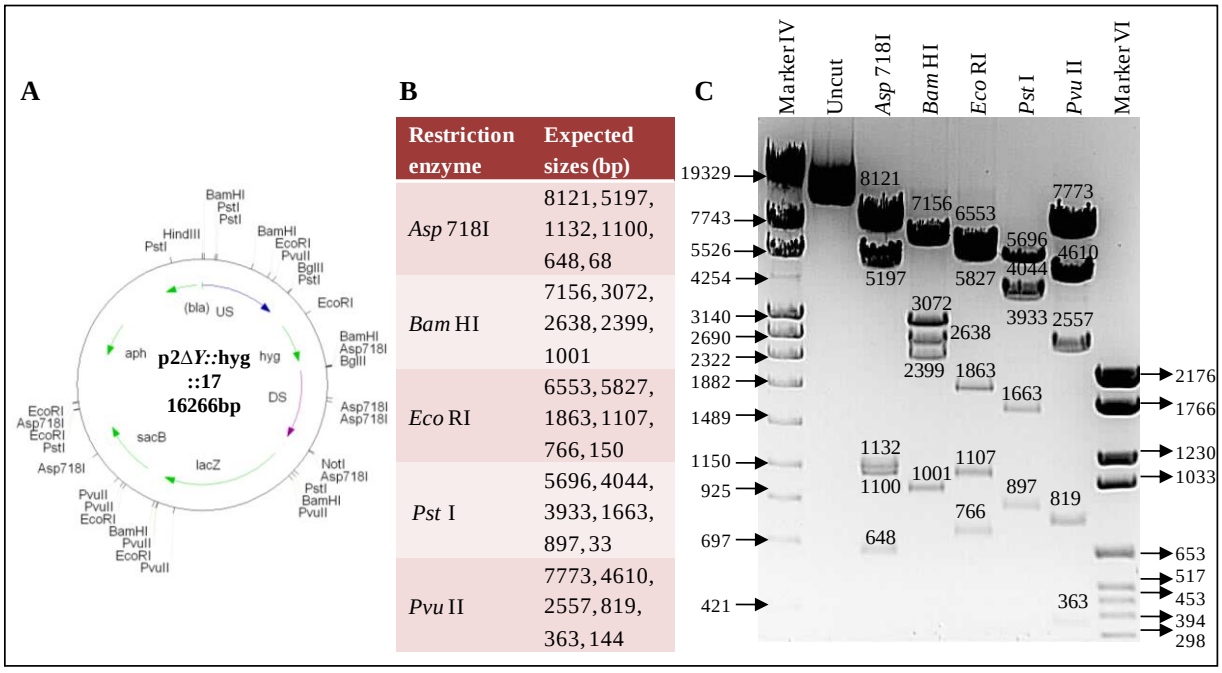


Figure 3.7: Generation of marked suicide vector containing the *Pac* cassette from pGOAL17. A. Restriction map of the p2ΔY::hyg::17 vector with the *lacZ* and *sacB* genes from pGOAL 17 depicted in the green arrows. **B.** Expected sizes of fragments for a selection of restriction enzymes used for mapping. **C.** Restriction analysis of p2ΔY::hyg::17, the restriction digests corresponded to the expected sizes.

Several attempts were made to generate an unmarked deletion vector of *mutY* which required the cloning of the *lacZ, sacB* and the hygromycin cassette from pGOAL19 (Figure 5.2.5 Appendix B) also as a *Pac* I cassette (7939bp). However, this cloning was inefficient perhaps due to the larger size of the *Pac* I cassette compared to the marked construct. Therefore, as an alternate strategy an unmarked suicide vector was generated by deleting the hygromycin cassette

with *Bgl* II from the marked suicide vector (p2ΔY::hyg::17) and allowing the vector to recircularize (Figure 5.2.6 Appendix B). Positive clones were screened with several restriction enzymes and all the fragments corresponded to the expected sizes (Figure 3.8).

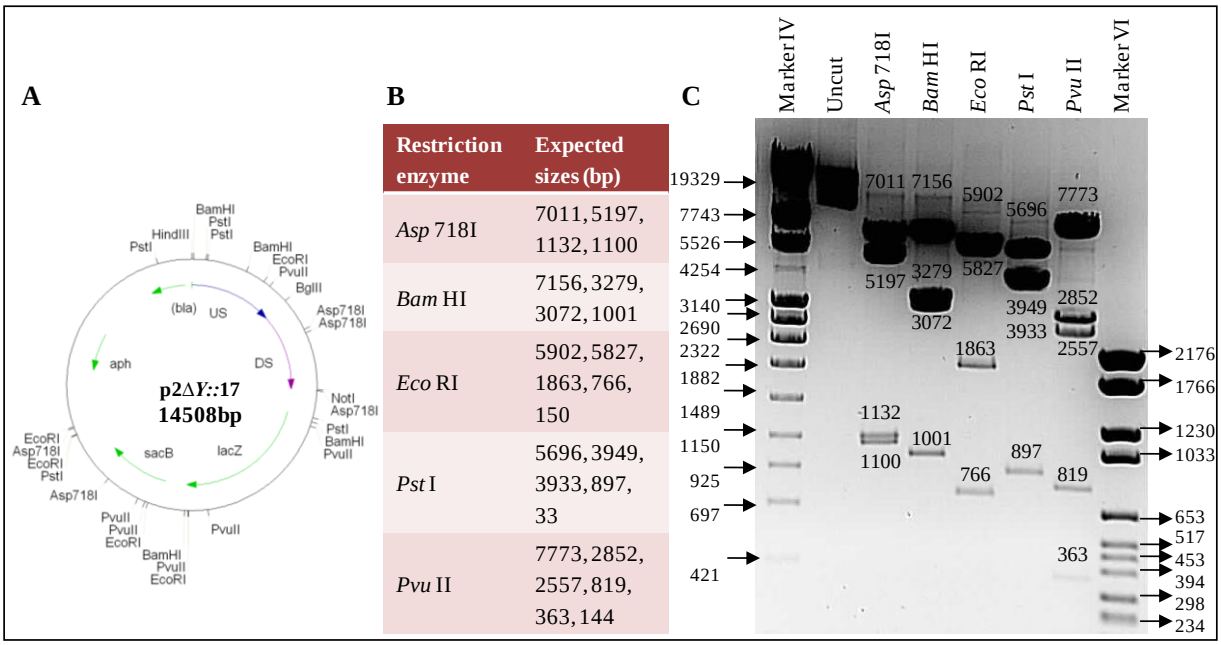


Figure 3.8: Generation of unmarked suicide vector p2Δ*Y*::17. A. Restriction map of the p2ΔY::17 vector in which the *hyg* cassette was removed. **B.** Expected sizes of fragments for a selection of restriction enzymes used for mapping. **C.** Restriction analysis of p2ΔY::17 in which the restriction digests all corresponded to the expected sizes.

Both the *E.coli* marked and unmarked *mutY* deficient mutant strains were tested for sucrose sensitivity and β-galactosidase activity to ensure that the *lacZ* and *sacB* (selectable and counter-selectable markers) genes were expressed. The respective clones were grown in LB media supplemented with kanamycin to mid-log phase and 5 µl of serial dilutions 10^1-10^6 were spotted on LA containing kanamycin, hygromycin and X-gal with (Figure 3.9 A and C) and without 5% sucrose (Figure 3.9 B and D). The spots were blue in colour due to the hydrolysis of the X-gal by β-galactosidase, encoded by the *lacZ* gene thus confirming that the gene was expressed. In the presence of sucrose, the *sacB* gene expresses levansucrase which converts sucrose to a toxic compound that prevents bacterial growth indicating that the bacteria are

sensitive. There was at least a 6 log kill for the marked strain and about 5 log kill for the unmarked strain (Figure 3.9). The sucrose sensitivity and the restriction analysis confirmed that both the marked and the unmarked vectors were correct however, only the unmarked construct purified from *E.coli* was used to generate *mutY* deficient mutants in *M. smegmatis*.

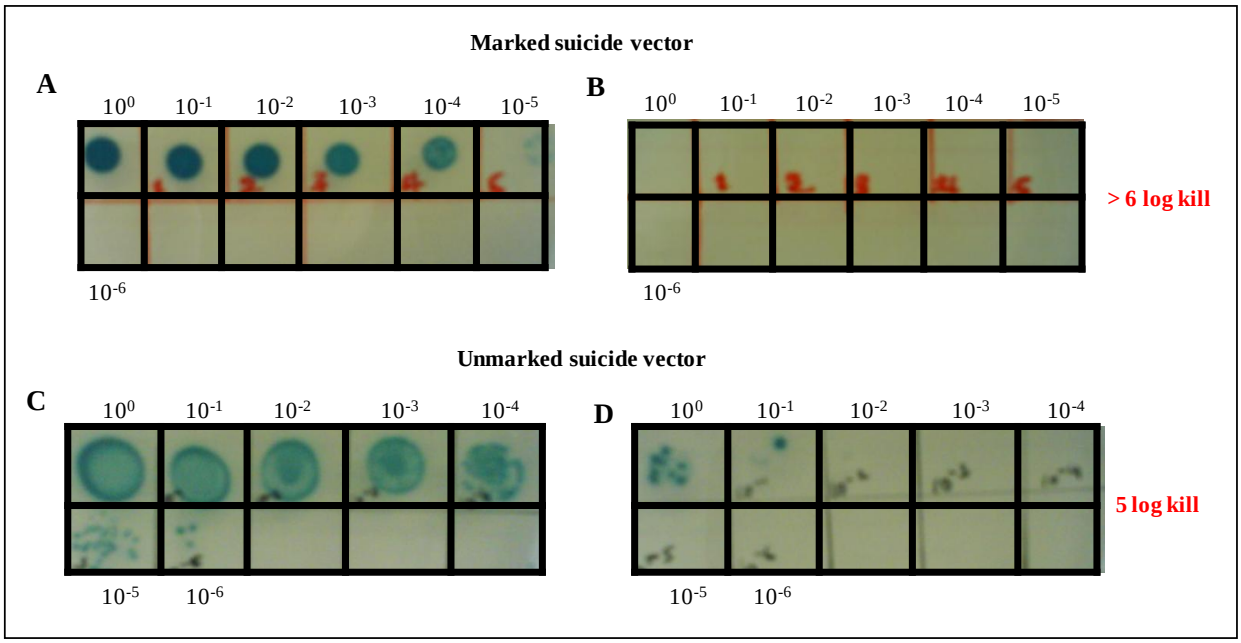


Figure 3.9: Sucrose sensitivity test of the marked and unmarked suicide vectors. The marked and unmarked suicide vectors displayed higher than 10^6 log kill and 10^5 log kill respectively. **A**- LA containing kanamycin, hygromycin and X-gal; **B**- LA containing sucrose, kanamycin, hygromycin and X-gal; **C**- LA containing kanamycin and X-gal; **D**- LA containing sucrose, kanamycin and X-gal).

3.3 Identification of SCOs and DCOs

The unmarked suicide vector (p2ΔY::17) was electroporated into the *M. smegmatis* parental (WT) mc^2155 strain and in the Fpg/Nei deficient mutant strains (Q1 and Q47 Table 2.1) that were previously generated in the laboratory (Goosens *et al.*, unpublished). Kanamycin resistant, blue transformants represented single cross over (SCO) mutants containing the functional copy and a deleted allele of *mutY* together with the selectable and counter-selectable vector marker genes (Figure 2.1). The genotype of the SCO was confirmed by Southern blot analysis before the SCO mutants were subjected to undergo a second cross over event in the

presence of sucrose to generate a double cross over (DCO) mutant. This two step selection process ensures that only colonies that have lost the *sacB* cassette will be able to survive in the presence of sucrose since the expression of the *sacB* gene produces levansucrase which is toxic to the cell hence, colonies that retain the *sacB* cassette will not be viable in the presence of sucrose. Loss of the *sacB* gene results in the simultaneous loss of the *lac*Z gene hence, the resulting colonies will be white and sucrose resistant which are possible DCOs (*mutY* mutants). These white colonies can also be wild type revertants if the cross over occurred on the same side as the SCO (depicted in Figure 2.1).

The possible DCOs were further confirmed to have lost the vector sequences carrying the marker genes by spot testing 20-30 white sucrose resistant colonies from each strain onto 7H10 media containing X-gal with and without kanamycin and sucrose (Gordhan and Parish, 2001). The possible DCOs from the various mutant strains (ΔY, Q1ΔY and Q47ΔY) were white and displayed sensitivity to sucrose and kanamycin, thus confirming the loss of the vector sequence carrying the marker and antibiotic genes (Figure 3.10). Hence, these positive clones (DCOs) were screened genotypically initially by PCR followed by Southern blot analysis to identify and confirm site specificity of the *mutY* deletion in the mutant strains.

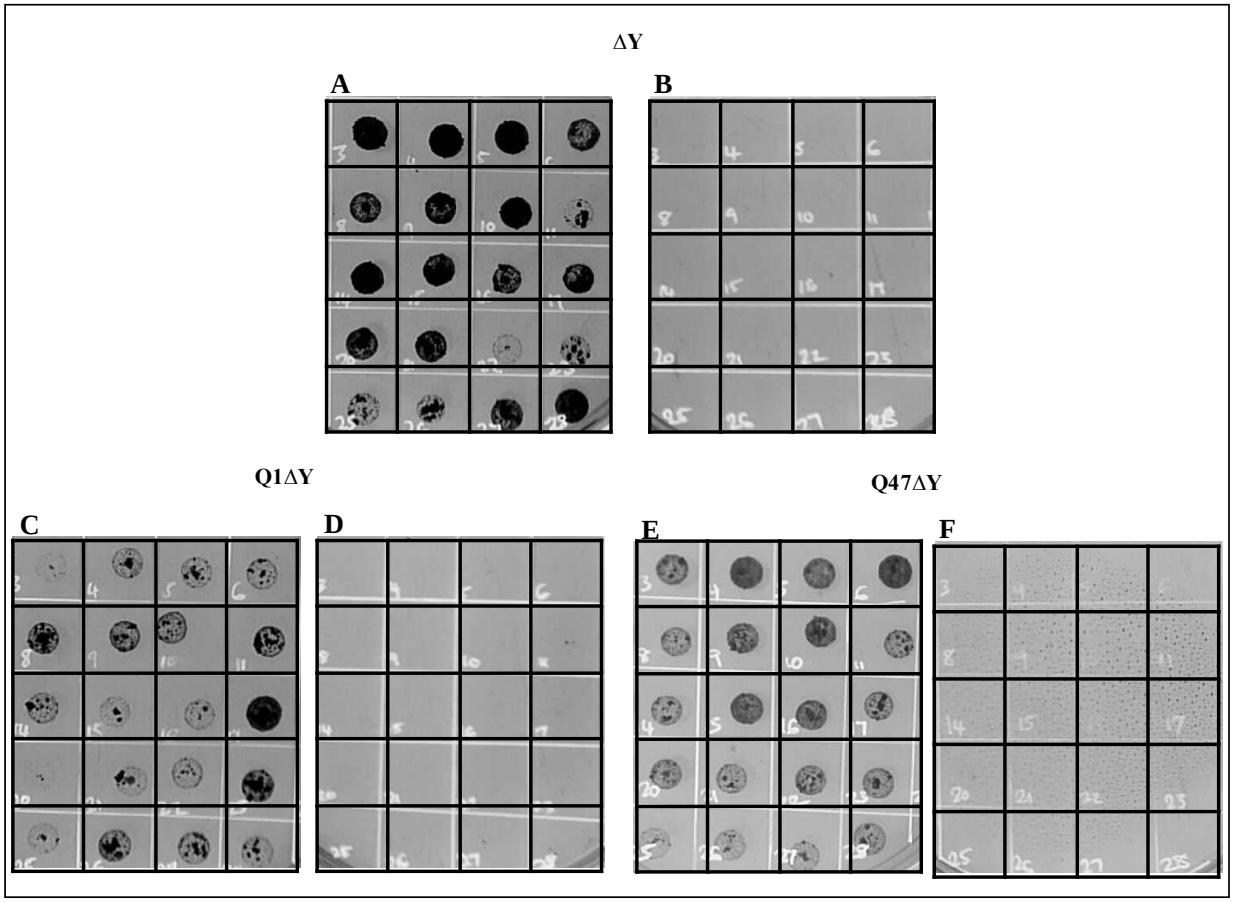


Figure 3.10: Spot testing of possible DCOs (*mutY* mutants). Re-suspended white resistant colonies from the ΔY, Q1ΔY and Q47ΔY strains were spotted on 7H10 media plates containing X-gal with and without kanamycin and sucrose. **A, C and E** - 7H10 containing X-gal; **B, D and F** - 7H10 containing X-gal, kanamycin and sucrose.

3.4 Genotypic verification of parental and mutant strains

3.4.1 Identification of mutants by PCR

PCR screening was performed as shown in Figure 3.11 using the primer set (mutYFwd and mutYRev) listed in Table 5.1.2 (Appendix A) to discriminate between clones containing an intact copy of the *mutY* allele and mutants that had a disrupted copy of the *mutY* allele. Chromosomal DNA from several (4-6) possible DCO clones isolated from the different mutant strains was amplified and the expected size for the intact *mutY* allele (1481bp) was obtained in the parental strain (WT) and the Fpg/Nei deficient mutant strains (Q1 and Q47) whilst a smaller 733bp fragment was observed when the functional copy was replaced with a disrupted copy of

the gene (ΔY, Q1ΔY and Q47ΔY). For a SCO mutant, both the intact copy and the mutant allele of the *mutY* gene were amplified as shown in Figure 3.11 (WTSCO). Several potential DCOs were identified for each of the strains but only one representative from each strain was analyzed by Southern blotting to ensure that gene replacement was specific and the sequence upstream and downstream of the deletion was genotypically correct i.e. no rearrangements had occurred during the homologous recombination process during the mutant generation.

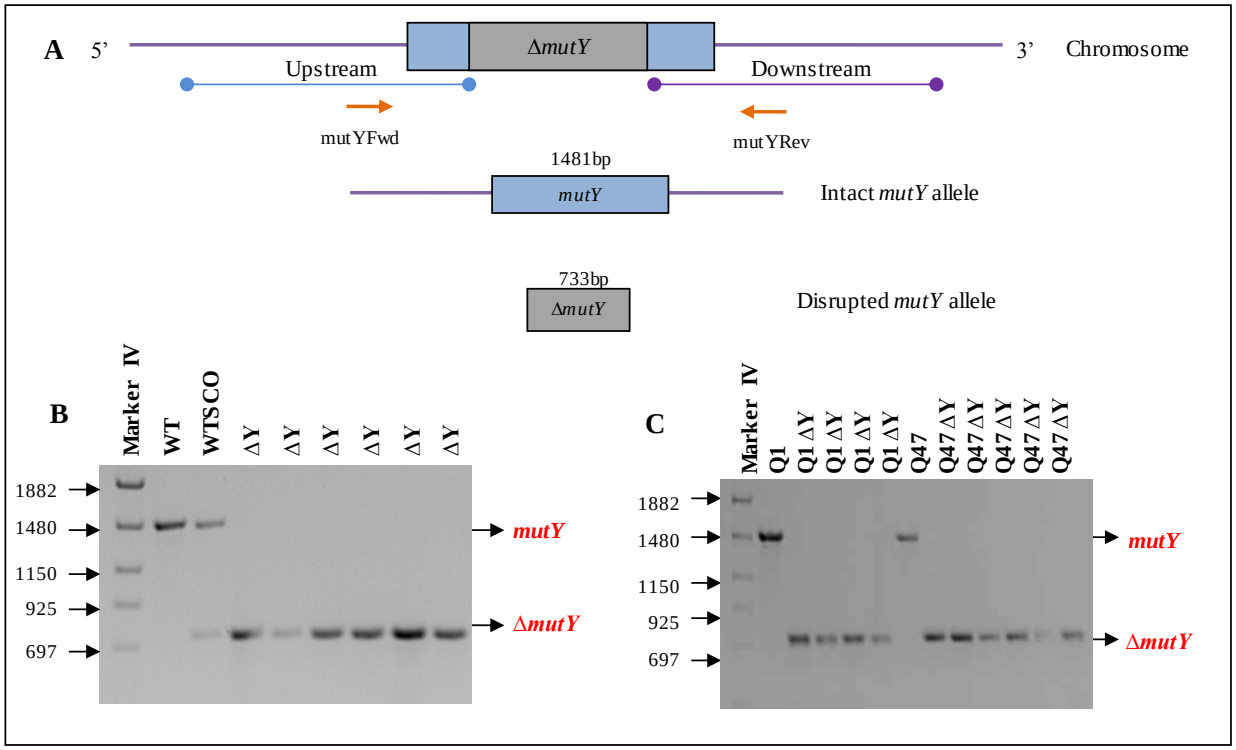


Figure 3.11: PCR screening strategy for the confirmation of DCOs. A. Primer set indicated in the orange arrows allowed for the distinction between an intact copy of the gene (1481bp) and a disrupted copy of the gene (733bp). **B and C.** Taq polymerase PCR-amplification of chromosomal DNA from the wild type (WT) or Fpg/Nei deficient mutants (Q1 and Q47) and their respective *mutY* deficient mutants (ΔY, Q1ΔY and Q47ΔY).

3.4.2 Southern blot analysis

Southern blotting was carried out as detailed in section 2.3.11 which is a more accurate analysis of site-specific integration of the deleted allele and it further confirms the genomic stability of the sequences upstream and downstream of the deletion. The US and DS homologous

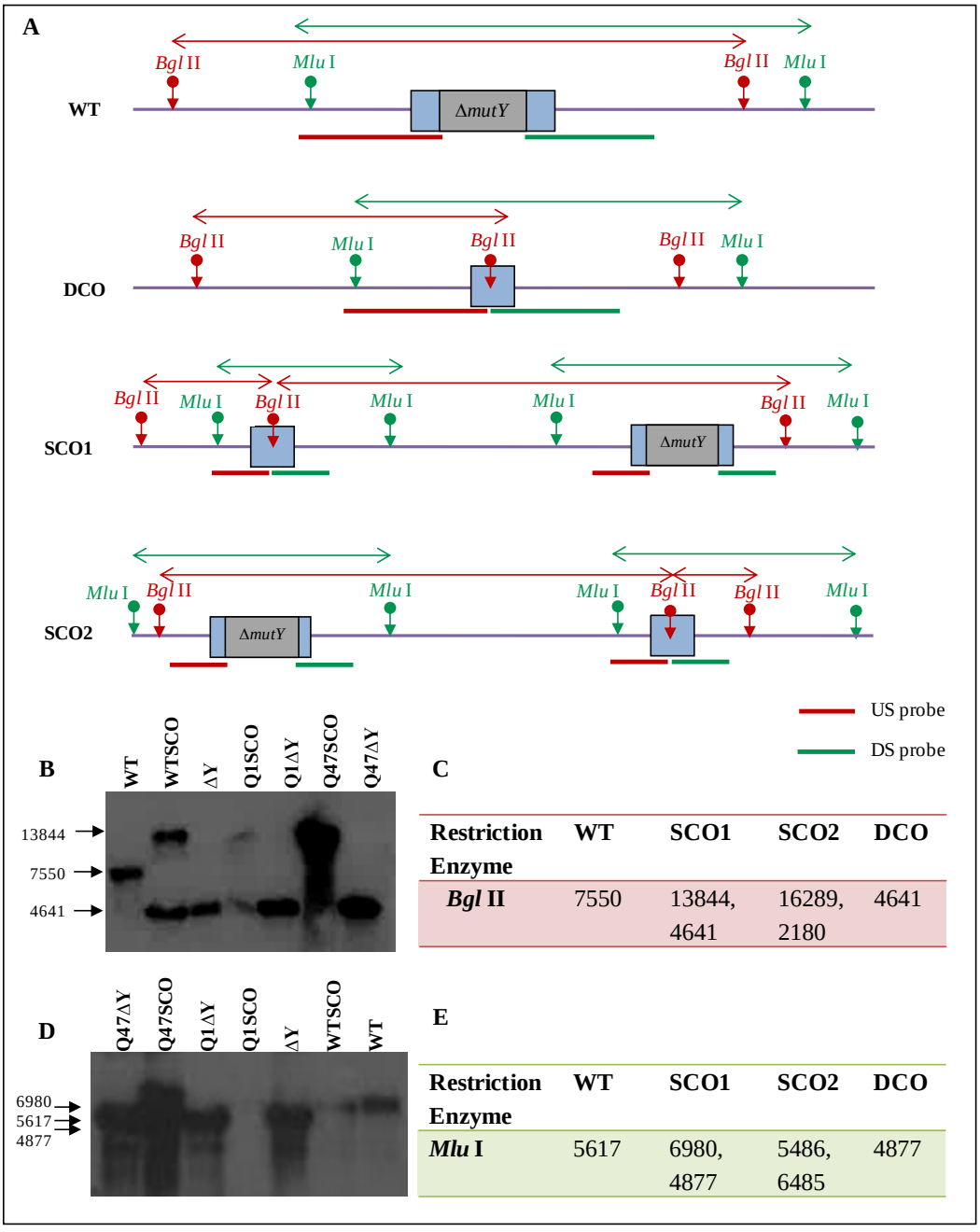
regions were probed separately with two different probes generated by the amplification of the US and DS homologous fragments (described in section 2.3.11.2 and shown in Figure 2.1).

For the US region the genomic DNA was digested with *Bgl* II and the separated DNA was transferred onto a nitrocellulose membrane and probed with a chemiluminescently labeled US probe to identify the respective predicted sizes (Figure 3.12A). The DS region was screened in the same manner but the genomic DNA was digested with the restriction enzyme *Mlu* I and probed with the DS probe (Figure 3.12A).

The expected fragment sizes identified by the US probe were in accordance with the expected fragments as predicted by the sequence map. The WT resulted in a 7550bp fragment and all the DCOs (*mutY* mutants) produced a smaller fragment of 4641bp confirming the replacement of the intact *mutY* allele with a nonfunctional allele (Figure 3.12B). However, for the SCOs the expected fragment sizes were obtained only for the WTSCO (SCO1- 13844bp and 4641bp) while the genomic DNA for the Q47SCO was partially digested hence the expected fragment sizes were not observed (Figure 3.12B).

With the DS probe, the WT produced the expected 5617bp fragment and all the DCOs produced a smaller fragment size of 4877bp confirming that the *mutY* allele was deleted in the respective strain backgrounds (Figure 3.12D). Once again the digested DNA for all the SCO mutants did not resolve very well making it difficult to confirm the expected fragment sizes, even though several attempts were made to re-extract the chromosomal DNA to obtain higher yield of DNA that will allow for detection of the fragments. Since the SCO mutants were correct by PCR analysis (data not shown) and were required only for the generation of the DCO in the two step selection process during homologous recombination, the Southern blot analysis of the

SCO was not pursued further as the resulting DCOs were all correct with both the US and DS regions in the correct genotypic context.



C

Restriction Enzyme	WT	SCO1	SCO2	DCO
***Bgl* II**	7550	13844, 4641	16289, 2180	4641

E

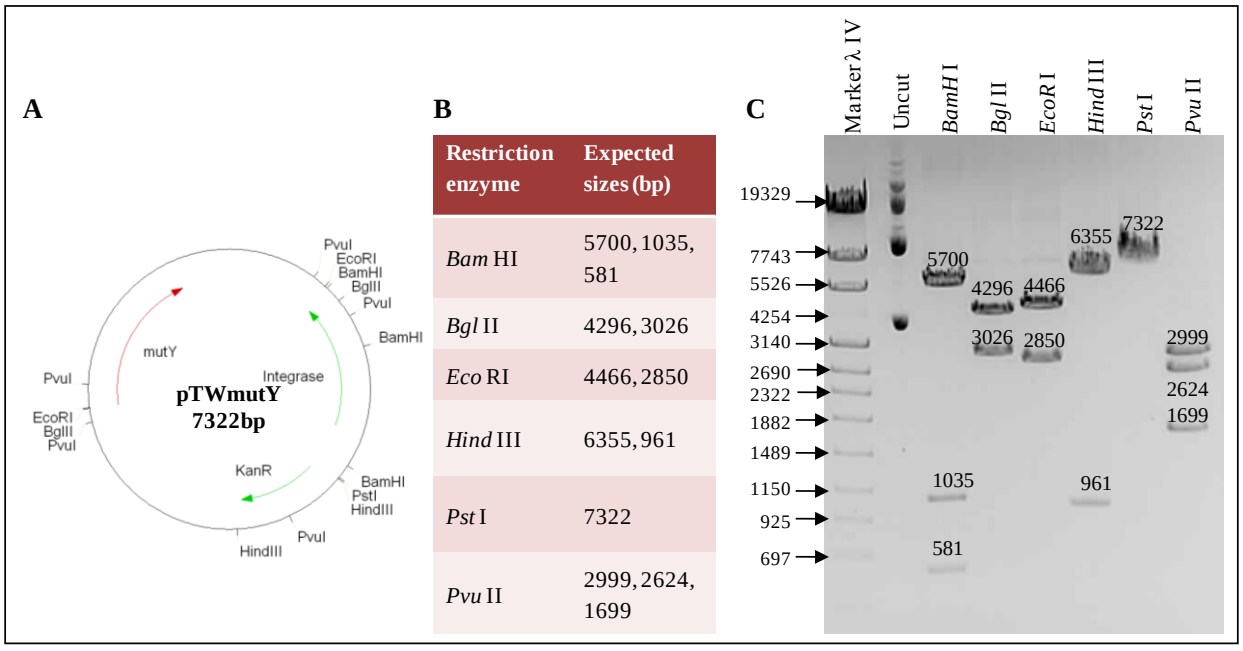
Restriction Enzyme	WT	SCO1	SCO2	DCO
***Mlu* I**	5617	6980, 4877	5486, 6485	4877

Figure 3.12: Southern blot strategy and genotypic verification of the upstream and downstream regions. A. The wild type, *mutY* mutants (DCOs), SCO1 and SCO2 strains generated after the single recombinational event in the respective strains. The chromosomal DNA of the parental, SCO and DCO was digested with either *Bgl* II (probed with the upstream fragment indicated by the red line – US probe) or *Mlu* I (probed with the downstream fragment indicated by the green line – DS probe) for confirmation of the upstream and downstream regions respectively. **B** – Southern blot analysis of the upstream region and **D –** Southern blot analysis of the downstream region. The chromosomal DNA of the different strains were digested with the respective enzyme and their expected sizes are indicated in the table alongside each blot (**C and E**).

Since, the mutants generated in this study were undergoing phenotypic characterization using several *in vitro* assay conditions, it was important to complement the mutant strains with an intact copy of the *mutY* gene to confirm that any observed phenotype is due to the deleted *mutY* allele and not due to downstream polar effects.

3.4.3 Complementation of *mutY* derivative

The complementation vector carrying the intact copy of the *mutY* gene was constructed as described in section 2.5.6. Primers (Table 5.1.2 Appendix A) were designed 200bp upstream of the *mutY* start codon to ensure that the native promoter region for expression of the gene was included. The amplified product was ligated at the *Sca* I site of the pTWEETY integrating vector (Figure 5.2.7 Appendix B), followed by transformation into DH5α and a positive clone was selected and screened by restriction map analysis. All the digests corresponded to the expected sizes (Figure 3.13).



Restriction enzyme	Expected sizes (bp)
Bam HI	5700, 1035, 581
Bgl II	4296, 3026
Eco RI	4466, 2850
Hind III	6355, 961
Pst I	7322
Pvu II	2999, 2624, 1699

Figure 3.13: Complemented vector containing the *mutY* gene. A. Restriction map of the pTWmutY vector, the *mutY* gene depicted in red was cloned into the pTWEETY integrating vector. **B.** Expected sizes of fragments for a selection of restriction enzymes used for mapping. **C.** Restriction digests corresponding to the expected sizes.

Furthermore, the vector DNA was sequenced using the primers listed in Table 5.1.2 (Appendix A) to ensure that no mutations were introduced during the PCR amplification (data not shown). Once the integrating vector with the functional *mutY* gene was confirmed to have no mutations it was electroporated into the following strains ΔY, Q1ΔY and Q47ΔY and the integrants were selected on kanamycin. The complemented strains (ΔY:Y, Q1ΔY:Y and Q47ΔY:Y), parental (WT) as well as their respective *mutY* deficient mutants (ΔY, Q1ΔY and Q47ΔY) were confirmed by PCR using the same strategy as described in Figure 3.11 in which the wild type and mutants *mutY* alleles were distinguished based on size difference in the complemented strains (Figure 3.14). As the PCR preferentially amplified the smaller deleted allele which is expected when the sizes between the two bands is almost double, semi-quantitative RT PCR was carried out to further confirm that the *mutY* gene was expressed in the various complemented strains (Figure 3.15).

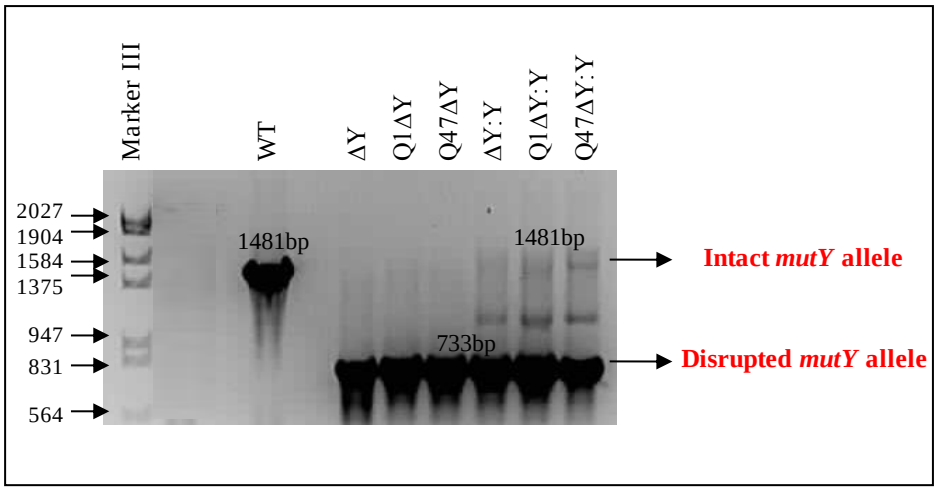


Figure 3.14: PCR confirmation of complemented strains. Represented as the intact copy of the *mutY* gene (1481bp) and the disrupted copy of the gene (733bp) is amplified.

RT PCR was conducted as described in section 2.5.7. RNA from each of the strains (parental, the Fpg/Nei deficient mutants, *mutY* deficient mutants and complemented strains) was

extracted and the reverse transcribed cDNA was PCR amplified using Taq polymerase. Expression of the *mutY* gene in the wild type or Fpg/Nei deficient mutant strains was to a level equivalent to 10^6 genomes (Figure 3.15), while the *mutY* deficient mutant strains (ΔY, Q1ΔY and Q47ΔY) showed no expression confirming the loss of *mutY* in these strains and that the flanking 100bp retained during the homologous recombination process is not sufficient for the expression of a functional *mutY* gene. Interestingly, the ΔY:Y complemented strain displayed a lower level of expression which was equivalent to $10^5 - 10^4$ genomes this could possibly be due to one of two reasons, firstly the PCR amplification of the intact allele did not amplify as compared to the Q1ΔY:Y and Q47ΔY:Y complemented strains. Secondly, the functional gene was not integrated at the original locus thus the reduced expression. However, in comparison the same construct restored *mutY* expression in the Q1ΔY:Y and Q47ΔY:Y strains to equivalent levels as observed for the Fpg/Nei deficient mutant strains (Q1 and Q47). The reason for this differential expression of the same gene in the three complemented strains is unknown and requires further investigation of additional clones.

Three sets of controls were included, the no reverse transcriptase (-cDNA) to indicate that the RNA samples had not been reverse transcribed, the no template control and the *sigA* housekeeping gene was expressed as a positive control. No amplification was detected for the two negative controls indicating that there was no genomic DNA contamination introduced during the RNA extraction process or during the PCR preparation.

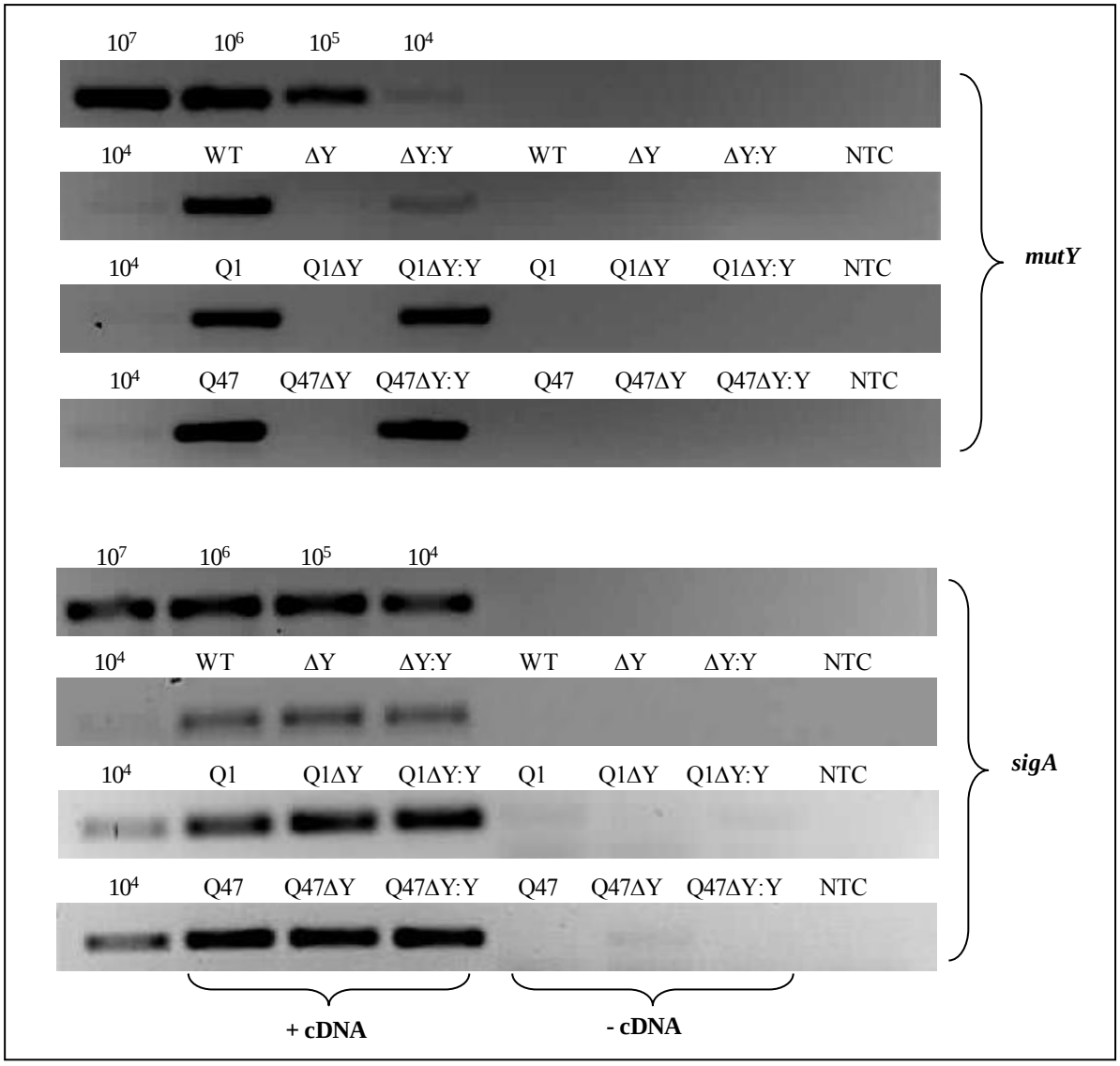


Figure 3.15: Real Time PCR assessment of *mutY* gene expression. The various *mutY* mutant strains (ΔY, Q1ΔY and Q47ΔY) showed no expression of the *mutY* gene as compared to their respective wild type (WT) or Fpg/Nei deficient mutants (Q1 and Q47) and complemented strains (ΔY:Y, Q1ΔY:Y and Q47ΔY:Y). The expression of the *sigA* gene was used as a positive control.
+cDNA (reverse transcribed to cDNA) and –cDNA (not reverse transcribed to cDNA) (control)

3.5 Phenotypic characterization

The Fpg/Nei deficient mutant strains Q1 and Q47 were generated previously (Goosens *et al.*, unpublished) from two different lineages. In the Q1 strain the two Fpg DNA glycosylases were inactivated first followed by the deletion of the two Nei glycosylases and in the Q47 strain

the order in which these glycosylases were inactivated was reversed (Table 2.1). These mutant strains showed no notable differences in growth under normal culture conditions, but mutants lacking three and four of these glycosylase homologues displayed a 2-3 log difference in survival under oxidative stress conditions compared to the parental strain. Lack of this phenotype in the single and double mutants suggests that in the absence of the Fpg DNA glycosylases the Nei glycosylases are able to take over the function and vice versa to protect the organism against oxidative damage.

Sequence analysis of the *E. coli* genome has shown that the *mutM* gene of the GO system and the *fpg* gene in the Fpg/Nei family of DNA glycosylases are identical thus these genes may have overlapping and interchangeable roles in the BER pathway and the GO system (Michael *et al*., 1991). Furthermore, in *E.coli* when both *mutY* and *mutM* (*fpg*) were deleted there was an increased mutator phenotype compared to the individual mutations (Michael and Miller, 1992). Hence, given the redundancy of the mycobacterial Fpg/Nei family of DNA glycosylases it was imperative to assess the role of the duplicated Fpg and Nei homologues in context of the single *mutY* gene in genome maintenance and to establish whether MutY may be playing a compensatory role in the absence of Fpg/Nei repair enzymes. Hence, *mutY* was deleted in the parental *M. smegmatis* strain and in both the Fpg/Nei deficient mutant strains (Q1 and Q47) to assess its role in mutagenesis.

These various *mutY* deficient mutant strains generated in this study were assessed *in vitro* for differences in growth kinetics under normal culture conditions as well as for their sensitivity to oxidative stress as generated by hydrogen peroxide and cumene peroxide. Progressive loss of DNA repair enzymes is expected to increase mutations, hence the spontaneous mutation rates to rifampicin was also determined for each of the mutant strains and the fold increase was

calculated in relation to the rate obtained for the parental strain. The mutational spectra for a representative number of rifampicin resistant colonies isolated from each of the mutant strains was analyzed to identify the types of mutations that resulted in rifampicin resistance. Collectively, these data were anticipated to provide further insights into the mechanisms and the role of these various DNA glycosylases in DNA repair under oxidative stress conditions.

3.5.1 Growth kinetics

Log phase cultures of the parental, Fpg/Nei deficient mutants and the various *mutY* deficient mutant strains were diluted a 100 fold in 7H9 and the growth monitored by optical density measurements every 3 hrs for a period of 36 hrs (described in section 2.6.1). In two independent experiments the growth of the wild type (WT), Fpg/Nei deficient mutants (Q1 and Q47) and the *mutY* deficient mutant strains (ΔY, Q1ΔY and Q47ΔY) were comparable (Figure 3.16), suggesting that loss of the entire Fpg/Nei family of DNA glycosylases and/or the *mutY* glycosylase does not affect the growth of *M. smegmatis*. The *in vitro* culture conditions were optimal for growth and probably caused only a basal amount of DNA damage as part of the organism's respiratory and metabolic activities, thus the remaining repair enzymes in the BER pathway maybe compensating and allowing the organism to efficiently handle the damage. In addition mycobacteria possess several genes that detoxify reactive oxygen species (ROS) such as catalases, peroxidases and superoxide dismutases (SODs). *katG* is an example of a catalase-peroxidase that is continuously primed in the pathogen to neutralize free radicals that are generated during normal aerobic growth within the host environment (Dussurget *et al*., 1998). Thus, prior to DNA repair these enzymes probably contribute as the first line of defense and survival of the organism under oxidative stress conditions.

It was important to understand the role of Fpg/Nei and MutY DNA glycosylases under oxidative conditions as encountered by the organism in the human host during pathogenesis. Hence, the various mutant strains were tested under *in vitro* oxidative culture conditions as mimicked by hydrogen peroxide.

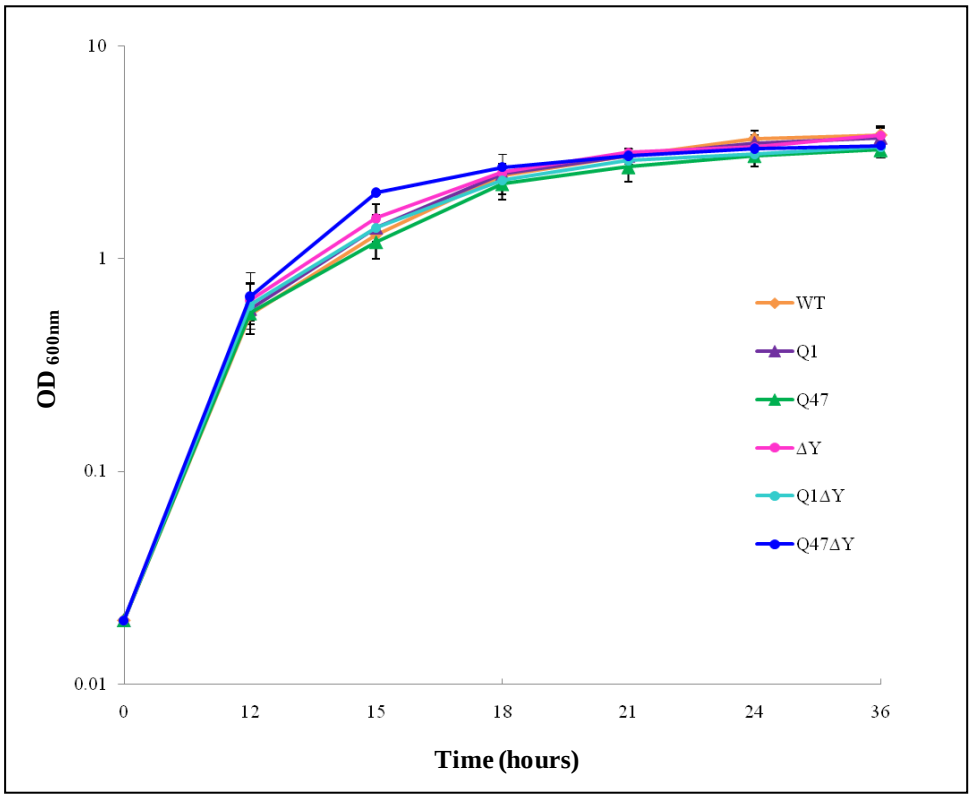


Figure 3.16: Growth kinetics of the parental and mutant strains under normal culture conditions. The wild type (WT), Fpg/Nei deficient mutants (Q1 and Q47), MutY deficient mutant (ΔY) and Fpg/Nei/MutY deficient mutants (Q1ΔY and Q47ΔY) growth was monitored every three hours as cfu/ml over a thirty six hour period. The data shows the mean of two independent experiments, with error bars representing the standard error.

3.5.2 Survival kinetics under oxidative stress induced by hydrogen peroxide

Hydrogen peroxide although a light sensitive and a relatively unstable compound is the most commonly used chemical agent to generate oxidative stress in microorganisms. During the generation of the Fpg/Nei deficient mutants the hydrogen peroxide concentration that resulted in a 3 log kill in the parental *M. smegmatis* strain was established as 2.5 mM (Goosens, 2008).

Hence, in this study log phase cultures (OD_{600nm} of 0.35) of the parental, Fpg/Nei deficient mutants, the various *mutY* deficient mutants and complemented strains were exposed to 2.5 mM hydrogen peroxide as described in section 2.6.3. The treated cells were monitored for survival by assessing cell viable counts (log cfu/ml) every 2 hrs over a 6 hr period. The *mutY* deficient mutant (ΔY) strain showed sensitivity to hydrogen peroxide relative to the wild type (WT) strain with a 2 log reduction of viable cells after 6 hrs (Figure 3.17). The ΔY:Y complemented strain only partially restored the phenotype of the ΔY mutant strain possibly due to the reduced level of *mutY* expression in the complemented strain compared to the parental strain as assessed by RT PCR shown in Figure 3.15 above. The expression data showed about a 50 fold reduction in *mutY* expression for the complemented strain compared to the parental strain and there is a correlated approximate 50% complementation of the growth recovery with the ΔY:Y complemented strain.

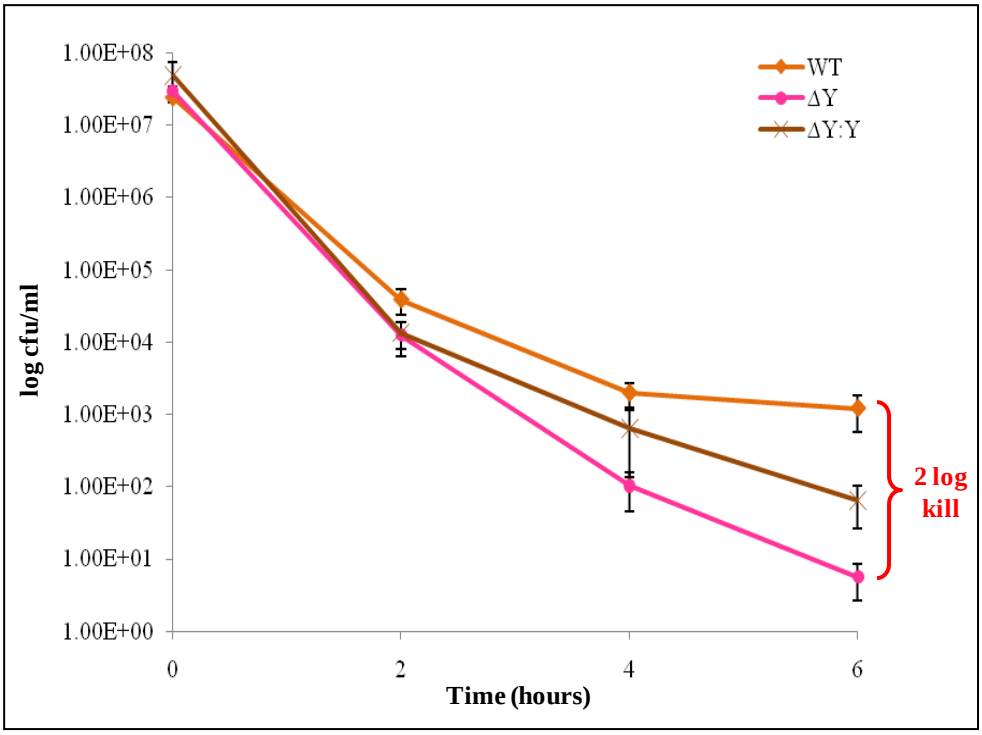


Figure 3.17: Survival kinetics of the parental and mutant strains under oxidative stress conditions. The wild type (WT), MutY deficient mutant (ΔY) and the complemented strain (ΔY:Y) were treated with 2.5 mM hydrogen peroxide and growth monitored every two hours as cfu/ml over a six hour period. The error bars represent the standard error between 5 technical repeats for the WT and ΔY strain while the ΔY:Y strain represents 3 technical repeats.

The Fpg/Nei deficient Q1 mutant also showed an approximate 2 log reduction in cfu compared to the WT strain in the presence of 2.5 mM hydrogen peroxide. This was exaggerated by a further log reduction in cfu (Figure 3.18), with the inactivation of *mutY* (Q1ΔY) in this background suggesting an antimutator role for MutY. The Q1ΔY:Y complemented strain rescued the reduced survival phenotype almost equivalent to the WT *M. smegmatis* strain although the survival was expected to be similar to that of the Q1 strain.

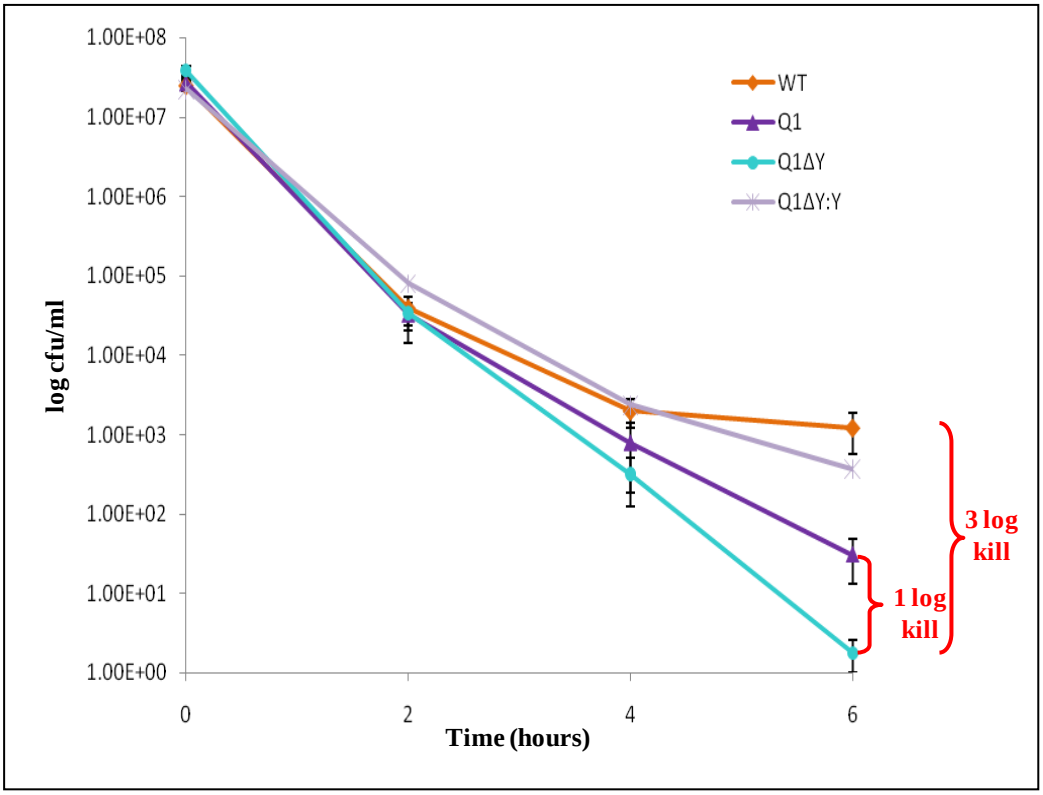


Figure 3.18: Survival kinetics of the parental and mutant strains under oxidative stress conditions. The wild type (WT), Fpg/Nei deficient mutant (Q1), Fpg/Nei/MutY deficient mutant (Q1ΔY) and the complemented strain (Q1ΔY:Y) were treated with 2.5 mM hydrogen peroxide and growth monitored every two hours as cfu/ml over a six hour period. The error bars represent the standard error between 5 technical repeats for the WT, Q1 and Q1ΔY while the complement strain represents a single experiment.

Similarly, Q47ΔY also displayed a greater reduction (3 log reduction) in cfu compared to Q47. Q47 showed less than a log difference in cfu in comparison to the WT strain. The Q47ΔY:Y complemented strain was able to restore the growth defect to that of Q47 (Figure 3.19) as expected. These data clearly suggest that *mutY* plays a more significant role in DNA damage repair induced under oxidative stress conditions. It is also possible that in the absence of both the Fpg/Nei homologues *mutY* is able to deal with the DNA damage, but when *mutY* is inactivated together with the Fpg/Nei family of DNA glycosylases the repair system becomes less effective. The differential survival of the two Fpg/Nei deficient mutants (Q1 and Q47) generated from

different lineages under oxidative stress conditions is also indicative that a hierarchy may exist amongst these DNA repair enzymes.

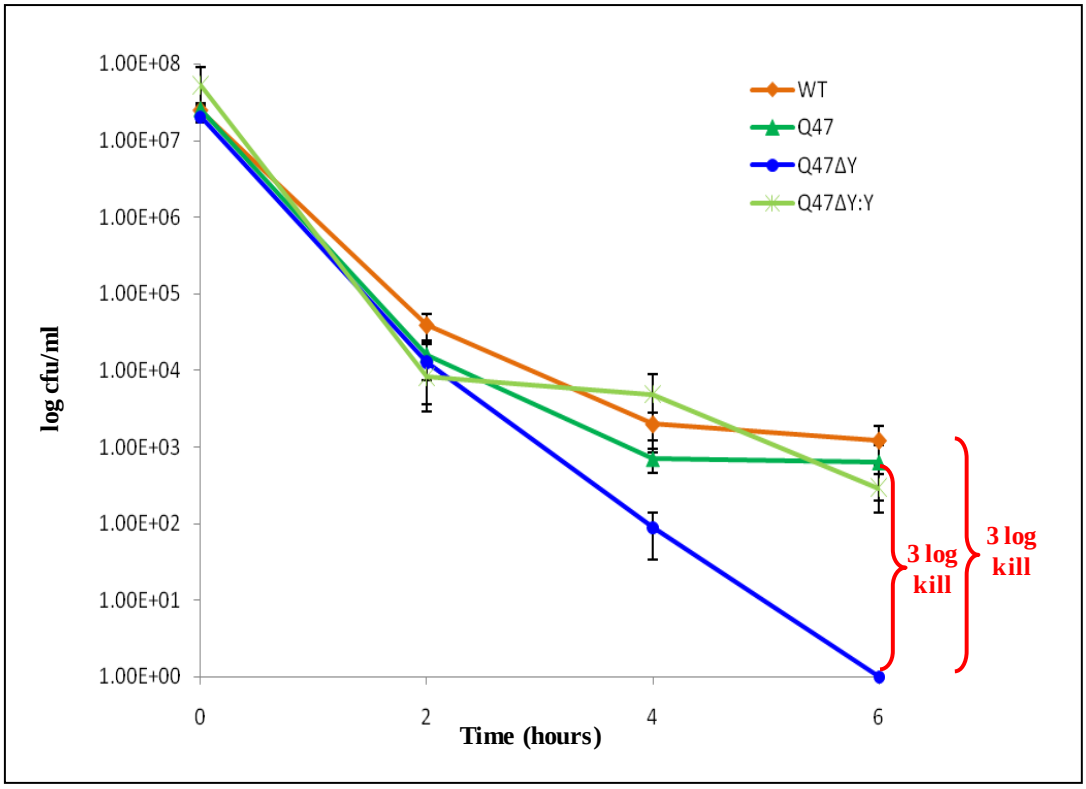


Figure 3.19: Survival kinetics of the parental and mutant strains under oxidative stress conditions. The wild type (WT), Fpg/Nei deficient mutant (Q47), Fpg/Nei/MutY deficient mutant (Q47ΔY) and the complemented strain (Q47ΔY:Y) were treated with 2.5 mM hydrogen peroxide and growth monitored every two hours as cfu/ml over a six hour period. The error bars represent the standard error between 5 technical repeats for the WT, Q47 and Q47ΔY, while the Q47ΔY:Y strain represents 3 technical repeats.

Hydrogen peroxide as an oxidative reagent that requires a high level of technical standardization and experiments had to be repeated at least 5 or more times to obtain uniform consistent data. Hence, an alternate agent, cumene peroxide which is a more stable agent was considered for *in vitro* induction of oxidative stress.

3.5.3 Survival kinetics under oxidative stress induced by cumene peroxide

A concentration of cumene peroxide that results in at least a three log kill in the parental strain was established by spot testing an early log phase culture on solid 7H10 media containing

different concentrations of cumene peroxide as described in section 2.6.2. At 2.0 mM there was about a one log kill of the parental strain which was greatly enhanced when the concentration of cumene peroxide was increased to 2.5 mM resulting in greater than 5 log kill of *M. smegmatis* (Figure 3.20). The kill kinetics of the parental strain was then assessed in liquid media at 2.5 mM concentration to establish whether the constituent of the media can influence the uptake of the compound by the cells. Surprisingly, 2.5 mM cumene peroxide in liquid media showed no significant killing effect in liquid media (data not shown) and therefore a higher concentration (3.8 mM) was assessed.

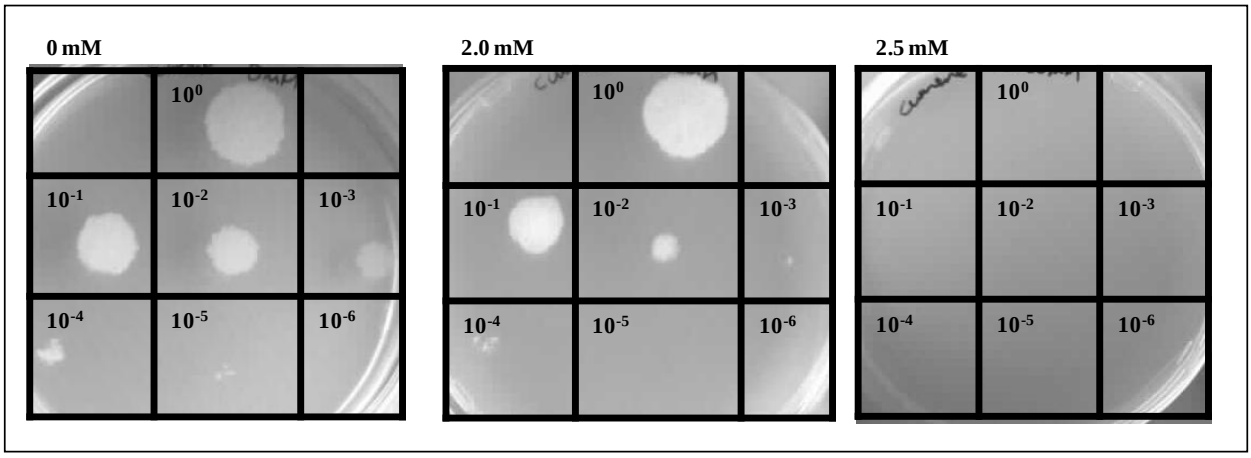


Figure 3.20: Growth of *M. smegmatis* on various concentrations of cumene peroxide. An early log phase culture of the parental (mc^2155) was diluted and spotted on 7H10 media plates containing 0, 2.0 and 2.5 mM cumene peroxide.

Early log phase (OD_{600nm} of 0.2) cultures of the parental, Fpg/Nei deficient mutants, the *mutY* mutants and complemented strains were exposed to 3.8 mM cumene peroxide (described in section 2.6.3) and the growth monitored at 0, 3 and 6 hrs as viable counts (cfu/ml). There was only a ½ log difference between the WT and ΔY mutant strain compared to the 2-3 log kill observed for these strains under hydrogen peroxide. Once again the ΔY:Y complemented strain was unable to completely restore the phenotype to that of the WT strain (Figure 3.21) as

observed with the hydrogen peroxide assay, confirming that the lack of complementation corresponds to the reduced expression of *mutY* and is not restricted to specific mutants.

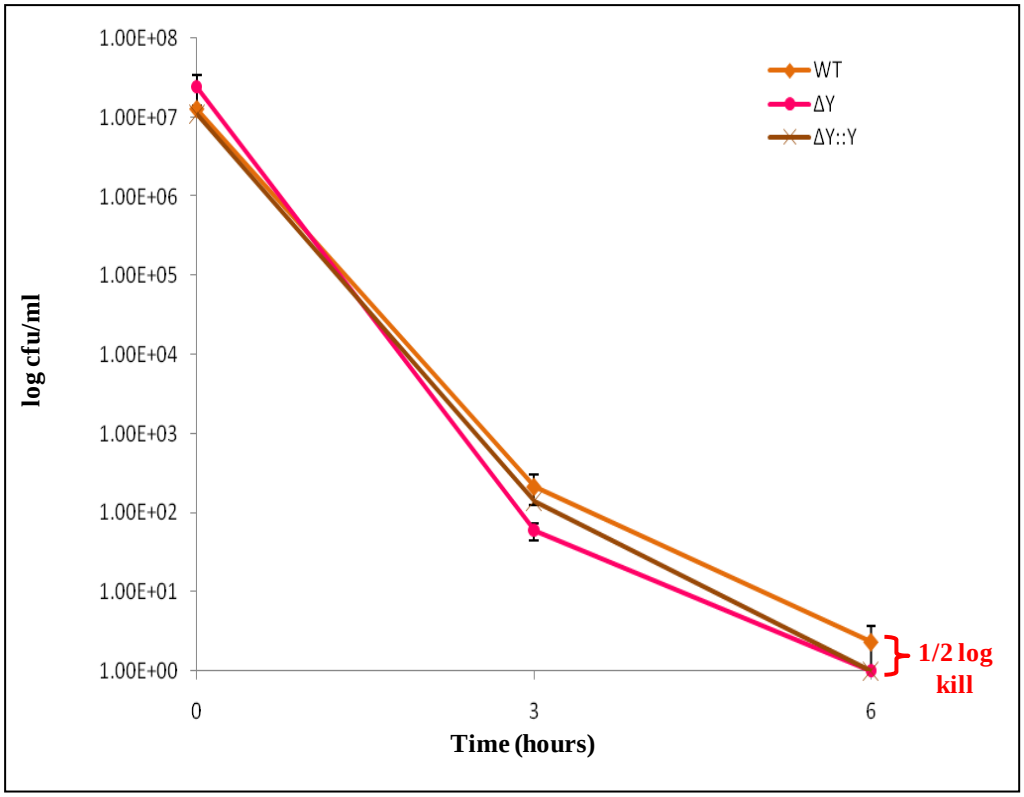


Figure 3.21: Survival kinetics of the parental and mutant strains under oxidative stress conditions. The wild type (WT), MutY deficient mutant (ΔY) and the complemented strain (ΔY:Y) were treated with 3.8 mM cumene peroxide and growth monitored every three hours as cfu/ml over a six hour period. The error bars represent the standard error between 3 technical repeats for the WT and ΔY strain, while the ΔY:Y strain represents a single experiment.

The Q1ΔY mutant relative to the Q1 and WT strains displayed similar growth survival kinetics as the ΔY mutant strain when exposed to cumene peroxide (Figure 3.22) and the Q1ΔY:Y complemented strain restored the growth comparable to Q1. Under these conditions the combinatorial deletion of the Fpg/Nei and MutY glycosylases did not impact on the survival of the organism, suggesting that anti-oxidant mechanisms are able to detoxify the oxidative radicals or perhaps that other repair enzymes in the BER pathway are poised to deal with the induced damage. It is possible that the cumene peroxide is generating DNA damage lesions that are

recognized and dealt with by BER and NER repair enzymes other than *fpg/nei* and *mutY*. Cumene peroxide is definitely entering the mycobacterial cell wall and causing the necessary damage as there is a 5 log kill irrespective of the genotype of the strains after 3 hrs of exposure, with a notable effect after 6 hrs as the mycobacteria are unable to survive the damage.

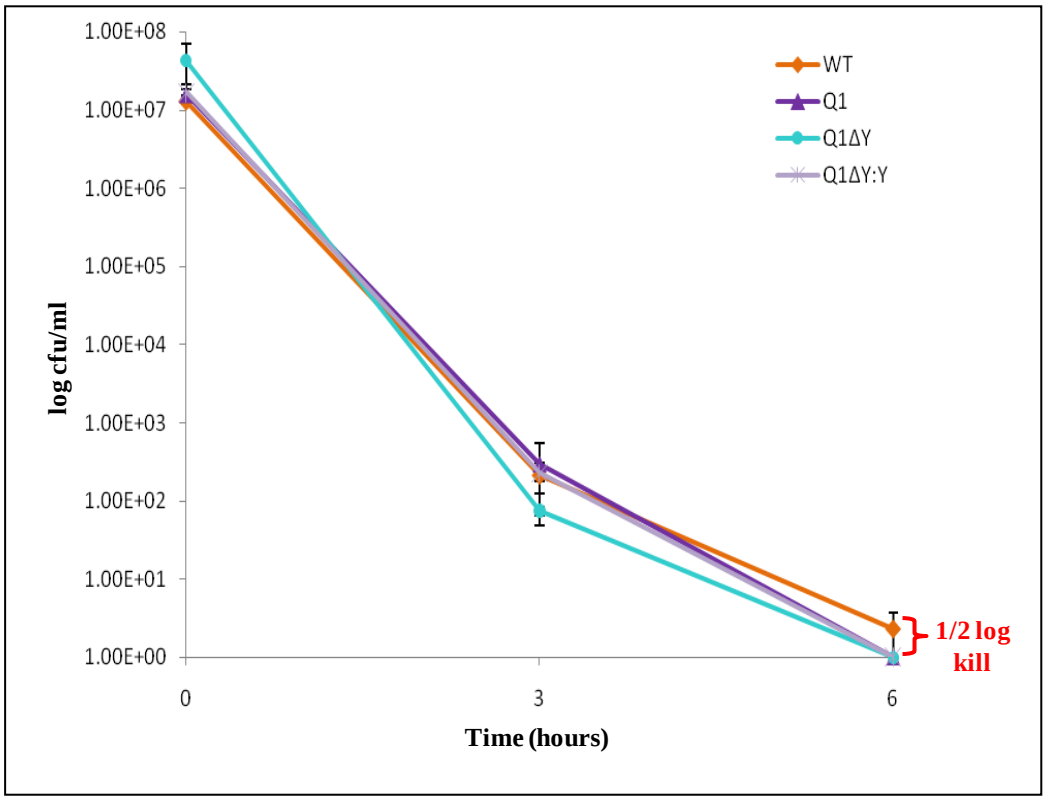


Figure 3.22: Survival kinetics of the parental and mutant strains under oxidative stress conditions. The wild type (WT), Fpg/Nei deficient mutant (Q1), Fpg/Nei/MutY deficient mutant (Q1ΔY) and the complemented strain (Q1ΔY:Y) were treated with 3.8 mM cumene peroxide and growth monitored every three hours as cfu/ml over a six hour period. The error bars represent the standard error between 3 technical repeats for the WT and Q1, 2 technical repeats for the Q1ΔY, while the Q1ΔY:Y strain represents 3 technical repeats.

However, in contrast the Q47ΔY combinatorial mutant showed increased sensitivity to cumene peroxide with a 1.5 and 2 log decrease in survival after 3 hrs of exposure relative to the Q47 and WT strains respectively (Figure 3.23). This growth defect was restored by the Q47ΔY:Y complemented strain to levels observed for the Q47 strain. The Q47 Fpg/Nei deficient mutant displayed the same survival phenotype in response to both cumene peroxide and

hydrogen peroxide indicating that the loss of the Fpg/Nei genes in this lineage seems to have a greater impact on survival of *M. smegmatis*. Further investigations are needed such as whole genome sequencing of the mutants to uncover the underlying molecular mechanisms that make the two Fpg/Nei deficient mutants different.

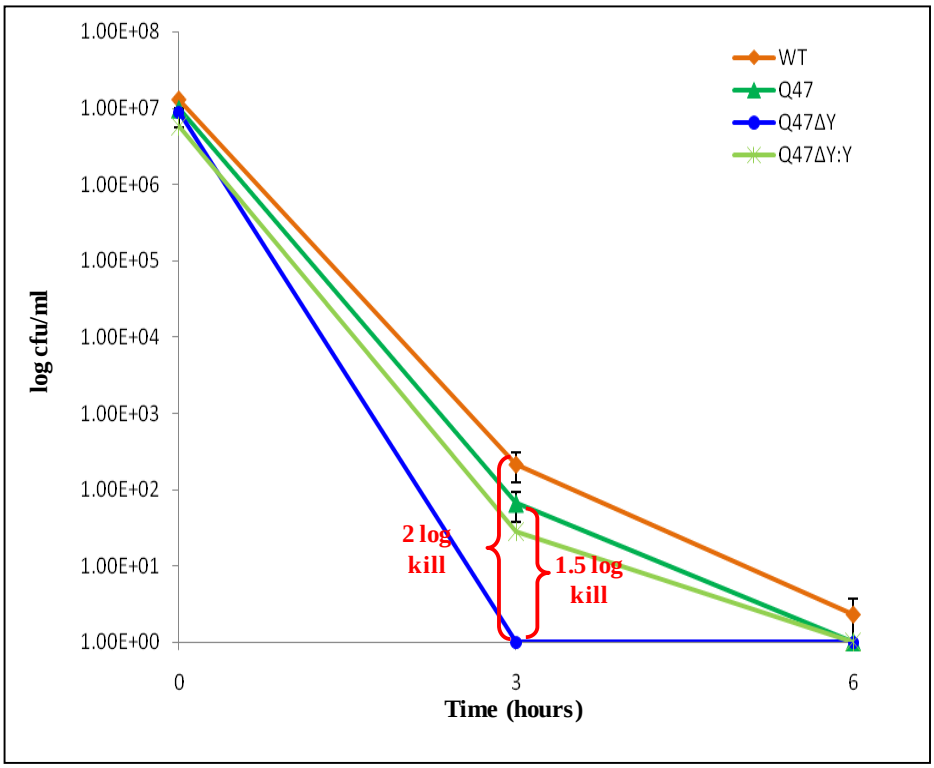


Figure 3.23: Survival kinetics of the parental and mutant strains under oxidative stress conditions. The wild type (WT), Fpg/Nei deficient mutant (Q47), Fpg/Nei/MutY deficient mutant (Q47ΔY) and the complemented strain (Q47ΔY:Y) were treated with 3.8 mM cumene peroxide and growth monitored every three hours as cfu/ml over a six hour period. The error bars represent the standard error between 3 technical repeats for the WT and Q47, 2 technical repeats for the Q47ΔY, while the Q47ΔY:Y strain represents 3 technical repeats.

The exaggerated decrease in survival observed with deletion of the *mutY* gene in combination with the Fpg/Nei DNA glycosylases confirms that MutY plays a more significant role as an antimutator in DNA repair for mycobacterial genome stability and maintenance. The duplication of the Fpg/Nei homologues may be a deliberate event as an enhanced survival mechanism for mycobacteria to protect their genome under harsh environments.

Other oxidative stress agents such as menadione (a generator of superoxide) and tert-butyl hydroperoxide (a membrane-permeant oxidant) were also tested because of their stable properties but more importantly to observe if these compounds display a better phenotypic effect of DNA damage tolerance by the various DNA repair enzyme deficient mutants.

3.5.4 Optimization of menadione and tert-butyl hydroperoxide

As described in section 2.6.2 the menadione and tert-butyl hydroperoxide optimum concentrations were determined by spot testing dilutions of an early log phase culture of the parental *M. smegmatis* on solid 7H10 media containing different concentrations of these agents. For menadione there was a 4 log kill with 10 µM whilst 100 µM tert-butyl hydroperoxide displayed a 3 log reduction in the *M. smegmatis* parental strain (data not shown). Interestingly, menadione displayed the same kill activity in liquid media at the established concentration of 50 and 100 µM with a 3.5 and 4 log kill respectively (Figure 3.24). Due to time constraints the various mutants could not be tested under these conditions. Since the killing activity for tert-butyl hydroperoxide differed to that established on solid media it requires further optimization (data not shown).

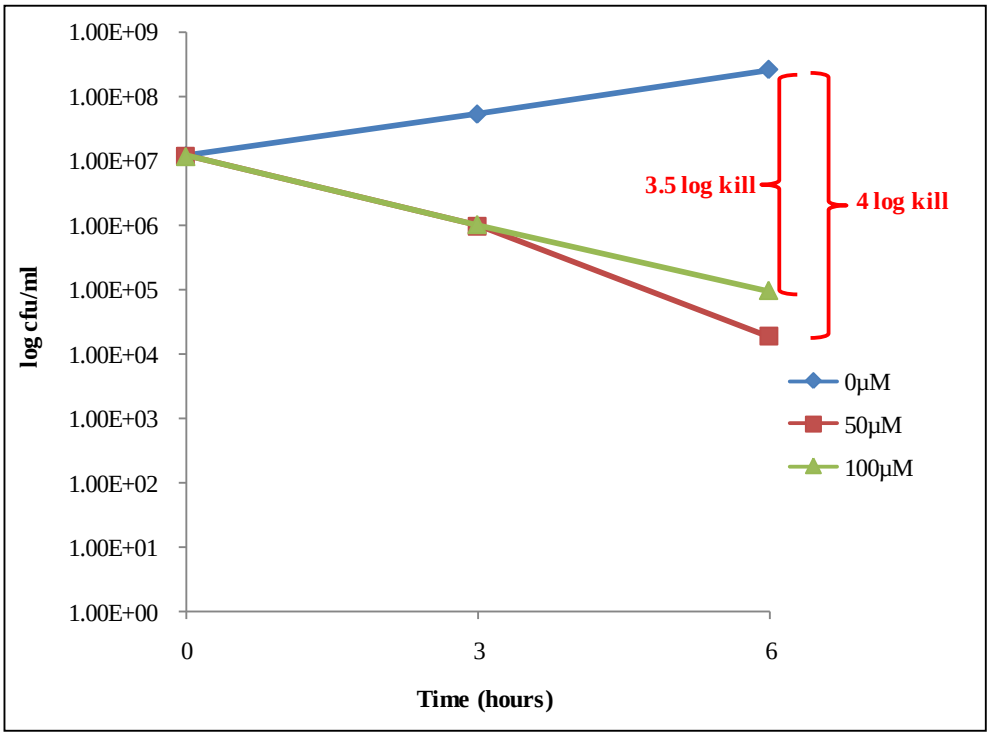


Figure 3.24: The optimal concentration of menadione in *M. smegmatis* liquid culture. Early log phase cultures were exposed to 0, 50 and 100 µM menadione.

Loss of DNA repair enzymes is expected to increase the number of mutational events in the genome with subsequent replication cycles until eventually the organism is unable to maintain genome integrity leading to the death of the organism. Hence, it was important to assess the spontaneous mutagenesis levels of the various DNA glycosylase deficient mutants constructed in this study together with the previously generated Fpg/Nei deficient mutants. In addition, the mutational spectra of the resistant mutants were also analyzed to establish whether there was a correlation between the progressive loss of DNA glycosylases and an increase in the mutation types resulting in resistance.

3.5.5 Analysis of mutation rates

The mutation rates were calculated using the Luria-Delbrück fluctuation analysis (as described in section 2.6.4) which is based on the probability of an organism developing a mutation per cell division in the lifetime of the bacterial cell. The fluctuation assay was set up with a number of parallel cultures (30 tubes) each containing the same number of cells which were allowed to grow and 5 of the 30 tubes were selected to determine the total number of cells. The remaining culture was plated on 7H10 media containing rifampicin. The mutation rates were calculated for each strain using the computer software program developed in the laboratory (Machowski *et al.*, 2007) which is based on the fluctuation analysis equations in Figure 3.25. The P naught method (P_0) was used since the m values were between 0.3 and 2.3 and the mutation rate (μ) was calculated using Equation 3, where m is the number of mutations per culture (Equation 2) which is estimated from P_0 (Equation1) and Nt is the final number of cells in the culture (Roche and Foster, 2000).

Equation 1:	**Equation 2:**	**Equation 3:**
$P_0 = \frac{\text{number of plates containing zero colonies}}{\text{total number of plates}}$	$m = -\ln x\ P_0$	$\mu = \frac{m \times \ln 2}{Nt}$

Figure 3.25: The P_0 method of calculating mutation rates uses equation 1, 2 and 3. (Rosche and Foster, 2000)

Mutation rates for the parental and various mutant strains were calculated seperately for three independent experiments (Table 3.1). Each of the experiments displayed some variability in the mutation rates which could possibly be due to differences in the Nt values which greatly influences the calculation of the mutation rates. Both the Q1 and Q47 mutants with the entire Fpg/Nei family of DNA glycosylases inactivated showed an increased mutation rate compared to the WT in two of the experiments, but the *mutY* deficient mutant strain (ΔY, indicated with a red

circle) displayed increased mutagenesis in experiment 2 only. The mutation rate is significantly increased with the inactivation of *mutY* in the Q1 mutant background (indicated by the orange circles) for two experiments, but simultaneous deletion of *mutY* in the Q47 mutant strain displayed a marginal added effect on the mutation rate in experiment 3 only (Q47ΔY, indicated by a blue circle). The Q47ΔY showed reduced survival under oxidative stress conditions therefore, this lack of increased mutagenesis with this mutant was unexpected and needs further investigation. However, the combined data obtained for Q1ΔY provides strong evidence that *mutY* plays an important role during mutagenesis and DNA repair. The data further suggests that there probably is a heirarchy between the various DNA repair enzymes in the BER pathway.

Table 3.1: **Mutation rates calculated from three independent fluctuation assay experiments.** These were determined using the P_0 method.

	Experiment 1		Experiment 2		Experiment 3	
Strain	**Mutation rate/10^7**	**Fold increase**	**Mutation rate/10^7**	**Fold increase**	**Mutation rate/10^7**	**Fold increase**
WT	1.1	1	0.05	1	2.2	1
ΔY	**1.2**	**1.1**	**0.6**	**12**	**2.2**	**1**
ΔY:Y	-	-	0.06	1.2	3.0	1.4
Q1	1.1	1	2.5	50	3.6	1.6
Q1ΔY	**1.0**	**0.9**	**8.0**	**160**	**5.6**	**2.5**
Q1ΔY:Y	-	-	8.3	166	1.2	0.5
Q47	2.7	2.5	5.6	112	2.4	1.1
Q47ΔY	**1.5**	**1.4**	**5.5**	**110**	**3.9**	**1.8**
Q47ΔY:Y	-	-	2.2	44	4.0	1.8

3.5.6 Mutation spectrum

Since there was a general increase in the mutation rates with progressive loss of DNA glycosylases it was important to understand whether there was a correlating increase in the type of mutations that led to rifampicin resistance. Between 16-20 rifampicin resistant colonies isolated in the fluctuation assay for the various Q47 lineage mutants were selected for analysis since in this mutant background, deletion of *mutY* resulted in a far more dramatic decrease in

survival under oxidative stress conditions, although there was no significant correlated increase in mutation rates. Deletion of *mutY* in the Q1 lineage resulted in a one log additive difference in survival under oxidative stress conditions, but displayed a more dramatic increase in mutation rates, therefore these mutants also requires future spectral analysis to delineate the underlying mechanisms responsible for the difference in phenotypes for the two lineages of Fpg/Nei deficient mutants. The rifampicin resistant colonies were PCR amplified and the products were sequenced to identify the type of mutations in the Rifampicin Resistance Determining Region (RRDR) within the *rpoB* gene that resulted in rifampicin resistance. Most of the spontaneous rifampicin resistant mutations have been reported to occur in the 81bp RRDR, between codons 507 and 533 within the *rpoB* gene with the most common mutations shown in red blocks in Figure 3.26 below.

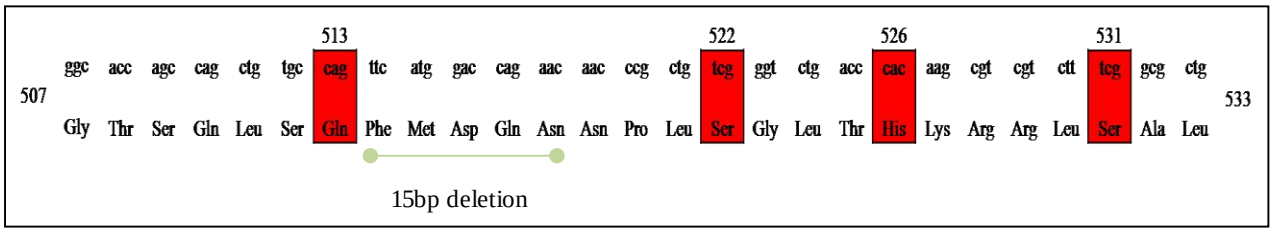


Figure 3.26: Mutations found in the 81bp RRDR within the *rpoB* gene in *M. smegmatis*. The red shaded blocks indicate common mutations within this region and the green line indicates where a 15bp deletion occurs.

There was an increase in C → A mutations (CAG → AAG Gln513 → Lys) from 5% in the WT to 25% in the ΔY strain (highlighted with a red circle Table 3.2), which was in concordance with the data obtained by Kurthkoti *et al.*, (2010). Interestingly, there was a substantial increase of this mutation in the Q47ΔY strain (indicated with a blue circle) with a 70% frequency compared to the 6.25% in the Q47 strain with no mutations in this amino acid for the Q47ΔY:Y complemented strain. There was also an increase of 15% in C → A mutations (CAC → AAC; His526 → Asn) in the Q47ΔY strain as compared to the WT, Q47 and Q47ΔY:Y

strains. These findings correlate with the 3 log kill observed for this mutant under induced oxidative stress conditions (hydrogen and cumene peroxide) and is the first report of such a significant increase in C → A mutations for strains lacking both *mutY* and *mutM* (Q47ΔY) in mycobacteria. Furthermore, the ΔY strain displayed the highest percentage of A → C mutations (75% CAC → CCC His526 → Pro) compared to the WT strain which accounts for the 12 fold increase in mutation rate for this mutant (Table 3.1). This mutation type was not observed in the in the ΔY:Y complemented strain.

Table 3.2: Rifampicin resistant mutations found in the RRDR region displayed as a percentage. A total number of mutants screened for each strain are indicated in brackets and dashed lines indicates that no mutation occurred. Mutations that occur outside of the 81bp region were recorded as unknown. Red rings indicate significant increases in C to A and A to C mutations.

Mutation/Substitution	WT (20)	ΔY (20)	ΔY:Y (20)	Q47 (16)	Q47ΔY (20)	Q47ΔY:Y (20)
AGC → CCC Ser509 → Pro	-	-	-	6.25	-	-
CAG → CTG Gln510 → Leu	-	-	-	6.25	-	-
CAG → AAG Gln513 → Lys	5	25	-	6.25	70	-
CAG → CCG Gln513 → Pro	10	-	-	-	-	-
CAG → GAG Gln513 → Glu	5	-	-	-	-	-
AAC → AAA Asn519 → Lys	-	-	-	-	-	5
TCG → TCT Ser522 → Ser	-	-	-	6.25	-	-
TCG → TGG Ser522 → Trp	-	-	-	-	-	5
ACC → TCC Thr525 → Ser	5	-	-	-	-	-
CAC → TAC His526 → Tyr	15	-	25	18.75	15	10
CAC → AAC His526 → Asn	-	-	-	-	15	-
CAC → CCC His526 → Pro	-	75	-	-	-	-
CAC → CGC His526 → Arg	-	-	45	6.25	-	5
CAC → GAC His526 → Asp	5	-	-	-	-	45
AAG → AAT Lys527 → Asn	-	-	-	6.25	-	-
CGT → TGT Arg529 → Cys	15	-	-	-	-	-
CGT → AGT Arg529 → Ser	5	-	-	-	-	-
TCG → TTG Ser531 → Leu	5	-	25	-	-	-
TCG → TGG Ser531 → Trp	-	-	-	-	-	10
GCG → GTG Ala532 → Val	-	-	-	6.25	-	-
15 bp deletion	5	-	-	-	-	-
Unknown	30	-	5	37.5	-	20

The collective data obtained in this study clearly shows that *mutY* is important in maintaining the level of C → A and A → C mutations in the genome of *M. smegmatis*. MutY could be classified as playing an antimutator role for genome stability in mycobacteria.

4 Discussion

TB has become a widespread disease globally and the increasing emergence of drug resistance has augmented the need for novel and/or improved current drug candidates. The hostile environment that the host macrophages create is an important immune response generating reactive species such as superoxide, hydrogen peroxide and hydroxyl radicals responsible for a bacteriocidal effect on the pathogen. Mycobacteria possess protective mechanisms such as catalases, peroxidases and superoxide dismutases against free radicals as well as repair mechanisms for damaged DNA such as the BER and GO repair systems to prevent cell death.

Enzymes involved in the BER and GO systems provides an interesting area of study in mycobacteria particularly since MutM (Fpg/Nei glycosylases) displays a high level of gene redundancy suggesting an evolutionary maintenance of these genes for the continued survival of the pathogen. Moreover, MutM is central to the BER and GO systems signifying an overlapping role of this enzyme in both repair pathways. Since mycobacterial genomes are G+C rich, the guanine base is more prone to oxidative damage thus generating 8-oxoG lesions. During DNA replication an 8-oxoG can mispair with an adenine (A) base which is generally recognized and repaired by MutM and MutY glycosylases to prevent 8-oxoG·A mispairs in subsequent DNA replication cycles (Nghiem *et al*., 1988).

The DNA glycosylases that make up the BER pathway have been extensively studied in *E. coli*. One of the landmark studies showed that a *mutM mutY* double mutant results in extremely high levels of GC → TA transversions compared to the individual mutants, suggesting that these glycosylases in combination are important as part of an error avoidance system that protects the cell from the mutagenic effects of 8-oxoG (Michaels *et al*., 1992). While studies in

mycobacteria, have shown a moderate increase in the spontaneous mutation frequency for a *M. smegmatis fpg* deficient mutant (Jain *et al*., 2007). The mutation spectrum revealed an increase in the A → G mutations instead of the expected C → A mutations, indicative of other glycosylases compensating in the absence of *fpg* (Jain *et al*., 2007). The genome and protein sequence alignment of *M. tuberculosis* revealed the presence of four Fpg/Nei DNA glycosylases; *M. smegmatis* also has homologues to these glycosylases with high amino acid sequence similarity of 75-82% (Cole *et al*., 1998; Olsen *et al*., 2009). Biochemical characterization of the Fpg protein of *M. tuberculosis* has shown that the length of the intergenic repeat found upstream of the gene influences the expression of the gene and since mycobacteria possess multiple copies of the *fpg/nei* genes they might be differentially expressed under different stress conditions (Olsen *et al*., 2009). Furthermore, there is biochemical evidence that the mycobacterial Fpg/Nei glycosylases have substrate overlap with their *E. coli* counterparts in recognizing oxidatively damaged DNA (Guo *et al*., 2010).

Recent work in our laboratory showed that only the loss of three or more of the Fpg/Nei glycosylases resulted in reduced survival under *in vitro* oxidative stress conditions (Goosens *et al.*, unpublished). Additionally, a recent study by Kurthkoti *et al.* (2010) reported that a *M. smegmatis mutY* deficient mutant strain displayed no significant susceptibility to oxidative stress nor was there an increase in the mutation rate suggesting that other repair enzymes must be compensating in the absence of MutY. Therefore, with the increasing biochemical and *in vitro* evidence on these glycosylases it was important to investigate the role of *mutY* in combination with the Fpg/Nei family of DNA glycosylases for a better understanding of the mutual and/or overlapping interaction of these DNA glycosylases during DNA repair.

Bioinformatics analysis has shown that *mutY* is in the same chromosomal context in all sequenced mycobacterial genomes thus *M. smegmatis*, a close relative of *M. tuberculosis* was an attractive model organism to elucidate the role of these DNA glycosylases given its fast growth and non-pathogenicity. Furthermore, the *mutY* gene is not only in the same chromosomal context in both *M. tuberculosis* and *M. smegmatis* but is also conserved and well studied in most other bacteria such as *E. coli*, *H. pylori* and *B. subtilis*, hence MutY must be an important antimutator in these organisms to maintain genome integrity (Nghiem *et al*., 1988; Eutsey *et al*., 2007; Debora *et al*., 2011). It is remarkable that *M. leprae* which has lost most of its genes to maintain a minimal genome has managed to retain *mutY*, a single *fpg* and the *nth* genes involved in the BER pathway suggesting the significance of these genes in DNA repair (Young and Robertson, 2001). More recently, a *M. tuberculosis mutY* deletion mutant (identified in a transposon mutant library) was found to have increased susceptibility to hydrogen peroxide as growth was totally inhibited with 9 mM hydrogen peroxide while the parental and the *mutM* mutant strains continued to grow under these conditions hence, signifying an important role of MutY for the continued survival of the organism under oxidative stress conditions (Mestre *et al*., 2013). Taken together all these studies clearly reflect an important role for *mutY* in genome maintenance.

As several combinatorial mutants deficient in the Fpg/Nei family of DNA glycosylases were previously generated (Goosens *et al*., unpublished), these provided the necessary *mutM* deficient background strains to further inactivate the *mutY* homologue using homologous recombination. The loss of *mutY* in the parental (ΔY) and in the Fpg/Nei deficient mutant strains (Q1ΔY and Q47ΔY) displayed no growth differences under normal culture conditions as the growth kinetics were comparable to the parental (WT) and Fpg/Nei deficient mutant strains (Q1 and Q47). Endogenous agents such as hydrogen peroxide and cumene peroxide were used to

mimic the *in vivo* oxidative environment found within activated macrophages that releases ROS as an immune response during infection. ROS has been an area of extensive study in TB pathogenesis in order to understand the role of the multiple DNA repair pathways and genes involved in counteracting the damaging effects that these reactive species have on DNA. The various mutant strains were assessed for their ability to survive under oxidative stress conditions as generated by hydrogen peroxide and cumene peroxide. A deficiency in *mutY* (ΔY) alone showed a detectable reduced survival phenotype under oxidative stress conditions in comparison to the parental (WT) strain. In contrast, Kurthkoti *et al.*, (2010) did not observe a reduced survival phenotype for the *M. smegmatis mutY* deficient mutant generated in their study. This difference in phenotype could probably be due to the lower concentration (1.5 mM) of hydrogen peroxide used which may have been insufficient to create the appropriate oxidative stress environment or the mycobacterial endogenous anti-oxidative system was able to control the ROS levels to prevent DNA damage. Another contributing factor could be the difference in growth assessment methodologies. Kurthkoti *et al.*, (2010) monitored survival as OD_{600nm} measurements as opposed to the more accurate cfu/ml measurements used in our study.

Inactivation of the *mutY* gene together with the Fpg/Nei family of DNA glycosylases (Q1ΔY and Q47ΔY) resulted in an increased sensitivity of the mutant strains suggesting that the inactivation of *mutY* has an additive effect on the loss of DNA repair function. Interestingly, the loss of the four *fpg/nei* genes had a lesser effect on cell survival than the loss of *mutY* alone indicating that *mutY* may play a more significant antimutator role in maintaining genome integrity. It is also possible that the Fpg/Nei DNA glycosylases are unable to compensate for mispaired bases and are specifically restricted to repairing 8-oxoG lesions. Complementation of the various mutants with a functional *mutY* gene was able to restore the growth deficient

phenotype to wild type levels, confirming that the observed reduced survival phenotype was due to the loss of *mutY* and not due to downstream polar effects.

The *in vitro* oxidative stress conditions created by cumene peroxide a more stable agent than hydrogen peroxide were not as effective in reducing the survival of the various mutants. The reason for this observation could be two fold; perhaps the appropriate oxidative stress environment was not created or the agent was too toxic to the cell and the organism was unable to overcome the damage as both the parental and mutant strains could not sustain growth after six hours of exposure. Alternatively, the amount of ROS levels generated by cumene peroxide was detoxified by the endogenous mycobacterial anti-oxidant mechanisms involving catalases, peroxidases and superoxide dismutases encoded by the *katG*, *ahpC* and *sodC* genes (Bartos *et al*., 2004). A recent study by Saikolappan *et al*. (2011) showed that the protein OsmC (osmotically induced bacterial protein C) plays a role in organic peroxide metabolism thus protecting the organism from oxidative stress. This gene is conserved across mycobacterial species and has functional and structural similarities to an organic hydroperoxide reductase (Ohr) but its mechanism of protection is unknown (Saikolappan *et al*., 2011). In another study a *M. tuberculosis ahpC* mutant strain exposed to different concentrations of cumene peroxide and hydrogen peroxide displayed increased sensitivity to cumene peroxide compared to the parental strain suggesting a detoxification role of AhpCs (Springer *et al*., 2001). These various studies clearly show that mycobacteria have a robust detoxification system.

Houghton *et al.* (2012) investigated the sensitivity of an *M. tuberculosis uvrD1* deletion mutant to ROS as generated by menadione and tert-butyl hydroperoxide. The study showed increased sensitivity of the mutant to these compounds and suggested that these compounds are more stable than hydrogen peroxide. We therefore decided to explore these agents as alternate

exogenous sources for generating oxidative stress conditions. Preliminary data showed that both these agents are able to kill wild type *M. smegmatis* strain at the optimized concentrations. However, due to time constraints the various DNA glycosylase deficient mutant strains generated in this study were not tested under these conditions but further optimization with these agents may lead to better *in vitro* oxidative stress assays.

With increasing loss of repair genes it was expected that there will be an increase in the number of mutational events. The spontaneous mutation rates to rifampicin assessed for the various mutant strains generated in this study were shown to increase relative to the wild type strain. Loss of *mutY* (ΔY) alone in *M. smegmatis* did not show an increase in mutation rate for two out of the three experiments (Table 3.1) which was consistent with the finding of Kurthkoti *et al.* (2010) in which an independent *M. smegmatis mutY* deficient mutant also displayed no significant increase in spontaneous mutation frequency. Our study further showed that the deletion of *mutY* in conjunction with the Fpg/Nei deficient mutant (Q1ΔY) displayed increased mutation rates for two out of three experiments which was consistent with the findings in *E. coli* that showed an increased mutation frequency only when both *mutM* and *mutY* are absent and not with the individual single mutants (Michaels *et al.*, 1992; Fowler *et al.*, 2003). Mutation frequency is defined as the frequency at which mutants arise within a population and is determined by the ratio of mutants that arise divided by the total cells in the population. Even though it is an easier method to determine mutations it is not very accurate since any pre-existing mutations will bias the frequency. In contrast a mutation rate is defined as the probability of a cell obtaining a mutation during its lifetime which is a more accurate and reproducible measure of acquired mutations compared to mutation frequency (Rosche and Foster, 2000). This study is

the first reported case of increased mutation rates in mycobacteria with the loss of five DNA glycosylases.

It was also important to understand whether an increase in the mutation rate resulted in an increase in the corresponding expected mutations. Mutations in the *rpoB* gene of *M. tuberculosis* is well documented, 17 mutations are found within the 81bp region called the rifampicin resistance determining region (RRDR) that include 14 point mutations, 2 deletions and 1 insertion (Miller *et al*., 1994). *M. smegmatis* shares 85% sequence similarity with the *rpoB* gene of *M. tuberculosis* therefore, the RRDR region is comparable in both organisms and similar amino acid changes are responsible for rifampicin resistance (Miller *et al.*, 1994 and Kurthkoti *et al*., 2010). Due to cost constraints only the Q47 lineage mutants were selected for analysis since it displayed higher sensitivity to hydrogen peroxide when *mutY* was inactivated. The RRDR region for DNA purified from spontaneous rifampicin resistant mutants isolated in the fluctuation assay for the wild type (WT), Fpg/Nei deficient mutant (Q47), *mutY* deficient mutants (ΔY and Q47ΔY) and complemented (ΔY:Y and Q47ΔY:Y) strains was PCR amplified and the amplicons sequenced.

The mutation spectrum data showed an increase of 25% in C → A mutations for the *mutY* deficient mutant which was consistent with the data reported by Kurthkoti *et al*., (2010). In the absence of mycobacterial *mutY*, DNA polymerases has the ability to incorporate an adenine to base pair with 8-oxoG and thus the observed increase in C → A mutations. There was a significant 75% increase in A → C mutations which is similar to that obtained for *E. coli mutY* deficient mutants which promotes GC → TA transversions in the absence of *mutY* (Nghiem *et al*., 1988). In addition, studies with *H. pylori* a common gastric pathogen that is constantly exposed to ROS showed a significant increase (98%) in GC → TA transversions in the absence

of *mutY* compared to a 51% increase for an *E. coli mutY* deficient strain (Eutsey *et al*., 2007). Similarly, *mutY* mutants in neisserial species *Neisseria meningitides* and *Neisseria gonorrhoeae* showed increased mutation rates of between 20 and 140 fold with corresponding increases in C → A transversions (Davidsen *et al*., 2005). Hence, all these studies provide sufficient evidence that *mutY* definitely plays a crucial role in preventing GC → AT transversions to protect the genome from oxidative DNA damage during pathogenesis. Furthermore, a significant increase in mutation rates was observed when *mutY* was absent together with the Fpg/Nei deficient mutant strain (Q47ΔY) with a dramatic 70% increase in C → A mutations compared to the 25% observed in the absence of *mutY* alone (ΔY). Several *E. coli* studies have shown mutator phenotypes that result in an increase in GC → TA transversions only with a *mutM* and *mutY* double deletion, suggesting a combinatory functional role of the two enzymes to prevent mutations caused by 8-oxoG (Michael *et al.*, 1992; Michael *et al.*, 1991; Guo *et al.*, 2010). The data generated for *M. smegmatis* mutants in this study definitely indicates a crippling effect on the survival of the organism in the absence of MutY and the entire Fpg/Nei family of DNA glycosylases.

4.1 Future studies

Future studies would involve expanding the phenotypic characterization of the various mutant strains under different oxidative stress conditions. Menadione and tert-butyl hydroperoxide seems to be promising agents thus; further optimization of these agents is required to assess the various *mutY* deficient strains. It would also be interesting to identify the type of DNA adducts that result after exposure to exogenous oxidative causing agents to correlate the mutational lesions to the DNA glycosylase. The *M. smegmatis* data generated thus far for the various DNA glycosylases provides an informative platform for streamlining the

investigation of these important enzymes in *M. tuberculosis* to better understand the physiological role of these BER DNA glycosylases on the survival of the pathogen under oxidative conditions.

4.2 Limitations

Although there was a general increase in mutation rate with increased loss of DNA glycosylases, the data was only for a small representative sample which could have biased the outcome and for statistical accuracy it would be ideal to have a larger sample size. In addition it would be imperative to assess the mutator phenotype of the Q1ΔY mutant strain and its associated parental and complemented strains to understand the underlying mechanisms that resulted in the differential phenotypes observed for the two Fpg/Nei deficient lineages. Due to the cost of sequencing, the sample size and the selection of strains for analysis was limited.

4.3 Concluding remarks

DNA repair is important for the survival of the pathogen thus it is a key biological area for understanding the development of drug resistance during TB therapy. This study provides valuable information regarding the importance of several glycosylases in the BER pathway. Progressive loss of the Fpg/Nei and MutY glycosylases rendered the organisms more sensitive to oxidative damage and interestingly the type of mutations that arose in the absence of these glycosylases were as expected, suggesting that *mutY* plays a more significant role in genome maintenace and no other repair genes seem to be able to compensate for the damage. The use of *M. smegmatis* as a model organism has provided useful information that adds to the current knowledge on DNA glycosylases in mycobacteria and provides the basis for streamlining similar studies in the pathogenic *M. tuberculosis* strain to better understand mycobacterial pathogenesis and resistance.

5 Appendices

5.1 Appendix A: PCR and sequencing primers

Table 5.1.1: Primers used in this study

Primer	Sequence (5'- 3')*	Annealing temperature for Taq or Phusion DNA polymerase	Expected (bp)
	***mutY* Upstream deletion region (Upstream)**		
MutYUSF MutYUSR	GGCCAAGCTTAGCTGATCGTTCGCGACTAC GGCCAGATCTGCATGGTCGTACCAACTCCAA	Taq = 63.1˚C Phusion = 63.1˚C	1932
	***mutY* Downstream deletion region (Downstream)**		
MutYDSF MutYDSR	GGCCAGATCTGACTCGCTGCTGGTGGAC GGCCGCGGCCGCGCTTCTGCGCAAGAGGTTTA	Taq = 58.8˚C Phusion = 58.8˚C	1775
	Confirmation of DCOs / Complementation of MutY		
mutYFwd mutYRev	GCCTCGAATTCGTCGACAC CGTGACACCGTCGGACTAC	Taq = 57˚C Phusion = 60˚C	Parental (1481) Mutant (733)
	Real Time PCR		
RTmutYfwd1 RTmutYrev1	TGGCGCTCGGCCGGATTCC GTCCACCAGCAGCGAGTC	Taq = 60˚C	198
	RRDR region		
MsmrpoBF2 MsmrpoBR2	TGGTGCGTCTGCACGAGGGTC ATCTGGTCGGTGACCACACCG	Taq = 60˚C	184

*Underlined sequences are restriction sites incorporated within the primers

Table 5.1.2: Sequencing primers used in this study

Primer	Sequence (5'- 3')*	Description
M13Fpuc(-40)	GTTTTCCCAGTCACGAC	Universal M13 sequencing primer
US1	AGTTGCACCACGAGGTTGTAG	Upstream internal sequencing primers
US2	CGGTCTCATTGGACTGATCG	
US3	CTTGCGTCCCAGCAGGAT	
DS1	GTTGGCGACGATCCAGTC	Downstream internal sequencing primers
DS2	ACCCACCAGGTCTTCCAGA	
DS3	AGCCGTTTCGAGCAGCAC	
pTWF	GCTGAAGCCAGTTACCTTCG	Complementation vector and internal sequencing primers
pTWR	TTTGCCTTCCTGTTTTTGCT	
mutYFwd	GCCTCGAATTCGTCGACAC	
mutYRev	CGTGACACCGTCGGACTAC	
mutYM	CATGTCGCCGAGGCCCTGGT	
MsmrpoBF1	GCAGACCCTGATCAACATCC	RRDR sequencing primer

5.2 Appendix B: Plasmid maps, cloning strategies and molecular weight markers

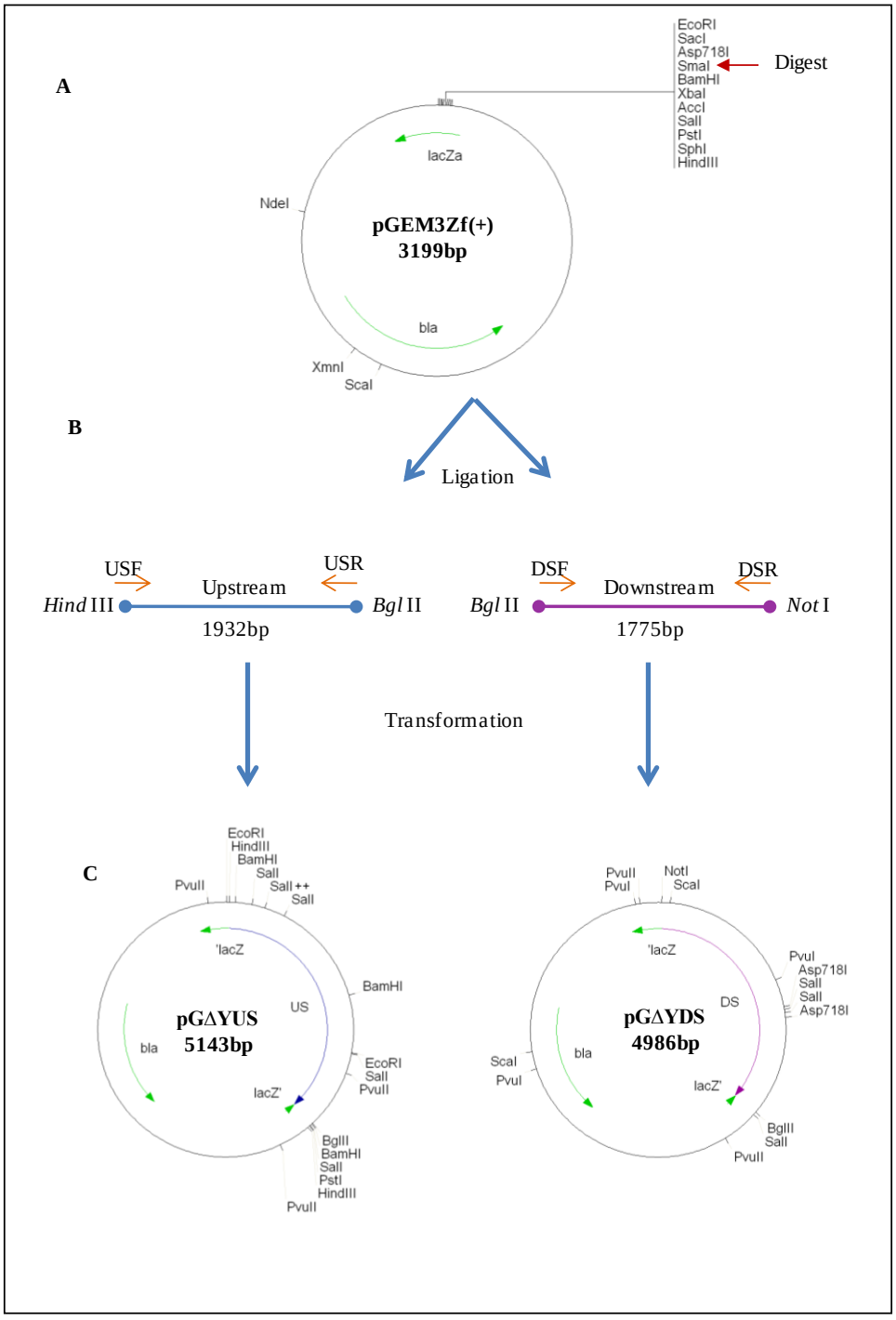


Figure 5.2.1: Cloning of upstream and downstream homologous regions into pGEM3Zf(+). A. The vector was digested at the *Sma* I site indicated by the red arrow followed by **B**. Ligation of the homologous fragments separately into pGEM3Zf(+) at the *Sma* I site. The orange arrows represent the forward and reverse primers with the unique restriction sites indicated to amplify the upstream (blue) and downstream (purple) homologous fragments respectively, followed by transformation. **C.** Restriction map of pGΔYUS and pGΔYDS.

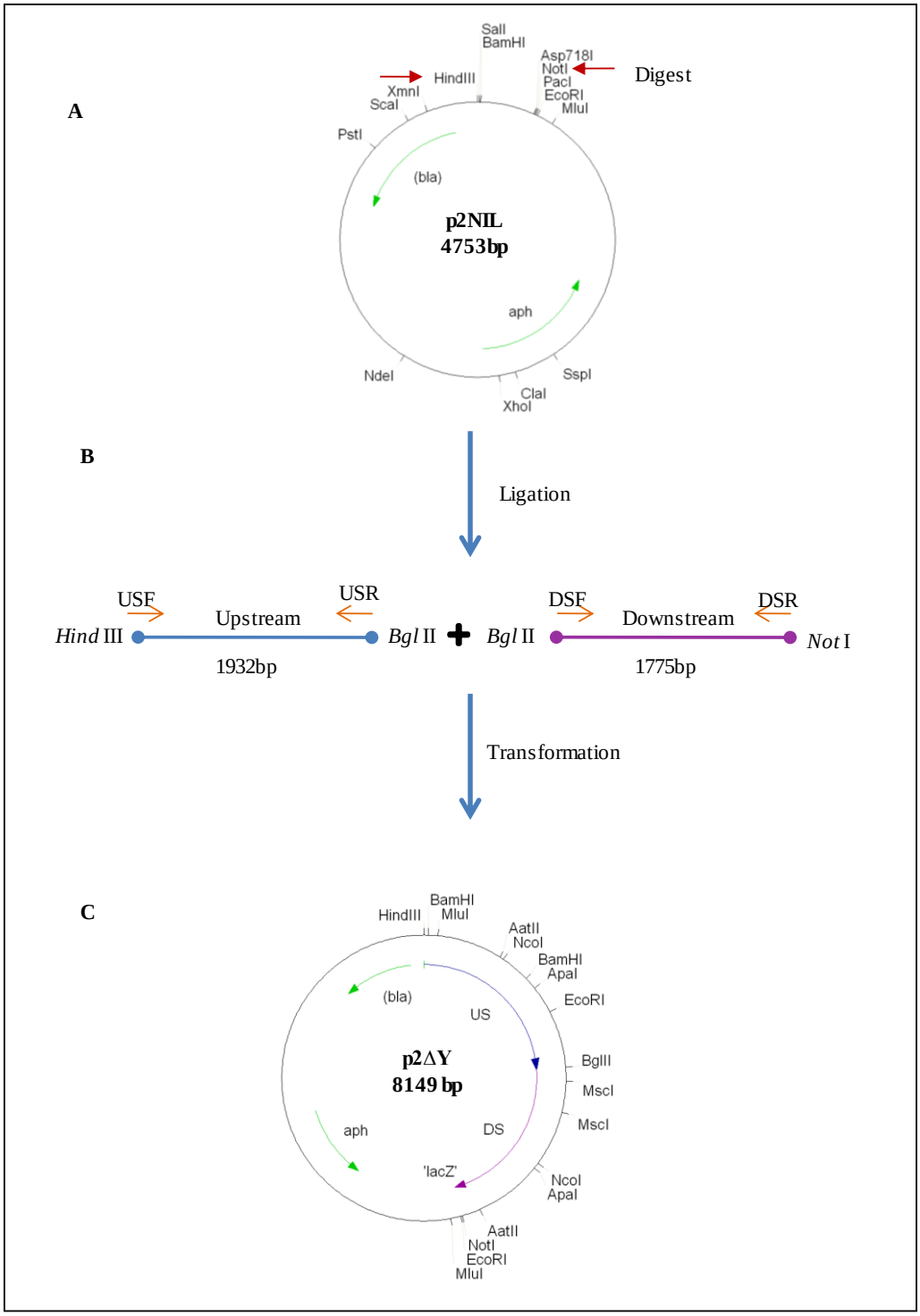


Figure 5.2.2: Three way cloning into p2NIL. A. Digested vector with *Hind* III and *Not* I indicated by the red arrows followed by **B**. The upstream and downstream homologous fragments were excised from pGΔYUS and pGΔYDS respectively, followed by ligation into p2NIL at the *Hind* III/*Not* I site. The orange arrows represent the forward and reverse primers with the unique restriction sites indicated to amplify the upstream (blue) and downstream (purple) homologous fragments respectively, followed by transformation. **C.** Restriction map of p2ΔY.

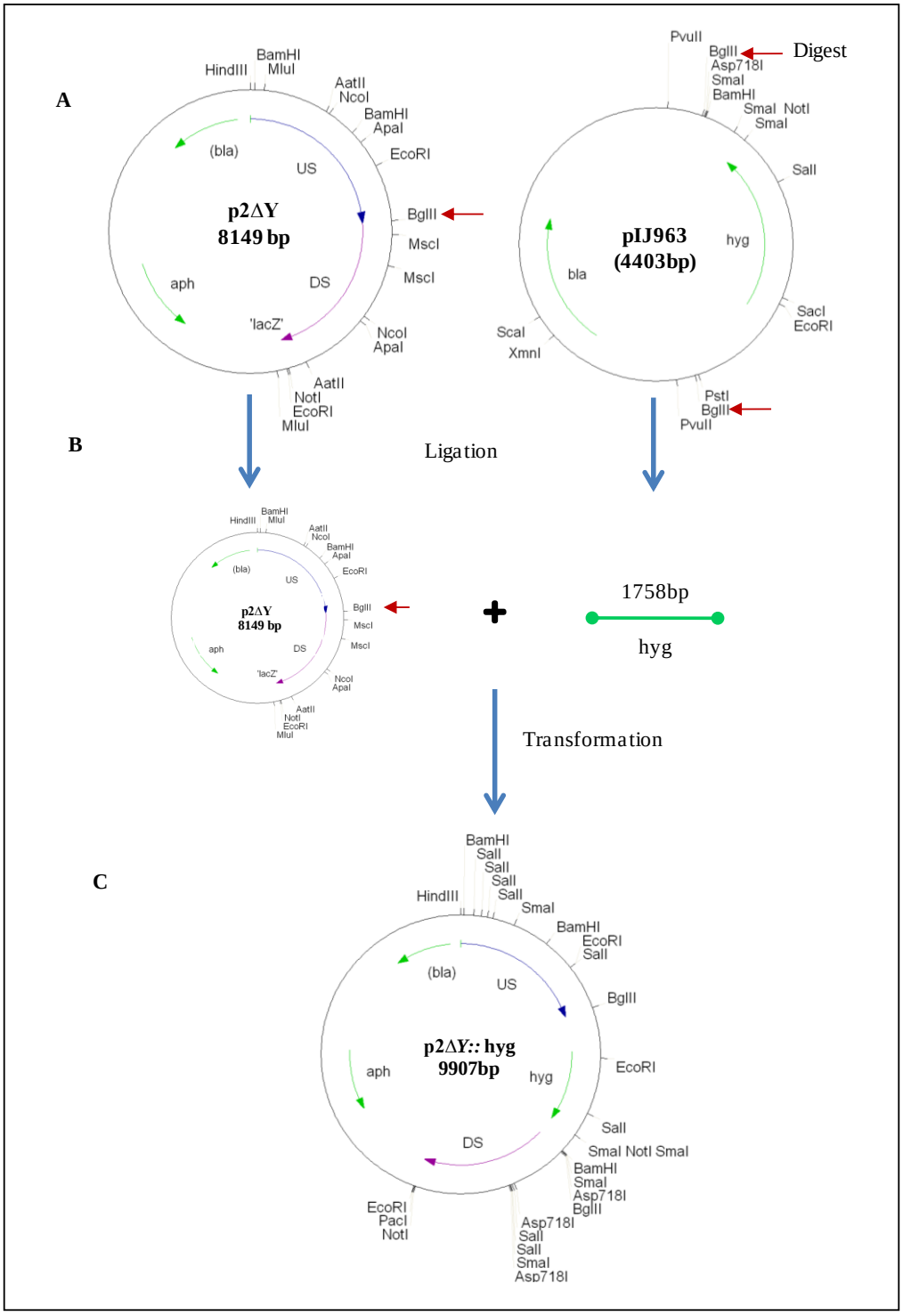


Figure 5.2.3: Cloning strategy for a marked suicide vector. A. p2NIL suicide vector (p2ΔY) and pIJ963 were digested with *Bgl* II (indicated by the red arrows) to linearize p2ΔY and to release the hygromycin cassette (depicted by the green arrow) respectively. **B**. Ligation of the hygromycin cassette (1758bp) at the *Bgl* II site of p2ΔY (ligated in between the upstream homologous fragment (blue) and downstream homologous fragment (purple), followed by transfomation. **C.** Restriction map of p2Δ*Y*::hyg

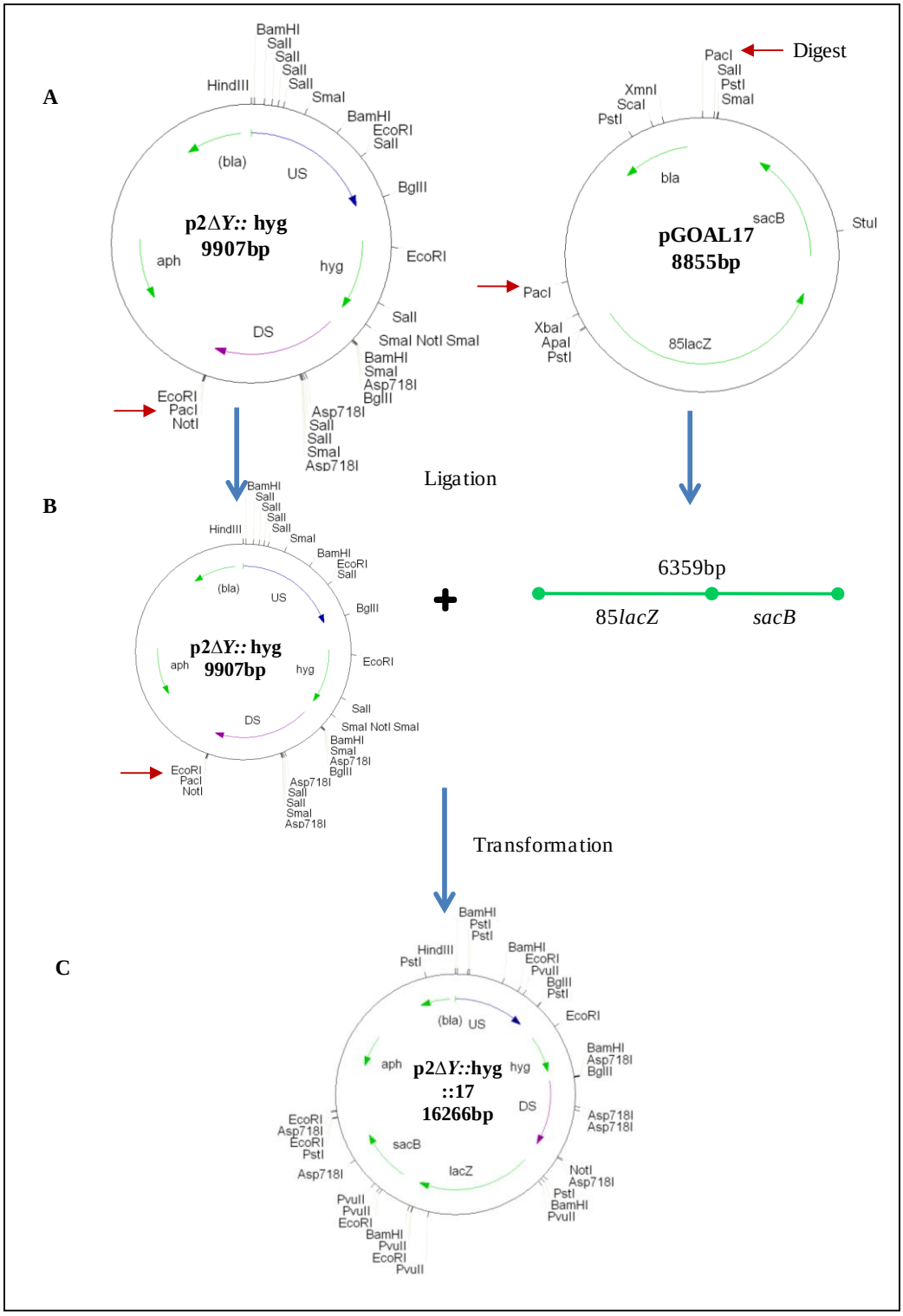


Figure 5.2.4: Cloning strategy for a marked suicide vector containing the *Pac* cassette. A. Digested the marked vector and pGOAL17 at the *Pac* I site (indicated by the red arrows) to linearize p2Δ*Y::* hyg and to release the *Pac* cassette (*lacZ* and *sacB* depicted by the green arrows). **B**. Ligation of the *Pac* cassette at the *Pac* I site of the marked suicide vector followed by transformation. **C.** Restriction map of p2Δ*Y*::hyg::17.

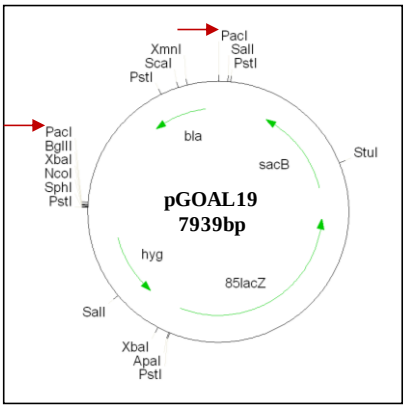


Figure 5.2.5: pGOAL19 containing the *Pac* cassette. *Pac* cassette for insertion into the p2NIL suicide vector to generate an unmarked suicide vector.

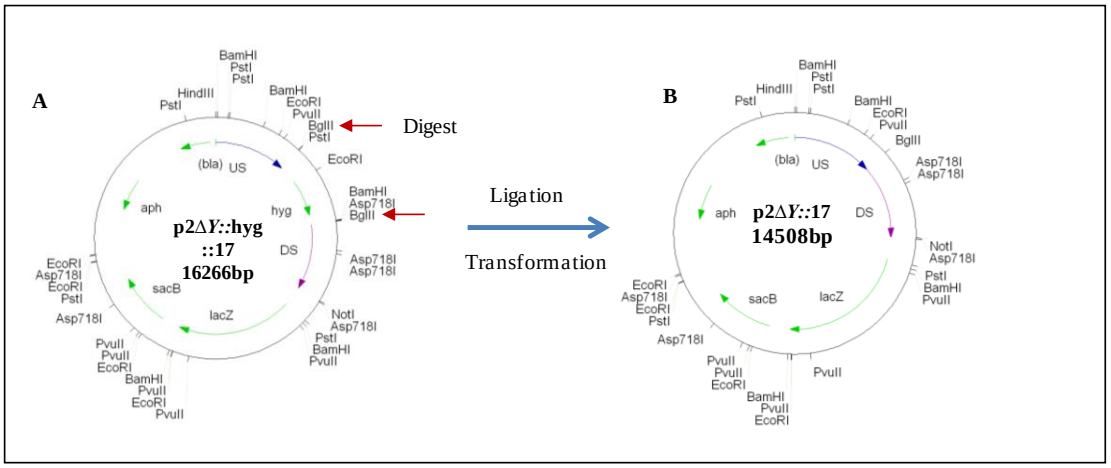


Figure 5.2.6: Cloning strategy for an unmarked suicide vector. A. Digested p2ΔY::hyg::17 (marked suicide vector) containing the marker cassettes with *Bgl* II (indicated by the red arrows) to remove the hygromycin cassette (green arrow). Followed by ligation and transformation. **B**. Restriction map of p2ΔY::17.

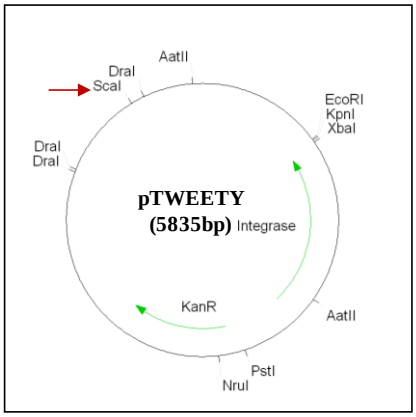


Figure 5.2.7: pTWEETY integrating vector. The *mutY* gene was amplified using mutYFwd and mutYRev primers (Table 5.1.1) and inserted at the *Sca* I site of pTWEETY (indicated by the red arrow).

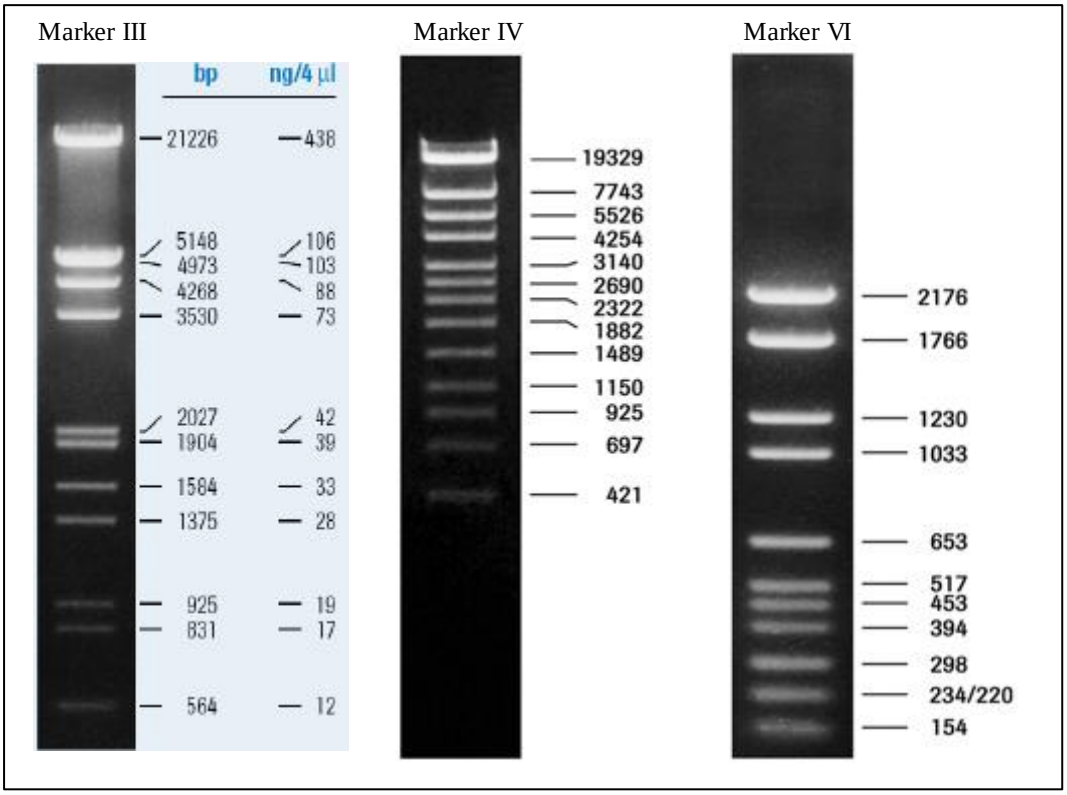


Figure 5.2.8: DNA molecular weight markers III, IV and VI. Marker III was used to quantify DNA.

6 References

1. Bartos, M., Falkinham, J. O. and Pavlik, I. 2004. Mycobacterial catalases, peroxidases and superoxide dismutases and their effects on virulence and isoniazid susceptibility in mycobacteria - a review. *Veterinary Medical – Czech* 49(5): 161-170.
2. Blondelet-Rouault, M. H., Weiser, J., Lebrihi, A., Branny, P. and Pernodet, J. L. 1997. Antibiotic resistance gene cassettes derived from the omega interposon for use in *E. coli* and Streptomyces. *Gene* 190(2): 315-317.
3. Cabrera, M., Nghiem, Y. and Miller, J.H. 1988. *mutM*, a second mutator locus in *Escherichia coli* that generates G·C → T·A transversions. *Journal of Bacteriology* 170(11): 5405–5407.
4. Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S. V., Eiglmeier, K.,Gas, S., Barry, C. E., 3rd, Tekaia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Moule, S., Murphy, L., Oliver, K., Osborne, J., Quail, M. A., Rajandream, M. A., Rogers, J., Rutter, S., Seeger, K., Skelton, J., Squares, R., Squares, S., Sulston, J. E., Taylor, K., Whitehead, S. and Barrell, B. G. 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* 393(6685): 537-544.
5. Cole, S. T., Eiglmeier, K., Parkhill, J., James, K. D., Thomson, N. R., Wheeler, P. R., Honore, N., Garnier, T., Churcher, C., Harris, D., Mungall, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R. M., Devlin, K., Duthoy, S., Feltwell, T., Fraser, A., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Lacroix, C., Maclean, J., Moule, S., Murphy, L., Oliver, K., Quail, M. A., Rajandream, M. A., Rutherford, K. M., Rutter, S.,

Seeger, K., Simon, S., Simmonds, M., Skelton, J., Squares, R., Squares, S., Stevens, K. Taylor, K., Whitehead, S., Woodward, J. R and Barrell, B. G. 2001. Massive gene decay in the leprosy bacillus. *Nature* 409(6823): 1007–1011.

6. Davidsen, T., Bjoras, M., Seeberg, E. C. and Tonjum, T. 2005. Antimutator role of DNA glycosylase MutY in pathogenic *Neisseria* species. *Journal of Bacteriology* 187(8): 2801–2809.
7. Debora, B. N., Vidales, L. E., Ramirez, R. Ramirez, M., Robleto, E. A. Yasbin, R. E. and Pedraza-Reyes, M. 2011. Mismatch repair modulation of MutY activity drives *Bacillus subtilis* stationary-phase mutagenesis. *Journal of Bacteriol*ogy 193(1): 236-245.
8. Dos Vultos, T., Blazquez, J., Rauzier, J., Matic, I., Gicquel, B. 2006. Identification of Nudix hydrolase family members with an antimutator role in *Mycobacterium tuberculosis* and *Mycobacterium smegmatis*. *Journal of Bacteriology* 188(8): 3159-3161.
9. Dos Vultos, T., Mestre, O., Tonjum, T. and Gicquel, B. 2009. DNA repair in *Mycobacterium tuberculosis* revisited. *FEMS Microbiology Review* 33(3): 471-487.
10. Dussurget, O., Rodriguez, M. and Smith, I. 1998. Protective role of the *Mycobacterium smegmatis* IdeR against reactive oxygen species and isoniazid toxicity. *Tubercle and Lung Disease* 79(2): 99-106.
11. Ehrt, S. and Schnappinger, D. 2009. Mycobacterial survival strategies in the phagosome: Defense against host stresses. *Cell Microbiology* 11(8):1170-1178.
12. Emile, J. F., Patey, N., Altare, F., Lamhamedi, S., Jouanguy, E., Boman, F., Quillard, J., Lecomte-Houcke, M., Verola, O., Mousnier, J.F., Dijoud, F., Blanche, S., Fischer, A., Brousse, N. and Casanova, J. L. 1997. Correlation of granuloma structure with clinical

outcome defines two types of idiopathic disseminated BCG infection. *Journal of Pathology* 181(1): 25-30.

13. Eutsey, R., Wang, G. and Maier, R. J. 2007. Role of a *mutY* DNA glycosylase in combating oxidative DNA damage in *Helicobacter pylori*. *DNA Repair (Amst)* 6(1): 19–26.
14. Fowler, R. G., White, S. J., Koyama, C., Moore, S. C., Dunn, R. L. and Schaaper, R. M. 2003. Interactions among the *Escherichia coli* MutT, MutM, and MutY damage prevention pathways. *DNA Repair (Amst)* 2(2): 159-173.
15. Friedberg, E. C., Walker, G. C. and Siede, W. 1995. DNA Repair and Mutagenesis. Washington D.C, ASM Press: 135-181.
16. Fromme, J. C., Banerjee, A. and Verdine, G. L. 2004. DNA glycosylase recognition and catalysis. *Current Opinion in Structural Biology* 14(1): 43–49
17. Gandhi, N. R., Moll, A., Sturm, A. W., Pawinski, R., Govender, T., Lalloo, U., Zeller, K., Andrews, J. and Friedland, G. 2006. Extensively drug-resistant tuberculosis as a cause of death in patients co-infected with tuberculosis and HIV in a rural area of South Africa. *Journal of Acquired Immune Deficiency Syndrome* 50(1): 37-43.
18. Garnier, T., Eiglmeier, K., Camus, J. C., Medina, N., Mansoor, H., Pryor, M., Duthoy, S., Grondin, S., Lacroix, C., Monsempe, C., Simon, S., Harris, B., Atkin, R., Doggett, J., Mayes, R., Keating, L., Wheeler, P. R., Parkhill, J., Barrell, B. G., Cole, S. T., Gordon, S. V. and Hewinson, R. G. 2003. The complete genome sequence of *Mycobacterium bovis*. *Proceedings of the National Academy of Science USA* 100(13): 7877–7882.
19. Goosens, V. 2008. MSc Dissertation. Faculty of Health Sciences, University of the Witwatersrand.

20. Goosens, V., Mizrahi, V. and Gordhan, B.G. 2013. The role of Fpg DNA glycosylases and Nei endonucleases in mutagenesis in *Mycobacterium smegmatis*. *In preparation*.
21. Gordhan, B. G. and Parish, T. 2001. Gene replacement using pretreated DNA. *Methods in molecular medicine Mycobacterium tuberculosis protocols*. Parish, T. and Stoker, N. G. New Jersey, Humana Press: 77-90.
22. Guo, Y., Bandaru, V., Jaruga, P., Zhao, X., Burrows, C. J. Iwai, S., Dizdaroglu, M., Bond, J. P. and Wallace, S. S. 2010. The oxidative DNA glycosylases of *Mycobacterium tuberculosis* exhibit different substrate preferences from their *Escherichia coli* counterparts. *DNA Repair* 9(2): 177-190.
23. Hanahan, D. 1983. Studies on transformation of *Escherichia coli* with plasmids. *Journal of Molecular Biololgy* 166(4): 557-580.
24. Harth, G. and Horwitz, M. A. 1999. Export of recombinant *Mycobacterium tuberculosis* superoxide dismutase is dependent upon both information in the protein and mycobacterial export machinery. *The American Society for Biochemistry and Molecular Biology* 274(7): 4281-4292.
25. Houghton, J., Townsend, C., Williams, A. R., Rodgers, A., Rand, L., Walker, K. B., Bottger, E. C., Springer, B. and Davis, E. O. 2012. Important role for *Mycobacterium tuberculosis* UvrD1 in pathogenesis and persistence apart from its function in nucleotide excision repair. *Journal of Bacteriology* 194(11): 2916-2917.
26. Jain, R., Kumar, P. and Varshney, U. 2007. A distinct role of formamidopyrimidine DNA glycosylase (MutM) in down-regulation of accumulation of G, C mutations and protection against oxidative stress in mycobacteria. *DNA Repair* 6(12): 1774-1786.

27. Koul, A., Arnoult, E., Lounis, N.,Guillemont, J. and Andries, K. 2011. Review: The challenges of new drug discovery for tuberculosis. *Nature* 469(7331): 483-490.
28. Krokan, H.E., Standal, R., and Slupphaug, G. 1997. DNA glycosylases in the base excision repair of DNA.*Biochemistry Journal* 325(1): 1-16.
29. Kumar, A. Farhana, A., Guidry, L., Sani, V., Hondalus, M. and Steyn, A.J.C. 2011. Redox homeostasis in mycobacteria: the key to tuberculosis control? *Expert Reviews In Molecular Medicine* 13(39): 1-25.
30. Kurthkoti, K. and Varshney U. 2011. Base excision and nucleotide excision repair pathways in mycobacteria. *Tuberculosis* 91(6): 533-543.
31. Kurthkoti, K. and Varshney, U. 2012. Distinct mechanisms of DNA repair in mycobacteria and their implications in attenuation of the pathogen growth. *Mechanisms of Ageing and Development* 133(4): 138-146.
32. Kurthkoti, K., Srinath, T., Kumar, P., Malshetty, V. S., Sang, P. B., Jain, R., Manjunath, R. and Varshney, U. 2010. A distinct physiological role of MutY in mutation prevention in mycobacteria. *Microbiology* 156(1): 88-98.
33. Li, X. and Lu, A. L. 2001. Molecular cloning and functional analysis of the MutY homolog of *Deinococcus radiodurans*. *Journal of Bacteriology* 183(21): 6151-6158.
34. Lienhardt, C., Raviglione, M., Spigelman, M., Hafner, R., Jaramillo, E., Hoelscher, M., Zumla, A. and Gheuens, J. 2012. New drugs for the development of Tuberculosis: Needs, challenges, promise, and prospects for the future. *Journal of infectious diseases* 205(2): 241-249.
35. Machowski, E. E., Barichievy, S., Springer, B., Durbach, S. I. and Mizrahi, V. 2007. *In vitro* analysis of rates and spectra of mutations in a polymorphic region of the Rv0746

PE_PGRS gene of *Mycobacterium tuberculosis*. *Journal of Bacteriology* 189(5): 2190-2195.

36. Manuel, R. C., Czerwinski, E. W. and Lloyd, R. S. 1996. Identification of the structural and functional domains of MutY, an *Escherichia coli* DNA mismatch repair enzyme. *The Journal of Biological Chemistry* 271(27): 16218–16226.
37. Mestre, O., Hurtado-Ortiz, R., Dos Vultos, T., Namouchi, A., Cimino, M., Pimentel, M., Neyrolles, O. and Gicquel, B. 2013. High throughput phenotypic selection of *Mycobacterium tuberculosis* mutants with impaired resistance to reactive oxygen species identifies genes important for intracellular growth. *Plos one* 8(1): 1-10.
38. Meya, D. B. and McAdam, K. P. 2007. The TB pandemic: an old problem seeking new solutions. *Journal of International Medicine* 261(4): 309-329.
39. Michaels, M. L. and Miller, J. H. 1992. The GO system protects organisms from the mutagenic effect of the spontaneous lesion 8-hydroxyguanine (7, 8-dihydro-8-oxoguanine). *Journal of Bacteriology* 174(20): 6321-6325.
40. Michaels, M. L., Cruz, C., Grollman, A. P. and Miller, J. H. 1992. Evidence that MutY and MutM combine to prevent mutations by an oxidatively damaged form of guanine in DNA. *Proceedings of the National Academic Science USA* 89(15): 7022-7025.
41. Michaels, M. L., Pham, L., Cruz, C. and Miller, J. H. 1991. MutM, a protein that prevents GC → TA transversions, is formamidopyrimidine-DNA glycosylase. *Nucleic acids Research*. 19(13): 3629-3632.
42. Miller, L.P., Crawford, J.T. and Shinnick, T. M. 1994. The *rpo*B gene of *Mycobacterium tuberculosis*. *Antimicrobial Agents and Chemotherapy* 38(4): 805-811.

43. Mizrahi, V. and Andersen, S. J. 1998. DNA repair in *Mycobacterium tuberculosis*. What have we learnt from the genome sequence? *Molecular Microbiology* 29(6): 1331-1339.

44. Nghiem, Y., Cabrera, M., Cupples, C.G. and Miller, J.H. 1988. The *mutY* gene: a mutator locus in *Escherichia coli* that generates G:C to T:A transversions. *Proceedings of the National Academic Science USA* 85(8): 9163–9166.

45. O' Sullivan, D.M., McHugh, T.D. and Gillespie, S.H. 2005. Analysis of *rpo*B and *pnc*A mutations in the published literature: an insight into the role of oxidative stress in *Mycobacterium tuberculosis* evolution? *Journal of Antimicrobial Chemotherapy* 55(5): 674-679.

46. Oliver, A., Sanchez, J.M. and Blazquez, J. 2002. Characterization of the GO system of *Pseudomonas aeruginosa*. *FEMS Microbiology Letters* 217(1): 31-35.

47. Olsen, I., Balasingham, S. V., Davidsen, T., Debebe, E., Rodland, E. A., Van Soolingen, D., Kremer, K., Alseth, I. and Tonjum, T. 2009. Characterization of the major formamidopyrimidine-DNA glycosylase homolog in *Mycobacterium tuberculosis* and its linkage to variable tandem repeats. *FEMS Immunology and Medical Microbiology* 56(2): 151-161.

48. Parish, T. and Stoker, N. G. 2000. Use of a flexible cassette method to generate a double unmarked *Mycobacterium tuberculosis tlyA plcABC* mutant by gene replacement. *Microbiology* 146(8): 1969-1975.

49. Pham, T. T., Jacobs-Sera, D. Pedulla, M. L. Hendrix, R. W. and Hatfull, G. F. 2007. Comparative genomic analysis of mycobacteriophage Tweety: evolutionary insights and construction of compatible site-specific integration vectors for mycobacteria. *Microbiology* 153(8): 2711-2723.

50. Rosche, W. A. and Foster, P. L. 2000. Determining mutation rates in bacterial populations. *Methods* 20(1): 4-17.

51. Saikolappan, S., Das, K., Saindran, S. J., Jagannath, C. and Dhandayuthapani, S. 2011. OsmC proteins of *Mycobacterium tuberculosis* and *Mycobacterium smegmatis*. *Tuberculosis* 91(1): 119-127.

52. Slupska, M., Baikalov, M., Luther, W. M., Chiang, J., Wei, Y. and Miller, J.H. 1996. Cloning and sequencing a human homolog (*hMYH*) of the *Escherichia coli mutY* gene whose function is required for the repair of oxidative DNA damage. 1996. *Journal of Bacteriology* 178(13): 3885-3892.

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

54. Springer, B., Master, S., Sander, P., Zahrt, T., McFalone, M., Sang, J., Papavinasasundaram, K. G., Colston, M. J., Boettger, E. and Deretic, V. 2001. Silencing of oxidative stress response in *Mycobacterium tuberculosis*: Expression patterns of aphC in virulent and avirulent strains and effect of *aphC* inactivation. *Infection and Immunity* 69(10): 5967-5973.

55. Teixeira, H. D. R., Schumacher, R. and Meneghini, R. 1998. Lower intracellular hydrogen peroxide levels in cells overexpressing Cu Zn-superoxide dismutase. *Proceedings of the National Academic Science USA* 95(14):7872–7875.

56. Velayati, A. A., Masjedi, M. R., Farnia, P., Tabarsi, P., Ghanavi, J., Ziazarifi, A. H. and Hoffner, S. E. 2009. Emergence of new forms of totally drug-resistant tuberculosis

bacilli: super extensively drug-resistant tuberculosis or totally drug-resistant strains in Iran. *Chest* 136(2): 420-425.

57. Wallace, S. S., Bandaru, V., Kathe, S. D. and Bond, J. P. 2003. The enigma of endonuclease VIII. *DNA Repair (Amst)* 2(5): 441-53.
58. World Health Organization. 2010/2011 TUBERCULOSIS global facts. www.who.int/tb/data. Accessed 14 February 2011.
59. Young, D. and Robertson, B. 2001. Genomics: leprosy - a degenerative disease of the genome. *Current Biology* 11(10): 381-383.