

https://dspace.org/

# Construction and phenotypic characterization of Mycobacterium smegmatis mutants deficient in DNA glycosylases

**Moolla, Nabiela**

**2014-02-18**
http://hdl.handle.net10539/13841

Downloaded from a DSpace repository. https://dspace.org/

# Construction and Phenotypic Characterization of *Mycobacterium smegmatis* Mutants Deficient in DNA Glycosylases

Nabiela Moolla

A dissertation submitted to the Faculty of Health science, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Master of Science in Medicine.

Johannesburg 2013

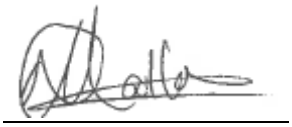
I dedicate this work to my family.

‘Allah! There is no God but Him, the Alive, the Eternal. Neither slumber nor sleep overtaketh Him. Unto Him belongeth whatsoever is in the heavens and whatsoever is in the earth. Who is he that intercedeth with Him but by His leave? He knoweth that which is in front of them and that which is behind them, while they encompass nothing of His knowledge but what He will. His throne includeth the heavens and the earth, and He is never weary of preserving them. He is the Sublime, the Tremendous.’

Holy Qur'an 2:255

# Declaration

I declare that this dissertation is my own, unaided work. It is being submitted for the degree of Masters 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.



(Nabiela Moolla)

27 day of May 2013

# Abstract

The causative pathogen of tuberculosis, *Mycobacterium tuberculosis* (Mtb), is equipped with several DNA repair mechanisms for continued survival within the host. One such mechanism is Base Excision Repair (BER) that repairs DNA damage caused by reactive oxygen and nitrogen species (ROS/RNS) generated by the host immune cells during infection. BER is dependent on DNA glycosylases namely: formamidopyrimidine (Fpg/MutM/Fapy), endonucleaseVIII (Nei) and endonucleaseIII (Nth) with Nei being structurally similar to Fpg but functionally similar to Nth. Bioinformatics analysis of the genome sequences of Mtb and its non-pathogenic relative *Mycobacterium smegmatis* (Msm) identified a unique duplication of Fpg and Nei glycosylases and a single *nth* gene in the same chromosomal context in both organisms. Previously, it has been shown that the lack of Fpg/Nei glycosylases in Msm display no differences in growth and survival under normal and oxidative stress conditions with no increase in spontaneous mutation rates as compared to the parental strain, suggesting that *nth* maybe significant for mycobacterial genome maintenance. Hence, in this study the *nth* gene was site specifically inactivated by homologous recombination in the parental Msm strain and in selected combinatorial mutant strains deficient in the Fpg/Nei glycosylases. Loss of the *nth* allele in the panel of mutants was genotypically confirmed by PCR and southern blot analyses. Inactivation of the *nth* gene did not affect the *in vitro* growth of the mutant strains under normal culture conditions. Interestingly, UV induced DNA damage of the single *nth* mutant resulted in a dramatic increase in mutation frequency that was not observed in any of the mutants. The progressive loss of *fpg*, *nei* and *nth* genes showed exaggerated reduced survival under oxidative stress. The subsequent deletion of *nth* in mutants deficient in *fpg*/*nei* resulted in a dramatic increase in spontaneous mutation rates and frequencies, implying that *nth* is

integral for the repair of both spontaneous and induced DNA damage. Undoubtedly, these results indicate that Msm *nth* encoding the Nth glycosylase is involved in DNA repair and has anti-mutator properties. Furthermore, *nth* together with *fpg* and *nei* is part of a robust DNA repair system that maintains the integrity of the mycobacterial genome.

# Acknowledgements

I would like to express my foremost gratitude to my supervisor **Dr. B.G. Gordhan** for her wisdom, guidance, patience, understanding and encouragement throughout my project. Also, to **Dr. B.D. Kana** thank you for your innovative ideas, support and motivation towards my project. Thank you to all my **colleagues at the CBTBR** with whom I have shared this experience and developed strong friendships. Everything in this project was possible with all your (everyone at the CBTBR) help and support.

I would like to thank the **NRF for the post graduate bursary** and the **University of the Witwatersrand for the Post Merit Award**. The NRF, NHLSRT and **SATBAT** (Grant # 1U2RTW007370/3) **are acknowledged for research funding awarded to B.G.G.**

I would like to extend a big thank you to the following people for their invaluable contribution to my project:

**Prof. M. Gulumian, Mrs. N. Matiwane** and **Mr. K. Boodhia** from the National Institute of Occupational Health (NIOH) for their hospitality and assistance with the comet assay and allowing me to use the fluorescence microplate reader. I will always remember the kindness I was received with.

**Prof. B. Weiss** and **Prof. R.P. Cunningham** for the *Escherichia coli* (*E. coli*) *nth* strains. **Prof. B. Weiss** from EMORY University your continuous support and input has motivated me in trying times for which I am truly grateful and honoured.

**Prof. N. P. Singh** from the University of Washington for advice on the comet assay, your work has inspired me.

**Mrs. D. van Rooi** (**Aunty Dolly**) from the Mycobacteriology unit at the National Health Laboratory Service (NHLS) for assistance with the sonicator. Your spirit and zest for life has lifted mine.

Lastly I give the greatest thanks to **Almighty** the creator, to my **Parents** and **Sisters** for their support and love that carried me through tough times; to my **Babies** for their innocence has reminded me of who I am, my **Twin** for your never ending help and laughs and **Gerald** thank you for your support, love, and faith (in me). I am humbled and blessed.

# Presentations

Aspects of this work have been presented at the following conferences:

**Moolla, N**. and Gordhan, B. G. (2012). Construction and Phenotypic Characterization of *Mycobacterium smegmatis* Mutants Deficient in DNA Glycosylases. South African Society of Biochemistry and Molecular Biology/Federation of African Societies of Biochemistry and Molecular Biology (SASBMB/FASBMB) congress hosted at the Drakensberg Champagne Sports Resort, Drakensberg, South Africa from 29 January to 01 February 2012 (poster presentation).

**Moolla, N**. and Gordhan, B. G. (2012). Construction and Phenotypic Characterization of *Mycobacterium smegmatis* Mutants Deficient in DNA Glycosylases. Wits Faculty of Health Sciences Research Day & Postgraduate Expo 2012 hosted at the University of the Witwatersrand Medical School, Parktown, South Africa on the 19 September 2012 (poster presentation).

**Moolla, N**. and Gordhan, B. G. (2012). Construction and Phenotypic Characterization of *Mycobacterium smegmatis* Mutants Deficient in DNA Glycosylases. Medical Research Council (MRC) early career scientist conference 2012 hosted at the MRC, Cape Town, South Africa from the 24 to 25 October 2012 (oral presentation).
**Awarded first prize for best oral presentation in the masters category with a sponsorship to the value of R10 000.**

# Table of Contents

Declaration....iii

Abstract....iv

Acknowledgements....vi

Presentations....viii

Table of Contents....ix

List of figures....xiv

List of Tables....xvii

Nomenclature....xviii

1. Introduction....1

1.1 The Tuberculosis scourge....1

1.2 Available TB treatment....1

1.3 Host response to *Mycobacterium tuberculosis* (Mtb) infection....4

1.4 Effects of reactive oxygen species (ROS)....5

1.5 *Mycobacterium tuberculosis* (Mtb) defense mechanisms against reactive oxygen species (ROS)....7

1.5.1 Detoxifying systems against ROS....7

1.5.2 DNA repair systems....9

1.6 DNA glycosylases of base excision repair (BER) in *Mycobacterium tuberculosis* (Mtb)....13

1.7 Nth DNA glycosylase....16

1.8 Gene inactivation by homologous recombination....21

1.9 Aim and objectives of this study: ....22
2. Materials and methods....24
2.1 Bioinformatics analyses....24
2.2 Bacterial strains, plasmids and maintenance of strains ....25
2.2.1 Bacterial strains and culture conditions ....25
2.2.2 Cloning vectors (or plasmids)....28
2.2.3 Assessment of cell viability ....29
2.3 General Molecular Techniques....30
2.3.1 DNA isolation ....30
2.3.1.1 Genomic DNA extraction from Msm ....30
2.3.1.1.1 Large scale genomic DNA extraction ....30
2.3.1.1.2 Small scale genomic DNA extraction ....31
2.3.1.2 Plasmid DNA extraction from *Escherichia coli* (Ec) ....31
2.3.1.2.1 Large scale plasmid DNA extraction ....31
2.3.1.2.2 Small scale plasmid DNA extraction ....31
2.3.2 DNA manipulation....32
2.3.2.1 Enzymatic modification of DNA ....32
2.3.2.1.1 Polymerase Chain Reaction (PCR) amplification....32
2.3.2.1.2 Restriction endonuclease digestion ....33
2.3.2.1.3 Phosphorylation ....34
2.3.2.1.4 De-phosphorylation....34
2.3.2.1.5 Ligation ....34
2.3.3 DNA precipitation....35
2.3.4 The separation of DNA fragments by agarose gel electrophoresis ....35
2.3.5 DNA fragment extraction and purification from agarose gel ....36
2.3.6 Quantification of DNA ....36

2.3.7 Southern blot analyses ..... 36
2.3.8 Transformation of *E. coli* (Ec) DH5α cells ..... 39
2.3.8.1 Preparation of Ec DH5α competent cells using rubidium chloride ..... 39
2.3.8.2 Transformation of chemically competent Ec DH5α cells ..... 39
2.3.9 Electroporation into Msm and Ec ..... 40
2.3.9.1 Preparation of electro-competent Msm and Ec cells ..... 40
2.3.9.2 Electroporation ..... 40
2.3.10 DNA sequencing ..... 41
2.4 The construction and identification of site specific deletion mutants in Msm ..... 41
2.4.1 Construction of the *nth* knockout vector ..... 41
2.4.1.1 Three way cloning ..... 42
2.4.1.2 PacI cassette cloning ..... 42
2.4.2 Identification of single crossover (SCO) and double crossover (DCO) mutants 43
2.5 Genetic complementation ..... 44
2.5.1 Complementation vector ..... 44
2.5.2 SQ RT-PCR ..... 45
2.5.2.1 RNA isolation ..... 45
2.5.2.2 RT or first strand cDNA synthesis ..... 46
2.5.2.3 cDNA amplification ..... 47
2.6 Phenotypic characterization of mutant strains ..... 47
2.6.1 Growth kinetics ..... 47
2.6.2 Sensitivity to oxidative stress using hydrogen peroxide ($H_2O_2$) susceptibility assays ..... 48
2.6.2.1 Assessment of cell viability post $H_2O_2$ treatment ..... 48
2.6.2.2 Measuring intracellular $H_2O_2$ levels ..... 49
2.6.3 Mutational assessments ..... 49

2.6.3.1 Spontaneous mutagenesis ..... 50

2.6.3.1.1 Assessment of mutation frequency ..... 50

2.6.3.1.2 Assessment of mutation rates ..... 50

2.6.3.2 Induced mutagenesis ..... 55

2.6.3.2.1 DNA damage-induced mutagenesis assay ..... 55

3. Results ..... 56

3.1 Bioinformatics ..... 56

3.1.1 Identification of the Nth glycosylase ..... 56

3.2 Generation of the EcΔ*nth*::Msm*nth* strain ..... 59

3.2.1 Assessment of cell viability post $H_2O_2$ treatment of Ec *nth* deficient mutant and its complements ..... 61

3.2.2 Mutation frequency ..... 63

3.3 The construction and identification of site specific deletion mutants in *Mycobacterium smegmatis (*Msm) ..... 64

3.3.1 Construction of the *nth* knockout vector ..... 64

3.3.1.1 Three way cloning ..... 66

3.3.1.2 Cloning of selectable and counter-selectable markers ..... 68

3.3.2 Identification of single crossover (SCO) and double crossover (DCO) events ..... 76

3.4 Genetic complementation ..... 81

3.4.1 Construction of the complementation vector ..... 81

3.5 Phenotypic characterization ..... 84

3.5.1 Growth kinetics ..... 84

3.5.2 Assessment of cell viability post $H_2O_2$ treatment for Msm *nth* mutants ..... 85

3.5.3 Involvement of *nth* in mutagenesis ..... 92

3.5.3.1 Mutation rates to assess spontaneous mutations ..... 93

3.5.3.1.1 Mutation spectrum ..... 95

3.5.3.2 UV induced mutagenesis ..... 100

4. Discussion and Conclusion ..... 104

5. Appendices ..... 116

5.1 Culturing of Bacteria ..... 116

5.1.1 Culture media ..... 116

5.1.2 DNA molecular weight markers ..... 117

5.2 PCR primers used in this study and amplification ..... 118

5.2.1 PCR primers used in this study ..... 118

5.2.2 PCR amplification ..... 120

5.3 Vectors maps used in this study ..... 123

5.3.1 Vectors ..... 123

5.3.2 PacI cassette nucleic acid sequence ..... 125

5.4 Bioinformatic analysis of the Nth DNA glycosylases ..... 126

6. References ..... 127

# List of figures

Fig. 1.1: The effect of ROS and defence mechanisms involved for cell survival.....6

Fig. 1.2: Universal BER pathway.....13

Fig. 1.3: Three-dimensional structure of the Ec Nth showing HhH and FCL motifs..... ..18

Fig. 2.1: An example of the estimation of the *m* value by the Luria-Delbrück distribution.....52

Fig. 3.1: Structural characteristics of the Nth protein.....56

Fig. 3.2: Alignments of the amino acid sequences of the Nth protein from various microorganisms..... .....57

Fig. 3.3: Genomic context of the *nth* DNA glycosylase in Msm, Mtb and Ec.....58

Fig. 3.4: Restriction mapping for pTWEETY::*nth*.....60

Fig. 3.5: PCR verification of four potential Msm *nth* complements in the Ec *nth* mutant strain using NthCF and NthCR primers..... .....61

Fig. 3.6: Assessment of complementation of Ec *nth* mutant by the Msm *nth* gene..... ....62

Fig. 3.7: Complementation of the Ec *nth* knockout mutant (Ec*Δnth*) with Msm *nth*..... .....63

Fig. 3.8: Schematic representation of PCR amplification of upstream and downstream regions of the *nth* gene…….....64

Fig. 3.9: Restriction map analyses of pGEM*Δnth* Us……...65

Fig. 3.10: Restriction map analyses of pGEM*Δnth* Ds……...66

Fig. 3.11: Directional cloning of upstream (Us) and downstream (Ds) regions into the suicide vector (p2NIL).....67

Fig. 3.12: Restriction mapping for p2NIL*Δnth*..... .....68

Fig. 3.13: Addition of PacI cassette containing selectable and counter-selectable markers... ...69

Fig. 3.14: Restriction mapping for a potential p2NILΔ*nth*::pGOAL19 clone……..… ...70

Fig. 3.15: Restriction mapping of p2NILΔ*nth*::2XpGOAL19..... ...71

Fig. 3.16: Restriction mapping for p2NILΔ*nth*::pGOAL19.....72

Fig. 3.17: Restriction mapping of an un-annotated *Msc*I site in the PacI fragment of pGOAL19.....74

Fig. 3.18: *Msc*I restrictions of various isolates of pGOAL19 plasmid DNA..... ....75

Fig. 3.19: Sucrose sensitivity test of p2NILΔ*nth*::2XpGOAL19 and p2NILΔ*nth*::pGOAL19 knockout constructs.....75

Fig. 3.20: Strategy for generating Nth glycosylase deficient mutants.....76

Fig. 3.21: Allelic replacement by homologous recombination…..... .....77

Fig. 3.22: Confirmation of possible *nth* DCOs by PCR amplification.....78

Fig. 3.23: Southern blot upstream analysis of Δ*nth* mutants.....79

Fig. 3.24: Southern blot downstream analysis of Δ*nth* mutants.....80

Fig. 3.25: SQ RT-PCR analysis to confirm *nth* expression of Msm Nth integrants..... ...83

Fig. 3.26: Comparative growth kinetics..... .....85

Fig. 3.27: Assessment of wild type Msm with 2.5 mM $H_2O_2$ at $OD_{600}$ of 0.70, 0.50 and 0.35 over a period of 6 hrs.....86

Fig. 3.28: Assessment of wild type Msm, single *nth* mutant and its respective complemented strain with and without 2.5 mM $H_2O_2$ over a period of 6 hrs..... ..87

Fig. 3.29: Intracellular $H_2O_2$ concentration of untreated and treated Msm, single *nth* mutant and its respective complement strains over a 6 hr period.....89

Fig. 3.30: The survival of early log-phase *nei/fpg/nth* deficient mutants compared to the parental Msm strain when treated with 2.5 mM $H_2O_2$ over a period of 6 hrs..... .....90

Fig. 3.31: The survival of early log-phase *fpg/nei/nth* deficient mutants compared to the parental Msm strain when treated with 2.5 mM $H_2O_2$ over a period of 6 hrs..... .....91

Fig 3.32: Amino acid changes in the RRDR region of the *rpoB* gene in rif resistant mutants isolated from Msm, Fpg/Nei deficient mutants, *nth* deletion mutants and complements.....98

Fig. 3.33: The percentage frequency of amino acid changes identified within the RRDR region of the various *nth* deletion mutants as compared to Msm, Fpg/Nei deficient mutants and complemented strains..... .....99

Fig. 3.34: UV induced mutation frequency of combinatorial Fpg/Nei/Nth deficient mutants...101

Fig. 5.1.2.1: The DNA molecular weight markers..... .....117

Fig. 5.2.2.1: PCR amplification of the Fpg/Nei deficient mutants in the absence and presence of Nth.....121

Fig. 5.2.2.2: PCR amplification of chromosomal DNA from possible Msm *nth* complements using NthCF and NthCR primers.....122

Fig. 5.3.1.1: pTWEETY vector used for constructing the complementation vector using the *Ecl136*II site.....123

Fig. 5.3.1.2: pGEM3Zf(+) vector used to clone PCR amplified Msm *nth* upstream and downstream fragments at the *Sma*I site.....123

Fig. 5.3.1.3: p2NIL suicide vector used to clone upstream and downstream *nth* fragments simultaneously to generate the *nth* deleted allele at the *Hind*III and *Kpn*I sites..... .....124

Fig. 5.3.1.4: pOLYG replicating vector used in electroporations as a positive control to determine the transformation efficiency.....124

Fig. 5.3.2.1: PacI cassette nucleic acid sequence..... ..125

# List of Tables

Table 2.1: Bacterial strains used in this study..... .....26

Table 2.2: Bacterial plasmids used in this study..... .....28

Table 3. 1: Spontaneous mutation rates resulting in rif resistance of the various *nth* knockout mutant strains compared to Msm as determined by fluctuation analysis..... .....95

Table 5.2.1.1: Primers used for PCR Amplification.....118

Table 5.2.1.2: Sequencing primers used in this study..... .....119

Table 5.2.1.3: Primers used to distinguish between wild type and mutants.....119

Table 5.2.1.4: Primers used for gene expression analysis by SQ RT-PCR.....120

Table 5.4.1: Similarity of the Msm Nth protein compared to several related microorganisms using BLASTp..... .....126

# Nomenclature

| | |
|---|---|
| 8OxoG | 8-oxo-7,8-dihydroguanine |
| A | Adenine |
| AIDS | Acquired Immune Deficiency Syndrome |
| amp | Ampicillin |
| ANOVA | Analysis Of Variance |
| AP | Apurinic/Apyrimidinic |
| *aph* | Gene encoding kanamycin resistance |
| BCG | Bacille Calmette-Guerin |
| BLAST | Basic Local Alignment Search Tool |
| bp/s | Base pair/s |
| BSA | Bovine Serum Albumin |
| C | Cytosine |
| DIG | Dioxygen |
| DNA | Deoxyribonucleic acid |
| DNase | Deoxyribonuclease |
| dNTPs | Deoxyribonucleotide triphosphates |
| dTTP | Deoxythymine triphosphate |
| dUTP | Deoxyuridine triphosphate |
| EDTA | Ethylenediaminetetra acetic acid |
| Fapy | Formamidopyrimidine |
| Fig/s. | Figure/s |
| *fpg1* | *f1* |
| *fpg2* | *f2* |
| G | Guanine |

| | |
|---|---|
| HIV | Human Immuno-deficiency Virus |
| hr/s | hour/s |
| $kan^R$ | Kanamycin resistance |
| Kb | Kilobase pairs |
| *lacZ* | Gene encoding β galactosidase activity |
| Mb | Megabase pairs |
| min | minutes |
| NADPH | Nicotinamide Adenine Dinucleotide Phosphate |
| *nei1* | *n1* |
| *nei2* | *n2* |
| NY | New York |
| Ori | Origin of replication |
| P | Phosphate |
| PCR | Polymerase Chain Reaction |
| RNA | Ribonucleic acid |
| RNase | Ribonuclease |
| rpm | Revolutions per minute |
| RRDR | Rifampicin Resistance Determining Region |
| *sacB* | Gene encoding levansucrase |
| T | Thymine |
| TE | Tris EDTA |
| Tg | *cis*- and *trans*-thymine glycol |
| Tween | Polyoxyethylene sorbitan monooleate |
| UV | Ultraviolet |
| USA | United States of America |
| v/v | Volume/volume |

| | |
|---|---|
| WHO | World Health Organization |
| X-gal | 5-bromo-4-chloro-3-indolyl- β-galactoside |
| β | Beta |
| μ | Mutation rate |

# 1. Introduction

## 1.1 The Tuberculosis scourge

In 1882 Robert Koch identified the bacillus *Mycobacterium tuberculosis* (Mtb) that causes tuberculosis (TB) leading to death in the human host (Koch, 1998). TB incidences from the 1950's were greatly reduced due to the discovery of antibiotics: isoniazid (1952), pyrazinamide (1954), ethambutol (1962) and rifampicin (1963) which are still currently used in the Directly Observed Treatment, Short course (DOTS) program for TB control (Herzog, & Basel, 1998). Today TB is the world's second most deadly infectious disease responsible for 1.4 million mortalities in 2011 despite TB being a fully treatable disease (WHO, 2012). The staggering rate of infection can be attributed to: an antiTB vaccine that has variable efficiency, a chemotherapeutic regimen that is more than 30 years old and ineffective against Multi- and Extreme Drug Resistant (MDR and XDR respectively) strains, HIV co-infection and the ability of Mtb to survive within the host despite the immune response.

## 1.2 Available TB treatment

The only preventative measure against TB is the Bacille Calmette-Guerin (BCG) vaccine derived from an attenuated *Mycobacterium bovis* (*M. bovis*) strain. BCG is an inexpensive, safe and the most administered human vaccine in the world that protects infants from primary TB infection (Bloom *et al*., 1994). For such a widely used century old vaccine the mechanism of protective immunity is unknown and the efficacy varies from 0 to 80% according to high or low TB burdened areas, host genetics, immunology and age (Bloom & Murray, 1992; Bloom *et al*., 1994). The protective effect of BCG is highest at the age of vaccination and weakened with age failing to protect adults against TB infection therefore

relying on antiTB agents to treat and eradicate the disease (Fine, 1989a; b; Rodrigues *et al*., 1993; Colditz *et al*.,1994).

The antiTB drugs used in the short-course chemotherapy remains to date the most effective agents against active TB and are hence termed first line drugs (Dye *et al*., 1999; Fox *et al*., 1975). Briefly, the six month drug regimen includes two months of ethambutol, pyrazinimide, isoniazid and rifampicin followed by four months of isoniazid and rifampicin and is part of a multifaceted program known as DOTS (Iseman *et al*.,1993; Bayer, 1995; Du Toit *et al*., 2006). The cornerstone and definition of DOTS is the supervision of standardized chemotherapy aimed to prevent relapse and resistance to antiTB drugs by improving patient adherence to the chemotherapy (Frieden *et al*., 2007; Garner *et al*., 2007).

The efficacy of DOTS is measured by cure rates associated with patient adherence to drug therapy. In South East Asia where DOTS is correctly implemented cure rates are 80% (Kamolratanakul *et al*.,1999). In contrast the poor implementation of the program is reflected by the low cure rates due to patient non-compliance to treatment regardless of the geographical settings with high (Zwarenstein *et al*.,1998; Zwarenstein *et al*., 2000; Walley *et al*., 2001) or low TB burden (Chaisson *et al*., 2001; Malotte *et al*., 2001). Patient non-compliance to the lengthy drug regimen has lead to acquired drug resistance. This form of drug resistant TB harbors drug resistant bacilli that can be transmitted to healthy individuals developing primary drug resistant disease (Chaulet *et al*., 1995; Muttil *et al*., 2009).

Recently, South Africa was identified as one of 27 countries with an escalating number of drug resistant MDR and XDR TB cases reported from 2007 (WHO, 2011). These resistant forms of TB are resistant to first line drugs and hence need to be treated with second line drugs (Meya & McAdam, 2007). There are six classes of second line drugs: aminoglycosides (amikacin and kanamycin), polypeptides (capreomycin, viomycin and enviomycin), fluoroquinolones (ciprofloxacin, levofloxacin and moxifloxacin), thioamides (ethionamide and prothionamide), cycloserine (closerin) and terizidone (Du Toit *et al*., 2006; Meya & McAdam, 2007). MDR is described as resistance to isoniazid and rifampicin, the two most effective front line drugs used for active TB, while XDR is defined as Mtb that is resistant to the first line drugs: isoniazid, rifampicin, a fluoroquinolone and to any of these injectable drugs: amikacin, capreomycin, or kanamycin (Meya & McAdam, 2007). Second line drugs are more toxic to the patient with reduced efficacy hence, prolonged treatment is required which is expensive (Du Toit *et al*., 2006; Meya & McAdam, 2007).

A further assault to the current TB treatment is the lethal synergy of TB and HIV co-infection, particularly in sub-Saharan Africa where 73% of new active TB incidences are associated with HIV infection (Harrington, 2010). TB is a cofactor in fuelling the progression of HIV infection to AIDS by increasing HIV replication at sites of TB infection (Nakata *et al.*, 1997; Hoshino *et al.*, 2002; Hoshino *et al.*, 2007). While HIV impairs the ability of the host immune system to control and contain the TB infection by reducing CD4 cells (Glynn *et al.*, 2008; Geldmacher *et al*., 2010) and manipulating host immune cells that aggregate around the site of TB infection (Patel *et al*., 2009; Kumawat *et al*., 2010). The above factors corroborate the demand for improved treatment through the

development of novel drugs by understanding the mechanisms of Mtb survival within the host.

## 1.3 Host response to *Mycobacterium tuberculosis* (Mtb) infection

Primary infection is initiated by the inhalation of infectious aerosols coughed into the environment by an individual with active TB (Kaufmann *et al*., 2005; Ramakrishnan, 2012). In response to invading bacilli in the lung tissue, alveolar macrophages elicit a mild inflammation triggering the production of dendritic cells and macrophages (Kaufmann *et al*., 2005; Ramakrishnan, 2012). These antigen presenting immune cells stimulate the aggregation and organization of an array of immune cells at the site of infection to contain the bacilli, creating a granuloma, a hallmark of Mtb infection (Flynn & Chan, 2001; Fenton *et al*., 2005; Cooper, 2009; Ramakrishnan, 2012). Ironically various stresses such as: hypoxia, low pH, starvation and reactive nitrogen and oxygen species (RNS and ROS) within the granuloma do not kill the bacilli, instead the Mtb bacilli infect these immune cells (Anderson, 2006; Ramakrishnan, 2012). Intraphagosomal bacilli use the granuloma as a protective niche for intracellular replication and growth (Kaufmann *et al.*, 2005; Ulrichs & Kaufmann, 2006; Ramakrishnan, 2012). Only 5% of infected individuals develop active TB, whereas the remaining 95% develop latent or asymptomatic TB (i.e. no disease symptoms) (Anderson, 2006). The granuloma and intraphagosomal bacilli can persist within the host until the immune system is compromised inducing unimpeded replication of quiescent Mtb leading to secondary infection and active TB disease (Anderson, 2006; Ramakrishnan, 2012).

Interestingly, RNS and ROS generated by activated macrophages during infection induce Mtb stress mechanisms that combat the harsh environment and allow continued survival

and proliferation within the host (Ehrt & Schnappinger, 2009). Stimulated macrophages produce inducible nitric oxide synthase and phagocyte oxidase that form nitric oxide and hydrogen peroxide ($H_2O_2$) to generate RNS and ROS (Ehrt & Schnappinger, 2009). Nitric oxide in addition to hypoxia, low pH and carbon monoxide are all stresses within the activated macrophage that induce the DosR/S/T regulon which consists of about 47 genes regulating fatty acid metabolism (Dasgupta *et al*., 2000; Kumar *et al*., 2008; Leistikow *et al*., 2010; Voskuil *et al*., 2011). The metabolic shift assists bacilli to enter into a non-replicating state defined as dormancy (Leistikow *et al*., 2010). This transition ensures the prolonged survival of Mtb within the host despite its harsh environment. However, $H_2O_2$ induces the expression of several genes belonging to detoxification systems and robust DNA repair pathways to defend against the bactericidal effects of ROS (Voskuil *et al*., 2011).

## 1.4 Effects of reactive oxygen species (ROS)

During an Mtb infection ROS is generated from host immune cells (Kaufmann *et al*., 2005; Kumar *et al*., 2011) and as a consequnce of aerobic respiration during Mtb growth and replication (Fig. 1.1). Aerobic respiration uses oxygen as the final electron acceptor and when completely reduced i.e. oxygen accepts four electrons, produces water as a by-product (Scandalios, 2005; Kumar *et al*., 2011). When oxygen is not completely reduced (i.e. oxygen accepts one or two or three electrons) then ROS in the form of radicals such as: superoxide anions ($O^-$), $H_2O_2$ and hydroxyl ions ($OH^-$) are produced (Scandalios *et al.*, 1997; Scandalios, 2005; Kumar *et al*., 2011).

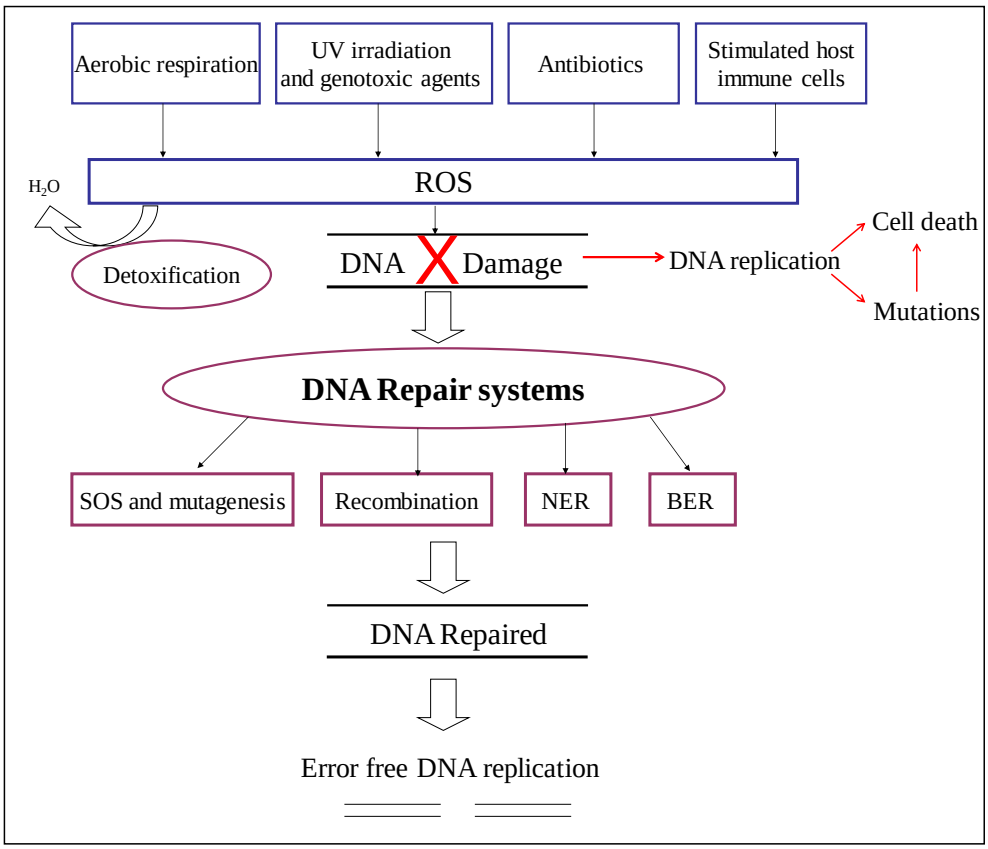



**Fig. 1.1: The effect of ROS and defence mechanisms involved for cell survival.** Failure of detoxification and DNA repair systems to rectify the effects of ROS ultimately results in cell death (Adapted from Krokan *et al*., 1997).

The level or concentration of ROS is maintained within the organism by neutralising mechanisms or ROS scavenging systems (Scandalios, 2005; Kumar *et al*., 2011). However, when these detoxifying systems fail to reduce ROS within the organism, ROS is accumulated and this state is known as oxidative stress (Scandalios, 2005; Kumar *et al*., 2011). High levels of ROS destroy cell protein and lipids and cause numerous types of DNA damage including single- and double strand breaks, AP sites and base modification (Demple & Harrison, 1994). If this DNA damage is not detected and repaired by DNA repair systems, replication occurs via error prone DNA repair systems leading to either cell death or chromosomal mutagenesis (Fig. 1.1). In addition, chromosomal mutagenesis can also be as a result of ROS production in response to exposure to antimicrobial agents such as antibiotics (Dwyer *et al*., 2009). Increased ROS production in response to antibiotic

treatment was shown to cause metabolic (Kohanski *et al*., 2007) and physiological (Dwyer *et al*., 2007) changes that induced antibiotic resistance and cell death. Kohanski *et al*., 2010 showed an increase in ROS levels increased the mutation rate in *Escherichia coli* (Ec) following antibiotic treatment with β-lactams, fluoroquinolones and aminoglycosides. This shows that oxidative stress is not the only contributor to mutagenesis but also implicates DNA repair systems and failure thereof. Therefore, understanding mechanisms of combating oxidative stress in Mtb is necessary for understanding mechanisms of mutagenesis and the prolonged survival within the host.

## 1.5 *Mycobacterium tuberculosis* (Mtb) defense mechanisms against reactive oxygen species (ROS)

### 1.5.1 Detoxifying systems against ROS

Since bacilli are exposed to various sources of ROS (i.e. $H_2O_2$, $O^-$ and $OH^-$) the survival of bacilli depends on detoxifying defence mechanisms against ROS and against damaged DNA (Fig. 1.1). *Escherichia coli* (Ec) is equipped with two regulators: SoxRS and OxyR that induce the expression of various reactive oxygen scavenging proteins upon oxidative stress (Farr & Kogoma, 1991; Imlay, 2008). The SoxRS regulates the expression of two superoxide dismutases SodA and SodB that dismutate $O^-$ anions (McCord & Fridovich, 1969; Fridovich, 1989; Touati, 1988). While OxyR regulates the expression of three catalases KatG, KatE and KatF that degrade $H_2O_2$ (Claiborne *et al.*, 1979; Loewen & Triggs-Raine, 1984; Triggs-Raine & Loewen, 1987), and two alkyl hydroperoxide reductases AhpC and AhpF that neutralise various organic hydroperoxides (Greenberg & Demple, 1988; Storz *et al.*, 1989). In addition, Ec produce glutathione to create a strong reducing intracellular environment (Farr & Kogoma, 1991; Imlay, 2008). Glutathione

serves as an antioxidant by the oxidizing electron carriers such as NADPH and reducing $H_2O_2$ (Farr & Kogoma, 1991; Imlay, 2008).

In contrast, Mtb is deficient in SoxRS and OxyR regulators (Cole *et al*., 1998) but constitutively expresses KatG (Wayne & Diaz, 1982; Zhang *et al*., 1992), AhpC (Sherman *et al*., 1999) and SodC and SodA (Kusunose *et al*., 1976; Dussurget *et al*., 2001), implying that Mtb is always equipped for detoxifying oxidative stress. Furthermore, low molecular weight thiols namely ergothioneine and mycothiol act as redox buffers defending mycobacteria against oxidative stress (Kumar *et al*., 2011). In particular, mycothiol is the antioxidant produced by Mtb that can be oxidised and reduced in a similar manner as glutathione (Newton & Fahey, 2002; Buchmeier *et al*., 2003; Kumar *et al*., 2011). Mycobacteria has four genes *mshA-D* that contribute to the production of mycothiol and the deletion of any of the four genes results in increased sensitivity to oxidative stress (Newton *et al*., 1999; Buchmeier *et al*., 2006; Rawat *et al*., 2007; Kumar *et al*., 2011). In particular *mshC* mutants are not viable, implying that mycothiol is required for the survival of Mtb (Sareen *et al*., 2003; Buchmeier *et al*., 2003; Kumar *et al*., 2011). Both mycobacteria and Ec are equipped with thioredoxin, another antioxidant that operates in the same manner as mycothiol and glutathione to detoxify ROS (Prinz *et al*., 1997; Arner *et al*., 2000; Kumar *et al*., 2011). Ec consists of two thioredoxins (TrxA and TrxC) and a reductase (Prinz *et al*., 1997; Arner *et al*., 2000), while Mtb encodes three (TrxA, TrxB and TrxC) thioredoxins with a single reductase (Zhang *et al*., 1999; Kumar *et al*., 2011). However, the deletion of any of the mycobacterial thioredoxin encoding genes does not hinder Mtb survival (Akif *et al*., 2008), suggesting that thioredoxin is not the primary mechanism of ROS neutralisation (Prinz *et al*., 1997; Arner *et al*., 2000; Kumar *et al*., 2011). Therefore, it is evident that Ec and Mtb are equipped with oxidative stress neutralising mechanisms that work differently.

To date there exists no neutralising mechanism for $OH^-$ ions which are implicated in the damage of purine and pyrimidine bases (Scandalios, 2005).

### 1.5.2 DNA repair systems

Bioinformatic analyses of the Mtb genome (Cole *et al*., 1998), identified homologs for several DNA repair systems (Fig. 1.1) including genes for: SOS repair, recombination, mutagenesis, nucleotide excision repair (NER) and base excision repair (BER) but lacks homologs for mismatch repair (Mizrahi & Anderson, 1998). Despite this heritable DNA repair deficiency, mycobacteria however, show no signs of hypermutability when cultured *in vitro* under standard culture conditions (David & Newman, 1971).

A plethora of DNA lesions and double and single strand breaks are the after effects of an organism being exposed to genotoxic compounds, alkylating agents or UV irradiation (Wallace, 2002; Chow & Courcelle, 2007). In Ec high fidelity replicative polymerases such as PolA cannot resume replication past these lesions creating a block or halt in replication, a consequence that is lethal (Friedberg *et al*., 1995; Friedberg *et al*., 2002; Goodman, 2002). Two DNA repair systems the SOS response and mutagenesis promotes error prone replication for the survival of the organism (Friedberg *et al*., 1995; Little & Mount, 1982; Goodman, 2002). These repair systems uses DNA polymerases with low replication fidelity that are characterized in Ec as Pol III, IV and V (Friedberg *et al*., 2002; Friedberg *et al*., 2005). Mtb possess homologs of the DNA polymerases Pol II, III and IV (Cole *et al*., 1998). In response to DNA damage a Pol III polymerase DnaE consisting of two functional homologs (DnaE1 and DnaE2) were upregulated (Cole *et al*., 1998; Boshoff *et al*., 2003) but not the PolIV DinB (Kana *et al*., 2010). Surprisingly, DnaE1 was shown to be the high fidelity replicative polymerase that is essential in Mtb as its deficiency is fatal

(Boshoff *et al*., 2003). However, DnaE2 was shown to be the low fidelity polymerase as it was associated with inducing mutagenesis for survival of the organism (Boshoff *et al*., 2003). The deletion of DnaE2 minimised the organisms' ability to tolerate DNA damage resulting in drug resistance (Boshoff *et al*., 2003). Furthermore, imuA and B proteins are equally required for the functioning of DnaE2 (Warner *et al*., 2010). In Mtb the SOS repair system expresses genes encoding the following proteins: RecA, RecBCD complex, RuvABC complex and UvrA, UvrB, UvrC, UvrD (Cole *et al*., 1998; Mizrahi & Anderson, 1998) that are involved in recombination pathways and NER.

The Recombination repair pathway has two repair systems that repair double strand breaks as a result of DNA damage: homologous recombination and non-homologous end joining (Kowalczykowski *et al*., 1994; Warner & Mizrahi, 2011). Homologous recombination repair involves the exchange of two complementary strands of DNA. This repair system relies on a homologous chromosomal copy as template for DNA synthesis and hence is considered not to be an error prone system (Kowalczykowski *et al*., 1994; Warner & Mizrahi, 2011). The RecBCD complex is made up of a helicase and nuclease proteins that unwinds double stranded DNA into single strands, which are cut off at a specific nucleotide sequence known as the Chi site (Kowalczykowski *et al*., 1994; Singleton *et al*., 2004; Warner & Mizrahi, 2011). The single stranded DNA is bound by the RecA protein that acts as a signal to induce the SOS response; in particular the RuvABC complex and searches for homologous duplex DNA (Kowalczykowski *et al*., 1994). Upon recognition of homologous DNA, RecA coated single stranded DNA invades the duplex recipient DNA (Kowalczykowski *et al*., 1994; Warner & Mizrahi, 2011). One of the recipient strands forms a bulge which is cleaved by the endonuclease from the RuvABC complex (Kowalczykowski *et al*., 1994). The complex then exchanges the strands which forms a

structure known as a holiday junction followed by the ligation of swapped strands (Kowalczykowski *et al*., 1994).

Whereas, non-homologous end joining repair is a less complicated process that ligates two strands of DNA together irrespective of whether it is homologous or not (Shuman & Glickman, 2007; Warner & Mizrahi, 2011). In Mtb a DNA binding protein known as Ku, is required for the ligase (LigD) to join and seal the two pieces of DNA together (Della *et al*., 2004; Shuman & Glickman, 2007; Warner & Mizrahi, 2011). Ku is vital for LigD function, as the deletion of the Ku DNA binding protein drastically reduces the efficiency of this process (Gong *et al*., 2005; Shuman & Glickman, 2007; Warner & Mizrahi, 2011). LigD is a multifunctional enzyme that catalyzes polymerase, ligase and phophoesterase reactions (Della *et al*., 2004). Hence, LigD can modify the ends of two strands until compatible making this repair system error prone (Shuman & Glickman, 2007).

Nucleotide excision repair (NER) identifies and removes bulky DNA lesions generated from UV irradiation and genotoxic compounds (Friedberg *et al*., 1995; Kurthkoti & Varshney 2011; Kurthkoti & Varshney, 2012). This repair pathway executes its function using several Uvr proteins (Cole *et al*., 1998). The Uvr proteins have low base specificity that act in combination usually forming a complex because it removes multiple lesions (Friedberg *et al*., 1995; Kurthkoti & Varshney 2011; Kurthkoti & Varshney, 2012). UvrA and UvrB identify and bind to the DNA lesion forming a UvrAB complex (Sancar, 1996; Kurthkoti & Varshney 2011; Kurthkoti & Varshney, 2012). UvrA is released when UvrC binds to UvrB (Sancar, 1996; Kurthkoti & Varshney 2011; Kurthkoti & Varshney, 2012). The UvrBC complex cleaves and excises the surrounding upstream and downstream regions of the DNA damage (Sancar, 1996; Kurthkoti & Varshney 2011; Kurthkoti &

Varshney, 2012). Thereafter, UvrD removes the UvrBC complex with the lesion surrounded by about 12-13 nucleotides and creates a gap (Sancar, 1996; Kurthkoti & Varshney 2011; Kurthkoti & Varshney, 2012). The gap is filled by DNA polymerase and ligase (Sancar, 1996; Kurthkoti & Varshney 2011; Kurthkoti & Varshney, 2012). In Mtb the *uvrB* gene is essential *in vivo* (in mice) for infection and *in vitro* (human macrophages) for survival (Darwin & Nathan, 2005; Kurthkoti & Varshney 2011; Kurthkoti & Varshney, 2012). Mtb possess two homologues of UvrD proteins - UvrD1 and UvrD2 (Cole *et al*., 1998; Mizrahi & Anderson, 1998). UvrD1 is a Ku dependent protein and is bifunctional (Sinha *et al*., 2007; Curti *et al*., 2007; Guthlein *et al*., 2008). It has helicase and DNA dependent ATPase activities that allow this protein to repair double and single strand breaks (Sinha *et al*., 2007; Curti *et al*., 2007; Guthlein *et al*., 2008). Hence, UvrD1can function within NER and non-homologous end joining repair (Shuman & Glickman, 2007; Kurthkoti & Varshney 2011; Kurthkoti & Varshney, 2012). UvrD2 on the other hand is a helicase specific to NER that is essential for Mtb survival when exposed to UV irradiation (Sinha *et al*., 2007; Boshoff *et al*., 2003; Kurthkoti & Varshney 2011; Kurthkoti & Varshney, 2012).

Another excision repair system is base excision repair (BER) that is a multi-protein, multi-step repair pathway (Friedberg *et al*., 1995; Kurthkoti & Varshney 2011; Kurthkoti & Varshney, 2012). BER involves the identification and excision of oxidized bases (represented by X, Fig. 1.2A) by DNA glycosylases generating an AP (apurinic/apyrmidinic) site, the conversion of the 3' termini to a hydroxyl group (OH) by AP endonucleases (Fig. 1.2B) and the removal of the phosphate (P) backbone by a diesterase (dRpase, Fig. 1.2C) (Kurthkoti & Varshney 2011; Kurthkoti & Varshney, 2012). These steps provide the substrate for DNA polymerase and ligase (Fig. 1.2D) to insert new

correct nucleotides thereby repairing the damaged DNA (Kurthkoti & Varshney 2011; Kurthkoti & Varshney, 2012). Since DNA glycosylases initiate BER and in mycobacteria several of these genes are duplicated (Kurthkoti & Varshney 2011; Kurthkoti & Varshney, 2012) suggesting an important role for these enzymes in DNA repair which demands their further elucidation.

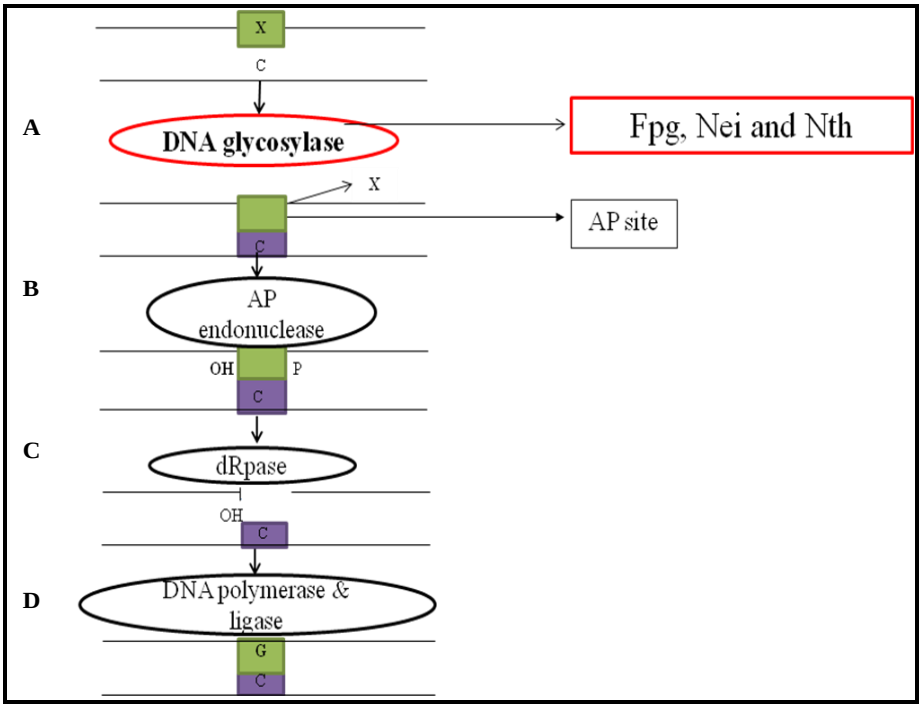



**Fig. 1.2: Universal BER pathway. A.** DNA glycosylases such as Fpg, Nei and Nth (section 1.6) detect and remove the DNA modification or lesion represented as X creating an AP (apurinic/apyrimidic) site. **B.** AP endonucleases cleave the AP site generating an OH group. **C.** dRpase excises the phosphate. **D.** DNA polymerase and ligase synthesizes and seals the correct base (modified from Friedberg *et al*., 1995).

## 1.6 DNA glycosylases of base excision repair (BER) in *Mycobacterium tuberculosis* (Mtb)

The first line of defence against oxidatively damaged DNA lesions are DNA glycosylases (Fig. 1.2A). Generally DNA glycosylases are proteins that consist of two domains that make contact with the oxidatively damaged DNA lesion and a catalytic residue that lies deep within these domains to possibly remove the lesion (Krokan *et al*., 1997). DNA

glycosylases either belong to the helix-hairpin-helix (HhH) or the Fpg/Nei super-family based on sequence and structural homology (Eisen & Hanawalt, 1999).

The endonuclease III (Nth) (Thayer *et al*., 1995; Kuo *et al*., 1992a; b), alkyladenine (AlkA) (Hollis *et al*., 2000) and MutY (Fromme *et al*., 2004) glycosylases belong to the HhH super-family as all these proteins have a HhH motif. These glycosylases are structurally similar but have different substrate specificities. The Nth glycosylase mainly repairs oxidized pyrimidine lesions (Katcher & Wallace, 1983; Breimer & Lindahl, 1984; Dizdaroglu *et al*., 1993; Thayer *et al*., 1995; Dizdaroglu *et al*., 2000) while AlkA repairs alkylated bases such as methyl-adenine (Au *et al*., 1989; Hollis *et al*., 2000) and MutY recognizes A mis-paired with 8-oxo-7,8-dihydroguanine (8-oxoG) (Masoaka *et al*., 1999; Fromme *et al*., 2004).

Similarly, the Fpg/Nei super-family contains the formamidopyrimidine (Fpg/MutM/Fapy) and endonuclease VIII (Nei) glycosylases. Both these glycosylases have an N terminal that consists of β sheets and a C terminal which has a helix-two turn-helix (H2TH) and zinc finger motif (Zharkov *et al*., 2003). The main substrates of Fpg are oxidized purine bases (i.e. G and A) such as 8-oxoG which is the most common mutagenic lesion in the mycobacterial high GC rich genomes producing G→T and G→C mutations (Shibutani *et al*., 1991; Cheng *et al*., 1992). Nei is known to also repair oxidized purine lesions (Blaisdell *et al*., 1999) but mainly recognizes oxidized pyrimidines (i.e. T and C) such as *cis*- and *trans*-thymine glycol (Tg), which is a lesion that is not mutagenic (Hayes *et al*., 1988) but blocks DNA replication polymerases (Ide *et al*., 1985; McNulty *et al*., 1998). Both Nei and Nth recognize oxidized pyrimidine lesions and therefore, these glycosylases are functionally similar. This substrate overlap between the various glycosylases suggests a

possible adaptation which perhaps allows for increased survival of Mtb under oxidative stress conditions.

Bioinformatics analysis of the genomic sequence for the pathogenic Mtb and its non-pathogenic relative *Mycobacterium smegmatis* (Msm) showed a unique duplication of the Fpg/Nei homologs. Two Fpg (encoded by *fpg*I and *fpg*II) and two Nei (encoded by *nei*I and *nei*II) DNA glycosylases were identified in both Mtb and Msm in the same chromosomal context (Sidorenko *et al*., 2008; Goosens, 2008 MSc; Olsen *et al*., 2009; Guo *et al*., 2010). These glycosylases have maintained the necessary domains for Fpg and Nei protein function and showed strong resemblances to each other as well as to previously characterised Fpg family of DNA glycosylases (Zharkov *et al*., 2003). However, the 5' end of *fpg*II is truncated in Mtb and hence it lacks the structural properties that confer glycosylase activity even though the protein can bind to the substrate (Sidorenko *et al*., 2008).

The fifth addition to the unique duplication of Fpg/Nei glycosylases identified in Mtb was the Nth glycosylase (Guo *et al*., 2010). Bioinformatic analyses of the Mtb Nth glycosylase encoded by a single *nth* gene showed structural similarities to the Nth super-family (Guo *et al*., 2010). Biochemical characterisation of the Fpg/Nei/Nth glycosylases in Mtb verified that these glycosylases repair oxidized purine and pyrimidine lesions and resemble their Ec counterparts (Guo *et al*., 2010). Physiological analyses of Fpg/Nei glycosylases in mycobacteria show that they are involved in cellular defence against oxidative stress. However, physiological characterization of the Nth glycosylase has not yet been attempted in mycobacteria. The fact that Mtb have multiple homologs of DNA glycosylases

(Fpg/Nei/Nth) with some duplication suggests that this bacterium is well equipped to survive changing environments and stressful conditions in the granuloma.

## 1.7 Nth DNA glycosylase

The Nth glycosylase has been conserved in achaea, bacteria and eukaryotes (Eisen & Hanawalt, 1999) suggesting that this repair protein was mandatory for survival in early evolution. Organisms with a minimal genome sequence only retain housekeeping genes that are essential for survival and maintenance so as an example *Mycobacterium leprae* (*M. leprae*) which has a minimal genome size of 3.3 Mb has retained *nth* but not *nei* (Cole *et al*., 2001). Similarly, the opportunistic microorganism *Haemophilus influenzae* with a 1.8 Mb genome has also retained the *nth* and *fpg* genes but not the *nei* gene (Fleischmann *et al*., 1995). The intracellular pathogen *Helicobacter pylori* (*H. pylori*) with a genome of 1.7 Mb only encodes for the Nth glycosylase within the BER system and lacks homologs of Fpg and Nei glycosylases (Tomb *et al*., 1997). Deletion of Nth in *H. pylori* reduces intraphagosomal survival and colonisation of the host (i.e. murine model) (O'Rourke *et al*., 2003).

DNA repair systems have evolved to accommodate increased environmental stresses such as oxidative stress imposed on organisms. As an adaptation many organisms have more than one Nth homolog and a deletion in any one homolog has extreme consequences. The deletion of one of the two Nth glycosylases in *Saccharomyces cerevisiae* results in hypersensitivity to oxidizing agents (Eide *et al*., 1996; Augeri *et al*., 1997; You *et al*., 1998). Interestingly, *Deinococcus radiodurans* an extremophilic bacterium that can survive extreme DNA damage has three Nth homologs which are all required for genome maintenance under oxidative stress (Hua *et al*., 2012).

Originally, the Nth glycosylase was described as an endonuclease for its nicking activity on X- irradiated (Strniste & Wallace, 1975) and heavily UV-irradiated DNA in Ec (Radman, 1976; Gates & Linn, 1977). The endonuclease activity was shown to be due to DNA glycosylase and AP lyase catalytic events (Baily & Verly, 1987; Kim & Linn, 1988). The DNA glycosylase identifies and removes the damaged base by cleaving the glycosidic bond resulting in an AP site and a free base-sugar which is then cleaved by an AP lyase via β elimination (Baily & Verly, 1987; Kim & Linn, 1988; Mazumder *et al*., 1991). The subsequent activity of this bifunctional enzyme triggers a cascade of repair enzymes starting with an AP endonuclease and diesterase generating a substrate for DNA polymerase and ligase to replace the damaged base with the correct nucleotide.

The over expression of the *nth* gene in Ec (Cunningham & Weiss, 1985) and the use of heavily UV irradiated DNA substrate (Radman, 1976) were techniques performed to isolate the Nth protein for structural (Kuo *et al*., 1992a; b) and biochemical characterization (Demple & Linn, 1980; Katcher & Wallace, 1983; Breimer & Lindahl, 1984; Armel *et al*., 1977; Gates & Linn, 1977). Structurally, the Nth protein consists of α helices forming continuous loops to possibly facilitate binding to DNA (Fig. 1.3, Kuo *et al*., 1992a; b; Thayer *et al*., 1995). This structure may be the reason that this is the only glycosylase protein that binds to the strand containing the oxidatively damaged lesion as well as the complementary strand. A cleft between the N and C domains is known as the DNA binding pocket (illustrated in Fig. 1.3, Thayer *et al*., 1995). Within the DNA binding pocket is the helix hairpin helix loop (HhH, Thayer *et al*., 1995) that contains a glycine-proline-aspartate (GPD) domain followed by the four iron- four sulfur cluster loop (FCL, Asahara *et al*., 1989; Thayer *et al*., 1995) which are characteristic features of the Nth super-family (illustrated in Fig.1. 3). The sharp bends of DNA binding pocket tries to

separate the oxidatively damaged pyrimidine base from the DNA backbone (Thayer *et al*., 1995).

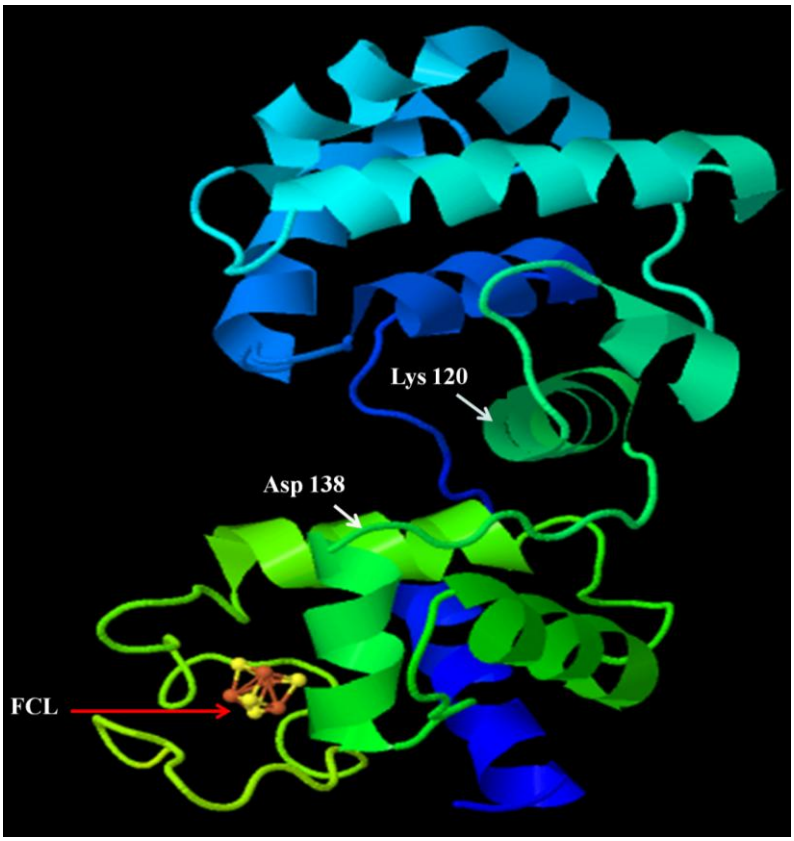



**Fig. 1.3: Three-dimensional structure of the Ec Nth showing HhH and FCL motifs.** The HhH motif is coloured in dark green containing the Lys 120 active residue, while the GPD is highlighted in bright green with the Asp 138 catalytic residue. The FCL motif is shown as a yellow and orange cluster (adapted from Jmol (PDB accession number 2ABK) - http://www.rcsb.org/pdb/explore/jmol.do?structureId=2ABK&bionumber=1).

For the removal of the damaged DNA base the HhH motif harbors two catalytic residues for this purpose - Lysine (Lys) 120 and Aspartate (Asp) 138 that is within the GPD domain (Fig. 1.3, Kuo *et al*., 1992a; b; Thayer *et al*., 1995). Importantly, this motif allows the catalytic Lys 120 and Asp 138 residues that lie deep within the DNA binding pocket to make contact with the damaged base (Thayer *et al*., 1995; Cunningham, 1997). As a result the Asp 138 residue deprotonates the Lys 120 residue producing a nucleophile (Thayer *et al*., 1995; Krokan *et al*., 1997). The nucleophile binds to the C1 of the damaged base

generating a Schiff's base and the the damaged base is removed by the nicking of the glycosidic bond (Thayer *et al*., 1995; Krokan *et al*., 1997). Protonation of the Lys 120 residue allows the protein to detach the DNA backbone followed by β elimination of the adjacent base (C2) (Thayer *et al*., 1995; Krokan *et al*., 1997). Therefore, the HhH motif is essential for protein-DNA interactions and the removal of damaged bases (Kuo *et al*., 1992a; b; Thayer *et al*., 1995).

The FCL motif consists of four irons and four sulfur residues that are linked by cysteines (Fig. 1.3, Cunningham *et al*., 1989). It is located a distance away from the HhH motif indicating that this motif is not involved in the binding and removal of damaged bases. Fu *et al*., (1992) verified this observation by analysing changes in the structure of the FCL motif as damaged pyrimidine bases were added. As shown by Raman resonance spectroscopy the addition of substrate to the motif did not change the structure of the FCL motif (Fu *et al*., 1992). In addition, oxidation and reduction of the cluster showed no shift in redox potential and was deemed as redox inactive (Fu *et al*., 1992). Rather the structure of the motif was thought to play a role in facilitating DNA binding and protein stabilization during enzymatic events. Point mutations within the cluster (Thayer *et al*., 1995) and structural analysis of the FCL motif bound and unbound to DNA (Fromme & Verdine, 2003) revealed that the properties of the cluster had a strong association with the DNA backbone and could potentially stabilise the structure of the protein during catalysis.

Recently, in addition to the structural function of the FCL motif; electrochemical analysis of this cluster showed that it is redox active when bound to DNA (Boal *et al*., 2009) and participates in DNA mediated charge transport (Núñez *et al*., 1999). When DNA is damaged it causes disturbances to the stacking of the DNA and an electron or charge is

then associated with the damaged DNA (Núñez *et al*., 1999). These charges are transferred within the DNA and can travel for long distances to signal FCL motifs containing proteins such as Nth (Núñez *et al*., 1999). The FCL motif of the Nth protein becomes active and changes to a state of oxidation (Boon *et al*., 2003; Boal *et al*., 2009). In this oxidized state it is lead toward the site of damage where repair takes places (Núñez *et al*., 1999). After catalysis it becomes inactive returning to its reduced state and can activate another FCL motif containing enzyme such as MutY (Boon *et al*., 2003; Boal *et al*., 2009). Therefore, the FCL motifs of Nth and MutY locate damaged DNA lesions and distinguishes them from correct or undamaged bases using DNA mediated charge transport thus making BER an efficient DNA repair pathway (Núñez *et al*., 1999; Boon *et al*., 2003; Boal *et al*., 2009).

Nth recognises a broad class of saturated, fragmented or contracted rings as a result of oxidized thymine and cytosine resides such as: 5,6-dihydrothymine (DHT), 6-hydroxy-5,6-dihydrothymine, thymine glycol, 5-hydroxy-5- methylhydantoin, methyltartonyl urea, and urea (Armel *et al*., 1977; Gates & Linn, 1977; Demple & Linn, 1980; Katcher & Wallace, 1983; Breimer & Lindahl, 1984; Cunningham & Weiss, 1985; Breimer & Lindahl, 1985; Dizdaroglu *et al*., 1993; Hatahet *et al*., 1994; Dizdaroglu *et al*., 2000). To a lesser extent Nth can also excise oxidized purines such as 8-OxoG oligonucleotides (Matsumoto *et al*., 2001) incorporated opposite G or A. The Nth glycosylase has a remarkable ability to repair a plethora of damaged pyrimidine bases as a consequence of exposure to oxidizing agents and X- and UV-irradiation.

This is evident in Ec where a single *nth* deletion does not have an effect in the survival of the organism under oxidative stress as Ec like Msm are equipped with other DNA glycosylases to compensate for the loss of *nth* (Cunningham &Weiss, 1985; Jiang *et al*.,

1997; Saito *et al*., 1997). Similarly, Msm mutants deficient in individual and/or both the *fpg/nei* genes, displayed no variation in growth kinetics under normal culture conditions, no changes in survival under oxidative stress conditions with no increase in spontaneous mutation rates, when compared to the parental strain, suggesting that these Fpg, Nei and Nth glycosylases must work in concert to provide effective repair of damaged DNA in a genotoxic environment (Goosens, 2008 MSc). Supporting this, mutator phenotypes only became apparent in Ec (Blaisdell *et al*., 1999) and *Salmonella typhimurium* (*S.typhimurium*) (Suvarnapunya *et al*., 2003; Suvarnapunya & Stein, 2005) when the *nth* encoding DNA glycosylase was inactivated in addition to the *fpg* and *nei* genes, thereby implying that the *nth* DNA glycosylase may have a vital role during DNA repair. Hence, in this study the *nth* gene was inactivated in the parental Msm strain and Msm mutants deficient in the entire Fpg/Nei family of glycosylases (Goosens, 2008 MSc) to understand the role of the Nth DNA glycosylase in DNA repair and mutagenesis in Msm.

## 1.8 Gene inactivation by homologous recombination

Homologous recombination is fundamental to eukaryotes, prokaryotes, archea and viruses as it is implicated in biological processes such as DNA repair, genetic variation and horizontal gene transfer (Kowalczykowski *et al*., 1994). It is a process that takes advantage of two homologous strands of DNA produced during replication, whereby if one strand is damaged the homologous strand can crossover and serves as a template for strand re-synthesis and correction (Kowalczykowski *et al*., 1994). This replacement or exchange of genetic information of two similar or identical DNA sequences is known as homologous recombination (Kowalczykowski *et al*., 1994).

Molecular biology has used this biological mechanism as a tool for studying gene function by gene inactivation. Particularly, in mycobacteria gene inactivation or knockout is a process that introduces site specific gene deletion in the desired organism via a suicide vector with the disrupted gene which retains small amounts of the 5' and 3' ends of the targeted gene while majority of the targeted gene is deleted (Gordhan & Parish, 2001). Upon delivery of the inactivated gene into the target organism via the suicide vector, a single crossover (SCO) event occurs incorporating the suicide vector into the chromosome (Gordhan & Parish, 2001). For the suicide vector to be maintained in the organism, it integrates into the chromosome either before or after the targeted gene (Gordhan & Parish, 2001). During the second crossover event referred to as a double crossover (DCO) the chromosomal target gene is replaced with the disrupted gene generating a mutant. Mutants are thus generated by a two step selection process (Gordhan & Parish, 2001). The suicide vector is engineered with selectable and counter-selectable antibiotic markers that aid in the process of selection of mutants as explained further in section 3 (Fig. 3.21, Gordhan & Parish, 2001).

In this study, a panel of mutant strains deficient in the *nth* gene were generated by homologous recombination which were physiologically assessed to determine the role of the Nth DNA glycosylase in Msm.

## 1.9 Aim and objectives of this study:

Previous work in our laboratory has shown that Msm mutants lacking Fpg/Nei glycosylases showed no reduced survival in response to hydrogen peroxide ($H_2O_2$) and no mutator phenotype (Goosens, 2008 MSc). However, mutator phenotypes were observed when all three glycosylases Fpg/Nei/Nth were deleted in *Escherichia coli* (Ec) and *S.*

*typhimurium* (Blaisdell *et al*., 1999; Suvarnapunya *et al*., 2003; Suvarnapunya & Stein, 2005), suggesting that the *nth* gene encoding the Nth glycosylase is vital for cellular survival and is involved in combating oxidative stress and its DNA damaging effects. Therefore, the aim of this study was to ascertain the individual role of the Nth DNA glycosylase and its relationship with the Fpg/Nei family of DNA glycosylases in response to oxidative stress and induced mutagenesis in Msm. In order to achieve the aim the objectives were the following:

**1.9.1 Inactivation of the *nth* gene in the parental (Msm) strain and mutant strain devoid of the Fpg/Nei family of DNA glycosylases by the:**

1.9.1.1 construction of suicide or knockout vector for the *nth* gene;

1.9.1.2 generation of Msm mutants deficient in *nth* in the presence and absence of *fpg* and *nei* by homologous recombination.

**1.9.2 Complementation of the various *nth* deleted mutants with the functional copy of the *nth* genes by:**

1.9.2.1 construction of a complementation or integration vector containing the *nth* gene;

1.9.2.2 generation of *nth* complemented strains and verifying expression of *nth* complements.

**1.9.3 Phenotypic characterization of combinatorial Fpg/Nei/Nth deficient mutants with their respective complemented strains:**

1.9.3.1 for growth and survival under normal and oxidative conditions as generated by $H_2O_2$;

1.9.3.2 for changes in spontaneous mutation rates to rifampicin as measured by the fluctuation assay and analysis of the mutational spectra;

1.9.3.4 for increased changes in mutation frequencies induced by UV irradiation DNA damage.

# 2. Materials and methods

## 2.1 Bioinformatics analyses

The SmegmaList database (http://mycobrowser.epfl.ch/smegmalist.html) and the Comprehensive Microbial Resource (CMR) database of the J. Craig Venter Institute (JCVI; http://cmr.jcvi.org/tigr-scripts/CMR/CmrHomePage.cgi) were used to retrieve the upstream and downstream genomic sequences and the protein sequence of the *nth* gene in *Mycobacterium smegmatis* (Msm). The protein database of Clusters of Orthologous Groups (COG database http://www.ncbi.nlm.nih.gov/COG) was used to identify the Nth glycosylase of several related microorganisms: *Escherichia coli* (Ec), *Helicobacter pylori* (*H. pylori*), *Salmonella typhimurium* (*S. typhimurium*), *Corynebacterium efficiens* (*C. efficiens*), *Corynebacterium jeikeium* (*C. jeikeium*), *Bacillus subtilis* (*B. subtilis*), *Mycobacterium leprae* (*M. leprae*), *Mycobacterium avium* (*M. avium*), *Mycobacterium avium paratuberculosis* (*M. avium paratuberculosis*), *Mycobacterium bovis* (*M. bovis*), *Mycobacterium tuberculosis* (Mtb) and Msm. The CMR database was used to retrieve the protein sequences for all of these microorganisms. A multiple sequence alignment was performed using Clustlal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo/) to establish whether the structural integrity of the Nth glycosylase was conserved amongst these microorganisms. The structural properties of the Nth protein were obtained from Pfam (http://pfam.sanger.ac.uk/) and Prosite (http://prosite.expasy.org/) databases. The crystal structure of the Nth protein from Ec was obtained from the protein databank (PDB) accession number 2ABK. The structure of the protein was formatted in ribbons using an interactive website Jmol:

http://www.rcsb.org/pdb/explore/jmol.do?structureId=2ABK&bionumber=1. In addition, COG database and BLASTp

(http://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins&PROGRAM=blastp&BLAST_PROGRAMS=blastp&PAGE_TYPE=BlastSearch&SHOW_DEFAULTS=on&BLAST_SPEC=blast2seq) were used to assess the similarity represented as E values between the Msm Nth glycosylase and the various other microorganisms.

CMR was used to compare the genomic context of the *nth* gene of Msm, Mtb and Ec. The percent identity of the Msm Nth protein sequences compared to protein sequences of Mtb and Ec were assessed by BLASTp. The algorithm parameters were default settings and the scoring parameters were matrix: BLOSUM62, gap costs existence: 11 and extension: 1 with a conditional compositional score matrix adjustment.

## 2.2 Bacterial strains, plasmids and maintenance of strains

### 2.2.1 Bacterial strains and culture conditions

Bacteria in this study were cultured using media described in Appendix 5.1.1. All Ec strains were cultured overnight at 37 °C with shaking (Labcon Shaking Incubator) in Luria-Bertani broth (LB) supplemented with the appropriate antibiotics. Liquid cultures of Ec strains were spread on Luria-Bertani agar plates (LA) containing the appropriate antibiotics and incubated at 37 °C (Incotherm Labotec Incubator). Ec containing plasmids ≥ 8000 bp were grown at 30 °C with shaking (New Brunswick Scientific Innova 400 incubator shaker) for 48 hrs to prevent plasmid rearrangements. For the selection and counter selection of transformants the following antibiotics: ampicillin (amp) 100 µg/ml, hygromycin (hyg) 200 µg/ml, kanamycin (kan) 50 µg/ml or rifampicin (rif) 100 - 200 µg/ml and supplements: 5% w/v sucrose and 40 µg/ml 5-bromo-4-chloro-3-indolyl- β-galactoside (X-gal) were added where necessary.

Msm strains were cultured in Middlebrook 7H10 agar or 7H9 broth at 37 °C shaking. Where appropriate hyg 50 μg/ml, kan 25 μg/ml, rif 200 μg/ml, 2% w/v sucrose and/or 40 μg/ml X-gal was added to the media.

All bacterial strains and plasmids used in this study are listed in Tables 2.1 and 2.2 and were stored at -70 °C and -20 °C respectively. Ec strains were stored in 33% glycerol (v/v), while Msm strains were stored in the spent culture medium.

**Table 2.1: Bacterial strains used in this study.**

| Strains | Characteristics | Origin |
|---|---|---|
| ***Escherichia coli* (Ec)** | | |
| DH5α | *supE44 ΔlacU169 (80 1acZΔM15 hsdR17 recA1endA1 gyrA96 thi-1 relA1* | Promega, Madison, WI |
| AB1157 (Ec Wild type) | *thr-1, araC14, leuB6*(Am), *Δ(gpt-proA) 62, lacY1, tsx-33, qsr'-0, glnV44*(AS), *galK2*(Oc), $\lambda^-$, *Rac-0, hisG4*(Oc), *rfbC1,mgl51,rpoS396*(Am), *rpsL31*(strR) *kdgK51, xylA5, mtl-1, argE3*(Oc), *thiE1* | Dewitt & Adelberg, 1962. |
| BW415 (EcΔ*nth*) | Derivative of AB1157 carrying an in-frame unmarked deletion in *nth*, Δ(*manA-nth*) *84 ksgB1* | Cunningham & Weiss, 1985. |
| BW531 (Ec*nth*) | Derivative of MM294 carrying the *nth* gene from *Escherichia coli* in pBR322, Δ(*srl-RecA*) 306, *endA1, hsdR17, glnV44*(AS), *thi-1, spoT1, rfbD1*, $\lambda^-$, *creC510*, $amp^R$ | Chan &Weiss, 1987. |
| EcΔ*nth*::Msm*nth* | Derivative of BW415 (EcΔ*nth*) carrying the *nth* gene from *Mycobacterium smegmatis* with the native upstream promoter in pTWEETY, $kan^R$ | This study |
| ***Mycobacterium smegmatis* (Msm)** | | |
| $mc^2155$ (Msm | *ept*-1, highly efficient plasmid transformation mutant of | Snapper *et al.*, |

| | | |
|---|---|---|
| wild type) | mc$^2$6 | 1990. |
| Δ*f1f2* | Derivative of mc$^2$155 with deletions in MSMEG_2419 and MSMEG_5545 | Goosens, 2008 (MSc). |
| Δ*n1n2* | Derivative of mc$^2$155 with deletions in MSMEG_1756 and MSMEG_4683 | Goosens, 2008 (MSc). |
| Δ*f1f2n1n2* | Derivative of mc$^2$155 with deletions in MSMEG_2419, MSMEG_5545, MSMEG_1756 and MSMEG_4683 | Goosens *et al*., (unpublished). |
| Δ*n1n2f1f2* | Derivative of mc$^2$155 with deletions in MSMEG_1756, MSMEG_4683, MSMEG_2419 and MSMEG_5545 | Goosens *et al*., (unpublished). |
| **Msm *nth* deficient mutants** | | |
| Δ*nth* | Derivative of mc$^2$155 with a deletion in MSMEG_6187 | This study |
| Δ*f1f2nth* | Derivative of mc$^2$155 with deletions in MSMEG_2419, MSMEG_5545 and MSMEG_6187 | This study |
| Δ*n1n2nth* | Derivative of mc$^2$155 with deletions in MSMEG_1756, MSMEG_4683 and MSMEG_6187 | This study |
| Δ*f1f2n1n2nth* | Derivative of mc$^2$155 with deletions in MSMEG_2419, MSMEG_5545, MSMEG_1756, MSMEG_4683 and MSMEG_6187 | This study |
| Δ*n1n2f1f2nth* | Derivative of mc$^2$155 with deletions in MSMEG_1756, MSMEG_4683, MSMEG_2419, MSMEG_5545 and MSMEG_6187 | This study |
| **Msm *nth* complemented strains** | | |
| Δ*nth*::*nth* | Derivative of mc$^2$155 with deletion in MSMEG_6187carrying pTWEETY::*nth* integrated at the *attP* phage attachment site | This study |
| Δ*f1f2nth*::*nth* | Derivative of mc$^2$155 with deletions in MSMEG_2419, MSMEG_5545 and MSMEG_6187 respectively carrying pTWEETY::*nth* integrated at the *attP* phage attachment site | This study |
| Δ*n1n2nth*::*nth* | Derivative of mc$^2$155 with deletions in MSMEG_1756, MSMEG_4683 and MSMEG_6187 respectively carrying pTWEETY::*nth* integrated at the *attP* phage attachment site | This study |

| | | |
|---|---|---|
| Δ*f1f2n1n2nth*::*nth* | Derivative of $mc^2155$ with deletions in MSMEG_2419, MSMEG_5545, MSMEG_1756, MSMEG_4683 and MSMEG_6187 respectively carrying pTWEETY::*nth* integrated at the *attP* phage attachment site | This study |
| Δ*n1n2f1f2nth*::*nth* | Derivative of $mc^2155$ with deletions in MSMEG_1756, MSMEG_4683, MSMEG_2419, MSMEG_5545 and MSMEG_6187 respectively carrying pTWEETY::*nth* integrated at the *attP* phage attachment site | This study |

### 2.2.2 Cloning vectors (or plasmids)

All bacterial plasmids used and generated in this study are listed in Table 2. 2. All corresponding maps are shown in Appendix 5. 3.

**Table 2.2: Bacterial plasmids used in this study.**

| Plasmids | Characteristics | Size (bp) | Origin |
|---|---|---|---|
| pGEM3Zf(+) | Ec cloning vector, $amp^R$, *lac*Z-alpha, *ori*E | 3199 | Promega |
| p2NIL | Ec cloning vector and mycobacterial suicide plasmid; $kan^R$, *ori*E | 4753 | Parish & Stoker, 2000. |
| pGOAL19 | Plasmid carrying *lac*Z, *hyg* and *sacB* genes as a PacI cassette; $amp^R$, *ori*E | 10435 | Parish & Stoker, 2000. |
| pOLYG | Mycobacterial and Ec shuttle vector; $hyg^R$, *ori*E | 5315 | Ó Gaora *et al*., 1997. |
| pTWEETY | Mycobacterial integrating vector; $kan^R$, *ori*E, *int* | 5835 | Pham *et al*., 2007. |
| pGEMΔ*nth* Us | Derivative of pGEM3Zf(+) cloning vector containing the deleted *nth* upstream region; $amp^R$, *lac*Z-alpha, *ori*E | 4396 | This study |

| | | | |
|---|---|---|---|
| pGEMΔ*nth* Ds | Derivative of pGEM3Zf(+) cloning vector containing the deleted *nth* downstream region; amp$^R$, *lac*Z-alpha, *ori*E | 4381 | This study |
| p2NILΔ*nth* | Derivative of p2NIL suicide vector containing the in-activated *nth* allele isolated from pGEMΔ*nth* Us and pGEMΔ*nth* Ds plasmids and ligated into p2NIL; kan$^R$, *ori*E | 6803 | This study |
| p2NILΔ*nth*:: pGOAL19 | Derivative of p2NILΔ*nth* suicide vector containing selectable and counter selectable markers as a PacI cassette; kan$^R$, hyg$^R$, *lac*Z, *sacB*, *ori*E | 14742 | This study |
| p2NILΔ*nth*:: 2XpGOAL19 | Derivative of p2NILΔ*nth* suicide vector containing two selectable and counter selectable markers (2X PacI cassette); kan$^R$, 2X hyg$^R$, 2X *lac*Z, 2X *sacB*, *ori*E | 22681 | This study |
| pTWEETY::*nth* | Mycobacterial integrating vector containing the intact Msm *nth* gene with its native promoter; kan$^R$, *ori*E, *int* | 6985 | This study |

### 2.2.3 Assessment of cell viability

Cell viability was assessed by optical density (OD) or by enumerating colony forming units (CFUs) and in some cases both approaches were used. ODs were measured using either a Shimadzu UV-1601 UV-visible spectrophotometer or a WPA Biowave C0800 cell density meter at 600 nm ($OD_{600nm}$) and cultures with $OD_{600nm}$ above 0.80 were diluted 1:10 with the appropriate broth and re-measured for accuracy. For measuring viable cell counts ten-fold serial dilutions were performed by mixing 100 µl of cell suspension ($10^0$) into 900 µl of broth creating a $10^{-1}$ dilution. 100 µl of $10^{-1}$ dilution was added to 900 µl making a $10^{-2}$ dilution, which continued until the entire dilution series $10^{-1}$ to $10^{-7}$ was created. Thereafter, 100 µl of the selected or entire dilution series ($10^{-1}$ to $10^{-7}$) was spread in

duplicate onto solid media and incubated at 37 °C for 3-4 days. Colonies that emerged after incubation were enumerated and viable cells per ml of culture were calculated as follows:

$$\text{CFU/ml} = \frac{\text{number of colonies on plate} \times \text{dilution factor}}{\text{volume plated}}$$

## 2.3 General Molecular Techniques

### 2.3.1 DNA isolation

#### 2.3.1.1 Genomic DNA extraction from Msm

##### 2.3.1.1.1 Large scale genomic DNA extraction

Genomic DNA from Msm was extracted according to Larsen *et al*., (2000) with minor modifications. Approximately two loop full of cells grown on 7H10 media were scraped with an inoculation loop, re-suspended in 500 µl of TE buffer and rigorously mixed for 1 min. The homogenous cell suspension was heat killed at 65 °C for 35-45 min and placed on ice for cooling. After which 50 µl of lysozyme (10 mg/ml) was added and incubated at 37 °C for 1 hr. Subsequently, 70 µl of 10% sodium dodecyl sulphate (SDS) and 6 µl proteinase K (10 mg/ml) was added and incubated further at 65 °C for 2 hrs. Next, 100 µl of 5 M sodium chloride and 80 µl CTAB (cetyltrimethylammonium bromide)/sodium chloride solution (10% CTAB in 0.7 M sodium chloride) was added, mixed by inversion of the tubes and incubated at 65 °C for 10 min. Following incubation, 750 µl of chloroform: isoamyl alcohol (24:1 v/v) was added and mixed. The aqueous and organic phases were separated by centrifugation at 13000 rpm (Beckman Coulter Microfuge® 16 centrifuge) for 5 min. The aqueous layer containing the DNA was collected into sterile Eppendorf tubes and precipitated with 450 µl isopropanol on ice for 30 min. The DNA was pelleted by centrifugation for 20 min at 13000 rpm, washed with 70% ethanol and dried in a vacuum

centrifuge (SpeedVac, Savant, Farmingdale NY, USA). The DNA pellet was re-suspended in 50-100 µl of sterile $dH_2O$ (distilled water), at 37 °C for 1 hr and the quality and quantity of DNA isolated was assessed as detailed in section 2.3.6.

#### 2.3.1.1.2 Small scale genomic DNA extraction

Single Msm colonies were re-suspended in 100 µl of sterile $dH_2O$ and boiled for 10 min at 65 °C. The heat killed suspension was treated with 50 µl of 100% chloroform and boiled for 10 min at 65 °C. Thereafter, the suspension was centrifuged at 13 000 rpm for 10 min. The aqueous phase containing the DNA was collected into a sterile Eppendorf tube and used for PCR amplification.

### 2.3.1.2 Plasmid DNA extraction from *Escherichia coli* (Ec)

#### 2.3.1.2.1 Large scale plasmid DNA extraction

100 ml cultures grown overnight were harvested by centrifugation at 4500 rpm (Beckman Coulter Allegra™ X-22R centrifuge) for 10 min at 4 °C. Purified DNA was extracted with the use of the NucleoBond kit (Macherey-Nagel) in accordance with manufacturer's instructions. The DNA was allowed to dissolve at 37 °C for 1 hr and quantified as described in section 2.3.6.

#### 2.3.1.2.2 Small scale plasmid DNA extraction

Plasmid DNA extraction from Ec was performed according to Sambrook *et al*., 1989 with some modifications. 1 ml of an overnight culture was centrifuged at 13 000 rpm for 5 min and the cells re-suspended in 100 µl solution I (0.5 M glucose, 0.5 M EDTA, tris-hydrochloric acid pH 8.0). Thereafter, 200 µl of solution II (10 M sodium hydroxide, 10% SDS) was added and the suspension mixed by gentle inversion before adding 150 µl of

solution III (5 M potassium acetate, glacial acetic acid). The mixture was incubated on ice for 5 min and centrifuged at 13 000 rpm for 5 min. The supernatant containing plasmid DNA was collected into a sterile Eppendorf tube and precipitated with 350 µl isopropanol. The DNA pellet was rinsed with 70% ethanol and vacuum dried before it was reconstituted in 9-18 µl sterile $dH_2O$ after which it was treated with 1-2 µl RNase A (10 mg/ml) for 1 hr at 37 °C. The quality and quantity of DNA isolated was assessed as detailed in section 2.3.6.

### 2.3.2 DNA manipulation

All DNA manipulations were performed according to standard protocols (Sambrook *et al*., 1989; Sambrook & Russell, 2001). Various enzymatic reactions were used to amplify, clone and analyze the DNA required for the construction of the final knockout vectors. Ec DH5α was the host for the cloning and propagation of final knockout vectors.

#### 2.3.2.1 Enzymatic modification of DNA

##### 2.3.2.1.1 Polymerase Chain Reaction (PCR) amplification

Primer3 (http://frodo.wi.mit.edu/) was used to identify oligonucleotides within the specified region and provided possible annealing temperatures for the optimization of amplification by PCR. Primers were obtained from Inqaba Biotech Ltd. and Whitehead Scientific Ltd. as detailed in the Appendix 5.2.1.

Faststart Taq DNA Polymerase, dNTPack kit (Roche Diagnostics; Mannheim, Germany) was used to optimize cycling conditions for amplification of genomic DNA with the primers listed in Appendix 5.2.1. Following optimization, the high fidelity Phusion Polymerase was used for the amplification of genomic DNA to generate blunt ended

amplicons for cloning and to minimize the introduction of errors during PCR amplification.

PCR reactions with Faststart Taq DNA Polymerase were carried out using the following cycling conditions; 5 min denaturation (94 °C), 30-35 cycles of 15 s of denaturation (94 °C), 30 s of annealing (63.5 °C) and 40-45 s elongation (72 °C), with a final elongation step for 5-7 min. Standard conditions for the Phusion polymerase were as follows; 1 min denaturation (98 °C), 30-35 cycles of denaturation (98 °C) for 10 s, 15 s of annealing (63.5 °C) and 15-30 s elongation (72 °C), with a final elongation cycle of 5-7 min.

The PCR reactions were set up in 20 µl total volume. Three control reactions were included in all PCR experiments to rule out non-specific amplification or DNA contamination, namely, a reaction without DNA referred to as a no template control (NTC) with both forward and reverse primers, a reaction lacking the forward primer and a reaction lacking the reverse primer.

#### 2.3.2.1.2 Restriction endonuclease digestion

Restriction endonucleases for the verification of clones by restriction mapping and cloning were obtained from Roche Biochemicals, Fermentas or New England Biolabs. Restriction enzyme digestions were performed as per the manufacturers' instructions using the specified buffers and when necessary bovine serum albumin (BSA) was added. The concentration of plasmid or genomic DNA to be digested was determined by using the Nanodrop spectrophotomer with a UV absorbance at 260 nm (NanoDrop Technologies see section 2.3.6). Generally about 0.5-1 µg plasmid DNA was digested in a total volume of 10-20 µl for 1 hr at 37 °C (unless otherwise stated), while 2-4 µg of genomic DNA was

digested overnight at 37 °C in a total volume of 20-40 µl. Double digests were performed in a single step with a compatible buffer.

#### 2.3.2.1.3 Phosphorylation

Blunt ended Phusion DNA polymerase generated amplicons lacking 5' phosphate groups which are required for ligation into the respective vector. Hence, the amplicons were phosphorylated using T4 polynucleotide kinase (Fermentas) according to the manufacturer's instructions prior to ligation. 1 µl of T4 polynucleotide kinase was added to 15-18 µl of the PCR product, incubated at 37 °C for 15 min and then heat denatured at 95 °C for 1 min. The phosphorylated PCR product was separated on an agarose gel (section 2.3.4) and purified by gel extraction (section 2.3.5).

#### 2.3.2.1.4 De-phosphorylation

Linearized plasmids with compatible restriction endonuclease ends can self ligate during ligation to yield a circular vector again. Antarctic Alkaline phosphatase (Fermentas) was used as per the manufacturer's instructions to prevent re-circularization of the linear vector. 1 µl phosphatase and its appropriate buffer was added to linearized vector DNA and incubated overnight at 37 °C. The phosphatase was heat inactivated at 65 °C for 5 min and the DNA was separated on agarose gels (section 2.3.4) and purified by gel extraction (section 2.3.5)

#### 2.3.2.1.5 Ligation

Digested purified vector and insert DNA was ligated with T4 DNA ligase (Roche Applied Science) according to the manufacturer's instructions. 1 µl ligase, 1.5 µl 10 mM ATP and 1.5 µl of the supplied ligation buffer was added to the vector and insert reaction in a total

volume of 15 µl. The reaction was incubated overnight at room temperature and then heat denatured at 65 °C for 5-10 min. 50 ng of vector was used for all ligation reactions. In order to optimize cloning events the vector to insert ratios of 1:1, 1:2 or 1:3 was calculated as follows:

$$\text{Concentration of insert DNA (ng)} = \frac{\text{the size of the insert (bp)} \times \text{50 ng of vector}}{\text{total size of the vector (bp)}}$$

The ligation reactions were transformed into competent Ec DH5α cells (section 2.3.8).

### 2.3.3 DNA precipitation

The ethanol salt precipitation method was used to concentrate genomic and plasmid DNA. Briefly, 0.1 times volume of 3 M sodium acetate (pH 5.2) and 2.5 volumes of ice cold 100% ethanol were added to the DNA solution. The DNA was allowed to precipitate at -20 °C for 20 min before harvesting the pellet at 13 000 rpm for 20 min. The pellet was washed with 70% ethanol, vacuum-dried and re-suspended in sterile $dH_2O$. The DNA was allowed to dissolve at 37 °C for 1 hr and the quality and quantity of DNA precipitated was estimated as detailed in section 2.3.6.

### 2.3.4 The separation of DNA fragments by agarose gel electrophoresis

Agarose gels (1-2%) depending on the molecular weight of the DNA fragments were prepared in 1x TAE (Tris- acetate-EDTA) buffer to which 0.5 µg/ml of ethidium bromide was added. For high molecular weight DNA fragments 1% agarose gels were prepared whilst 2% agarose gels were used for DNA fragments ≤1 kb. The DNA fragments were separated at 80-100 volts in a Mini-Sub Cell GT minigel horizontal submarine unit (BIO-RAD). Lambda DNA molecular weight markers (Roche Biochemicals) in Appendix 5.1.2 (Fig. 5.1.2.1) were used to assess fragment sizes of samples and to estimate the

concentration of DNA. Gels were visualized by a gel fluorescence (UV-light) imaging system; G:BOX and GeneSnap image acquisition software (Syngene).

### 2.3.5 DNA fragment extraction and purification from agarose gel

After gel electrophoresis, the correct DNA fragment was excised from the agarose gel and purified using the PCR clean-up Gel extraction kit (NucleoSpin Extract II, Macherey-Nagel) as per the manufacturer's instructions. Briefly, the excised gel slice was melted at 65 °C and the suspension loaded onto a column which bound the DNA. The column was then washed before the DNA was eluted with 20-40 µl of sterile $dH_2O$. The quality and quantity of DNA recovered was evaluated as detailed in section 2.3.6.

### 2.3.6 Quantification of DNA

DNA was quantified on the NanoDrop ND-1000 Spectrophotometer (NanoDrop Technologies) used in conjunction with software (Coleman Technologies) provided by the manufacturer and on agarose gels using lambda DNA molecular weight markers (Roche Biochemicals). The DNA bands of the lambda molecular weight markers Fig. 5.1.2.1 (Appendix 5.1.2) have known concentrations therefore the band intensities of the DNA samples can be compared to the molecular marker concentration to estimate the approximate concentrations of the sample DNA.

### 2.3.7 Southern blot analyses

The genotype of mutants at the deletions site was analyzed by Southern blot analysis. This technique involved the separation of digested DNA by agarose gel electrophoresis, the transfer of DNA onto a nitrocellulose membrane followed by hybridization with specific chemiluminescence labelled probes to the DNA on the membrane and visualization of complementary hybridized fragments on X-ray film.

Approximately 2-4 µg of genomic DNA from the relevant strains was digested overnight at 37 °C with the appropriate restriction endonuclease(s) and genomic fragments were separated by gel electrophoresis at 80 V in a 1% agarose gel. The gel was washed with depurination solution (0.25 M hydrochloric acid) for 15 min, rinsed with sterile $dH_2O$ twice and washed with denaturing solution (0.5 M sodium hydroxide, 1.5 M sodium chloride) for 30 min after which it was equilibrated in 1x TBE (Tris- borate-EDTA) buffer (Sigma-Aldrich) for 5 min. The agarose gel was then covered with the same size Hybond$^{TM}$ – N nitrocellulose membrane, sandwiched between two 3 mm Whatmann filter papers and two pre-soaked sponges and placed in a transfer cassette. The cassette was placed in the OmniPAGE Electroblotting Unit (Cleaver Scientific Ltd CS-300V) and the DNA transferred with 1x TBE buffer onto the nitrocellulose membrane at 0.6 A, 130 V for 2 hrs at 4 °C. After the transfer the DNA was crosslinked onto the nylon membrane at 2500 mJ/cm$^2$ (UV Stratalinker 1800, Stratagene).

The DNA High Prime DNA labelling and Detection Starter Kit II (Roche Biochemicals) was used for hybridization, labelling and detection as per the manufacturer's instructions. Pre-hybridization of the membrane was performed using the DNA High Prime DNA labelling and Detection Starter Kit II. Pre-hybridization was performed in the Hybaid HB-OV-BM roller bottles at 52.5-55.0 °C for 30 min in the hybridization oven (Hybaid Micro-4) by incubating the membrane in 10 ml of DIG Easy Hyb solution (0.5% SDS, 6x saline-sodium citrate [SSC], 5x Denhardt's solution, 50% deionized formamide). Following incubation, the DIG Easy Hyb solution was decanted into a sterile 50 ml Falcon tube and stored at room temperature for future use. The PCR DIG Probe Synthesis Kit replaces dTTP with digoxigenin-labelled dUTP (DIG- dUTP) in a PCR amplification reaction. Upstream and downstream DIG-labelled DNA probes were generated using the PCR DIG Probe Synthesis Kit, upstream and downstream primers in Table 5.2.1.1 (Appendix 5.2.1)

and the p2NILΔ*nth*::pGOAL19 knockout construct (Table 2.2) as template according to the manufacturers' manual. To confirm DIG-dUTP incorporation both the DIG labelled and unlabeled upstream and downstream PCR probes were analyzed on a 1% agarose gel. The DIG-dUTP labelled probes had a higher molecular weight compared to the unlabeled upstream and downstream PCR amplicons.

The hybridizing temperature for each of the probes was calculated according to the formula in the manufacturers' handbook. The probes were heat-denatured at 95 °C for 2-5 min, immediately submerged in ice, then added to the pre-hybridized membrane in a new 10 ml aliquot of DIG Easy Hyb solution and incubated overnight at 52.5-55.0 °C. After hybridization the probes were decanted into sterile 50 ml falcon tubes and stored at -20 °C for further use. The membrane was washed twice at room temperature for 5 min with solution I (2x SSC, 0.1% SDS) and twice at 68 °C for 15 min in solution II (0.5x SSC, 0.1% SDS). The membrane was removed from the tubes into a deep container for equilibration with 50 ml wash buffer (0.01 M maleic acid, 0.015 M sodium chloride , 0.3% triton X-100, pH 7.5) for 5 min after which 125 ml of 1x blocking solution (0.01 M maleic acid, 0.015 M sodium chloride, 10 × blocking solution supplied, pH 7.5) was added and the membrane incubated for 30-40 min, followed by a further incubation of 30-40 min with 25 ml of 1x blocking solution mixed with 1 µl of Anti-DIG antibody per blot. The membrane was then, washed twice for 15 min with 100 ml of wash buffer after which it was immersed in 20 ml detection buffer (50 mM magnesium chloride, 0.1 M Tris· hydrochloric acid, 1 M sodium chloride, pH 9.0). The membrane was placed into a hybridization bag (Roche Biochemicals) with 1 ml CSPD (Disodium 2-chloro-5-(4-methoxyspiro (2-dioxetane-3,2 (2-dioxetane-3,2'-(5'-chloro)-tricyclo[3.3.1.1. 3, 7.]decan)-. 4-yl)-1-phenyl phosphate, supplied), incubated at 37 °C for 5-10 min and exposed to X-

ray film (AGFA CP-G Plus Medical X-ray film). The antarctic phosphatase attached to DIG-dUTP labelled probes de-phosphorylates the CSPD and hence chemiluminescence is emitted and detected at a wavelength of 477 nm which can be captured on X-ray film.

### 2.3.8 Transformation of *E. coli* (Ec) DH5α cells

#### 2.3.8.1 Preparation of Ec DH5α competent cells using rubidium chloride

100 ml of Ec DH5α cells were grown to mid-log phase an $OD_{600nm}$ to approximately 0.48 - 0.52 in LB at 37 °C with shaking. The culture was cooled on ice for 15 min before harvesting the cells by centrifugation at 3500 rpm for 5 min. The bacterial pellet was re-suspended in 0.4 volume TfbI (15% v/v glycerol, 30 mM potassium acetate, 100 mM rubidium chloride, 10 mM calcium chloride, 50 mM manganese chloride) and chilled on ice for 15 min. The cells were pelleted again and the pellet was re-suspended in 0.04 volume TfbII (10 mM 3-(N-morpholino) propanesulfonic acid [MOPS], 75 mM calcium chloride) and incubated on ice for 15 min. The cells were used immediately or frozen in 500 µl aliquots at -70 °C for future use.

#### 2.3.8.2 Transformation of chemically competent Ec DH5α cells

Chemically competent Ec DH5α cells were thawed on ice and 100 µl of cells were added to each ligation reaction. After incubation on ice for 15 min the mixture was heat shocked at 42 °C for 90 s to allow for DNA uptake and placed on ice for a further 10-15 min. 1 ml 2XTY was added for cellular replication and recovery of the cells at 37 °C for 1 hr. Thereafter, the cells were spread on LA plates containing appropriate antibiotics and incubated overnight to allow for colony growth. A no DNA control was included to verify that the competent cells were not contaminated and a positive control transformed with 1 µg of a replicating plasmid were included for every transformation to test the

transformation efficiency of the competent cells which was determined to be between 1 X $10^7$ – 1 X $10^8$ transformants/µg DNA.

### 2.3.9 Electroporation into Msm and Ec

#### 2.3.9.1 Preparation of electro-competent Msm and Ec cells

100 ml of Msm or Ec culture was grown to $OD_{600nm}$ 0.40 - 0.70 in 7H9 or LB at 37 °C with shaking, incubated on ice for 30-90 min and centrifuged at 4000 rpm for 10 min at 4 °C. The pellet was re-suspended in 40 ml chilled 10% glycerol and centrifuged as above. Thereafter, the pellet was re-suspended in 20 ml chilled 10% glycerol, re-pelleted and finally re-suspended in 2 ml chilled 10% glycerol. The cell suspension was kept on ice and used immediately.

#### 2.3.9.2 Electroporation

400 µl of chilled electro-competent Msm or Ec cells were placed into a 0.2cm electroporation cuvette (Gene pulser; Bio-Rad Laboratories) and mixed with 2-6 µg of plasmid DNA before electroporation in the shock pod of the Gene Pulser Xcell (Bio-Rad Laboratories) with the following settings for Msm: 2500 V, 25 µF and1000 Ω and for Ec: 1800 V, 25 µF and 200 Ω. Immediately after electroporation 800 µl 2XTY was added for recovery and replication and incubated overnight at 37 °C. Following incubation Msm or Ec cells were spread on 7H10 or LA plates containing appropriate antibiotics and incubated to allow for colony growth. A no DNA control was included to verify that the competent cells were not contaminated and a positive control which involved electroporation of 1 ng of replicating plasmid pOLYG in Fig. 5.3.1.4 (Appendix 5.3.1) was used to calculate the transformation efficiency. The transformation efficiency of the Msm

electro-competent cells was between 1 X $10^3$ – 1 X $10^5$ transformants/μg DNA while for Ec electro-competent cells it was between 1 X $10^5$ – 1 X $10^7$ transformants/μg DNA.

### 2.3.10 DNA sequencing

Bulk purified pGEM plasmids with cloned inserts were sent to Inqaba Biotec (Pretoria, RSA) for sequencing using the M13 universal primers. The integrating plasmid and RpoB PCR products were sent to the Central Analytical Facilities-DNA Sequencing Unit of Stellenbosch University for sequencing. Sequencing data was analyzed using the SeqMan module of the Lasergene suite of DNASTAR programs. Sequencing primers were designed using the bioinformatic programs Primer3. All sequencing primers used in this study are listed in Table 5.2.1.2 (Appendix 5.2.1).

## 2.4 The construction and identification of site specific deletion mutants in Msm

The knockout vectors and mutant strains of Msm were constructed as previously described (Gordhan & Parish, 2001). Maps of vectors used in this study are illustrated in the Appendix 5.3.

### 2.4.1 Construction of the *nth* knockout vector

For allelic exchange homologous upstream and downstream regions 1 kb in length including about 100 bp of the *nth* gene (illustrated in section 3.7) were generated by PCR amplification using Phusion polymerase (section 2.3.2.1.1) and primers listed in Table 5.2.1.1 (Appendix 5.2.1). Both blunt ended upstream and downstream amplicons were phosphorylated (section 2.3.2.1.3) and gel purified (sections 2.3.4 and 2.3.5) before cloning into the replicating plasmid pGEM3Zf(+) (Fig. 5.3.2). The pGEM3Zf(+) vector was digested with *Sma*I to produce a blunt linearized fragment which was

dephosphorylated (section 2.3.2.1.4) to avoid self-circularization. Gel purified Us and Ds fragments and the linearized pGEM3Zf(+) vector were ligated (section 2.3.2.1.5) and transformed separately into competent Ec DH5α cells (section 2.3.8). Transformants were selected on LA plates supplemented with 100 µg/ml ampicillin and X-gal. Potential clones were selected and screened by PCR to verify insertion of the upstream and downstream fragment. Two representative clones from each cloning (pGEMΔ*nth* Us and pGEMΔ*nth* Ds Table 2.2) were further verified by restriction endonucleases to ensure that no rearrangements had occurred during the cloning. In addition, both clones were sequenced to confirm that no mutations were introduced during PCR amplification. Mutations in homologous regions could have downstream polar effects that could have potential phenotypic consequences.

#### 2.4.1.1 Three way cloning

Once both clones were verified to contain no mutations and no rearrangements, the upstream and downstream inserts were excised from the pGEMΔ*nth* Us and pGEMΔ*nth* Ds clones with *Hind*III, *Bgl*II and *Kpn*I restriction enzymes, gel purified and both upstream and downstream fragments were ligated simultaneously into the suicide plasmid p2NIL (Fig. 5.3.3) to generate the *nth* deleted construct (illustrated in Fig. 3.11). Following transformation positive clones were confirmed by restriction endonuclease analysis and only one clone (p2NILΔ*nth*) was subjected to further analysis.

#### 2.4.1.2 PacI cassette cloning

The PacI cassette containing additional marker genes *sacB*, *hyg* and *lacZ*, from the pGOAL19 vector (Parish & Stoker, 2000) was cloned into the p2NILΔ*nth* vector at the *Pac*I site to facilitate the isolation of Msm mutants with an inactivated *nth* allele. The *lacZ*

selectable marker allows for identification of Msm single cross over mutants (SCOs) as blue colonies since the *lacZ* gene expresses β galactosidase that degrades X-gal producing a blue precipitate within transformed colonies. The *sacB* counter selectable marker encodes a sucrose degrading enzyme, levansucrase that generates a toxic by-product hence, only Msm transformants that have lost the vector carrying the marker genes will survive in the presence of sucrose (Gordhan & Parish, 2001). The PacI cassette was isolated from the pGOAL19 vector and ligated to *Pac*I digested p2NILΔ*nth* (Fig. 3.13). Blue transformants were screened with restriction endonucleases. In addition, log phase cultures of the confirmed knockout constructs were serially diluted from $10^0$-$10^{-7}$ (section 2.2.3) and 100 µl of each dilution was spread onto LA with sucrose (5%) and X-gal (40 µg/ml) and LA containing X-gal (40 µg/ml). The ratio of viable cells in the presence of sucrose to the viable cells without sucrose determined the sucrose sensitivity of the knockout constructs.

### 2.4.2 Identification of single crossover (SCO) and double crossover (DCO) mutants

The final knockout construct was electroporated into electro competent Msm cells. Positive clones were observed after seven to fourteen days as blue colonies resistant to kan and hyg. A few blue colonies representing single cross over (SCO) mutants were picked and the DNA isolated by small scale genomic isolation (section 2.3.1.1.2) for PCR amplification to confirm the genotype of the SCOs (illustrated in Fig. 3.22) mutants.

Once a SCO was confirmed it was grown in broth to log phase without antibiotics to prevent selection of the vector. To allow for the second crossover event to occur to generate a double cross over mutant that has lost the vector carrying the selectable and counter selectable markers (illustrated in Fig. 3.21), a log phase culture of the SCO was serially diluted ($10^0$-$10^{-7}$) and 100 µl of each dilution plated onto 7H10 with sucrose (2%)

and X-gal (40 µg/ml). There was a 10 000 fold reduction in viable cells in the presence of sucrose compared to the total number of viable cells on media lacking sucrose indicating that the *sacB* gene was functional. The white colonies that emerged on the 7H10 sucrose plates after 3-7 days of incubation were re-suspended in sterile $dH_2O$. 5 µl of the cell suspension was spotted onto 7H10 with sucrose and X-gal, 7H10 kan, hyg and X-gal and 7H10 X-gal, to distinguish between spontaneous *sacB* and *lacZ* mutants and possible DCOs. Clones with spontaneous mutations in the *sacB* and *lacZ* genes can be mistaken for DCOs as these mutants are sucrose resistant and white in colour, but grow in the presence of kan and hyg as they retain the vector sequence. Hence, only white colonies that grew on sucrose and X-gal plates but were susceptible to kan and hyg were examined by PCR to distinguish between wild type revertants and *nth* deletion mutants. During the second selection process the wild type allele can either be restored if the cross over event occurs on the same side as the first cross over event. A cross over event on the second side results in the complete loss of the vector to generate a mutant strain containing the inactivated allele (illustrated in Fig. 3.21). Potential DCOs identified by PCR were further confirmed by Southern blot analysis to ensure genotypic site-specificity of the homologous region. Details of the Nth (Nth F1 and Nth R1) primers and amplification conditions used for PCR confirmation are shown in Table 5.2.1.3 (Appendix 5.2.1).

## 2.5 Genetic complementation

### 2.5.1 Complementation vector

Primers in Table 5.2.1.1 (Appendix 5.2.1) were designed to PCR amplify (Phusion) a functional copy of the *nth* gene which incorporated approximately 350 bp upstream of the start codon to ensure that the promoter region was included. The pTWEETY Ec - *Mycobacterium* integrating shuttle vector (Table 2. 2) that has an *attP* phage site for

integration into the *attB* locus on the mycobacterial chromosome was linearized with *Ecl*136II to generate blunt ends (Appendix 5.3.1, Fig. 5.3.1.1). The blunt linearized vector was dephosphorylated (section 2.3.2.1.4) and gel purified (sections 2.3.4 and 2.3.5). The *nth* amplicon was cloned into the multiple cloning site (MCS) of the linearized Ec pTWEETY vector. Potential clones were screened by a restriction endonuclease and once a positive clone was identified it was screened thoroughly with multiple restriction endonucleases before sequencing with primers in Table 5.2.1.2 (Appendix 5.2.1). The complemented vector was isolated from Ec and 2 μg of plasmid was electroporated into both the Ec *nth* mutant (Ec*Δnth*) and Msm *nth* mutant strains (section 2.3.9). Integration of the complemented vector in the mutant strains was confirmed by PCR using the primers in Table 5.2.1.1 (Appendix 5.2.1). In addition, Msm integrants were analysed by semi-quantitative reverse transcriptase-polymerase chain reaction (SQ RT-PCR section 2.5.2) for the expression of the Nth DNA glycosylase.

### 2.5.2 SQ RT-PCR

To test the expression of *nth* samples saved during the growth kinetics experiment (section 2.6.1) were used to isolate the RNA to avoid variations in expression as the growth conditions and starting inoculums were the same for all the strains. 10 ml of culture grown for 15 hrs with an OD of 0.35 was used.

#### 2.5.2.1 RNA isolation

RNA was isolated using the NucleoSpin® RNAII kit with some modifications. Briefly 10 ml of culture was centrifuged at 4500 rpm for 15 min and the supernatant discarded. The pellet was re-suspended in 500 μl of TE buffer containing 1 mg/ml lysozyme and incubated at 37 ºC for 1 hr. The cell suspension was mixed with RA1 buffer and β-

mercaptoethanol (Ambion) and transferred to Lysing Matrix B tubes (Qbiogene). The cells were ribolysed three times for 45 s at speed 6 using the Savant Fastprep FP120 ribolyser, with 2 min intervals between pulses when the cells were cooled on ice. The supernatant was transferred to a column which binds the RNA and treated with DNase on the column to remove any residual DNA. The column was then washed with RA2 and RA3 buffers before the RNA was eluted with 40 µl of sterile $dH_2O$. The quantity of RNA isolated was evaluated as detailed in section 2.3.6. As a precautionary measure the 2 µg of RNA was digested with Turbo DNase (Ambion) at 37 ºC for 1 hr as per the manufacturer's instructions to remove any contaminating DNA. The reaction was stopped with 14 µl of resin (Ambion), centrifuged at 13000 rpm for 10 min and the supernatant transferred to a sterile PCR tube. The isolated RNA and the treated 2 µg of RNA was either stored at -70 ºC for future use or kept on ice for immediate use.

#### 2.5.2.2 RT or first strand cDNA synthesis

RT or first strand cDNA synthesis was performed using Invitrogen products and instructions. The 2 µg of treated RNA was added to 50 ng/µl random hexamers and 10 mM dNTP mix, incubated at 65 ºC for 5 min and then chilled on ice for 5 min. This mixture was then added to: 1 X RT buffer, 25 mM magnesium chloride and 0.1 M Dithiothreitol [DTT]. Next, 200 U/µl SuperScript™ III RT was added to half of the mixture (+ samples) for cDNA synthesis while, sterile $dH_2O$ was added to the other half of the reactions which served as a control to check for DNA contamination (- samples). Both + and – reactions were incubated in the thermocycler with the following cycling conditions: 25 ºC for 10 min, 50 ºC for 50 min followed by 85 ºC for 5 min. The samples were either stored at -20 ºC for future use or kept on ice for immediate use.

### 2.5.2.3 cDNA amplification

cDNA amplification was performed using 2 µl of + and – samples generated during RT (section 2.5.2.2) with Faststart Taq DNA Polymerase using primers listed in Table 5.2.1.4 (Appendix 5.2.1). A no DNA control referred to as no template control (NTC) was included in the PCR to check for DNA contamination. To create a genomic standard to serve as a positive control, 10 µl of 40 ng of genomic DNA ($10^7$) was added to 90 µl of sterile $dH_2O$ to create a $10^6$ standard. Then 10 µl of the $10^6$ dilution was added to 90 µl of sterile $dH_2O$ to make a $10^5$ standard and this process continued until $10^1$. 2 µl of each standard ($10^7$-$10^1$) was amplified with primers listed in Table 5.2.1.4 (Appendix 5.2.1). However, only the $10^4$ genomic standard which best represented the data was loaded onto a 2% agarose gel together with all PCR products and electrophoresed (section 2.3.4).

## 2.6 Phenotypic characterization of mutant strains

### 2.6.1 Growth kinetics

All Msm cultures were grown to log phase in 7H9 broth and used to inoculate fresh broth with a starting $OD_{600nm}$ of 0.02. Growth of these cultures was monitored every 3 hours (hrs) as OD measurements. 10 ml aliquots were sampled at an $OD_{600nm}$ of 0.35 at 15 hrs centrifuged at 4500 rpm for 10 min and were frozen at – 70 ºC immediately for SQ RT-PCR analysis (section 2.5.2). All experiments were repeated at least three times for statistical significance. Growth was represented graphically as absorbance ($_{600nm}$) vs. time (hrs).

### 2.6.2 Sensitivity to oxidative stress using hydrogen peroxide ($H_2O_2$) susceptibility assays

$H_2O_2$ (Calbiochem® 9.3 M) was stored in the dark at 4 °C. All cultures treated with $H_2O_2$ were covered in foil to prevent light exposure to minimise $H_2O_2$ degradation which could have an impact on cell viability resulting in variable susceptibility.

#### 2.6.2.1 Assessment of cell viability post $H_2O_2$ treatment

Previously Goosens (2008, MSc) reported that the bactericidal effect of $H_2O_2$ on Msm was growth phase dependent. Hence, the growth phase at which the bactericidal activity of $H_2O_2$ was most effective was determined in the parental Msm strain. Msm cultures in early ($OD_{600nm}$ 0.35), mid ($OD_{600nm}$ 0.50) and late ($OD_{600nm}$ 0.70) log phase were treated with 2.5 mM $H_2O_2$ (concentration previously optimized by Goosens, 2008 MSc). Cell viability was monitored every 2 hrs over a 6 hr period by plating 100 µl of serially diluted cultures ($10^0$-$10^{-7}$ section 2.2.3) onto 7H10 plates in duplicate, incubated at 37 °C and growth scored after 3 days. Once the growth phase was optimized in the parental Msm strain, all mutant strains were grown under these conditions. In some instances untreated cultures of the respective Msm strains were included as a negative control. In addition, 1 ml aliquots of Msm cultures were frozen at -70 ºC immediately to determine the intracellular $H_2O_2$ levels (section 2.6.2.2).

While, Ec cultures grown to log phase in LB from pre-cultures were treated with 10 mM $H_2O_2$ and incubated at 37 °C with shaking according to Cunningham & Weiss (1985). Cell survival was monitored every 10 min for a period of 30 min by plating 100 µl of serially diluted ($10^0$-$10^{-7}$ section 2.2.3) cultures onto LA plates in duplicate, incubated at 37 °C and colonies scored after 1 day. In all instances survival post $H_2O_2$ treatment was represented

graphically as survival (log CFU/ml) vs. exposure to $H_2O_2$ (min). One way ANOVA was performed to assess the statistical significance between Msm, Fpg/Nei mutants, *nth* mutants and the respective complemented strains using GraphPad Prism Software. These experiments were repeated at least three times for statistical significance.

#### 2.6.2.2 Measuring intracellular $H_2O_2$ levels

OxiSelect™ In Vitro ROS/RNS Assay Kit (STA 347 Cell Biolabs) was used to measure the concentration of intracellular ROS induced by 2.5 mM $H_2O_2$ as per the manufacturer's instructions. Briefly, frozen cells of Msm cultures (section 2.6.2.2) were thawed on ice, centrifuged at 13000 rpm for 10 min and the pellet re-suspended in 1 X PBS. The cell suspension was lysed using either the Transsomic T460 Elma® or Ultrasonics Gen-probe® ultra-sonicator water bath for 15 min. The lysed suspension was centrifuged at 13000 rpm for 10 min. Thereafter, 50 µl of cell supernatant was mixed with a fluorescent probe that binds $H_2O_2$ generated radicals and $H_2O_2$. The probe fluoresced at 480 nm excitation and the signal was captured at 530 nm which was visualized using the Microplate fluorescent reader (FLX 800 Bio-Tel instruments, INC) with KC4™ Data analysis software. A $H_2O_2$ standard curve and a no sample control were included to calculate the concentration of $H_2O_2$ for treated and untreated cultures.

### 2.6.3 Mutational assessments

The antibiotic used to determine resistance was rifampicin (rif). Rif inhibits bacterial transcription when it binds to the *rpoB* gene encoding the β subunit of a DNA dependent RNA polymerase. Four well characterized rifampicin resistance ($rif^R$) mutations are mapped to a region known as the rifampicin resistant determining region (RRDR) within the *rpoB* gene (Ramaswamy *et al*., 1998; Lee *et al*., 2005; McCammon *et al*., 2005; Sheng

*et al*., 2008). Spontaneous $rif^R$ mutants were randomly selected, their genomic DNA extracted and amplified across the RRDR region and the amplicons sequenced to identify the types of lesions that resulted in $rif^R$.

### 2.6.3.1 Spontaneous mutagenesis

#### 2.6.3.1.1 Assessment of mutation frequency

Mutation frequency measures all the mutants present within the entire population and was performed according to Li & Lu, (2001) with some modifications. Briefly, 50 ml Ec cultures were grown in LB to $OD_{600nm}$ 0.70 from pre-cultures. 100 µl of each culture was serially diluted ($10^0$-$10^{-7}$) and plated onto LA to determine the cell titre. Another 100 µl of each culture was spread onto LA rif (100 µg/ml) to determine the number of spontaneous $rif^R$ mutants. The frequency was calculated as the ratio of $rif^R$ mutants to the total number of cells.

#### 2.6.3.1.2 Assessment of mutation rates

Mutation rates are a measure of the probability of a cell gaining a mutation during its lifetime and the Luria-Delbrück fluctuation analysis is the most accurate method of measuring mutation rates as it 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 & Foster, 2000).

Measuring mutation rates (µ) in Msm was performed as previously described by Rosche and Foster (2000) & Machowski *et al*., (2007). Single colonies of the Msm strains were cultured in 7H9 broth to log phase (~$10^8$ cells/ml), of which a dilution series ($10^0$ to $10^{-7}$) was spread in duplicate on 7H10 for the initial inoculum ($N_0$) and 1 ml of each culture was spread on 7H10 rif (200 µg/ml) to account for any pre-existing mutants (*r*). To exclude

pre-existing mutants the cell density was reduced from ~$10^8$ cells/ml to ~$10^2$ cells/ml of which 2 ml of this inoculum was distributed into 30 culture tubes with continuous agitation using a magnetic stirrer (FMH Instruments Heat stirrer, Stuart heat stir SB162) to ensure homogeneity of the inoculum. The parallel cultures were incubated at 37 °C with shaking until log phase. Thereafter, a dilution series ($10^0$ to $10^{-7}$) from 5 tubes was spread in duplicate on 7H10 to determine the *Nt* value; while the entire culture (2 ml) from the 25 remaining tubes was plated onto 7H10 rif (200 µg/ml) to identify spontaneous rif mutations. The plates were incubated at 37 °C and scored after 3 days for total cell count (*Nt*) or after 7 days for $rif^R$ mutants.

Mutation rate is the probability that a cell will sustain a mutation during its lifetime and depends on the number of observed mutations per culture (*m*) and the total cell population (*Nt*). The initial mutation rate (µ) value can be calculated by dividing *m* by *Nt* (Equation 5). However, the *m* value depends on the distribution of mutations per culture and was estimated by the Luria-Delbrück distribution (Rosche & Foster, 2000). A combination of goal standard methods the Ma-Sandri-Sarkar Maximum Likelihood Estimator (MSS-MLE) and the $P_o$ method as described by the Luria-Delbrück distribution were used to estimate the *m* value. Since the expected number of mutations per culture was low (i.e. *m* value is between 0.3 - 2.3) and some equalled zero the $P_o$ method was used (Equations 1 and 2). The MSS-MLE method calculates the *m* value from the entire data set and not the median as described by the Lea-Coulson method (Equations 3 and 4). The extrapolated results from Equation 3 correspond to all possible *m* values, taking into account the initial *m* value. These *m* values are plotted on the x-axis and the extrapolated results from Equation 4 correspond to all possible y-axis values. These probable x and y values are then plotted to give a Luria- Delbrück distribution, an example of which is shown in Fig. 2.1. The peak

of this distribution is the most probable m-value which is used to determine the mutation rate (Equation 5).

| | Equation 3<br>*m* value estimates | Equation 4 |
|---|---|---|
| m1 | -1.0 | -1.0E+00 |
| m2 | -0.8 | -9.5E-03 |
| m3 | -0.6 | -2.0E-05 |
| m4 | -0.4 | -7.1E-09 |
| m5 | -0.2 | 3.0E-17 |
| m6 | 0.0 | 2.8E-16 |
| m7 | 0.2 | 3.0E-18 |
| m8 | 0.4 | 3.5E-15 |
| initial m | 0.6 | 4.6E-15 |
| m9 | 0.8 | 1.6E-15 |
| m10 | 1.0 | 2.8E-16 |
| m11 | 1.2 | 3.0E-17 |
| m12 | 1.4 | 2.2E-18 |
| m13 | 1.6 | 1.3E-19 |
| m14 | 1.8 | 6.1E-21 |
| m15 | 2.0 | 2.5E-22 |
| m16 | 2.2 | 8.7E-24 |

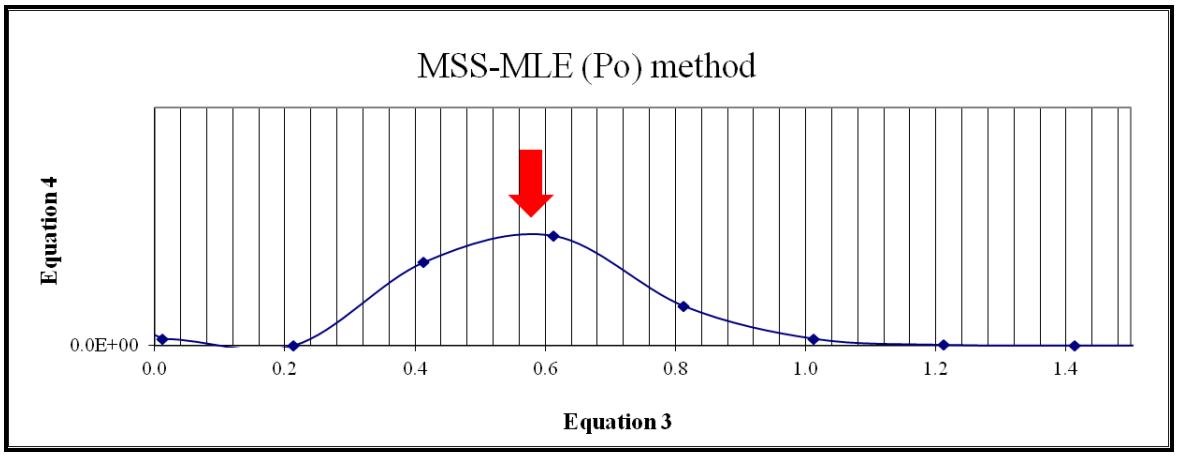



**Fig. 2.1: An example of the estimation of the *m* value by the Luria-Delbrück distribution.** The table lists the results of Equation 3 and Equation 4, which when plotted onto a graph allows for the identification of the most probable *m* value at the peak of the graph (indicated by the red arrow) used to calculate the mutation rate.

Mutation rates were assessed at least three times to ensure statistical accuracy. The mutation rates were calculated using Microsoft Office Excel spreadsheets developed at the CBTBR by Dr. E. Machowski (Machowski *et al*., 2007).

The calculations were as follows:

**Equation 1** - *Po* value estimation

$$Po = \frac{\text{number tubes with zero}}{\text{total number of tubes}}$$

**Equation 2** - *Po m* value equation

$$po = e^{-m}$$

$$m = -ln\ po$$

**Equation 3**

$$po = e^{-m};\ pr = \frac{m;}{r} \sum_{i=0}^{r-1} \frac{Pi}{r - I + 1}$$

**Equation 4**

$$f(r|m) = \text{II}\ f(r_i|m)$$

Where $f(r|m) = pr$ from equation 3

**Equation 5**

$$\mu = \frac{m \times ln2}{Nt}$$

### 2.6.3.2 Induced mutagenesis

#### 2.6.3.2.1 DNA damage-induced mutagenesis assay

UV induced mutagenesis assay was performed as previously described by Boshoff *et al*., (2003). All strains were grown in 50 ml 7H9 broth to early log phase to an $OD_{600nm}$ of 0.35. 100 µl of cultures were serially diluted ($10^0$-$10^{-7}$) and spread on 7H10 to determine cell viability. In addition 1 ml was spread on 7H10 rif (200 µg/ml) to determine spontaneous $rif^R$ mutants. The cultures were split into two equal cultures then centrifuged at 4500 rpm for 10 min at room temperature and the pellets re-suspended in 5 ml broth. Each of the cell suspensions were transferred to a sterile petri dish and one aliquot of cells was exposed to 250 $mJ/cm^2$ UV using the UV Stratalinker 1800 (Stratagene), whilst the other set served as a negative control. The treated and untreated cells were then transferred into separate flasks, containing 50 ml of fresh broth and incubated at 37 °C with shaking for recovery for 3 hrs. Thereafter, both the treated and untreated cultures were serially diluted ($10^0$ to $10^{-7)}$ and 100 µl of each dilution was spread in duplicate on 7H10 media to determine the CFU/ml. In addition 1 ml of culture for each of the treated and untreated strains was spread in duplicate on 7H10 rif (200 µg/ml) every 3 hrs for a period of 6 hrs to isolate spontaneous $rif^R$ mutants that resulted from UV induced mutagenesis. 7H10 plates were incubated at 37 °C and scored after 3 days while 7H10 rif plates were scored after 7 days. The mutation frequency was calculated as the number of $rif^R$ colonies (CFU/ml) divided by the total number of cells (CFU/ml) and a one way ANOVA was performed to assess the statistical significance between treated and untreated Msm strains using GraphPad Prism Software.

# 3. Results

## 3.1 Bioinformatics

### 3.1.1 Identification of the Nth glycosylase

Pfam and Prosite databases identified the structural features of the Nth protein: the Helix hairpin Helix (HhH) motif containing two active sites and the glycine-proline-aspartate (GPD) domain and the four iron four sulphur (FCL) cluster loop (illustrated in Fig. 3.1).

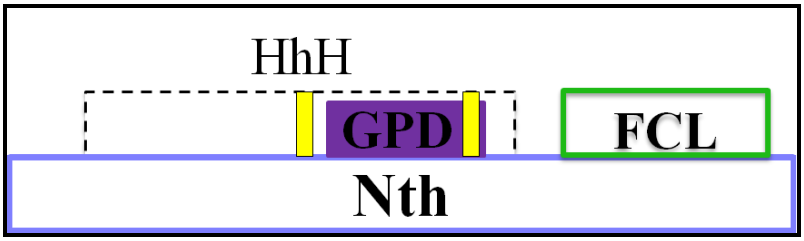



**Fig. 3.1: Structural characteristics of the Nth protein.** The HhH loop represented as a dashed black box contains the GPD domain in purple and two catalytic residues indicated as yellow boxes. The FCL is indicated by the green box (adapted from Pfam and Prosite databases).

The protein database of Clusters of Orthologous Groups (COG database) was used to identify the Nth glycosylase of several related microorganisms. The Nth protein is recognized in Actinomycetes, Enterobacteriaceae, Campylobacterales and Bacillales to possibly play a significant role in DNA repair. Interestingly, the following microorganisms: *Mycobacterium tuberculosis* (Mtb), *Mycobacterium bovis* (*M. bovis*), *Mycobacterium leprae* (*M. leprae*), *Mycobacterium avium paratuberculosis* (*M. avium paratuberculosis*), *Mycobacterium avium* (*M. avium*), *Corynebacterium jeikeium* (*C. jeikeium*), *Helicobacter pylori* (*H. pylori*), *Salmonella typhimurium* (*S. typhimurium*) and certain serotypes of *Escherichia coli* (Ec), belonging to the above mentioned orders are intracellular pathogens. COG database and BLASTp were used to compare the Msm Nth glycosylase to these microorganisms which showed significant similarity (low E-values) in Table 5.4.1 (Appendix 5.4). The structural integrity of the Nth protein of the above

mentioned microorganisms was analysed by Clustal Omega alignments (Fig. 3.2). CMR was used to retrieve protein sequences of all the microorganisms.

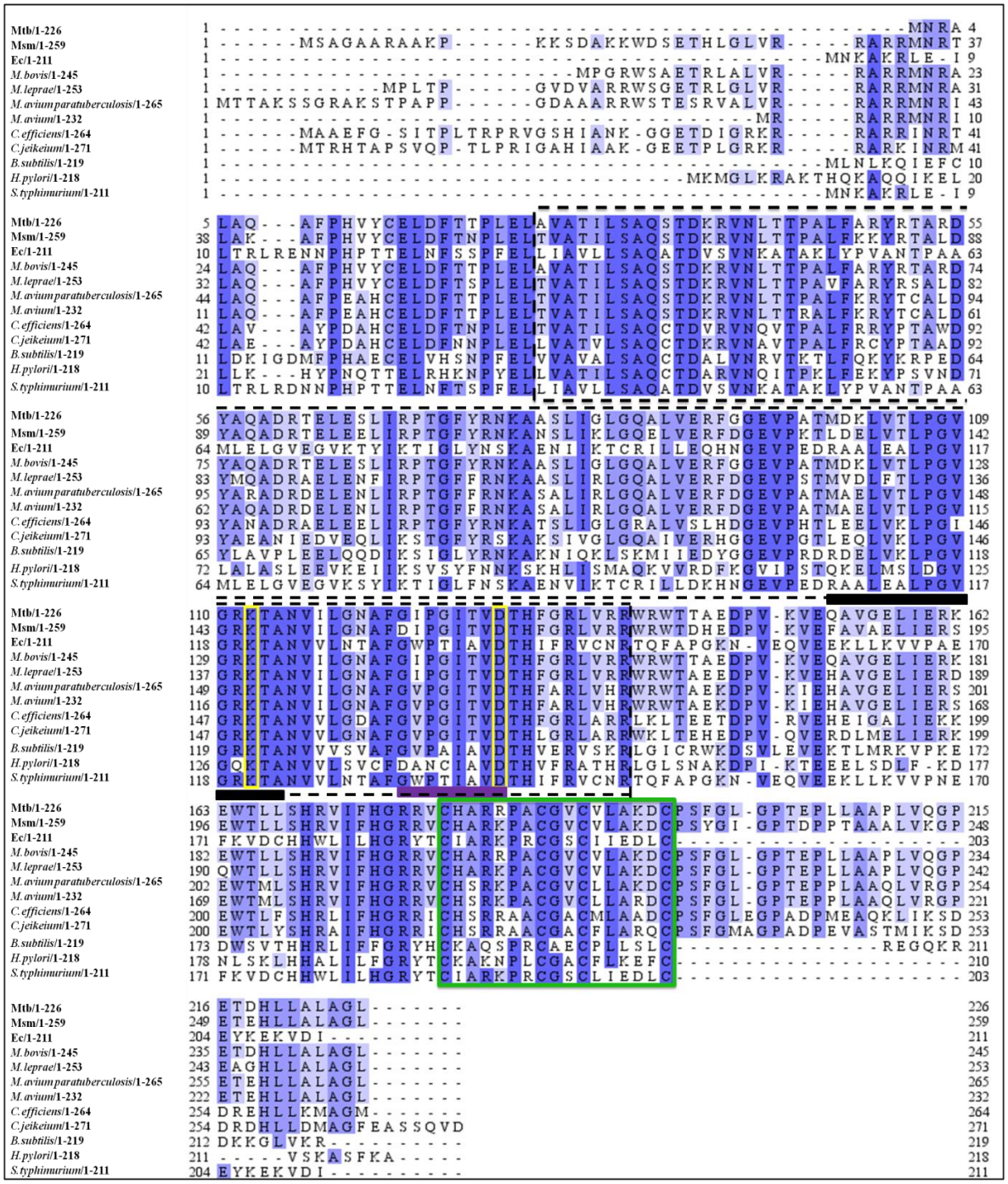



**Fig. 3.2: Alignments of the amino acid sequences of the Nth protein from various microorganisms.** Sequences were downloaded from CMR for Mtb, Msm, Ec, *M. bovis*, *M. leprae*, *M. avium paratuberculosis*, *M. avium*, *C. efficiens*, *C. jeikeium*, *B. subtilis*, *H. pylori and S. typhimurium*. The HhH motif is outlined in black dashed boxes and elements of the HhH motif is underlined in black. The GPD domain is underlined in purple, while active sites are highlighted in yellow boxes. The FCL motif is indicated in the green box.

The residues of the HhH motif were well conserved and the active site (lysine) at position 120 was highly conserved (Fig. 3.2 – underlined in black). In the GPD domain, at position 131 there was an aspartate/aspartic acid instead of a glycine residue in Msm and *H. pylori* (Fig. 3.2 – underlined in purple). The proline amino acid at position 133 was not conserved for *H. pylori* (Fig. 3.2 – underlined in purple). However, the catalytic residue at position 138 (aspartate) was highly conserved in all microorganisms (Fig. 3.2 – underlined in purple). Also, the FCL domain with the four cysteine residues was highly conserved in all microorganisms (Fig. 3.2 – green box). The structural domains of the Nth protein were well conserved in sequences of the microorganisms analysed indicating that these domains are essential for the functioning of the Nth glycosylase and imply that this DNA glycosylase is probably vital for DNA repair in intracellular pathogens.

The CMR database was used to assess the genomic context of the *nth* gene in Msm, Mtb and Ec (Fig. 3.3). The *nth* gene is co-transcribed as the terminal gene in Ec while in Msm and Mtb it is the first gene in an operon and interestingly in Msm the downstream genes in the operon encode for proteins that are related to DNA metabolism (Fig. 3.3).

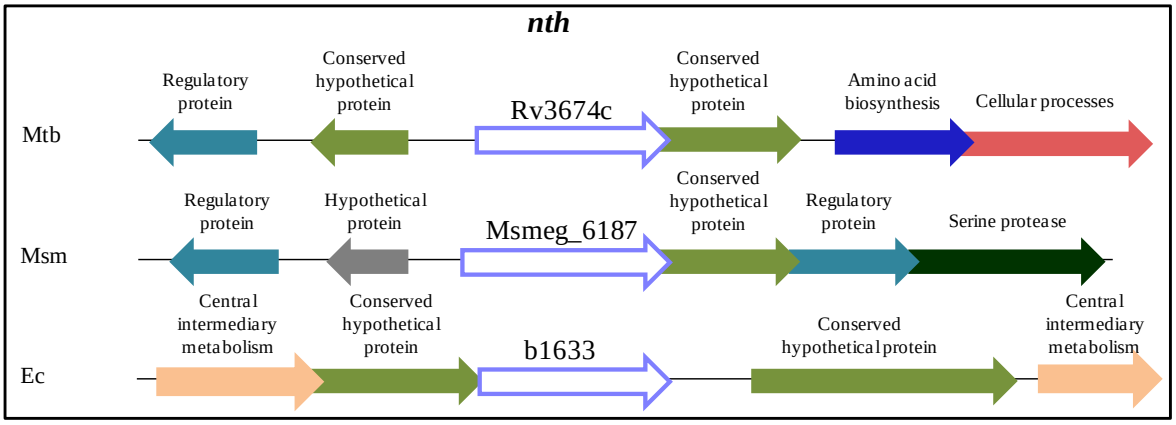


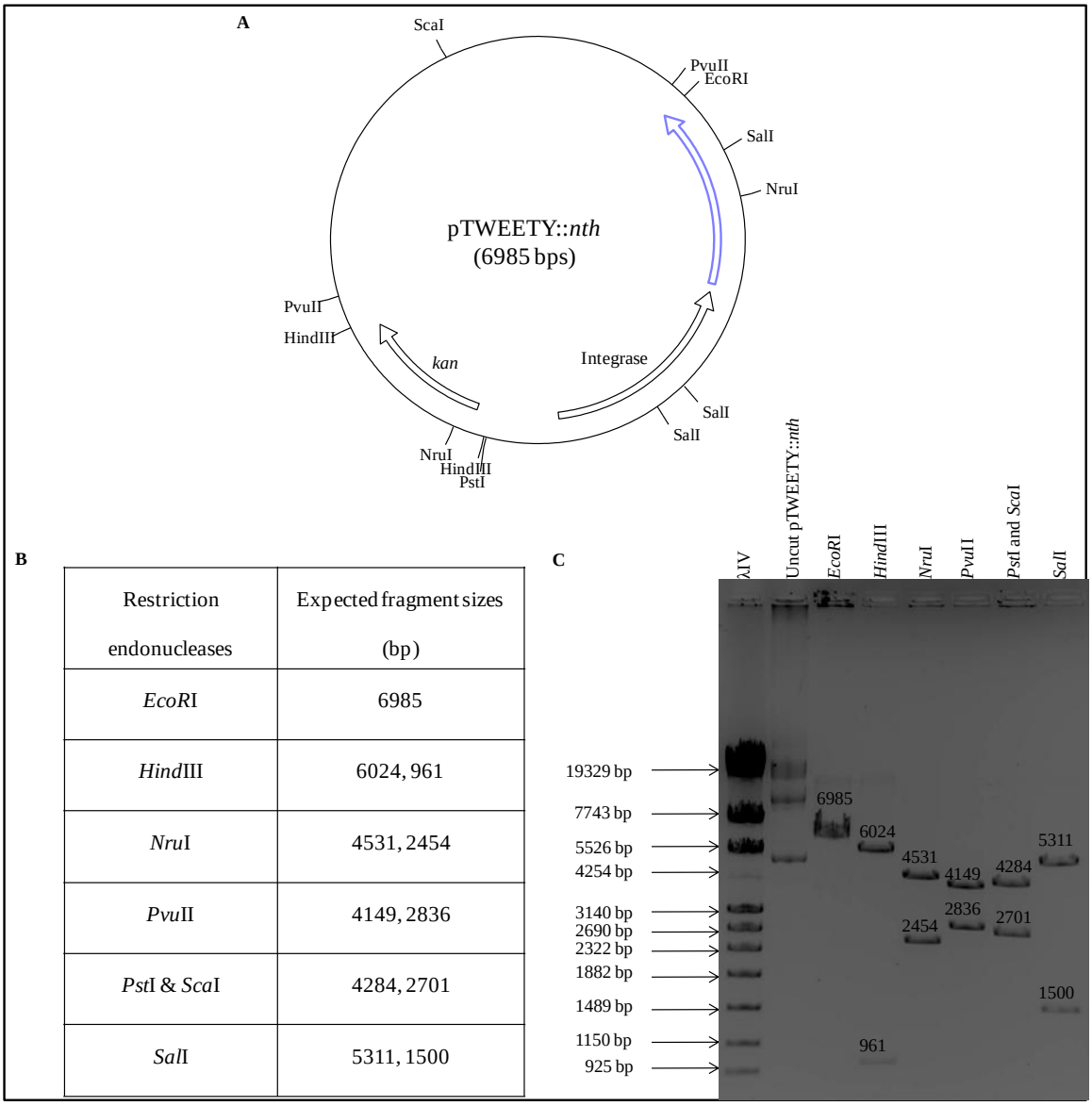

**Fig. 3.3: Genomic context of the *nth* DNA glycosylase in Msm, Mtb and Ec.** The *nth* gene is indicated in lilac while other homologous genes are denoted as block arrows in the same colour. The white open arrows are genes not in the same chromosomal context (adapted from CMR).

Since the *nth* gene was in the same genomic context for Msm, Mtb and Ec the percentage identity of the Nth protein between these microorganisms was predicted using BLASTp which was 39% ($1e^{-41}$) between Msm and Ec and 89% ($9e^{-149}$) between Msm and Mtb shown in Table 5.4.1 (Appendix 5.4).

Bioinformatic analyses showed that the Ec Nth glycosylase has significant similarity to the Msm Nth glycosylase with all the active residues for catalysis and domains necessary for DNA binding conserved in both organisms. To determine if the Msm Nth glycosylase functions similar to the Ec Nth glycosylase, the Msm *nth* allele was integrated into an Ec *nth* deficient mutant (EcΔ*nth*) to assess the functionality of the Msm *nth* gene.

## 3.2 Generation of the EcΔ*nth*::Msm*nth* strain

A functional Msm *nth* was amplified by PCR using primers listed in Table 5.2.1.1 (Appendix 5.2.1) and cloned into an integrating vector (pTWEETY). The resulting vector was verified by restriction mapping (Fig. 3.4) and sequenced to check for the absence of mutations that could have occurred during amplification (sequencing data not shown).




| Restriction endonucleases | Expected fragment sizes (bp) |
|---|---|
| *Eco*RI | 6985 |
| *Hind*III | 6024, 961 |
| *Nru*I | 4531, 2454 |
| *Pvu*II | 4149, 2836 |
| *Pst*I & *Sca*I | 4284, 2701 |
| *Sal*I | 5311, 1500 |

**Fig. 3.4: Restriction mapping of pTWEETY::*nth*. A.** pTWEETY::*nth* map showing the cloned intact Msm *nth* allele (labelled in lilac) cloned into pTWEETY. **B.** Table showing the expected sizes (bp) of digested pTWEETY::*nth* with multiple restriction endonucleases. **C.** Restriction analysis of pTWEETY::*nth* with various restriction enzymes.

No mutations were identified in the cloned amplicon hence, the complementation vector containing the Msm *nth* gene was transformed into competent EcΔ*nth* mutant cells. Potential kan$^R$ integrants were verified by PCR (Fig. 3.4) using the Msm NthCF and NthCR primers listed in Table 5.2.1.1 (Appendix 5.2.1).

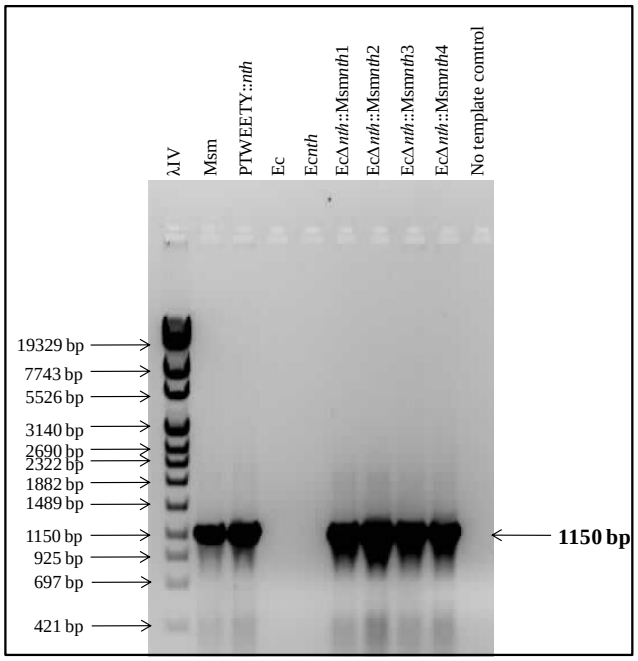



**Fig. 3.5: PCR verification of Msm *nth* complemented Ec *nth* mutant strains.** The Msm NthCF and NthCR primers are described in Table 5.2.1.1 (Appendix 5.2.1).

All four integrants (EcΔ*nth*::Msm*nth*1-4) produced the expected size amplicon (i.e. 1150 bp) as for Msm and the complementation vector (pTWEETY::*nth*), indicating that the Msm *nth* gene is present in the Ec *nth* mutant (Fig. 3.5). Since, the Msm primers were specific to Msm (in particular the Msm *nth* gene) these primers were not expected to produce amplicons for Ec (wild type) or EcΔ*nth*::Ec*nth* (Fig. 3.5). Only one of the EcΔ*nth*::Msm*nth* clones was used to ascertain the Msm Nth glycosylase functionality by monitoring the survival of the parental, the mutant and complemented strains in the $H_2O_2$ susceptibly assay (Cunningham & Weiss, 1985) and assessing the spontaneous mutation frequency of these various mutants to rifampicin (Li & Lu, 2001).

### 3.2.1 Assessment of cell viability post $H_2O_2$ treatment of Ec *nth* deficient mutant and its complements

To determine whether the Msm *nth* gene is involved in the cellular response to DNA damage the survival of the Msm *nth* complement (EcΔ*nth*::Msm*nth*) was compared to Ec

and Ec*nth* under oxidative stress (Fig. 3.6). Ec strains were grown to log phase and treated with 10 mM $H_2O_2$. Cell viability was monitored every 10 min for a period of 30 min.

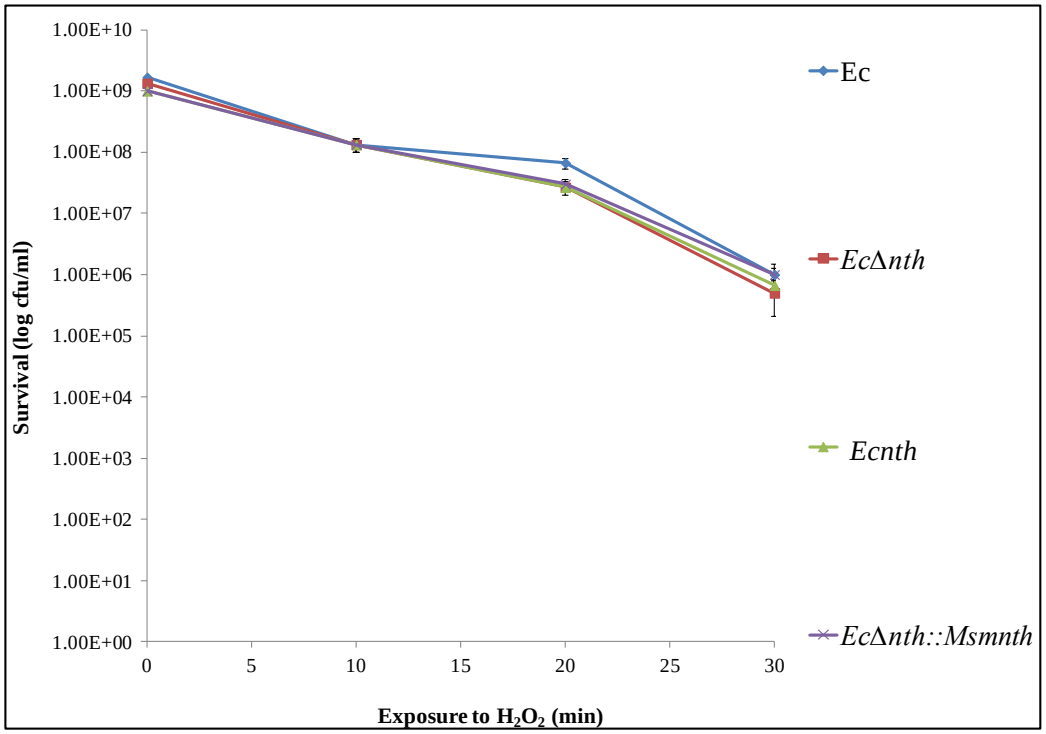



**Fig. 3.6: Assessment of complementation of Ec *nth* mutant by the Msm *nth* gene.** The mutant strains were treated with 10 mM $H_2O_2$ and survival monitored as CFU/ml over a 30 min period. Values are the mean of 3 independent biological replicate assays ± standard error (SE).

The differences in survival of the EcΔ*nth*::Msm*nth* mutant compared to wild type Ec, EcΔ*nth* and its Ec producing *nth* (Ec*nth*) were marginal under oxidative stress (Fig. 3.6). All Ec strains showed a similar response to oxidative stress indicating that this assay did not provide sufficient evidence to conclude that the Msm *nth* gene had Nth functional properties and was able rescue the Ec *nth* deficient phenotype. Hence, to further investigate whether the Msm and Ec *nth* genes had complementary function, the spontaneous mutation frequency to rifampicin of the various Ec strains was measured.

### 3.2.2 Mutation frequency

The spontaneous mutation frequency to rifampicin for the parental Ec, the Ec *nth* mutant (EcΔ*nth*), Ec *nth* mutant complemented with Ec *nth* (Ec*nth*) and the EcΔ*nth* mutant complemented with Msmnth (EcΔ*nth*::Msm*nth*) was measured to assess if the Msm *nth* gene can restore the Nth glycosylase function in the EcΔ*nth* mutant (Fig. 3.7).

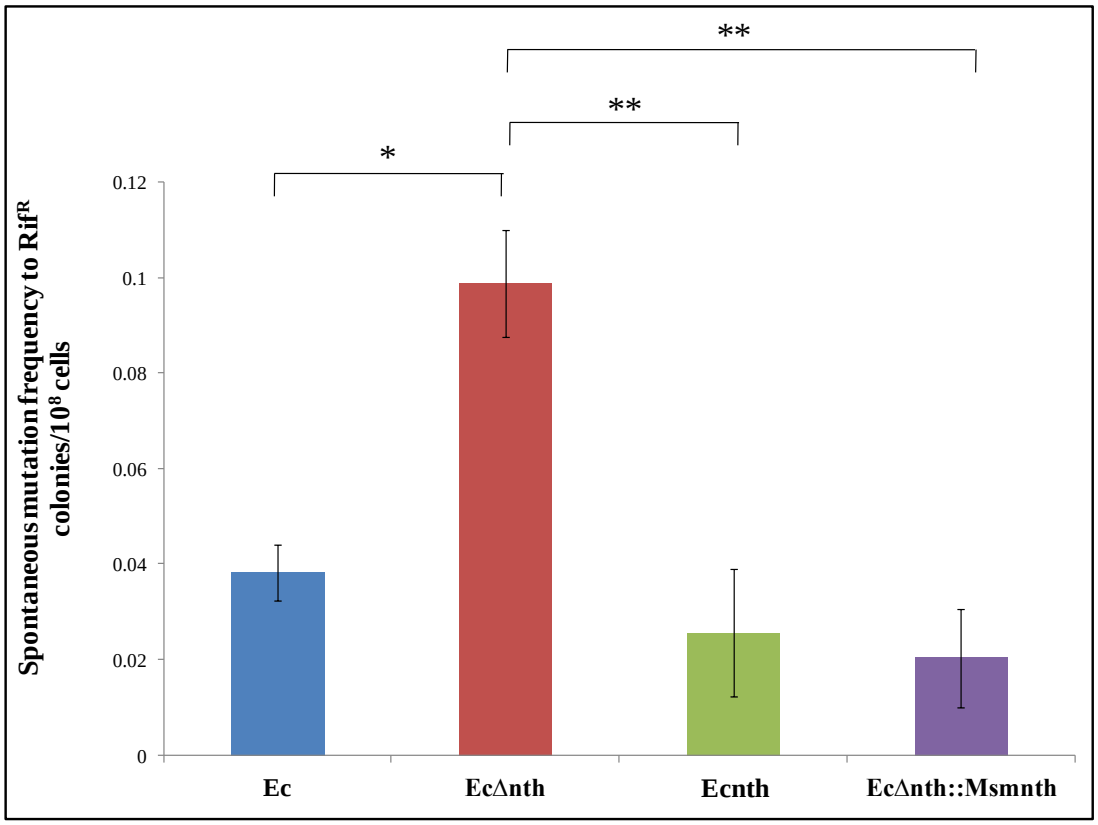



**Fig. 3.7: Complementation of the Ec *nth* knockout mutant (EcΔ*nth*) with Msm *nth*.** All Ec strains were assessed for spontaneous resistance to rifampicin. The data represented is the mean of 3 independent experiments ± SE. P-values were determined by the one way ANOVA where * = $p<0.05$ and ** = $p<0.01$ using GraphPad Prism Software.

There was a three fold increase in the spontaneous mutation frequency in the EcΔ*nth* mutant compared to the wild type Ec, while both *nth* expressing strains reduced the mutation frequency to levels equivalent to that of the Ec enzyme, thus confirming that the Msm *nth* gene is functioning as a DNA repair enzyme (Fig. 3.7). The statistical difference between the wild type Ec strain and the EcΔ*nth* was $p<0.05$ while it was $p<0.01$ between

the EcΔ*nth* and the Ec expressing *nth* and the Msm *nth* complemented strain (Fig. 3.7). From these data it is suggestive that Msm *nth* encoding the Nth glycosylase functions as a DNA glycosylase during DNA repair.

## 3.3 The construction and identification of site specific deletion mutants in *Mycobacterium smegmatis* (Msm)

### 3.3.1 Construction of the *nth* knockout vector

Primers engineered with unique restriction sites shown in Table 5.2.1.1 (Appendix 5.2.1) to facilitate the cloning process were used to amplify homologous flanking upstream (5') and downstream (3') sequences which included only about 100 bp of the 780 bp *nth* sequence on either side (illustrated in Fig. 3.8).

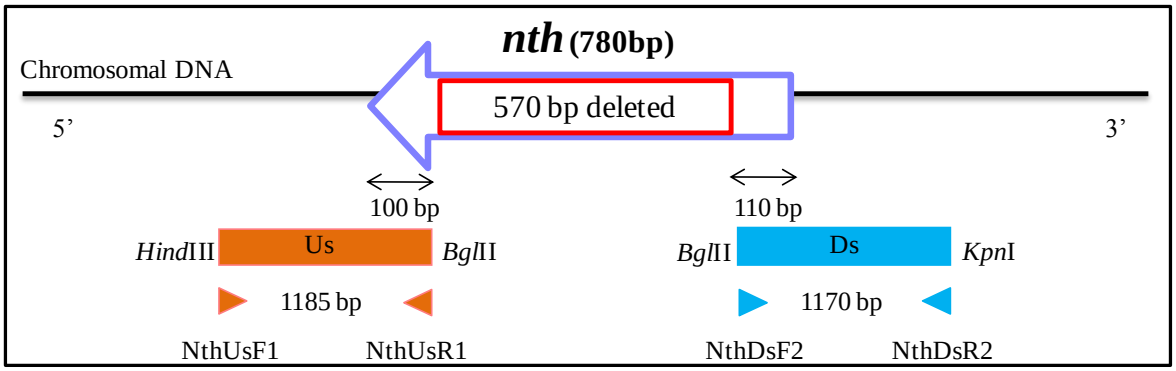
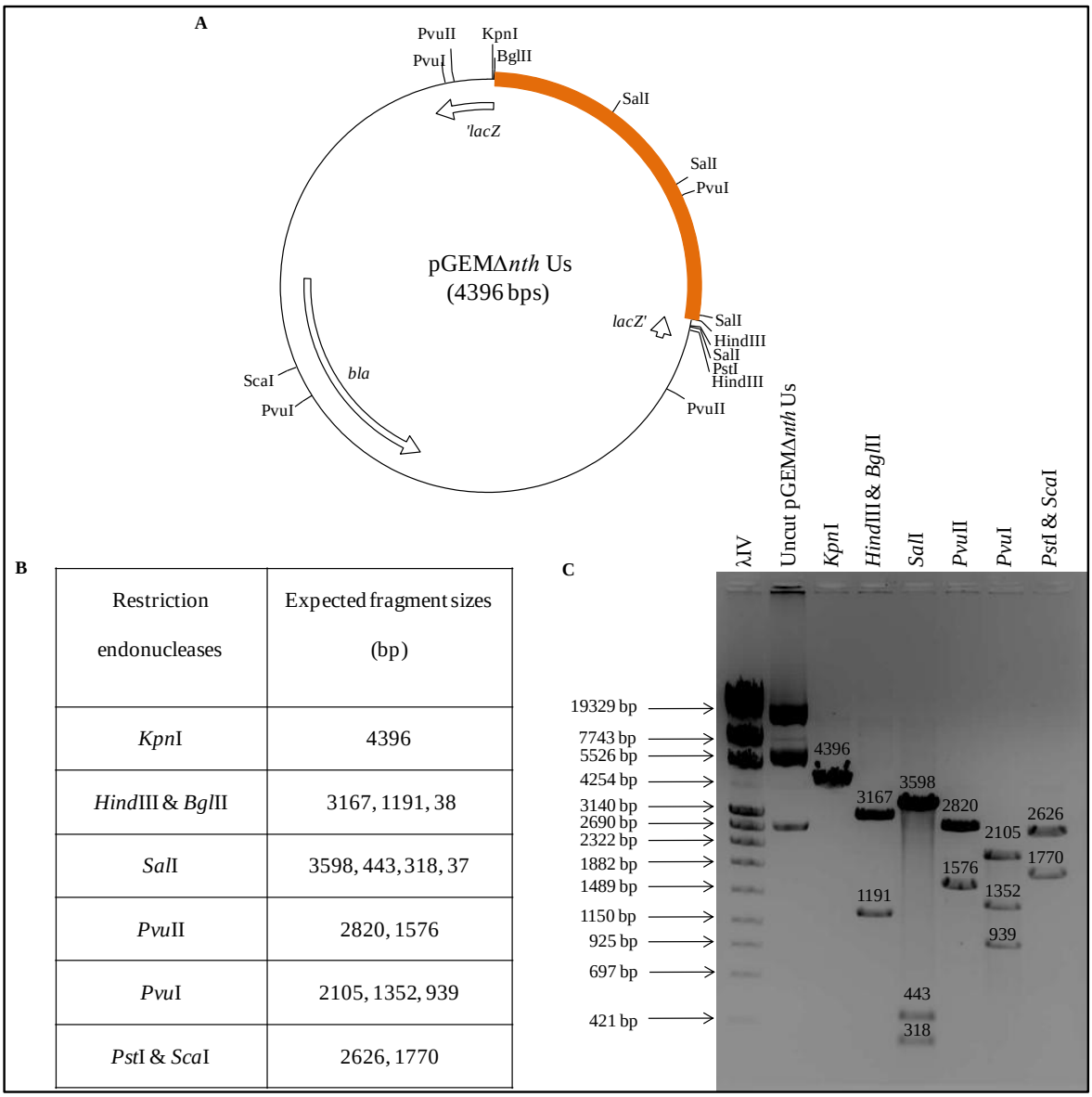



**Fig. 3.8: Schematic representation of PCR amplification of upstream and downstream regions of the *nth* gene.** Arrow heads represent forward and reverse primers and the targeted gene is highlighted and labelled in purple. Us - upstream and Ds - downstream fragments are shaded in orange and blue respectively. The "out- of- frame" deletion has been indicated as a red box.

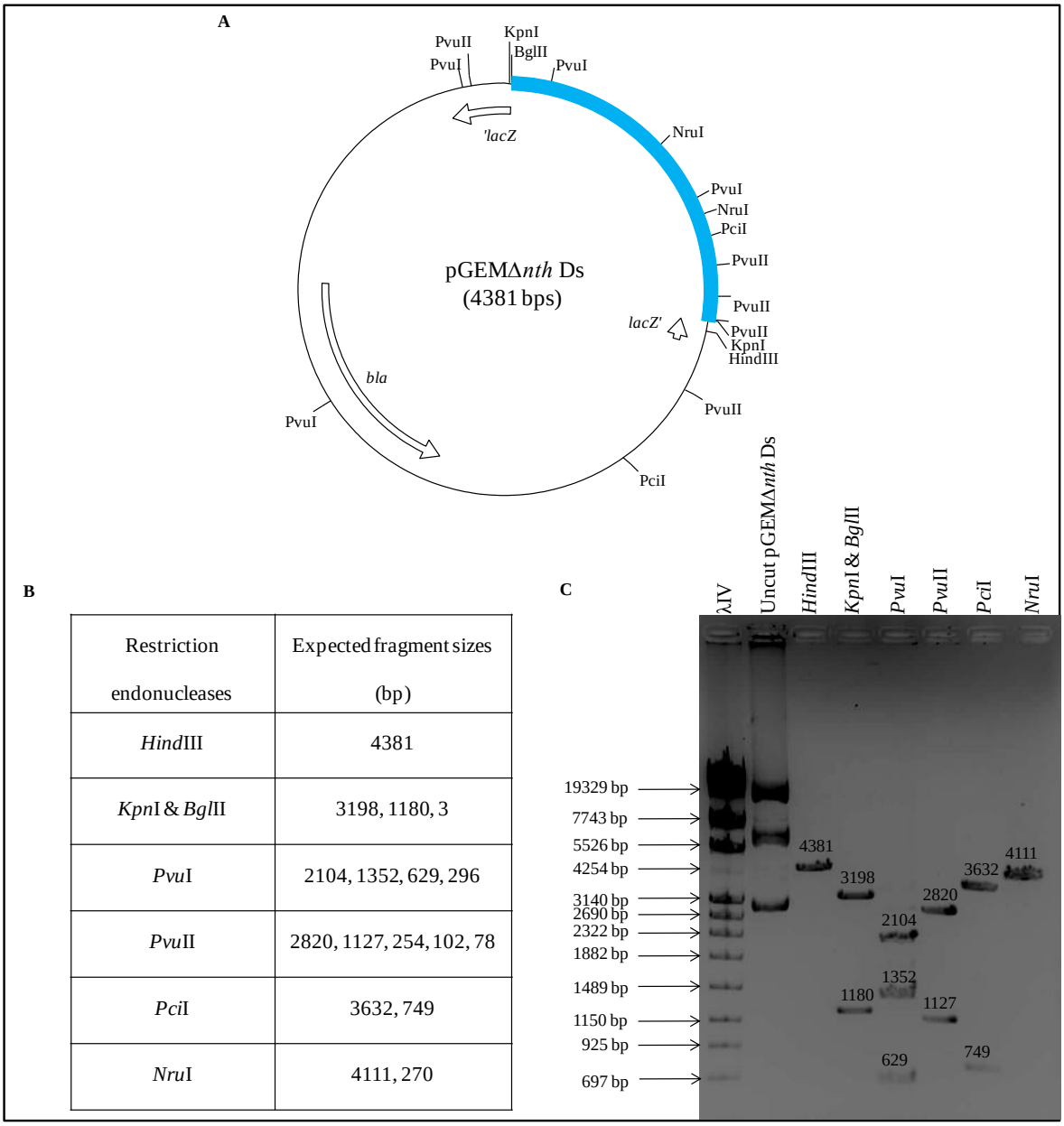
The upstream and downstream (Us and Ds) amplicons were cloned separately into the *Sma*I site of the replicating plasmid pGEM3Zf(+)in Fig. 5.3.1.2 (Appendix 5.3) to generate the vectors pGEMΔ*nth* Us and pGEMΔ*nth* Ds. The integrity of the resulting clones were confirmed by restriction analyses (Figs. 3.9 & 3.10) and one confirmed clone for the Us

and Ds was sequenced to ensure that no mutations were introduced during PCR amplification (sequencing data not shown).




| Restriction endonucleases | Expected fragment sizes (bp) |
|---|---|
| *Kpn*I | 4396 |
| *Hind*III & *Bgl*II | 3167, 1191, 38 |
| *Sal*I | 3598, 443, 318, 37 |
| *Pvu*II | 2820, 1576 |
| *Pvu*I | 2105, 1352, 939 |
| *Pst*I & *Sca*I | 2626, 1770 |

**Fig. 3.9: Restriction map analyses of pGEMΔ*nth* Us. A.** pGEMΔ*nth* Us map showing the cloned upstream region shaded in orange into pGEM3Zf(+). **B.** Table showing the expected sizes (bp) of the digested pGEMΔ*nth* Us with multiple restriction endonucleases. **C.** Restriction analysis of pGEMΔ*nth* Us with various restriction endonucleases.




| Restriction endonucleases | Expected fragment sizes (bp) |
|---|---|
| *Hind*III | 4381 |
| *Kpn*I & *Bgl*II | 3198, 1180, 3 |
| *Pvu*I | 2104, 1352, 629, 296 |
| *Pvu*II | 2820, 1127, 254, 102, 78 |
| *Pci*I | 3632, 749 |
| *Nru*I | 4111, 270 |

**Fig. 3.10: Restriction map analyses of pGEMΔ*nth* Ds. A.** pGEMΔ*nth* Ds map showing the cloned downstream region shaded in blue into pGEM3Zf(+). **B.** Table showing the expected sizes (bp) of the digested pGEMΔ*nth* Ds with multiple restriction endonucleases. **C.** Restriction analysis of pGEMΔ*nth* Ds with various restriction endonucleases.

### 3.3.1.1 Three way cloning

Sequencing (data not shown) and restriction endonuclease analyses (Figs. 3.9 and 3.10) confirmed no mutations in the Us and Ds regions and therefore, these fragments were excised with *Hind*III/*Bgl*II and *Kpn*I/*Bgl*II respectively from the replicating vector and cloned concurrently into the suicide vector p2NIL, in Fig. 5.3.1.3 (Appendix 5.3) with

compatible ends (illustrated in Fig. 3.11) to generate an *nth* deleted construct (p2NILΔ*nth*) which was verified by restriction mapping (Fig. 3.12).

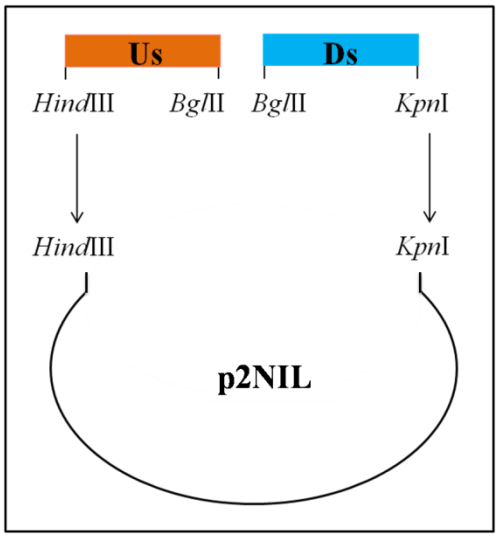
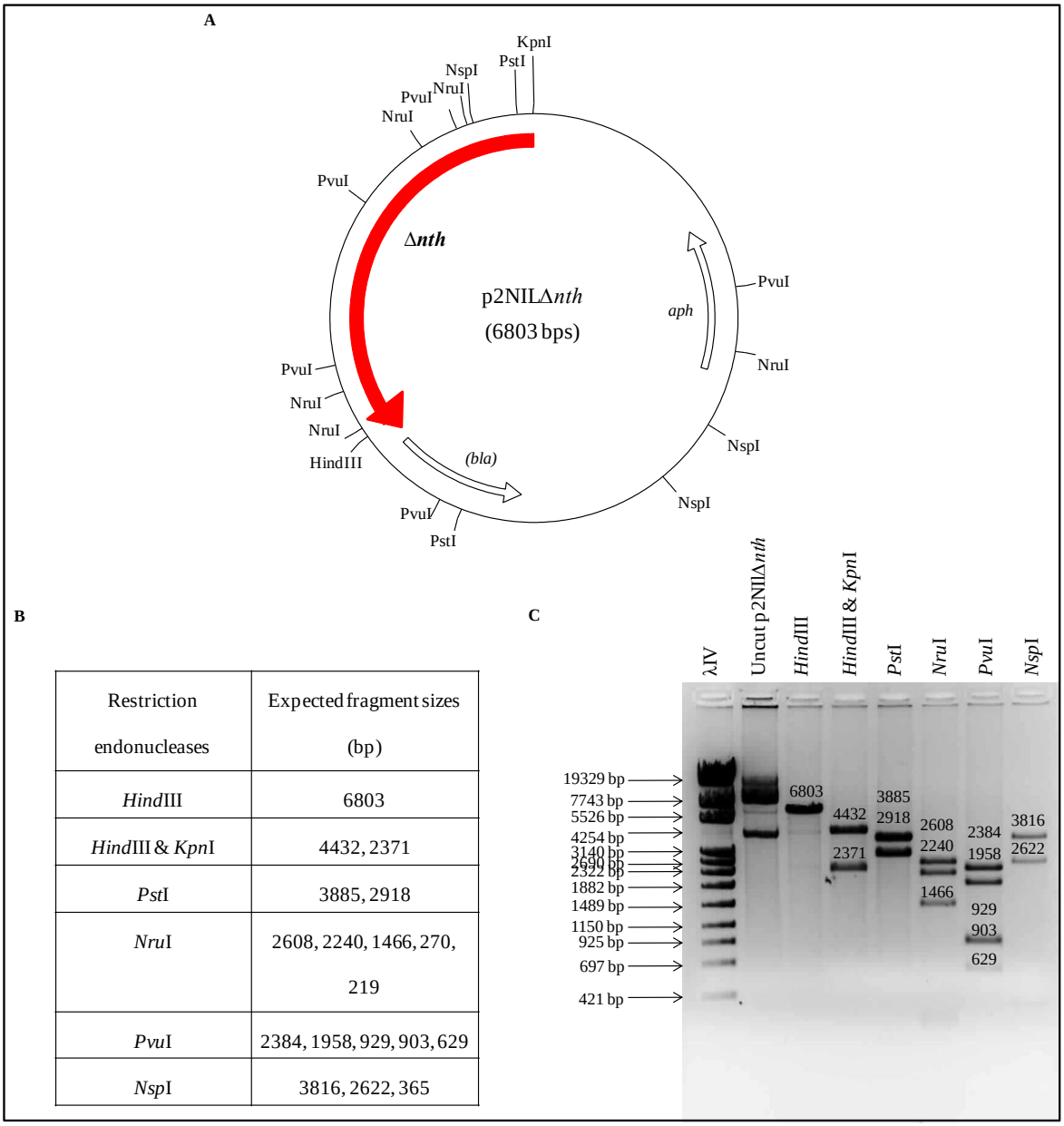



**Fig. 3.11: Directional cloning of upstream (Us) and downstream (Ds) regions into the suicide vector (p2NIL).** The *nth* inactivated allele is a product of the upstream (Us) and downstream (Ds) regions fused at the *Bgl*II site. The *Hind*III and *Kpn*I generate sites allowed for ligation of the *nth* deletion allele into the suicide vector (p2NIL).




| Restriction endonucleases | Expected fragment sizes (bp) |
|---|---|
| *Hind*III | 6803 |
| *Hind*III & *Kpn*I | 4432, 2371 |
| *Pst*I | 3885, 2918 |
| *Nru*I | 2608, 2240, 1466, 270, 219 |
| *Pvu*I | 2384, 1958, 929, 903, 629 |
| *Nsp*I | 3816, 2622, 365 |

**Fig. 3.12: Restriction mapping of p2NILΔ*nth*. A.** p2NILΔ*nth* map showing the *nth* deletion allele shaded in red into p2NIL. **B.** Table showing the expected sizes (bp) of the digested p2NILΔ*nth* with multiple restriction endonucleases. **C.** Restriction analysis of p2NILΔ*nth* with various restriction endonucleases.

### 3.3.1.2 Cloning of selectable and counter-selectable markers

Finally, additional selectable and counter-selectable markers harboured in the plasmid pGOAL19 (Fig. 3.13) was excised as a PacI cassette and introduced into the p2NILΔ*nth* vector backbone (illustrated in Fig. 3.13) to facilitate the selection of transformants with an

inactivated *nth* allele during the homologous recombination process. Initially, only one clone was isolated and verified by restriction mapping (Fig. 3.14).

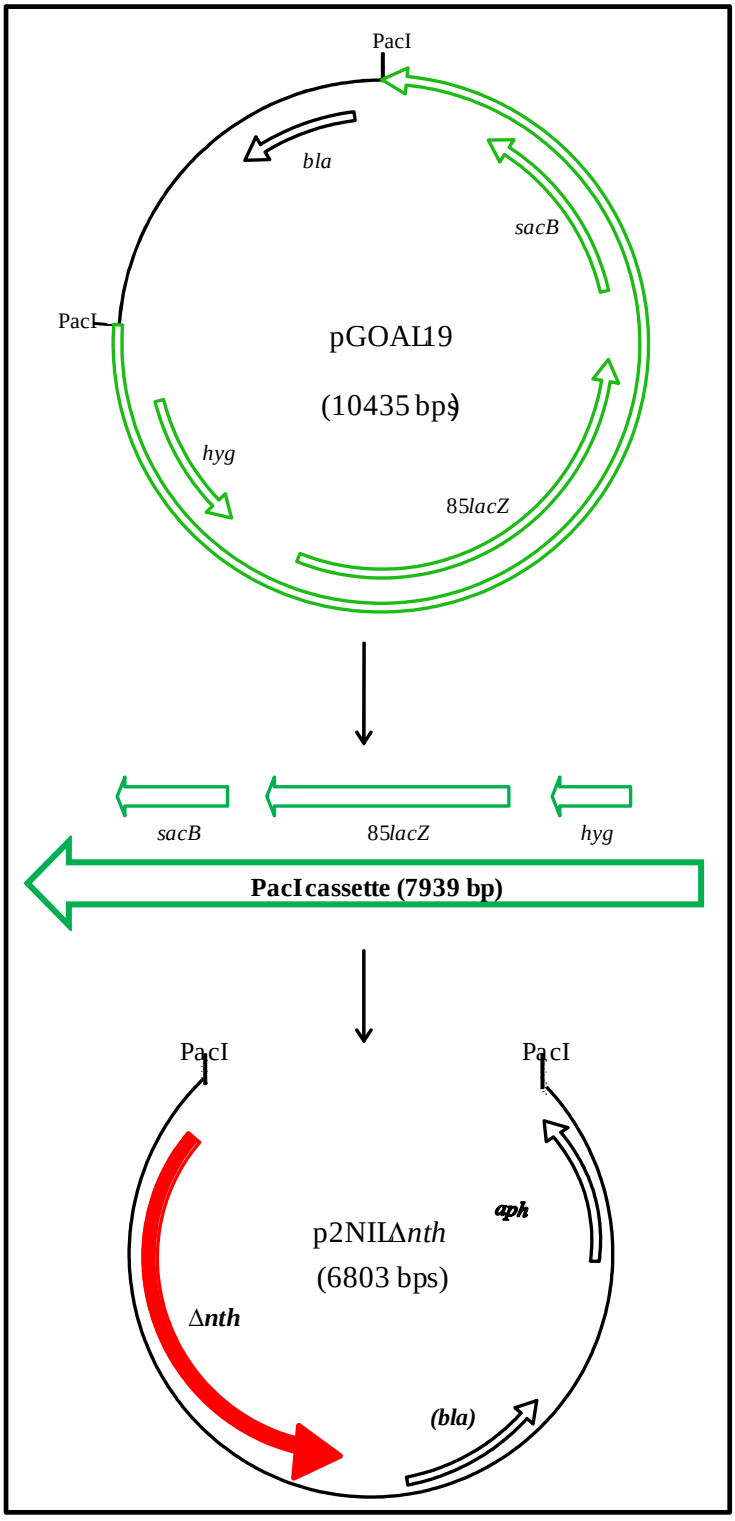
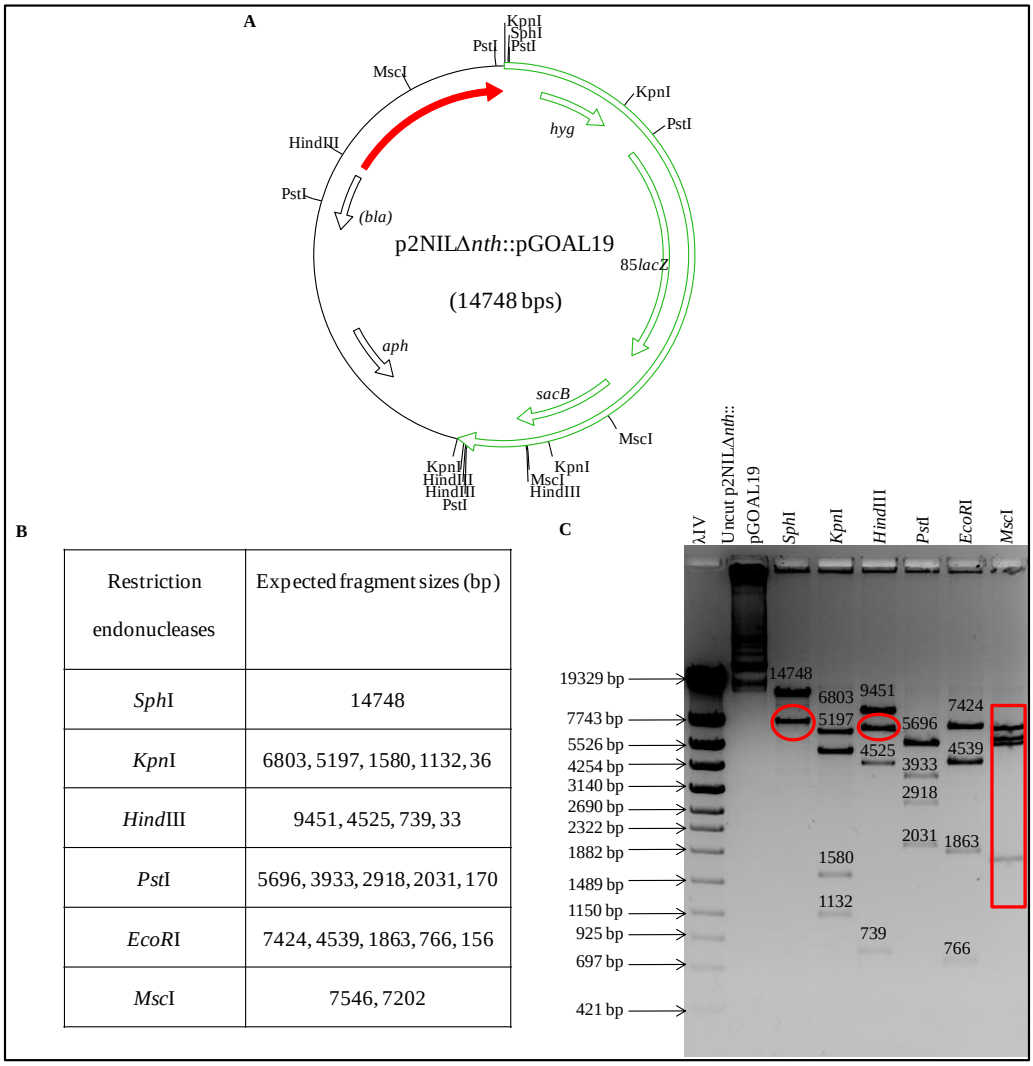



**Fig. 3.13: Cloning of PacI cassette containing selectable and counter-selectable markers.**

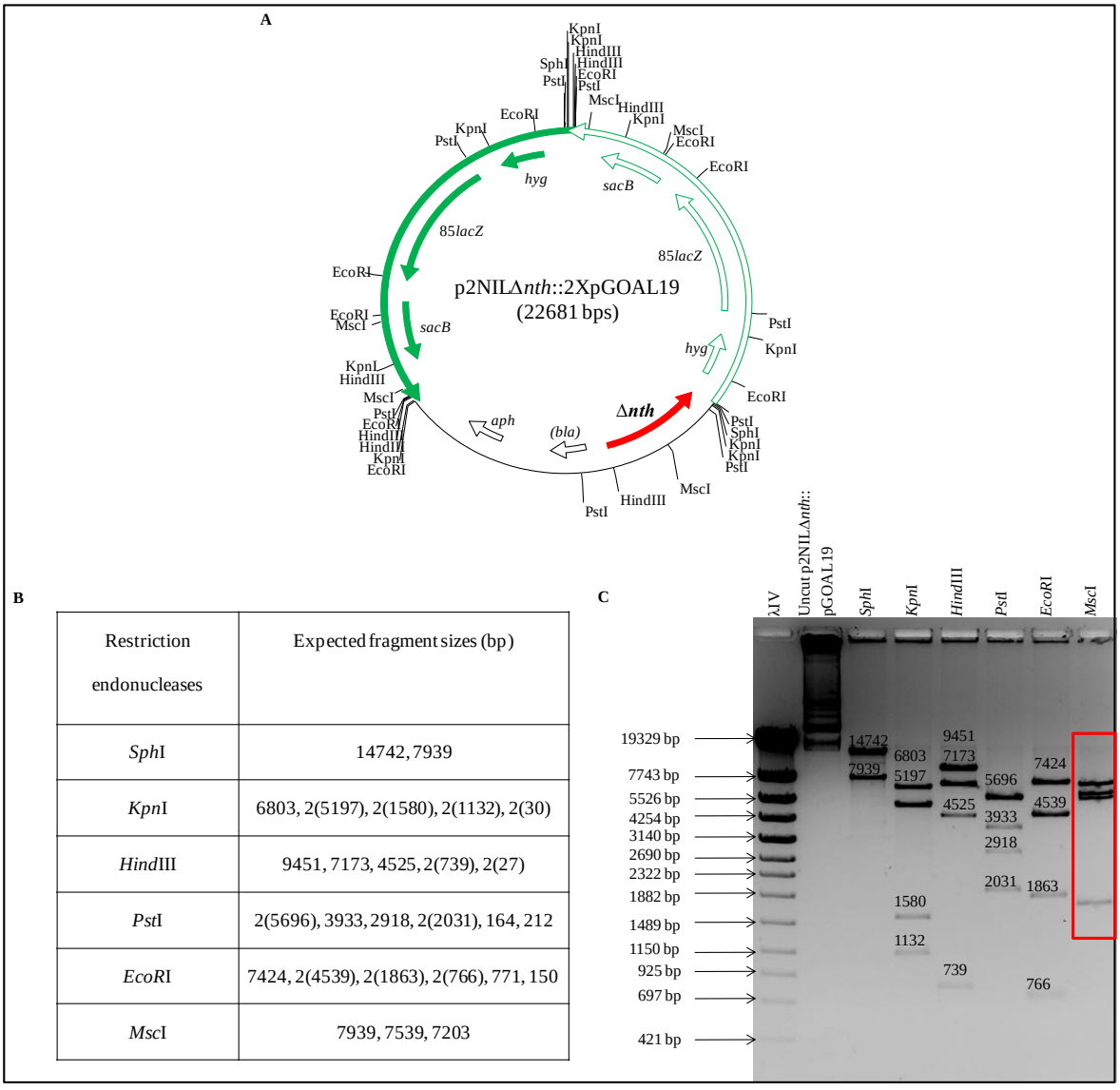



| Restriction endonucleases | Expected fragment sizes (bp) |
|---|---|
| *Sph*I | 14748 |
| *Kpn*I | 6803, 5197, 1580, 1132, 36 |
| *Hind*III | 9451, 4525, 739, 33 |
| *Pst*I | 5696, 3933, 2918, 2031, 170 |
| *Eco*RI | 7424, 4539, 1863, 766, 156 |
| *Msc*I | 7546, 7202 |

**Fig. 3.14: Restriction mapping for a potential p2NILΔ*nth*::pGOAL19 clone. A.** p2NILΔ*nth*::pGOAL19 map showing the cloned PacI cassette indicated by a green arrow into p2NILΔ*nth*. **B.** Table showing the expected sizes (bp) of the digested p2NILΔ*nth* with multiple restriction endonucleases. **C.** Restriction analysis of p2NILΔ*nth* with various restriction endonucleases. The red circles and box outline bands that do not correspond to the expected sizes in the table.

The sizes observed in Fig. 3.14 for the restriction digests: *Sph*I, *Hind*III and *Msc*I (bands outlined in red) of the potential p2NILΔ*nth*::pGOAL19 clone were not according to the expected sizes in the table in Fig. 3.14. The extra band in the *Sph*I digest was approximately the size of the PacI cassette (7939 bp shown in Fig. 3.13), suggesting that this clone could possibly have two PacI cassettes and hence a map with two PacI cassettes was constructed (Fig. 3.15) to make sense of the restriction digests. This clone was

renamed p2NILΔ*nth*::2XpGOAL19 and the expected sizes with the same restriction digest were predicted using the map (Fig. 3.15).

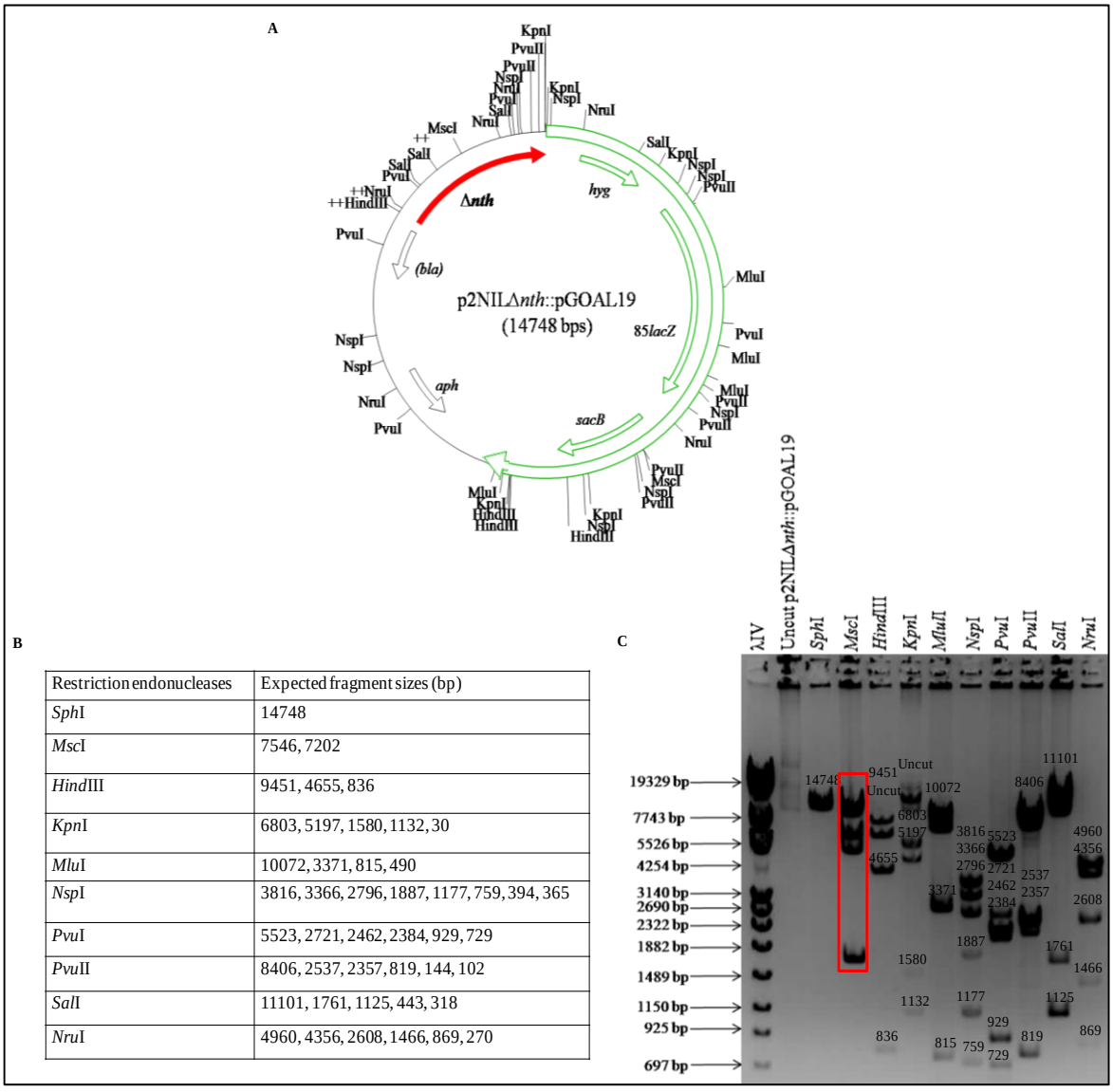



| Restriction endonucleases | Expected fragment sizes (bp) |
|---|---|
| *Sph*I | 14742, 7939 |
| *Kpn*I | 6803, 2(5197), 2(1580), 2(1132), 2(30) |
| *Hind*III | 9451, 7173, 4525, 2(739), 2(27) |
| *Pst*I | 2(5696), 3933, 2918, 2(2031), 164, 212 |
| *Eco*RI | 7424, 2(4539), 2(1863), 2(766), 771, 150 |
| *Msc*I | 7939, 7539, 7203 |

**Fig. 3.15: Restriction mapping of p2NILΔ*nth*::2XpGOAL19. A.** p2NILΔ*nth*::2XpGOAL19 map showing two PacI cassettes shown in green cloned into p2NILΔ*nth*. **B.** Table showing the expected sizes (bp) of the digested p2NILΔ*nth*::2XpGOAL19 with multiple restriction endonucleases. **C.** Restriction analysis of p2NILΔ*nth*::2XpGOAL19 with various restriction endonucleases. The red box outlined bands that did not correspond to the expected sizes in the table.

All the sizes observed in Fig. 3.15 corresponded to the predicted sizes in the table (Fig. 3.15) confirming that the p2NILΔ*nth*::2XpGOAL19 clone contained two PacI cassettes. However, the *Msc*I digest still did not give the predicted size fragments which were 7939,

7539 and 7203 bp (Fig. 3.15) and resulted in an extra band. Another potential clone (p2NILΔ*nth*::pGOAL19) which was obtained from a second transformation experiment upon restriction enzyme analysis also showed an extra fragment with *Msc*I but had a single PacI cassette (Fig. 3.16).




| Restriction endonucleases | Expected fragment sizes (bp) |
|---|---|
| *Sph*I | 14748 |
| *Msc*I | 7546, 7202 |
| *Hind*III | 9451, 4655, 836 |
| *Kpn*I | 6803, 5197, 1580, 1132, 30 |
| *Mlu*I | 10072, 3371, 815, 490 |
| *Nsp*I | 3816, 3366, 2796, 1887, 1177, 759, 394, 365 |
| *Pvu*I | 5523, 2721, 2462, 2384, 929, 729 |
| *Pvu*II | 8406, 2537, 2357, 819, 144, 102 |
| *Sal*I | 11101, 1761, 1125, 443, 318 |
| *Nru*I | 4960, 4356, 2608, 1466, 869, 270 |

**Fig. 3.16: Restriction mapping for p2NILΔ*nth*::pGOAL19. A.** p2NILΔ*nth*::pGOAL19 map showing the cloned PacI cassette in green into p2NILΔ*nth*. **B.** Table showing the expected sizes (bp) of the digested p2NILΔ*nth*::2pGOAL19 with multiple restriction endonucleases. **C.** Restriction analysis of p2NILΔ*nth*::pGOAL19 with various restriction endonucleases. The red box outlined bands that did not correspond to the expected sizes in the table.

To account for the sizes observed in the *Msc*I digest of the knockout constructs (Figs. 3.15 and 3.16); the p2NILΔ*nth*::pGOAL19, p2NILΔ*nth*::pGOAL19, p2NILΔ*nth* clones, pGOAL19 and the PacI cassette (illustrated in Fig. 3.13) were digested with *Msc*I (Fig. 3.17B). One extra band was common to pGOAL19 and the PacI cassette which off course as expected was also present in the p2NILΔ*nth*::2XpGOAL19 and p2NILΔ*nth*::pGOAL19 vectors, indicating that there was an un-annotated *Msc*I site in the PacI cassette of pGOAL19 (Fig. 3.17B). The PacI cassette was digested with restriction endonucleases *Pst*I and *Msc*I (Fig. 3.17B) and based on the map (Fig.3. 17B) the location of the *Msc*I site was predicted. The restriction digests (Fig. 3.17A) showed that the *Msc*I site was close to the *Pst*I site (Fig. 3.17B) in the un-annotated region of the PacI cassette in Fig. 5.3.2.1 (Appendix 5.3.2). In addition, purified pGOAL19 plasmid DNA obtained from four colleagues (1- 4 pGOAL19) in the laboratory was also digested with *Msc*I and compared to my pGOAL19 *Msc*I digest (N pGOAL19) to confirm that pGOAL19 indeed has two *Msc*I sites and the additional site observed was not due to a mutation or re-arrangement during the cloning of the PacI cassette into the p2NILΔ*nth* construct (Fig. 3.18). Therefore, the expected fragment sizes (bp) of p2NILΔ*nth*::2XpGOAL19 and p2NILΔ*nth*::pGOAL19 clones were 7203, 6356, 5538, 2(1792) and 7540, 5416, 1792 respectively when digested with *Msc*I.

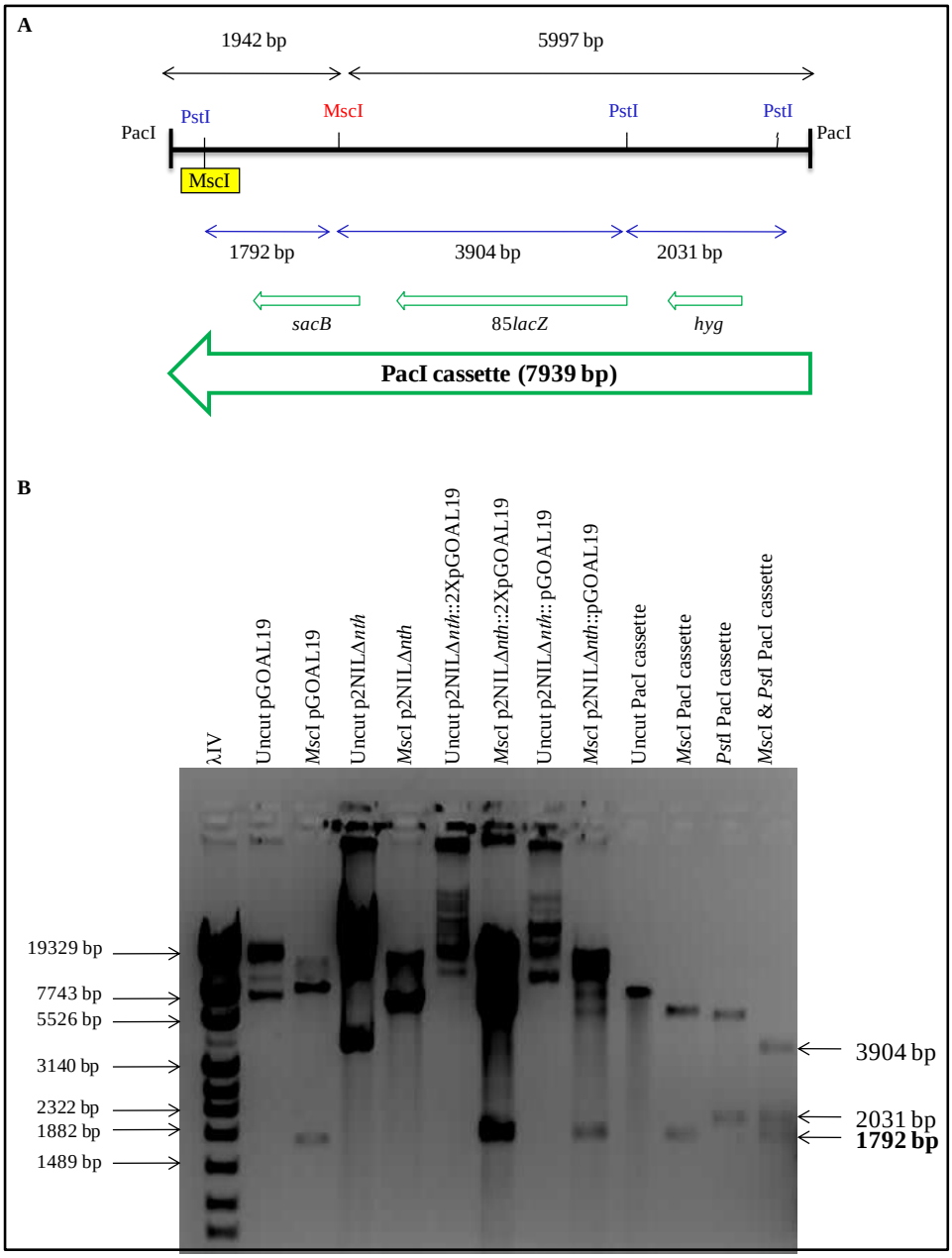



**Fig. 3.17: Restriction mapping of an un-annotated *Msc*I site in the PacI fragment of pGOAL19. A.** Restriction mapping with *Msc*I and/or *Pst*I to position the un-annotated *Msc*I site. **B.** A map of the PacI cassette (7939 bp) from pGOAL19 showing the predicted location of the previously un-annotated *Msc*I site shaded in yellow.

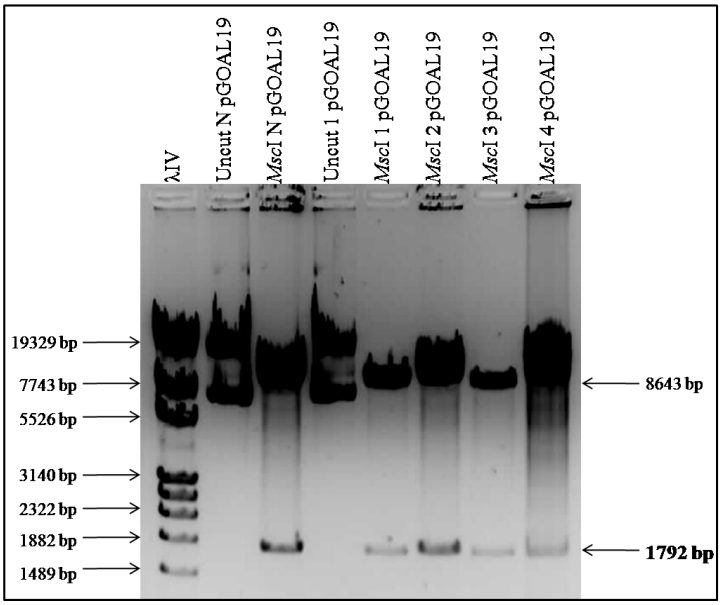



**Fig. 3.18: *Msc*I restrictions of various isolates of pGOAL19 plasmid DNA.** The restriction digests of various sources of pGOAL19 plasmid DNA with *Msc*I.

Both knockout constructs p2NILΔ*nth*::2XpGOAL19 and p2NILΔ*nth*::pGOAL19 were tested for sucrose sensitivity (explained in section 2.4.4) to ensure that the *sacB* was functional and to assess whether the duplication had a more dramatic effect. Both the strains showed a 6 log reduction in survival in the presence of sucrose (Fig. 3.19); despite the p2NILΔ*nth*::2XpGOAL19 construct having two PacI cassettes.

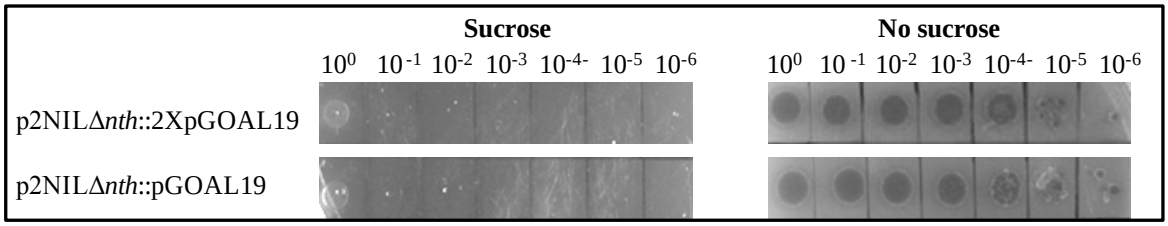



**Fig. 3.19: Sucrose sensitivity test of p2NILΔ*nth*::2XpGOAL19 and p2NILΔ*nth*::pGOAL19 knockout constructs.**

This suggested that both constructs were suitable vehicles for the sequential deletion of *nth* in the parental Msm strain and in selected mutant strains deficient in Fpg and/or Nei that were previously generated in our laboratory (Goosens, 2008 MSc; Goosens *et al*., unpublished). The strategy of mutant generation is illustrated in Fig. 3.20.

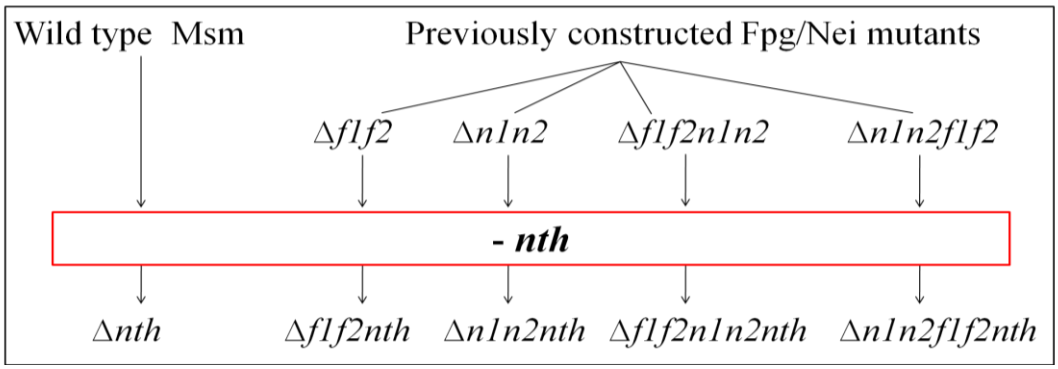



**Fig. 3.20: Strategy for generating Nth glycosylase deficient mutants.** The Δ*f1f2* and Δ*n1n2* mutant strains were generated by V. Goosens (2008, MSc) while, the Δ*f1f2n1n2* (Fpg lineage) and Δ*n1n2f1f2* (Nei lineage) mutant strains were generated by B. Gordhan (Goosens *et al*., unpublished). The sequential deletion mutants of *nth* in the previously constructed *fpg/nei* mutant strains and Msm were generated in this study.

Prior to generating the various *nth* mutant strains by homologous recombination the Fpg, Nei and Fpg/Nei mutant strains (Goosens, 2008 MSc; Goosens *et al*., unpublished) were verified by PCR using primers in Table 5.2.1.3 (Appendix 5.2.1) specific for each of the DNA glycosylases to ensure that the mutants were correct. The primers amplified different size fragments for the intact and deleted genes in the wild type and the various mutant strains as shown in Fig. 5.2.2.1 (Appendix 5.2.2).

### 3.3.2 Identification of single crossover (SCO) and double crossover (DCO) events

A two step selection process previously described by Gordhan & Parish (2001) was used for allelic exchange mutagenesis for the generation of mutants lacking the *nth* gene in the wild type strain and the Fpg/Nei deficient mutants (illustrated in Fig. 3.21). Briefly, 2-6 µg of p2NILΔ*nth*::2XpGOAL19 or p2NILΔ*nth*::pGOAL19 DNA was used to electroporate (section 2.3.9.1.2) a mid-log phase culture of the respective strain and plated on 7H10 plates containing kan and X-gal for selection of transformants. During the homologous recombination process, a SCO event generated a strain containing both the wild type *nth* gene as well as the inactivated allele p2NILΔ*nth* with the marker genes in the vector

backbone (illustrated in Fig. 3.21). SCOs were identified as blue colonies that were subjected to a second selection on 7H10 media containing sucrose which is converted to a toxic by-product by levansucrase, an enzyme encoded for by the *sacB* gene. Hence, during the second recombination event, only mutants that lost the vector backbone together with the marker genes grew in the presence of sucrose. The DCO mutant therefore, had the functional copy of the *nth* gene replaced with the inactivated copy of the gene (illustrated in Fig. 3.21, Gordhan & Parish, 2001). The SCO and DCO mutants were genetically verified using PCR (Fig. 3.22) and Southern blot analyses (Figs. 3.23 & 3.24) to confirm the integrity of the regions Us and Ds of the deleted Δ*nth* allele.

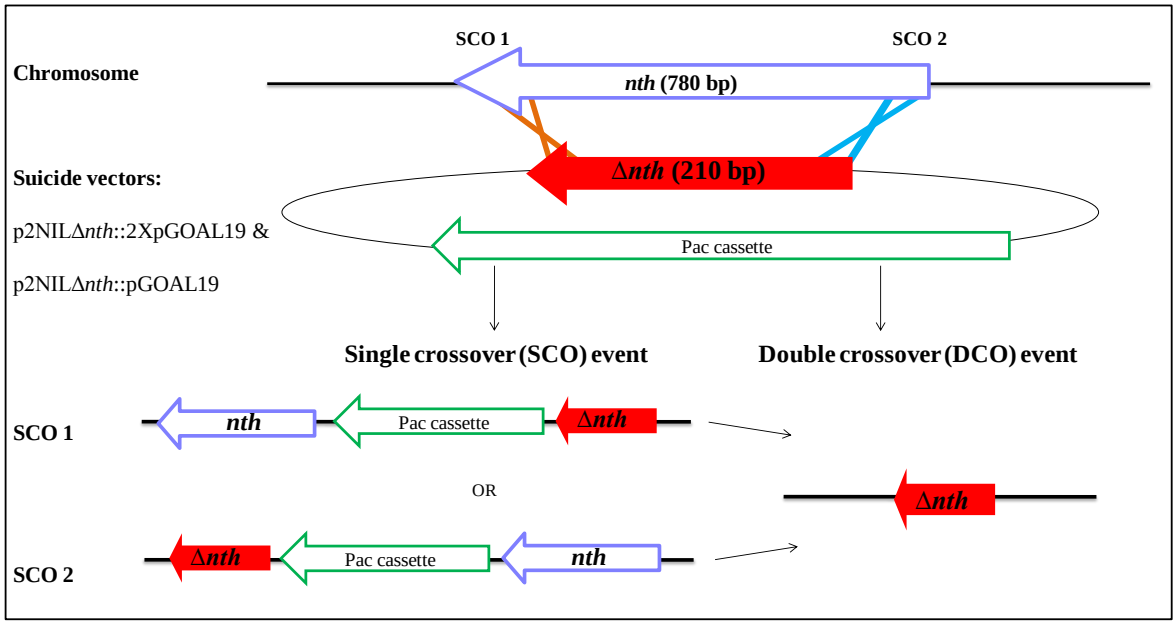



**Fig. 3.21: Allelic replacement by homologous recombination** (Adapted from Gordhan & Parish, 2001). Homologous recombination involves the genetic exchange of the chromosomal *nth* gene with the inactivated *nth* gene delivered by the *nth* knockout construct (p2NILΔ*nth*::2XpGOAL19 or p2NILΔ*nth*::pGOAL19) resulting in the deletion of the Nth DNA glycosylase. During the homologous recombination process, a single cross over (SCO) event occurred either upstream (SCO 1) or downstream (SCO 2) of the chromosomal *nth* gene and the suicide vector carrying the deleted allele of the *nth* gene, generating a strain containing both the wild type *nth* gene as well as the inactivated allele, with the marker genes in the vector backbone. The SCOs underwent a second cross over event to generate double crossover (DCO) mutants, which had the functional copy of the *nth* gene replaced with the inactivated copy of the gene (Gordhan & Parish, 2001).

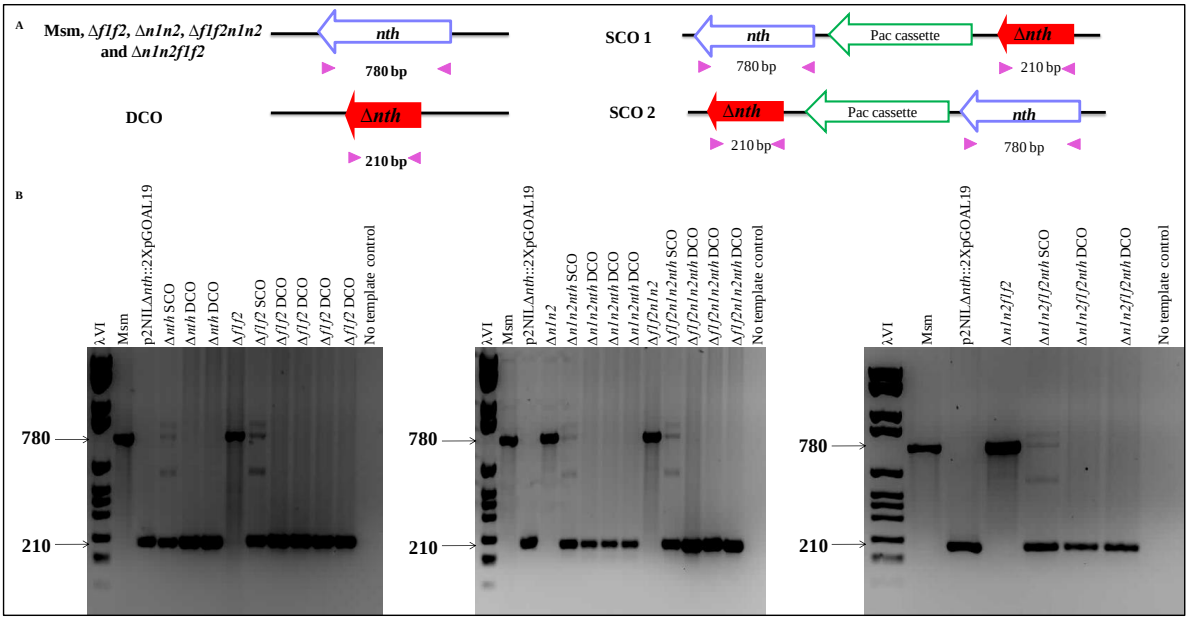



**Fig. 3.22: Confirmation of possible *nth* DCOs by PCR amplification. A.** Illustration for distinguishing wild type revertants, SCOs and DCOs using primers indicated as pink arrow heads described in Table 5.2.1.3 (Appendix 5.2.1). **B.** PCR amplification of possible Δ*nth* knockout mutants following allelic exchange mutagenesis. Msm and the knockout construct p2NILΔ*nth*::2XpGOAL19 served as positive controls to distinguish between DCOs, SCOs and wild type revertants.

PCR confirmed that the *nth* gene was intact in the Fpg/Nei mutants and deleted in the DCOs or *nth* mutants (Fig. 3.22). SCOs showed the expected two amplicons with a preference for the deleted allele as it is smaller in addition two other nonspecific bands. Only one PCR verified DCO was analysed further by Southern blot analyses to ensure that the mutant was genotypically intact (Figs. 3.23 & 3.24).

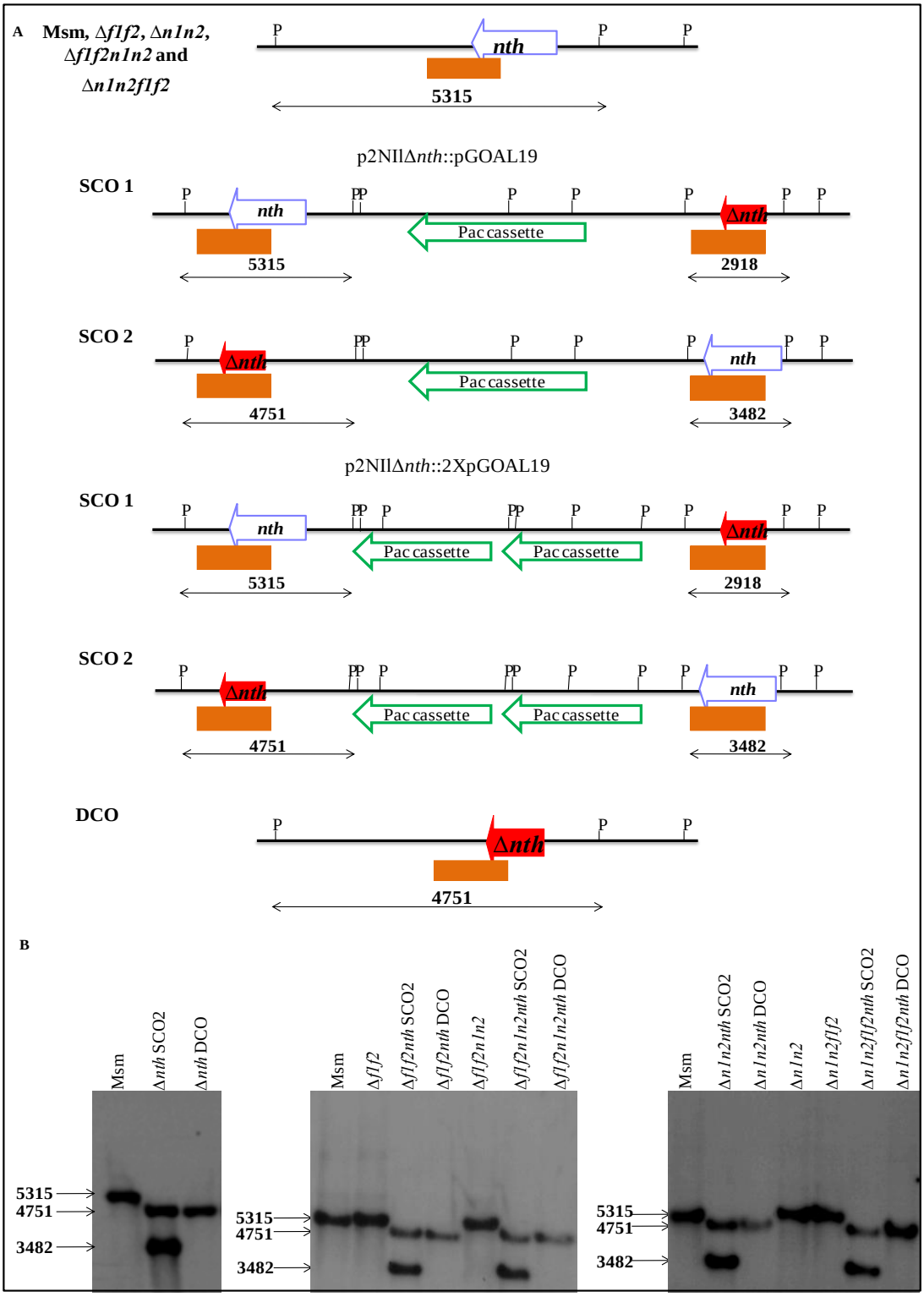



**Fig. 3.23: Southern blot upstream analysis of Δ*nth* mutants. A.** Southern blot strategy to distinguish the wild type, the SCO and DCO mutants. **B.** Chromosomal DNA from the parental, SCO and DCO mutant strains was digested with *Pst*I and probed with the Us PCR amplicon (1185 bp) represented as the orange rectangle of either knockout constructs p2NILΔ*nth*::2XpGOAL19 or p2NILΔ*nth*::pGOAL19. The expected fragment sizes to distinguish mutants from the parental strain are shown in bp.

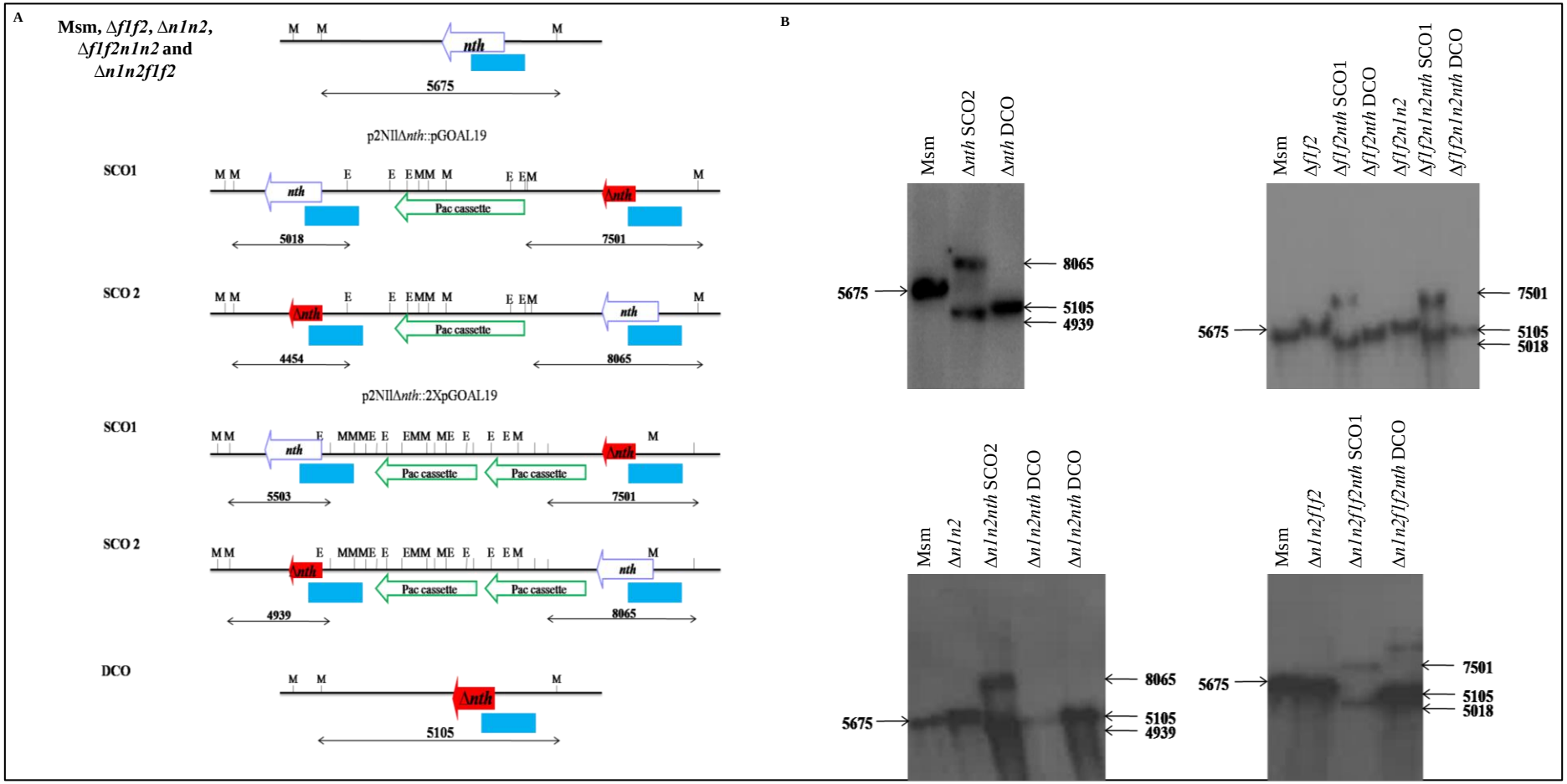



**Fig. 3.24: Southern blot downstream analysis of Δ*nth* mutants. A.** Southern blot strategy to distinguish the wild type, the SCO and DCO mutants. **B.** Chromosomal DNA from the parental, SCO and DCO mutant strains was digested with *Mlu*I and *Eco*RI and probed with the Ds PCR amplicon (1170 bp) represented as the orange rectangle of either knockout constructs p2NILΔ*nth*::2XpGOAL19 or p2NILΔ*nth*::pGOAL19. The expected fragment sizes to distinguish mutants from the parental strain are shown in bp.

Both the Us (Fig. 3.23) and Ds (Fig. 3.24) regions for each of the mutants showed no re-arrangements and all the genes were maintained in the same chromosomal context indicating that the *nth* deletion was site specific and did not influence the surrounding genes. These genotypically verified mutants referred to as: Δ*nth*, Δ*f1f2nth*, Δ*n1n2nth*, Δ*f1f2n1n2nth* and Δ*n1n2f1f2nth* were assessed in all downstream phenotypic analysis.

## 3.4 Genetic complementation

Preliminary data have shown that Msm mutants lacking two Fpg/Nei glycosylases showed no reduced survival in response to $H_2O_2$ and no mutator phenotype (Goosens, 2008 MSc). However, mutator phenotypes were observed when all three Fpg/Nei/Nth glycosylases were deleted in *Escherichia coli* (Ec) and *S.typhimurium* (Blaisdell *et al*., 1999; Suvarnapunya *et al*., 2003; Suvarnapunya & Stein, 2005). The deletion of the Nth DNA glycosylase together with the Fpg/Nei family of DNA glycosylases was expected to exaggerate the mutator phenotype. Hence, it was essential to genetically complement the various *nth* deficient mutants with the functional *nth* gene to ensure that the observed phenotype was due to loss of the *nth* gene function and not due to downstream polar effects during the homologous recombination process.

### 3.4.1 Construction of the complementation vector

The complementation vector was constructed by PCR amplification of the *nth* gene including approximately 350 bp upstream of the start codon to include the native promoter region using primers in Table 5.2.1.1 (Appendix 5.2.1). The amplified sequence was cloned into an *E. coli – Mycobacterium* integrating shuttle vector shown in Fig. 5.3.1.1 (pTWEETY Appendix 5.3.1) that has an *attP* site for integration into the *attB* locus on the mycobacterial chromosome (Pham *et al*., 2007). The complementation vector was analysed

by restriction enzyme digestions (Fig. 3.4) and sequenced (data not shown). Post sequencing once the fragment was confirmed to be error free, the complementation vector was electroporated into the appropriate mutant strains and the resulting colonies were assessed for the intact *nth* gene by PCR in Fig. 5.2.2.2 (Appendix 5.2.2) using NthCF and NthCR primers in Table 5.2.1.1 (Appendix 5.2.1). Only one PCR verified integrant per strain was further analysed for expression of the Nth glycosylase by semi-quantitative reverse transcriptase-polymerase chain reaction (SQ RT-PCR; Fig. 3.25).

Briefly, RNA was isolated from the Msm strain, the Fpg/Nei mutants, *nth* deficient mutant and *nth* complemented strains, converted to cDNA (illustrated in Fig. 3.25) which was amplified using the SigA or Nth primers in Table 5.2.1.4 (Appendix 5.2.1). The results (Fig 3.25B) showed that there were equal amounts of RNA converted to cDNA for amplification with either SigA or Nth primers, hence cDNA samples amplified with SigA primers served as an appropriate positive control to compare expression of the Nth glycosylase (Fig. 3.25B). The no template control (NTC) and no reverse transcriptase controls for each of the samples produced no amplicons suggesting that there was no genomic contamination in the PCR reagents and in the RNA isolation. Therefore, the amplicons/products observed in Fig. 3.25B represent the expression of the Nth glycosylase in each of the strains. The expression observed in the wild type and Fpg/Nei mutants suggested that this glycosylase was functional under normal growth conditions and hence is constitutively expressed. Various *nth* deficient mutants showed the lack of expression of the *nth* gene mimicking the negative controls. The intensity of the Msm integrant bands amplified by Nth primers in each of the five panels corresponded to the intensity of SigA products (Fig. 3.25B). However, the bands in Δ*n1n2nth*::*nth* and Δ*f1f2n1n2nth*::*nth* strains were more intense as compared to the intensity of bands shown in the respective Msm

(wild type) and Δ*n1n2* and Δ*f1f2n1n2* mutant strains (Fig. 3.25B) implying that there was more *nth* expression in these strains. While, in some strains (Δ*f1f2* and Δ*n1n2f1f2*) the band intensity was greater than Msm and the genomic standard (Fig.3.25B) suggesting that there was more *nth* expression in these strains. The verified *nth* deficient mutants and complemented derivatives were phenotypically characterised to determine whether the *nth* is involved in DNA repair (section 2.6).

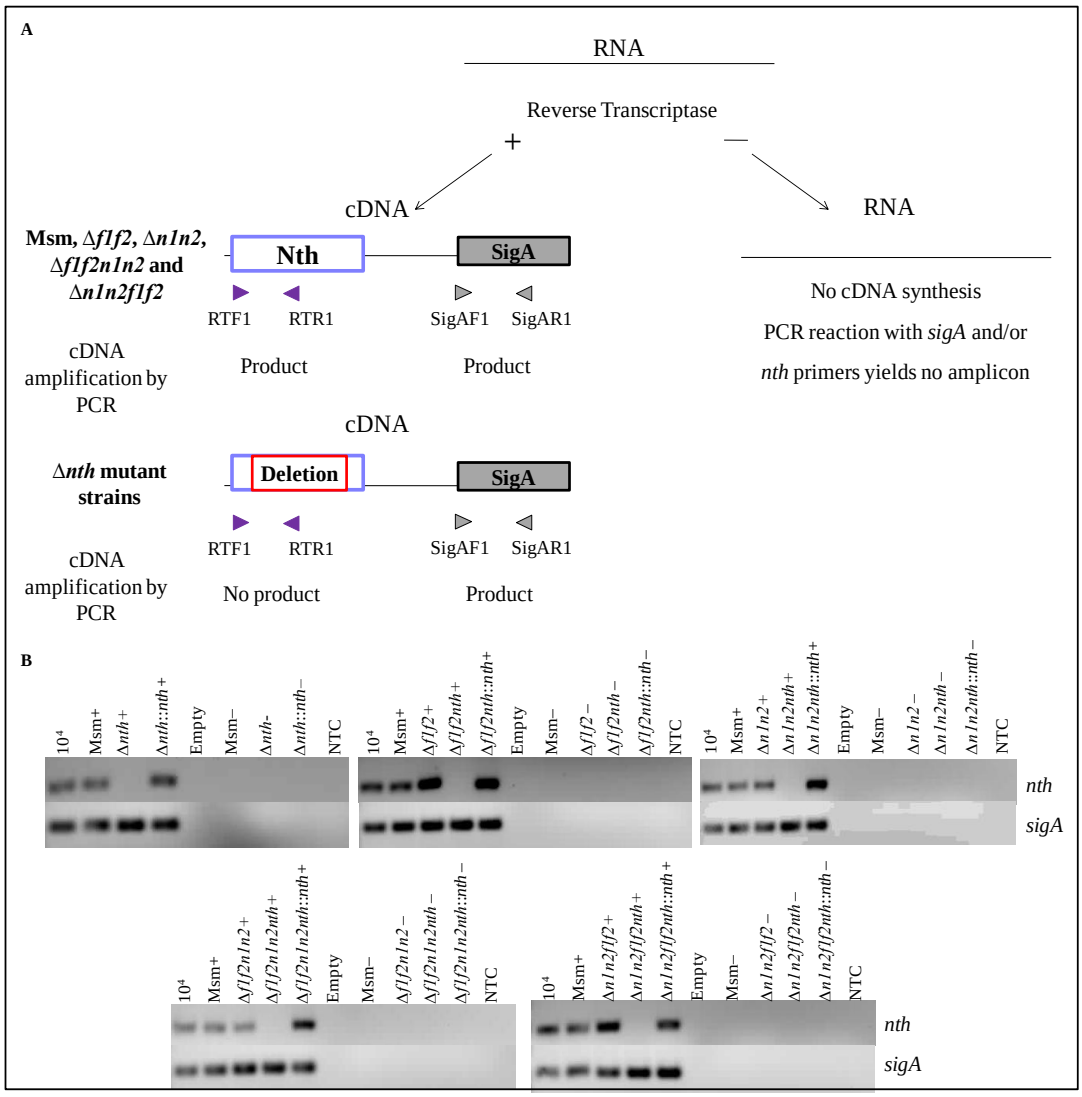



**Fig. 3.25: SQ RT-PCR analysis to confirm *nth* expression of Msm Nth integrants. A.** Illustration of SQ RT-PCR protocol. RNA extraction and reverse transcription was performed as previously described (Section 2.5.2) and 5 µl of the amplified cDNA using specific primers in table were loaded onto a 2% agarose gel. **B.** Each of the 5 panels shows the expression of the *nth* gene in the Msm integrants. cDNA samples were amplified to detect the *nth* transcripts and no reverse transcriptase (-) samples were used as negative controls to detect DNA contamination. The $10^4$ genomic standard was included to compare band intensities and the no template control (NTC) was used to detect DNA contamination from PCR reagents.

## 3.5 Phenotypic characterization

Genetically verified mutants lacking the *nth* gene and their respective complemented strains were assessed for (1) growth under normal culture conditions, (2) survival under oxidative stress conditions and (3) increase in mutator phenotypes as compared to the wild type Msm strain to allow for a better understanding of the role of the Nth DNA glycosylase in the absence and presence of the Fpg/Nei family of DNA glycosylases. Colonies of all strains recovered from the freezer for phenotypic experiments were PCR verified for the loss of the various Fpg/Nei DNA glycosylases as shown in Fig. 5.2.2.1 (Appendix 5.2.2) using primers described in Table 5.2.1.3 (Appendix 5.2.1) to ensure once again that the strains were all correct.

### 3.5.1 Growth kinetics

Msm mutants lacking the Nth glycosylase, the Fpg/Nei mutant strains together with the parental Msm were grown to log phase in 7H9 broth and used to inoculate fresh media (section 2.6.1). Growth of these cultures was monitored every 3 hours (hrs) as optical density (OD) measurements at 600nm over a 40 hour (hr) period. No significant differences in growth kinetics was observed between the Msm strain, Fpg/Nei mutants, the various *nth* mutants and the complemented strains (Fig. 3.26) as all strains grew at the same rate indicating that these genes were not essential for the growth of Msm.

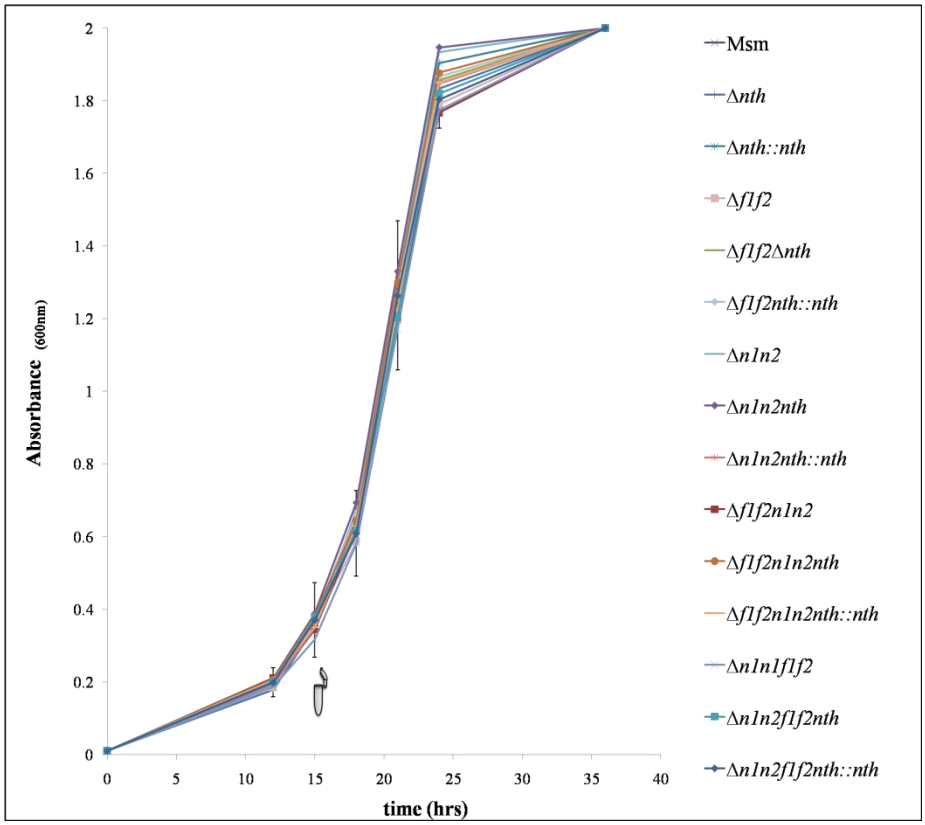



**Fig. 3.26: Comparative growth kinetics.** The growth of the various strains was monitored as $OD_{600nm}$ measurements. Data represent averages and SE from 3 independent experiments. The grey eppendorf indicates that a sample was taken at $OD_{600nm}$ of 0.35 for all strains for SQ RT-PCR. One way ANOVA analysis showed no significant differences between strains (P-values >0.1 for all points).

### 3.5.2 Assessment of cell viability post $H_2O_2$ treatment for Msm *nth* mutants

The optimal hydrogen peroxide ($H_2O_2$) concentrations that caused about a three log kill in Msm was previously optimised (Goosens, 2008 MSc) to be 2.5 mM. However, in this study the effect of growth phase of the cells was not determined. Hence, different stages of early (0.35), mid (0.50) and late (0.70) log phase Msm cultures were treated with 2.5 mM $H_2O_2$ and assessed for cell viability to determine the optimal growth phase for effective $H_2O_2$ killing (section 2.6.2.2). The growth phase at which the bactericidal activity of $H_2O_2$ that was most effective was established to be at an $OD_{600nm}$ of 0.35 (Fig. 3.27) and therefore, these conditions were used in all further $H_2O_2$ susceptibility assays.

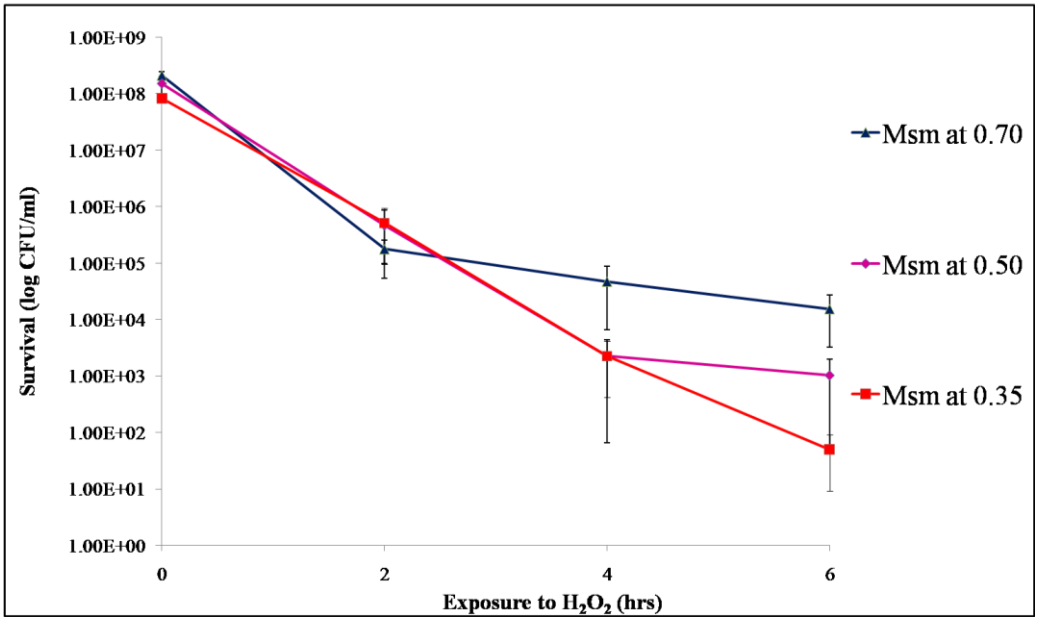



**Fig. 3.27: Assessment of wild type Msm with 2.5 mM $H_2O_2$ at $OD_{600}$ of 0.70, 0.50 and 0.35 over a period of 6 hrs.** Values are the means of 3 biological replicates (± SE).

Msm, the single *nth* mutant and the respective complemented strains were grown in 7H9 to early log phase after which the cultures were split into two equal amounts. The one set of cultures were treated with 2.5 mM $H_2O_2$ and the growth was regularly monitored over a period of 6 hrs (referred as treated). The second set of cultures served as negative controls since they were untreated but allowed to grow for the same length of time (referred as untreated). In each instance, 100 µl of dilutions from $10^0$ to $10^{-8}$ were spread in duplicate on 7H10 plates every 2 hrs for 6 hrs to assess the susceptibility of the Msm *nth* mutant to oxidative damage compared to the wild type Msm strain. The loss of *nth* in Msm did not show a reduced survival phenotype under oxidative stress conditions as growth of the mutant strain was comparable to the parental and the complemented strain. However, there was almost a 5 log difference in survival between the treated and untreated cultures for all the strains as analysed with a paired t-test ($p<0.001$) (Fig. 3.28).

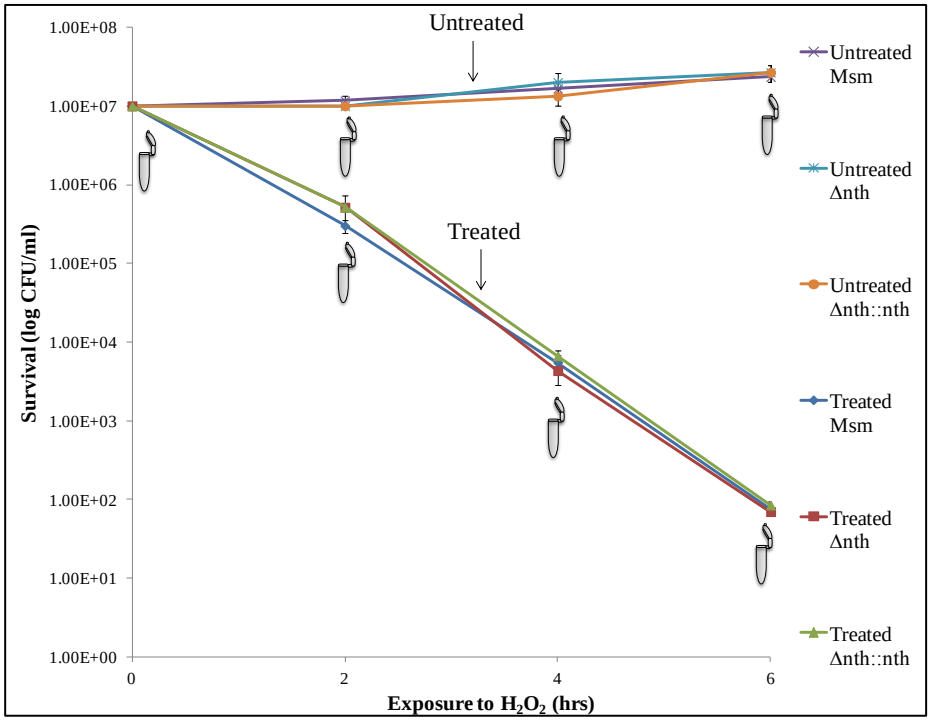



**Fig. 3.28: Assessment of wild type Msm, single *nth* mutant and its respective complemented strain with and without 2.5 mM $H_2O_2$ over a period of 6 hrs.** Data represent averages and SE from 3 independent experiments. Untreated cultures were significantly different as compared to the treated cultures using a $p< 0.001$ paired t-test. Grey eppendorfs indicates time points at which samples were taken for measuring intracellular $H_2O_2$ concentrations.

To ascertain whether the Nth glycosylase has the ability to reduce the accumulation of intracellular $H_2O_2$, the intracellular concentrations of $H_2O_2$ for the treated and untreated cultures of Msm, the single *nth* mutant and its respective complemented strain was measured as described in section 2.6.2.3 (Fig. 3.29). Intracellular $H_2O_2$ or its derivatives in the form of ROS were kept at low levels for both untreated and treated Msm indicating that wild type Msm has adept mechanisms to deal with oxidative stress (Fig. 3.29). However, this was not the case for the treated single *nth* mutant where levels of ROS increased with time (Fig. 3.29) indicating that the ROS generated by the exogenous agent overwhelmed the system and the organisms was unable to maintain the ROS to below toxic levels. It is unclear whether Nth which is classified as a DNA repair glycosylase also has a role in ROS detoxification. The concentration of ROS for the untreated single *nth* mutant doubled until 4 hrs but then decreased at 6 hrs (Fig. 3.29) suggesting that the organisms'

endogenous detoxification mechanisms were controlling the amount of ROS. The untreated complemented strain showed ROS concentrations similar to the parental Msm strain over the 6 hr period but the treated complemented strain showed ROS levels almost equivalent to that observed for the *nth* deficient mutant (Fig. 3.29). This was unexpected as the complemented strain should have resulted in ROS levels comparable to Msm. These data imply that *nth* was not complemented but the expression data (Fig. 3.25) negate this argument as *nth* is expressed in the complemented strain to wild type levels. Based on the data for both the treated and untreated *nth* deletion mutant, it is possible that the Nth glycosylase can detect high levels of ROS and control it as part of the mycobacterial detoxification system. However, in the absence of reversion of ROS levels in the complemented strain, the data needs to be interpreted and extrapolated with caution.

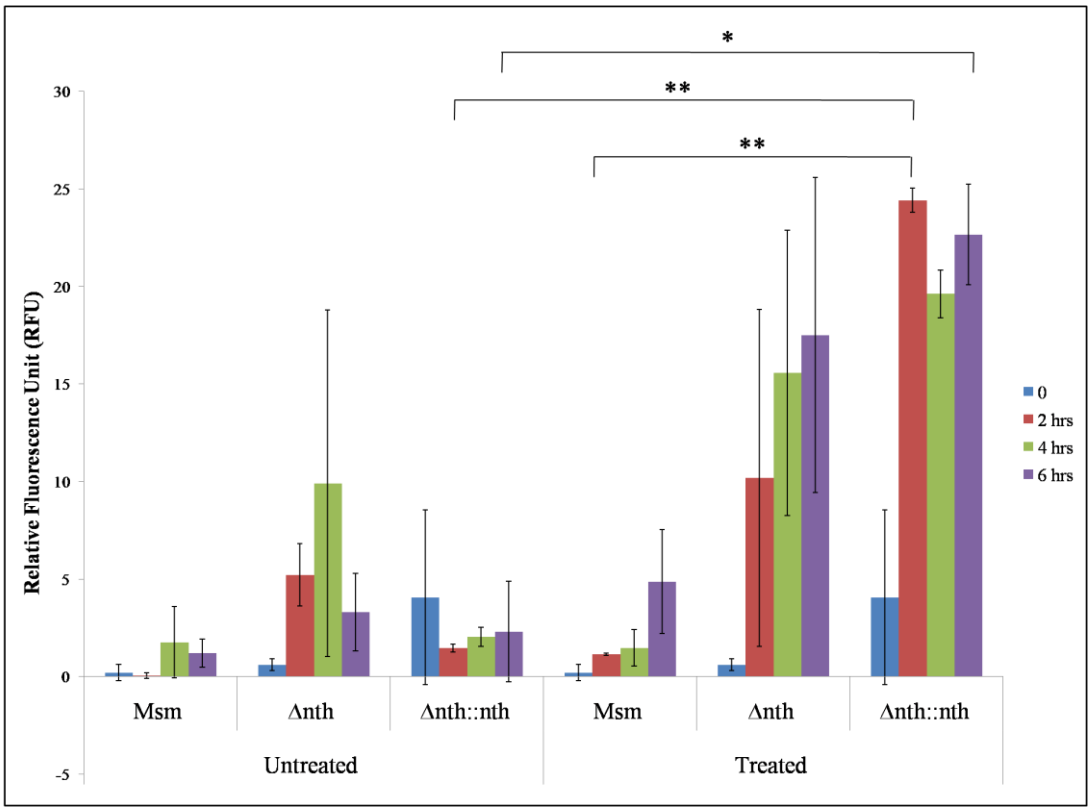



**Fig. 3.29: Intracellular $H_2O_2$ concentration of untreated and treated Msm, single *nth* mutant and its respective complement strains over a 6 hr period.** Values are the means of 3 biological replicates (± SE). P-values were determined by the one way ANOVA where *p<0.05 and **p<0.01 using GraphPad Prism Software.

Lack of a reduced survival phenotype for the single *nth* mutant compared to the parental strain under genotoxic stress (Fig. 3.28) suggests that other enzymes in the Base Excision Repair (BER) pathway are able to deal with the DNA damage caused by genotoxic stress. Previously, in our laboratory much effort was invested to investigate particularly the individual and or combined role(s) of the Fpg/Nei family of DNA glycosylases (Goosens, 2008 MSc; Goosens *et al.*, unpublished) and therefore several combinatorial Fpg/Nei deletion mutants were available which were useful tools to establish whether this family of glycosylases were in any way compensating for the loss of the Nth glycosylase. In Ec the Nei glycosylase can compensate for the loss of the Nth glycosylase under oxidative stress (Jiang *et al*., 1997). To ascertain if the same observation applies to Msm, mutants lacking Nth, Fpg and/or Nei glycosylases were assessed for survival under oxidative stress (Fig.

3.30). All cultures were grown in 7H9 broth to early log phase to which 2.5 mM $H_2O_2$ was added. Thereafter, 100 µl of dilutions from $10^0$ to $10^{-8}$ were spread in duplicate on 7H10 at 2 hrs intervals for a period of 6 hrs to assess the susceptibility of the various *nth* mutants to oxidative damage compared to the wild type Msm strain.

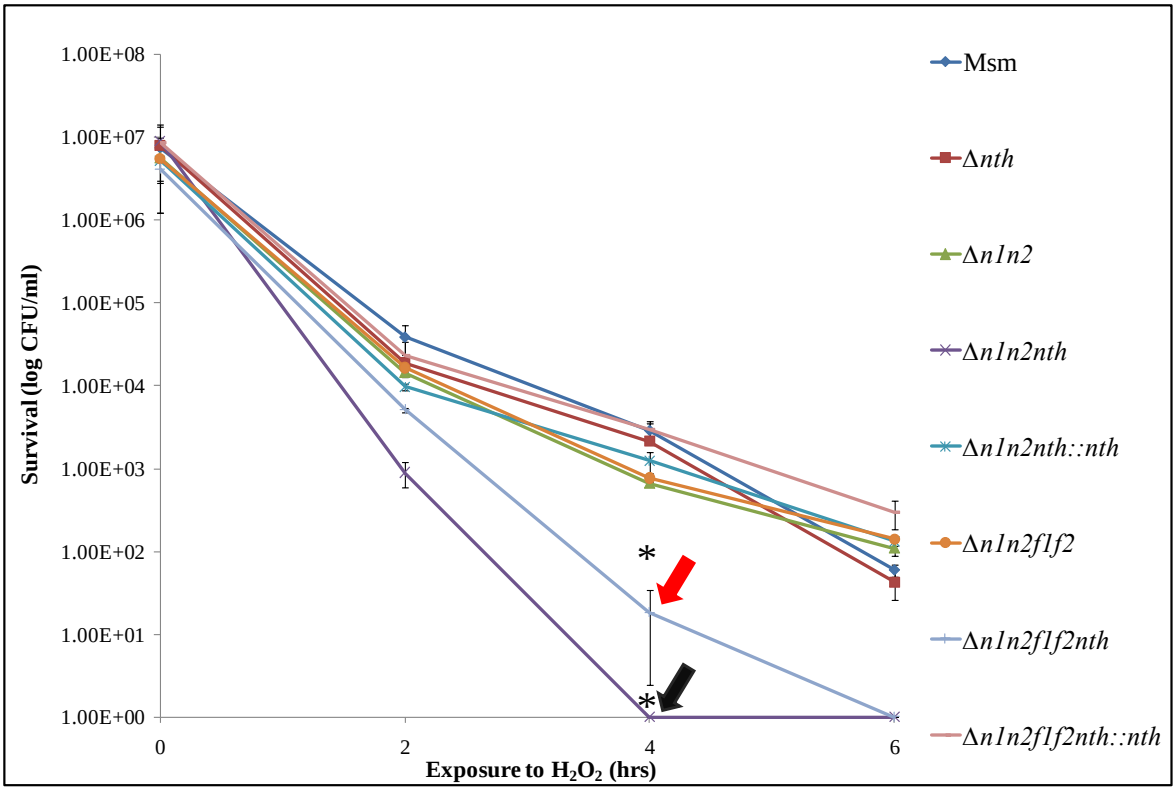



**Fig. 3.30: The survival of early log-phase *nei/fpg/nth* deficient mutants compared to the parental Msm strain when treated with 2.5 mM $H_2O_2$ over a period of 6 hrs**. Data represent averages and SE from 3 independent experiments. One way ANOVA analysis showed a significant difference between Msm and Δ*n1n2nth* and Δ*n1n2f1f2nth* at 4 hrs post $H_2O_2$ treatment (* $p<0.05$). Black and red arrows show survival of the Δ*n1n2nth* and Δ*n1n2f1f1nth* mutant strains at 4 hrs post $H_2O_2$ treatment respectively.

Nei deficient mutants in combination with and without Fpg and Nth glycosylases were assessed for survival under oxidative stress (Fig. 3.30). The parental Msm strain, the single *nth* mutant (Δ*nth*), Nei deficient mutant strains (Δ*n1n2* and Δ*n1n2f1f2*) and the complemented strains (Δ*n1n2nth*::*nth* and Δ*n1n2f1f2nth*::*nth*) responded to oxidative stress in a similar manner (Fig. 3.30). The Nei deficient and Nei/Fpg deficient mutants showed a gradual decline in survival similar to that observed for the parental Msm strain and the

single *nth* deletion mutant when exposed to oxidative stress (Fig. 3.30). The sequential deletion of the Nth glycosylase in the Nei deficient (Δ*n1n2nth*) and Nei/Fpg deficient (Δ*n1n2f1f2nth*) mutants significantly reduced survival by 3-4 and 2-3 log respectively after 4 hours of exposure (indicated by black and red arrows in Fig. 3.30) to 2.5 mM $H_2O_2$ as compared to Msm. To verify the above observation and to assess whether the Fpg glycosylase can compensate for the loss of Nth, Fpg deficient mutants in combination with and without Nei and Nth glycosylases were assessed for survival under oxidative stress (Fig. 3.31).

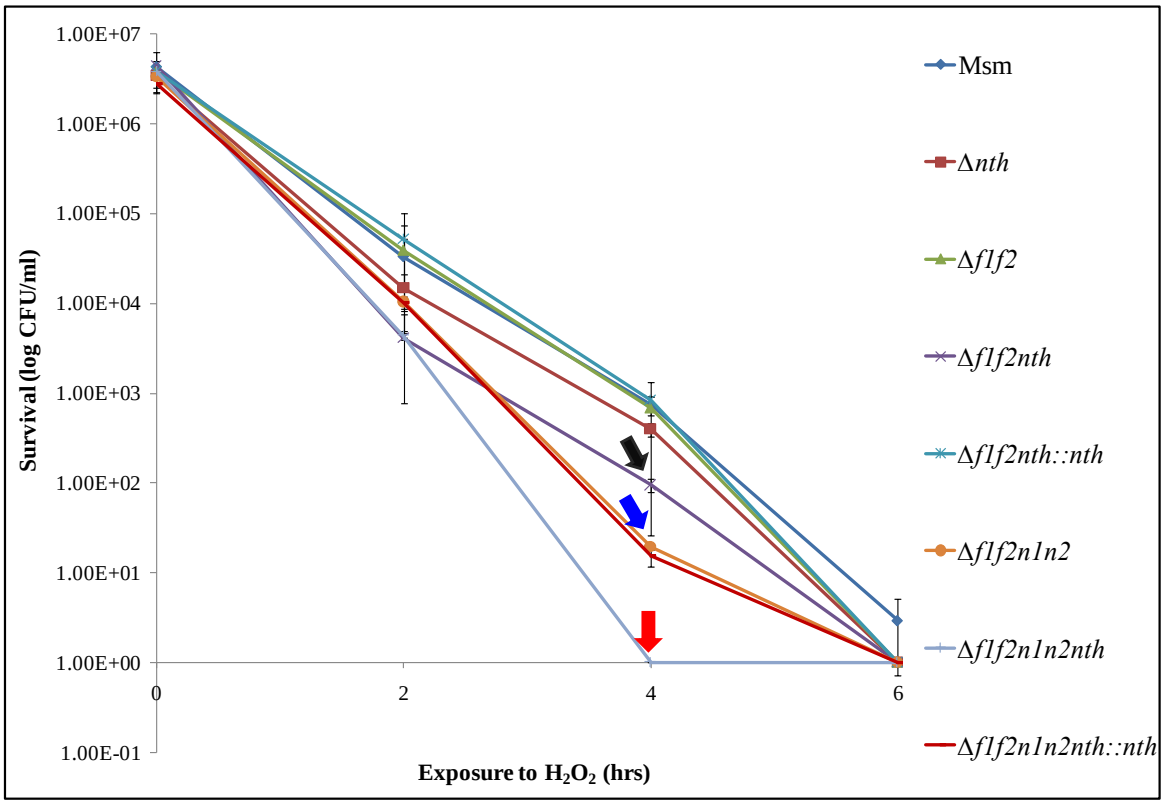



**Fig. 3.31: The survival of early log-phase *fpg/nei/nth* deficient mutants compared to the parental Msm strain when treated with 2.5 mM $H_2O_2$ over a period of 6 hrs**. Data represent averages and SE from 3 independent experiments. One way ANOVA analysis showed no significant differences between strains (P-values >0.1 for all points). Black, blue, and red arrows show survival of the Δ*f1f2nth*, Δ*f1f2n1n2* and Δ *f1f1n1n2nth* mutant strains at 4 hrs post $H_2O_2$ treatment respectively.

The parental Msm strain, the single *nth* mutant (Δ*nth*), Fpg deficient mutant strains (Δ*f1f2* and Δ*f1f2n1n2*) and the complemented strains (Δ*f1f2nth*::*nth* and Δ*f1f2n1n2nth*::*nth*)

responded to oxidative stress in a similar manner, but deletion of the Nth glycosylase in the Fpg deficient mutant background (Δ*f1f2nth*) showed an approximate 1 log reduction in survival compared to the parental strain after 4 hrs of exposure (indicated by the black arrow in Fig. 3.31) to 2.5 mM $H_2O_2$ indicating that the Nei glycosylase may be compensating for the loss of Fpg/Nth glycosylases. The mutant deficient in the entire Fpg/Nei family of DNA glycosylases (Δ*f1f2n1n2*) resulted in a 2 log reduction in survival compared to the parental (indicated by the blue arrow in Fig. 3.31). There was an exaggerated reduction in survival (a further log difference) with the sequential deletion of the Nth glycosylase in this background resulting in a 3 log total reduction in survival (indicated by red arrow in Fig. 3.31) compared to the parental strain after 4 hrs of exposure to 2.5 mM $H_2O_2$, suggesting that the *nth* glycosylase may be playing a greater role than the Fpg/Nei DNA glycosylases in DNA repair under genotoxic conditions. Collectively these data suggests that Nei and Nth most probably function in a similar manner under genotoxic stress conditions and have the ability to compensate for each other as observed by the added killing effect when *nth* is deleted in combination with the *nei* DNA glycosylases.

### 3.5.3 Involvement of *nth* in mutagenesis

As observed in Ec (Blaisdell *et al*., 1999) and *S. typhimurium* (Survarypunya & Stein, 2005), Msm mutants deficient in both the Fpg/Nei and the Nth DNA glycosylases is anticipated to have an increased spontaneous mutation rate compared to the single *nth* mutant. The spontaneous mutation frequencies of the following Nth deficient strains Δ*n1n2nth*, Δ*n1n2f1f2nth* and Δ*f1f2n1n2nth* that were not exposed to UV increased with time as observed in Figs. 3.34A and B, indicating that a mutator phenotype was associated with the deletion of the Nth glycosylase. However, this observation could have been due to a single pre-existing mutant that doubles with each generation also known as “jackpots” as

a mutation frequency determines rifampicin resistance ($rif^R$) mutants within the entire population. To accurately determine the effect of the loss of the Nth glycosylase on spontaneous mutagenesis in Msm, a fluctuation assay (section 2.6.3.1.2) was performed to determine the mutation rates of the various mutant strains. Mutation rates exclude pre-existing mutants and only consider spontaneous $rif^R$ mutants that arise during growth. In addition the spectral analysis of *rpoB* mutations within the $rif^R$ resistant mutants (explained in section 2.6.3) was tested for the parental Msm strain and the various *nth* deficient mutants.

#### 3.5.3.1 Mutation rates to assess spontaneous mutations

Mutation rates are a measure of the probability of a cell gaining a mutation during its lifetime and the Luria-Delbrück fluctuation assay is the most accurate method of measuring mutation rates (Rosche & Foster, 2000) described in section 2.6.4.2.

Mutation rates (μ) are dependent on the distribution of mutations (*m*) in the entire population (*Nt*). A combination of methods MSS-MLE and $P_o$ were used to create the Luria-Delbrück distribution to estimate the *m* value required for the mutation rate calculation. These extrapolations consider the entire data set and are dependent on the *m* and *Nt* values. Therefore, statistical analyses that measure significant difference of the mutation rates in each of the strains and standard error of each of the biological replicates cannot be determined (Rosche & Foster, 2000).Hence, if the mutation rates are drastically dissimilar between strains in every experiment then one can conclude that there is a difference. To investigate the role of Nth during mutagenesis the spontaneous mutation rates to rifampicin of the various *nth* mutants were assessed and compared to the parental strain as indicated below (Table 3.1). The sequential deletion of the Nth glycosylase in combination with the Fpg DNA glycosylases display no increase in mutation rate (Table

3.1 – outlined in blue) which clearly suggest that the Nei homologs are able to compensate for Nth. In contrast, deletion of *nth* together with the *nei* homologs results in a 3-6 fold increase in mutation rate (Table 3.1– outlined in purple) which further confirms that the Fpg homologs do not share functionality with *nth* and *nei*. Interestingly, deletion of *nth* in the two Fpg/Nei deletion mutants lacking all four homologs derived from two different lineages displayed a significant increase in the mutation rate suggesting *nth* plays an anti-mutator role and is able to maintain genome stability (Table 3.1– outlined in orange and red). However, the two Fpg/Nei lineages displayed differential fold increases in mutation rates with the Nei lineage mutants resulting in almost a 2-3 fold greater increase than the Fpg lineage mutants (Table 3.1). These phenotypic differences may be due to other confounded genetic changes in the genome that could have occurred during the homologous recombination process used to sequentially inactivate multiple DNA glycosylases in the same background strain. The two genotypic Fpg/Nei mutant strains will require further analysis such as whole genome sequencing to understand the underlying molecular mechanisms. These data also suggest that there is hierarchy between *fpg*, *nei* and *nth* genes in terms of function and redundancy.

**Table 3.1: Spontaneous rifampicin resistance mutation rates of the various *nth* knockout mutant strains compared to Msm as determined by fluctuation analysis.** Various *nth* deleted mutants are outlined in blue, purple, orange and red for comparison.

| Strain | Mutation rates calculated after each fluctuation assay (X $10^{-10}$) | | | Fold change relative to Msm after each fluctuation assay | | |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 1 | 2 | 3 |
| Msm | 10.5 | 8.5 | 6.3 | 1.0 | 1.0 | 1.0 |
| Δ*nth* | 1.2 | 3.6 | 3.2 | 0.1 | 0.4 | 0.5 |
| Δ*nth::nth* | 1.1 | 4.1 | 2.5 | 0.1 | 0.5 | 0.4 |
| Δ*f1f2* | 1.3 | 1.4 | 4.8 | 0.1 | 0.2 | 0.8 |
| Δ*f1f2nth* | 9.1 | 5.2 | 9.4 | 0.9 | 0.6 | 1.5 |
| Δ*f1f2nth::nth* | 8.2 | 2.2 | 6.0 | 0.8 | 0.3 | 0.9 |
| Δ*n1n2* | 1.1 | 5.3 | 5.1 | 0.1 | 0.6 | 0.8 |
| Δ*n1n2nth* | 41.1 | 32.7 | 41.5 | 3.9 | 3.8 | 6.6 |
| Δ*n1n2nth::nth* | 8.5 | 2.5 | 3.8 | 0.8 | 0.3 | 0.6 |
| Δ*f1f2n1n2* | 18.3 | 5.6 | 2.6 | 1.7 | 0.7 | 0.4 |
| Δ*f1f2n1n2nth* | 37.6 | 24.6 | 21.4 | 3.6 | 2.9 | 3.4 |
| Δ*f1f2n1n2nth::nth* | 7.3 | 4.6 | 2.0 | 0.7 | 0.5 | 0.3 |
| Δ*n1n2f1f2* | 13.5 | 20.9 | 5.9 | 1.3 | 2.4 | 0.9 |
| Δ*n1n2f1f2nth* | 63.2 | 68.4 | 65.3 | 6.0 | 8.0 | 10.4 |
| Δ*n1n2f1f2nth::nth* | 9.3 | 13.9 | 12.2 | 0.9 | 1.6 | 1.9 |

#### 3.5.3.1.1 Mutation spectrum

The diluted inoculum distributed into parallel tubes as described in the fluctuation assay (section 2.6.4.2) minimizes the probability of $rif^R$ colonies (pre-existing rif mutants) originating from a common source (Rosche & Foster, 2000). Therefore, genomic DNA isolated from 10 randomly selected $rif^R$ colonies from the first fluctuation assay for Msm, Fpg/Nei deficient mutants, combinatorial Fpg/Nei/Nth deficient mutants and

complemented strains was PCR amplified across the RRDR region using primers in Table 5.2.1.1 (Appendix 5.2.1) and the amplicon sequenced with primers in Table 5.2.1.2, Appendix 5.2.1 (sequencing data not shown) to identify the types of lesions that resulted in rif resistance. The common amino acids positions where mutations occurred in the RRDR regions are shown in Fig. 3.32 and the resultant amino acid changes which were recorded in these regions are shown in Fig. 3.33.

From Fig. 3.32 more amino acid mutations occur within the RRDR region when *nth* is deleted than in the region outside of the RRDR region of the *rpoB* gene. A higher percentage of mutations occur in the region outside of the RRDR region of the *rpoB* gene in the parental and complemented strains (Fig. 3.32). Surprisingly, wild type Msm showed a 70% C→A and 30% C→T amino acid changes suggesting that under normal culture conditions endogenous ROS oxidizes purines and pyrimidines respectively. The single *nth* and Δ*n1n2f1f2nth* mutants showed a 60% increase in C→T transitions as compared to Msm (Fig. 3.33). The Δ*f1f2nth* mutant showed no increase in C→T transitions while Δ*n1n2nth* mutant had a 10% increase in C→T and 30% increase in G→A transitions compared to Msm (Fig. 3.33) confirming that *nth* and *nei* have compensatory and/or interchangeable roles. The Δ*f1f2n1n2nth* mutant showed a 30% increase in C→T transition compared to Msm (Fig. 3.33). The deletion of the Fpg and Nei glycosylases (Δ*f1f2*, Δ*n1n2*, Δ*n1n2nth* and Δ*n1n2f1f2*) resulted in an increase in G→T amino acid changes irrespective of the absence or presence of the *nth* gene in comparison to Msm (Fig. 3.33). While G→C transversions (10 -20%) was observed in Δ*f1f2*, Δ*n1n2* and Δ*n1n2nth* mutant strains the frequency increased to a 60% in Δ*f1f2n1n2* suggesting that Fpg and Nei probably work in concert and have a close supportive association during DNA repair (Fig. 3.33). However, the same phenotype was not observed for the Δ*n1n2f1f2* lineage mutant indicating that the

order in which the Fpg/Nei genes were deleted influenced the integrity of the organism and the molecular mechanisms involved needs further investigation. Overall the overwhelming amino acid alteration observed in all mutants deficient in *nth* was C→T transitions which was the expected mutation (Figs. 3.32 and 3.33). This data suggests that *nth* has an anti-mutator role by primarily repairing oxidized cytosines induced under oxidative stress conditions to maintain genome integrity in Msm.

A

| | | | | | | | 513 | | | | | | | | | 522 | | | 526 | | | | | 531 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 507 – | ggc | acc | agc | cag | ctg | tgc | cag | ttc | atg | gac | cag | Aac | aac | ccg | ctg | tcg | ggt | ctg | acc | cac | aag | cgt | cgt | ctt | tcg | gcg | ctg | – 533 |
| | Gly | Thr | Ser | Gln | Leu | Ser | Gln | Phe | Met | Asp | Gln | Asn | Asn | Pro | Leu | Ser | Gly | Leu | Thr | His | Lys | Arg | Arg | Leu | Ser | Ala | Leu | |

B

| | | | Frequency of mutations (%) | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Mutation | Alteration | Msm | *Δnth* | *Δnth ::nth* | *Δf1f2* | *Δf1f2 nth* | *Δf1f2nth ::nth* | *Δn1n2* | *Δn1n2 nth* | *Δn1n2nth ::nth* | *Δf1f2n1 n2* | *Δf1f2n1n2 nth* | *Δf1f2n1n2 nth::nth* | *Δn1n2f1 f2* | *Δn1n2f1f2 nth* | *Δn1n2f1f2 nth::nth* |
| cag→tag | Gln513→Ter | | | | | 10 | | | | | | | | | | | |
| cag→ctg | Gln513→Leu | | | | | 30 | | | | | | | | | | | |
| cag→ccg | Gln513→Pro | | | | 20 | | | | | | | | | | | | |
| cag→cat | Gln513→His | | | | 20 | | | | | | | | | | | | 10 |
| cag→caa | Gln513→Gln | | | | | | | 20 | 30 | | | | | | | | |
| cag→aag | Gln513→Lys | 70 | | 20 | | 10 | | 20 | | | 10 | 20 | 10 | 10 | 10 | 10 | |
| tcg→tcc | Ser522→Ser | | | | | | | 10 | 20 | 10 | | | | | | | |
| tcg→tgg | Ser522→Trp | | | | | 10 | | 10 | | | | | | 10 | | | |
| cac→cgc | His526→Arg | | 10 | 60 | | | | | | | 10 | 20 | 10 | 30 | | | |
| cac→tac | His526→Tyr | 30 | 20 | | | | | | 10 | 10 | | | 40 | 10 | 90 | 20 | |
| cac→ccc | His526→Pro | | | | | | 10 | 10 | 10 | 10 | | | 10 | | | | |
| cac→caa | His526→Gln | | | | | | | | | | | | | | | | 10 |
| tcg→ttg | Ser531→Leu | | 70 | 20 | 20 | | | 10 | 10 | 10 | 20 | 40 | 20 | 20 | | | |
| tcg→tgg | Ser531→Trp | | | | | | 20 | 10 | | | | | | | | | 20 |
| tcg→tct | Ser531→Ser | | | | | 10 | | 10 | | | | | | | | | |
| Unknown | | | | | 40 | 30 | 70 | | 20 | 60 | 60 | 20 | 10 | 20 | | 30 | |

**Fig 3.32: Amino acid changes in the RRDR region of the *rpoB* gene in rif resistant mutants isolated from Msm, Fpg/Nei deficient mutants, *nth* deletion mutants and complemented strains. A.** The codon and corresponding amino acid sequence of the RRDR region. Areas where mutations occurred are highlighted in orange boxes. **B.** Table indicating the specific amino acid alterations and percentage frequency of the mutations for the *nth* deficient mutants compared to Msm. Only 10 $rif^R$ mutants were isolated from each of the mutant, parental and Msm strains. The mutations which did not map to the RRDR region are labelled as unknown.

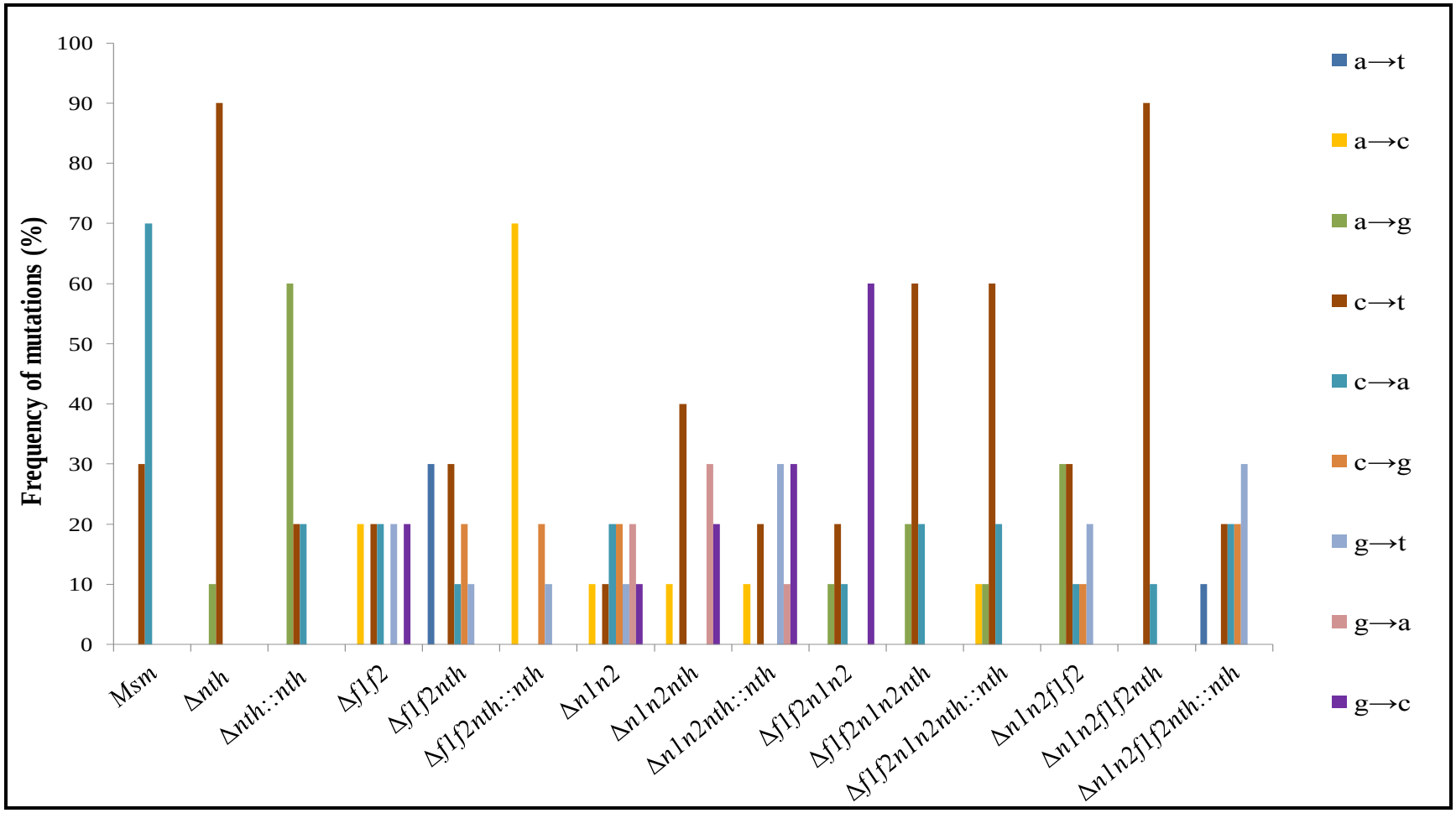



**Fig. 3.33: The percentage frequency of amino acid changes identified within the RRDR region of the various *nth* deletion mutants as compared to Msm, Fpg/Nei deficient mutants and complemented strains.**

### 3.5.3.2 UV induced mutagenesis

To further ascertain the role of the Nth DNA glycosylase during DNA repair an alternate mutagen in the form of UV was used to assess if Fpg and/or Nei DNA glycosylases have a compensatory function for the Nth DNA glycosylase (Fig. 3.34). The various *nth* deficient mutants in the absence and presence of the Fpg/Nei glycosylases were grown in 7H9 to early log phase and split into two equal cultures. One aliquot of cells was exposed to 250 $mJ/cm^2$ UV, whilst the other set served as a negative control. Cell viability was assessed at 3 hr and 6 hr intervals to obtain CFU/ml. In addition, 1 ml of culture for each of the strains was spread in duplicate on 7H10 rif (200 µg/ml) to isolate spontaneous rif resistant mutants that resulted from UV induced mutagenesis. The mutation frequency for each of the strains was calculated as the ratio of the number of spontaneous rifampicin resistant mutants to viable cell counts for the various treated and untreated *nth* deficient mutant strains and compared to the Msm under the same conditions (Fig. 3.34).

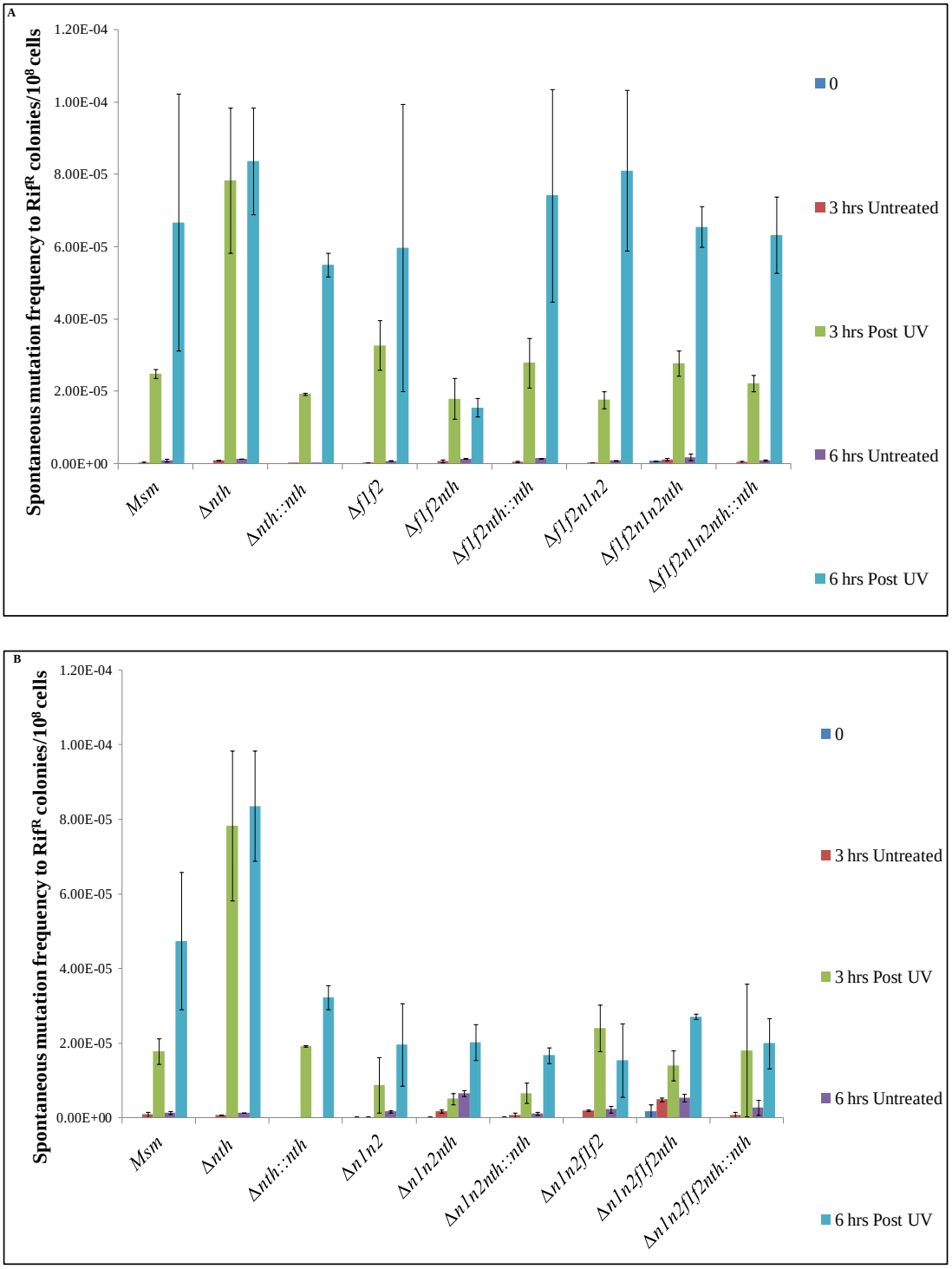



**Fig. 3.34: UV induced mutation frequency of combinatorial Fpg/Nei/Nth deficient mutants. A.** UV induced mutation frequency of *fpg/nei/nth* deficient mutants to rif$^R$. Untreated cultures showed statistical difference to the treated cultures $p<0.00009$ at 3 hrs and $p<0.0001$ at 6 hrs (paired t test). **B.** UV induced mutation frequency of *nei/fpg/nth* deficient mutants to rif$^R$. Untreated cultures showed statistical difference to the treated cultures $p<0.05$ at 3 hrs and $p<0.001$ at 6 hrs (paired t test). Data represent averages and SE from 3 independent experiments.

Deletion of only the Nth DNA glycosylase resulted in an exaggerated mutation frequency 3 and 6 hrs post UV (250 mJ/cm$^2$) exposure as compared to Msm, the respective complemented strains and the Fpg/Nei/Nth and Nei/Fpg/Nth combinatorial deficient mutants (Fig. 3.34). The single *nth* deletion mutant showed a significant increase in rif resistant mutants 3 hrs post UV exposure compared to all the combinatorial mutants irrespective of the order in which the genes were inactivated (p values were $p<0.001$; Figs. 3.34A and B). In Fig. 3.34A the Fpg/Nth deficient mutant (Δ*f1f2nth*) was not sensitive to UV exposure as the mutation frequency was low suggesting that the intact Nei DNA glycosylase in this mutant strain could compensate for the loss of Fpg and Nth DNA glycosylases. Deletion of the entire Fpg/Nei family of DNA glycosylases (Δ*f1f2n1n2*) showed an increase in mutation frequency at 6 hrs post UV exposure (Fig. 3.34A). Surprisingly, the subsequent deletion of the Nth glycosylase in this genotypic background (Δ*f1f2n1n2nth*) did not have an additive effect but rather displayed mutagenic frequencies comparable to Msm at 6 hrs post UV treatment (Fig. 3.34A), indicating that other DNA repair mechanisms can possibly compensate for the loss of *fpg*, *nei* and *nth*. Whereas, in Fig. 3.34B when the Nei glycosylases were deleted before the Fpg glycosylases all the combinatorial mutant strains and their respective complemented strains (Δ*n1n2*, Δ*n1n2nth,* Δ*n1n2nth*::*nth*, Δ*n1n2f1f2* and Δ*n1n2f1f2nth*::*nth)* showed low mutation frequencies after 6 hrs of UV exposure relative to Msm and the single *nth* deletion mutant (Fig. 3.34B). From these data it is clear that the two Fpg/Nei deficient mutants have different phenotypic characteristics that may be linked to genotypic changes currently unknown and requires further investigations.

Similarly, the fluctuation analysis assay in conjunction with the spontaneous mutation frequencies of the Nei lineage of mutants not exposed to UV displayed a more dramatic rif

resistant phenotype than the respective Fpg lineage of mutants, further corroborating that the two Fpg/Nei deficient mutant backgrounds are genotypically different (Fig. 3.34, Table 3.1). The deletion of the Fpg/Nei family of DNA glycosylases demonstrated a marginal increase in spontaneous mutation frequencies of Fpg/Nei deficient mutants not exposed to UV compared to wild type Msm (Fig. 3.34), in contrast to the fluctuation analysis that displayed a decreased mutation rate of these mutants relative to Msm (Table 3.1). Analysis of both mutation frequencies and rates of the various mutants indicate that combinatorial mutants deficient in *fpg*, *nei* and *nth* genes have elevated mutator phenotypes relative to the parental strain Msm (Fig. 3.34, Table 3.1 – outlined in blue, purple, orange and red), the Fpg/Nei glycosylase deficient mutants and the complemented strains demonstrating the functional redundancy of these glycosylases. Interestingly, the progressive deletion of *nth* in the two lineages Fpg and Nei displayed differential increases in mutation frequencies and rates. Furthermore, spectral analysis of rif resistant mutants showed a variety of lesions with the loss of Fpg/Nei glycosylases as compared to Msm. This frequency of C→T transitions was dramatically increased with subsequent deletion of *nth* in Fpg/Nei glycosylase deficient mutants. This implies that *nth* is primarily responsible for the repair of oxidized pyrimidine lesions. Therefore, both methods used to assess the anti-mutator properties of *nth* confirmed that *nth* together with *fpg* and *nei* genes contributes to mycobacterial genome maintenance.

# 4. Discussion and Conclusion

The host immune system generates reactive nitrogen and oxygen species (RNS and ROS) to destroy the lipids, protein and DNA of invading alveolar Mtb bacilli during an infection. However, Mtb is equipped with DNA repair systems such as base excision repair (BER) that employs DNA glycosylases to identify, remove and correct the DNA base modification or damage caused by oxidative stress (Krokan *et al*., 1997; Friedberg *et al*., 1995; Kurthkoti & Varshney 2011; Kurthkoti & Varshney, 2012). In this manner these glycosylases initiate repair and maintain the genomic integrity of the organism. Previously, the effect of the Fpg/Nei glycosylases deficiency in mycobacteria was shown to have no differences in survival of the mutant strains as compared to the wild type Msm strain (Goosens, 2008 MSc). In *Escherichia coli* (Ec) and *Salmonella typhimurium* (*S. typhimurium*) a mutator and reduced survival phenotype was observed when *fpg*, *nei* and *nth* were deleted (Blaisdell *et al*., 1999; Suvarnapunya *et al*., 2003; Suvarnapunya & Stein, 2005), indicating that the Nth glycosylase of the BER pathway may be important for cellular survival.

Bioinformatic analysis showed the presence of *nth* in several related microorganisms suggesting that the *nth* gene was conserved through evolution. Organisms with a small genome retain only genes required for survival and examples of such organisms are *Helicobacter pylori* (*H. pylori*) and *Mycobacterium leprae* (*M. leprae*) both of which have retained the *nth* gene indicating that this gene may be integral for survival (Tomb *et al*., 1997; Cole *et al*., 2001). Interestingly the *nth* gene is also present in several intracellular pathogens that are exposed to stressful conditions such as *H. pylori*, *S. typhimurium*, *Corynebacterium jeikeium* (*C. jeikeium*), *Mycobacterium avium paratuberculosis* (*M. avium paratuberculosis*), *M. leprae*, *Mycobacterium bovis* (*M. bovis*) and *Mycobacterium*

*tuberculosis* (Mtb), indicating that *nth* is potentially important for DNA repair and the continued survival of the organism within the host.

The Nth glycosylase of Ec was shown to have the following structural characteristics: the helix hairpin helix loop (HhH) containing the lysine (Lys) 120 and aspartate 138 (Asp) active sites, the glycine-proline-aspartate (GPD) domain and the four iron and four sulfur cluster loop (FCL) that classify it as a DNA glycosylase with AP lyase activity capable of identifying and removing damaged or modified base lesions (Cunningham *et al*., 1989; Kuo *et al*., 1992a; b; Thayer *et al*., 1995; Cunningham, 1997). Bioinformatic analysis in this study showed that the aforementioned structural features of the Nth protein were all conserved within mycobacteria. In particular, the percentage identity of the Nth protein of Mtb and Msm was high (i.e. 89%) and moreover, the *nth* gene in both organisms is in the same genomic context. This indicates that Msm which is non pathogenic and faster growing than Mtb lends itself as a good model organism to elucidate the role of the Nth glycosylase in BER to better understand the complexities of the DNA repair systems of Mtb.

Despite the low similarity (i.e. 39%) of Nth proteins observed between Ec and Msm the structural domains and active sites of the Nth protein which are integral for the functioning of the Nth glycosylase in DNA repair are conserved. Both the Msm Nth and the Ec Nth have the structural machinery to function as Nth glycosylases. Therefore, to assess whether the Msm *nth* encoding the Nth glycosylase has similar functions to the Ec Nth glycosylase, the Ec *nth* deficient mutant (EcΔ*nth*) was complemented with the Msm *nth* to generate the EcΔ*nth*::Msm*nth* strain that was assessed for survival under oxidative stress conditions and for mutagenesis to determine if the mycobacterial *nth* gene encoding the Nth glycosylase could compensate for the Nth deficiency in Ec.

The survival kinetics of the Ec strains with and without *nth* showed equivalent reduced survival phenotype as the wild type Ec strain under hydrogen peroxide ($H_2O_2$) induced oxidative stress conditions, suggesting that other DNA repair glycosylases are compensating for the loss of *nth* (Blaisdell *et al*., 1999; Cunningham & Weiss, 1985; Jiang *et al*., 1997; Saito *et al*., 1997). Since, this assay did not establish that mycobacterial Nth could restore the Nth deficiency in Ec, the mutation frequencies of the Ec strains was assessed which clearly showed that the Msm *nth* gene was capable of significantly reducing the spontaneous mutation frequency in the Ec*Δnth* strain indicating an overlapping role of Msm and Ec *nth* gene. Thus, these data clearly supports that the Msm Nth glycosylase is structurally and functionally suited as a DNA glycosylase in DNA repair. To determine the function of the *nth* gene encoding the Nth glycosylase in Msm, a panel of *nth* deficient mutants in wild type Msm and the Fpg/Nei deficient mutants were generated by homologous recombination and assessed for differences in growth, survival under oxidative stress conditions and increased mutagenesis compared to the parental Msm, the Fpg/Nei deficient mutants and complemented strains.

The various *nth* deleted mutants were firstly assessed for changes in growth under normal culture conditions. The combinatorial *nth* deficient mutants displayed equivalent growth kinetic as the parental strain, Fpg/Nei deficient mutants and complemented strains, indicating that the Nth glycosylase is not necessary for growth under normal culture. Therefore, the various *nth* deficient mutants were exposed to endogenous DNA damaging agents such as $H_2O_2$ and UV irradiation treatments. Due to the unstable nature of $H_2O_2$ upon exposure to light it was of concern that it may be degraded before it enters the cell, resulting in minimal phenotypic differences between the *nth* mutant and the Msm parental strain. To reduce light exposure all flasks were covered during experimental procedures.

Furthermore, to determine whether the 2.5 mM concentration of $H_2O_2$ used in this study (previously optimised by Goosens, 2008 MSc) was sufficient to generate genotoxic radicals that were not overcome by the mycobacterial detoxification system, the intracellular $H_2O_2$ concentration of Msm was measured and compared to the untreated cells. The elevated levels of intracellular $H_2O_2$ in treated Msm as compared to untreated Msm indicated that considerable levels of $H_2O_2$ and its derivatives were available to potentially cause DNA damage and the endogenous detoxification systems were unable to control the amount of ROS generated.

The single *nth* deletion mutant (Δ*nth*) showed no differences in survival under genotoxic conditions compared to its respective parental and complemented strain which was in concordance with the observed phenotype for Ec *nth* deficient mutants (Cunningham & Weiss *et al*., 1985; Jiang *et al*., 1997; Saito *et al*., 1997). Therefore, this negligible difference in survival of the Ec and Msm single *nth* mutants compared to wild type can be attributed to the ability of Nei to compensate for the loss of Nth under genotoxic conditions. It has also been reported that deletion of the Nei glycosylase also showed no reduction in survival under genotoxic conditions (Goosens, 2008 MSc.). However, in this study when both the *nei* and *nth* genes were deleted (Δ*n1n2nth*) the combinatorial mutant showed an exaggerated 3 log reduction in growth under oxidative stress conditions compared to the single *nth* mutant and the double *nei* glycosylases (Δ*n1n2*). Similarly, the same hypersensitivity was observed for Ec (Jiang *et al*., 1997; Saito *et al*., 1997) and *S. typhimurium* (Suvarnapunya *et al*., 2003; Suvarnapunya & Stein, 2005) mutants that were deficient in both *nei* and *nth* genes but was absent in the respective single mutants. These data clearly indicate that *nei* and *nth* share similarity and are able to compensate for each other.

Interestingly, deletion of the Fpg glycosylase together with Nth (Δ*f1f2nth*) resulted in only a 1 log difference in survival compared to Msm and the single *nth* mutant. However, when both Fpg and Nei glycosylases were deleted (Δ*f1f2n1n2*) survival was reduced by 2 log compared to Msm and the single *nth* mutant suggesting that the Nth glycosylase was able to compensate for both the loss of Fpg and Nei glycosylases. The data also confirms that the Fpg glycosylase is unable to compensate for the loss of Nei or Nth glycosylases and is unable to recognize and repair the same damaged substrates as *nth* and *nei*. However, deletion of *nth* in the Fpg/Nei deficient mutants Δ*n1n2f1f2nth* and Δ*f1f2n1n2nth* mutant showed an exaggerated reduction in survival further confirming the interchangeable role of Nei and Nth. Surprisingly the two Fpg/Nei deficient mutants display variable survival rates suggesting that the two mutants are genotypically different. These genotypic differences could have been introduced during the sequential deletion of the Fpg/Nei homologs to generate the various combinatorial mutants. At present the molecular mechanisms underlying these differences are unclear and require further investigation such as whole genome sequencing to identify any genomic differences in the two lineages of mutants.

Taken together these data clearly showed that the Nth glycosylase of Msm can compensate for the loss of the Nei glycosylase, implying that Nth and Nei glycosylases are able to recognise similar DNA lesions (Jiang *et al*., 1997; Saito *et al*., 1997) as generated by $H_2O_2$. This is the first reported study in mycobacteria that shows functional overlap between two DNA repair glycosylases under oxidative stress even though they do not share structural similarity. In support of this result Guo *et al*., (2010) showed biochemically that both Nth and Nei excise similar lesions.

UV irradiated DNA was used as a substrate to identify and extract the Nth protein in Ec (Radman, 1976; Gates & Linn, 1977), demonstrating that this glycosylase can repair DNA

lesions generated by UV. To determine the role of the Nth glycosylase in the repair of UV induced lesions, various *nth* deficient mutants were irradiated with UV (250 mJ/cm$^2$). Of all the mutants assessed for DNA damage induced mutagenesis the single Δ*nth* mutant showed the highest mutation frequencies at 3 hrs post UV treatment. Weiss & Duker (1986) found that UV irradiated DNA predominantly generated modified pyrimidine bases such as thymine glycols, pyrimidine dimers, urea, uracils and cytosines. Both uracil and cytosine lesions are further oxidised to form premutagenic lesions such as: 5-hydroxycytosines, uracil glycols and 5-hydroxyuracil (Purmal *et al*., 1994). The Nth glycosylase in Ec was found to nick cytosine lesions generated by UV (Weiss & Duker, 1986) thus reducing the formation of premutagenic lesions. UV irradiation also generated to a lesser extent modified purine premutagenic lesions (Weiss & Duker, 1986) that were presumably repaired by the Fpg glycosylase as in this study the Δ*f1f2* mutant showed an increase in mutation frequencies at 3 hrs post UV exposure.

We have demonstrated that the Nei glycosylase was able to compensate for the loss of the Fpg glycosylase as the mutation frequency of the Δ*f1f2nth* mutant was kept lower than Msm at 3 and 6 hrs post UV exposure suggesting that the Nei glycosylase can identify modified purine premutagenic lesions. However, neither Fpg nor Nei glycosylases were able to reduce the mutation frequency of the single *nth* mutant post UV treatment. The Fpg and Nei glycosylases can identify and repair to a lesser extent 5-hydroxycytosines and 5-hydroxyuracil premutagenic lesions in the absence of *nth* (Hatahet *et al*., 1994; Guo *et al*., 2010). Therefore, the Nth glycosylase is primarily responsible for repairing UV induced DNA lesions namely: thymine glycols, urea, uracil, 5-hydroxycytosines and 5-hydroxyuracil (Armel *et al*., 1977; Gates & Linn, 1977; Demple & Linn, 1980; Katcher &

Wallace, 1983; Breimer & Lindahl, 1984; Dizdaroglu *et al*., 1993; Dizdaroglu *et al*., 2000; Dizdaroglu, 2003; Guo *et al*., 2010).

Surprisingly, the escalated mutation frequency of Δ*nth* at 3 hrs after UV exposure was not doubled at 6 hrs post UV exposure and the mutation frequency of the Fpg/Nei/Nth (Δ*f1f2n1n2nth*) mutant did not show a dramatic increase as expected but resembled that of wild type, suggesting that other DNA repair systems can repair UV photoproducts. It is well known that the SOS response is induced upon UV exposure upregulating DNA repair genes belonging to nucleotide excision repair (NER) and homologous recombination (Caldeira de Araujo & Favre, 1986; Boshoff *et al*., 2003; Krishna *et al*., 2007; Janion, 2008). NER mainly recognises and repairs pyrimidine dimers but can also identify thymine glycols in the absence of *nth* (Kow *et al*., 1989). Pyrimidine dimers and thymine glycols are not premutagenic (Ide *et al*., 1985; Hayes *et al*., 1988) but are known to block DNA synthesis such as replication, transcription and translation resulting in a lethal outcome (McNulty *et al*., 1998; Ide *et al*., 1985). In Mtb NER is essential and aids in the repair of UV irradiated DNA with or without *fpg* and *nei* to promote cellular survival (Sinha *et al*., 2007; Boshoff *et al*., 2003). Mutants with the Nei glycosylases deleted first displayed low mutation frequencies as compared to Msm and the single *nth* mutant (Δ*nth*), suggesting that in addition to SOS induced repair systems another repair mechanism is at play. The *nei* gene in Msm, MSMEG_1756 (Goosens, 2008 MSc) and Mtb, Rv3297 (Guo *et al*., 2010) are co-transcribed with the *lhr* gene encoding an ATP dependent DNA helicase. In mycobacteria this gene is induced by UV irradiation and in the macrophage environment (Boshoff *et al*., 2003; Schnappinger *et al*., 2003). Therefore, we assume that UV exposure may have induced the *lhr* gene to stimulate other DNA repair pathways (Krishna *et al*.,

2007; Janion, 2008) that ultimately reduce premutagenic lesions resulting in lowered mutation frequencies.

The added deletion of *nth* in the following mutant strains: Δ*n1n2*, Δ*n1n2f1f2* and Δ*f1f2n1n2* which were not exposed to UV irradiation showed increased spontaneous mutation frequencies to rif as compared to untreated Msm, Fpg/Nei deficient mutants and complemented strains. Boshoff *et al*., (2004) reported that the *nth* gene was upregulated when there was inhibition of DNA synthesis not associated with DNA damage, emphasizing that *nth* is integral for the repair of both spontaneous and induced DNA damage.

Since a mutation frequency considers all $rif^R$ mutants present within the entire population this method does not consider that mutations can emerge early and continue doubling during growth. Therefore, a single mutational event can be amplified which is known as a "jackpot" resulting in an exaggerated mutation frequency. To verify the results of the spontaneous mutation frequency observed in the UV assay that imply that the *nth* gene does contribute to genomic maintenance a more accurate method of determining mutagenesis, the fluctuation assay which measures mutation rates was applied to the various DNA glycosylase deficient mutants.

A mutation rate uses the Luria-Delbrück fluctuation analysis to determine the mutational events that occur within a cell and excludes pre-existing mutants that result in "jackpots" (Rosche & Foster, 2000). In contrast to the mutation frequency of *nth*, *f1f2* and *n1n2* deficient mutants which was marginally more than Msm, the fold changes in the mutation rates for these mutants was less than the wild type Msm strain suggesting that mutation rate determination is a more stringent method of assessing mutagenesis compared to mutation frequencies (Rosche & Foster, 2000). In addition the low mutator phenotype

associated with *nth* (Blaisdell *et al*., 1999; Jiang *et al*., 1997), *f1f2* and *n1n2* (Goosens 2008, MSc; Blaisdell *et al*., 1999; Jiang *et al*., 1997) strongly corroborates the functional redundancy associated with these glycosylases.

Minor mutator phenotypes were associated with the loss of Fpg/Nei function as these mutants displayed a marginal increase in mutation rates relative to Msm. However, the subsequent deletion of *nth* in combination with Fpg/Nei showed a dramatic 3-6 fold increase in mutator phenotypes relative to Msm. Collectively, our data using two different methods for measuring DNA damage unreservedly showed that the deletion of the *nth* gene in combination with *fpg* and *nei* genes contributes to mutagenesis and the Nth glycosylase has anti-mutator properties which assists in microbial genome maintenance (Blaisdell *et al*., 1999; Jiang *et al*., 1997; Guo *et al*., 2010).

To gain further insight into the impact of the loss of *nth* in mutagenesis we identified the DNA alterations leading to $rif^R$. The *rpoB* gene of $rif^R$ colonies was amplified across the RRDR region and the changes analysed and compared with previously characterised mapped modifications (Ramaswamy *et al*., 1998; Lee *et al*., 2005; McCammon *et al*., 2005; Sheng *et al*., 2008). In Msm C→A mutations occurred with the highest frequency followed by C→T amino acid changes indicating that under normal culture conditions purine bases are more susceptible to endogenous ROS than pyrimidine bases. It has been reported in mycobacteria that the Fpg glycosylases do not repair or recognise 8oxoG lesions paired with A unlike Ec and also repairs to a lesser extent 8oxoG lesions paired with C which often results in an increase in C→A (or G→T) mutations (Michaels & Miller, 1992; Michaels *et al*., 1992; Grollman & Moriya, 1993; Jain *et al*., 2007; Guo *et al*., 2010). However, in the absence of Fpg/Nei glycosylases a low C→A (or G→T) mutation frequency was observed which we attribute to the repair action of the MutY DNA

glycosylase that is part of the GO system and is able to remove the A, G and T but not C paired with 8oxoG lesions (Kurthkoti *et al*., 2010). Following replication the 8oxoG lesions are paired with G resulting in A→G mutations which were apparent in mutants deficient in Fpg/Nei glycosylases and in the single *nth* complemented strain, indicating that Fpg/Nei glycosylases efficiently repair 8oxoG lesions (Blaisdell *et al*., 1999) and 8oxoG lesions paired with G (Guo *et al*., 2010; Jain *et al*., 2007).

The Δ*f1f2n1n2* mutant showed a 60% frequency of G→C transversions that was also present in Δ*f1f2* and Δ*n1n2* mutants. Likewise, Jain *et al*., (2007) reported the production of G→C transversions when wild type Msm was exposed to $H_2O_2$. Further oxidation of 8oxoG produces spiroiminodihydantoin and guanidinohydantoin lesions which pair with A or T forming G→C or G→T transversions (Duarte *et al*., 1999; Kornyushyna *et al*., 2002). Therefore, Fpg/Nei glycosylases repair derivatives of 8oxoG, i.e. spiroiminodihydantoin and guanidinohydantoin lesions as previously noted by Hailer *et al*., (2005), Guo *et al*., (2010), Jain *et al*., (2007) and Leipold *et al*., (2000). The subsequent deletion of *nth* in Msm and Fpg/Nei deficient mutant backgrounds resulted in an overall increase in C→T or G→A transitions indicating that damaged pyrimidine bases and oxidised cytosines are mainly recognised by the Nth glycosylase (Blaisdell *et al*., 1999; Jiang *et al*., 1997; Guo *et al*., 2010).

Recently, Boal *et al*., (2009) through electrochemical methods showed that the previously classified redox inactive four iron and four sulfur cluster loop (FCL) of Nth was in fact redox active suggesting a secondary role of the structural motif of Nth to combat oxidative stress in addition to DNA repair. In light of this the concentration of intracellular ROS was measured in Msm, the single *nth* mutant (Δ*nth*) and its respective complemented strain (Δ*nth*::*nth*) using a fluorescent probe that binds $H_2O_2$ and its generated radicals.

Treated wild type (Msm) has detoxifying mechanisms that reduce and maintain ROS levels to a minimum unlike the treated Δ*nth* mutant. Detoxification systems in the Δ*nth* mutant exposed to $H_2O_2$ was not robust enough to reduce exogenous ROS but was capable of reducing endogenous ROS as shown by the untreated Δ*nth* mutant. However, the addition of a functional *nth* gene into the mutant strain did not reduce ROS levels. This result was surprising as the expression data showed comparable expression for both the complemented *nth* mutant (Δ*nth*::*nth*) and Msm. To date the FCL motif of Nth has not been implicated in redox homeostasis, but these data implies that Nth can potentially detect ROS and may be part of the detoxification system that maintains the redox balance of Msm.

Iron-sulfur clusters are involved with sensing of iron and oxygen which are elements fundamental for survival of Mtb (Rouault & Klausner, 1996). Mycobacteria has an abundance of proteins with iron-sulfur clusters as they are involved with pyrimidine biosynthesis, the electron transport chain, the tricarboxylic acid cycle and required for KatG activity for the detoxification of ROS (McCready & Ratledge, 1978; Rouault & Klausner, 1996; De Voss *et al*., 1999). One of the oxygen sensing mechanisms of Mtb is the WhiB3 protein that consists of a four iron and four sulfur cluster which is associated with sensing increased levels of nitric oxide in addition to oxygen (Singh *et al*., 2007). This protein uses the iron-sulfur cluster to create an environment (reducing or oxidising) that facilitates the transport of fatty acids and lipids thus regulating the polyketide synthesis and fatty acid metabolism (Singh *et al*., 2009). The WhiB3 protein facilitates redox homeostasis, antibiotic resistance virulence and persistence of bacilli (Morris *et al*., 2005; Singh *et al*., 2007; Singh *et al*., 2009). Thus, other possible functions that may be associated with Nth include redox sensing of ROS or oxygen that stimulates mechanisms

for detoxification and persistence of bacilli in addition to DNA repair. Hence, to determine this reverse transcription polymerase chain reaction (RTPCR) can be performed to monitor changing expression levels of *nth* in response to stresses such as: hypoxia, low pH, starvation, ROS and RNS that governs intracellular environment of the activated macrophage or granuloma.

In conclusion, these results confirm that the Msm *nth* gene encoding the Nth glycosylase is a DNA glycosylase that is involved in DNA repair under normal and DNA damaging conditions. The progressive loss of *nth* in mutants deficient in *fpg* and *nei* showed reduced survival under oxidative conditions suggesting a collective role of these DNA glycosylases in genome maintenance as part of a robust DNA repair system. In the absence of *fpg*, *nei* and *nth* another DNA repair system possibly NER was able to repair UV induced DNA damage. The functional redundancy between DNA glycosylases and between DNA repairs systems, contributes to the understanding of mechanisms involved in ensuring Mtb survival within the hostile host environment. This study has provided the basis for investigating the interaction of these various DNA glycosylases in Mtb to better understand mycobacterial pathogenesis and long term survival in the host.

# 5. Appendices

## 5.1 Culturing of Bacteria

### 5.1.1 Culture media

Middlebrook 7H10 minimal media plates (MM):

19 g 7H10 powder, 5 ml glycerol, 0.2% glucose, 0.85% sodium chloride and 0.05% Tween

Middlebrook 7H9 minimal media (MM):

4.7 g 7H9 powder, 2 ml glycerol, 0.2% glucose, 0.85% sodium chloride

Luria-Bertani broth (LB):

10 g tryptone, 10 g sodium chloride, 5 g yeast

Luria-Bertani agar plates (LA):

10 g tryptone, 10 g sodium chloride, 5 g yeast extract, 15 g agar

2XTY media:

16 g tryptone, 10 g yeast extract, 5 g sodium chloride

Media was sterilized by autoclaving at 121°C for 10-15 min

### 5.1.2 DNA molecular weight markers

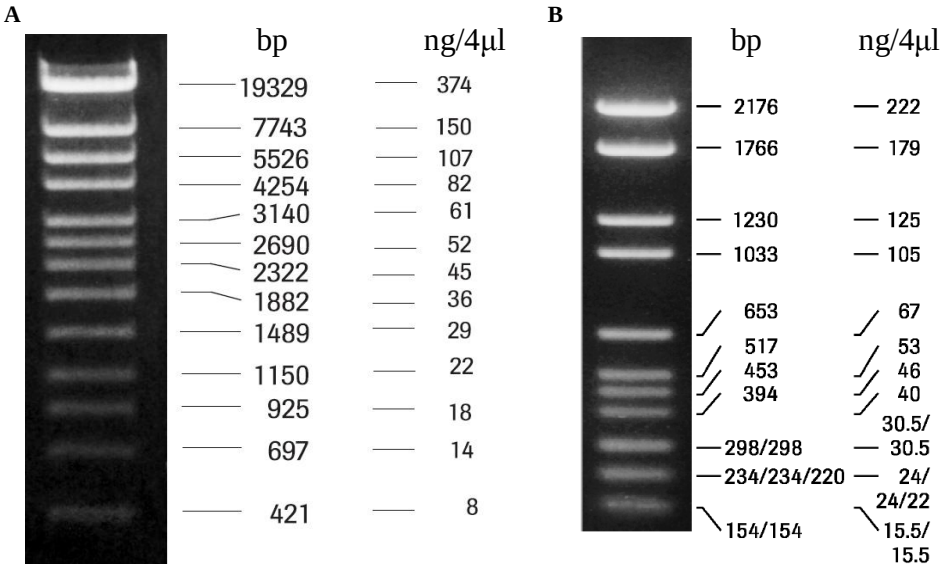



**Fig. 5.1.2.1: The DNA molecular weight markers. A.** Marker λIV. **B.** Marker λVI used in this study (Supplied by Roche Biochemicals).

## 5.2 PCR primers used in this study and amplification

### 5.2.1 PCR primers used in this study

**Table: 5.2.1.1 Primers used for PCR amplification.**

| Name of fragment | Name of primers | Oligonucleotide sequence (5'- 3') | Annealing temperature and DNA polymerase | Amplicon properties |
|---|---|---|---|---|
| Msm *nth* | NthCF | AGCTTGACGGTCTGAAGGACC AGT | Taq/Phusion = 63.0 $^{o}$C | 1150 bp *nth* amplicon generated from wild type Msm genomic DNA including 350 bp upstream of the gene. |
| | NthCR | GGTACCCGAGACGCTCATCGT GTG | | |
| Us | NthUsF1 | *AAGCTT* GTGTGCGTGCTGGCCAAG | Taq/Phusion = 63.5 $^{o}$C | 1185 bp amplicon generated from wild type genomic Msm DNA retaining 100 bp upstream of the *nth* gene |
| | NthUsR1 | *AGATCT* CACACCGGTGATCGTCTT | | |
| Ds | NthDsF2 | *AGATCT* ACGACGCACGAGACCCAG | Taq/Phusion = 63.5 $^{o}$C | 1170 bp amplicon generated from the wild type Msm *nth* allele retaining 110 bp downstream of the gene. |
| | NthDsR2 | *GGTACC* AGCTGGGCGATCTCTTCC | | |
| RRDR region of *rpoB* | MsmrpoB F2 | TGGTGCGTCTGCACGAGGGTC | Taq = 60 $^{o}$C | 184bp amplicon of *rpoB* allele. The RRDR region was sequenced using the internal primer. |
| | MsmrpoB R2 | ATCTGGTCGGTGACCACACCG | | |

**Table 5.2.1.2: Sequencing primers used in this study.**

| Template | Name of Primer | Sequence (5' – 3') |
|---|---|---|
| pTWEETY::*nth* | pTWEETY F | GCTGAAGCCAGTTACCTTCG |
| | pTWEETY R | TTTGCCTTCCTGTTTTTGCT |
| | Nth M | CGAACTCGAAGAGCTGATCC |
| | Nth F | AGCTTGACGGTCTGAAGGACCAGTT |
| | Nth R kpn | GGTACCCGAGACGCTCATCGTGTG |
| *rpoB* | Msmrpo BF1 | GCAGACCCTGATCAACATCC |

**Table 5.2.1.3: Primers used to distinguish between wild type and mutants.**

| Temp-late | Name of primers | Oligonucleotide sequence (5'- 3') | Annealing temperature and DNA polymerase | Amplicon Sizes (bps) |
|---|---|---|---|---|
| *f1* | F1F | CCGCAGCTATTCGTATGAGA | Taq = 65.0 $^{o}$C | Msm = 1554 |
| | F1R | AGATCTAATGGAAGGTGTGCATGTCC | | Mutant = 875 |
| *f2* | F2F | AGATCTGTCACCGCAGCAGGATAAC | Taq = 64.0 $^{o}$C | Msm = 1549 |
| | F2R | GCCGATGCTGTCGATGTACT | | Mutant = 1248 |
| *n1* | N1F | AGATCT GTATCACGCCGTGGTGCTC | Taq = 64.0 $^{o}$C | Msm = 1178 |
| | N1R | AAGCTTCTGATCCCACGAGCGACTAT | | Mutant = 692 |
| *n2* | N2F | AAGCTTACGACTACCCGGCGTTCT | Taq = 65.0 $^{o}$C | Msm = 1181 |
| | N2R | GGTACCCCGACGACGACAAGTAACAA | | Mutant = 544 |
| *nth* | NthF1 | GTCAGCCGTGGTCAGTACC | Taq = 63.0 $^{o}$C | Msm = 780 |
| | NthR1 | CGAGACGCTCATCGTGTG | | Mutant = 210 |

**Table 5.2.1.4: Primers used for gene expression analysis by SQ RT-PCR.**

| Name of primers | Oligonucleotide sequence (5'- 3') | Annealing temperature and DNA polymerase | Amplicon properties |
|---|---|---|---|
| RTF1 | TTCCTGGCCGAGCTTGAT | Taq = 56.5 $^{o}$C | Amplification of fragment from position 135-360 in *nth* |
| RTR1 | ACAAGCGGGTCAACCTGA | | |
| SigAF1 | GGGCGTGATGTCCATCTCCT | Taq = 56.0 $^{o}$C | Amplification of fragment from position 367-488 *sigA* |
| SigAR1 | GTATCCCGGTGCATGGTC | | |

## 5.2.2 PCR amplification

A

| Name of primers | Amplicon Sizes (bps) |
|---|---|
| F1 | Msm = 1554<br>Mutant = 875 |
| F2 | Msm = 1549<br>Mutant = 1248 |
| N1 | Msm = 1178<br>Mutant = 692 |
| N2 | Msm = 1181<br>Mutant = 544 |
| Nth | Msm = 780<br>Mutant = 210 |

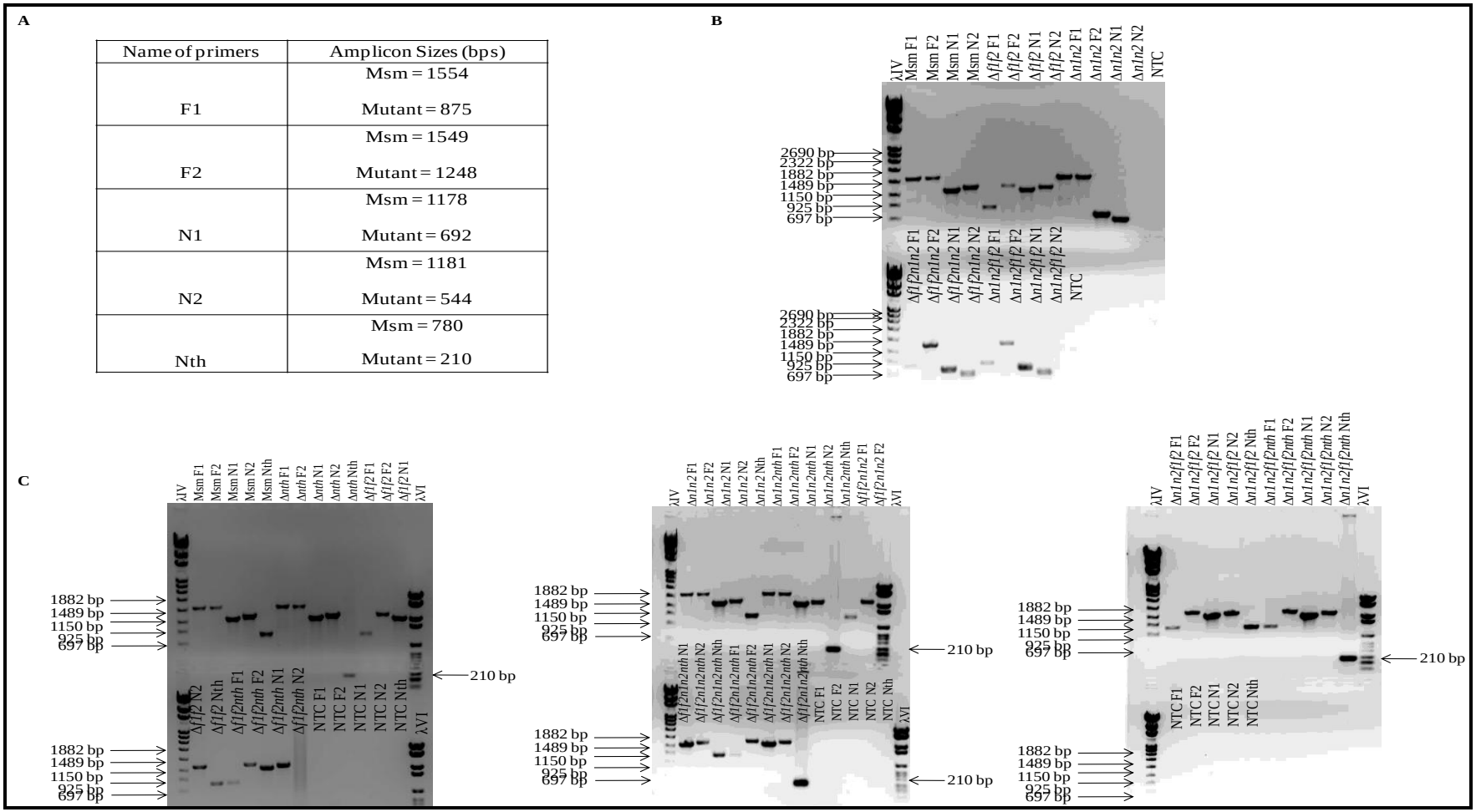



**Fig. 5.2.2.1: PCR amplification of the Fpg/Nei deficient mutants in the absence and presence of Nth. A.** Table of the primers used to distinguish between wild type Msm and mutants (a full description of these primers are in Table 5.2.1.3, Appendix 5.2.1). **B.** PCR screen of the Fpg/Nei deficient mutants before generating the Nth mutants. **C.** PCR screen of the Fpg/Nei/Nth deficient mutants recovered from the freezer for phenotypic experiments. A no template control (NTC) of each primer set described as NTC F1, F2, N1, N2 and Nth was included to show there was no DNA contamination.

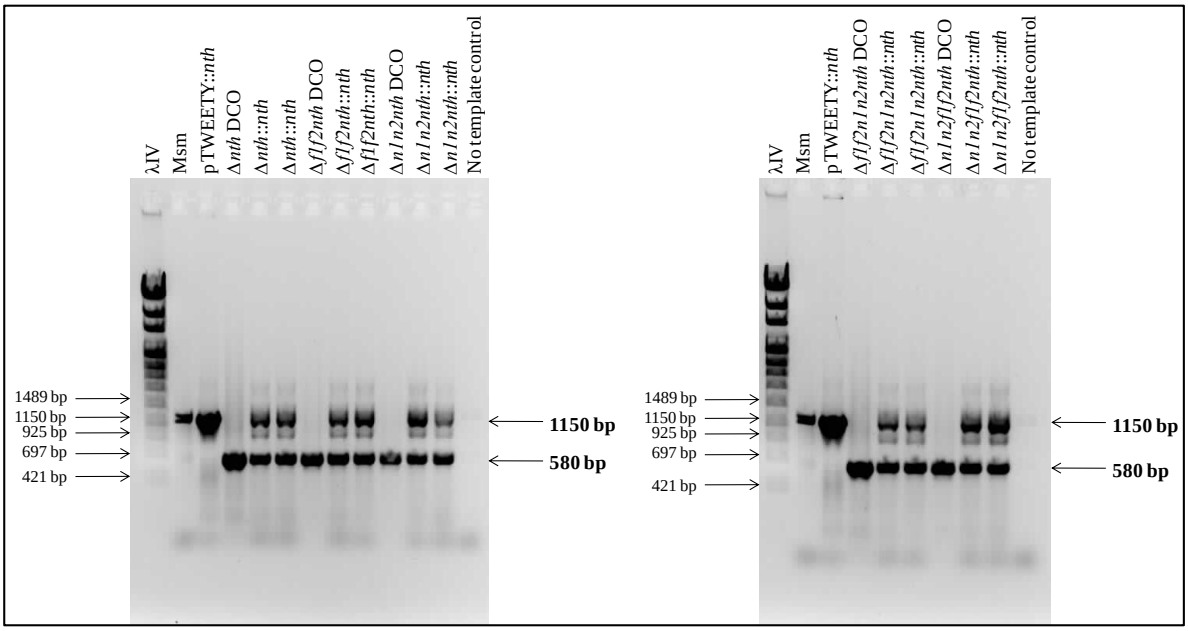



**Fig. 5.2.2.2: PCR amplification of chromosomal DNA from possible Msm *nth* complements using NthCF and NthCR primers in Table 5.2.1.1 (Appendix 5.2.1).**

## 5.3 Vectors maps used in this study

### 5.3.1 Vectors

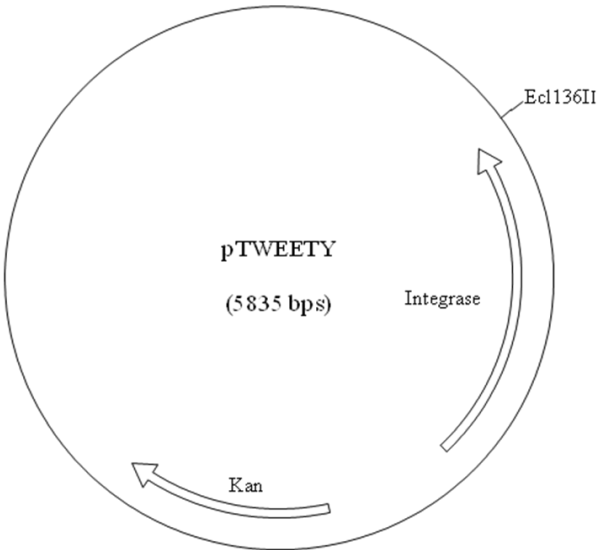



**Fig. 5.3.1.1: pTWEETY vector used for constructing the complementation vector using the *Ecl136*II site**.

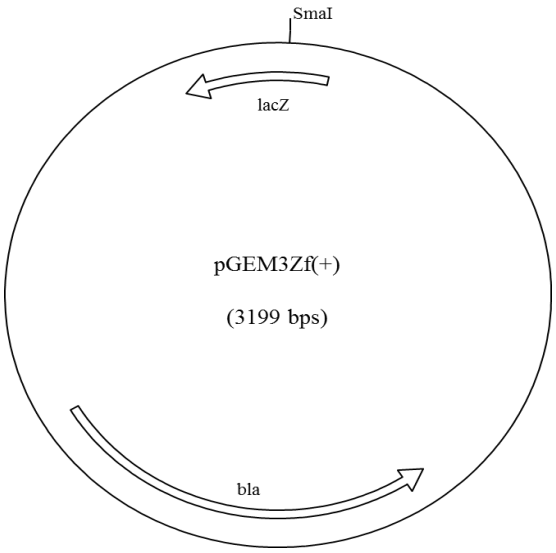



**Fig. 5.3.1.2: pGEM3Zf(+) vector used to clone PCR amplified Msm *nth* upstream and downstream fragments at the *Sma*I site.**

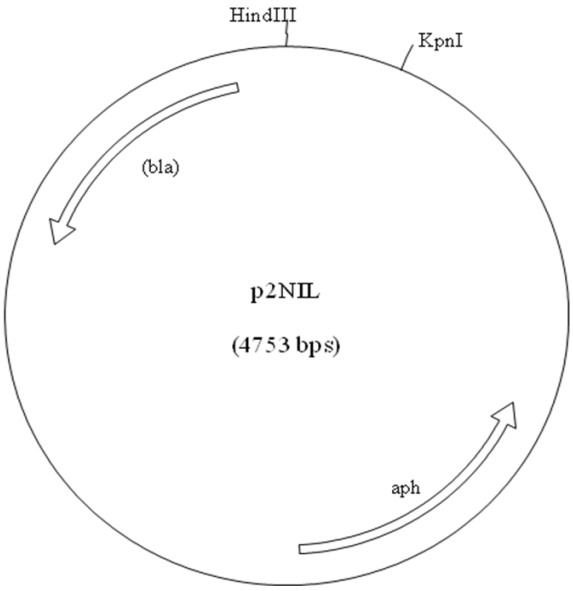



**Fig. 5.3.1.3: p2NIL suicide vector used to clone upstream and downstream *nth* fragments simultaneously to generate the *nth* deleted allele at the *Hind*III and *Kpn*I sites.**

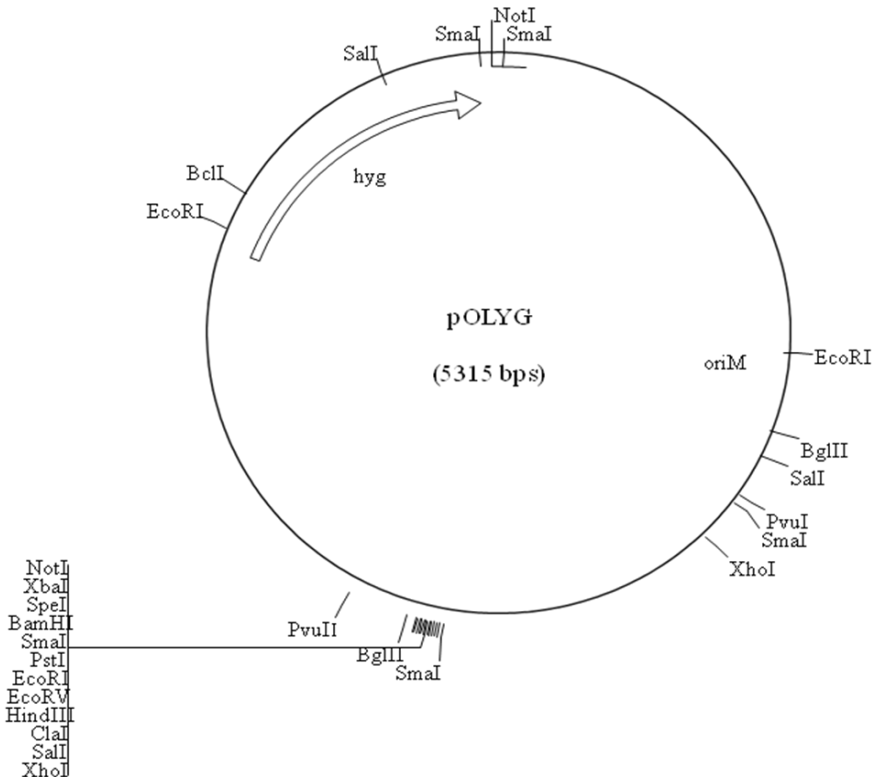



**Fig. 5.3.1.4: pOLYG replicating vector used in electroporations as a positive control to determine the transformation efficiency**.

## 5.3.2 PacI cassette nucleic acid sequence


CTGCAGCCCGGGGGATCGAAAAGGTTAGGAATACGGTTAGCCATTTGCCTGCTTTTATATAGTTATATGGGATTCACCTTTATGTTGATAAGAAATAAAAGAAAATGCCAATAGGATN
NNNN

## 5.4 Bioinformatic analysis of the Nth DNA glycosylases

**Table 5.4.1: Similarity of the Msm Nth protein compared to several related microorganisms using BLASTp.** The locus of Nth in each of the microorganisms is indicated in brackets.

| | Mtb (Rv3674c) | Ec (b1633) | *M. bovis* (BCG-3732c) | *M. leprae* (ML2301) | *M. avium paratuber-culosis* (MAP_0400) | *C. efficiens* (CE0289) | *C. jeikeium* (JK1971) | *B. subtilis* (BSU-22340) | *H. pylori* (HP0585) | *S. typhimurium* (STM1453) |
|---|---|---|---|---|---|---|---|---|---|---|
| Msm (MSMEG_6187) | $9e^{-149}$ | $1e^{-41}$ | $3e^{-150}$ | $3e^{-156}$ | $2e^{-148}$ | $1e^{-116}$ | $7e^{-119}$ | $3e^{-56}$ | $2e^{-55}$ | $5e^{-41}$ |

# 6. References


Akif, M., Khare, G., Tyagi, A. K., Mande, S. C. & Sardesai, A. A. (2008). Functional studies of multiple thioredoxins from *Mycobacterium tuberculosis*. *Journal of Bacteriology, 190,* 7087-7095.

Andersen, P. (2006). Vaccine strategies against latent tuberculosis infection. *Trends in Microbiology, 15*, 7-13.

Armel, P. R., Strniste, G. F. & Wallace, S. S. (1977). Studies on *Escherichia coli* x-ray endonuclease specificity. Roles of hydroxyl and reducing radicals in the production of DNA lesions. *Radiation Research, 69,* 328–338.

Arner, E. S. J. & Holmgren, A. (2000). Physiological functions of thioredoxin and thioredoxin reductase. *Eur. J. Biochem, 267,* 6102-6109.

Asahara, H., Wistort, P. M., Bank, J. F., Bakerian, R. H. & Cunningham, R. P. (1989). Purification and characterization of *Escherichia coli* endonuclease III from the cloned *nth* gene. *Biochemistry, 28,* 4444–4449.

Au, K. G., Clark, S., Miller, J. H. & Modrich, P. (1989). *Escherichia coli mutY* gene encodes an adenine glycosylase active on G-A mispairs. *Proceedings of the National Academy of Sciences of the United States of America , 86,* 8877-8881.

Augeri, L. Lee, Y. M., Barton, A. B., & Doetsch, P. W. (1997). Purification, characterization, gene cloning, and expression of *Saccharomyces cerevisiae* redoxyendonuclease, a homolog of *Escherichia coli* endonuclease III. *Biochemistry, 36* , 721–729.

Baily, V. & Verly, W. G. (1987). *Escherichia coli* endonuclease III is not an endonuclease but a β-elimination catalyst. *Biochemistry, 242*, 565-572.

Bayer, R. & Wilkinson, D. (1995). Directly observed therapy for tuberculosis: history of an idea. *Lancet, 345,* 1545–1548.

Blaisdell, J. O., Hatahet, Z. & Wallace, S. S. (1999). A novel role for *Escherichia coli* endonuclease VIII in prevention of spontaneous G→T transversions. *Journal of Bacteriology, 181,* 6396–6402.

Bloom, B. R. & Murray, C. J. (1992). Tuberculosis: commentary on a reemergent killer. *Science, 257,* 1055-1064.

Bloom, B. R. & Fine, P. E. (1994). In Bloom, B.R (Ed) *Tuberculosis: Pathogenesis, Protection and Control* (pp. 531-537). ASM press.

Boal, A. K., Genereux, J. C., Sontz, P. A., Gralnick, J. A., Newman, D. K. & Barton, J. K. (2009). Redox signaling between DNA repair proteins for efficient lesion detection. *Proc Natl Acad Sci USA., 106,* 15237–15242.

Boon, E. M., Livingston, A. L., Chmiel, N. H., David, S. S. & Barton, J. K. (2003). DNA-mediated charge transport for DNA repair. *Proc Natl Acad Sci USA., 100,* 12543–12547.

Boshoff, H. I., Reed, M. B., Barry, C. E. & Mizrahi, V. (2003). DnaE2 polymerase contributes to in vivo survival and the emergence of drug resistance in *Mycobacterium tuberculosis*. *Cell, 113,* 183–193.

Boshoff, H. I., Myers, T. G., Copp, B. R., McNeil, M. R., Wilson, M. A. & Barry III, C. E. (2004). The Transcriptional Responses of *Mycobacterium tuberculosis* to Inhibitors of Metabolism. *Journal of Biological Chemistry, 279,* 40174–40184.

Breimer, L. H. & Lindahl, T. (1984). DNA Glycosylase Activities for Thymine Residues Damaged by Ring Saturation, Fragmentation, or Ring Contraction Are Functions of Endonuclease III in *Escherichia coli*. *Journal of Biological Chemistry, 259,* 5543-5548.

Breimer, L. H. & Lindahl, T. (1985). Enzymatic excision of DNA bases damaged by exposure to ionizing radiation or oxidizing agents. *Mutation Research, 150,* 85-89.

Buchmeier, N. A., Newton, G. L., Koledin, T. & Fahey, R. C. (2003). Association of mycothiol with protection of *Mycobacterium tuberculosis* from toxic oxidants and antibiotics. *Molecular Microbiology, 47,* 1723-1732.

Buchmeier, N. A., Newton, G. L. & Fahey, R. C. (2006). A mycothiol synthase mutant of *Mycobacterium tuberculosis* has an altered thioldisulfide content and limited tolerance to stress. *Journal of Bacteriology, 188,* 6245-6252.

Caldeira de Araujo, A. & Favre, A. 1986. Near ultraviolet DNA damage induces the SOS responses in *Escherichia coli*. *The EMBO Journal, 5,* 175- 179.

Chaisson, R. E., Barnes, G. L., Hackman, J., Watkinson, L., Kimbrough, L., Metha, S.,Cavalcante, S. & Moore, R.D.(2001). A randomized, controlled trial of interventions to improve adherence to isoniazid therapy to prevent tuberculosis in injection drug users. *American Journal of Medicine, 110,* 610-615.

Chan, E. & Weiss, B. (1987). Endonuclease IV of *Escherichia coli* is induced by paraquat. *Proc. Natl. Acad. Sci. USA , 84,* 3189-3193.

Chaulet, P., Boulahbal, F. & Grosset, J. (1995). Surveillance of drug resistance for tuberculosis control: why and how? *Tubercle and Lung Disease, 76,* 487–492.

Cheng, K. C., Cahill, D. S., Kasai, H., Nishimura, S. & Loeb, L. A. (1992). 8-Hydroxyguanine, an abundant form of oxidative DNA damage, causes G-->T and A-->C substitutions. *Journal of Biological Chemistry, 267,*166–172.

Chow, K. H. & Courcelle, J. (2007). RecBCD and RecJ/RecQ initiate DNA degradation on distinct substrates in UV-irradiated *Escherichia coli*. *Radiation Research, 168,* 499–506.

Claiborne, A., Malinowski, D. P. & Fridovich, I. (1979). Purification and characterization of hydroperoxidase II of *Escherichia coli*. *Journal of Biological Chemistry, 254,*11664-11668.

Colditz, G. A., Brewer, T. F., Berkey, C. S., Wilson, M. E., Burdick, E., Fineberg, H. V., & Mosteller,F.(1994). Efficacy of BCG vaccine in the prevention of tuberculosis. Meta-analysis of the published literature. *JAMA, 271*, 698-702.

Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., & Harris, D. 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. & Barrell, B. G. (1998).Deciphering the biologyof *Mycobacterium tuberculosis* from the complete genome sequence. *Nature, 393*, 537–544.

Cole, S. T., Eiglmeier, K., Parkhill, J., James, K. D., Thomson, N. R., Wheeler, P. R., Honoré, 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. & Barrell, B.G. (2001). Massive gene decay in the leprosy bacillus. *Nature, 409*, 1007-1011.

Cooper, A. M. (2009 ). Cell-mediated immune responses in tuberculosis. *Annual Review of Immunology, 27*, 393–422.

Cunningham, R. P. & Weiss, B. (1985). Endonuclease III (*nth*) mutants of *Escherichia coli*. *Proc Natl Acad Sci USA., 82*, 474–478.

Cunningham, R. P., Asahara, H., Bank, J. F., Scholes, C. P., Salerno, J. C., Surerus, K., Munck, E., McCracken, J., Peisach, J. & Emptage, M.H., (1989). Endonuclease III is an iron-sulfur protein. *Biochemistry, 28*, 4450-4455.

Cunningham, R. (1997). DNA glycosylases. *Mutation Research, 383*, 189-196.

Curti, E., Smerdon, S. J. & Davis, E. O. (2007). Characterization of the helicase activity and substrate specificity of *Mycobacterium tuberculosis* UvrD. *Journal of Bacteriology, 189*, 1542–1555.

Darwin, K. H. & Nathan, C. F. (2005). Role for nucleotide excision repair in virulence of *Mycobacterium tuberculosis*. *Infection and Immunology, 73*, 4581–4587.

Dasgupta, N. (2000). Characterization of a two-component system, devR-devS, of *Mycobacterium tuberculosis*. *Tubercle and Lung Disease, 80*, 141-159.

David, H. L. & Newman, C. M. (1971). Some observations on the genetics of isoniazid resistance in the tubercle bacilli. *The American Review of Respiratory Disease, 104*, 508-515.

De Voss, J. J., Rutter, K., Schroeder, B. G. & Barry, C. E. (1999). Iron acquisition and metabolism by mycobacteria. *Journal of bacteriology, 181*, 4443-4451.

DeWitt, S. K. & Adelberg, E. A. (1962). The occurrence of a genetic transposition in a strain of *Escherichia coli*. *Genetics, 47,* 577-585.

Della, M., Palmbos, P. L., Tseng, H. M., Tonkin, L. M., Daley, J. M., Topper, L. M., Pitcher, R.S., Tomkinson, A.E., Wilson, T.E. & Doherty, A.J.(2004). Mycobacterial Ku and ligase proteins constitute a twocomponent NHEJ repair machine. *Science, 306,* 683–685.

Demple, B. & Harrison, L. (1994). Repair of oxidative damage to DNA: Enzymology and biology. *Annual Review of Biochemistry, 63,*915-948.

Demple, B. & Linn, S. (1980). DNA N-glycosylases and UV repair. *Nature, 287,* 203–208.

Dizdaroglu, M., Laval, J. & Boiteux, S. (1993). Substrate specificity of *Escherichia coli* endonuclease III: excision of thymine and cytosine-derived lesions in DNA produced by ionizing radiation-generated free radicals. *Biochemistry, 32,* 12105–12111.

Dizdaroglu, M., Bauche, C., Rodriguez, H. & Laval, J. (2000). Novel substrates of *Escherichia coli* Nth protein and its kinetics for excision of modified bases from DNA damaged by free radicals. *Biochemistry, 39,* 5586–5592.

Dizdaroglu, M. (2003). Review: Substrate specificities and excision kinetics of DNA glycosylases involved in base-excision repair of oxidative DNA damage. Mutation Research, 531, 109–126.

Duarte, V., Muller, J.G. & Burrows, C.J. (1999). Insertion of dGMP and dAMP during *in vitro* DNA synthesis opposite an oxidized form of 7,8-dihydro-8-oxoguanine. *Nucleic Acids Research, 27,* 496-502.

Du Toit, L.C., Pillay, V. & Danckwerts, M. P. (2006). Tuberculosis chemotherapy: current drug delivery approaches. *Respiration Research*, *7*, 1-18.

Dussurget, O., Stewart, G., Neyrolles, O., Pescher, P., Young, D. & Marchal, G. (2001). PescheRole of *Mycobacterium tuberculosis* copper-zinc superoxide dismutase. *Infection & Immunity, 69,*529–533.

Dwyer, D. J., Kohanski, M. A., Hayete, B. & Collins, J. J. (2007). Gyrase inhibitors induce an oxidative damage cellular death pathway in *Escherichia coli*. *Mol Syst Biol, 3,* 91.

Dwyer, D. J., Kohanski, M. A. & Collins, J. J. (2009). Role of Reactive Oxygen Species in Antibiotic Action and Resistance. *Current Opinion in Microbiology*, *12* , 482–489.

Dye, C., Scheele, S., Dolin, P., Pathania, V. & Raviglione, M. C. (1999). Consensus statement. Global burden of tuberculosis: estimated incidence, prevalence, and mortality by country. WHO Global Surveillance and Monitoring Project. *JAMA ,282,* 677-686.

Ehrt, S. & Schnappinger, D. (2009). Microreview. Mycobacterial survival strategies in the phagosome: defense against host stresses. *Cellular Microbiology, 11,* 1170–1178.

Eide, L., Bjoras, M., Pirovano, M., Alseth, I., Berdal, K. G. & Seeberg, E. (1996). Base excision of oxidative purine and pyrimidine DNA damage in *Saccharomyces cerevisiae*

by a DNA glycosylase with sequence similarity to endonuclease III from *Escherichia coli*. *Proc Natl Acad Sci USA., 10*, 10735-10740.

Eisen, J. A. & Hanawalt, P. C. (1999). A phylogenomic study of DNA repair genes, proteins, and processes. *Mutation Research, 435*,171-213.

Farr, S. B. & Kogoma, T. (1991). Oxidative Stress Responses in *Escherichia coli* and *Salmonella typhimurium*. *Microbiological reviews, 55*, 561-585

Fenton, M. J., Riley, L. W. & Schlesinger, L. S. (2005). Receptor-mediated recognition of *Mycobacterium tuberculosis* by host cells. In K. E. S.T. Cole, *Tuberculosis and the Tubercle Bacillus* (pp. 405-426). NY: ASM Press.

Fine, P. E. (1989a). The BCG story: lessons from the past and implications for the future. *Review Infectious Diseases, 11 (Suppl. 2),*S353–S359.

Fine, P. E., Ponnighaus, J. M., & Maine, N. (1989b). The distribution and implications of BCG scars in northern Malawi. *Bulletin of the World Health Organisation, 67*, 35-42.

Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., Kerlavage, A. R., Bult, C. J., Tomb, J.-F., Dougherty, B. A., Merrick, J. M., McKenney, K., Sutton, G., FitzHugh, W., Fields, C., Gocayne, J. D., Scott, J., Shirley, R., Liu, L.-I., Glodek, A., Kelley, J. M., Weidman, J. F., Phillips, C. A., Spriggs, T., Hedblom, E., Cotton, M. D., Utterback, T.R., Hanna, M. C., Nguyen, D. T., Saudek, D. M., Brandon, R. C., Fine, L. D., Fritchman, J.L., Fuhrmann, J. L., Geoghagen, N. S. M., Gnehm, C. L., McDonald, L. A., Small, K. V., Fraser, C. M., Smith, H. O. & Venter, J. C. (1995). Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science, 269,* 496-512.

Flynn, J. L. & Chan, J. (2001). Immunology of tuberculosis. *Annual Review of Immunology, 19*, 93–129.

Fox, W. & Mitchison, D. A. (1975). Short-course chemotherapy for pulmonary tuberculosis. *The American Review of Respiratory Disease*, 325-353.

Fridovich, I. (1989). Superoxide dismutase: an adaptation to a paramagnetic gas. *Journal of Biological Chemistry*, *264,*7761-7764.

Friedberg, E. C., Walker, G. C. & Siede, W. (1995). DNA Repair and Mutagenesis. *American Society for Microbiology*, 91-225.

Friedberg, E. C., Wagner, R. & Radman, M. (2002). Specialized DNA polymerases, cellular survival, and the genesis of mutations. *Science, 296*, 1627–1630.

Friedberg, E. C., Lehmann, A. R. & Fuchs, R. P. (2005). Trading places: how do DNA polymerases switch during translesion DNA synthesis? *Mol Cell, 18*, 499–505.

Frieden, T. R., & Sbarbaro, J. A. (2007). Promoting adherence to treatment for tubercuoisis: the importance of direct observation. *Bulletin of the World Health Organization, 85*, 407-409.

Fromme, J. C., & Verdine, G. L. (2003). Structure of a trapped endonuclease III–DNA covalent intermediate. *EMBO J, 22*, 3461-3471.

Fromme, J.C., Banerjee, A., Huang, S.J. & Verdine, G.L. (2004). Structural basis for removal of adenine mispaired with 8-oxoguanine by MutY adenine DNA glycosylase. *Nature, 427*, 652-656.

Fu, W., O'Handley, S., Cunningham, R. P. & Johnson, M. K. (1992). The role of the iron-sulfur cluster in *Escherichia coli* endonuclease III A resonance Raman study. *Journal of Biological Chemistry, 267*, 16135–16137.

Garner, P., Smith, H., Munro, S. & Volmink, J. (2007). Promoting adherence to tuberculosis treatment. *Bulletin of the World Health Organization, 85*, 404-409.

Gates, F. T. & Linn, S. (1977). Endonuclease from *Escherichia coli* that acts specifically upon duplex DNA damaged by ultraviolet light, osmium tetroxide, acid, or x-rays. *Journal of Biological Chemistry, 252*, 2802–2807.

Geldmacher, C., Ngwenyama, N., Schuetz, A., Petrovas, C., Reither, K., Heeregrave, E. J., Casazza, J.P., Ambrozak, D.R., Louder, M., Ampofo, W., Pollakis, G., Hill, B., Sanga, E., Saathoff, E., Maboko, L., Roederer, M., Paxton, W.A., Hoelscher, M. & Koup, R.A. (2010). Preferential infection and depletion of *Mycobacterium tuberculosis*-specific CD4 T cells after HIV-1 infection. *Journal of Experimental Medicine, 207*, 2869–2881.

Glynn, J. R., Murray, J., Bester, A., Nelson, G., Shearer, S. & Sonnenberg, P. (2008). SEffects of duration of HIV infection and secondary tuberculosis transmission on tuberculosis incidence in the South African gold mines. *AIDS*, 1859–1867.

Gong, C., Bongiorno, P., Martins, A., Stephanou, N. C., Zhu, H., Shuman, S. & Glickman, M. S. (2005). Mechanism of non homologous end-joining in mycobacteria: a low-fidelity repair system driven by Ku, ligase D and ligase C. *Nature Structural and Molecular Biology, 12*, 304–312.

Goodman, M. (2002). Error-prone repair DNA polymerases in pro karyotes and eukaryotes. *Annual Review of Immunology, 71*, 17–50.

Goosens, V. J. (2008). *Msc dissertation: The Construction and Phenotypic Characterization of Mycobacterial Mutants Deficient in DNA Glycosylases.* Faculty of Science, University of the Witwatersrand.

Goosens, V., Mizrahi, V., & Gordhan, B. G. (2012). Goosens, V., Mizrahi, V. & Gordhan, B.G. 2012. The role of Fpg DNA glycosylases and Nei endonucleases in mutagenesis in *Mycobacterium smegmatis*. In preparation. *Unpublished*.

Gordhan, B. G., & Parish, T. (2001). Gene replacement using pretreated DNA. In e. T. Parish & N.G. Stoker, *Methods in Molecular Medicine* (pp. 77-91). NJ: Humana Press Inc.

Greenberg, J. T., & Demple, B. (1988). Overproduction of peroxide-scavenging enzymes in *Escherichia coli* suppresses spontaneous mutagenesis and sensitivity to redox-cycling agents in oxyR mutants. *EMBO J, 7*, 2611-2617.

Grollman, A.P. & Moriya, M. (1993). Mutagenesis by 8-oxoguanine: an enemy within. *Trends in Genetics, 9,* 246–249.

Guo, Y., Bandaru, V., Jaruga, P., Zhao, X., Burrows, C. J., Iwai, S., Dizdaroglu, M., Bond, J. P., & Wallace, S.S. (2010). The oxidative DNA glycosylases of *Mycobacterium tuberculosis* exhibit different substrate preferences from their *Escherichia coli* counterparts. DNA Repair (Amst). *DNA Repair (Amst), 9,* 177-190.

Guthlein, C., Wanner, R. M., Sander, P., Davis, E. O., Bosshard, M., Jiricny, J., Bottger, E.C., & Springer, B. (2008). Characterisation of the mycobacterial NER system reveals novel functions of uvrD1 helicase. *Journal of Bacteriology, 191,* 555–562.

Hailer, M. K., Slade, P. G., Martin, B. D. & Sugden, K.D. (2005). Nei deficient *Escherichia coli* are sensitive to chromate and accumulate the oxidized guanine lesion spiroiminodihydantoin. *Chemical Research in Toxicology, 18,* 1378–1383.

Harrington, M. (2010). From HIV to tuberculosis and back again: a tale of activism in 2 pandemics. *Clinical and Infectious Diseases, 50(Suppl. 3),*S260–S266.

Hatahet, Z., Kow, Y. W., Purmal, A. A., Cunningham, R. P. & Wallace, S. S. (1994). New substrates for old enzymes. 5-Hydroxy-2'-deoxycytidine and 5-hydroxy-2'-deoxyuridine are substrates for *Escherichia coli* endonuclease III and formamidopyrimidine DNA N-glycosylase, while 5-hydroxy-2'-deoxyuridine is a substrate for uracil DNA N-glycosylase. *Journal of Biological Chemistry, 269,* 18814-18820.

Hayes, R. C., Petrullo, L. A., Huang, H. M., Wallace, S. S. & LeClerc, J. E. (1988). Oxidative damage in DNA. Lack of mutagenicity by thymine glycol lesions. *Journal of Molecular Biology, 201,*239–246.

Herzog, H. & Basel, H. (1998). History of Tuberculosis. *Respiration, 65,* 5-15.

Hollis, T., Ichikawa, Y., & Ellenberger, T. (2000). DNA bending and a flip-out mechanism for base excision by the helix–hairpin–helix DNA glycosylase, *Escherichia coli* AlkA. *The EMBO J, 19,* 758-766.

Hoshino, Y., Nakata, K., Hoshino, S., Honda, Y., Tse, D. B., Shioda, T.,Rom, W.N. & Weiden, M. (2002). Maximal HIV-1 replication in alveolar macrophages during tuberculosis requires both lymphocyte contact and cytokines. *Journal of Experimental Medicine, 195,* 495–505.

Hoshino, Y., Hoshino, S., Gold, J. A., Raju, B., Prabhakar, S., Pine, R., Rom, W.N., Nakata, K. & Weiden, M. (2007). Mechanisms of polymorphonuclear neutrophilmediated induction of HIV-1 replication in macrophages during pulmonary tuberculosis. *Journal of Infectious Diseases, 195,*1303–1310.

Hua, X., Xu, X., Li, M., Wang, C., Tian, B. & Hua, Y. (2012). Three *nth* homologs are all required for efficient repair of spontaneous DNA damage in *Deinococcus radiodurans*. *Extremophiles, 16,* 477–484.

Ide, H., Kow, Y. W. & Wallace, S. S. (1985). Thymine glycols and urea residues in M13 DNA constitute replicative blocks in vitro. *Nucleic Acids Research, 13,* 8035–8052.

Imlay, J. A. (2008). Cellular defences against superoxide and hydrogen peroxide. *Annual Review of Biochemistry, 77,* 755–776.

Iseman, M. D., Cohen, D. L. & Sbarbaro, J. A. (1993). Directly observed treatment of tuberculosis. We can't afford not to do it. *New England Journal of Medicine, 328,*576–578.

Jain, R., Kumar, P. & 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,* 1774–1785.

Janion, C. 2008. Review: Inducible SOS Response System of DNA Repair and Mutagenesis in *Escherichia coli*. *International Journal of Biological Science, 4,* 338-344.

Jiang, D., Hatahet, Z., Blaisdell, J. O., Melamede, R. J. & Wallace, S. S. (1997a). *Escherichia coli* endonuclease VIII: cloning, sequencing, and overexpression of the *nei* structural gene and characterization of *nei* and *nei nth* mutants. *Journal of Bacteriology, 179,* 3773-3782.

Kamolratanakul, P., Sawert, H., Lertmaharit, S., Kasetjaroen, Y., Akksilp, S. & Tulaporn, C. (1999). Randomized controlled trial of directly observed treatment (DOT) for patients with pulmonary tuberculosis in Thailand. *Transactions of the Royal Society of Tropical Medicine and Hygiene, 93,* 552-557.

Kana, B.D., Abrahams, G.L., Sung, N., Warner, D.F., Gordhan, B.G., Machowski, E.E., Tsenova, L., Sacchettini, J.C., Stoker, N.G., Kaplan, G. & Mizrahi, V. (2010) Role of the DinB Homologs, Rv1537 and Rv3056, in *Mycobacterium tuberculosis*. *Journal of Bacteriology, 192,* 2220–2227.

Katcher, H. L. & Wallace, S. (1983). Characterization of the *Escherichia coli* X-ray endonuclease, endonuclease III. *Biochemistry, 22,* 4071-4082.

Kaufmann, S. H., Cole, S. T., Mizrahi, V., Rubin, E. & Nathan, C. (2005). *Mycobacterium tuberculosis* and the host response. *Journal of Experimental Medicine, 201,*1693–1697.

Kim, J. & Linn, S. (1988). The mechanisms of action of *E.coli* endonuclease III and T4 UV endonuclease (endonuclease V) at AP sites. *Nucleic Acids Research,16,*1135–1141.

Koch, R. (1998). Die Aetiologie der Tuberculose. Berl. Klin. Wchnschr., xix: 221-230. In T. D. Brock, *In Milestones in Microbiology, 1556 to 1940* (p. 109). ASM press.

Kohanski, M. A., Dwyer, D. J., Hayete, B., Lawrence, C. A. & Collins, J. J. (2007). A common mechanism of cellular death induced by bactericidal antibiotics. *Cell, 130,* 797-810.

Kohanski, M. A., DePristo, M. A., & Collins, J. J. (2010). Sublethal Antibiotic Treatment Leads to Multidrug Resistance via Radical-Induced Mutagenesis. *Molecular Cell, 37,* 311–320.

Kornyushyna, O., Berges, A.M., Muller, J. G. & Burrows, C. J. (2002). In Vitro Nucleotide Misinsertion Opposite the Oxidized Guanosine Lesions

Spiroiminodihydantoin and Guanidinohydantoin and DNA Synthesis Past the Lesions Using *Escherichia coli* DNA Polymerase I (Klenow Fragment). *Biochemistry, 41*, 15304-15314.

Kow, Y.W., Wallace, S.S. & van Houten, B. (1989). UvrABC nuclease complex repairs thymine glycol, an oxidative DNA base damage. *Mutation Research, 235*,147–156.

Kowalczykowski, S. C., Dixon, D. A., Eggleston, A. K., Lauder, S. D. & Rehrauer, W. M. (1994). Biochemistry of homologous recombination in *Escherichia coli*. *Microbiology Reviews, 58*, 401–465.

Krishna, S., Maslov, S. & Sneppen, K. (2007). UV-Induced Mutagenesis in *Escherichia coli* SOS Response: A Quantitative Model. *PLoS Comput Biol, 3,* 451-462.

Krokan, H. E., Standal, R. & Slupphaug, G. (1997). DNA glycosylases in the base excision repair of DNA. *Journal of Biochemisrty, 325*,1-16.

Kumar, A., Deshane, J. S., Crossman, D. K., Bolisetty, S., Yan, B. S., Kramnik, I., Argawal, A. & Steyn, A. J. (2008). Heme oxygenase-1-derived carbon monoxide induces the *Mycobacterium tuberculosis* dormancy regulon. *Journal of Biological Chemistry, 283*, 18032-18039.

Kumar, A., Farhana, A., Guidry, L., Saini, V., Hondalus, M. & Steyn, A. J. (2011). Redox homeostasis in mycobacteria: the key to tuberculosis control? *Expert reviews in molecular medicine, 13,* 1-25.

Kumawat, K., Pathak, S. K., Spetz, A. L., Kundu, M. & Basu, J. (2010). Exogenous Nef is an inhibitor of *Mycobacterium tuberculosis*-induced tumor necrosis factor-alpha production and macrophage apoptosis. *Journal of Biological Chemistry, 285,* 12629–12637.

Kuo, C. F., McRee, D. E., Cunningham, R. P. & Tainer, J. A. (1992a).Crystallization and crystallographic characterization of the iron-sulfur containing enzyme endonuclease III from *Escherichia coli*. *Journal of Molecular Biology, 227,*347-351.

Kuo, C. F., McRee, D. E., Fisher, C. L., O'Handley, S. F., Cunningham, R. P. & Tainer, J. A. (1992b). Atomic structure of the DNA repair [4Fe-4S] enzyme endonuclease III. *Science, 258,* 434 440.

Kurthkoti, K., Srinath, T., Kumar, P., Malshetty, V.S., Sang, P.B., Jain, R., Manjunath, R. & Varshney, U. (2010). A distinct physiological role of MutY in mutation prevention in mycobacteria. *Microbiology, 156,* 88–98.

Kurthkoti, K. & Varshney U. (2011). Base excision and nucleotide excision repair pathways in mycobacteria. *Tuberculosis, 91,* 533-543.

Kurthkoti, K. &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,* 138-146.

Kusunose, E., Ichihara, K., Noda, Y. & Kusunose, M. (1976). Superoxide dismutase from *Mycobacterium tuberculosis*. *Journal of Biochemistry*, 1343–1352.

Larsen, M. H. (2000). Appendix 1. In G. F. Hatfull, & W. R. Jacobs Jr, *Molecular genetics of mycobacteria* (pp. 313-320). Washington, DC: ASM Press.

Lee, A. S., Lim, I. H., Tang, L. L., & Wong, S. Y. (2005). High frequency of mutations in the rpoB gene in rifampin-resistant clinical isolates of *Mycobacterium tuberculosis* from Singapore. *Journal of Clinical Microbiology, 43,* 2026–2027.

Leipold, M.D., Muller, J.G., Burrows, C.J. & David, S.S. (2000). Removal of hydantoin products of 8-oxoguanine oxidation by the *Escherichia coli* DNA repair enzyme, FPG. *Biochemistry, 39,* 14984–14992.

Leistikow, R. L. (2010). The *Mycobacterium tuberculosis* DosR regulon assists in metabolic homeostasis and enables rapid recovery from nonrespiring dormancy. *Journal of Bacteriology, 192,* 1662.

Li, X. & Lu, A. L. (2001). Molecular Cloning and Functional Analysis of the MutY Homolog of *Deinococcus radiodurans*. *Journal of Bacteriology, 183,* 6151-6158.

Little, J. W. & Mount, D. W. (1982). The SOS regulatory system of *Escherichia coli*. *Cell, 29,* 11–22.

Loewen, P. C. & Triggs, B. L. (1984). Genetic mapping of katE a locus that with katE affects the synthesis of a second catalase species in *E. coli*. *Journal of Bacteriology, 160,* 668-675.

Machowski, E. E., Barichievy, S., Springer, B., Durbach, S. I. & 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,* 2190-2195.

Malotte, C. K., Hollingshead, J. R. & Larro, M. (2001). Incentives vs outreach workers for latent tuberculosis treatment in drug users. *Malotte CK, Hollingshead JR, Larro M. IncentiAmerican Journal of Preventive Medicine, 20,* 103–107.

Masaoka, A., Terato, H., Kobayashi, M., Honsho, A., Ohyama, Y. & Ide, H. (1999). Enzymatic repair of 5-formyluracil. I. Excision of 5-formyluracil site-specifically incorporated into oligonucleotide substrates by alka protein (*Escherichia coli* 3 methyladenine DNA glycosylase II). *The Journal of biological chemistry, 274,* 25136-25143.

Matsumoto, Y., Zhang, Q. M., Takao, M., Yasui, A., & Yonei, S. (2001). *Escherichia coli* Nth and human hNTH1 DNA glycosylases are involved in removal of 8-oxoguanine from 8-oxoguanine/guanine mispairs in DNA. *Nucleic Acids Research, 29,*1975–1981.

Mazumder, A., Gerlt, J. A., Absalon, M. J., Stubbe, J., Cunningham, R. P., Withka, J. & Bolton, P.H. (1991). Stereochemical studies of the beta-elimination reactions at aldehydic abasic sites in DNA, endonuclease III from *Escherichia coli*, sodium hydroxide, and Lys-Trp-Lys. *Biochemistry, 30,*1119–1126.

McCammon M. T., Gillette, J. S., Thomas, D. P., Ramaswamy, S. V., Graviss, E. A., Kreiswirth, B. N., Vijg, J. & Quitugua, T. N. (2005). Detection of rpoB mutations

associated with rifampin resistance in *Mycobacterium tuberculosis* using denaturing gradient gel electrophoresis. *Antimicrob Agents Chemother, 49,* 2200–2209.

McCord, J. M. & Fridovich, I. (1969). An enzymic function for erythrocuprein. *Journal of Biological Chemistry,244,* 6049-6055.

McCready, K.A. & Ratledge, C. (1978). Amounts of iron, heme and related compounds in *Mycobacterium smegmatis* grown in various concentrations of iron. *Biochem Soc Trans, 6,* 421–423.

McNulty, J. M., Jerkovic, B., Bolton, P. H. & Basu, A. K. (1998). Replication inhibition and miscoding properties of DNA templates containing a site-specific cis-thymine glycol or urea residue. *Chemical Research in Toxicology, 11,* 666–673.

Meya, D. B. & McAdam, K. P. (2007). The TB pandemic: an old problem seeking new solutions. *International Journal of Medicine, 261,* 309-329.

Michaels, M. L. & 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,* 6321–6325.

Michaels, M.L., Tchou, J., Grollman, A.P. & Miller, J.H. (1992). A repair system for 8-oxo-7,8- dihydrodeoxyguanine. *Biochemistry, 31,* 10964–10968.

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

Morris, R. P. (2005). Ancestral antibiotic resistance in *Mycobacterium tuberculosis*. *Proc Natl Acad Sci USA, 102,* 12200-12205.

Muttil, P., Wang, C. & Hickey, A.J. (2009). Inhaled drug delivery for tuberculosis therapy. *Pharmaceutical research, 26,* 2401-2416.

Nakata, K., Rom, W. N., Honda, Y., Condos, R., Kanegasaki, S., Cao, Y., & Weiden, M. (1997). *Mycobacterium tuberculosis* enhances human immunodeficiency virus-1 replication in the lung. *American Journal of Respiratory and Critical Care Medicine, 155,* 996–1003.

Newton, G.L., Unson, M.D., Anderberg, S.J., Aguilera, J.A., Oh, N. N., delCardayre, S.B., Av-Gay, Y. & Fahey, R.C. (1999). Characterization of Mycobacterium smegmatis Mutants Defective in 1-dmyo-inosityl-2-amino-2-deoxy-alpha-dglucopyranoside and Mycothiol Biosynthesis. *Biochemical and biophysical research communications,* 255, 239-244.

Newton, G.L. & Fahey, R.C. 2002. Mycothiol biochemistry. *Arch. Microbiology, 178,* 388-394.

Núñez, M. E., Hall, D. B., & Barton, J. K. (1999). Long-range oxidative damage to DNA: Effects of distance and sequence. *Chemistry & biology, 6,* 85–97.

Ó Gaora, P. S., Barnini, C., Hayward, E., Filley, G., Rook, D., Young, J. & Thole, J. (1997). Mycobacteria as immunogens: development of expression vectors for use in multiple mycobacterial species. *Med. Princ. Pract. 6,* 91-96.

Olsen, I., Balasingham, S. V., Davidsen, T., Debebe, E., Rodland, E. A., van Soolingen, D., Kremer, K., Alseth, I., & Tonju, T. (2009). Characterization of the major formamidopyrimidine DNA glycosylase homolog in *Mycobacterium tuberculosis* and its linkage to variable tandem repeats. *Federation of European Microbiological Societies Immunology and Medical Microbiology, 56,*151–161.

O'Rourke, E. J., Chevalier, C., Pinto, A. V., Thiberge, J. M., Ielpi, L., Labigne, A. & Radicella, J.P. (2003). Pathogen DNA as target for host-generated oxidative stress: role for repair of bacterial DNA damage in *Helicobacter pylori* colonization. *Proc Natl Acad Sci USA, 100,* 2789-2794.

Parish, T. & 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,* 1969-1975.

Patel, N. R., Swan, K., Li, X., Tachado, S. D. & Koziel, H. (2009).Impaired *M. tuberculosis-mediated* apoptosis in alveolar macrophages from HIV persons: potential role of IL-10 and BCL-3. *Journal of Leukocyte Biology, 86,* 53–60.

Pham, T. T., Jacobs-Sera, D., Pedulla, M. L., Hendrix, R. W. & Hatfull, G. F. (2007). Comparative genomic analysis of mycobacteriophage Tweety: evolutionaryand construction of compatible integration vectors for mycobacteria. *Microbiology, 153,*2711–2723.

Prinz, W.A., Aslund, F., Holmgren, A. & Beckwith J. (1997). The role of the thioredoxin and glutaredoxin pathways in reducing protein disulfide bonds in the *Escherichia coli* cytoplasm. *J. Biol. Chem, 272,* 15661-15667.

Purmal A. A., Kow,Y. W. & Wallace, S. S. (1994). Major oxidative products of cytosine, 5-hydroxycytosine and 5-hydroxyuracil, exhibit sequence context-dependent mispairing *in vitro*. *Nucleic Acids Research, 22,* 72–80.

Radman, M. (1976). An Endonuclease from *Escherichia coli* That Introduces Single Polynucleotide Chain Scissions in Ultraviolet-irradiated DNA. *Journal of Biological Chemistry, 251,* 1438-1445.

Ramakrishnan, L. (2012). Revisiting the role of the granuloma in tuberculosis. *Nature Reviews Immunology, 12*, 352-366.

Ramaswamy, S., & Musser, J. M. (1998). . Molecular genetic basis of antimicrobial agent resistance in *Mycobacterium tuberculosis*: update. *Tuber Lung Dis, 79,* 3–29.

Rawat, M., Johnson, C., Cadiz, V. & Av-Gay, Y. (2007). Comparative analysis of mutants in the mycothiol biosynthesis pathway in *Mycobacterium smegmatis*. *Biochemical and Biophysical Research Communications, 363,* 71-76.

Rodrigues, L. C., Diwan, V. K. & Wheeler, J. G. (1993). Protective effect of BCG against tuberculous meningitis and military tuberculosis: a meta-analysis. *Int. J. Epidemiol, 22* , 1154–1158.

Rosche, W.A. & Foster, P.L. (2000). Determining mutation rates in bacterial populations.

*Methods, 20,* 4-17.

Rouault, T. A., & Klausner, R. D. (1996). Iron-sulfur clusters as biosensors of oxidants and iron. *Trends in Biochemical Sciences, 21,* 174-177.

Saito, Y., Uraki, F., Nakajima, S., Asaeda, A., Ono, K., Kubo, K. & Yamamoto, K. (1997). Characterization of endonuclease III (*nth*) and endonuclease VIII (*nei*) mutants of *Escherichia coli* K-12. . *J Bacteriol , 179,* 3783-5.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). Molecular Cloning. A laboratory manual Second Edition. Cold Spring Harbour, New York: Cold Spring Harbour Laboratory Press.

Sambrook, J. & Russell, D. W. (2001). Molecular Cloning. A laboratory manual. Third Edition. Cold Spring Harbour, New York: Cold Spring Harbour Laboratory Press.

Sancar, A. (1996). DNA excision repair. *Annual Review of Biochemistry, 65,* 43–81.

Sareen, D., Newton, G.L., Fahey, R.C. & Buchmeier, N.A. (2003). Mycothiol Is Essential for Growth of *Mycobacterium tuberculosis* Erdman. *Journal of Bacteriology, 185,* 6736–6740.

Scandalios, J. G., Guan, L. & Polidoros, A. N. (1997). Catalases in plants: gene structure, properties, regulation, and expression. In e. JG Scandalios, *Oxidative Stress and the Molecular Biology of Antioxidant Defenses* (pp. 343-406). NY: Cold spring habour laboratory press.

Scandalios, J. G. (2005). Oxidative stress: molecular perception and transduction of signals triggering antioxidant gene defences. *Brazilian Journal of Medical and Biological Research, 38,* 995-1014.

Schnappinger, D., Ehrt, S., Voskuil, M. I., Liu, Y., Mangan, J. A., Monahan, I. M., Dolganov, G., Efron, B., Butcher, P. D., Nathan, C., Schoolnik, G. K. (2003). Transcriptional Adaptation of *Mycobacterium tuberculosis* within Macrophages: Insights into the Phagosomal Environment. *J Exp Med, 198,* 693–704.

Sheng, J., Li, J., Sheng, G., Yu, H., Huang, H., Cao, H., Lu, Y. & Deng, X. (2008). Characterization of rpoB mutations associated with rifampin resistance in *Mycobacterium tuberculosis* from eastern China. . *J Appl Microbiol, 105,* 904–911.

Sherman, D. R., Mdluli, K., Hickey, M. J., Barry, C. E. & Stover, C. K. (1999). AhpC, oxidative stress and drug resistance in *Mycobacterium tuberculosis*. *Biofactors, 10,* 211–217.

Shibutani, S., Takeshita, M. & Grollman, A. P. (1991). Insertion of specific bases during DNA synthesis past the oxidation-damaged base 8-oxodG. *Nature, 349,*431–434.

Shuman, S. & Glickman, M. S. (2007). Bacterial DNA repair by nonhomologous end joining. *Nat Rev Microbio , 5,* 852–861.

Sidorenko, V. S., Rot, M. A., Filipenko, M. L., Nevinsky, G. A. & Zharkov, D. O. (2008). Novel DNA glycosylases from *Mycobacterium tuberculosis*. *Biochemistry, 73,*442–450.

Singh, A., Guidry, L., Narasimhulu, K.V., Mai, D., Trombley, J., Redding, K.E., Giles, G.I., Lancaster, J.R. & Steyn, A.J. (2007). *Mycobacterium tuberculosis* WhiB3 responds to $O_2$ and nitric oxide via its [4Fe-4S] cluster and is essential for nutrient starvation survival. *Proc Natl Acad Sci USA.*, *104*, 11562-11567.

Singh, A., Crossman,D. K., Mai, D., Guidry, L., Voskuil, M. I., Renfrow, M. B. & Steyn, A. J.(2009). *Mycobacterium tuberculosis* WhiB3 maintains redox homeostasis by regulating virulence lipid anabolism to modulate macrophage response. *PLoS Pathogens, 5*.

Singleton, M. R., Dillingham, M. S., Gaudier, M., Kowalczykowski, S. C. & Wigley, D. B. (2004). Crystal structure of RecBCD enzyme reveals a machine for processing DNA breaks. *Nature, 432*, 187–193.

Sinha, K. M., Stephanou, N. C., Gao, F., Glickman, M. S. & Shuman, S. (2007). Mycobacterial UvrD1 is a Ku-dependent DNA helicase that plays a role in multiple DNA repair events, including double-strand break repair. *J Biol Chem, 282*, 15114–15125.

Snapper, S.B., Melton, R.E., Mustafa, S., Kieser, T. & Jacobs, W.R., Jr. (1990) Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. *Mol Microbiol 4*, 1911-1919.

Storz, G., Jacobson, F. S., Tartaglia, L. A., Morgan, R. W., Silveira, L. A. & Ames, B. N. (1989). An alkyl hydroperoxide reductase induced by oxidative stress in *Salmonella typhimurium* and *Escherichia coli*: genetic characterization and cloning of *ahp*. *Journal of Bacteriology, 171*, 2049-2055.

Strniste, G. F. & Wallace, S. S. (1975). Endonucleolytic Incision of X-Irradiated Deoxyribonucleic Acid by Extracts of *Escherichia coli*. *Proc Natl Acad Sci USA.*, *72*, 1997–2001.

Suvarnapunya, A. E., Lagassé, H. A. D. & Stein, M. A. (2003). The role of DNA base excision repair in the pathogenesis of *Salmonella enteric* serovar *Typhimurium*. *Mol Microbiol, 48,* 549–559.

Suvarnapunya, A. E. & Stein, M. A. (2005). DNA base excision repair potentiates the protective effect of Salmonella Pathogenicity Island 2 within macrophages. *Microbiology, 151,*557-567.

Thayer, M. M., Ahern, H., Xing, D., Cunningham, R. P. & Tainer, J. A. (1995). Novel DNA binding motifs in the DNA repair enzyme endonuclease III crystal structure. *EMBO J, 14,* 4108– 4120.

Tomb, J. F., White, O., . Kerlavage, A. R., Clayton, R. A., . Sutton, G. G., Fleischmann, R. D.,Ketchum, K.A., Klenk, H.P., Gill, S., Dougherty, B.A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E.F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H.G., Glodek, A., McKenney, K., Fitzegerald, L.M., Lee, N., Adams, M.D., Hickey, E.K., Berg, D.E., Gocayne, J.D., Utterback, T.R., Peterson, J.D., Kelley, J.M., Cotton, M.D., Weidman, J.M., Fujii, C., Bowman, C., Watthey, L., Wallin, E., Hayes, W.S., Borodovsky, M., Karp, P.D., Smith, H.O., Fraser, C.M. & Venter, J.C. (1997). The

complete genome sequence of the gastric pathogen *Helicobacter pylori*. *Nature, 388,* 539-547.

Touati, D. (1988). Molecular genetics of superoxide dismutases. *Free Radicals of Biological Medicine, 5,* 393-402.

Triggs-Raine, B. L. & Loewen, P. C. (1987.). Physical characteristicsof *katG,* encoding catalase HPI of *Escherichia coli*. *Gene, 52* , 121-128.

Ulrichs, T., & Kaufmann, S. H. (2006.). New insights into the function of granulomas in human tuberculosis. *Journal of Pathology, 208,* 261–269.

Voskuil, M.I ., Bartek, I. L., Visconti, K. & Schoolnik, G. K. (2011). The response of *Mycobacterium tuberculosis* to reactive oxygen and nitrogen species. *Front. Microbiol, 2,* 1-12.

Wallace, S. S. (2002). Biological consequences of free radical-damaged DNA bases. *Free Radical Biology & Medicine, 33,*1–14.

Walley, J. D., Khan, M. A., Newell, J. N. & Khan, M. H. (2001). Effectiveness of the direct observation component of DOTS for tuberculosis: a randomised controlled trial in Pakistan. *Lancet, 357,* 664-669.

Warner, D.F., Ndwandwe, D.E., Abrahams, G.L., Kana, B.D., Machowski, E.E., Venclovas, C. & Mizrahi, V. (2010). Essential roles for imuA- and imuB-encoded accessory factors in DnaE2-dependent mutagenesis in *Mycobacterium tuberculosis*. *Proc Natl Acad Sci USA, 107,* 13093-13098.

Warner, D.F. & Mizrahi, V. (2011). Making ends meet in mycobacteria. *Molecular microbiology, 79,* 283-287.

Wayne L.G. & Diaz, G. (1982). Serological, taxonomic, and kinetics studies of the T and M classes of mycobacterial catalase. *International Journal of Systemic Bacteriology, 32,* 296–304.

Weiss, R. B. & Duker, N. J. (1986). Photoalkylated DNA and ultraviolet-irradiated DNA are incised at cytosines by endonuclease III. *Nucleic Acids Research, 14,* 6621–6631.

WHO. (2012). Global Tuberculosis Report. *WHO.*

WHO. (2011). Progress report on the global plan to stop tuberculosis. *Geneva: WHO, Stop TB Dept.*

You, H. J., Swanson, R. L. & Doetsch, P. W. (1998). *Saccharomyces cerevisiae* possesses two functional homologues of *Escherichia coli* endonuclease III. *Biochemistry, 37,* 6033–6040.

Zhang, Y., Heym, B., Allen, B., Young, D. & Cole, S. (1992). The catalase-peroxidase gene and isoniazid resistance of *Mycobacterium tuberculosis*. *Nature, 358,* 591–593.

Zhang, Z., Hillas, P.J. & Ortiz de Montellano, P.R. (1999). Reduction of peroxides and dinitrobenzenes by *Mycobacterium tuberculosis* thioredoxin and thioredoxin reductase. *Arch. Biochem. Biophys, 363,* 19-26.

Zharkov, D. O., Shoham, G. & Grollman, A. P. (2003). Structural characterization of the Fpg family of DNA glycosylases. *DNA repair, 2,* 839-862.

Zwarenstein, M., Schoeman, J. H., Vundule, C., Lombard, C. J. & Tatley, M. (1998). Randomised controlled trial of self-supervised and directly observed treatment of tuberculosis. *Lancet, 352,*1340–1343.

Zwarenstein, M., Schoeman, J. H., Vundule, C., Lombard, C. J. & Tatley, M. (2000). A randomised trial of lay health workers as direct observers for treatment of tuberculosis. *International Journal of Tuberculosis and Lung Disease, 4*, 550–554.