The Molecular Epidemiology of *Mycobacterium tuberculosis* in North London

By

Robert John Shorten

A thesis submitted to University College London in fulfilment of the requirement for the degree of Doctor of Philosophy

Centre for Clinical Microbiology
Department of Infection
University College London
Royal Free Campus
Rowland Hill Street
London NW3 2PF

Declaration

I, Rob Shorten, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed:

Abstract

It is estimated that one third of the world's population are infected with the bacterium *M. tuberculosis* and approximately 9.4 million new cases of tuberculosis were diagnosed globally in 2008. Molecular tools, developed over the previous two decades, have allowed further in-depth study of this historic disease. Genotyping *M. tuberculosis* allows the study of evolutionary relationships and well as the routes of transmission of the organism between hosts. The pairing of genotyping with demographic data allows the analysis of the current trends of disease within a given patient population. Two genotyping methodologies (IS*6110* RFLP and Mycobacterial Interspersed Repetitive Units – Variable Number Tandem Repeats (MIRU-VNTR)), alongside patient demographic data have been utilised for the TB population at the Royal Free Hospital, London. The data shows a patient population that is largely concordant with that seen in London as a whole; with the majority of TB patients being born outside of the UK and drug resistance rates that are higher than the national average.

The acquisition of drug resistance-conferring mutations by *M. tuberculosis* is often presumed to be associated with a fitness cost. Here we investigate the fitness of isolates from two outbreaks involving large numbers of drug resistant strains. The first group of strains was found to be part of the ongoing north London isoniazid resistant outbreak. The data suggests that this outbreak consists of successful, closely related, circulating strains with heterogeneous resistance profiles and mutations and little or no associated fitness cost.

The relationship between the biology of the organisms isolated from a cohort of TB patients who underwent a novel diagnostic interferon gamma release assay was investigated. Genotyping, ESAT-6 gene sequencing and gene expression assays suggest that varying immune response between individuals is driven by host factors rather than a characteristic of the strain of *M. tuberculosis*.

This study shows that this geographic location in London is home to a diverse population of *M. tuberculosis* with a low rate of transmission. The demographic characteristics of the TB patients are largely concordant with the UK as a whole. These findings show that *M. tuberculosis* genotyping is critically important in the identification of suspected outbreaks and contamination events as well as the investigation of successful related strains and characteristic disease phenotypes. Genotyping can augment classic epidemiology and clinical practice to provide a holistic approach in the investigation, treatment and control of the global health problem. Understanding the biology and genomics of *M. tuberculosis* is how the enormous disease burden will be tackled.

Acknowledgements

Warm thanks go to my supervisors, Dr. Tim McHugh and Dr. Claire Jenkins for their continued support, advice and encouragement throughout my PhD. I am additionally grateful to Prof. Stephen Gillespie for his encouragement and input when he supervised my work in the early stages of my research alongside Tim.

I am especially grateful to Simon Rattenbury and the Department of Medical Microbiology, Royal Free Hampstead NHS Trust for funding reagents and fees during my studies. Additional funding was received from the Peter Samuel Fund at the Royal Free Hampstead NHS Trust. I am grateful to the board of trustees of this fund for their generous award.

I would also like to thank a number of colleagues for their support and input to this work. I am grateful to Dr. Ronan Breen for immunological and clinical data, especially in Chapter 4. Dr. Marc Lipman and the TB multidisciplinary team, in particular the TB nurses provided excellent support, for which I am grateful. Dr. Al McGregor worked with me during his Master's degree on Chapter 5. I am grateful for his practical help. Research assistants Selina Bannoo and Laura Wright were extremely helpful in the analysis of MIRU-VNTR samples in Chapter 3, for which I am very grateful. I am also thankful to Dr. Clare Ling, who provided support and warm encouragement during the previous seven years. I am additionally grateful to Andy Thomas for assistance with data analysis and to Dr. Steven Platt at the Department of Bio Informatics, Health Protection Agency Centre for Infections for the minimum spanning tree in chapter 3.

I am grateful to the scientists, and in particular Malcolm Yates, at the Mycobacteria Reference Unit (MRU), Whitechapel. All phenotypic susceptibility data in this work were performed at the MRU. Malcolm was especially helpful in arranging the transport of isolates back to the Royal Free.

I am forever indebted to my family and friends for their support, encouragement and understanding during the preceding seven years. Finally, special thanks must go to Dr. Harriet Unsworth for her unstinting support, encouragement, understanding, tolerance and patience. I could not have done this without you.

Declaration 2
Abstract 3
Acknowledgements 5
List of Tables........ 11
List of Figures........13
Abbreviations........16

Chapter 1: Introduction 19
1.1 Disease 19
1.2 Organism 21
1.3 Genome, Evolution and the Global Phylogeny of M. tuberculosis 22
1.4 Host-Pathogen Interaction 37
1.5 Disease Presentation and Diagnosis 50
1.5.1 Immunological Diagnosis 50
1.5.2 Microscopy 53
1.5.3 Culture 54
1.5.4 Nucleic Acid Amplification Tests 55
1.5.5 Antimicrobial Susceptibility Testing 57
1.5.6 Imaging 58
1.6 TB Epidemiology in UK and London 59
1.7 Treatment, Drug Resistance and Fitness 67
1.7.1 Treatment 67
1.7.2 Modes of Antimicrobial Action and Resistance 67
1.7.3 Fitness 72
1.7.4 Epidemiology of Drug Resistance 74
1.8 TB and HIV 80
1.9 Genotyping 83
1.9.1 Contamination 84
1.9.2 Re-infection vs Relapse 85
1.9.3 Defining the Relatedness of Strains by Genotyping: Clusters, Lineages and Superfamilies 87
1.9.4 Genotyping Methodologies 88
1.9.4.1 IS6110 RFLP 88
1.9.4.2 Mycobacterial Interspersed Repetitive Units – Variable Number Tandem Repeats (MIRU-VNTR) 96
1.9.4.3 Spoligotyping 103
1.9.4.4 Polymorphic GC-rich Repetitive Sequence (PGRS) Genotyping 105
1.9.4.5 Genome Sequencing and Genotyping 108
1.9.5 The Utilisation of Genotype Data 108
1.10 Aims of Thesis 114

Chapter 2. Materials and Methods 116
2.1 M. tuberculosis IS6110 RFLP database 116
2.1.1 Organism Selection and Culture Methods 116
2.1.2 Data collection 117
2.1.3 Preparation of DNA probe by PCR 117

2.1.4 Random labelling of DNA probe 118
2.1.5 Isolation of genomic DNA from M. tuberculosis 119
2.1.6 Restriction enzyme digestion of chromosomal DNA 120
2.1.7 Separation of DNA fragments by electrophoresis and Southern blotting...... 120
2.1.8 Hybridisation and detection. 121
2.1.9 Inputting images into BioNumerics 122
2.2 – Mycobacterial Interspersed Repetitive Units – Variable Number Tandem Repeats (MIRU-VNTR) 123
2.2.1 Isolate selection 123
2.2.2 Isolation of genomic DNA from M. tuberculosis 123
2.2.3 MIRU Polymerase Chain Reaction 123
2.2.4 Fragment analysis of the MIRU PCR products 126
2.2.5 Inputting MIRU Data into BioNumerics. 126
2.3 ESAT-6 and the Host Response 127
2.3.1 Patient selection 127
2.3.2 esxA PCR and Sequencing 127
2.3.3 Sequence Analysis 129
2.3.4 esxA mRNA Reverse Transcriptase PCR 130
2.3.4.1 Evaluation of optimal esxA expression 130
2.3.4.2 RNA Extraction 130
2.3.4.3 RNA Cleanp........132
2.3.4.4 Reverse Transcriptase q-PCR........132
2.4 The Fitness of Over-Represented strains and Acquired Isoniazid Resistance...... 134
2.4.1 Isoniazid Resistance Gene Sequencing 134
2.4.2 Sequence Analysis 136
2.4.3.Fitness Assays 137
2.5 Statistical analysis........137
2.5.1 Chi squared and Fisher's Exact Test........137
2.5.2 One-way Anova with Kruskal-Wallis post test........138
2.5.3 Two-way Anova........139

Chapter 3: Development and Characterisation of an M. tuberculosis database for North London 140
3.1 Introduction........140
3.1.1 Investigating Outbreaks 142
3.1.2 Transmission rates 143
3.1.3 Laboratory contamination 144
3.1.4 Investigating Over-represented Strains 145
3.2 Results 147
3.2.1 Database coverage 148
3.2.2 Database demographics 152
3.2.3 Drug susceptibilities 157
3.2.4 Site of infection 160
3.2.5 Molecular characteristics 160
3.2.6 Laboratory contaminiation rate and transmission events........164
3.2.7 Secondary Typing: MIRU-VNTR Data........165

3.2.7.1 Secondary MIRU-VNTR typing of >70% IS*6110* lineages……………165
3.2.7.2 Secondary MIRU-VNTR typing of IS6110 low copy number strains…171
3.3 Discussion 172
3.3.1 MIRU-VNTR Discussion……………………………………………..185

Chapter 4: ESAT-6 and the Host Response……………………………………190
4.1 Introduction……………………………………………………………190
4.2 Results 200
4.2.1 IFN-γ release assay 200
4.2.2 IS*6110* RFLP genotyping 200
4.2.3 *esx*A PCR and sequencing 201
4.2.4 Reverse transcriptase quantitative polymerase chain reaction (RT qPCR)…202
4.2.4.1 Establishing a reference gene.. 203
4.2.4.2 *esxA* Expression Time Trial 204
4.2.4.3 *esx*A expression by Q RT PCR……………………………………204
4.3 Discusion………………………………………………………………206

Chapter 5: The Relative Fitness of Over-Represented Strains and Acquired Isoniazid Resistance 211
5.1 Introduction.. 211
5.2 Results 216
5.2.1 Lineage 15 217
5.2.2 Extended Family with TB 224
5.3 Discussion 227
5.3.1 Lineage 15 227
5.3.2 Extended Family with TB 231

Chapter 6: The Role of Molecular Epidemiology in Clinical Practice 234
6.1 Introduction.. 234
6.2 Example 1: Paradoxical reaction and Strain Type 234
6.3 Example 2: Suspected Contamination due to a Faulty Bronchoscope 236
6.4 Example 3: Differentiation of relapse and re-infection 239
6.5 Example 4: Genotype and Contract tracing 241
6.6 Discussion 250

Chapter 7: Final Discussion 254

References: ……………………………………………………………………..269

Appendices: Published letters………………………………………………………..318

Perrin, F. M., Breen, R. A., Lipman, M. C., Shorten, R. J., Gillespie, S. H., & McHugh, T. D. 2005, "Is there a relationship between Mycobacterium tuberculosis strain type and TB paradoxical reaction?", Thorax, vol. 60, no. 8, pp. 706-707.

Shorten, R. J., Gillespie, S. H., Sule, O., Lipman, M., & McHugh, T. D. 2005, "Molecular strain typing of M. tuberculosis isolates from a suspected outbreak involving a faulty bronchoscope", J.Hosp.Infect., vol. 61, no. 1, pp. 86-87

List of Tables

Chapter 1. Introduction

1.1 Comparison of the percentage of TB cases as a result of active transmission ...112

Chapter 2. Materials and Methods

2.1 Primers for the amplification of the INS probe for IS6110 RFLP hybridisation…………118
2.2 MIRU primer sets…………124
2.3 MIRU PCR master mixes (1-5)…………...125
2.4 MIRU PCR master mixes (6-8)…………...125
2.5 Primers for the amplification of the *esx*A gene…………...127
2.6 Primers for the amplification of the *esx*A gene plus flanking region…………..128
2.7 Primers and probes for RT-qPCR…………133
2.8 Primers for amplifying isoniazid-conferring mutations…………..135

Chapter 3. Development and Characterisation of an *M. tuberculosis* database for North London

3.1 The age, gender and site of disease of the Royal Free and University College London Hospital patients…………151
3.2 Country of birth of 162 TB patients…………153
3.3 Comparison of pulmonary smear and HIV status…………153
3.4 Comparison of HIV status and ethnicity…………..156
3.5 Comparison of pulmonary smear status and ethnicity…………156
3.6 Comparison of pulmonary smear status and country of birth…………..157
3.7 Break down of antimicrobial resistance…………...158
3.8 Comparison of drug susceptibility and HIV status…………..159
3.9 Comparison of drug susceptibility and ethnicity…………159
3.10 Comparison of drug susceptibility and country of birth…………..159
3.11 Comparison of site of disease and HIV status…………160
3.12 Summary of 28 100% clusters…………162
3.13 Number and size of lineages…………163
3.14 Summary of the six most populated 70% lineages…………..164
3.15 MIRU-VNTR genotypes of two low-copy IS*6110* isolates…………171
3.16 Comparison of patient ages between this data set and (Maguire et al. 2002)…..174
3.17 Comparison of ethnicity in this dataset and that seen nationally…………175
3.18 Summary of the number of isolates in the IS6110 RFLP lineages that were confirmed as being 70% related by MIRU-VNTR…………..184
3.19 Summary of the variability of MIRU-VNTR loci…………...186

Chapter 4: ESAT6 and the Host Response

4.1 Summary of the effects of the products of extRD1……………………………….193
4.2 Summary of the sensitivity and specificity of commercial IGRA and TST in the diagnosis of active TB………………………………………………...197
4.3 Summary of INF-γ responses, PCR results and sequencing data………………202
4.4 Comparison of mean CT values for *esx*A, 16S and *sig*A for three patient isolates………………………………………………………………………..203
4.5 Change in *esxA* expression over time………………………………………..204
4.6 Fold change of *esxA* expression in each clinical change relative to *M. tuberculosis* H37Rv compared to ESAT/PPD IFN-γ response……………..205

Chapter 5: The relative Fitness of Over-Represented Strains and Acquired Isoniazid Resistance

5.1 The six most populated 70% lineages in the RFH database……………………214
5.2 Resistance profiles of strains in lineage RFH-15 and corresponding isoniazid resistance mutations………………………………………………..215
5.3 MIRU-VNTR genotypes of the isolates in lineage 15………………………….216
5.4 Resistance profiles of strains in lineage RFH-15, corresponding isoniazid resistance mutations and relative fitness………………………………………..217
5.5 Fitness data of the strains isolated from the extended family with a high rate of active TB and latent TB infection……………………………………….222

Chapter 6: The Role of Molecular Epidemiology in Clinical Practice

6.1 Summary of MIRU-VNTR type for the four isolates in the suspected bronchoscope outbreak…………………………………………………………238
6.2 15 loci MIRU-VNTR typing results of the two patients……………………….249

Chapter 7: Final Discussion

7.1 The advantages, disadvantages and applications of IS6110, MIRU-VNTR, spoligotyping and whole genome sequencing………………………………….256

List of Figures

Chapter 1: Introduction

1.1 Estimated TB incidence rates, 2008……………………………………………...20
1.2 Mycobacterial cell wall……………………………………………………..23
1.3 A diagrammatic representation of the *M. tuberculosis* H37Rv genome…………25
1.4 Diagram of the evolution of *M. tuberculosis* complex…………………………..29
1.5 Global Phylogeny of *M. tuberculosis*…………………………………………34
1.6 Examples of spoligotype and IS*6110* patterns of Beijing/W strains…………….36
1.7 The possible outcomes of an *M. tuberculosis* infected macrophage…………….38
1.8 a) H&E stained section of lung tissue x100……………………………………..40
b) Ziehl-Neelsen stained section of lung tissue x400…………………………41
1.9 The interaction of the infected antigen-presenting cell and an antigen specific T cell after infection…………………………………………………43
1.10 NICE guidance: close contacts of all cases of active TB………………………..52
1.11 Chest x-ray demonstrating cavitation…………………………………………60
1.12 CT scan of the same patient in figure 1.11……………………………………61
1.13 TB notifications from 1913 to 2008……………………………………………..62
1.14 Tuberculosis case reports and rates, UK, 2000-2008…………………………....63
1.15 Tuberculosis case reports and rates by region, England, 2008…………………64
1.16 Three-year average tuberculosis case rates by primary care organisation, UK, 2006-2008……………………………………………………………………..65
1.17 MDR-TB among new cases 1994-2007………………………………………75
1.18 XDR-TB among MDR-TB cases 2002-2007…………………………………77
1.19 Proportion of tuberculosis cases with first-line drug resistance, UK, 2000-2008.78
1.20 HIV notifications in the UK…………………………………………………...82
1.21 IS*6110* RFLP of sequential isolates from an individual patient…………………86
1.22 Physical map of IS*6110*……………………………………………………….90
1.23 Image of the IS*6110* RFLP genotyping patterns of multiple clinical strains……91
1.24 Frequency distribution of number of copies of IS6110 for 2490 clinical isolates of *M. tuberculosis* in London between 1995 and 1997…………………92
1.25 Location of MIRU loci in *M. tuberculosis* H37Rv genome……………………..98
1.26 12-loci MIRU amplicons on an agarose gel with ethidium bromide staining….101
1.27 GeneMapper Fragment analysis………………………………………………..102
1.28 Hybridization patterns (spoligotypes) of amplified mycobacterial DNAs of 35 *M. tuberculosis* and 5 *M. bovis* strains…………………………………...104
1.29 Representation of IS*6110*, MIRU and DR regions within the *M. tuberculosis* genome………………………………………………………...107

Chapter 2. Materials and Methods

2.1 Flow diagram of isolates genotyped and included in subsequent analysis…….117

Chapter 3. Development and Characterisation of an *M. tuberculosis* database for North London

3.1 Greater London boroughs……………………………………………………146
3.2 Example of *IS6110* RFLP image……………………………………………….149
3.3 Example of IS*6110* RFLP dendrogram………………………………………….150
3.4 Notifications of TB from 2002 – 2007…………………………………………150
3.5 Distribution of patients by age…………………………………………………152
3.6 Distribution of patients by ethnicity……………………………………………154
3.7 Site of disease……………………………………………………………………..155
3.8 Distribution of strains with varying numbers of IS*6110*……………………….161
3.9 MIRU-VNTR dendrogram of IS*6110* RFLP lineage 15……………………….166
3.10 MIRU-VNTR dendrogram of IS*6110* RFLP lineage 19……………………….167
3.11 MIRU-VNTR dendrogram of IS*6110* RFLP lineage 41……………………….168
3.12 MIRU-VNTR dendrogram of IS*6110* RFLP lineage 43……………………….168
3.13 MIRU-VNTR dendrogram of IS*6110* RFLP lineage 54……………………….169
3.14 MIRU-VNTR dendrogram of IS*6110* RFLP lineage 61……………………….169
3.15 Minimum spanning tree (MST) of all isolates typed by MIRU-VNTR………..170
3.16 MIRU-VNTR dendrogram of all low IS*6110* copy number strains……………172

Chapter 4: ESAT6 and the Host Response

4.1 The genes present in the RD-1 along with the flanking genes are shown as an extended RD-1 (extRD-1) locus of *M. tuberculosis*………………………...192
4.2 Dendrogram showing the relatedness of the patient isolates…………………...201
4.3 Regression analysis plotting relative change in *esxA* expression and ESAT-6:PPD response…………………………………………………………….206

Chapter 5: The relative Fitness of Over-Represented Strains and Acquired Isoniazid Resistance

5.1 Dendrogram of IS6110 fingerprints of strains within lineage 15………………213
5.2 MIRU-VNTR dendrogram of lineage 15………………………………………215
5.3 Generation times of the isolates in lineage 15………………………………….218
5.4 Split tree analysis - Relationship of the strains in lineage 15…………………..219
5.5 Family tree of extended family with a high rate of TB…………………………221
5.6 Generation times of the isolates from the extended family…………………….222

Chapter 6: The Role of Molecular Epidemiology in Clinical Practice

6.1 Dendrogram of patient strains isolated from patients with paradoxical reaction……….236
6.2 The IS*6110* RFLP genotypes of the isolates obtained from the patient on two occasions, approximately 17 months apart ……….240
6.3 Chest x-ray three weeks after presentation……….242
6.4 Chest CT three weeks after presentation……….243
6.5 Chest x-ray at the completion of treatment……….244
6.6 Chest x-ray of male contact at day 8 of the patient's treatment……….245
6.7 Male contact's chest x-ray 10 weeks after his partner had started anti-TB treatment……….246
6.8 Chest CT of male contact at week 12 of the patient's treatment……….247
6.8(a) Magnified area of CT in the left lower zone of figure 6.8……….248
6.9 NICE guidance for contact tracing……….250

Abbreviations

A	Adenine
AFB	acid fast bacilli
APC	antigen presenting cell
Arg	arginine
ATS	American Thoracic Society
BHIVA	British HIV Association
BMI	body mass index
BTS	British Thoracic Society
C	Cytosine
CD	Cluster of Differentiation
CDC	Centers for Disease Control
CFP10	culture filtrate protein 10
CRP	C reactive protein
CT	computerised tomography
DC	dendritic cell
DLV	double locus variant
DNA	Deoxyribose nucleic acid
DOT	directly observed therapy
DR	Direct Repeat
DVR	direct variant repeats
ESAT6	Early Secretory Antigen Target 6
ETR	exact tandem repeats
G	Guanine
H&E	hematoxylin and eosin
HAART	highly active antiretroviral therapy
HIV	Human Immunodeficiency Virus
HLA	human leucocyte antigens
HPA	Health Protection Agency
IFNγ	interferon gamma
IGRA	interferon gamma release assay
IL	interleukin
IRIS	immune reconstitution inflammatory syndrome
IS	Insertion sequence
kDa	kiloDalton
LAM	lipoarabinomannan
LED	light emitting diode

Leu	leucine
LSP	Large sequence polymorphism
M. bovis BCG	Mycobacterium bovis Bacillus Calmette-Guérin
MDR	multidrug resistant
MGIT	Mycobacteria growth indicator tube
MHC	major histocompatibility complex
MIC	Minimum Inhibitory Concentration
MIRU-VNTR	Mycobacterial Interspersed Repetitive Units – Variable Number Tandem Repeats
MODS	Microscopic-observation drug-susceptibility
MPTR	Major Polymorphic Tandem Repeats
MRU	Mycobacteria Reference Unit
NAAT	nucleic acid amplification technique
NADPH	nicotinamide adenine dinucleotide phosphate
NHS	National Health Service
NICE	National Institute for Health and Clinical Excellence
NK	natural killer cells
NMRL	National Mycobacteria Reference Laboratory
nsSNP	non-synonymous single nucleotide polymorphism
PAMP	pathogen-associated molecular pattern
PAS	para-aminosalicylic acid
PCR	polymerase chain reaction
PCT	Primary Care Trust
PE	proline-glutamate
PGL	phospohoglycolipid
PGRS	polymorphic GC-rich repetitive sequence
PPD	purified protein derivative
PPE	proline-proline-glutamate
QFT-G	QuantiFERON-TB Gold
RD	region of difference
rDNA	ribosomal deoxyribose nucleic acid
RFLP	Restriction Fragment Length Polymorphism
RNA	ribose nucleic acid
RRDR	rifampicin resistance determining region
rRNA	ribosomal ribose nucleic acid
SDA	strand displacement amplification
Ser	serine
sfam	superfamily

SLV	single locus variant
SNP	Single nucleotide polymorphism
spoligotyping	spacer-oligonucleotide typing
sSNP	synonymous single nucleotide polymorphism
T	Thymine
TB	Tuberculosis
TbD1	*M. tuberculosis* specific deletion 1
TCR	T-cell receptor
Th	helper T cells
Thr	thrionine
TLR	toll-like receptor
TLV	triple locus variant
TNFα	tumour necrosis factor alpha
TST	Tuberculin skin test
WHO	World Health Organisation
XDR	extensively drug resistant
ZN	Ziehl-Neelsen

Chapter 1: Introduction

1.1 Disease

Tuberculosis (TB) is an ancient disease that currently presents an immense global health challenge. It is estimated that one third of the world's population (approximately 2 billion people) are infected with the tubercle bacilli. However, infection by the organism does not necessarily lead to disease and only 5-10% of these individuals will progress to active disease each year (WHO 2007). The remaining 90% of infected individuals will initially be asymptomatic and experience latent infection, from which reactivation may occur. The World Health Organisation (WHO) estimate that globally in 2008 there were 9.4 million cases of active TB leading to 1.3 million deaths (WHO 2009).

The geographic incidence of infection and therefore disease burden varies greatly. Most of the estimated number of cases in 2008 occurred in Asia (55%) and Africa (30%). The 22 high-burden countries (in terms of absolute numbers of cases) account for 80% of all estimated cases worldwide. India and China alone account for an estimated 35% of TB cases worldwide (WHO 2009). Widely varying incidence rates are demonstrated in figure 1.1.

Living in or having visited a high incidence area are key risk factors, but host factors also play a role in the risk of patients progressing to active disease. The immunosuppressed, especially those individuals that are infected with HIV are particularly susceptible to TB disease. Individuals who have poor nutrition and general bad health, such as the homeless and those who misuse alcohol and drugs are also at increased risk. Other

immunocompromised patients, such as patients with chronic renal disease, neoplastic disorders and those receiving immunosuppressive therapy are also susceptible. The vaccination status of the individual may also play a role.

The HIV pandemic has greatly altered the global landscape of TB. WHO estimate that 1.2–1.6 million (13–16%) of the 9.4 million cases in 2008 were HIV-positive. Of these HIV-positive cases, 78% were in the African Region and 13% were in the South-East Asia Region (WHO 2009). Individuals infected with HIV are more than 20 times more likely to reactivate TB than those who are HIV negative (Girardi E et al. 2000).

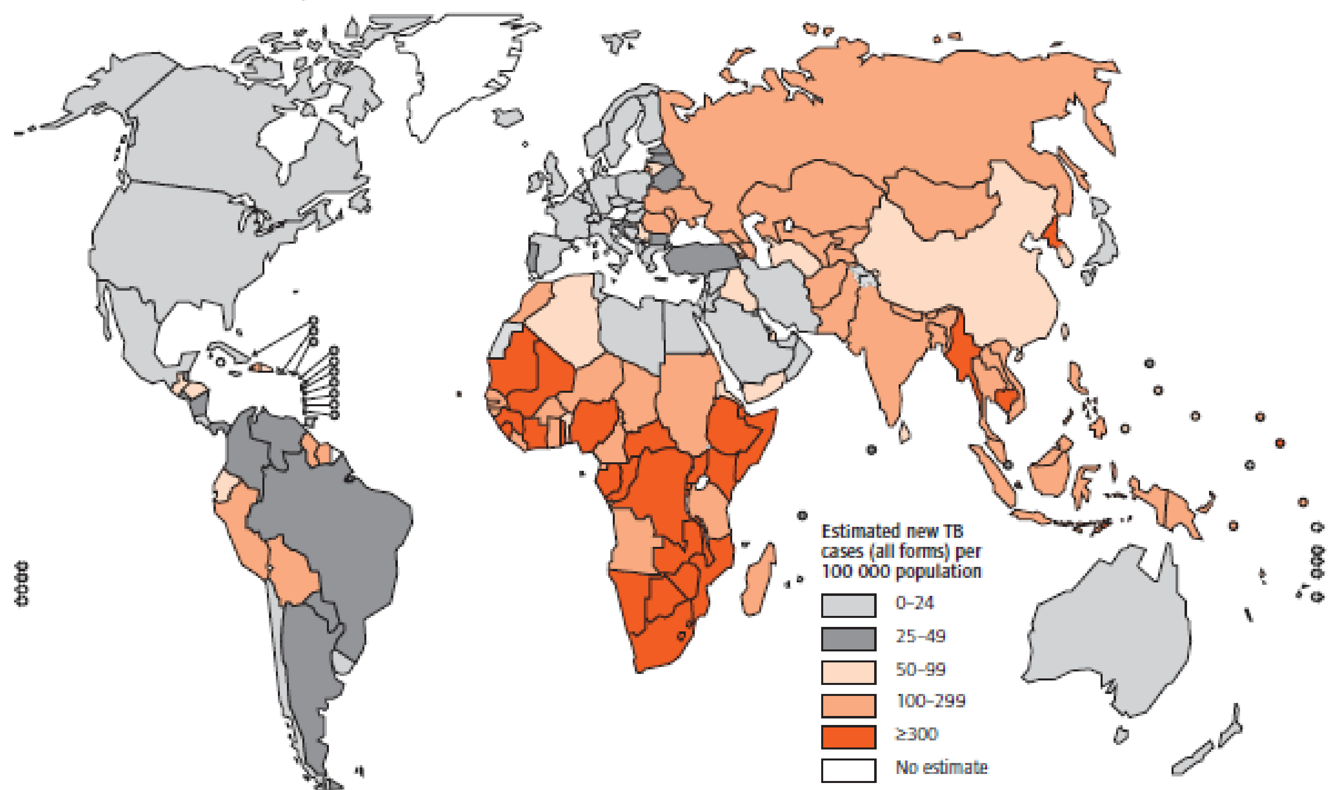


Figure 1.1. Estimated TB incidence rates, 2008 (WHO 2009)

TB in humans is caused predominantly by *Mycobacterium tuberculosis,* but also by *M. africanum* and *M. bovis*. These species, together with *M. microti, M. canettii* and other animal-adapted *M. bovis* ecotypes make up the *M. tuberculosis* complex. TB primarily affects the lungs but can also infect the central nervous system, lymph nodes, bones, joints, urinary tract and other sites. The disease is infectious via the respiratory route by inhalation of aerosolised bacilli from patients with active respiratory disease.

1.2 Organism

M. tuberculosis belongs to the order Actinomycetales. The organism is a non-motile, non spore forming, slowly reproducing aerobe. Slow cell division, resistance to detergents and certain antimicrobial agents, persistence in the environment and an acid/alcohol fast nature are due to the characteristic hydrophobic cell wall (figure 1.2). Working from the bottom layer of the envelope towards the external surface of the bacterium there is a triple plasma membrane thought to be similar to those in other bacteria. Next follows a cell wall skeleton consisting of peptidoglycan, complexed with mycolic acids unique to mycobacteria and arabinogalactan, which forms a hydrophobic permeability barrier. Transport proteins in the plasma membrane and porins in the mycolic acid layer allow for the import of hydrophilic substances. Above this layer lies a capsule (or pseudocapsule as it is not covalently bound to the rest of the envelope) consisting of polysaccharides, proteins and lipid. (Draper and Daffe 2005).

The lipid rich cell wall does not take up stain readily, but resists decolourisation when destained with an acid-alcohol wash. Mycobacteria are often stained with fluorescent

stains, such as auramine-O (a diamylmethane dye) or with more conventional stains for light microscopy, such Ziehl-Neelsen or Kinyoun using carbol fuchsin.

1.3 Genome, Evolution and the Global Phylogeny of *M. tuberculosis*

The complete genome of *M. tuberculosis* H37Rv was published in 1998 (Cole ST et al. 1998). This revealed a circular genome in the region of 4.4 million base pairs containing approximately 4000 genes – see figure 1.3. The genome of *M. tuberculosis* is GC rich (65.6%), a feature that has been associated more often with aerobic prokaryotes (Naya et al. 2002). This was updated four years later and the function of 2058 genes (52%) was predicted (Camus et al. 2002).

It was demonstrated that *M. tuberculosis* possesses all the genes necessary for synthesising essential amino acids, vitamins and enzyme co-factors. It was noted that there are a high proportion of genes encoding enzymes involved in lipogenesis and lipolysis. Additionally, *M. tuberculosis* possesses the genes to synthesise glycolytic enzymes and enzymes necessary for the anabolic pentose phosphate pathway, which generates NADPH and pentose sugars, the catabolic Kreb's cycle and the glyoxylate cycle, which synthesises carbohydrates from lipids. The tubercle bacillus also possesses enzymes used in aerobic, microaerophilic and anoxic electron transfer and is capable of surviving in a number of different environments including the oxygen rich lung, the macrophage and at the centre of caceous granuloma (Cole ST et al. 1998).

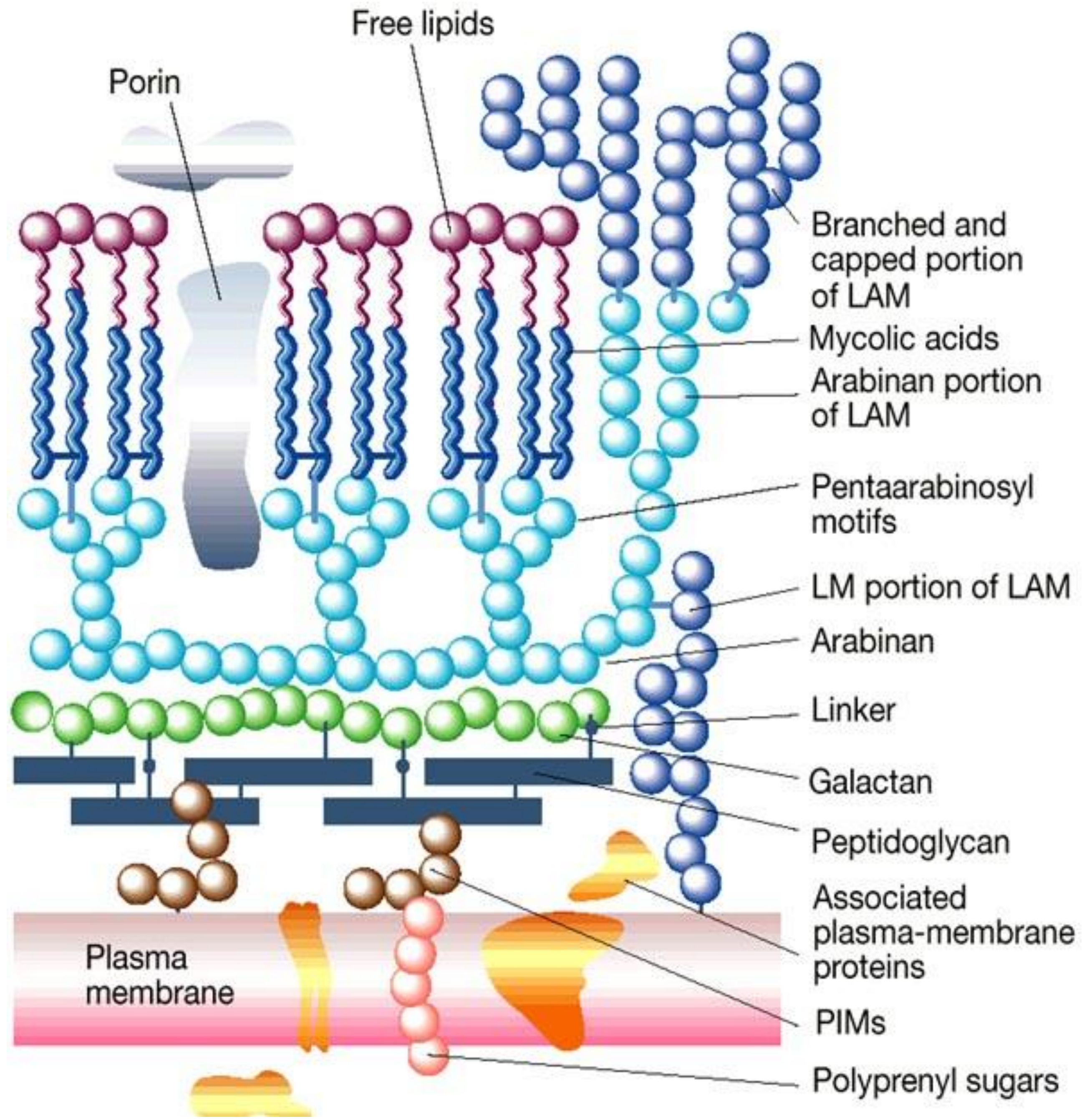


Figure 1.2 Mycobacterial cell wall (Park & Bendelac 2000)

There are a number of repetitive DNA sequences in the *M. tuberculosis* genome including insertion sequences (IS), the direct repeat (DR) region, the major polymorphic tandem repeats (MPTR) and the polymorphic GC-rich repetitive sequence (PGRS) (Poulet & Cole ST 1995b).

The genome contains two large protein families encoding acidic, asparagine- or glycine-rich proteins referred to PE (proline-glutamate – n = 99) and PPE (proline- proline-glutamate – n = 68) (Cole ST et al. 1998). Some of these genes are situated in the MPTR and the PGRS regions and are associated with antigenic variability and virulence (Karboul et al. 2006;Poulet & Cole ST 1995a;Poulet & Cole ST 1995b).

Previous studies of clinical and laboratory strains have shown that the *M. tuberculosis* complex possesses a high degree of sequence homogeneity, demonstrated by a greater than 99.9% similarity in rDNA sequences (Frothingham, Hills, & Wilson 1994; Sreevatsan et al. 1997). Sequence homology extends even between different species of the *M. tuberculosis* complex. There is less than 0.05% sequence divergence between *M. bovis* and *M. tuberculosis* (Garnier et al. 2003). This genetic similarity may be viewed as significant when compared to other species, for example, two strains of *Escherichia coli* have shown sequence diversity of 1.6% (Perna et al. 2001).

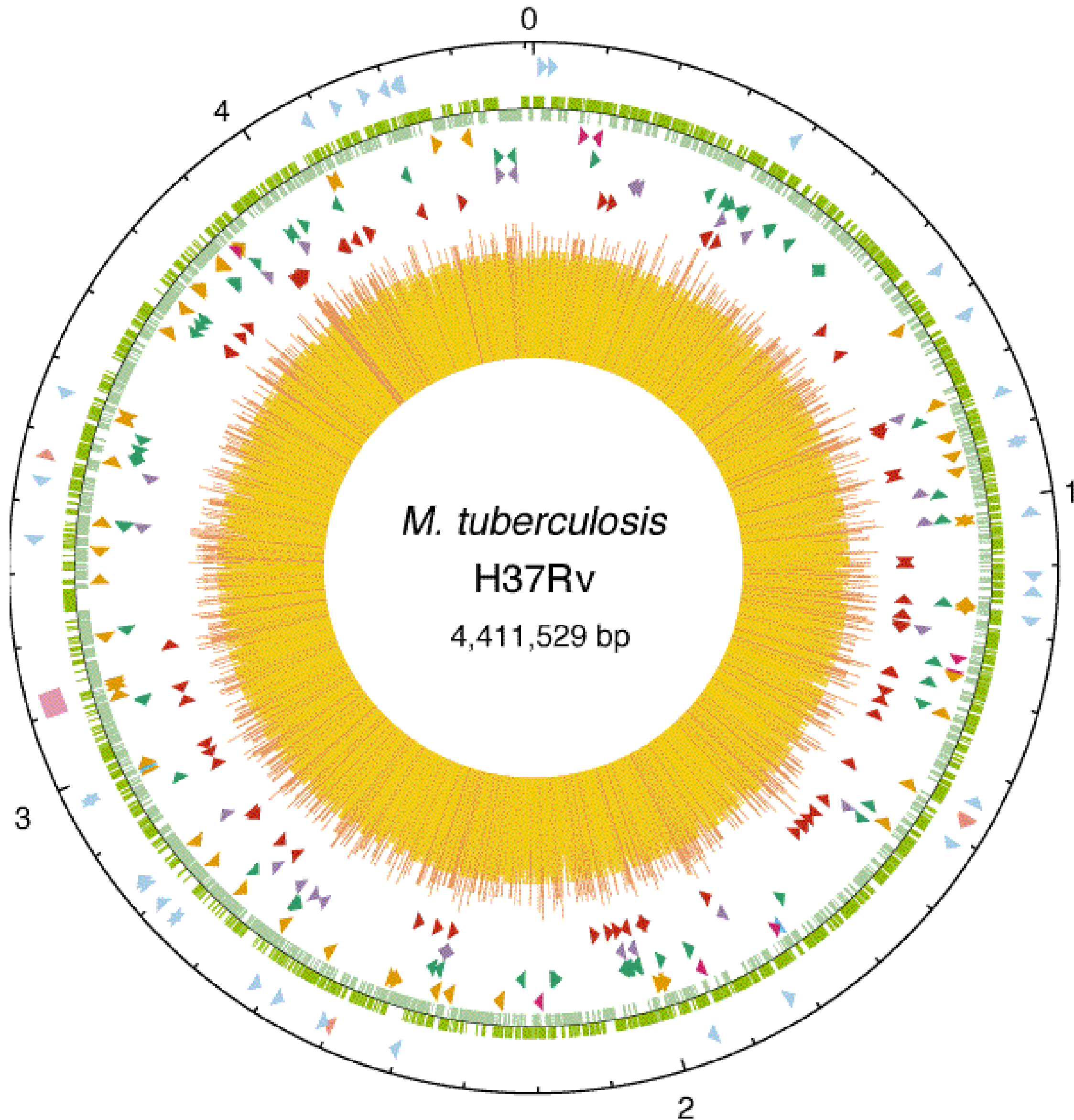


Figure 1.3. A diagrammatic representation of the *M. tuberculosis* H37Rv genome. The outer circle shows the scale in Mb, with 0 representing the origin of replication. The first ring from the exterior denotes the positions of stable RNA genes (tRNAs are blue, others are pink) and the direct repeat region (pink cube); the second ring inwards shows the coding sequence by strand (clockwise, dark green; anticlockwise, light green); the third ring depicts repetitive DNA (insertion sequences, orange; 13E12 REP family, dark pink; prophage, blue); the fourth ring shows the positions of the PPE family members (green); the fifth ring shows the PE family members (purple, excluding PGRS); and the sixth ring shows the positions of the PGRSsequences (dark red). The histogram (centre) represents G + C content, with $<65\%$ G + C in yellow, and $>65\%$ G + C in red (Cole ST et al. 1998).

This lack of variation in *M. tuberculosis* may have resulted from limited selective pressure or the hypothesis that its distribution is very recent, in evolutionary terms. Indeed, some groups hypothesised that the lack of genetic variation meant that *M. tuberculosis* was younger than 20,000 years old (Fleischmann et al. 2002; Frothingham 1999; Frothingham, Hills, & Wilson 1994; Kapur et al. 2000; Sreevatsan et al. 1997). In addition, there is little evidence of horizontal gene transfer in *M. tuberculosis* and so its evolution appears to have occurred in a clonal manner (Filliol et al. 2006; Hirsh et al. 2004; Supply et al. 2003). With the exception of *M. canetti* (Gutierrez et al. 2005) there is limited evidence for genetic recombination in the *M. tuberculosis* complex (Hughes AL, Friedman R, & Murray M 2002; Liu et al. 2006). Indeed, *M. tuberculosis* exhibits relatively low levels of genetic recombination, similar to other clonal organisms such as *Escherichia coli* and *Salmonella enterica* (Smith et al. 2003) when compared to the heterogeneity in housekeeping genes other organisms, such as *H. pylori* (Go et al. 1996; Suerbaum et al. 1998; Feil & Spratt 2001; Linz & Schuster 2007) and *Neisseria gonorrhoeae* (O'Rourke & Stevens 1993; O'Rourke & Spratt 1994;).

Although there is limited genetic diversity between strains, the differences that have been observed have been useful in the investigation of the global evolution of the *M. tuberculosis* complex. As *M. tuberculosis* rarely benefits from horizontal gene transfer, deletions and point mutations can be studied to track the global evolution of the organism. Due to the non-promiscuous nature of *M. tuberculosis*, the alterations within its genome are as a result of deletions or point mutations. Deletions in *M. tuberculosis* can be generated in three ways; 1) mobile genetic elements, such as insertion sequences,

2) homologous recombination of insertion sequences resulting in the deletion of intervening DNA and 3) deletions whose bordering genetic regions do not contain repetitive sequences (Brosch et al. 2002).

Further investigation into significant genes directed the study of the organism's evolution. The *kat*G gene encodes catalase peroxidae, and mutations in this gene can confer resistance to isoniazid in *M. tuberculosis* – see section 1.7. Mutations in the *gyr*A gene, which encodes DNA gyrase, may confer resistance to fluoroquinolones. Analysis of these genes implicated in drug resistance in over 800 geographically diverse isolates from the *M. tuberculosis* complex by Sreevatsan and colleagues identified three distinct groups of strains (Sreevatsan et al. 1997). It was surmised that the common precursor organism of the *M. tuberculosis* complex possessed codon *kat*G^{463} CTG (Leu) and *gyr*A^{95} ACC (Thr). Members of the *M. tuberculosis* complex that possessed these codons were considered to be the oldest and members were termed to be in Principal Group 1 (Sreevatsan et al. 1997). This group included some strains of *M. tuberculosis* as well as *M. africanum* and *M. microti*. The mutation to *kat*G^{463} C<u>G</u>G (Arg) allocated the strain into Principal Group 2, whilst the presence of that mutation together with the mutation *gyr*A^{95} A<u>G</u>C (Ser) placed the strain in Principal Group 3 (figure 1.4). This group then analysed approximately 6000 strains from Houston and New York and found no strains belonging to Group 3. It was reasoned that this group was either the most recently evolved and/or was less virulent than the strains in Groups 1 and 2.

The analysis of deletions further allowed the evolution of the *M. tuberculosis* complex to be mapped. Comparative genomic analysis showed 14 regions of difference (RD1-14), ranging in size from 2 to 12.7kb, that were present in *M. tuberculosis*, but absent in *M. bovis* BCG (Behr et al. 1999; Gordon et al. 1999). In addition, six deletions were identified in the reference strain *M. tuberculosis* H37Rv compared with other members of the *M. tuberculosis* complex (RvD1-5 and the *M. tuberculosis* specific deletion 1 (TbD1)) (Brosch et al. 1999; Gordon et al 1999). Geographically distributed isolates from the *M. tuberculosis* complex were examined for the presence or absence of TbD1, a number of regions of difference (RD) as well as single nucleotide polymorphisms (SNPs) in the gene sequences of *kat*G, *gyr*A, *oxy*R, *pnc*A and *mmp*l6 (Brosch et al. 2002). Results showed that *M. tuberculosis* as well as other members of the *M. tuberculosis* complex had evolved from a common precursor (figure 1.4). Strains that possessed TbD1 were termed 'ancient' whilst those lacking the region were labelled 'modern'. Additionally, the assumption that *M. tuberculosis* had evolved from *M. bovis* was discredited. This finding was also confirmed by Mostowy and colleagues (Mostowy et al. 2002).

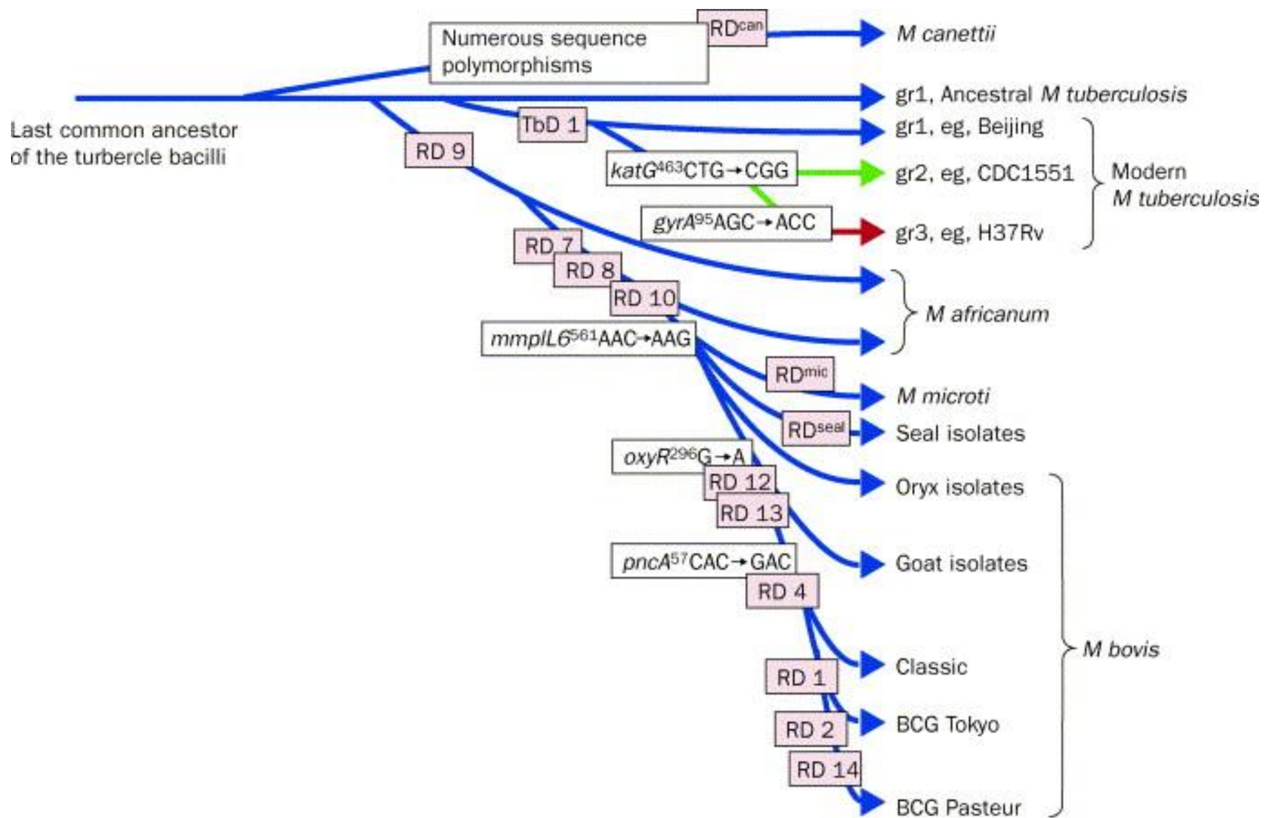


Figure 1.4 Diagram of the evolution of *M. tuberculosis* complex. Principle Groups 1, 2 and 3 are shown towards the top of the figure and are indicated by blue, green and red arrows, respectively. (Donoghue et al. 2004).

Further analysis of well-characterised laboratory strains, as well as clinical isolates has contributed to the study of the evolution of the *M. tuberculosis* complex. In addition to the presence or absence of TbD1 and the SNPs at *kat*G^{463} and *gyr*A^{95}, other SNPs and long sequence polymorphisms (LSPs) have been investigated. Baker and colleagues (Baker et al. 2004) investigated SNPs in *kat*G and *gyr*A, as well as a further five genes (*rpo*B, *oxy*R, *ahp*C, *pnc*A, and *rps*L). As sSNPs are neutral, they become fixed in a bacterial population and may be used to study its evolution.

The analysis of 37 neutral sSNPs allowed a phylogenetic tree to be constructed that divided the strains into four distinct groups (I – IV) (Baker et al. 2004). The analysis of *kat*G^{463} and *gyr*A^{95} showed congruence between the groups found in this study and the work by Sreevatsan and colleagues (Sreevatsan et al. 1997); Groups I, III and IV aligned with Principal Group 1. Group IV (along with *M. bovis, M. microti* and *M. africanum*) possessed TbD1, confirming that the deletion of TbD1 preceded the SNPs at *kat*G^{463} and *gyr*A^{95}. Group II strains aligned with Principal Groups 2 and 3. The observation that common nsSNPs in *rpo*B, *kat*G, *rps*L and *inh*A, causing antimicrobial resistance, were present in all groups indicates that these evolved independently due to selective pressure of drug treatment (Baker et al. 2004).

Epidemiological data showed that there were strong associations between the groups and the country of birth of the patients. Groups I, II and III were associated with South East Asia, Europe and the Indian subcontinent, respectively. Group IV strains were globally disseminated, but were negatively associated with Europe. A large proportion of these Group IV strains possessed only one copy of IS*6110* and were more likely to be drug susceptible. This, combined with the presence of TbD1 strongly suggested that this group was more closely related to the common ancestor than the other groups (Baker et al. 2004).

Filliol and Gutacker have used SNPs to further investigate the global phylogeny of *M. tuberculosis* (Filliol et al. 2006; Gutacker et al. 2006). Both studies allocated *M. tuberculosis* isolates into nine lineages that are largely congruent with others' findings in

the regards of the Principal Groups 1-3, based on SNPs at *kat*G^{463} and *gyr*A^{95}, other SNPs, LSPs and the country of origin of the patient.

Gangeux and colleagues demonstrated similar findings based on LSPs and SNPs. They made strong associations between each of the six characterised groups and geographic region; Indo-Oceanic, East Asian, East African/Indian, Euro-American, West African-1 and West African-2 (Gagneux et al. 2006). When investigating the diverse population of San Francisco, it was seen that >99% of strains were assigned to only three of the six groups: Indo-Oceanic, East Asian and European-American. Secondary cases in each of these groups were studied and it was shown that patients infected with a strain in the European-American group were three times more likely to give rise to a secondary case. Additionally, sympatric infections were far more common than allopatric infections. For example, US-born patients who defined themselves as Chinese were far more likely to be infected with a strain that was associated with patients born in China. However, allopatric infections were disproportionately seen in more susceptible hosts. For example, US-born TB patients of non-Chinese and non-Philipino ethnicity who were infected with Indo-Oceanic or East Asian strains were more likely to be HIV positive or homeless (Gagneux et al 2006).

Hirsh and colleagues studied unclustered isolates in San Francisco. Strains isolated from individuals from distinct geographical regions were genotypically linked regardless of whether the researchers believed that the individual had contracted the infection in the US or had reactivated disease from their home country (Hirsh et al. 2004). The strong

associations between phylogeographic lineage and a patient's country of origin have also been demonstrated by others (Dale et al. 2003b; Reed et al. 2009). These observations raise the hypothesis that some strains of *M. tuberculosis* may have adapted to infect individuals with a certain genetic background. For example, different human leucocyte antigens (HLA) of the major histocompatibility complex (MHC) have been associated with TB. A meta analysis of studies investigating HLA types and thoracic TB found a decreased risk of TB in carriers of B13, DR3 and DR7 antigens. Carriers of DR8 had an increased risk of thoracic TB (Kettaneh et al. 2006). Additionally, vitamin D deficiencies have been associated with TB. A meta analysis found that vitamin D levels were significantly lower in TB patients than in controls (Nnoaham & Clarke 2008). Associations between vitamin D receptor genotype and the response to TB treatment have also been made (Chocano-Bedoya & Ronnenberg 2009). A recent study showed, however, that the supplementation of anti-TB treatment with vitamin D did not improve outcome. The exception was individuals with the *tt* genotype of the *TaqI* vitamin D receptor polymorphism who became sputum culture negative more rapidly than controls (Martineau et al 2011). However, broader studies to separate the effect of social and epidemiological factors from the genetic background of the patient necessarily include large numbers of patients and few have been performed.

The sum total of these five studies (Sreevatsan, Baker, Filliol, Gutacker and Gagneux), together with the analysis of an extremely large spoligotyping (section 1.9.4.3.) database of almost 40,000 strains (Brudey et al. 2006) gives rise to the current understanding of the global phylogeny of *M. tuberculosis*. The evolution of the organism appears to have

occurred in a clonal manner with no evidence of horizontal gene transfer. The six main groups of strains are characterised by SNPs and deletions (LSPs) and are strongly associated with geographic location (figure 1.5; Gagneux & Small 2007).

Additionally, evolution of *M. tuberculosis* complex alongside that of humans has been suggested in some of these works. Gutierrez and colleagues suggest that an ancient ancestor of *M. tuberculosis* complex, '*M. prototuberculosis*' co-evolved with early hominids between 2.6 and 2.8 million years ago (Gutierrez et al. 2005). More recent evolution of *M. tuberculosis* complex can be mapped alongside the movement of humans from east Africa approximately 50,000 years ago (Hershberg et al. 2008). Wirth and colleagues suggest that the oldest common ancestor for the *M. tuberculosis* complex emerged from its progenitor in east Africa approximately 40,000 years ago (Wirth et al. 2008). Hershberg and colleagues further aligned the evolution of the *M. tuberculosis* complex into the six phylogeograpic lineages alongside worldwide human migrations and trade routes over the past few hundred years (Hershberg et al. 2008). These findings, along with the predominance of sympatric infection suggest a co-evolution between pathogen and host.

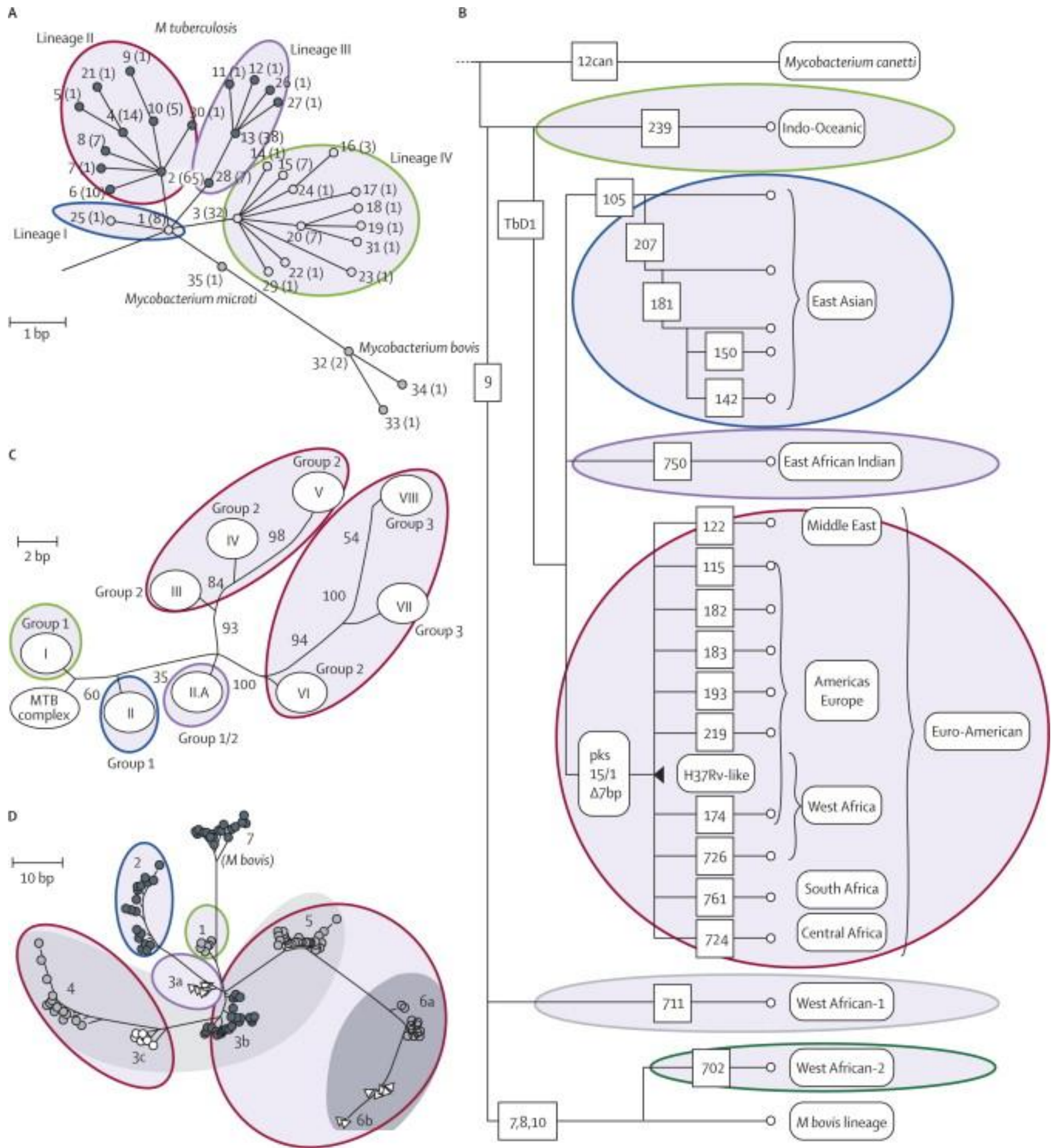


Figure 1.5 – Global Phylogeny of *M. tuberculosis*. A: Groups I-IV adapted from Baker 2004. B six phylogeographic groups adapted from Gagneux 2007. C: Nine groups with corresponding Principal Genetic Groups 1-3 (Srrevatsan 1997) adapted from Gutacker 2006. D: nine groups and *M. bovis* adapted from Filliol 2003. Coloured areas indicate corresponding groups of strains (Gagneux & Small 2007).

One group of strains has been investigated more than any other. The predominance of a family of strains, based on spoligotyping, was seen in the Beijing area of China (van Soolingen et al. 1995). Representatives of the Beijing family are now seen worldwide and have been intensively investigated. They predominate over all of East Asia, Russia and the countries of the former Soviet Union (Bifani et al. 2002). A review of worldwide literature shows prevalence rates in China between 1956 and 1999 of 83-100%, and virtually no geographic location is without representatives of this lineage (Glynn et al. 2002).

The pathogenesis, host immune response and disease progression have also been studied. Increased virulence as well as early, but short lived TNF-α induction in mouse models has been demonstrated (Lopez et al. 2003). The organisms' ability to divide rapidly in human macrophages (Zhang et al. 1999) and the observation of a highly mutable phenotype have also been documented (Ebrahimi-Rad et al. 2003). The Beijing lineage, together with closely related MDR strain 'W' first identified in New York, are characterised first by the deletion of TbD1. These strains are therefore described as 'modern'. Additionally, they are placed in Principal Group 1 (Sreevatsan et al. 1997). Further differentiation is characterised by the spoligotype S000034 (these strains posses only the spacers 35-43). There is a diverse range of IS*6110* patterns that fall within the Beijing/W lineage, but all possess large numbers of copies of the insertion sequence (15-26) (van Soolingen et al. 1995) – figure 1.6.

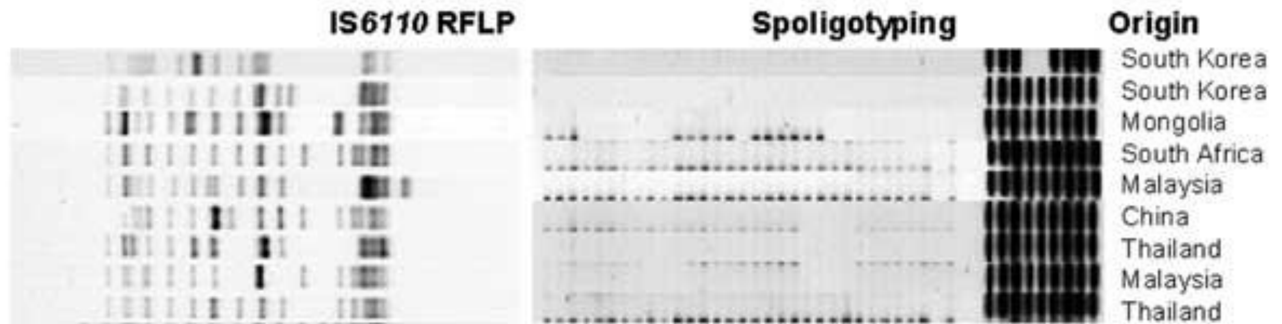


Figure 1.6 Examples of spoligotype and IS*6110* patterns of Beijing/W strains (Ebrahimi-Rad et al. 2003)

Beijing/W strains have been implicated in outbreaks and have sometimes been associated with drug resistance and HIV patients (Glynn et al. 2002; Moss et al. 1997). Drug resistance and radiologically advanced disease has also been associated with Beijing strains by Drobniewski and colleagues (Drobniewski et al 2005). Further characterisation of Beijing strains has been demonstrated by the examination of LSPs. Tsolaki and colleagues demonstrated that all Beijing/W strains had RD105 deleted, and these strains were further subdivided into four groups by the subsequent deletion of RDs 181, 150 and 142. It was noted that all strains retained an intact gene *pks*15/1, which is involved in phenolic glycolipd biosynthesis, and is a putative virulence factor (Reed et al. 2004b; Tsenova et al. 2005; Tsolaki et al. 2005). Additionally, Reed and colleagues found that Beijing/W strains accumulate large quantities of immunoregulatory triacylglycerides and have a dormancy advantage over other strains (Reed et al. 2007). Finally, the suspected route and time-scale of the co-migration and evolution of the Beijing/W lineage and humans from Africa has been described by comparing the bacterial genome to the non-recombining portion of the human Y chromosome. The migration of humans and *M. tuberculosis* from Africa to Asia approximately 45,000 years ago was concluded from this analysis (Mokrousov et al. 2005).

As well as Beijing/W, associations have been made between other genotypes and progression to active disease (de Jong et al. 2008), pulmonary cavitation (Kato-Maeda et al. 2001), host immune response (Lopez et al. 2003; Theus, Cave, & Eisenach 2005), growth rates in ex-vivo monocytes (Theus, Cave, & Eisenach 2005) and macrophages (Zhang JID 1999), the ability to cause disseminated disease (Caws et al. 2008; Garcia de Viedma et al. 2005; Thwaites et al. 2008) and pulmonary smear negativity (Dale et al. 2005a).

Whatever the degree of recombination, and the reasons for this, in the *M. tuberculosis* complex, it is generally agreed that the rate of recombination in the sequenced strains is negligible (Smith et al. 2006), genetic variation between strains is scarce and the pathogen has evolved in a clonal manner from a common ancestor approximately 35,000 years ago (Hughes AL, Friedman R, & Murray M 2002) and it is widely acknowledged that human TB is an ancient disease, in human terms at least (Donoghue et al. 2004; Donoghue 2009; Taylor, Young, & Mays 2005). Indeed, fossilised evidence of *M. tuberculosis* infection in *Homo erectus,* which evolved approximately 2 million years ago (Antón 2003), has been discovered in Turkey (Kappelman et al. 2008).

1.4 Host-Pathogen Interaction

The interaction of *M. tuberculosis* with the human host is central to the disease progression. TB is fundamentally an immune-related disease and infection occurs by entry into mononuclear phagocytes (macrophages or dendritic cells) by receptor-mediated phagocytosis (Schlesinger 1993). For review, see (Ernst 1998). This mucosal

interaction largely occurs in the lung after inhalation of aerosolised bacilli, but may also occur in the gut following the ingestion of *M. bovis* in unpasteurised milk. Once phagocytosed, the organism causes progression in three directions: clearance, disease or latency (figure 1.7). If the phagocytosed bacilli are rapidly killed by the innate immune system then the individual shows no sign of infection or disease and no apparent immune response (Dietrich & Doherty 2009). Infected dendritic cells migrate to local lymph nodes where they induce specific B and T cell responses. If the infection is not contained, then haematogenous spread can occur at this stage (Ehlers 2009).

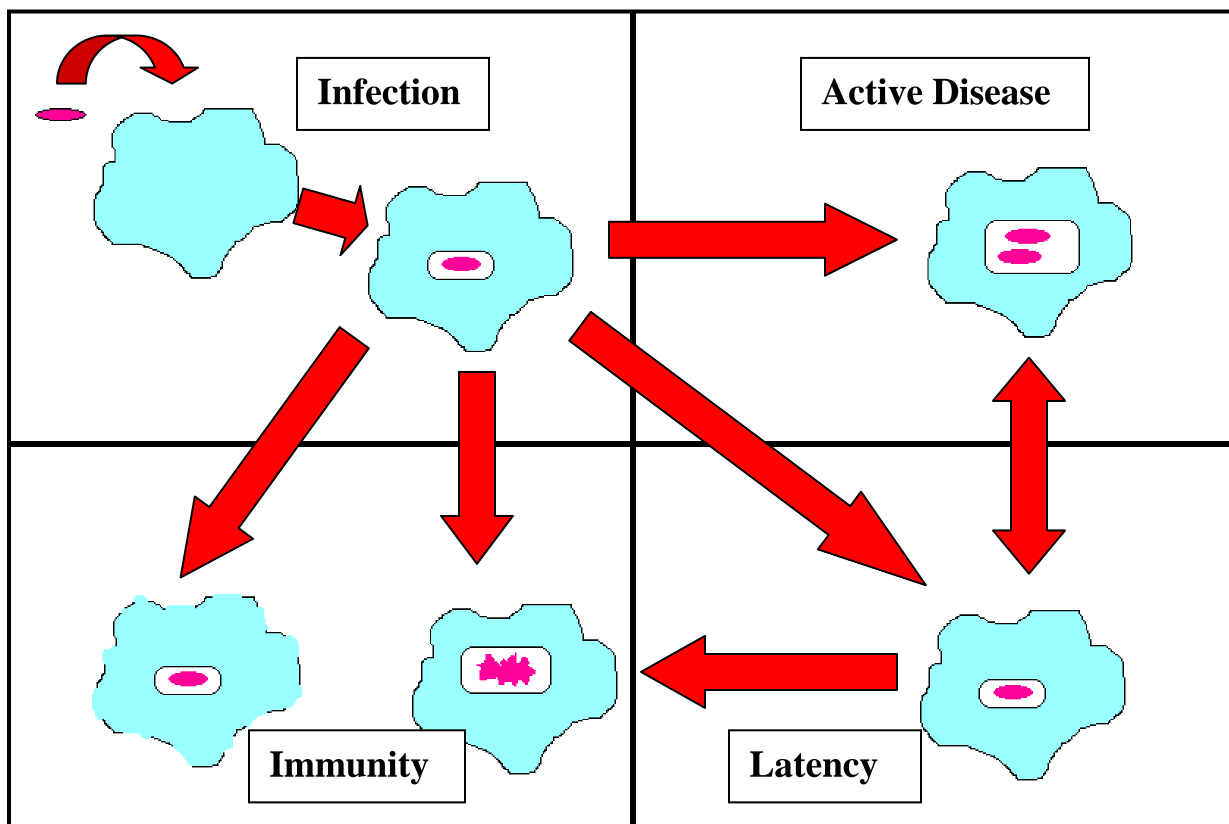


Figure 1.7. The possible outcomes of an *M. tuberculosis* infected macrophage. Infection: The bacillus is taken up by the antigen presenting cell and contained in the phagosome. Active Disease: The growth of the bacillus is not restricted. Latency: Bacillus growth is restricted, but not eradicated. Immunity: Apoptosis of the host cell or formation of the lysosome eradicates the bacillus. (Dietrich & Doherty 2009)

As already stated, latency occurs in the vast majority of infections. This occurs when the bacilli are taken up by antigen presenting cells (APC) and a rapid inflammatory immune response ensues. Activated T cells from the local lymph nodes migrate to the site of initial infection and create a cascade of cytokines that recruits a range of immune cells and a granuloma is formed. A granuloma consists of a small number of infected phagocytes in the centre, surrounded by activated monocytes and macrophages. Activated CD4, CD8 and small number of B cells then surround the mononuclear cells (Gonzalez-Juarrero et al. 2001). At this stage, there begins a balance between the host and the organism. If the host's innate immune response is able to control the infection, then the granuloma shrinks and may even disappear, causing calcification and T cell memory to *M. tuberculosis* antigens (Dietrich & Doherty 2009). Alternatively, bacterial replication leads to the granuloma increasing in size. Eventually the centre of the granuloma becomes necrotic and tissue destruction begins. Characteristic granulation and AFB may be seen in histological specimens (figure 1.8a & 18b, respectively). If the granuloma is close to the surface of the lung then the tissue destruction can breach the mucosa and cause cavitation. Cavitation is associated with advanced disease as well as increased bacterial burden, infectivity and mortality (Golub et al. 2006; Helke, Mankowski, & Manabe 2006; Kourbatova et al. 2006; Perrin et al. in press).

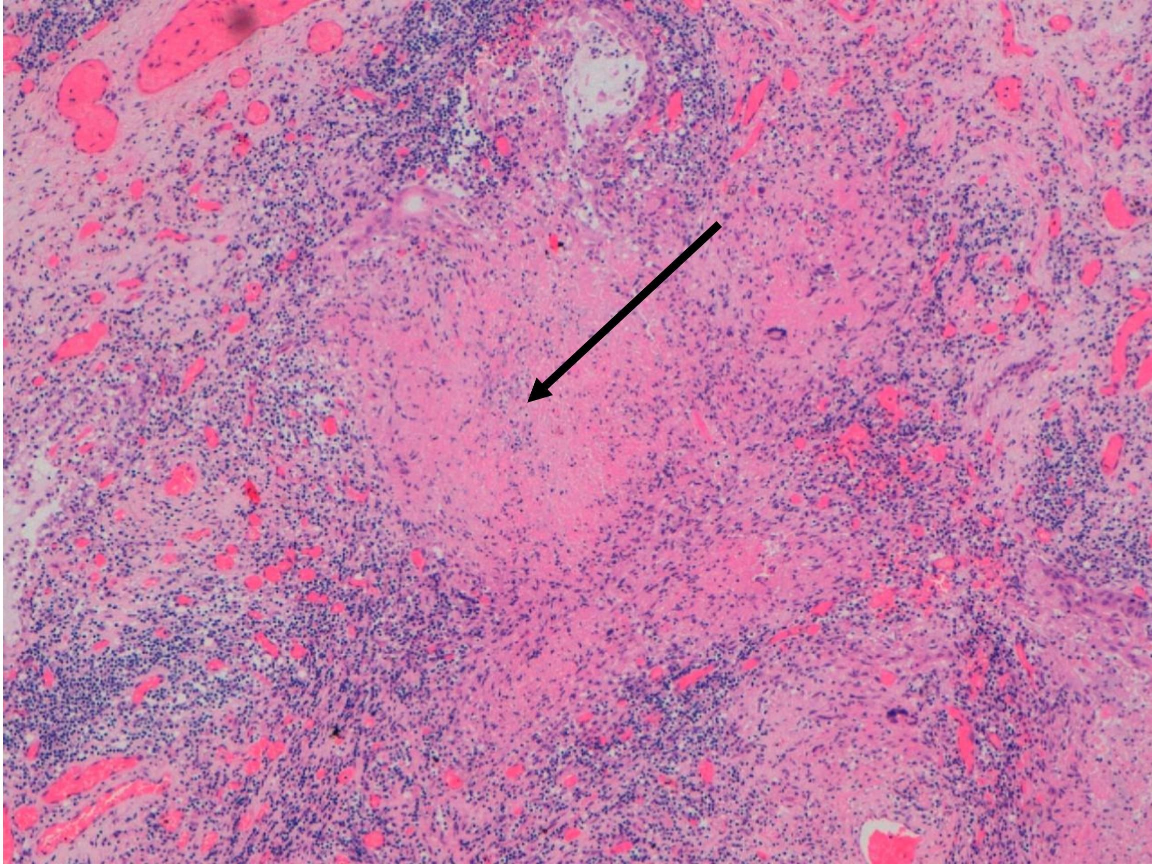


Figure 1.8a H&E stained section of lung tissue x100. The centre of the granuloma is indicated with an arrow. Courtesy Steve Davis, Histopathology, Royal Free Hospital.

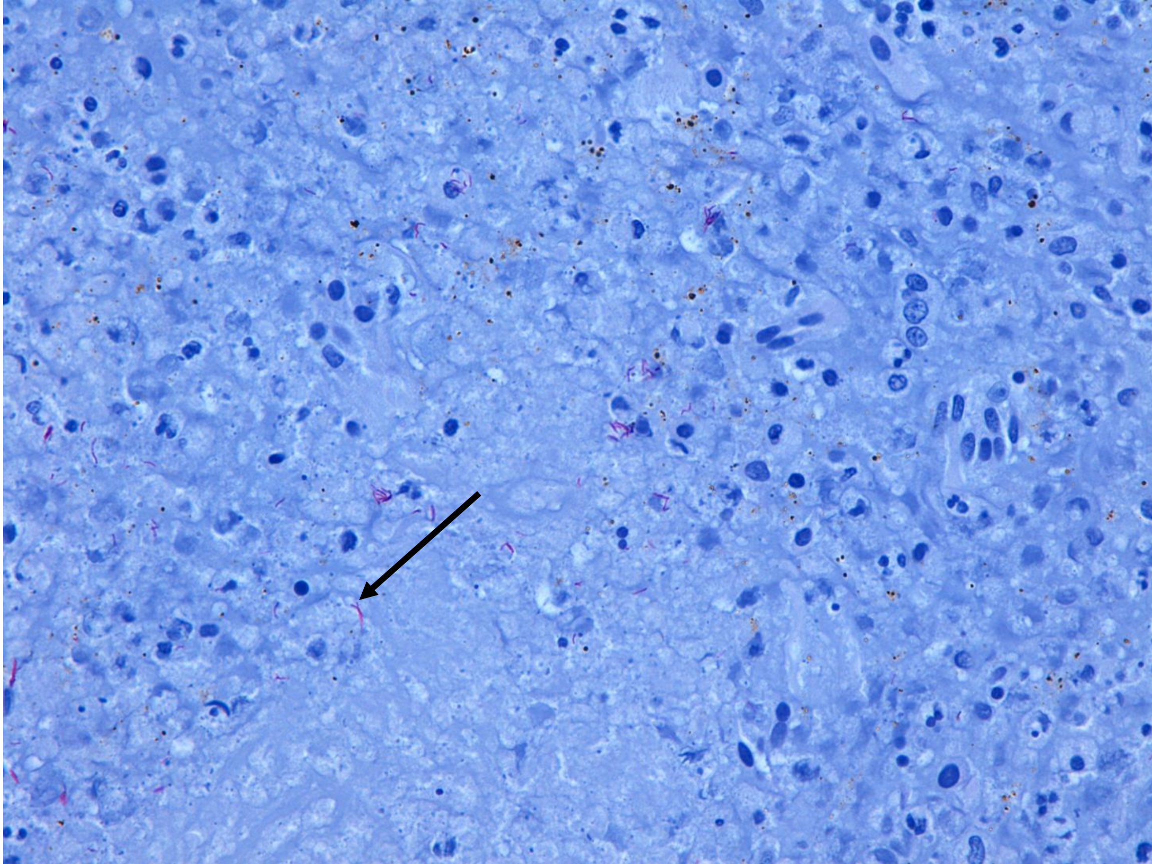


Figure 1.8b Ziehl-Neelsen stained section of lung tissue x400. AFB are indicated with an arrow. Courtesy Steve Davis, Histopathology, Royal Free Hospital.

The organism is central to the stimulation and inhibition of all stages of the host immune response and this appears to benefit the survival of the host and pathogen. The cell surface of *M. tuberculosis* contains a number of molecules that bind to the host's pathogen-associated molecular pattern (PAMP) receptors (Dietrich & Doherty 2009), such as the Toll-like receptor family (Means et al. 1999b; Means et al. 1999a). These receptors allow rapid and efficient activation of the innate immune system to protect the host (Quesniaux et al. 2004; Stenger & Modlin 2002), but *M. tuberculosis* expresses a number of molecules (both membrane bound and secreted) that stimulate these pathways. Most of these molecules do not appear to be essential for the survival of *M. tuberculosis*

so it may be concluded that their conservation is due to the beneficial outcome from the resulting granuloma formation and cavitation (Brosch et al. 2002).

M. tuberculosis is able to persist inside a host for extended periods due to its ability to survive inside macrophages and monocytes. These immune cells play a role in bacterial killing, but also stimulating an immune response by antigen presentation, which is summarised in figure 1.9. *M. tuberculosis* has co-evolved with humans to modulate these processes to its benefit. Stokes and colleagues showed that mannose derivatives on the surface of *M. tuberculosis* inhibit phagocytosis by activated macrophages and therefore may specifically target certain macrophage phenotypes (Stokes et al. 2004). The receptor DC-SIGN on the surface of dendritic cells is known to bind to lipoarabinomannan (LAM), which is a major cell wall component of *M. tuberculosis*. DC-SIGN is involved in dendritic cell maturation but LAM binding inhibits this as well as down-regulating IL-12 production whilst stimulating IL-10 secretion (Appelmelk et al. 2003; van Kooyk Y. & Geijtenbeek 2003). These modifications by *M. tuberculosis* inhibit antigen presentation and MHC expression. Levels of IL-10 have been shown to be elevated patients with active TB, which may support this (Jang, Uzelac, & Salgame 2008; Olobo et al. 2001; Redpath, Ghazal, & Gascoigne 2001).

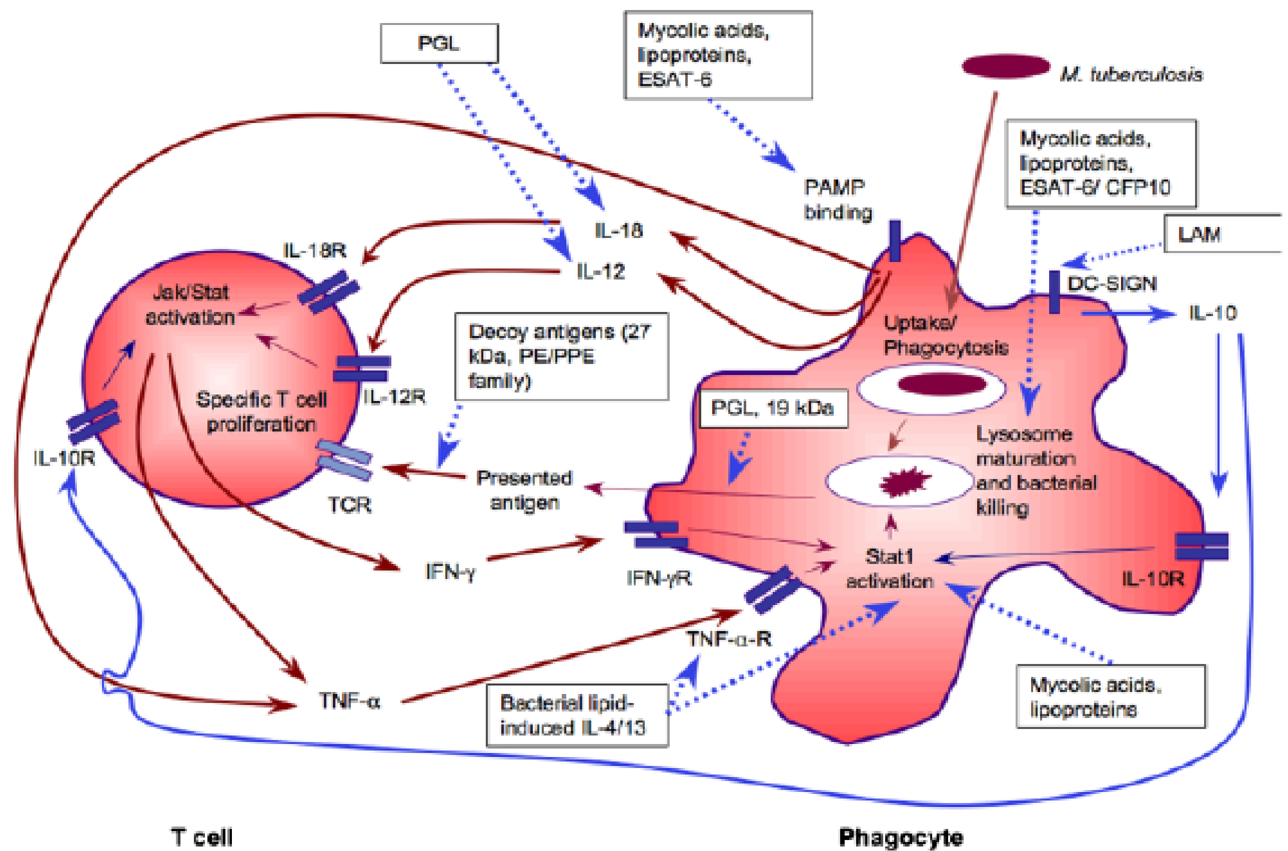


Figure 1.9. The interaction of the infected antigen-presenting cell and an antigen specific T cell after infection. The key pathways in the host's immune response are shown as solid arrows that can suppress (red) or enhance (blue) bacterial growth, together with the known bacterial products (white boxes, dotted arrows) that can interfere with the host's response (Dietrich & Doherty 2009).

Other cell wall components strongly stimulate the inflammatory response. Long chain fatty acids stimulate the host's immune response (Brennan 2003; Briken et al. 2004; Korf et al. 2005; Stenger & Modlin 2002), cause granuloma formation, up-regulation of antigen presentation and subsequent NK and T-cell responses (Hunter et al. 2006; Ryll, Kumazawa, & Yano 2001). If the host's immune response is allowed to proceed without restriction, the infection is rapidly be cleared by the host (Dietrich & Doherty 2009). The 19 kDa lipoprotein produced by *M. tuberculosis* has been shown to down-regulate the expression of IFN-γ by a third in murine macrophages (Pai et al. 2004), reduces antigen processing and MHC II expression (Fortune et al. 2004; Noss et al. 2001; Pai et al. 2004). Knocking out this lipoprotein attenuates *M. tuberculosis* in both immunocompetent mice and those deficient in the IFN- γ gene (Henao-Tamayo et al. 2007) and reduced immunity was seen in mice when expressed as a recombinant protein in *M. smegmatis* and *M. vaccae* (Yeremeev et al. 2000).

The proteins early secretory antigen target-6 (ESAT-6) and culture filtrate protein-10 (CFP-10) are co-transcribed, excreted by a novel secretory mechanism and are highly immunogenic (Berthet et al. 1998; Pallen 2002; Sorensen et al. 1995). ESAT-6 acts through TLR-2 and inhibits several pathways following TLR signalling (Pathak et al. 2007). It is a virulence factor that is absent from the attenuated vaccine strain *M. bovis* BCG (Harboe et al. 1996; Lewis et al. 2003; Pym et al. 2002). Additionally, the absence of these proteins has been shown to reduce immungenicity and virulence of *M. tuberculosis* in vivo and vitro (Brodin et al. 2006; Guinn et al. 2004).

Once inside the macrophage, *M. tuberculosis* further attempts to avoid immune clearance by inhibiting the phagosome maturation process. However, the organism must control its own environment to allow the acquisition of essential nutrients (especially iron) by the fusion of other vesicles and membrane remodelling (Deretic et al. 2006; Russell 2001; Steinberg & Grinstein 2008). In order to perform effective killing, the phagosome must mature and reduce the pH to provide and acidic environment. ESAT-6/CFP-10 (Tan et al. 2006), and SecA1/2 (Hinchey et al. 2007; Hou et al. 2008) interfere with this process by blocking the accumulation of ATPases and GTPases in the vacuole. *M. tuberculosis* further subverts host immune processes by molecular mimicry. The lipid manose-capped lipoarabinomannan (ManLAM) (Brennan 2003; Briken, Porcelli, Besra, & Kremer 2004; Steinberg & Grinstein 2008) appears to mimic host phosphatidylinositols, which are only usually present on the surface of a mature vacuole (Chua et al. 2004; Deretic et al. 2006). Also, *M. tuberculosis* produces a eukaryotic-like serine/threonine protein kinase G that can inhibit phagosome-lysosome fusion (Dietrich & Doherty 2009).

M. tuberculosis produces 'decoy' molecules that may be presented to stimulate a Th1 response. Animal models have shown that the 27 kDa lipoprotein produces a strong IFN-γ response that is antigen specific but non-protective and may even promote bacterial growth (Hovav et al. 2003; Hovav & Bercovier 2006). *M. tuberculosis* also utilises antigenic variation to evade host immune responses. The highly polymorphic PE-PGRS and PPE MPTR gene families are highly immunogenic, but although TB patients display significant IFN-γ responses to PGRS proteins, they are not protected (Singh et al. 2008; Zheng et al. 2008).

During the early stages of *M. tuberculosis* infection, the unconventional T-cell subsets (γδ, NK-T and CD-1 restricted cells) expand considerably (Ladel et al. 1995b; Ladel et al. 1995a; Schaible & Kaufmann 2000). The receptors of these cells display far less variation than the MHC I and II molecules of conventional T cells. In this way, they can act as a bridge between the innate and adaptive immune responses by 'kickstarting' cytokine production (Dietrich & Doherty 2009; Schaible & Kaufmann 2000). The secretion of IFN-γ by these cells activates APCs and stimulates the production of IL-12 and IL-18, which participate in a positive feedback loop for IFN- γ (Caccamo et al. 2006). The role of IL-12 in the susceptibility to TB is demonstrated by the protection of individuals with genotypes conferring high levels of the cytokine. Conversely, functional mutations in the IL-12 receptor have been associated with extreme susceptibility to mycobacterial disease (de Jong R. et al. 1998; Tso et al. 2004). The role of IL-12 in the proliferation of IFN-γ secreting CD4 T cells is important in immunity to TB. Increased susceptibility to disease has been shown in animal models when defects in CD4 T cell function (Saunders et al. 2002; Scanga et al. 2000), or with IFN-γ expression or recognition (Cooper et al. 1993; Flynn et al. 1993). Increased susceptibility has also been seen in such patients (Havlir & Barnes 1999; Ottenhoff et al. 2005).

The host's adaptive immune response plays a role in the control of *M. tuberculosis* infection, although the organism does manipulate this also. Some downstream effects of IFN-γ on cytotoxic CD8 T cells can be inhibited by either live *M. tuberculosis* or cell wall extracts. Even if IFN-γ is produced, its activity may be reduced (Kincaid & Ernst 2003; Manca et al. 2005; Ting et al. 1999). Sodhi and colleagues have demonstrated that

IFN-γ responses are poor in patients with advanced TB (Sodhi et al. 1997), while IL-4 is elevated (Jimenez-Martinez et al. 2004; Roberts et al. 2007). Studies of disease severity in patients (Jalapathy, Prabha, & Das 2004; Jimenez-Martinez et al. 2004) and susceptibility to disease in individuals exposed to *M. tuberculosis* (Ordway et al. 2004; Wassie et al. 2008) have shown associations with raised IL-4 expression. The *M. tuberculosis* phosphoglycolipids and the 19 kDa antigen that stimulate of the production of IL-4 and IL-13 also inhibit IFN-γ (Briken et al. 2004; Manca et al. 2004; Reed et al. 2004a). LAM binding to the DC-SIGN receptor stimulates IL-10 production, which also inhibits IFN-γ (Appelmelk et al. 2003; Redpath, Ghazal, & Gascoigne 2001; van Kooyk Y. & Geijtenbeek 2003). Progression to disease in patients exposed to *M. tuberculosis* is associated with a low IFN-γ/IL-4 ratio (Wassie et al. 2008) and a low IFN-γ/IL-10 ratio (Hussain et al. 2007; Sahiratmadja et al. 2007) and as such, phosphoglycolipids, the 19 kDa antigen and LAM can be considered to be evolved virulence factors that help to modulate the host immune system.

TNF-α plays a critical role in controlling *M. tuberculosis* infection (Jacobs et al. 2007). Patients with latent TB infection rapidly progress to active disease when treated with TNF-α antagonists, such as Infliximab (Anonymous 2004; Gomez-Reino et al. 2003). As such, patients should be thoroughly investigated for TB before commencing such therapy (British Thoracic Society 2005).

Another part of the hosts' arsenal in controlling infection with *M. tuberculosis* is the ability to remove infected cells from the site of infection either by apoptosis or necrosis.

Necrosis is the more disruptive process and results in the destruction of surrounding tissue and the release of viable bacilli (Bocchino et al. 2005), whereas apoptosis can remove infected cells while minimising death to adjacent cells (Gutierrez et al. 2004). Stenger and colleagues have shown that in addition to stimulating the maturation of dendritic cells, activating antimicrobial activity in macrophages and the orchestration of leucocyte movement, TNF-α plays an important role in the induction of apoptosis (Stenger 2005). Furthermore, a reduced apoptotic capacity in the macrophages of HIV-infected individuals is associated with an inability to control *M. tuberculosis* infection (Patel et al. 2007).

It is highly likely that latent disease will follow infection with *M. tuberculosis*. Once infection occurs, the detection of this event may only be possible by immune-mediated investigations. The tuberculin skin test (TST) involves injecting *M. tuberculosis* antigens in the form of purified protein derivatives (PPD) subcutaneously to illicit a delayed-type hypersensitivity response indicating long-term T cell memory (American Thoracic Society 2000). An interferon gamma release assay (IGRA) detects long-term T cell memory to specific *M. tuberculosis* antigens such as ESAT-6 and CFP-10 (Arend et al. 2007b; Ferrara et al. 2006; Granger 2006; Lalvani 2007; Richeldi 2006). This assay is discussed in more detail in section 1.5. There is still a lot of uncertainty surrounding the process of latency and the dormant bacillus. It is true that those individuals with a positive IGRA response are more likely to have encountered the tubercle bacillus than those with a positive TST due to the cross reactivity of antigens with non-tuberculous

mycobacteria. However, it is uncertain how many of these individuals will proceed to active disease and indeed how many of them actually have viable bacilli within them.

There are currently two major models for latency in tuberculosis, which are reviewed in (Ehlers 2009). The first describes slowly replicating organisms within granulomas that periodically interact with the immune system to recruit new cells to the lesion. This disruption of this intermittent interplay with the immune system causes reactivation. However, as this model describes free-living bacilli in the alveolar air-space, this contradicts the very definition of latency. The second argues that dormant, non-replicating bacilli are found in alveolar epithelial cells and adipocytes. Here they are adapted to hypoxic conditions and are characterised by lipid inclusions (Garton et al. 2008). Reactivation of disease is associated with the up-regulation by mycobacterial resuscitation promoting factors and the escape of bacilli into the alveoli and bronchi (Ehlers 2009). Surprisingly, the bacilli seen in the sputum of patients with active pulmonary TB contain lipid inclusions that characterise dormancy (Garton et al. 2008).

Vaccination with the attenuated *M. bovis* BCG (Bacillus Calmette-Guérin) strain has been shown to give good protection in children (Colditz et al. 1995) and enhanced protection against death, disseminated disease and meningitis (Colditz et al. 1994) but this immunity wanes with age (Sterne, Rodrigues, & Guedes 1998) and does not provide equal immunity globally, particularly in the areas with the greatest burden (Eickhoff 1977). It is thought that prior exposure to environmental mycobacteria, previous BCG vaccination or TB infection may account for the poor immunogenicity in some areas.

This theory is supported by the failure to boost immunity with repeat doses of BCG (Leung et al. 2001; Rodrigues et al. 2005). Indeed, previous infection with *M. tuberculosis* does not prevent subsequent re-infection. These findings highlight the scale of the challenge in developing novel TB vaccines.

1.5 Disease Presentation and Diagnosis

Although initial infection occurs in the lung, systemic effects are often seen in the clinical presentation of TB. Fever, weight loss and malaise are all common symptoms and are likely to be caused by cytokine release, especially tumour necrosis factor alpha (TNF-α) (Ernst 1998; Orme & Cooper 1999; Takashima et al. 1990; Valone et al. 1988). Indeed, such systemic symptoms are highlighted in UK guidance (NICE 2006). Biomarkers in peripheral blood are also seen. Anaemia and leukopenia may be present. Anaemia and pancytopaenia are more often seen when there is bone marrow involvement (Cameron 1974). Inflammatory markers such as the acute phase protein C-reactive protein (CRP) are used in the diagnosis of TB. However, a prospective study has shown that 15% of TB patients had normal CRP levels. This study also found that fever, sweats and weight loss were absent in 37%, 39% and 39% of TB patients, respectively (Breen et al. 2008).

1.5.1 Immunological Diagnosis

The tuberculin skin test (TST) is an *in vivo* assay that is used to detect latent TB infection. The TST, also known as an intradermal Mantoux test since 1910, is one of the oldest diagnostic test still used in modern medical practice (Richeldi 2006). The validity of the TST suffers from low specificity. Purified protein derivative (PPD) is a culture filtrate of tubercle bacilli (Seibert & Dufour 1948) that contains over 200 antigens that

are shared with *M. bovis* BCG and most non-tuberculous mycobacteria (Huebner, Schein, & Bass, Jr. 1993). PPD is inoculated subdermally and the scale of the delayed-type hypersensitivity response, indicating long-term T cell memory, is measured in the size of the induration generated. Therefore, interpretation of this test is hampered due to exposure to environmental mycobacteria, previous BCG vaccination (Tissot et al. 2005; Wang et al. 2002) and in the immunocompromised.

An improvement on the TST are two commercially available immunological assays that detect exposure to *M. tuberculosis*. The QuantiFERON-TB Gold (QFT-G; Cellestis, Carnegie, Australia) and the T-SPOT.*TB* (Oxford Immunotec, Oxford, UK) are *ex vivo* assays that measure the interferon gamma (IFN-γ) expressed by T cells in response to exposure to the antigens ESAT-6 and CFP-10, and as such, are collectively referred to as interferon gamma release assays (IGRA). The TB-SPOT.*TB* assay enumerates the IFN-γ secreting T cells in a venous blood sample, whilst the QFT-G assay quantifies the amount of IFN-γ secreted by patient derived, stimulated T cells (Lalvani 2007).

The use of these assays is recommended in the diagnosis of latent TB infection by the National Institute of Health and Clinical Excellence (NICE). Alongside the tuberculin skin test, or Mantoux, and chest x-ray these assays are used when contract tracing individuals who have been exposed to a confirmed case (NICE 2006). The algorithm for the use of these assays in contact tracing is displayed in figure 1.10.

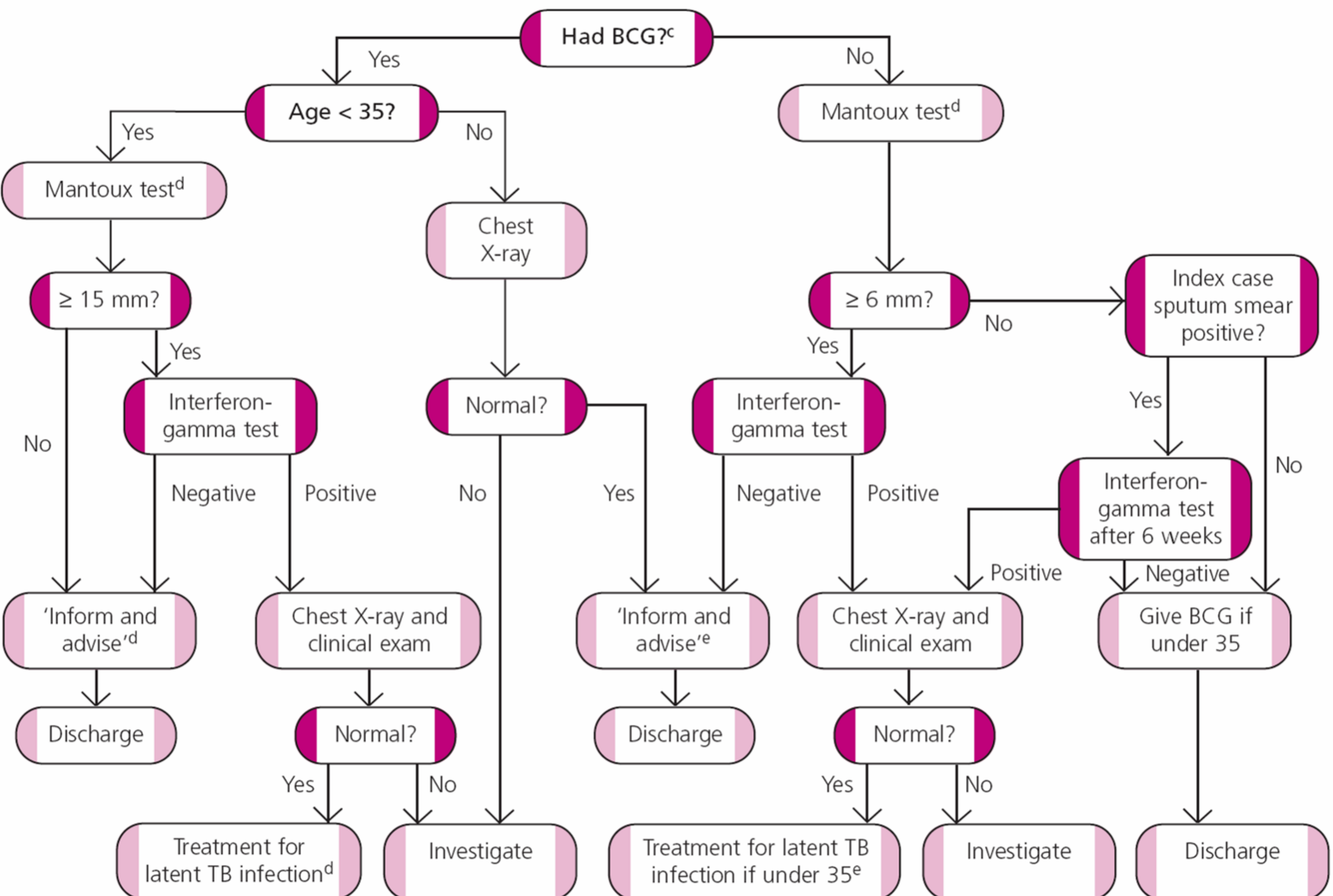


Figure 1.10. NICE guidance: close contacts of all cases of active TB (NICE 2006)

The antigens ESAT-6 and CFP-10 are much more specific to *M. tuberculosis*, than those used in the TST, and so the IGRAs are less likely to produce positive results in the case of BCG vaccination and exposure to environmental mycobacteria. Indeed, in one study of a contact tracing investigation, a positive IGRA result was associated with TB exposure whereas a positive TST result was not (Arend et al. 2007a). These assays will be discussed in more detail in chapter 5.

1.5.2 Microscopy

The early diagnosis of TB is central to its control. In the poorest parts of the world, where majority of the disease burden lies, the mainstay of diagnosis is microscopy. Microscopy of respiratory samples may be performed using either light microscopy with a Ziehl-Neelsen or Kinyoun stain or fluorescent microscopy with a stain such as auramine-O. Light microscopy is easier and cheaper and is used in the diagnosis of TB in many resource poor settings. In addition, it may be performed in areas with no electricity by using reflected sunlight. Fluorescent microscopy is more sensitive but conventionally requires an expensive fluorescent microscope and a dark room. The introduction of low-cost fluorescent LED technology has enabled this more sensitive diagnostic tool to be introduced into the areas where it is most needed (Mizuno et al. 2009).

In addition to the type of microscopy used, the number of samples and sampling time are crucial to improving the sensitivity of this diagnostic assay. Early morning samples are considered to contain more organisms and Ulukanligil and colleagues (Ulukanligil, Aslan, & Tasci 2000) showed that the sensitivity of sputum smears increased with repeat

sampling. The sensitivity of Ziehl-Neelsen and fluorescence microscopy was found to be 61% and 83% respectively when a single sample was submitted. When two were submitted the sensitivities were 66% and 83% and where three or more were submitted sensitivities were 80% and 92%, respectively.

It has been shown that between 5,000 and 10,000 bacilli per milliliter of sputum are required to allow visualisation. This is a far higher number than is required for culture isolation but the detection of a symptomatic patient with a positive pulmonary smear provides the clinician with a presumptive diagnosis and allows the patient to be placed in respiratory isolation.

1.5.3 Culture

Culture, although not always positive in patients who are treated for TB, is considered the gold-standard for diagnosis. Recent UK data shows that only 56% of notified TB patients were confirmed with a positive culture (Health Protection Agency 2009). The sensitivity of culture is far higher than microscopy and as few as 10 bacilli per milliliter have been detected (Yeager, Jr. et al. 1967). In addition to the increased sensitivity, culture should be performed on all specimens to provide definitive species identification, drug susceptibility testing and genotyping. Following the removal of contaminating non-mycobacterial organisms, the specimen is then cultured. It should be noted that, in the UK, members of the *M. tuberculosis* complex are designated to Hazard Group 3 Pathogens and therefore, such specimens and cultures must be handled using appropriate biological safety cabinets inside a containment level 3 laboratory. Traditional culture

media are either egg-based (Löwenstein-Jensen), agar based (Middlebrook 7H10 or 7H11) or liquid (Middlebrook 7H9, 7H12 or Kirshners).

More recent advances into automation have allowed the sensitivity and time to detection to improve. The automated liquid culture systems such as MB/BacT system (bioMerieux, France) and the MGIT 960 (Becton Dickinson, New Jersey, USA) have both been shown to increase sensitivity and reduce detection time when compared to solid media (Muyoyeta et al. 2009; Palomino et al. 2008; Parrish et al. 2009; Somoskovi et al. 2000; Sorlozano et al. 2009). The MB/BacT system (bioMerieux, France) utilises a modified Middlebrook 7H9 broth and a pH-dependent indicator that alters when the pH falls in the presecence of increased CO_2, which is produced by dividing organisms. The MGIT 960 (Becton Dickinson, New Jersey, USA) uses a similar culture media, but the detection method varies from the MB/BacT. A fluorophore is bound to oxygen in the tube. When the organisms divide and utilise the oxyen, the fluorphore is released and detected by the analyser. This latter method has been shown to recover more organisms from extrapulmonary samples than solid culture (Hilleman et al 2006). Furthermore, an increase in time to detection of serial samples in MGIT 960 has been shown to correlate to disease resolation in TB patients (Epstein et al 1998).

1.5.4 Nucleic Acid Amplification Tests

Despite the reduction in time to detection by the advances in culture techniques, the bacteriological diagnosis of TB can still take several days. Clinicians often wish to start treating patients who are strongly suspected of having active TB disease before culture

results are available. Several molecular tools have become available over the past 15 years to aid in the rapid diagnosis of TB. Various molecular targets have been used to amplify *M. tuberculosis* DNA to aid diagnosis. For example, the Qiagen Artus TB PCR (Crawley, UK) and the Roche COBAS TaqMan MTB assay (Indianapolis, USA) amplify an *M. tuberculosis* specific section of the 16S rDNA gene. The Cepheid Gene Xpert (Derby, UK) targets the *rpo*B gene whilst the insertion sequence IS*6110* is the target in the strand displacement amplification (SDA) assay by Becton Dickinson (New Jersey, USA).

Sensitivity and specificity for respiratory samples of 93% and 92%, respectively, have been reported for this latter assay (McHugh et al. 2004). It should be noted that most strains of *M. tuberculosis* possess multiple copies of IS*6110*, seemingly making it an excellent target for a diagnostic NAAT. However, a very small number of strains have no copies of IS*6110*, and therefore the use of this target should be considered carefully. Additionally, IS*6110* homologs have been reported in other non-tuberculous mycobacteria, so cross-reactivity may reduce the specificity of this assay (McHugh, Newport, & Gillespie 1997).

The observation that these assays perform best when used for respiratory samples was echoed in an evaluation of NAATs by Dinnes and colleagues (Dinnes et al 2007). However, due to a lack of direct comparisons it was not possible to recommend any one assay over another. A relatively recently introduced NAAT, the Cepheid GeneXpert, allows the detection of M. tuberculosis DNA and rifampicin resistance-causing mutations

simultaneously. Analysis of this assay found that the sensitivity was and similar to that of the Becton Dickinson assay but higher than that of the Roche Assay (Boehme et al 2010).

1.5.5 Antimicrobial Susceptibility Testing

Drug resistance is discussed further in section 1.7. The prompt detection of drug resistance is central to controlling the disease. A number of assays are in use for the phenotypic detection of drug resistance/susceptibility. The resistance ratio (RR) method is perhaps that the most accurate at measuring the drug susceptibilities of *M. tuberculosis*. (Heifets and **Cangelosi** 1999). The RR measures the Minimum Inhibitory Concentration (MIC – the lowest concentration of antimicrobial capable of inhibiting growth) of the strain, compared to reference strain $H_{37}Rv$ when measured in the same experiment. The MIC of tested strain is divided by the MIC of $H_{37}Rv$. If the RR is 2 or less the strain is susceptible. An RR of 8 or higher indicates resistance. The modified proportion methed is less labour expensive and may utilise automated liquid culture systems (Pfyffer et al 2002). The strain in question is incubated in the prescence of first line drugs. A growth control (1:100 dilution of the strain) is incubated in the absence of drugs. Should growth be detected in any of the drug-containing tubes prior to, or at the same time as the growth control, this indicates that at least 1% of the bacterial population is resistant to that agent.

The European Committee for Antimicrobial Susceptibility Testing (EUCAST) of the European Society of Clinical Microbiology and Infectious Diseases (ESCMID) has a Subcommittee on Antimicrobial Susceptibility Testing of Mycobacterium tuberculosis, which has reviewed these complex methods alongside molecular and novel assays. It

highlights that with regard to phenotypic antimicrobial susceptibility testing, accuracy is more important than speed and that stringent quality control measures are standardisation are required (Drobniewski et al 2007). Indeed, the WHO European Laboratory Strengthening Task Force has recommended standards for modern tuberculosis laboratory services in Europe to ensure quality (Drobniewski et al 2006).

The phenotypic detection of antimicrobial susceptibilities is the gold-standard, but this takes several days following the isolation of the organism. Microscopic-observation drug-susceptibility assays (MODS) allow the growth of resistant organisms in the presence of anti-tuberculosis drugs to be visualised. This assay may be performed directly on clinical samples and has shown some promise, particularly in resource poor settings (Moore et al. 2006).

The molecular detection of the most commonly occurring resistance genotypes provides a rapid diagnosis of drug-resistant TB. Nucleic acid amplification techniques (NAAT) and reverse hybridisation assays allow the detection of the most commonly occurring rifampicin resistance mutations direct from patient samples (Rossau et al. 1997). The detection of mutations conferring resistance to isoniazid and other anti-tuberculosis drugs are also available (Hillemann, Rusch-Gerdes, & Richter 2007).

1.5.6 Imaging

Imaging plays a significant role in the diagnosis TB. A multivariate analysis showed that upper lobe disease on a chest radiograph was positively associated with culture proven tuberculosis (odds ratio 14.6) (Wisnivesky et al. 2000). Chest radiographs (figure 1.11)

do not detect all cases of pulmonary TB however. Pepper and colleagues found that 9% of patients with culture-confirmed pulmonary TB had a normal chest x-ray (Pepper et al. 2008). NICE guidance states that patients with suspected pulmonary TB should undergo a posterior–anterior chest X-ray. In addition, all patients with non-respiratory TB should have a chest X-ray to exclude or confirm coexisting respiratory TB (NICE 2006). CT scanning (figure 1.12) can aid in the diagnosis of both pulmonary and extra-pulmonary, especially lymph node and skeletal disease. In addition, CT allows guided biopsies to be obtained for laboratory analysis.

1.6 TB Epidemiology in UK and London

The incidence of TB in England and Wales has increased since 1988, following a decline over the previous two centuries. Recorded notifications have fallen over the last hundred years (figure 1.13). The number of cases of TB in the UK was 8655 in 2008 (14.1 per 100,000). This was a slight increase from 2007 and numbers and rates had appeared to have stabilised since 2005 (Health Protection Agency 2009). The majority of disease is centred in major cities. London accounted for 45% of the total number of TB cases in England and Wales in 2008 compared with 28% in 1987 (Antoine D 2006;Health Protection Agency 2009). The rates and numbers of TB cases in England by region in 2008 are depicted in figures 1.14 and 1.15. The remainder of cases are distributed unevenly across the country, with particular hotspots in the Midlands and North West. The burden of disease is also not spread evenly across the capital. Some Primary Care Trusts (PCTs) in between 2006-2008 experienced rates that were approximately ten times greater than others. For example, Newham and Brent Primary Care Trusts (PCTs) have

rates of greater than 80 per 100,000 individuals whilst Richmond and Twickenham PCT has rates of 5-9 per 100,000 individuals (figure 1.16).

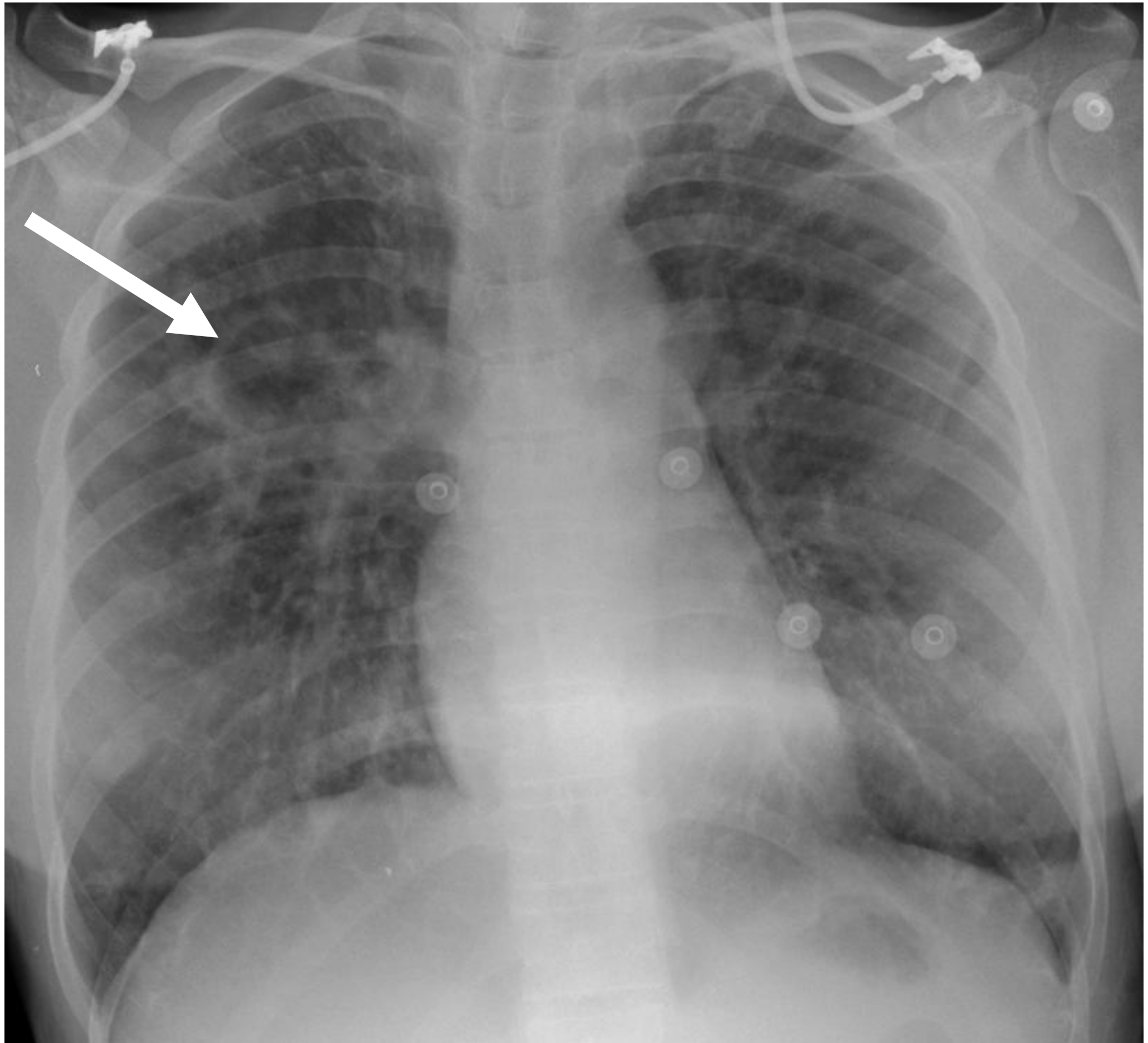


Fig 1.11 Chest x-ray demonstrating cavitation (indicated by arrow) in the upper right lobe and consolidation. Courtesy of Dr. Marc Lipman, Royal Free Hospital, London

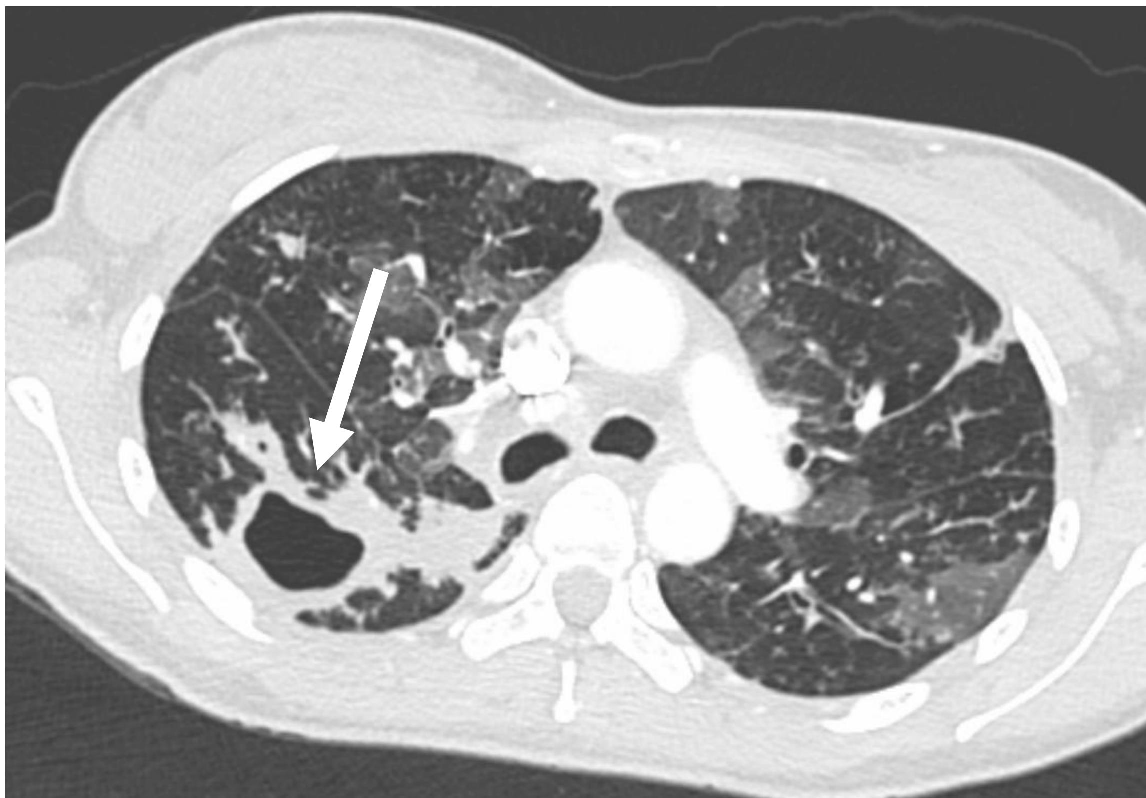


Figure 1.12 CT scan of the same patient in figure 1.11. Note the large cavity in the upper right lobe (indicated by arrow). Courtesy of Dr. Marc Lipman, Royal Free Hospital, London

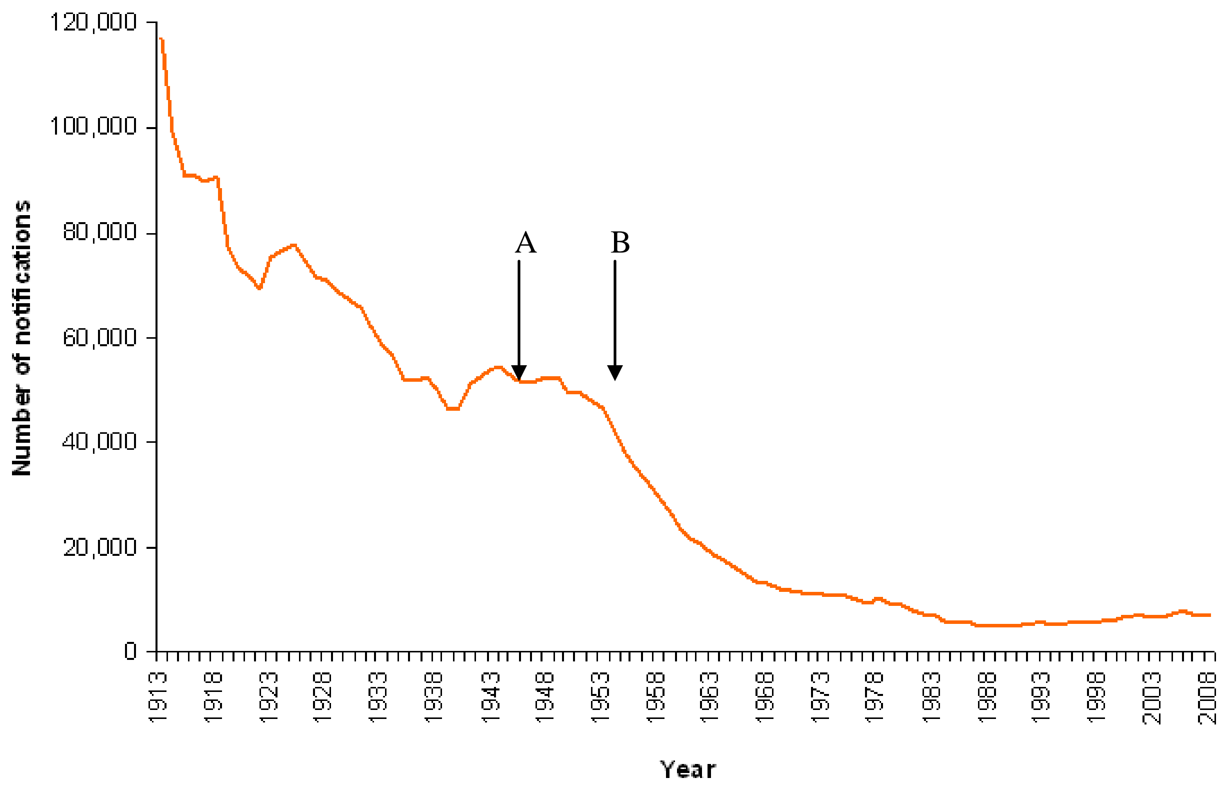


Figure 1.13 TB notifications from 1913 to 2008. A = introduction of streptomycin. B = Introduction of *M. bovis* BCG vaccination (Adapted from (Health Protection Agency 2008))

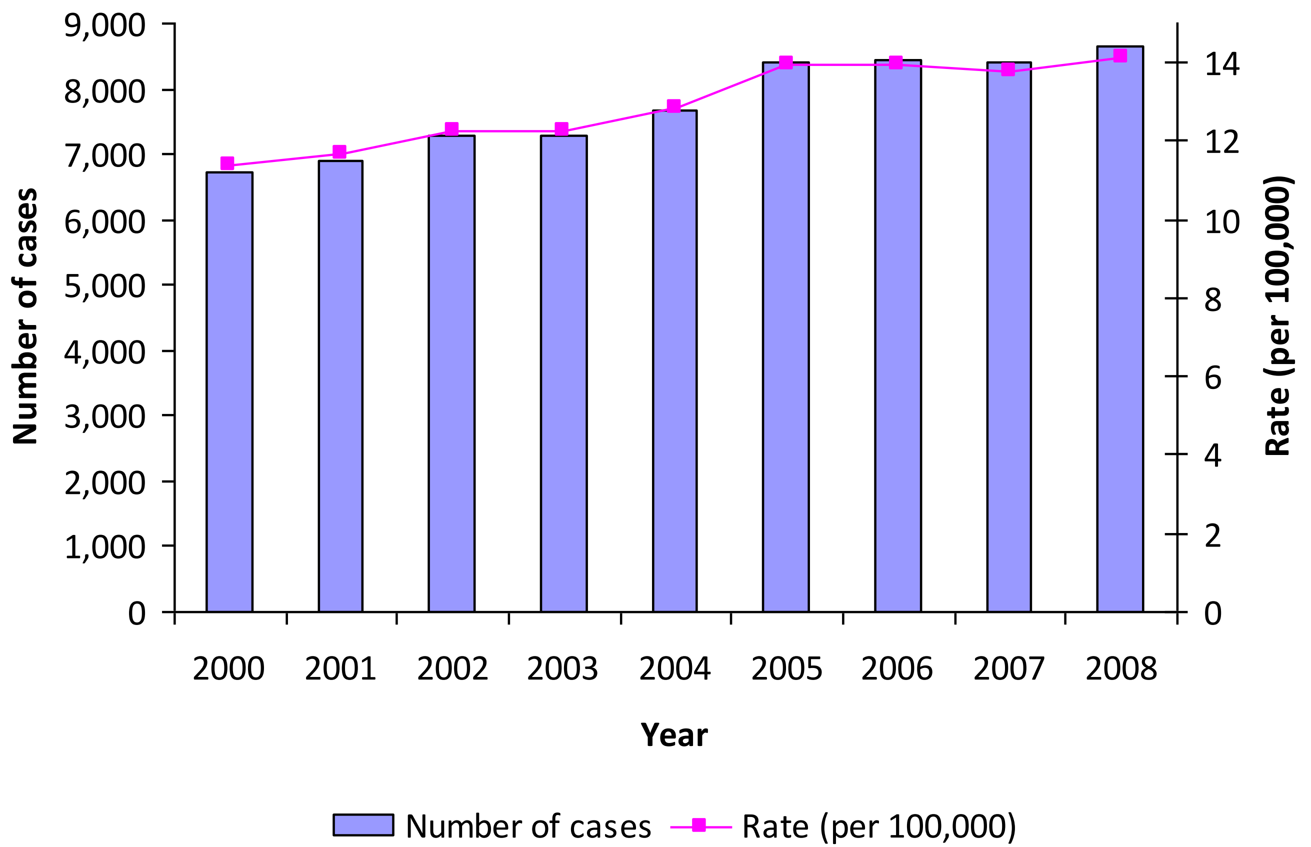


Figure 1.14. Tuberculosis case reports and rates, UK, 2000-2008 (Health Protection Agency 2009).

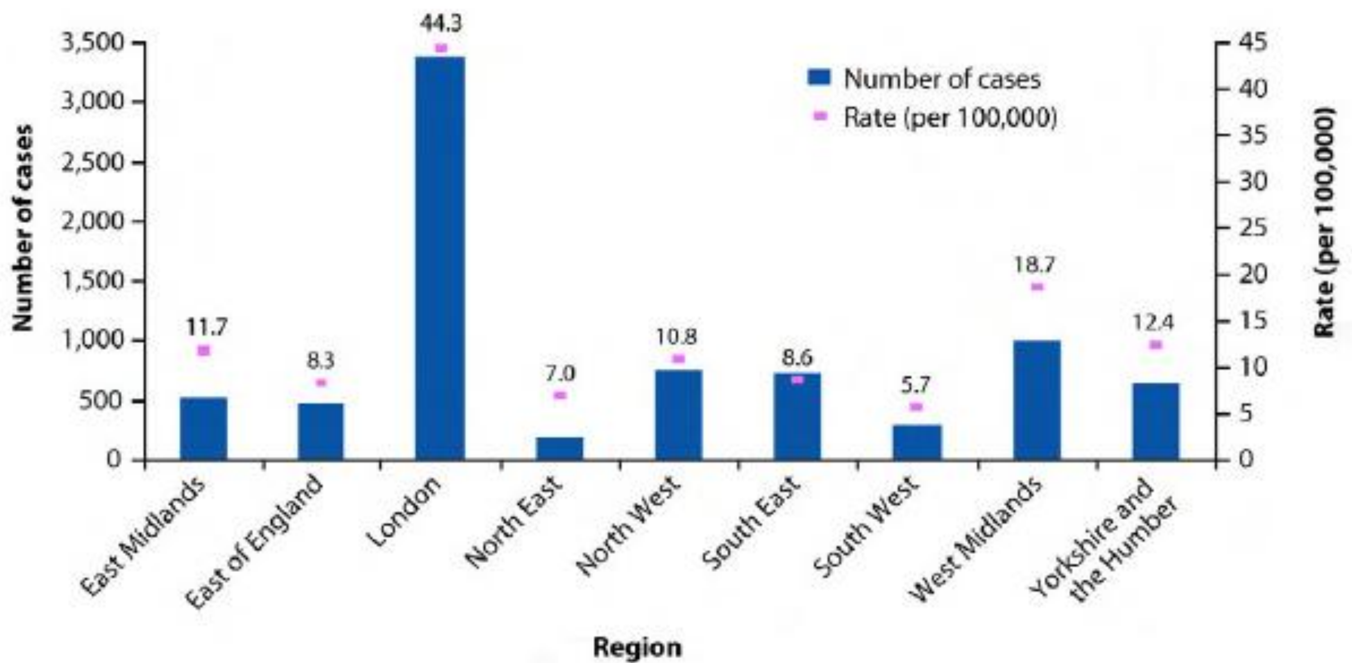


Figure 1.15 Tuberculosis case reports and rates by region, England, 2008 (Health Protection Agency 2009)

Rates of active transmission in London between 1995-1997 were low and the majority of cases were due to reactivation or importation of infection in recent immigrants (Maguire et al. 2002). In the UK in 2008 the majority of cases (72%) occurred in those born outside of the UK (Health Protection Agency 2009). A recent study has shown that of 21 European countries, only the UK, Sweden and Norway experienced an increase in rates of TB between 1996 and 2005. Approximately three quarters of TB patients in these three countries were foreign-born. From the data available, the UK possessed only third highest number of foreign nationals (behind Germany and France), but received had the highest number of migrants from countries with a TB incidence of ≥ 250 per 100,000 (Gilbert et al. 2009).

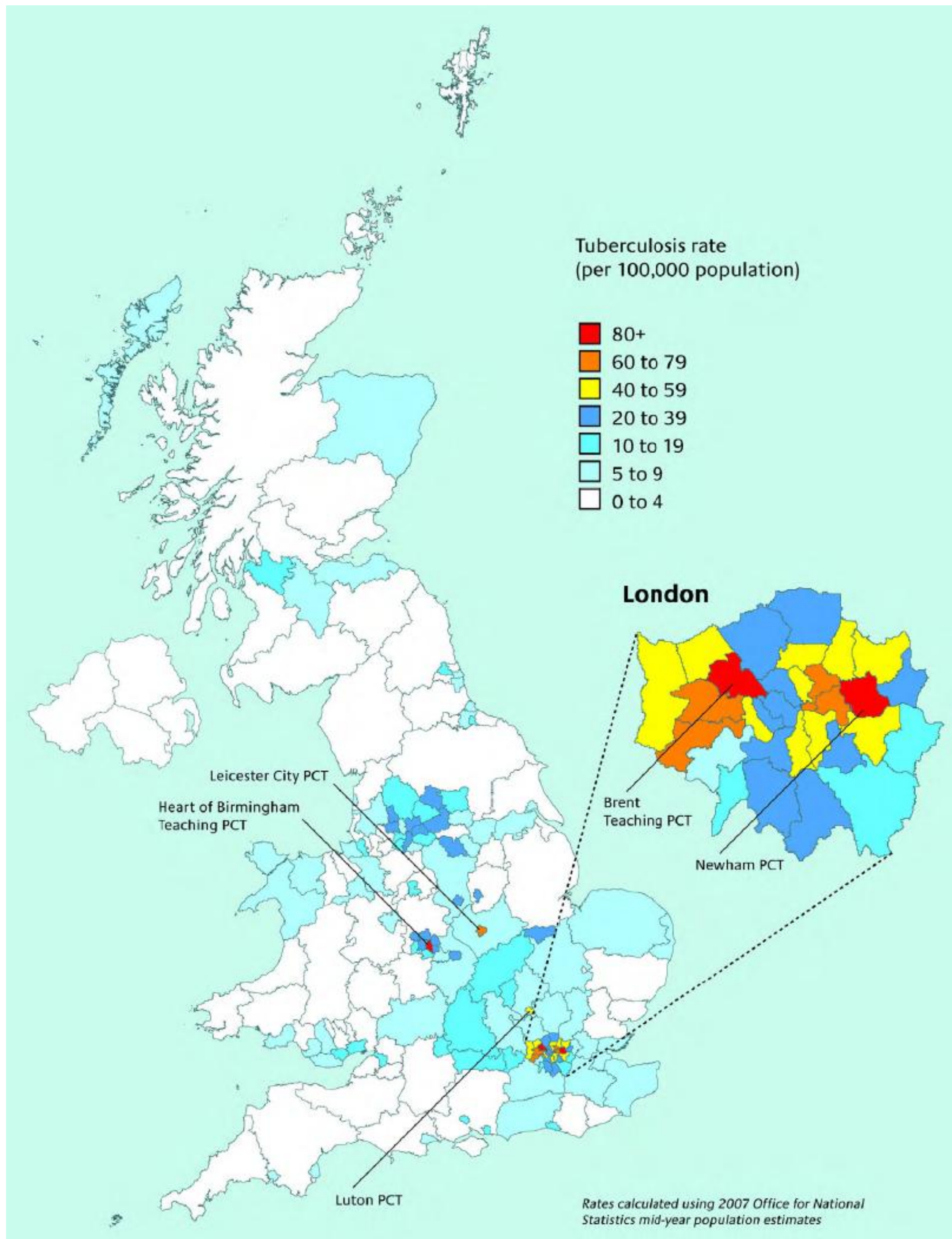


Figure 1.16 Three-year average tuberculosis case rates by primary care organisation*, UK, 2006-2008
(Health Protection Agency 2009)

Although the majority of cases in London are a result of immigration from high incidence areas, there are other groups of patients that present a challenge to control of the disease. A study by Story and colleagues (Story et al. 2007) of all TB patients who were or should have been on treatment in Greater London on July 1st 2003 showed strong associations with certain hard-to-reach groups. The overall prevalence of TB in this area was 27 per 100,000. Rates per 100,000 in the homeless, problem drug users and prisoners were 788, 354 and 208, respectively. Overall, these three groups accounted for 17% of TB cases in London and were more likely to be poorly compliant with treatment and lost to follow up. It has been further shown that pulmonary TB patients that used crack cocaine were 2.4 times more likely to have smear-positive disease, and therefore be more infective, than those who were not known to use drugs (Story, Bothamley, & Hayward 2008).

Of particular note is an outbreak of isoniazid resistant TB, centred on North London that was originally reported in 2004 (Ruddy et al. 2004). The demographics of the patients in this outbreak were unusual as the cases were more likely to be male and white or black Caribbean. Additionally, the outbreak was strongly associated with homelessness and prison detention. By the end of 2001 there were 70 confirmed cases and this was predicted to rise significantly. By 2006, the number of culture-confirmed cases had risen to 266, with 40 of these outside London. There were nine cases of MDR-TB. Four of these cases were a result of poor treatment adherence, while three cases were due to primary drug resistance by likely transmission in the community (Maguire H, Forrester, & Adam 2006).

1.7 Treatment, Drug Resistance and Fitness

1.7.1 Treatment

The treatment of TB was revolutionised in the 1940s with the discovery of streptomycin by Selman Waksman (Jones et al. 1944; Waksman, Reilly, & Schatz 1945), for which he received a Nobel Prize in 1952. However, resistance to streptomycin in *M. tuberculosis* was seen rapidly after the introduction of the drug (Crofton & Mitchison 1948). Indeed, the British Medical Research Council trial that investigated the efficacy of streptomycin showed that the majority of treated patients developed resistant strains (British Medical Research Council 1948). The discovery of para-aminosalicylic acid (PAS) by Jorgen Lehman in the same year that Waksman discovered streptomycin allowed combination therapy to be trialled (Lehman 1949). These trials showed that treatment with PAS and streptomycin combined was more effective than either agent when used alone (Turnbull et al. 1953).

1.7.2 Modes of Antimicrobial Action and Resistance

Streptomycin is an aminocyclitol glycoside and was the first anti-tuberculosis drug used. It has been shown in *E. coli* to bind to 16S rRNA, inhibit translational binding and therefore protein synthesis (Moazed & Noller 1987; Noller 1984). Resistance to streptomycin has been associated with mutations in the 16S rRNA gene (*rrs*) and *rps*L encoding ribosomal protein S12 (Cooksey et al. 1996; Douglass & Steyn 1993; Finken et al. 1993; Honore & Cole 1994; Meier et al. 1994; Meier et al. 1996; Nair et al. 1993; Sreevatsan et al. 1996). Unlike other bacteria, members of the *M. tuberculosis* complex only have one copy of *rrs* (Bercovier, Kafri, & Sela 1986; Suzuki et al. 1987). Therefore,

a single polymorphism has the potential to result in phenotypic streptomycin resistance. Most mutations in *rrs* are focused in two locations (nucleotides 530 and 915), however clinical streptomycin resistance is predominantly caused by mutations in *rps*L. The most common mutation occurs at codon 43 and results in a substitution of lysine with arginine (Cooksey et al. 1996; Finken et al. 1993; Honore & Cole 1994; Meier et al. 1994; Meier et al. 1996; Nair et al. 1993; Sreevatsan et al. 1996). Mutations are also seen at codon 88. A study of isolates from diverse geographic locations showed that 8% of resistant isolates were caused by mutations at nucleotides 530 and 915 in *rrs*. 54% were caused by missense mutations at codons 43 or 88 of rpsL whilst 25% showed no mutations in the 530/915 regions of *rrs* or in *rps*L (Sreevatsan et al. 1996).

Isoniazid is a bactericidal drug that is thought to prevent mycolic acid synthesis and therefore cell wall production in *M. tuberculosis*. Isoniazid is a prodrug that is converted to isonicotinic acid by catalse-peroxidase and is used in first line therapy against *M. tuberculosis*. An association was made in the 1950s between isoniazid resistance and a reduction in catalase-peroxidase activity (Cohn et al. 1954; Hedgecock & Faucher 1957; Middlebrook 1954; Middlebrook, Cohn, & Schaefer 1954). Missense mutations, or small deletions or insertions have been found in the gene encoding catalase-peroxidase (*kat*G) in ~50-60% of isoniazid resistant isolates (Musser 1995). The most common mutation is a serine to threonine at codon 315 (Musser et al. 1996), although this differs with geographic location (Marttila et al. 1996; Rouse et al. 1995).

Case finding and prompt, appropriate treatment to prevent further dissemination of the disease is key to controlling TB. This has been hampered, as with many other infectious diseases, by the continuing emergence of drug resistance. Drug resistance in bacteria may occur by several mechanisms. These include; the alteration of the drug target, modification of the antimicrobial agent, reduced permeability to the agent, its removal by efflux mechanisms or metabolic bypass by the utilisation of alternative pathways.

Generally, the genetic mechanisms of resistance in bacteria arise by a number of routes, including the acquisition of genes by plasmids or other transposable elements (Guiney, Jr. 1984; Lacey 1984; Maiden 1998; Ochman, Lawrence, & Groisman 2000), recombination of foreign DNA into the chromosome (Johnsborg, Eldholm, & Havarstein 2007) or spontaneous mutational events (Davies 1994).

Modification of the intended target of the antimicrobial agent is seen in the case of glycopeptides, fluoroquinolone and rifampicin resistance. Modification of the cell wall leads to glycopeptide resistance in Enterococci (Werner et al. 2008). DNA gyrase and topisomerase IV, the targets of fluoroquinolones, are altered by mutations in *gyr*A/B or *par*C/E, respectively (Chen & Lo 2003).

Some organisms, such as Staphylococci, produce enzymes known as β-lactamases. These enzymes cleave the β-lactam ring at the centre of penicillins and cephalosporins, making them inactive (Drawz & Bonomo 2010). The emergence of drug resistant organisms necessitates the development of additional antimicrobial agents. Several generations of

cephalosporins are now in clinical use in the UK, but the emergence of extended spectrum β-lactamases in Enterobacteriaceae has continued (Pitout & Laupland 2008).

An organism may reduce its exposure to an antimicrobial agent, and therefore have a reduced susceptibility, by either being less permeable to the agent, or by actively reducing the intracellular concentration via efflux pumps. Most antimicrobial agents gain intracellular access via porins. For example, imipenem resistant *Pseudomonas aeruginosa* have been shown to be deficient in the OprD porin (Wang & Mi 2006), which allows the uptake of imipenem. Additionally, high level chloramphenical resistance has been described in *Burkholderia cepacia* with reduced permeability of the outer membrane (Burns, Hedin, & Lien 1989). Efflux pumps actively remove toxic substances from both Gram positive and Gram negative organisms (Van Bambeke F., Balzi, & Tulkens 2000).

An organism may utilise an alternative pathway, upon which an antimicrobial agent may not have an effect. Resistance to sulphonamide and trimethoprim occurs by altered production of dihydropteroate synthetase and dihydropholate reductase, respectively (Then 1982).

Drug resistance in *M. tuberculosis* occurs spontaneously and at random (David 1970). Using the Luria and Delbrück fluctuation test (Luria & Delbruck 1943), Hugo David calculated the mutation rates for *M. tuberculosis* $H_{37}Rv$ to isoniazid, streptomycin, ethambutol and rifampicin. The mutation rate to isoniazid is approximately 100 times

higher per bacterium per generation than to rifampicin – 2.56 x 10^{-8} vs 2.25 x 10^{-10} (David 1970). However, mutation rates of *rpo*B of between 1.3 x10^{-7} and 3.3 x10^{-6} have been reported (O'Sullivan, McHugh, & Gillespie 2008). If the selective pressure of antimicrobial treatment is present, then these resistant organisms will become fixed in the bacterial population due to this advantage. The development of drug resistance is often seen in non-compliant patients (Mahmoudi & Iseman 1993).

Low level resistance to isoniazid, as well as to ethionamide (a structural analogue of isoniazid) is conferred by mutations in a two-gene operon *mab*A and *inh*A (Ramaswamy & Musser 1998). In addition, a study in the Netherlands found that 11 of 51 clinical isoniazid resistant strains showed a C→T substitution upstream of the presumed ribosome binding site of *mab*A. These strains lacked mutations at codon 315 in *kat*G (Musser, Kapur, Williams, Kreiswirth, van, & van Embden 1996). It is hypothesised that these changes result in an increase in InhA protein production and therefore isoniazid resistance by target saturation (Ramaswamy & Musser 1998).

Rifampicin is a bactericidal drug used in first line therapy against *M. tuberculosis*. Studies in *E. coli* have shown that rifampicin binds to the β-subunit of RNA polymerase, resulting in the inhibition of transcription initiation (McClure & Cech 1978). Rifampicin resistance has been well characterized to be the result of point mutations or small deletions in the RNA polymerase gene (*rpo*B) (Miller, Crawford, & Shinnick 1994; Musser 1995; Telenti et al. 1993). Importantly, >95% of resistance causing mutations are found in an 81bp region of *rpo*B (codons 507 – 533, encoding 27 amino acids) termed the

RRDR (rifampicin resistance determining region) (Donnabella et al. 1994; Heym et al. 1994; Kapur et al. 1995; Kapur et al. 1994; Morris et al. 1995; Musser 1995; Telenti et al. 1993). The most common mutation sites in the RRDR are at codon 531 (42%) and 526 (23%) (Musser 1995).

1.7.3 Fitness

The often-used term 'fitness' in relation to bacteria and drug resistance encompasses a number of overlapping factors. The term implies that there is a variation that is inheritable between different members of the same species. With regard to infectious pathogens, the term can be taken to be the representation of an organism's ability to survive, reproduce and be transmitted between hosts. The factors that influence the organism's ability in each of these areas are of course numerous. They include the growth characteristics, the ability to survive within (and outside) a host and the ability to be disseminated.

Some of the characteristics that contribute to an organisms' fitness are measurable in the laboratory, such as the determination of growth rates in liquid culture (Bennett, Dao, & Lenski 1990; Kugelberg et al. 2005), infectivity in animal models (Bjorkman et al. 1999; Bjorkman, Hughes, & Andersson 1998) or the ability to survive a particular challenge. The ability to spread between hosts in a given population can be measured epidemiologically. The dissemination of a resistant strain in a population can be compared with sensitive strains to determine whether the acquisition of drug resistance has affected its transmissibility.

It had previously been assumed that the acquisition of antimicrobial resistance results in a fitness cost (Gillespie & McHugh 1997). This has been shown to be partially true, but the full picture is far from clear. A spontaneous mutation that confers drug resistance will therefore provide an advantage in an appropriately selective environment. If the mutation affects an essential function and therefore incurs a metabolic cost, then it would be sensible to hypothesise that the mutant will be 'less fit' than its sensitive precursor. However, this is not always the case. A fitness cost is usually, but not always, seen initially. For example, Bjorkman and colleagues described streptomycin resistant *Salmonella typhimurium* strains with different mutations at the same locus in *rps*L. Some strains demonstrated a reduced fitness, whilst some were comparable to wild type. In this case, fitness was measured by both growth rates in murine models and optical density in liquid culture (Bjorkman et al 1999).

Another assumption was that when a selective pressure was removed, for example an antibiotic, then the less-fit resistant mutants would revert to sensitive wild type. Studies have shown that this does occur, but not often. For example, true reversion has been described in the *fus*A gene in fusidic acid resistant *Salmonella typhimurium* in murine models (Andersson & Levin 1999). More commonly, compensatory mutations occur which allow the organism to retain the resistance mutation whilst restoring fitness to a level that was previously present. These mutations may be within the same gene as the resistance-causing mutation (intragenic) or elsewhere (extragenic). Bjorkman and colleagues described intragenic compensatory mutations in the *rps*L, *gyr*A and *rpo*B genes of *S. typhimurium* resistant to streptomycin, naladixic acid and rifampicin,

respectively. In each case, the drug resistance was maintained (Bjorkman et al 1999). Extragenic compensatory mutations in *rps*D/E have been described in *E. coli* strains with streptomycin conferring mutations in *rps*L (Schrag, Perrot, & Levin 1997).

1.7.4 Epidemiology of Drug Resistance

With hindsight, the rapid acquisition of drug resistance, especially to mono-therapy, was inevitable. Resistance arises due to genetic mutations brought about by selective pressure due to inappropriate or incomplete antimicrobial treatment (Aziz & Wright 2005). This may be a result of erroneous prescribing practices or erratic dosing on the part of the patient (Mahmoudi & Iseman 1993). Fitness and drug resistance in *M. tuberculosis* are explored in chapter 6.

Standard first line therapy for sensitive respiratory disease in the UK consists of six months treatment with rifampicin and isoniazid, supplemented for the first two months with pyrazinamide and ethambutol or streptomycin. (NICE 2006). The most commonly seen drug resistance in clinical practice in the UK is isoniazid mono-resistance. The additional acquisition of resistance to rifampicin constitutes multi-drug resistance (MDR-TB) (Kruijshaar et al. 2008; Zager & McNerney 2008). The emergence of MDR-TB necessitated the use of other antimicrobial agents and of course, in Darwinian fashion, the development of further resistance.

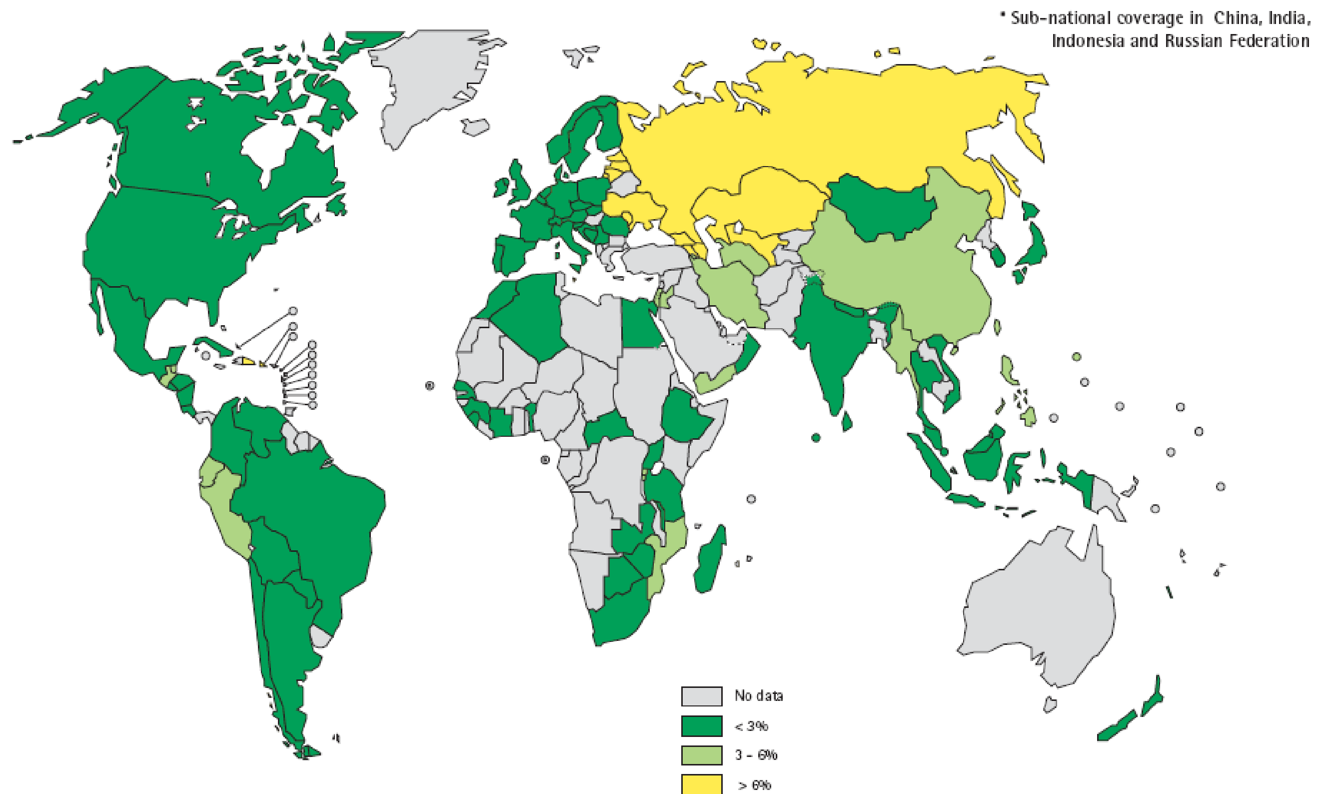


Figure 1.17. MDR-TB among new cases 1994-2007 (WHO 2008).

In 2000 the WHO's Stop TB partnership noted a number of cases that were resistant to all second line agents (Anonymous 2006). On this basis, the WHO defined those strains that are resistant to isoniazid and rifampin and at least three of the six main classes of second line agents (aminoglycosides, polypeptides, fluoroquinolones, thioamides, cycloserine, and para-aminosalicyclic acid) as extensively drug resistant (XDR). Using this definition, surveillance between 2000 and 2004 by CDC and WHO of over 17,000 strains in an international network of TB laboratories domonstrtaed rates of MDR-TB at 20% and XDR-TB at 2% (Anonymous 2006).

Clinicians in a rural hospital in the KwaZulu Natal province of South Africa in early 2005 noted high mortality rates in TB patients who were co-infected with HIV. Surveillance showed that of over 1500 TB patients, 221 had MDR strains and 53 of these had further resistances (Gandhi et al. 2006). These 53 strains were resistant to isoniazid, rifampicin and at least three of the six main classes of second-line drugs (aminoglycosides, polypeptides, fluoroquinolones, thiamides, cycloserine and PAS). The resistance profile was termed extensively drug resistant TB (XDR-TB). All 53 patients were HIV positive and over half claimed to have never been previously treated for TB, indicating primary acquisition of XDR-TB. All but one patient died with a mean survival of 16 days from the date sputum was first collected (Gandhi et al. 2006). XDR-TB was redefined by WHO in October 2006 to include the resistance to isoniazid, rifampicin, a fluoroquinolone and at least one of three injectable second-line drugs (amikacin, kapreomycin or kanamycin) (Anonymous 2007). This definition still stands today.

The global spread of MDR and XDR-TB pose a huge challenge to the control of the disease. Surveillance by the Global Project on Anti-Tuberculosis Drug Resistance between 2002 and 2007 showed that resistance to any drug was seen 11.1% (Wright et al 2009). Rates of MDR-TB increased in virtually all counties. Rates of 7% were seen in two provinces in China, between 6·8% and 22·3% in nine countries of the former Soviet Union, including 19·4% in Moldova and 22·3% in Baku, Azerbaijan. Five countries, all from the former Soviet Union, reported 25 cases or more of XDR-TB each.

Control of TB depends on early detection and treatment of cases and it should also be noted that the prevalence of drug resistance is higher amongst individuals who have

previously been treated for TB (WHO 2008). Additionally, the prevalence of MDR and XDR-TB in an area or country is inversely proportional to the quality of the TB programs (Matteelli et al. 2007). The most recent estimates of global prevalence of MDR (figure 1.17) and XDR-TB (figure 1.18) are shown.

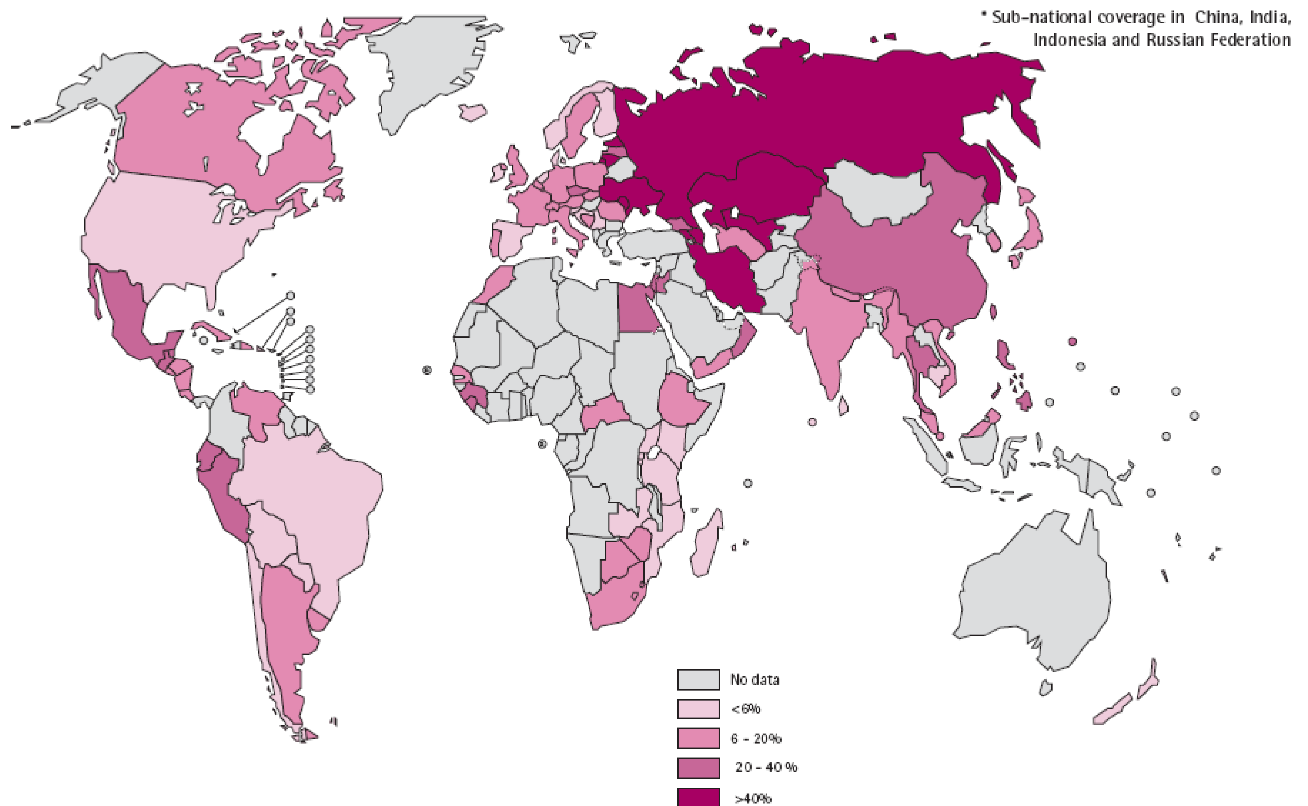


Figure 1.18. XDR-TB among MDR-TB cases 2002-2007 (WHO 2008)

Rates of drug resistance in the UK have been stable for the past ten years (figure 1.19) and in 2008, 6.8% of *M. tuberculosis* isolates in the UK were resistant to at least one first line drug. The majority (6%) were resistant to isoniazid, while 1.1% were resistant to both isoniazid and rifampicin, constituting MDR TB (Health Protection Agency 2009). The largest proportion of drug resistant cases in the UK (48.6% in 2008) is seen in London (Health Protection Agency 2009). However, it should be noted that although

resistance rates are stable, the overall rise in TB cases means that the absolute number of patients with drug resistant TB is increasing.

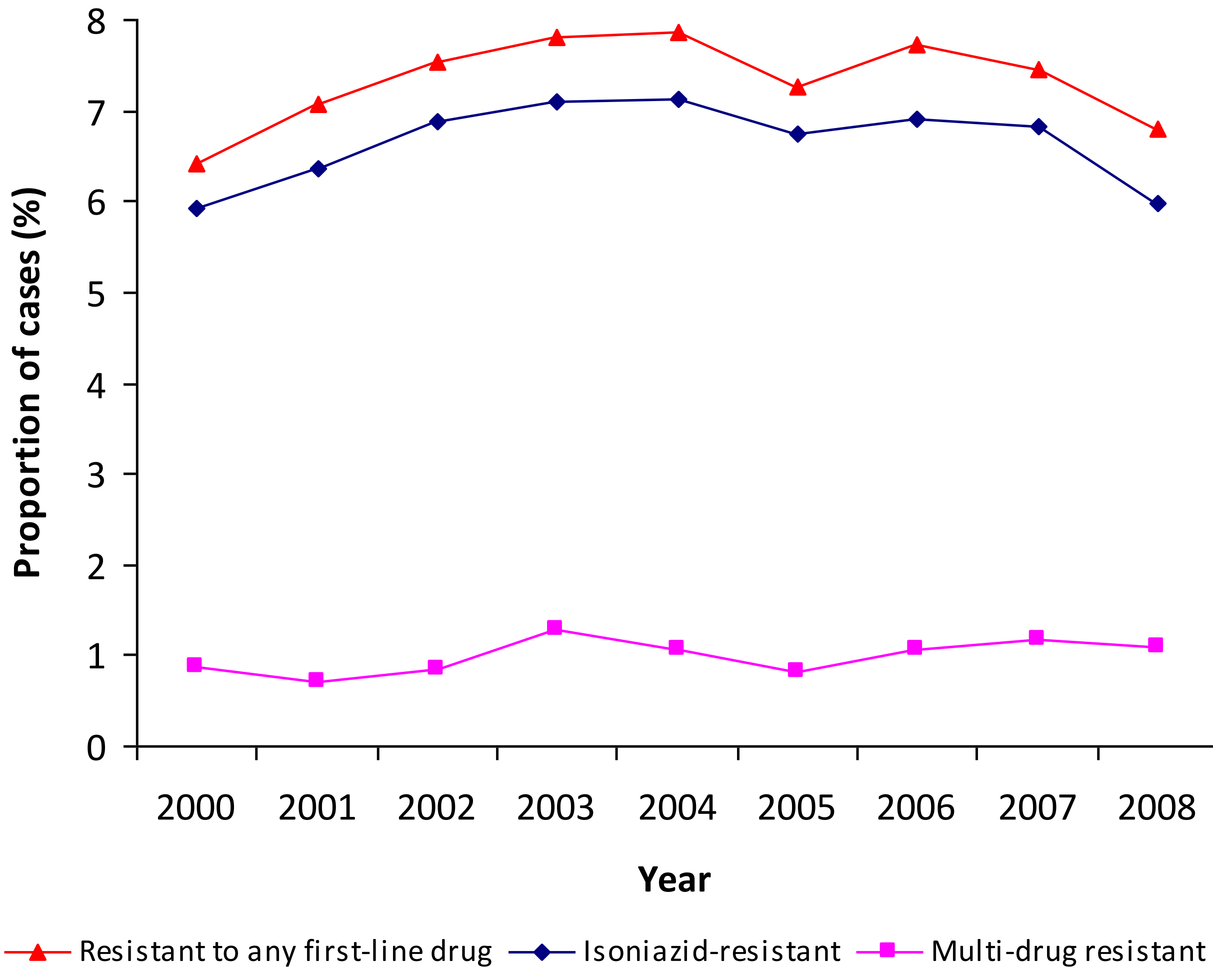


Figure 1.19. Proportion of tuberculosis cases with first-line drug resistance, UK, 2000-2008 (Health Protection Agency 2009).

It may occasionally be necessary to visually supervise the adherence to therapy for some patients to ensure treatment completion and minimise the risk of relapse. NICE guidance recommends that although the use of directly observed therapy (DOT) is not usually necessary in the management of most cases of active TB, all patients should have a risk

assessment for adherence to treatment. Some patients, such as street- or shelter-dwelling homeless people and those with likely poor adherence, in particular those who have a history of non-adherence are more likely to benefit from DOT (NICE 2006). Indeed, a study by Story and colleagues has shown that only 15% of homeless people, drug abusers and prisoners with TB start on DOT, but 46% end up on DOT due to poor compliance (Story et al. 2007).

It should be noted that the treatment of drug-resistant TB is less effective than first line therapy and causes more drug related toxicity. A number of studies comparing clinical outcomes in MDR have been performed – see (Chan & Iseman 2008) for review. However, due to differences in data collection, end points and methods of drug-susceptibility testing, it is not possible to directly compare these studies. Chan and colleagues did, however, identify risk factors that were associated with poor outcome in MDR and XDR-TB. These include: prior TB treatment, the number of anti-microbial agents that the strain was resistant to, resistance to (or prior use of) fluoroquinolones, HIV positivity, a history of imprisonment, a low BMI and the involvement of extrapulmonary sites (Chan & Iseman 2008).

Two studies from the same institution have demonstrated that the care of patients with drug-resistant TB has improved in some settings, however. Goble and colleagues (Goble et al. 1993) studied 171 patients with pulmonary MDR-TB between 1973 and 1983. Only 65% initially responded to treatment (as measured by three consecutive culture-negative sputum samples over three months). The long-term ‘cure’ rate at 51 months was 56% and

the overall mortality rate was 37%. Between 1984 and 1998 Chan and colleagues (Chan et al. 2004) studied 205 MDR-TB patients. The initial success rate was 85%, long-term cure rose to 75% and mortality had declined to 12%.

Furthermore, it is far more costly to treat a case of MDR TB compared to sensitive disease. In 2000 it was estimated that the cost of treating a single case of MDR TB was approximately £60,000 compared to £6,040 for susceptible disease (White & Moore-Gillon 2000). Additionally, the patients who do not respond to treatment but survive may extend the period of infectiousness, which may cause increased transmission (Goble et al. 1993).

As clinicians in the UK have limited experience managing drug resistant cases an MDRTB Service was established at the Cardiothoracic Centre in Liverpool and now the British Thoracic Society has been operational since 1 January 2008. The service has the support of the relevant professional bodies, including the British Thoracic Society, the British Infection Society, and the Health Protection Agency (Davies & Cullen 2008). This expert group can give direct advice and support nationwide via an electronic network.

1.8 TB and HIV

The HIV pandemic plays a hugely significant role the global landscape of TB. The infection of CD4 T cells seen in HIV is central to this phenomenon. There is a range of CD4 cells from the early activated cells that produce only IL-2, cells making IFN-γ and

multifunctional cells expressing IL-2, IFN and TNF. These multifunctional cells are associated with protection (Cooper 2009). The loss of these cells increases the likelihood of succumbing to the disease (Havlir & Barnes 1999).

Certain associations have been made regarding the phenotypes of TB disease in HIV patients. For example, Marshall and colleagues (Marshall et al 1999) noted a higher rate of pulmonary TB in HIV patients. This may be due to a higher rate of recently acquired disease. It has previously been shown that the diagnosis of TB in HIV infected individuals increases the risk of death (Perneger et al. 1995; Whalen et al. 1995), and there is even some evidence that suggests that *M. tuberculosis* may accelerate the progression of HIV disease (Wallis et al. 1993). An increased death rate from TB in HIV patients was also observed by Atun and colleagues (Atun et al 2005). However, there was no such association in a recent study in London (Rodger et al. 2010). This study also found no link between HIV and smear status in TB patients. Higher rates of drug-resistant TB were observed in HIV patients in the UK between 1993 and 1997 by Djuretic and colleagues (Djuretic et al 2002). However, this was not observed in later studies (French et al. 2009; Rodger et al. 2010).

TB was assigned as an AIDS-defining illness in 1998 and rates of HIV diagnosis increased partially due to this alongside the increase in immigration to the UK from areas with high rates of HIV (figure 1.20). Despite this, TB often occurs in HIV patients with more intact immune systems than other AIDS-defining illnesses (Antonucci et al. 1995). The proportion of TB patients who were co-infected with HIV increased in England and Wales from 3.1% in 1999 to 8.3% in 2003. HIV co-infected patients contributed to

almost a third of the increase in the number of cases of TB during this period (Ahmed et al. 2007). Indeed, universal testing for HIV in TB patients has been advocated for more than a decade. A recent study in London found a wide range in the percentage of TB patients being offered an HIV test. Overall, approximately 50% of patients were offered an HIV test. However, this varied widely between centres, with only one offering tests to over 80% of patients and more than half offering a test in fewer than 50% of cases (Bonora & Di 2008; Rodger et al. 2010).

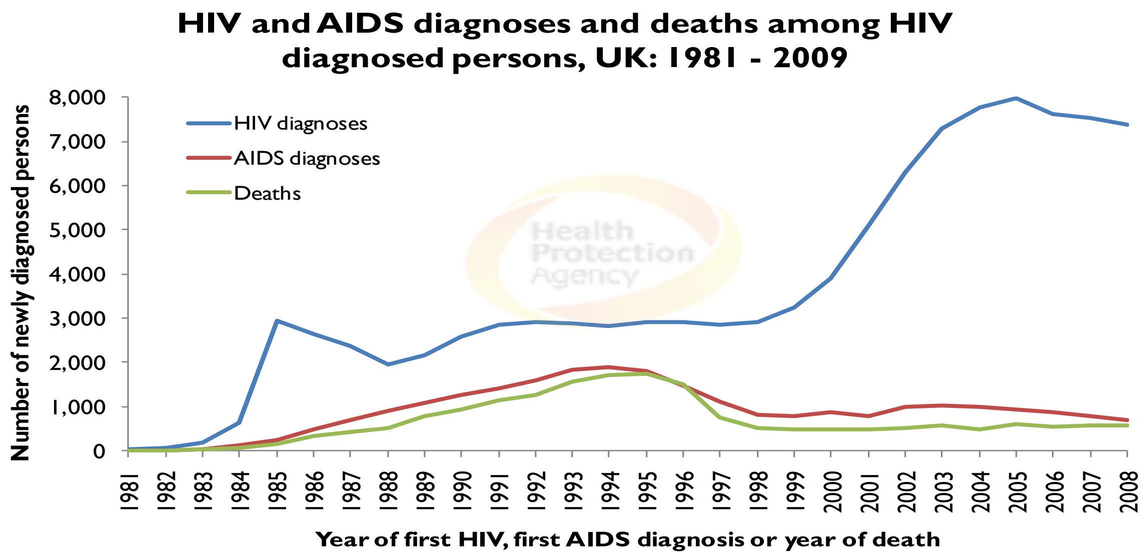


Figure 1.20 HIV notifications in the UK (Health Protection Agency HIV and AIDS reporting service)

Finally, there are the complex interactions between antituberculosis drugs and highly active antiretroviral therapy (HAART) to consider in this important patient group (Bonora & Di 2008). The British HIV Association has issued guidelines for the treatment of TB/ HIV co-infection 2009 (British HIV Association 2009). It is necessary to

effectively treat TB whilst reducing the HIV viral load and maintaining CD4 counts. This has been shown to be achievable (Breen et al. 2006).

One clinical presentation seen in these patients is the phenomenon of immune reconstitution inflammatory syndrome (IRIS). This syndrome contains a wide range of symptoms and is not restricted to infection to HIV/TB infection (French, Price, & Stone 2004; Lipman & Breen 2006). IRIS may be seen in HIV patients co-infected with *M. tuberculosis*, *M. avium* or *Cryptococcus neophormans* (Shelburne et al. 2005) and other infectious and autoimmune conditions (French, Price, & Stone 2004). *M. tuberculosis*-related IRIS in HIV usually presents within the first two months of commencing HAART but more commonly within the first 2–3 weeks. The most common presenting features are severe fever, intra-thoracic and cervical lymphadenopathy and pulmonary infiltrates. Extra-pulmonary disease is less commonly seen, but may include focal cerebritis, pleural effusions, hepatosplenomegaly and ascites (French, Price, & Stone 2004). It is also seen in non-HIV infected TB patients, although far less frequently (Breen et al. 2004), and may be compartmentalized, affecting only one area of the body (Wilkinson et al. 2005).

1.9 Genotyping

In addition to elucidating the evolutionary origins of *M. tuberculosis*, as already stated, genotyping of *M. tuberculosis* aids the investigation of known or suspected outbreaks, identifies laboratory cross contamination, facilitates the estimation of the extent of recent transmission and distinguishes between reactivation and re-infection. Identifying clusters and associated lineages can be useful in investigating virulence, transmissibility and drug

resistance. In essence, genotyping should measure genetic markers that are representative of the genome as a whole. The changes in these markers may then be used to hypothesise the degree of relatedness and evolutionary paths between strains.

Molecular fingerprinting has been used in the UK to confirm or refute outbreaks, contributing to the management of an outbreak of sensitive tuberculosis centred around a high school in Leicester (Drobniewski et al. 2003) and an ongoing outbreak of isoniazid resistant tuberculosis in London (Ruddy et al. 2004). The former outbreak was largely identified by conventional epidemiology although strain typing was used to definitively include or exclude patients.

1.9.1 Contamination

Molecular strain typing *M. tuberculosis* isolates is useful in the monitoring of laboratory cross contamination. Cross contamination is an ongoing problem due to the ability of the organism to survive in the environment for long periods and the sensitive culture systems used. Contamination rates of between 0 and 60% have been reported (Ruddy et al. 2002). However, the highest reported rates were in studies where the laboratory suspected that contamination was occurring and contamination rates of below 3% are more common (Burman & Reves 2000; Ruddy et al. 2002). The identification of a false-positive culture from a clinical sample may have serious consequences as unnecessary drug treatment may be commenced, which can therefore cause toxicity. Additionally, unnecessary investigations, hospitalisation and contact tracing may be performed.

1.9.2 Re-infection vs. Relapse

The presentation of a patient with symptoms consistent with TB following a previous episode usually falls into one of two categories. The patient may have been inadequately treated and a relapse with the original strain may have occurred. This may be complicated by the acquisition of drug resistance. Alternatively, the patient may have been successfully treated for their original infection, but has become re-infected with another strain.

These two scenarios are impossible to distinguish on clinical presentation alone. Genotyping of any organism isolated allows comparison to previous isolates. An indistinguishable genotype indicates relapse (with or without acquired drug resistance) whilst a different genotype would suggest re-infection. Figure 1.21 shows the sequential IS*6110* RFLP fingerprints from *M. tuberculosis* isolates cultured from a non-compliant TB patient. The patient re-presented over the course of over one year with recurrence of their disease. Genotyping was performed on each isolate to determine whether the patient has relapsed or been infected with a novel strain. The genotyping demonstrated that the patient was infected with the same strain and was relapsing. The asterisks indicate the time points at which drug-susceptibilities were performed. The organism was fully susceptible at week 31, was isoniazid mono-resistant at week 52 and additionally acquired rifampicin resistance at week 67. The acquisition of drug resistance did not affect the IS*6110* RFLP fingerprint.

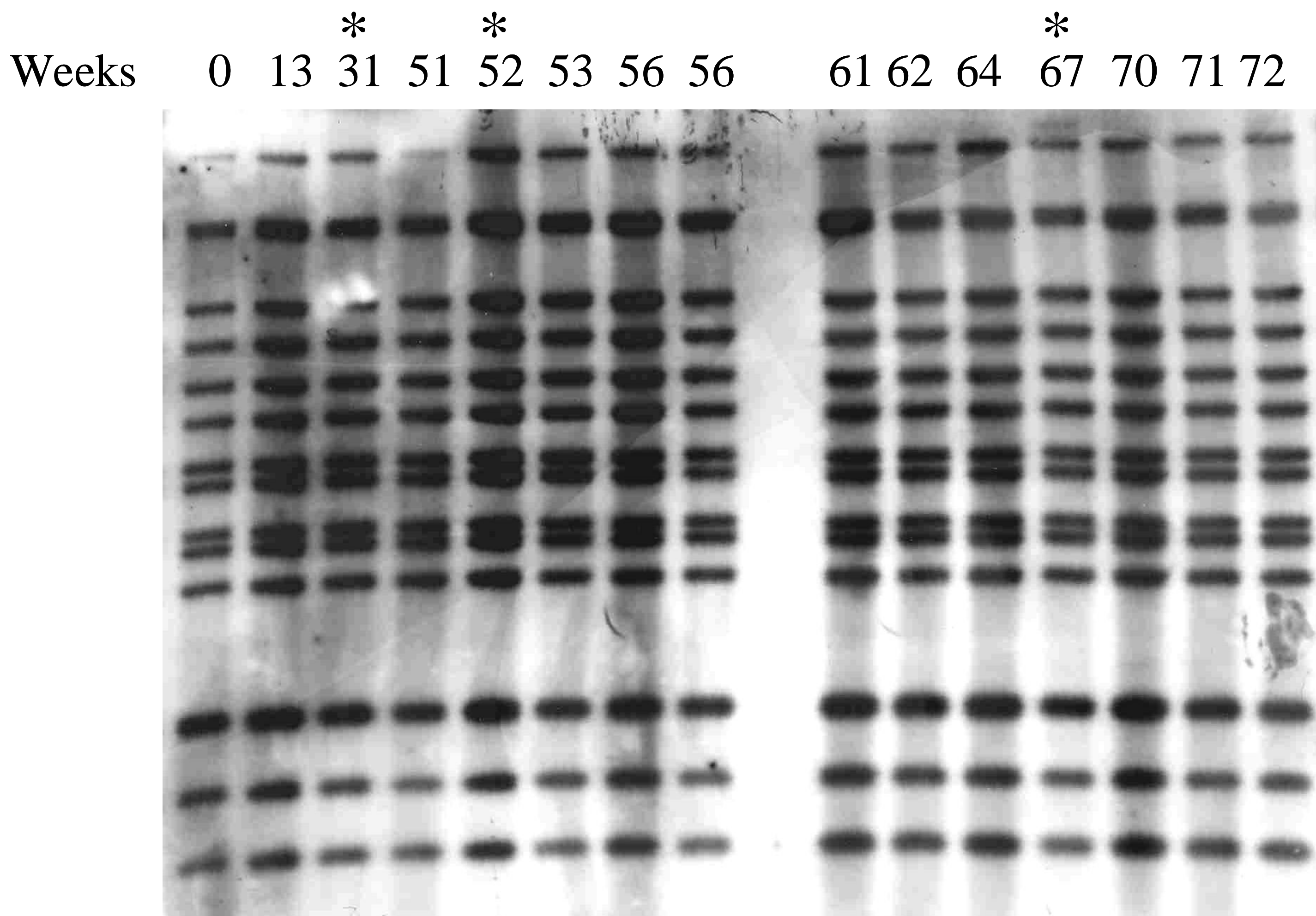


Figure 1.21. IS*6110* RFLP of sequential isolates from an individual patient. Genotyping confirmed that the patient was undergoing multiple relapses and acquiring drug resistance (identified at weeks 31, 52 and 67) rather than being reinfected with different strains. Courtesy Dr. Nicky Hutchinson, St. George's University of London.

1.9.3 Defining the Relatedness of Strains by Genotyping: Clusters, Lineages and Superfamilies

It is important to understand the differences between genotyping and epidemiological data, as well as acknowledging the wealth of knowledge that they can cumulatively provide. Epidemiological links between patients and the collection of demographic data provide the backbone in outbreak investigation. The collection of complete epidemiological data is essential for any study or investigation. Alongside a thorough clinical history, including the date of onset of symptoms, a range of patient demographics must be noted. These may include the place of residence and work, country of birth, travel history and social habits to ascertain potential exposure. These may then be analysed alongside genotyping data.

A number of different techniques have been used to genotype *M. tuberculosis* and these have been thoroughly reviewed (Barnes & Cave 2003a; Kanduma, McHugh, & Gillespie 2003; Mathema et al. 2006) for reviews. Several factors are important when considering a genotyping method. Reliable reproducibility and transportability are crucial to allow the comparison of data between centres. The technical difficulty of the method plays a role in the reproducibility obtained. The level of discrimination is very important when interpreting data. This discriminatory power varies between methodologies and is directly related to the number of subdivisions in a group of strains that a genotyping method generates. That is, a highly discriminatory method will subdivide a cluster of strains into more groups than one with a lower discriminatory power. This will be explored later in chapters 3 and 4.

It is not always necessary to use the genotyping method that is most discriminatory, however. If, for example, a method with a low discriminatory power refutes a suspected outbreak, there is no need to use a more discriminatory method (Shorten et al. 2005).

1.9.4 Genotyping Methodologies

Regardless of the method used, the discriminatory power, and therefore the degree of clustering of genotyped strains will depend directly on the rate of change of the molecular element. The basic methodology and usefulness of the most widely used genotyping tools and their molecular stability are described here.

1.9.4.1 IS*6110* RFLP

Repetitive elements were identified in the *M. tuberculosis* genome independently by Eisenach and colleagues (Eisenach, Crawford, & Bates 1988; Zainuddin & Dale 1989). The insertion sequence IS*6110* was first sequenced and found to be a member of the IS3 family by Thierry and colleagues (Thierry et al. 1990). This group described the insertion sequence as being specific to the *M. tuberculosis* complex, estimated that *M. tuberculosis* possessed between 10 – 20 copies scattered throughout the genome, and hypothesised on its usefulness in epidemiological studies. The insertion sequence was also amplified in clinical samples from patients with suspected and confirmed TB. Others independently identified the insertion sequences IS*986* (McAdam et al. 1990) and IS*987* (Hermans et al. 1990). These two insertion sequences differed from each other, and also from IS*6110*, by

only a few base pairs so were considered to be essentially the same element, designated IS*6110* (van Embden et al. 1993).

The insertion of IS*6110* has been shown to modify regulatory function in some strains. Safi and colleagues (Safi et al 2004) demonstrated that IS*6110* can upregulate downstream genes through an outward-directed promoter in its 3' end. Promotor activity was orientation dependent but upregualtion was seen in during growth in monocytes and in *in vitro* cultures. This suggests that insertion and deletion of IS*6110* may alter the phenotype of the organism as well as being a means by which its evolution may be studied.

Insertion sequences have been used to genotype *M. tuberculosis* in numerous outbreak situations (Cave et al. 1991; Daley et al. 1992; Edlin et al. 1992; Fomukong et al. 1992; Hermans et al. 1990; Mazurek et al. 1991; Otal et al. 1991; van Soolingen et al. 1991) and a consensus was required to allow the standardisation of the methodology that would allow comparison of strain types between centres. Molecular fingerprinting using IS*6110* relies on the cultivation of the organism, genomic DNA extraction, enzymatic restriction, agarose gel electrophoresis, Southern hybridisation and detection of the IS elements with a labelled probe. The three elements that required standardisation were the restriction enzyme, nature of the labelled probe and the molecular size-markers used. In collaboration, researchers who had been instrumental in the discovery and characterisation of IS*6110* concluded that *Pvu*II should be used, since IS*6110* possessed only one restriction site for this enzyme (figure 1.22). The probe was designed to bind

only to the right of the restriction site in order to allow the visualisation of only a single band per IS*6110* copy (fig 1.22). Finally, it was agreed that a molecular size marker ranging between 0.9 and 10kb be used in addition to a reference strain Mt14323 (van Embden et al. 1993).

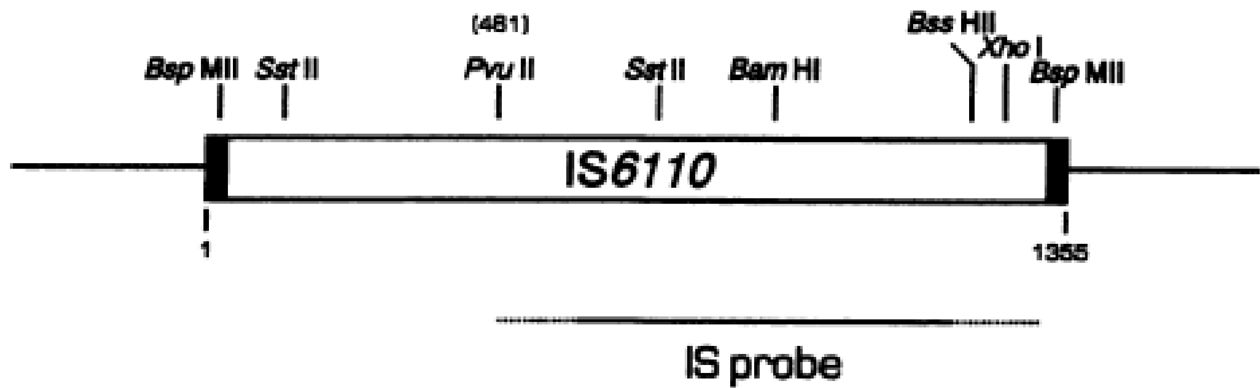


Figure 1.22 Physical map of IS*6110* (van Embden et al. 1993).

It was soon noted that some strains posses few copies of IS*6110* and therefore could not be adequately discriminated by this method (van Soolingen et al. 1993). This study also demonstrated one strain that lacked even a single copy of IS*6110*. This rare event may have a detrimental effect on the use of diagnostic NAATs that target this insertion sequence. Indeed, a respiratory sample from a symptomatic, smear-positive patient gave a false negative TB PCR result. The sample subsequently grew an organism that was identified as *M. tuberculosis* and was shown to lack IS*6110* (Shorten, unpublished data). An example of an IS*6110* RFLP image can be seen in figure 1.23, whilst an example of the distribution of strains with varying numbers of IS*6110* copies can be seen in figure 1.24.

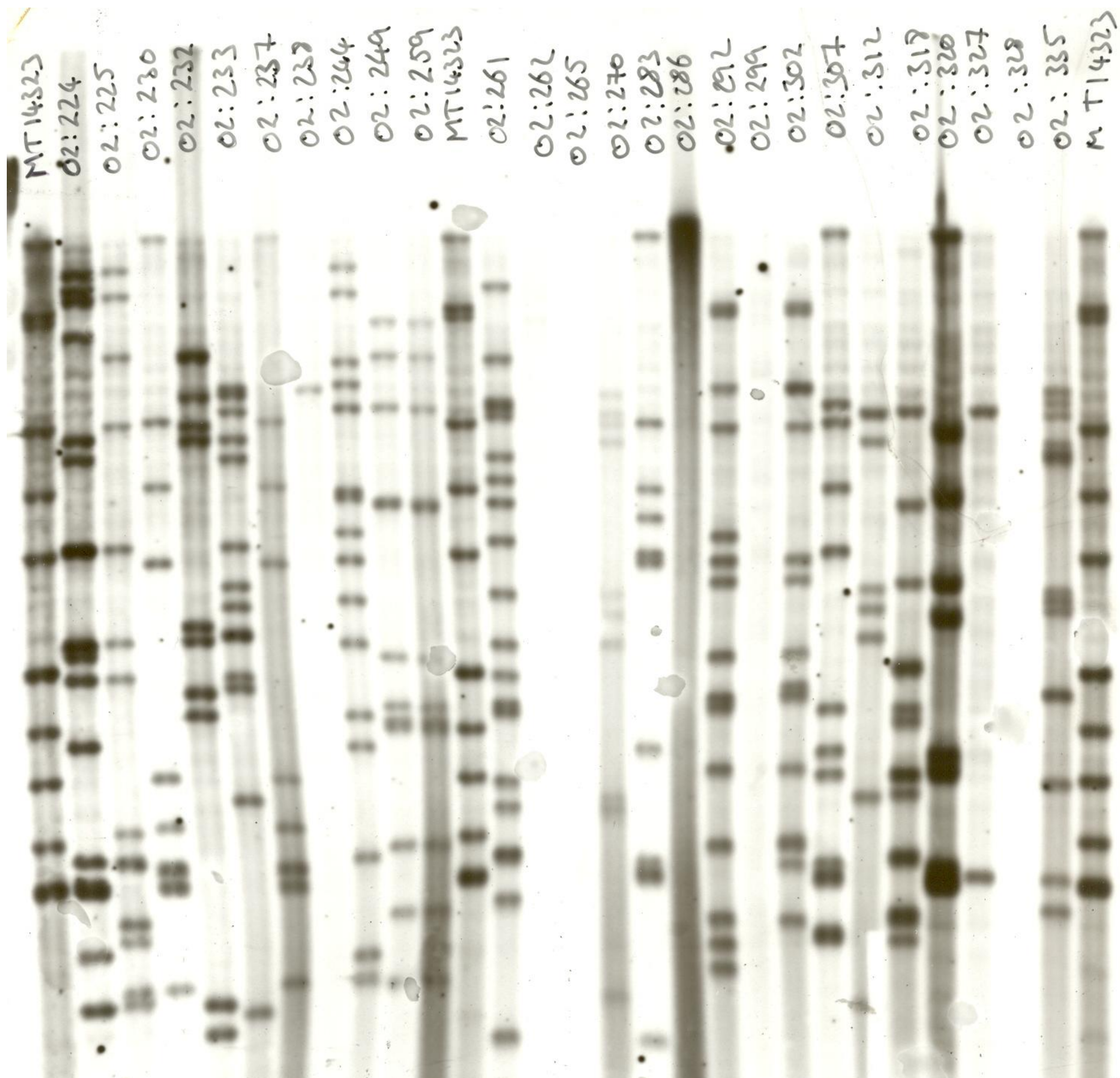


Figure 1.23. Image of the IS*6110* RFLP genotyping patterns of multiple clinical strains. Note the reference strain MT14323 in the first, twelfth and last lanes. Note also the low copy number strains (those possessing fewer than five copies of IS*6110*).

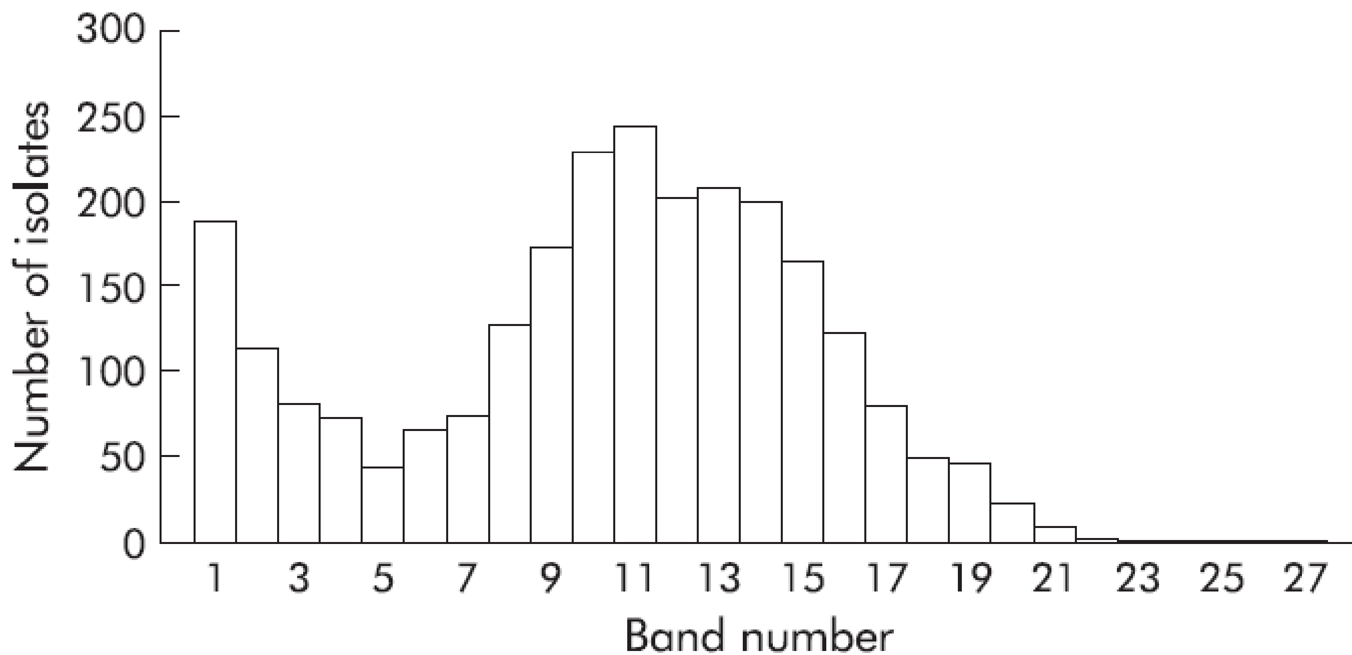


Figure 1.24. Frequency distribution of number of copies of IS6110 for 2490 clinical isolates of *M. tuberculosis* in London between 1995 and 1997 (Maguire et al. 2002)

Knowledge of the rate of change of this insertion sequence is critical when interpreting any genotyping data. The insertion sequences may be duplicated and inserted into the genome, moved to another location in the genome or excised and lost from the genome. These events have occasionally been termed birth, shift and death, respectively (Rosenberg, Tsolaki, & Tanaka 2003). If the rate of change is rapid, then a strain transmitted from one individual to another may produce a different genotype, which would result in this chain of transmission being missed, and an underestimation of ongoing transmission in a community. However, if the rate of change is slow, then organisms that are not closely epidemiologically linked would produce false clusters and would over estimate the rate of transmission.

Several studies have been performed that investigate the rate of change in IS*6110* banding patterns in serial isolates from individuals infected with *M. tuberculosis*. It is difficult to standardise and control such studies as repeat samples from TB patients are usually only obtained should the patient be unwell and re-present to their clinician.

Yeh and colleagues investigated 49 patients with two or more isolates of *M. tuberculosis* cultured from specimens taken at least 90 days apart (range 90 days – 3 years) (Yeh et al. 1998). Changes in the IS*6110* pattern were seen in 12 (25%) patients. IS*6110* changes were not related to HIV status, drug resistance or adherence to treatment. This group concluded that as there was an unexpectedly high rate of change, patients with IS*6110* patterns that differ by a single band should maybe be considered related and be investigated as a possible transmission event.

The investigation of serial isolates from 544 patients showed a much lower rate of change (de Boer et al. 1999). Only 25 (4.6%) of follow-up patient isolates differed from the original strain. This group calculated that in this cohort that there was a change in banding pattern (half-life) on average every 3.2 years (95% CI 2.1 – 5.0). There was an association with increased rate of change in strains cultured from patients with extrapulmonary as well as both pulmonary and extrapulmonary disease. The data collected by Yeh and colleagues (Yeh et al. 1998) was reanalysed by this group and the mean half-life was found to be 2 years (95% CI 1.2 – 3.5) (de Boer et al. 1999).

A longer half-life was described in the high-incidence setting of Cape Town, South Africa. Of 349 repeat patient samples, 14 (4%) demonstrated a variant IS*6110* pattern compared to the initial IS*6110* genotype (Warren et al. 2002). The percentage of altered strains was comparable with the findings of de Boer 4% vs. 4.6%, respectively, (de Boer et al. 1999), but the period between initial and subsequent samples being taken varied between 0 and 2000 days. Analysis showed a mean half-life of 8.74 years. Interestingly however, the 14 strains demonstrating variance in their IS*6110* genotype fell into two distinct groups. Eight of the 14 (57%) were sampled within 20 days of the initial isolate, whilst the remaining six (43%) were sampled after at least 250 days. This gave two distinct groups of half lives – 0.57 years (the early group) and 10.69 years (the late group) (Warren et al. 2002).

A slightly higher rate of change was seen in the study of 56 TB patients in Germany (Niemann, Richter, & Rusch-Gerdes 1999). This group observed that 5 out of 56 patients (9%) were infected with a strain that changed IS*6110* pattern upon repeat isolation. The length of time between first and subsequent isolates varied between 1 and 772 days. There was a correlation between a change in IS*6110* pattern and increased length of time between samples. However, there was no link between a change in IS*6110* pattern and acquisition of drug resistance.

The location of the insertion sequence has been shown to occur in a non-random manner. Certain ‘hotspots’ have been identified where insertion sequences are more likely to be found. Indeed, strains possessing few copies of IS*6110* were more likely to have

insertions in these locations (McHugh & Gillespie 1998). For this reason, isolates that possess fewer than five copies of IS*6110* are termed 'low copy number' strains and are not able to be differentiated by this method. Furthermore, low copy number strains that have been shown to possess the same insertion by sequence analysis have produced different IS*6110* RFLP patterns (Dale et al. 2003a). This may be due to polymorphisms elsewhere that insert or delete a *Pvu*II restriction site.

One should therefore take into account the finer details of this method, including the lack of discrimination of strains with few IS*6110* copies, the rate of change of the banding patterns obtained and the predisposition of the insertion sequence to insert into hotspots in a non-random fashion, when interpreting any data obtained. False clusters have been seen in epidemiologically unlinked isolates. A study in Northern Tanzania demonstrated that three strains of *M. tuberculosis* that were epidemiologically unlinked were indistinguishable by IS*6110* RFLP (Gillespie, Dickens, & McHugh 2000). Further analysis showed that six of the nine copies of IS*6110* that the strains possessed were in hotspots and polymorphic GC-rich repetitive sequence (PGRS) typing (see 1.10.4.4) revealed that they were distinct strains.

IS*6110* RFLP offers good discrimination but is slow, labour intensive, and requires large quantities of DNA (>200ng). Mature cultures are required to provide sufficient quantities of DNA, which makes the use of this method limited in real-time outbreak investigation. It is also unable to distinguish strains with few IS*6110* copies (< 5).

1.9.4.2 Mycobacterial Interspersed Repetitive Units – Variable Number Tandem Repeats (MIRU-VNTR)

Repetitive units of DNA in organisms are well described (Higgins, McLaren, & Newbury 1988; Lupski & Weinstock 1992) and have been used to genotype a large array of bacteria such as *Escherichia coli* (Manges et al. 2009), *Clostridium difficile* (Marsh et al. 2006; van den Berg et al. 2007), *Bacillus anthracis* (Keim et al. 1999; Le Fleche P. et al. 2001), *Yersinia pestis* (Klevytska et al. 2001; Le Fleche et al. 2001), *Listeria monocytogenes* (Lindstedt et al. 2008), *Shigella* spp. (Gorge et al. 2008). Additionally, VNTRs are also used in mapping human genes (Nakamura et al. 1987).

Supply and colleagues identified a novel group of such repetitive units that they named mycobacterial interspersed repetitive units (MIRU). They are mainly intergenic and are distributed throughout the genome. These units were unlike other repetitive units previously described as they contained no obvious palindromic sequences, were direct tandem repeats and were orientated in one direction relative to the transcription of the adjacent genes. Additionally, approximately half of the MIRUs described contained small ORFs and based on sequence homology of the *M. tuberculosis* sequences at the time, it was estimated that there may be 40-50 such MIRUs per genome (Supply et al. 1997). At approximately the same time, analysis of published sequences by Frothingham and Meeker-O'Connell identified a number of repetitive regions in *M. tuberculosis* $H_{37}Rv$. Some of these, termed exact tandem repeats A to F (ETR-A to F), were variable between strains (Frothingham & Meeker-O'Connell 1998).

Further work by Supply and colleagues identified a total of 41MIRU loci when analysing the complete $H_{37}Rv$ genome (figure 1.25) (Supply et al. 2000). These tandem repeats varied between loci and strains and may be amplified with the use of specific PCR primers to each MIRU-flanking region. If a strain lacked a single repeat at a MIRU locus a PCR product would still be generated. The sizing of these PCR products would enable the investigator to identify the number of repeats present at each locus and therefore generate a numerical genotype (figure 1.26).

There was congruence between five of these MIRU loci and ETR A-E previously described (Frothingham & Meeker-O'Connell 1998). The variability between strains of these 41 loci differ and analysis of a set of internationally obtained strains (Kremer et al. 1999) suggested that 12 of these loci exhibited a suitable degree of variability to be used in a genotyping methodology (Supply et al. 2000).

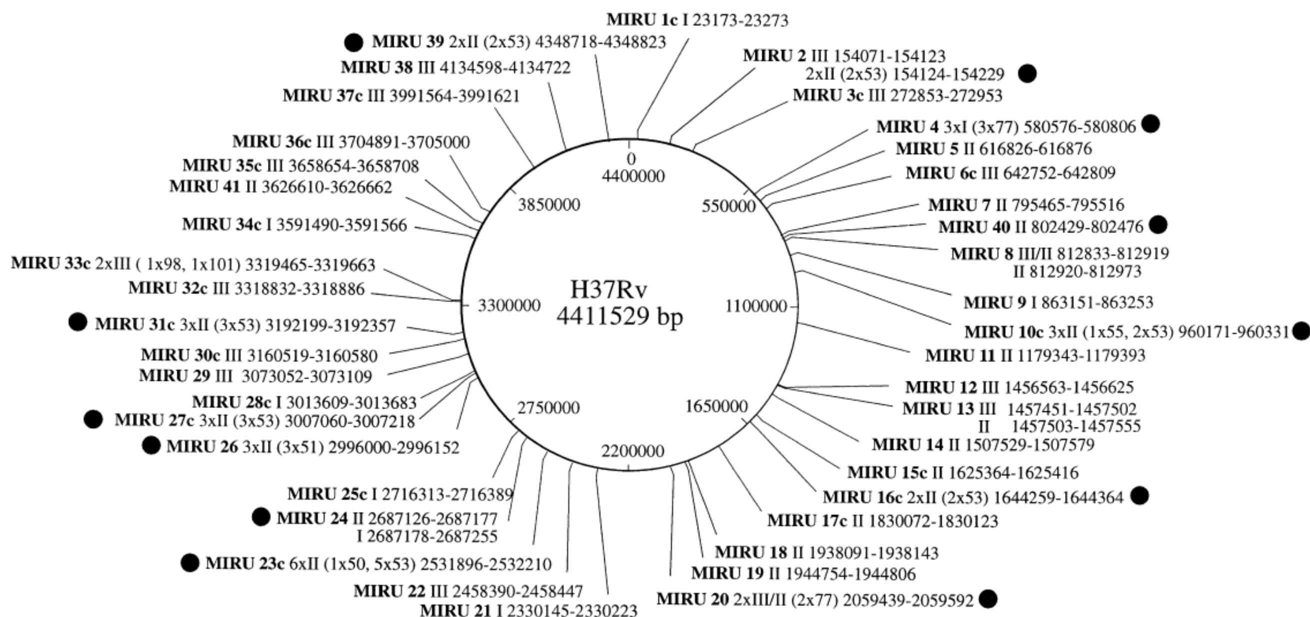


Figure 1.25 Location of MIRU loci in *M. tuberculosis* H37Rv genome (Supply 2000 Mol Micro)

The discriminatory power of this method may be altered by varying the number of MIRU loci included in the analysis. For example, when only five loci were used, Kremer and colleagues found VNTR to be less discriminatory than IS*6110* RFLP, spoligotyping (1.9.4.3) and PGRS (1.9.4.4) (Kremer et al. 1999). Kwara and colleagues analysed 64 *M. tuberculosis* strains clustered into ten groups by IS*6110* RFLP (Kwara et al. 2003). Spoligotyping and 12 loci MIRU were used to attempt to further differentiate these clusters. MIRU analysis differentiated these strains to a higher degree than both IS*6110* RFLP and spoligotyping. In another study, however, 12 loci MIRU alongside spoligotyping was found to be less discriminatory than IS*6110* RFLP (Supply et al. 2006). Five loci VNTR was found to be more discriminatory than IS*6110* RFLP in strains with few copies of IS*6110* (Barlow et al. 2001). The further discrimination of apparent

clusters of high IS*6110* copy number strains by 12-loci MIRU has also been described (van Deutekom et al. 2005). The relative levels of discrimination will of course vary depending on the number, and geographical representation of strains analysed.

As some loci are more variable than others, the discriminatory power of MIRU does not increase in a linear manner with the analysis of additional loci. The more discriminatory a method, the fewer clusters will be generated when strain typing a group of strains. The allelic diversity (h) of each locus may be calculated as by Selander and colleagues as follows (equation 1.1):

$$h = 1 - \Sigma x_i^2 \, [n/(n - 1)],$$

Equation 1.1 Calculation of allelic diversity ((Selander et al. 1986)

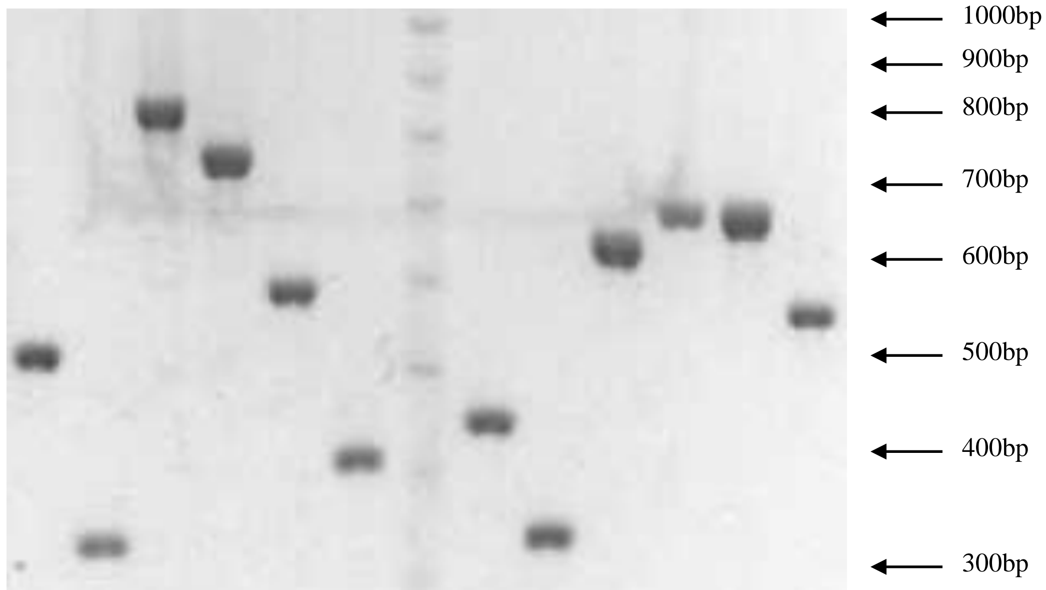
where n is the number of isolates and x_i is the frequency of the ith allele at the locus (Selander et al. 1986).

Various combinations of MIRU loci have been suggested for genotyping of *M. tuberculosis* (Frothingham & Meeker-O'Connell 1998; Goyal et al. 1994; Kam et al. 2006; Kremer et al. 2005; Le Fleche. et al. 2002; Magdalena et al. 1998; Magdalena, Supply, & Locht 1998; Mazars et al. 2001; Roring et al. 2002; Roring et al. 2004; Skuce et al. 2002; Smittipat et al. 2005; Supply et al. 2000; Surikova et al. 2005). Supply and colleagues showed by analysing 29 loci that five were unreliable for use in a genotyping methodology. A well characterised set of 90 strains (Kremer et al. 1999) were investigated with the remaining 24 loci (Supply et al. 2006). When using 24 loci, the 90

strains were separated into 89 distinct genotypes – $H_{37}Rv$ and $H_{37}Ra$ were indistinguishable by this method as well as by spoligotyping, and the addition of the five unreliable MIRU loci. Maximum resolution of these 89 genotypes could be achieved by the analysis of only nine loci (MIRU 04, 10, 16, 26 and 40 and VNTR 0577, 2163b, 2165 and 4052), thus indicating that, in this group of strains at least, these nine loci were the most variable. Analysis of nearly 500 isolates identified the loci that demonstrated the most allelic diversity and were most likely to exhibit single, double and triple locus variants (SLVs, DLVs and TLVs). The removal of the loci with the least diversity generated a subset of 15 MIRU loci with 96% of the discriminatory power of the full set of 24 loci (Supply et al. 2006). It was reasoned that the 15 loci set may be used for epidemiological studies, whilst the 24 loci set would be of use in phylogenetic studies. The Health Protection Agency Mycobacterium Reference Unit (HPA MRU) in Whitechapel, London has adopted the set of 24 loci for prospective analysis from 1st January 2010.

The stability of the original 12 MIRU loci (Supply et al. 2000) was assessed by comparing the MIRU genotypes obtained from serial isolates from patients over up to six years (isolates analysed in (Warren et al. 2002)). A total of 123 isolates, representing at least two isolates from 56 patients, were genotyped (Savine et al. 2002). The selected isolates belonged to a variety of distinct IS*6110* RFLP families. All 12 MIRU loci were identical within 55 of the 56 serial patient groups. One isolate showed a reduction in one repeat in a single locus. Conversely, 11 of the 56 serial patient isolates showed a variation

in their IS*6110* pattern. The MIRU genotype for these isolates remained unchanged, suggesting that 12 loci MIRU was more stable than IS*6110* RFLP.



MIRU locus	2	4	10	16	20	23	24	26	27	31	39	40
Approx. size (bp)	510	320	850	770	590	410	450	320	640	700	700	550
MIRU repeat	2	2	7	4	2	5	1	1	3	4	3	4

Figure 1.26.12-loci MIRU amplicons on an agarose gel with ethidium bromide staining. The first six lanes are MIRU loci 2, 4, 10, 16, 20 & 23, followed by a 1000bp ladder and the remaining six MIRU loci 24, 26, 27, 31, 39 & 40.

MIRU analysis has several technical advantages over IS*6110* RFLP. Only small amounts of genomic material are required as the method involves DNA amplification. The assay itself is extremely simple to perform when compared with the involved process of RFLP.

Moreover, as a digital genotype is generated, the data may be readily transported and compared in other centres. The method can be simplified even further by the utilisation of multiplex PCR and automated DNA sequencing technologies (Supply et al. 2001). An example of sizing PCR amplicons by automated sequencing may be seen in figure 1.27. Additionally, the method may be used to genotype *M. tuberculosis* strains that posses both high and low numbers of IS*6110*.

Furthermore, Brown and colleagues (Brown et al 2010) have shown that MIRU-VNTR types align with other phylogenetic markers and assign strains into the large phylogentic groups as previously reviewed (Gagneux & Small 2007). This group have also shown that there is >90% concordance between MIRU-VNTR types and phylogenetic groups by SNP analysis (Gibson et al 2005).

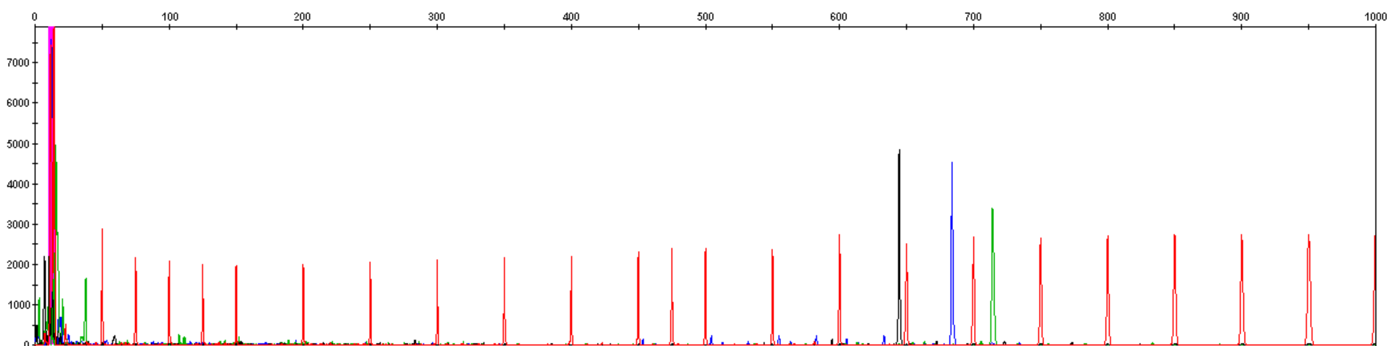


Figure 1.27. GeneMapper Fragment analysis. The x-axis demonstrates the size of the fragment, or amplicon, detected. The y-axis demonstrates the number of fluorescent units detected. The red peaks are fragments in the 1000bb, ROX-labelled, marker whilst the black, blue and green peaks correspond to the PCR products for MIRU loci 960, 1644 and 3192, respectively. The sizes of these three peaks are 645bp, 670bp and 710bp, which correspond to alleles 3, 2 and 4 at these loci.

1.9.4.3 Spoligotyping

Spacer-oligonucleotide typing (Spoligotyping) is a PCR-based typing method that relies on identifying polymorphisms in the spacer units in the direct repeat (DR) region of the genome. The DR region comprises multiple, virtually identical, 36bp regions interspersed with non-repetitive spacer sequences of a similar size (Groenen et al. 1993). As a nucleic acid amplification method, small amounts of DNA are required, and as such, spoligotyping has been used to detect and genotype *M. tuberculosis* from paraffin wax-embedded samples (van der Zanden et al. 1998). Examples of spoligotypes are displayed in figure 1.28.

Variation in these regions was thought to be caused by homologous recombination between adjacent or distinct DRs or by transformation due to the insertion of IS*6110*, which is almost invariably present in the DR region (Groenen et al. 1993). These direct variant repeats (DVRs) differ between strains, but their order remains conserved (van Embden et al. 2000). The DR region contains over 60 such spacers, but 43 were selected as a basis of this typing methodology. The presence or absence of these spacers is detected by the amplification and subsequent reverse hybridisation (Kamerbeek et al. 1997).

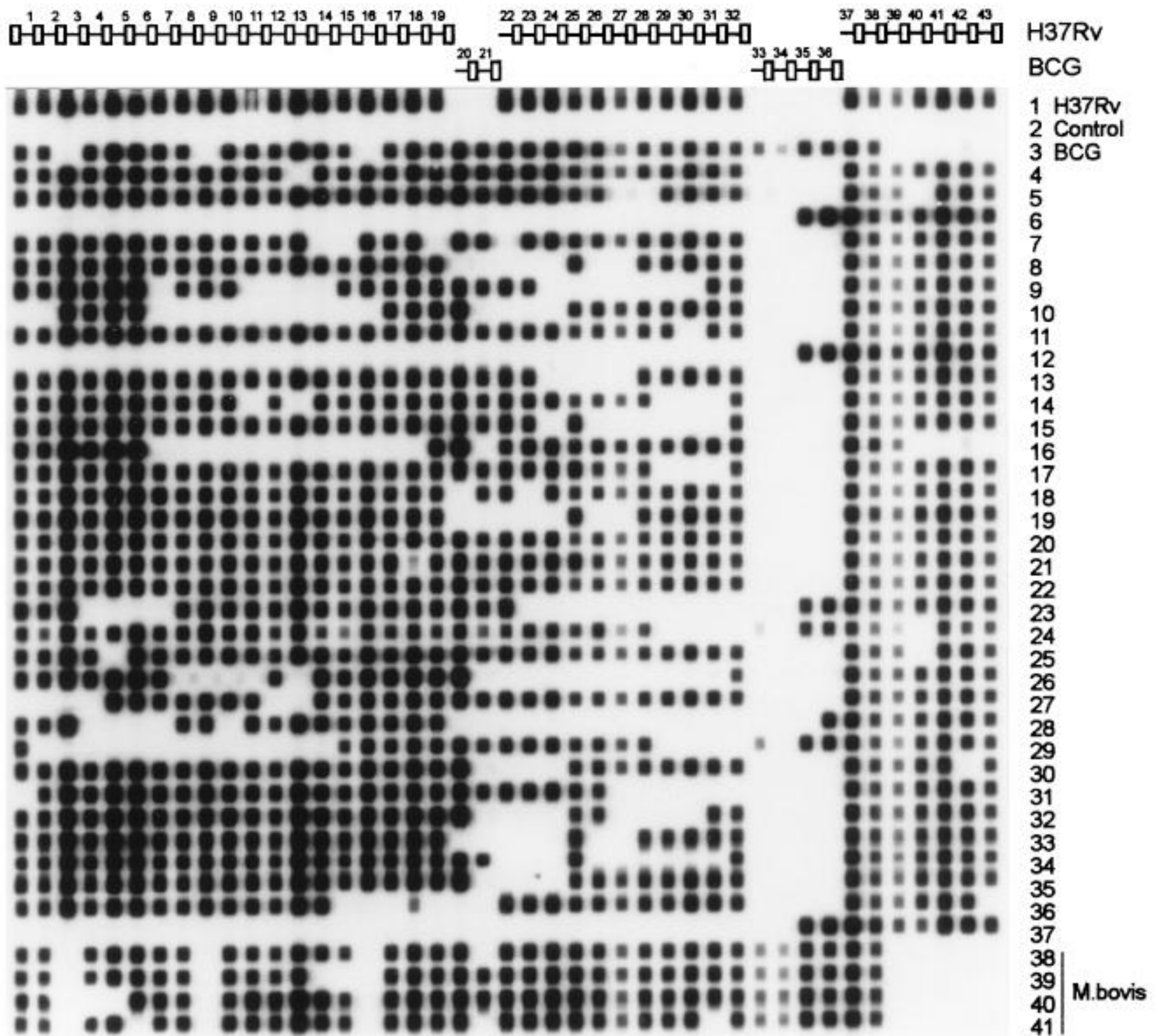


Figure 1.28. Hybridization patterns (spoligotypes) of amplified mycobacterial DNAs of 35 *M. tuberculosis* and 5 *M. bovis* strains. The order of the spacers on the filter corresponds to their order in the genome. Note that the spoligotype of strains 6, 12, and 37 corresponds to that of strains from the Beijing family, as described earlier (Kamerbeek et al. 1997).

This technique is widely used in both evolutionary and epidemiological studies. As such, an agreed nomenclature was proposed to allow comparison of genotypes (Dale et al. 2001). Spoligotyping is superior to IS*6110* RFLP when studying strains with few copies of IS*6110*. However, it is less discriminatory than IS*6110* typing for all other strains (Kremer et al. 1999; Soini et al. 2001).

The analysis of DVRs by Warren and colleagues revealed that the evolution of a spoligotype occurs by four processes; IS*6110*-mediated mutation, homologous recombination between repeat sequences leading to DVR deletion, strand slippage during replication leading to duplication of the DVR sequences and point mutation (Warren et al. 2002). Interestingly, it was demonstrated that the insertion of IS*6110* into the DR region lead to the failed hybridisation with some DVRs, despite their presence. DR RFLP, together with cloning and sequencing showed that these DVRs were present although spoligotyping indicated their deletion. The authors concluded that as the evolution of spoligotypes was not an independent process, that this method should not be relied upon alone for evolutionary studies in *M. tuberculosis*. They advocate its use alongside other methodologies. A diagrammatic representation of the genomic sites utilised in IS*6110* RFLP, MIRU-VNTR and Spoligotyping may be seen in figure 1.29.

1.9.4.4 Polymorphic GC-rich Repetitive Sequence (PGRS) Genotyping

The PGRS element is the most abundant repetitive sequence in the *M. tuberculosis* complex (Poulet & Cole ST 1995b). Like IS*6110* RFLP, PGRS typing utilises Southern blotting and probe hybridisation to differentiate strains of *M. tuberculosis*. This

technique, first described by Ross and colleagues, uses a probe of 3.4kb that is specific to the PGRS sequence and was cloned into the plasmid pTBN12 (Ross et al. 1992). Genomic DNA, restricted with *Alu*I, may be separated by agarose gel electrophoresis and hybridised with a labelled, PGRS-specific probe. This method has been shown to differentiate unrelated strains whilst clustering those that are epidemiologically linked (Ross et al. 1992; Yang et al. 1996). Some studies have shown that PGRS typing is more discriminatory than IS*6110* RFLP, especially when examining strains with few copies of the insertion sequence (Chaves et al. 1996; Rhee et al. 2000; Yang et al. 1996). Computer-based analysis of the fingerprints is possible, however, the intensity of the obtained bands may vary and this means that additional visual verification is required (Kremer et al. 1999).

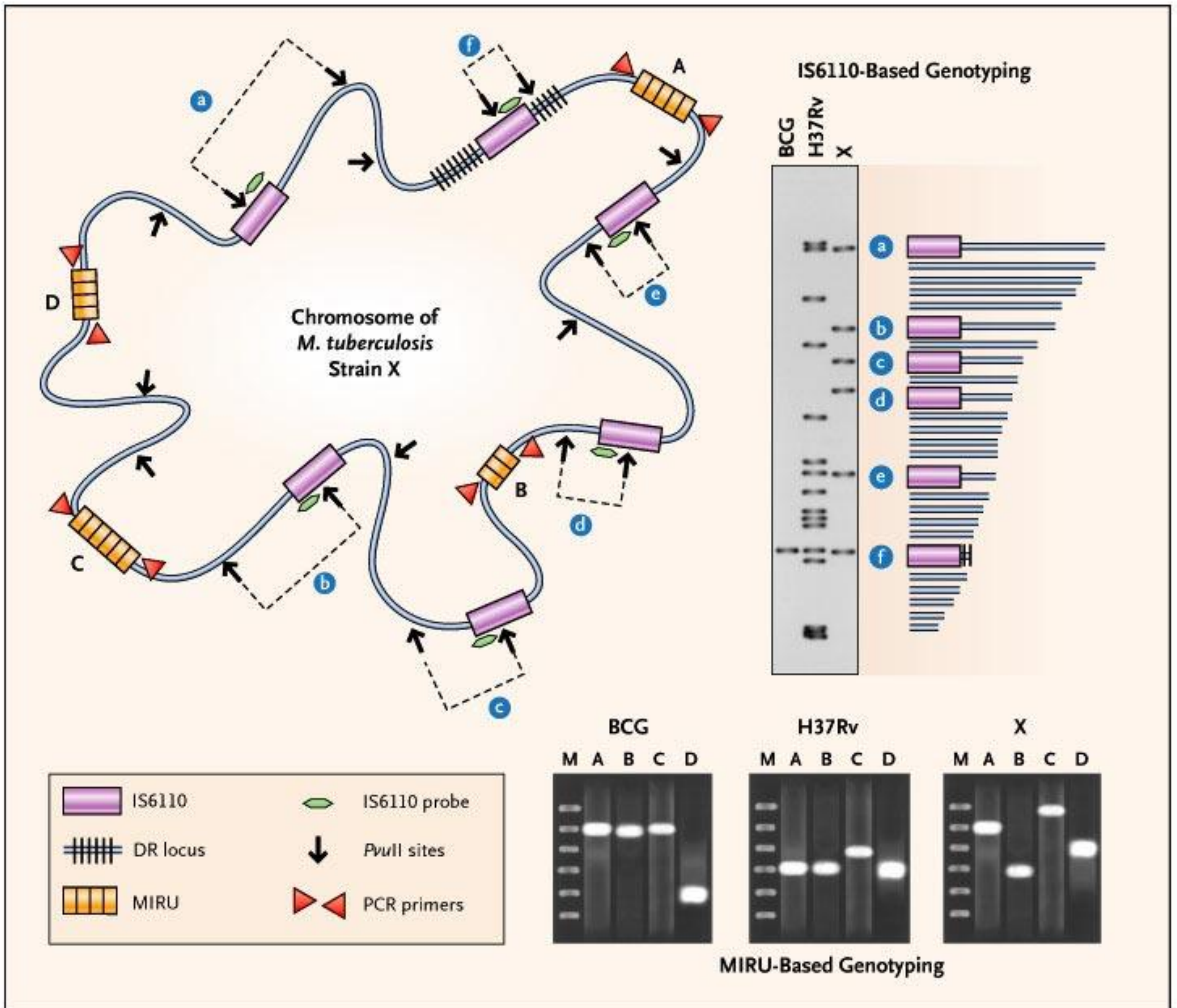


Figure 1.29 Representation of IS*6110*, MIRU and DR regions within the *M. tuberculosis* genome (Barnes & Cave 2003b)

1.9.4.5 Genome Sequencing and Genotyping

The level of discrimination of genotyping technologies varies. Ultimately, the way to definitively distinguish individual strains is by whole genome analysis. The genotyping methodologies described so far are a 'snap-shot' of the tubercle genome and act as a surrogate marker of their evolution.

The advancement in sequencing technology means that entire genomes may be analysed rapidly and at a fraction of the previous cost (Ansorge 2009). Whole genome analysis has been used to establish chains of transmission of *M. tuberculosis* when the genotyping by established methods had been unable to (Schurch et al. 2010)

1.9.5 The Utilisation of Genotype Data

As stated earlier, it is often beneficial, or even essential, to use more than one genotyping methodology. The various levels of discrimination and rate of change of the genetic elements mean that a combination of methods is often required to definitively differentiate strains. For example, the prevalence of strains possessing few copies of IS*6110* often prohibits its use, unaided. Likewise, the relatively low discriminatory power of spoligotyping means that it cannot be used alone to confirm related cases of TB.

The discriminatory power of a typing method is its ability to distinguish between unrelated strains. It is defined by the number of genotypes detected by the methodology and the relative frequencies of these genotypes.

Hunter and Gaston suggested the use of a single numerical index of discrimination (D), based on the probability that two unrelated strains sampled from the test population will be placed into different typing groups (Hunter & Gaston 1988). They calculated this probability by using Simpson's index of diversity, which was developed for the description of species diversity within an ecological habitat (Simpson 1949).

The aim is to assign a numerical value to a genotyping methodology using the following formula (equation 1.2):

$$D = 1 - \frac{1}{N(N-1)} \sum_{j=1}^{S} n_j(n_j - 1)$$

Equation 1.2 Calculation of the discriminatory power of a genotyping method using the Hunter Gaston index (Hunter & Gaston 1988)

where N is the total number of strains in the sample population, s is the total number of types described, and nj is the number of strains belonging to the jth type. This equation is derived as follows. The probability that a single strain sampled at random will belong to the jth group is njN. The probability that two strains sampled consecutively will belong to that group is nj(nj - 1)IN(N - 1) (Hunter & Gaston 1988).

The value calculated will be $\leq$1.0 - the higher the value, the more discriminatory the method. For example, take two hypothetical genotyping methods both differentiating 20 types in a population of 100 strains. In the first, assume a distribution of strains of 40 in type 1, 30 in type 2, 7 in types 3 and 4, and 1 in the remaining 16 types. In the second,

assume an ideal distribution of five strains in each type. Working through the first example, we have N = 100, s = 20, $n_1 = 40$, $n_2 = 30$, n_3 and $n_4 = 7$, n_5 to $n_{20} = 1$: D = 1- [(40 x 39 + 30 x 29 + 7 x 6 + 7 x 6 + 1 x 0 . . . 1 x 0)/(100 x 99)] = 1 - (2514/9900) = 0.746. Therefore, the first example has an index of 0.746. This index indicates that if two strains were sampled randomly from the population, then on 74.6% of occasions they would fall into different types. The second example is clearly more discriminating and has an index of 0.960 (Hunter & Gaston 1988).

It is worthwhile reiterating that it is not always essential to opt for the most discriminatory method. The benefits and limitations of an assay must be taken into account when assessing the question that needs to be answered. For example, isolates displaying divergent spoligotypes in a suspected outbreak are sufficient to dispute chains of transmission. Indistinguishable strains would need secondary typing to confirm their common source.

It has been shown that epidemiologically linked strains cluster together by IS*6110* RFLP, MIRU-VNTR and spoligotyping (Maguire et al. 2002; van Soolingen et al. 1999). Isolates that are indistinguishable by the typing method(s) used, that is are 100% similar, may therefore be due to recent transmission between individuals. The discriminatory power of the method used (alongside the epidemiological data) must, of course, be considered.

The number of clustered strains in a population may used to calculate a percentage of cases due to active transmission, rather than progression to disease following infection in the past. The percentage may be calculated as follows (equation 1.3; Small et al. 1994):

$$\text{Rate of transmission} = \frac{\text{No. of clustered strains} - \text{No. of clusters}}{\text{Total number of strains}} \times 100$$

Equation 1.3. Calculation for the rate of transmission of *M. tuberculosis* ((Small et al. 1994).

Indeed, typing is an important means of ruling out sources of recent transmission. It has been shown that despite large increases in the incidence of TB London between 1995 and 1997, only 14.4% of cases were due to active transmission. Reactivation of disease in an immigrant population was the main driving force behind the increase (Maguire et al. 2002). Rates of transmission vary greatly between study populations. The geographic area and length of study will of course play a critical role in the calculated percentage if transmission. A higher percentage of related strains are likely to be seen in a larger geographic area. If a study covers only a small region, then some TB patients in an active chain of transmission may seek medical attention elsewhere and are therefore excluded from the sample. This will therefore falsely decrease the level of transmission. Likewise, as the incubation period of TB is significantly longer than for other infectious diseases, a short study may not identify all cases in a chain of transmission or outbreak. The rates of transmissions from a number of studies are listed in table 1.1.

Country/City	**Date/Duration**	**Sample Size**	**Typing Method**	**Transmission %**	**Reference**
Norway	1994-2001	551	IS*6110* RFLP	11%	(Dahle et al. 2003)
London	1995-1997	2042	IS*6110* RFLP	14.4%	(Maguire et al. 2002)
Paris	1995	181	IS*6110* RFLP	14.9%	(Gutierrez et al. 1998)
New York	1989-1992	104	IS*6110* RFLP	26.0%	(Alland et al. 1994)
Seville	1993-1995	176	IS*6110* RFLP	27.3%	(Safi et al. 1997)
San Francisco	1991-1992	473		31.1%	(Small et al. 1994)
Netherlands	1993-1997	4266	IS*6110* RFLP	35%	(van Soolingen et al. 1999)
Amsterdam	1992-1994	459	IS*6110* RFLP	35.1%	(van Deutekom H. et al. 1997)
Dublin	1998-2002	142	IS*6110* RFLP & Spoligotyping	50.0%	(Fair et al. 2006)
Harare	1997	224	Spoligotyping & 5 loci VNTR	72.2%	(Easterbrook et al. 2004)

Table 1.1. Comparison of the percentage of TB cases as a result of active transmission.

As already stated, genotyping may be used to study the evolutionary relationships between strains. If 100% similarity is an indicator of recent transmission then isolates that are related at lower levels of similarity must therefore suggest longer evolutionary relationships. For example, spoligotyping was instrumental in identifying the Beijing family. This family shares a common spoligotype, but is less closely related by other typing methods (van Soolingen et al. 1995). The analysis of 90 strains from 38 countries by Kremer and colleagues identified several families by multiple typing methods (Kremer et al. 1999). The presence of the Beijing/W family was shown alongside two other novel families. The first comprised 13 strains originating from eight countries in

Asia, Europe and the Americas. Organisms showed at least 54% similarity by IS*6110* RFLP, but were indistinguishable by 5-loci VNTR and spoligotyping. This group was designated as the Haarlem family. The second group of eight strains displayed shared polymorphisms and originated from central Africa. This group was termed the Africa family (Kremer et al. 1999).

Taking into account the different levels of discrimination and the proposal to use more than one genotyping methodology, McHugh and colleagues (McHugh et al. 2005) defined a lineage as a group of strains that are $\geq$ 70% similar by two typing methods. It is reasoned that if such a lineage is genuinely related evolutionarily, then it should remain robust when genotyped by an alternative methodology.

The retrospective analysis of 2490 strains from the ethnically diverse population of London was performed to investigate their origins. Analysis of the IS*6110* genotypes differentiated the strains into 12 groups, or ‘superfamilies’ (sfams) related by at least 50% similarity (Dale et al. 2005b). Three sfams contained low copy number strains and were additionally investigated using spoligotyping and PGRS (Dale et al. 2003b). The superfamilies consisting of high copy number strains were designated sfam 1-9, whilst the three low copy number strains were named L1-3.

There were associations between some sfams and the site of disease, but stronger associations were where some sfams correlated with the country of birth of the TB patients. There was concordance between the findings of this study and previous work.

For example, the Haarlem family described by Kremer and colleagues mapped to sfam 1, whilst the Africa and Beijing families mapped to sfam 5 (Kremer et al. 1999). Additionally, the low copy superfamilies L1-3 were broadly associated with the East African Indian family, Africa and Europe, respectively (Dale et al. 2003b). It may be suggested that strains that display a similarity of at least 50% share a common evolutionary path.

1.10 Aims of Thesis

The aim of this thesis was to study the molecular epidemiology of *M. tuberculosis* isolates in TB patients at the Royal Free Hospital between 2002 and 2007. The hypothesis presented here is that the characterisation of the genotype of *M. tuberculosis* could be used to give an insight into the biology of the organism that relates to its clinical phenotype. The hypothesis was addressed by performing *M. tuberculosis* genotyping alongside the collection of demographic data for all culture positive TB patients at the Royal Free Hospital Hampstead NHS Trust, London. These isolates are studied in chapter 3 with the use of IS*6110* RFLP and MIRU-VNTR. The isolates and patient demographic characteristics are compared with national data to assess if any associations can be made genotype and phenotype.

An interesting cohort of patients who underwent a novel diagnostic interferon gamma release assay is investigated in chapter 4 to ascertain the relationship between immune response to *M. tuberculosis* and bacterial biology. The hypothesis is that the immunological response seen in the T cells of this group of TB patients is driven by the

infecting organism. The strains isolated from these patients will be investigated using genotyping, sequencing and gene expression studies to ascertain if this is the case.

Isolates obtained from two separate outbreaks are investigated in chapter 5. Both outbreaks involve drug resistant strains so the effect of this on their relative fitness is analysed. The hypothesis is that these drug-resistant strains have suffered a negligible fitness cost as they have been able to transmit between individuals readily. This will be assessed by the measurement of generation times of the isolates in vitro, together with molecular analysis of the resistance genes, themselves.

Finally, examples when genotyping has played an important clinical role are outlined in chapter 6. This chapter highlights the need for the integration of clinical, epidemiological and laboratory data to maximise the understanding of this complex global disease.

Chapter 2. Materials and Methods

2.1 *M. tuberculosis* IS*6110* RFLP database

2.1.1 Organism Selection and Culture Methods

Samples referred to the Department of Medical Microbiology, Royal Free Hampstead NHS Trust, London, for investigation of mycobacterial disease between 2002 and 2007 were processed. Samples that may be potentially contaminated with commensal organisms were decontaminated with 4% sodium hydroxide for 20 minutes, followed by buffering to neutral pH. All samples were concentrated by centrifugation prior to smear preparation. Slides were heat-fixed, stained with auramine-O and examined using a fluorescent microscope. Positive smears were over-stained with carbol fuchsin using a hot Ziehl-Neelsen stain.

Concentrated samples were inoculated into an automated liquid culture system; the MB BacT system (BioMerieux, France). Positive cultures were concentrated by centrifugation and confirmed by hot Ziehl-Neelsen stain. Confirmation of mycobacterial species was confirmed using Strand Displacement Amplification (SDA) on a ProbeTec (Becton Dickinson, USA) according to manufacturer's instructions. Positive cultures were additionally sub-cultured onto Lowenstein Jensen slopes.

Positive cultures were referred to the Health Protection Agency (HPA) Mycobacterium Reference Unit (MRU) (Whitechapel, London) for species confirmation and phenotypic susceptibility testing by either the resistance ratio or modified liquid proportion methods.

2.1.2 Data collection

Microbiological data was collated from the local information system (LIS), WinPath (Sysmed, UK). Patient demographic data was collated from WinPath and additionally from the London TB register (LTBR). The site of disease was surmised from the anatomical site(s) of the specimens collected. The number of isolates studied are outlined in figure 2.1.

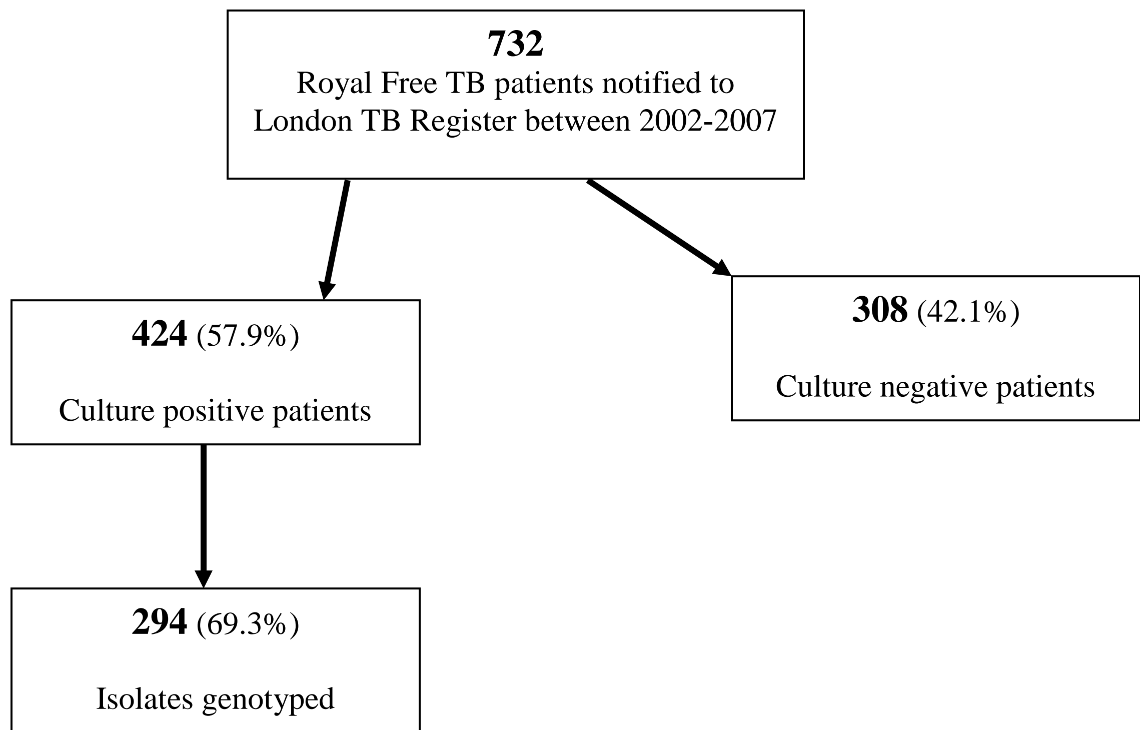


Figure 2.1 Flow diagram of isolates genotyped and included in subsequent analysis.

2.1.3 Preparation of DNA probe by PCR

The IS*6110*- specific INS probe is a PCR product of 245bp in length. The amplicon was generated with the following primer sequences, shown in table 2.1:

Target	Primer	Sequence	Reference
IS*6110*	INS-1 (forward)	5'CGT GAG GGC ATC GAG GTG GC 3'	(van Embden et al. 1993)
	INS-2 (reverse)	5'GCG TAG GCG TCG GTG ACA AA 3'	(van Embden et al. 1993)

Table 2.1 primers for the amplification of the INS probe for IS6110 RFLP hybridisation.

The PCR master mix was prepared with final reagent concentrations as follows; 1x KCl buffer, which contains $MgCl_2$ and gave a final $MgCl_2$ concentration of 1.5mM, 0.15mM dNTPs, 0.5µM of each primer and 0.025 units of Taq/µL (Bioline, London, UK). Genomic DNA from the *M. tuberculosis* control strain H37Rv was added. Thermocycling conditions of 94°C for 2 minutes, followed by 35 cycles of 94°C for 1 minute, 65°C for 1 minute and 72°C for 2 minutes. The final elongation step was 72°C for 5 minutes. The amplicon was detected using agarose gel electrophoresis with ethidium bromide staining. The PCR product was purified using the Qiagen QIAquick PCR Purification Kit Protocol following the manufacturer's instructions (Qiagen, Crawley, UK). The eluted solution containing the purified PCR product was stored at –20°C.

2.1.4 Random labelling of DNA probe

The purified PCR products were labelled using the Gene Images random prime labelling protocol (GE Healthcare, Chalfont St Giles, Bucks, UK) following the manufacturer's instructions. The labelled probe was stored in the dark at -20˚C for up to 6 months.

2.1.5 Isolation of genomic DNA from *M. tuberculosis*

A single isolate from each patient was selected unless relapse or re-infection was suspected. As *M. tuberculosis* is a hazard group 3 pathogen, all stages of the DNA extraction, up to and including the addition of chloroform, were performed in a class 1 biological safety cabinet in a containment level 3 laboratory.

A loopful of colonies from an LJ slope was added to 400µL 1x TE buffer and incubated at 80°C for 20 minutes. The tubes were cooled to room temperature and 50µL 10mg/mL lysozyme (Sigma, UK) was added. The tubes were incubated at 37°C overnight. The following morning, 70µL 10% SDS and 5µL 10mg/ml proteinase K (Sigma, UK) were added before incubation at 65°C for 10 minutes. Then, 100µL of 5M NaCl 100µL and CTAB/NaCl (pre-warmed to 65°C) were added and incubated at 65°C for 10 minutes. 750µL chloroform/isoamylalcohol (24:1) was added and the tubes were mixed by inversion. The tubes were centrifuged at 10,000g for 5 minutes and the upper aqueous phase containing the DNA was removed and added to 450µL of ice-cold isopropanol. These tubes were mixed gently before chilling at –20°C for at least 30 minutes. The tubes were centrifuged at 10,000g for 15 minutes at room temperature, the supernatant was removed and the DNA pellet was washed with 1mL ice-cold 70% ethanol. The tubes were centrifuged at 10,000g for 5 minutes at room temperature, the supernatant was removed and the pellet was allowed to air-dry. The pellet was rehydrated using TE buffer overnight at 4°C. The concentration of DNA obtained was estimated by comparing with serial dilutions of lambda DNA using agarose gel electrophoresis. DNA was also extracted from the *M. tuberculosis* reference strain MT14323, which was kindly provided

by Prof. Peter Hawkey and Jason Evans, West Midlands Public Health Laboratory, Health Protection Agency, Birmingham Heartlands Hospital.

2.1.6 Restriction enzyme digestion of chromosomal DNA

200ng of DNA was restricted at 37°C for 1 hour with 2.5μL 10x reaction buffer, 1.25μL *Pvu*II and 0.25μL BSA in a final volume of 25μL. To achieve a final image with standardised intensity, all samples were electrophoresed on an agarose gel and restricted DNA concentrations were compared visually to ensure that similar quantities are of DNA were loaded to the large agarose gel (step 2.1.7).

2.1.7 Separation of DNA fragments by electrophoresis and Southern blotting

The restricted DNA was applied to a 25 x 20cm 0.8% agarose gel using λ *Hind*III as a DNA marker. The restricted DNA of MT14323 was loaded into three wells of the gel (lane two, one off-centre lane in the middle and the final well) to allow the final image to be orientated and compared with others. An electric current was applied until the 2.2kb fragment of λ *Hind*III had migrated 11cm. The DNA in the gel was then depurinated in 200mL 0.2M HCl for 10 minutes with gentle agitation and the gel was briefly rinsed three times in distilled water. The DNA in the gel was denatured in 1.5M NaCl, 0.5M NaOH (200mL) for 40 minutes with gentle agitation and neutralised in 1M Tris (pH7.4), 1.5M NaCl (200mL) for 40 minutes with gentle agitation. The gel was rinsed briefly in 10x SSC before transferring the DNA from the gel to a nylon membrane – Hybond N+ (GE

Healthcare, Bucks, UK) by Southern blotting for 2 hours. The membrane was air-dried and the DNA cross-linked to the membrane by exposure to UV light for 3 minutes.

2.1.8 Hybridisation and detection.

The membrane was pre-wetted with 5x SSC before adding 80mL of hybridisation solution (5x SSC, 0.1%SDS, 5% liquid block and 5% dextran sulphate). Prehybridisation took place for 30 minutes at 50°C shaking at approximately 80 rpm on an orbital shaker. The INS probe was denatured at 95°C for 5 minutes and quenched on ice. 15µL of the denatured probe was added to 1mL of hybridisation solution, which was removed from the membrane, before returning this mixture to the bioassay tray (containing the hybridization solution and membrane), taking care to distribute the mixture evenly. The bioassay tray was sealed with Parafilm (Alpha Labs, Eastleigh, Hampshire, UK) to prevent the membrane drying out and hybridisation took place overnight at 50°C shaking at approximately 80rpm. Stringency washes were performed the following morning with pre-warmed solutions as follows: 300ml 1x SSC, 0.1% SDS for 7.5 minutes twice, followed by 300ml 0.5x SSC, 0.1% SDS for 7.5 minutes twice at 50°C with gentle agitation. The membrane was placed in a fresh bioassay tray and was washed briefly with Buffer A (100mM Tris base and 300mM NaCl), which was autoclaved prior to use. The membrane was blocked with 300mL of 1:10 Liquid Block in Buffer A for 90 minutes at room temperature with gentle agitation.

The membrane was incubated with 130mL of Buffer B (1:5000 anti-fluorescein alkaline phosphate-conjugate in 0.5% BSA in Buffer A) for 1 hour at room temperature with gentle

agitation. The membrane was washed with 300mL Buffer C (Buffer A, 0.3% Tween 20) for 10 minutes at room temperature with gentle agitation three times to remove any non-bound conjugate. The membrane was drained and 16mL of detection solution (GE Healthcare, Bucks, UK) was applied to the membrane 5 minutes. A piece of Hyperfilm (GE Healthcare, Bucks, UK) was aligned with the membrane in dark room conditions, developed and fixed to produce a molecular fingerprint.

2.1.9 Entry of images into BioNumerics

Images were scanned as greyscale TIFF files and uploaded into BioNumerics version 3.5 (Applied Maths, Sint-Martens-Latem, Belgium). The reference strain MT14323 lanes were assigned as reference markers by the software in order to normalise against other images. Each lane was carefully examined visually by the researcher to assign bands indicating an insertion sequence in the genome.

Once each isolate had been inserted into the database, patient demographics were added. These included patient identifiers, such as the isolate number and associated laboratory number, hospital number, date of birth, site of disease and drug susceptibility. If duplicate isolates from the same patients were identified and inspected to confirm similarity and removed from the final analysis. If multiple isolates from a given patient generated different genotypes, then the possibility of re-infection with a different strain, mutation of the incumbent strain or laboratory error were investigated.

The relatedness of the isolates in the database was assessed using the Dice coefficient and a band tolerance of 1.2% to generate a dendrogram. Isolates related at 100% similarity were termed clusters and were assigned an arbitrary number from the top of the dendrogram. Likewise, isolates related at ≥70% similarity were termed lineages and were assigned an arbitrary number from the top of the dendrogram.

2.2 – Mycobacterial Interspersed Repetitive Units – Variable Number Tandem Repeats (MIRU-VNTR)

2.2.1 Isolate selection

Isolates from the IS*6110* database were additionally typed using 24-MIRU if they were a member of a 100% cluster or a 70% lineage. Furthermore, all strains that possessed fewer than five copies of IS*6110* were MIRU typed.

2.2.2 Isolation of genomic DNA from *M. tuberculosis*

Genomic DNA was extracted as in 2.1.5. Occasionally there were insufficient volumes of DNA remaining following the processing of the IS*6110* RFLP. In this case, 20µL of PCR grade water was added to the empty tube and mixed gently.

2.2.3 MIRU Polymerase Chain Reaction

The 24 MIRU targets were amplified in eight triplex polymerase chain reactions using fluorescently labeled primers as described by Supply (Supply et al.). Reactions were set up in 96-well microtitre plates (Applied Biosystems, Warrington, UK) using a

QiaSymphony liquid handler (Qiagen, Crawley, Surrey, UK). The PCR primer sets and master mixes are listed below in tables 2.2 – 2.4.

Multiplex	Locus	Alias	Repeat unit length, bp[a]	PCR primer pairs (5' to 3', with labeling indicated[b])
Mix 1	580	MIRU 4	77	GCGCGAGAGCCCGAACTGC (FAM) GCGCAGCAGAAACGCCAGC
	2996	MIRU 26	51	TAGGTCTACCGTCGAAATCTGTGAC CATAGGCGACCAGGCGAATAG (VIC)
	802	MIRU 40	54	GGGTTGCTGGATGACAACGTGT (NED) GGGTGATCTCGGCGAAATCAGATA
Mix 2	960	MIRU 10	53	GTTCTTGACCAACTGCAGTCGTCC GCCACCTTGGTGATCAGCTACCT (FAM)
	1644	MIRU 16	53	TCGGTGATCGGGTCCAGTCCAAGTA CCCGTCGTGCAGCCCTGGTAC (VIC)
	3192	MIRU 31	53	ACTGATTGGCTTCATACGGCTTTA GTGCCGACGTGGTCTTGAT (NED)
Mix 3	424	42	51	CTTGGCCGGCATCAAGCGCATTATT GGCAGCAGAGCCCGGGATTCTTC (FAM)
	577	43	58	CGAGAGTGGCAGTGGCGGTTATCT (VIC) AATGACTTGAACGCGCAAATTGTGA
	2165	ETR A	75	AAATCGGTCCCATCACCTTCTTAT (NED) CGAAGCCTGGGGTGCCCGCGATTT
Mix 4	2401	47	58	CTTGAAGCCCCGGTCTCATCTGT (FAM) ACTTGAACCCCCACGCCCATTAGTA
	3690	52	58	CGGTGGAGGCGATGAACGTCTTC (VIC) TAGAGCGGCACGGGGGAAAGCTTAG
	4156	53	59	TGACCACGGATTGCTCTAGT GCCGGCGTCCATGTT (NED)
Mix 5	2163b	QUB-11b	69	CGTAAGGGGGATGCGGGAAATAGG CGAAGTGAATGGTGGCAT (FAM)
	1955		57	AGATCCCAGTTGTCGTCGTC (VIC) CAACATCGCCTGGTTCTGTA
	4052	QUB-26	111	AACGCTCAGCTGTCGGAT (NED) CGGCCGTGCCGGCCAGGTCCTTCCCGAT
Mix 6	154	MIRU 2	53	TGGACTTGCAGCAATGGACCAACT TACTCGGACGCCGGCTCAAAAT (FAM)
	2531	MIRU 23	53	CTGTCGATGGCCGCAACAAAACG (VIC) AGCTCAACGGGTTCGCCCTTTTGTC
	4348	MIRU 39	53	CGCATCGACAAACTGGAGCCAAAC CGGAAACGTCTACGCCCCACACAT (NED)
Mix 7	2059	MIRU 20	77	TCGGAGAGATGCCCTTCGAGTTAG (FAM) GGAGACCGCGACCAGGTACTTGTA
	2687	MIRU 24	54	CGACCAAGATGTGCAGGAATACAT GGGCGAGTTGAGCTCACAGAA (VIC)
	3007	MIRU 27	53	TCGAAAGCCTCTGCGTGCCAGTAA GCGATGTGAGCGTGCCACTCAA (NED)
Mix 8	2347	46	57	GCCAGCCGCCGTGCATAAACCT (FAM) AGCCACCCGGTGTGCCTTGTATGAC
	2461	48	57	ATGGCCACCCGATACCGCTTCAGT (VIC) CGACGGGCCATCTTGGATCAGCTAC
	3171	49	54	GGTGCGCACCTGCTCCAGATAA (NED) GGCTCTCATTGCTGGAGGGTTGTAC

Table 2.2 MIRU primer sets (Supply 2005)

mix	1	2	3	4	5
Loci	4-26-40	10-16-31	0424- 0577-2165	2401-3690-4156	2163b- 1955-4052
$MgCl_2$ final concentration	3mM	2mM	1,5 mM	3mM	1,5 mM
H2O	7,5	8,3	8,7	7,5	8,7
Buffer 10 X	2	2	2	2	2
Q Solution 5x	4	4	4	4	4
$MgCl_2$ 25 mM	1,2	0,4	0	1,2	0
DNTP 5mM	0,8	0,8	0,8	0,8	0,8
Primers EACH[a]	0,4	0,4	0,4	0,4	0,4
Hotstart DNA pol	0,08	0,08	0,08	0,08	0,08
Total premix	18	18	18	18	18

Table 2.3 MIRU PCR master mixes (1-5). [a] Six in total, *i.e.* one forward and one reverse primer for each of the 3 pairs. Initial concentration for all unlabeled primers: 20pmol/μL. Initial concentration for the labeled oligonucleotides: 2pmol/μL for locus 0577, 3690 and 1955, 8pmol/μL for locus 4052, 20pmol/μL for locus 4156 and 4pmol/μL for the other loci (Supply 2005).

mix	6	7	8
Loci	2-23-39	20-24-27	2347-2461-3171
$MgCl_2$ final concentration	2.5 mM	1,5 mM	2mM
H2O	7,9	8,7	8,3
Buffer 10 X	2	2	2
Q Solution 5x	4	4	4
$MgCl_2$ 25 mM	0.8	0	0,4
DNTP 5mM	0,8	0,8	0,8
Primers **EACH**[a]	0,4	0,4	0,4
Hotstart DNA pol	0,08	0,08	0,08
Total premix	18	18	18

Table 2.4 MIRU PCR master mixes (6-8). [a] Six in total, *i.e.* one forward and one reverse primer for each of the 3 pairs. Initial concentration for all unlabeled primers: 20pmol/μL. Initial concentration for all labeled oligonucleotides: 4pmol/μL (Supply 2005).

2.2.4 Fragment analysis of the MIRU PCR products

0.2µL of ROX labeled 1000bp DNA marker (MapMarker 1000, 8fmol/band/µL, Bioventures, Tennessee, USA) was added to each well in the 96-well microtitre plate.

Fragment Analysis software (Applied Biosystems, Warrington, UK) was used to assign amplicon sizes and therefore MIRU copy number for each well. The fluorescent signals generated in each well were examined by eye to verify that the software had called the fragments correctly and to identify any stutter-peaks generated by strand slippage during primer annealing or large fluorescent flares due to high amplicon concentrations. PCR products that generated large flares could be diluted and re-analysed.

2.2.5 Entry of MIRU Data into BioNumerics.

The 24-digit MIRU type was added for each isolate in the BioNumerics database. Relatedness was assessed using MIRU data alone to assess whether the IS*6110* clusters and lineages were corroborated. The relatedness of the IS*6110* low copy number strains was then assessed using MIRU to differentiate any clusters and lineages.

2.3 ESAT-6 and the Host Response

2.3.1 Patient selection

Isolates cultured from a cohort of TB patients described elsewhere (Breen et al 2008a) were selected for analysis. The relatedness of the isolates by IS*6110* and MIRU-VNTR was assessed as in sections 2.1 and 2.2, respectively.

2.3.2 *esx*A PCR and Sequencing

An initial PCR was designed, optimised and performed by the author. It amplified a 228bp product was designed to identify the presence or absence of the gene. The conditions were as follows (table 2.5):

Target	Primer	Sequence	Reference
*esx*A	Forward	3’ CCAGGGAAATGTCACGTCCATTCA 5’	This experiment
	Reverse	3’ AACATCCCAGTGACGTTGCCTT 5’	This experiment

Table 2.5. Primers for the amplification of the *esx*A gene

The PCR master mix was optimised to final reagent concentrations as follows; 1x PCR buffer, 2mM $MgCl_2$, 0.2mM dNTPs, 0.4µM of each primer and 0.5 units of Taq/µL (Invitrogen, Paisley, UK). 1µL of DNA of at least 0.5ng/mL was added. Thermocycling conditions of 95°C for 5 minutes, followed by 30 cycles of 95°C for 30 seconds, 50°C for 30 seconds and 72°C for 30 seconds. The final elongation step was 72°C for 5 minutes.

DNA amplification was detected by agarose gel electrophoresis with ethidium bromide staining.

The *esxA* gene was sequenced from PCR positive isolates, using a second PCR protocol (626bp product) to search for polymorphisms. The conditions were as follows:

Target	**Primer**	**Sequence**	**Reference**
*esx*A plus flanking region	Forward	3' CCAAGAAGCAGCCAATAAGC 5'	This experiment
	Reverse	3' GGAGCTTCCATACCTTCGTG 5'	This experiment

Table 2.6. Primers for the amplification of the *esx*A gene plus flanking region.

The PCR master mix was optimised to final reagent concentrations as follows; 1x PCR buffer, 2mM $MgCl_2$, 0.2mM dNTPs, 0.4µM of each primer and 0.5 units of Taq/µL (Invitrogen, Paisley, UK). 1µL of DNA of at least 0.5ng/mL was added. Thermocycling conditions of 95°C for 5 minutes, followed by 30 cycles of 95°C for 30 seconds, 50°C for 30 seconds and 72°C for 30 seconds. The final elongation step was 72°C for 5 minutes. The amplification of the gene was detected by agarose gel electrophoresis with ethidium bromide staining. Where a PCR product was detected, it was purified using the QIAGEN QIAquick PCR Purification Kit Protocol (Qiagen, Crawley, UK) following manufacturer's instructions.

A cycle sequencing reaction was set up using the purified PCR product. Four tubes were set up for each PCR product, consisting of two forward and two reverse reactions. A final concentration of 1.05µM of one primer (2µL), 2µL Big Dye RRM V3.1 (Applied

Biosystems, Warrington, UK), 10µL water and 5µL purified PCR product. Thermocycling conditions of 96°C for 1 minute, followed by 40 cycles of 96°C for 10 seconds, 50°C for 5 seconds and 72°C for 4 minutes.

Each 19µL cycle sequencing product was precipitated with the addition of 3µL 3mM NaCl pH4.6, 62.5µL 95% ethanol and 14.5µL water. The mixture was vortexed before being centrifuged at ≥15,000g for 15 minutes at 4°C. The supernatant was removed without disturbing the pellet. 200µL of 70% ethanol was added to each tube followed by inversion to wash the pellet. The tubes were centrifuged at ≥15,000g for 5 minutes at 4°C. The supernatant was removed and the pellet was allowed to dry completely. 15µL formamide was added to each tube before being loaded onto the automated plate sequencer (Applied Biosystems, Warrington, UK).

2.3.3 Sequence Analysis

The sequences obtained were analysed by the author using BioNumerics, version 3.5 (Applied Maths, Sint-Martens-Latem, Belgium). Consensus sequences were visually checked for correct base-calling by the researcher. These were then aligned with the *esx*A sequence of *M. tuberculosis* H37Rv using Clustal W (http://www.ebi.ac.uk/Tools/clustalw2/index.html).

2.3.4 *esx*A mRNA Reverse Transcriptase PCR

2.3.4.1 Evaluation of optimal *esx*A expression

Triplicate cultures of *M. tuberculosis* H37Rv (0.5 Macfarland initial turbidity) were incubated in Kirshner's medium (E&O Laboratories, Scotland) at 37°C for 7, 14 and 28 days, to represent lag, log and stationary phase of growth. RNA was extracted at these time-points by the method stated below. Reverse transcription, quantitative polymerase chain reaction (RT-qPCR) was performed on each RNA extract as stated below.

2.3.4.2 RNA Extraction

Cultures of the selected isolates were sub-cultured into Kirshner's broth (E&O Laboratories, Scotland) and incubated at 37ºC for 10 days. After this time, the growth was assessed by observing the turbidity of the broth culture. A turbidity of Macfarland 0.5 was necessary. Any broths that were more turbid than this were diluted with fresh Kirshner's broth (E&O Laboratories, Scotland) until Macfarland 1 was achieved. These broths were then incubated at 37 ºC for the required number of days to reach an optimal level of *esx*A transcription (see 2.3.3.1).

The RNA in each 10mL broth was stabilised by the addition of 35mL 5M guanine thiocyanate (GTC) and 245µL β-mercaptoethanol (0.7% final volume). This solution was centrifuged at 2000g for 30 minutes at room temperature. The supernatant was discarded and 1mL of RNApro (MP Biomedicals, Illkirch, France), was added to the deposit and was completely resuspended with gentle pipetting.

1mL of this suspension was added to a RNA extraction tube (MP Biomedicals, Illkirch, France) containing lysing matrix beads. These tubes were placed in a ribolyser (FastPrep 24, MP Biomedicals, Illkirch, France) on setting 6.0 for 40 seconds. The tubes were then centrifuged at ≥12,000g for 5 minutes at room temperature.

The supernatant (approximately 750µL) was transferred to a fresh microcentrifuge tube without disturbing the matrix interface. The transferred samples were allowed to stand at room temperature for 5 minutes to increase the RNA yield.

300µL of chloroform was added to each sample, vortexed for 10 seconds and allowed to stand at room temperature for 5 minutes to permit nucleoprotein disassociation and increase RNA purity. Tubes were then centrifuged at 12,000g for 5 minutes at room temperature and the upper phase was transferred to a fresh microcentrifuge tube without disturbing the interface. 500µL ice-cold absolute ethanol was added to the fresh tube and inverted 5 times. The samples were then placed at -80ºC for 15 minutes.

The samples were centrifuged at 12,000g for 20 minutes at 4ºC and the supernatant discarded. The nucleic acid pellet was washed by the addition of 500µL ice-cold 70% ethanol (prepared with diethylpyrocarbonate (DEPC) treated water to remove RNAase) and inversion. The sample was centrifuged at 12,000g for 15 minutes at 4ºC. The supernatant was removed completely and the pellet was allowed to air dry at 50ºC. The pellet was rehydrated with 100µL RNAase-free water and was used immediately or stored at -80ºC.

2.3.4.3 RNA clean-up

43µL of the extracted RNA was mixed with 1µL (1 unit) of Turbo DNase and 5µL 10x Turbo DNase buffer (Ambion, Turbo DNA-free kit, Applied Biosystems, Warrington, UK). This mixture was incubated at 37ºC for 30 minutes. A further 1µL of Turbo DNase was added and the mixture was incubated for an additional 30 minutes. 10µL of DNase Inactivation Reagent (Ambion, Turbo DNA-free kit, Applied Biosystems, Warrington, UK) was added and the mixture was incubated at room temperature for 5 minutes with occasional mixing. Finally, the mixture was centrifuged at 12,000g for 90 seconds and the supernatant removed to a fresh tube. This DNase-treated RNA could be used immediately or stored at -80ºC.

2.3.4.4 Reverse transcriptase qPCR

RT-qPCR was performed for the target gene *esx*A as well as the housekeeping, or reference genes 16S rDNA and *sig*A. Each reaction was performed in triplicate. The conditions were as shown in table 2.7.

The PCR master mix was optimised to final reagent concentrations as follows; 12.5µL Quantitect RT PCR mastermix (Qiagen, Crawley, UK), 0.4µM of each primer and 0.15µM of probe. 0.25µL reverse transcriptase (RT) mix was added to each tube. Water was added to make a final volume of 20µL. 5µL of RNA (diluted 1:10 from 2.3.3.3) was then added. Thermocycling conditions on a RotorGene 3000 (Qiagen, Crawly, UK) of 50°C for 30 minutes (RT step), 95°C for 15 minutes, followed by 40 cycles of 94°C for 15 seconds and 60°C for 60 seconds.

Target	Primer	Sequence	Reference
*esx*A	Forward	3’ CCAAGAAGCAGCCAATAAGC 5’	This experiment
	Reverse	3’ GGAGCTTCCATACCTTCGTG 5’	This experiment
	Probe	3’ CGCGGCAAGCGCAATCCAGGGAAATGT 5’ (FAM)	This experiment
16S	Forward	3’ GTGATCTGCCCTGCACTTC 5’	Amani Alnmir, personal communication
	Reverse	3’ ATCCCACACCGCTAAAGCG 5’	Amani Alnmir, personal communication
	Probe	3’ AGGACCACGGGATGCATGTCTTGT 5’ (ROX)	Amani Alnmir, personal communication
*sig*A	Forward	3’ AAACCATCTGCTGGAAGCCA 5’	(Hampshire et al. 2004)
	Reverse	3’ TTCTCGACCTGATCCAGGAAG 5’	O’Sullivan 2007
	Probe	3’ CCTGCGCCTGGTGGTTTCGC 5’ (JOE)	(Hampshire et al. 2004)

Table 2.7 Primers and probes for RT-qPCR

Each RT-qPCR was analysed as follows; slope correct was selected and the thresholds for *esx*A, 16S and *sig*A were optimised to 0.01, 0.05 and 0.1, respectively. The Ct of the reactions with and without RT were compared for gross DNA contamination. Once this was ruled out, the relative gene expression change was quantified using the ΔΔCT method (Livak and Schmittgen 2001). The formulae are as follows:

$$\text{Relative fold gene expression} = 2^{-\Delta\Delta Ct}$$

$$\text{Where: } \Delta\Delta Ct = \Delta Ct_{\text{selected strain}} - \Delta Ct_{\text{H37Rv}}$$

$$\text{and } \Delta Ct = (Ct_{\text{target gene}} - Ct_{\text{reference gene}}) \text{ for each selected strain}$$

Equation 2.1. Calculation of relative gene expression change using ΔΔCt (Livak and Schmittgen 2001).

2.4 The Fitness of Over-Represented strains and Acquired Isoniazid Resistance

2.4.1 Isoniazid Resistance Gene Sequencing

Genomic DNA was extracted as follows; organisms were inoculated into Kirshner's medium (E&O Laboratories, Scotland) and incubated until growth was visible. Cells were pelleted in a 2mL screw-cap tube by centrifugation and were washed in 1x TBE. The organisms were heat killed at ≥95˚C for 30 minutes. Glass beads (Becton Dickinson, Oxford, UK) were added and the tubes were placed in a Ribolyser (FastPrep 24, MP Biomedicals, Illkirch, France) on setting 6.0 for 45 seconds. The tube was then centrifuged at 12,000g for 5 minutes and the supernatant removed. DNA extraction was performed by Dr. Alastair McGregor as part of his MSc degree project under the supervision of the author.

Selected regions of the *inh*A and *kat*G genes of all strains were amplified using the following PCR conditions (table 2.8). PCR amplification was performed by Dr. Alastair McGregor as part of his MSc degree project under the supervision of the author.

The PCR master mix was previously optimised in this laboratory as follows; 1x KCl buffer (Bioline, London, UK), 0.04µM of each primer, 0.15mM dNTP and 2 units Taq polymerase (Bioline, London, UK). Water was added to make a final volume of 90µL. 10µL of extracted DNA was then added. Thermocycling conditions of 95°C for 5 minutes, followed by 30 cycles of 94°C for 1 minute, either 56°C (*kat*G) or 65°C (*inh*A) for 2 minutes and 72°C for 3 minutes. A final elongation step of 72°C for 7 minutes was used.

Target	**Primer**	**Sequence**	**Reference**
*kat*G (185bp fragment targeting codon 315)	Forward	5' GTC ACA CTT TCG GTA AGA CC 3'	(O'Sullivan 2007)
	Reverse	5' TTG TCC CAT TTC GTC GGG 3'	(O'Sullivan 2007)
*kat*G (296bp fragment targeting codon 463)	Forward	5' GCG AAG CCG AGA TTG CCA GC 3'	(O'Sullivan 2007)
	Reverse	5' ACA GCC ACC GAG CAC GAC GA 3'	(O'Sullivan 2007)
*inh*A (248bp fragment of the promoter region)	Forward	5' CCTCGCTGCCCAGAAAGGGA 3'	(Telenti et al. 1997)
	Reverse	5' ATCCCCCGGTTTCCTCCGGT 3'	(Telenti et al. 1997)

Table 2.8 Primers for amplifying isoniazid-conferring mutations.

Where a PCR product was detected by agarose gel electrophoresis with ethidium bromide staining, it was purified using the QIAGEN QIAquick PCR Purification Kit Protocol (Qiagen, Crawley, UK) following manufacturer's instructions.

Cycle sequencing reactions were set up using the purified PCR product. Four tubes were set up for each PCR product, consisting of two forward and two reverse reactions for each gene. A final concentration of 1.05µM of one primer (2µL), 2µL Big Dye RRM V3.1 (Applied Biosystems, Warrington, UK), 10µL water and 5µL purified PCR product. Thermocycling conditions of 96°C for 1 minute, followed by 40 cycles of 96°C for 10 seconds, 50°C for 5 seconds and 72°C for 4 minutes.

Each 19µL cycle sequencing product was precipitated with the addition of 3µL 3mM NaCl pH4.6, 62.5µL 95% ethanol and 14.5µL water. The mixture was vortexed before being centrifuged at ≥15,000g for 15 minutes at 4°C. The supernatant was removed without disturbing the pellet. 200µL of 70% ethanol was added to each tube followed by inversion to wash the pellet. The tubes were centrifuged at ≥15,000g for 5 minutes at 4°C. The supernatant was removed and the pellet was allowed to dry completely. 15µL formamide was added to each tube before being loaded onto the automated plate sequencer (Applied Biosystems, Warrington, UK). Sequencing was performed by Dr. Alastair McGregor as part of his MSc degree project under the supervision of the author.

2.4.2 Sequence Analysis

The sequences obtained were imported into BioNumerics, version 3.5 (Applied Maths, Sint-Martens-Latem, Belgium) and were analysed by the author. Consensus sequences were visually checked for correct base-calling by the researcher**.** These were then aligned with the corresponding gene sequences of *M. tuberculosis* H37Rv using Clustal W (http://www.ebi.ac.uk/Tools/clustalw2/index.html).

2.4.3 Fitness Assays

Isolates were incubated in Kirshner's medium (E&O Laboratories, Scotland) and incubated until 0.5 Macfarland of growth was visible. 0.5mL of each broth was inoculated into a MGIT (mycobacterial growth indicator tube) tube (Becton Dickinson, Oxford, UK) containing growth supplement, which contains bovine albumin, catalase, oleic acid, dextrose and polyoxyethylene stearate The MGIT tubes were incubated in the MGIT 960 system. Within 24 hours of these cultures signaling positive, 1:10 (termed dilution A) and 1:10,000 (termed dilution B) dilutions were prepared. A 0.5mL aliquot of each dilution was inoculated in triplicate into fresh MGIT tubes containing growth supplement. These tubes were incubated in the MGIT 960. This was performed in triplicate for all isolates.

A modified Youmans and Youmans method (Youmans and Youmans 1949) was used to calculate the growth rate constant, k; k=(logA – logB / t, where A is the largest inoculum (dilution A), B is the smallest inoculum (dilution B) and t is equal to the difference of the three replicates of the time to positivity (TTP) between dilution A and B. To determine the generation time G, the equation G = log 2/k was used. This process was repeated three times for each strain.

2.5 Statistical analysis

2.5.1 Chi squared and Fisher's Exact Test

These two methods calculate a P value from a population with an assumed normal distribution and assesses whether there is an association between the variables analysed,

particularly in a 2 x 2 table. When calculating the P value, one starts with the assumption about the population (the null hypothesis) and determines the probability of randomly selecting the samples with as large a difference as is observed. The P value is a figure between 0 and 1. For example, a P value of 0.034 means that if this null hypothesis is true, 3.4% of all possible experiments of this size would results in a difference between mean values as large (or larger) than was observed (Motulsky 1995). A P value of ≤ 0.05 is generally considered to be statistically significant. The Fisher's test will always give a theoretically correct P value, while the Chi squared test gives an approximation. The factor that decides whether a Chi Squared or a Fisher's exact test is used is largely sample size. Small samples (if total number of samples less than 40 or if the value in any cell is less than 5) must be analysed using a Fisher's exact test. This method is used in Chapter 3 to assess the statistical significance of the difference in gender RFH and UCLH patients, smear positivity and HIV status, smear status and country of birth.

Samples larger than this are analysed using a Chi squared test. This method is used to assess the significance of smear positivity and ethnicity, ethnicity and HIV status, age and site of disease in RFH and UCLH patients and drug resistance with HIV status, ethnicity and country of birth.

2.5.2 One-way Anova with Kruskal-Wallis post test

The one-way analysis of variance using the Kruskal–Wallis post-test compares the medians among groups. The Kruskal–Wallis test does not assume a normal population

distribution. It does, however, assume an identically-shaped and scaled distribution for each group, except for any difference in medians (Motulsky 1995).

This method is used to compare the median generation times of the isolates studied for variation of fitness in chapter 5.

2.5.3 Two-way Anova

This method compares the whether the means or medians of several groups (more than two) are all equal. It is comparable to performing multiple t-tests. This has the advantage over performing a statistical analysis on the groups separately as this would result in an increased chance accumulated error (Motulsky 1995).

This method is used to compare the ethnicities of patients from this data set, national data (Health Protection Agency 2009) and patients in London between 1995 and 1997 (Maguire et al. 2002).

Chapter 3: Development and Characterisation of an *M. tuberculosis* database for North London

3.1 Introduction

Typing databases are powerful tools for epidemiological and evolutionary studies. Collation of all strains from a given geographical area allows the investigator to understand the prevalence of strains within that population and gives an indication of levels of transmission as well as genotypes that are over-represented, suggesting a biological or clinical advantage. It is possible to identify and study outbreaks and this may contribute to early interventions that prevent further transmission (Kik et al. 2008).

This study will attempt to identify any clusters or lineages of strains that are particularly associated with a disease or bacterial phenotype. This will be achieved by genotyping isolates from the TB service at the Royal Free Hospital between 2002 and 2007 and investigating these alongside demographic data.

The highly discriminatory IS*6110* RFLP was used as the primary typing method for the isolates outlined in chapter 2. The addition of a secondary typing technique may augment the discriminatory power of an initial method. The discriminatory power of a genotyping method is displayed by the number of indistinguishable clusters generated from a group of strains (Kremer et al. 1999). Alternatively, clustered isolates may require secondary genotyping to confidently determine their relatedness. For example, strains that clustered by IS*6110* RFLP genotyping were shown to be unrelated using the polymorphic GC-rich repetitive-sequence (PGRS) method (Gillespie, Dickens, & McHugh 2000) and the

poorly discriminatory spoligotyping cannot be used in isolation in an outbreak setting; Beijing/W strains share a common spoligotype despite possessing varying IS*6110* RFLP fingerprints (Ebrahimi-Rad et al. 2003)

In addition to the limitations regarding the discriminatory power of genotyping methodologies, IS*6110* RFLP genotyping is unable to differentiate isolates that possess fewer than five copies of IS*6110*. A study in London between 1995 and 1997 showed that although the overall discriminatory power of RFLP typing was high, almost 20% of strains in the UK capital fell into this group (Maguire et al. 2002). Secondary typing of these strains using PGRS and spoligotyping was required to fully investigate the rate of active transmission these strains (Dale et al. 2003).

The robust characterisation of evolutionarily related strains clearly requires an integrated approach and the use of two genotyping methodologies has been proposed as a minimum standard for defining related lineages (McHugh et al. 2005). McHugh and colleagues suggest that isolates of *M. tuberculosis* are genuinely related in a lineage only if they are related at greater than 70% similarity by two typing methods. Therefore, any isolates related at ≥70% similarity by IS*6110* were secondarily typed using MIRU-VNTR.

MIRU-VNTR was selected as secondary typing method in this data set for a number of reasons. In contrast to other methods, MIRU-VNTR generates an objectively interpretable numerical genotype, it is possible to automate the method thus facilitating high sample throughput, and finally, the use of differing numbers of loci allows varying

discriminatory power. When only five loci were used, Kremer and colleagues found MIRU-VNTR to be less discriminatory than IS*6110* RFLP, PGRS and spoligotyping (Kremer et al. 1999). Twelve-loci MIRU-VNTR has been shown to be both more (Kwara et al. 2003), and less (Supply et al. 2006), discriminatory than IS*6110* RFLP.

A set of 24 loci is now widely used, but a subset of the 15 most discriminatory MIRU loci generates 96% of the discriminatory power of the full set of 24 loci (Supply et al. 2006). It was reasoned that the 15 loci set may be used for epidemiological studies, whilst the 24 loci set would be of use in phylogenetic studies. The Health Protection Agency has adopted the set of 24 loci for prospective analysis in January 2010.

3.1.1 Investigating Outbreaks

An outbreak may be defined as incidence of disease clearly in excess of normal experience or expectancy (Bannister et al 2006). With regards to TB, due to the often prolonged incubation period the increase in case numbers may not be sudden as seen in outbreaks of other infectious diseases, such as measles (Heathcock and Watts 2008) or Legionnaire’s disease (Coetzee et al. 2009). Nevertheless, an unexpected increase in cases may warrant investigation.

If there is no obvious common link between cases, such as attendance at the same school (Caley et al. 2010) or public house (Gaber et al. 2005), or there is no marked increase in cases, then genotyping can be useful in detected unknown outbreaks (Bifani et al. 1999;

Lopez-Calleja et al. 2009.; van Soolingen D. et al. 1999). Indeed, epidemiological links may only be investigated following the availability of genotyping data (de Vries.et al. 2009).

3.1.2 Transmission rates

The number of clustered strains in a population can be used to calculate a percentage of cases due to active transmission, rather than progression to disease following infection in the past. The percentage is calculated as follows (equation 3.1; Small et al. 1994):

$$\text{Percentage of transmission} = \frac{\text{No. of clustered strains} - \text{No. of clusters}}{\text{Total number of strains}} \times 100$$

Equation 3.1.Calculation for the rate of transmission of *M. tuberculosis* (Small et al. 1994).

Subtracting the number of clusters from the number of clustered strains is important so as not to double-count a transmission event. For example, a cluster of five patients with indistinguishable strains represents four transmission events.

The calculation of transmission events can be useful in tackling outbreaks and directing public health interventions, assessing the effectiveness of TB control measures (de Vries et al. 2009) or allocation of resources for contact tracing. Establishing the rate of transmission can also have wider social implications. For example, the knowledge that a large proportion of individuals diagnosed in London with TB were born outside the UK may lead to unfounded and inflammatory headlines. In comparison, the data showing that

the rate of transmission in London between 1995 and 1997 was low (14.4%) (Maguire et al. 2002) can allay fear of the uncontrolled propagation of the disease.

3.1.3 Laboratory contamination

The lipid–rich nature of the cell wall of *M. tuberculosis* makes the organism hardy and enables it to persist in a number of environments. For example, viable bacilli have been recovered from heat fixed sputum smears and from 0.9% sodium chloride decontamination solutions up to 3 weeks after inoculation (Aber et al. 1980). False positive cultures can occur due to contamination at any stage from specimen collection, such as from contaminated bronchoscopes, clerical and laboratory based errors. Faulty bronchoscopes may be incompletely decontaminated between patients and result in the inoculation of organisms into other samples. Samples may be labelled with the incorrect patient details and may lead to the wrong diagnosis. (Burman and Reves 2000). Additionally, the automated liquid culture systems used today enhance the possibility of laboratory cross contamination due to their increased sensitivity.

Currently at the Royal Free Hospital, laboratory contamination is suspected if a patient sample becomes culture positive for *M. tuberculosis* but does not fit with the clinical diagnosis of TB. On suspicion of a contamination event, the date that the sample was processed is identified; samples processed during the preceding seven days that subsequently grew *M. tuberculosis* are identified and genotyped. In addition, any *M. tuberculosis* strains isolated in the preceding seven days are genotyped. If any of these isolates are indistinguishable from the suspected isolate, then laboratory contamination is

presumed (de Boer et al 2002; Ruddy et al. 2002). In good practice contamination rates of below 3% are more common (Burman and Reves 2000; Ruddy et al. 2002).

It is important to ensure that false-positive cultures are identified as quickly as possible, as the prolonged combination therapy has several potential side effects, including gastrointestinal symptoms, rashes and most significantly, hepatotoxicity (Walker et al. 2008; Saukkonen et al. 2006). Overall incidence rates of side effects of 1.48 per 100 person-months have been reported (Yee et al. 2003 ;Walker et al. 2008; Saukkonen et al. 2006). Additionally, unnecessary investigations, hospitalisation and contract tracing may be performed. Three patients wrongly diagnosed with TB as a result of laboratory contamination in Massachusetts, USA in the late 1990s found that the cost was on average $10,873 (range $1,033 – $21,306) (Northrup et al. 2002).

3.1.4 Investigating Over-represented Strains

Genotyping may be used to identify evolutionarily-linked or over-represented groups of strains, such as the Beijing/W family (van Soolingen et al. 1995), the Kilimanjaro and Meru lineages in northern Tanzania (McHugh et al. 2005) and the Haarlem family (Kremer et al. 2005). The over-representation of a group of strains may indicate their success or fitness. The strains may have a selective advantage over other strains within the same geographic locus. The Beijing/W family is an example of this (Parwati et al. 2010).

A database was established in 2002 to correlate genotype and demographic data of all isolates of *M. tuberculosis* cultured in the Department of Medical Microbiology at the Royal Free Hampstead NHS Trust, London. IS*6110* RFLP was used as the primary typing tool (Chapter 2: van Embden et al. 1993). The TB service at the Royal Free Hospital has a catchment area of approximately 350,000 people and includes the London boroughs of South Barnet, Islington, North Camden and a small portion of Westminster (figure 3.1).

Figure 3.1. Greater London boroughs (londoncouncils.gov.uk).

3.2 Results

Samples referred to the Department of Medical Microbiology, Royal Free Hospital NHS Hampstead Trust, for mycobacterial investigations between 2002 and 2007 were processed. Positive cultures from the MB BacT system (bioMerieux, France) were sub-cultured onto Lowenstein Jensen slopes. The database was rationalised by including only a single isolate from each culture confirmed case of tuberculosis diagnosed at the laboratory. All stages of the RFLP analysis were performed by the author.

Microbiological data was collated from the local information system (LIS), WinPath (Sysmed, UK). Patient demographic data was collated from WinPath and additionally from the London TB register (LTBR). These data were collected for routine diagnostic use, but utilised by the author retrospectively for the purposes of this study.

The site of disease was surmised from the anatomical site(s) of the specimens collected. If *M. tuberculosis* was cultured only from respiratory samples (for example, sputum or broncho-alveolar lavage) then the patient was classified as having pulmonary disease, whilst if only samples from other sites were positive then the patient was characterised as having an extrapulmonary infection. A patient could be placed into a third category by having culture positive samples from respiratory and non-respiratory sites.

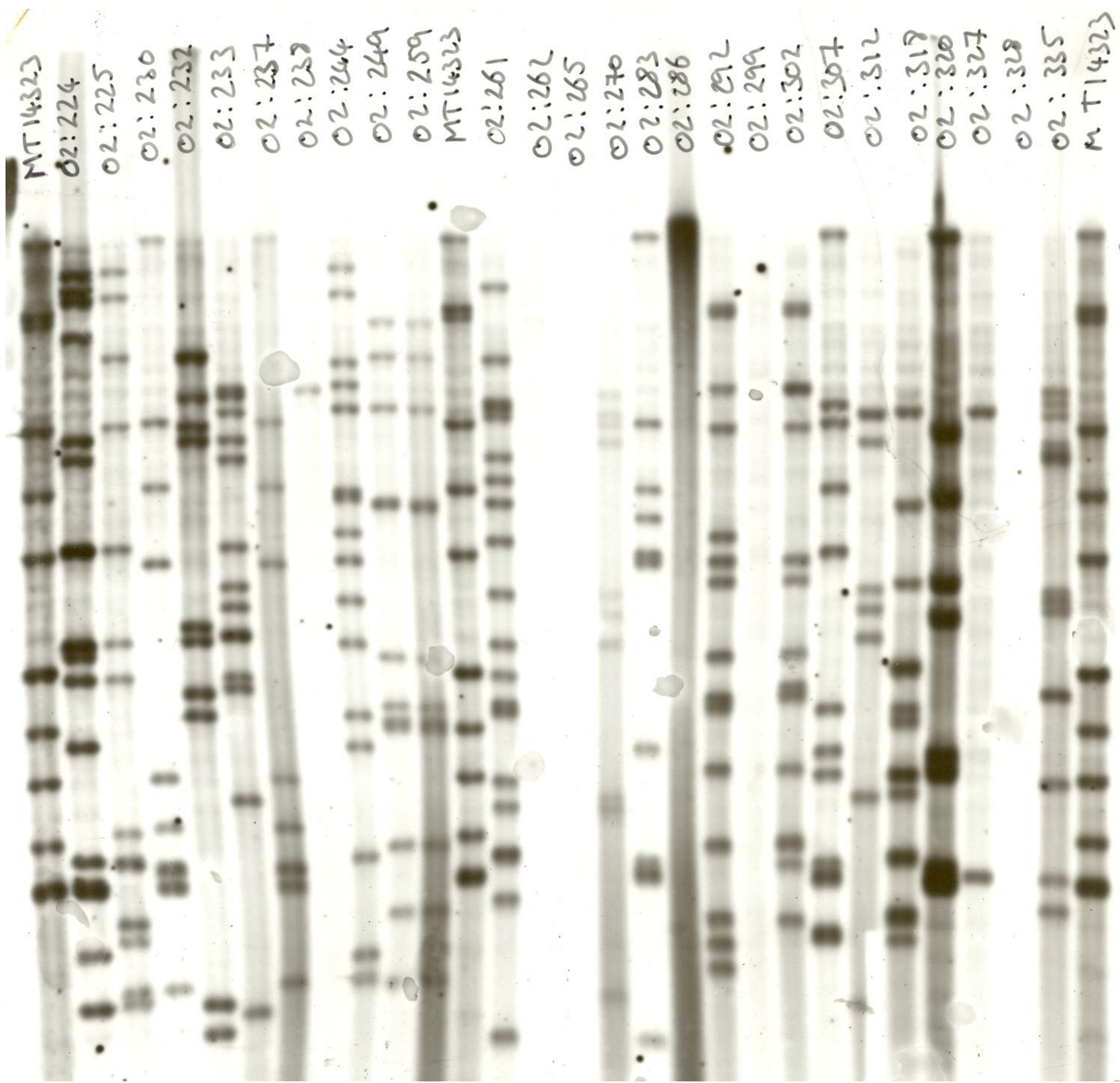
IS*6110* RFLP typing was already established in the laboratory to provide genotyping data for the diagnostic TB service. The images generated (figure 3.2) were analysed using BioNumerics version 3.5 (Applied Maths, Sint-Martens-Latem, Belgium).

The relatedness of the isolates in the database was assessed using the Dice coefficient as a measure of similarity of assigned bands and a band tolerance of 1.2% to generate a dendrogram, a small example of which is shown in figure 3.3. These settings are as described in the internationally agreed methodology (Van Embden et al 1993). Isolates related at 100% similarity were termed clusters and were assigned an arbitrary number from the top of the dendrogram. Likewise, isolates related at ≥70% similarity were termed lineages and were assigned an arbitrary number from the top of the dendrogram.

Mycobacterial interspersed repetitive units – variable number tandem repeats (MIRU-VNTR) was used as a secondary typing method in this study. This data was also analysed using BioNumerics and will be discussed in chapter 4.

3.2.1 Database coverage

At the Royal Free Hospital a total of 732 patients were notified to the local Health Protection Unit (HPU) over the six year period. Of these, 424 patients (57.9%) were culture positive for *M. tuberculosis*. An RFLP strain type (figure 3.2) was obtained on 294 isolates (69.3%). In turn, these fingerprints were used to establish the relationships between the isolates (figure 3.3). The number of notified cases that were culture positive for *M. tuberculosis* varied from 66% in 2005 to 50% in 2002. The proportion of isolates genotyped between 2002 and 2007 varied from 96% in 2003 to 30% in 2007. These differences are demonstrated in figure 3.4.



1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29

Figure 3.2 Example of an RFLP image. The reference strain MT14323 is present in lanes 1, 12 and 29. Note the variation in band intensity between strains, which is caused by slight differences in DNA concentration. Also note the low copy number strains, for example. 02:327 in lane 26.

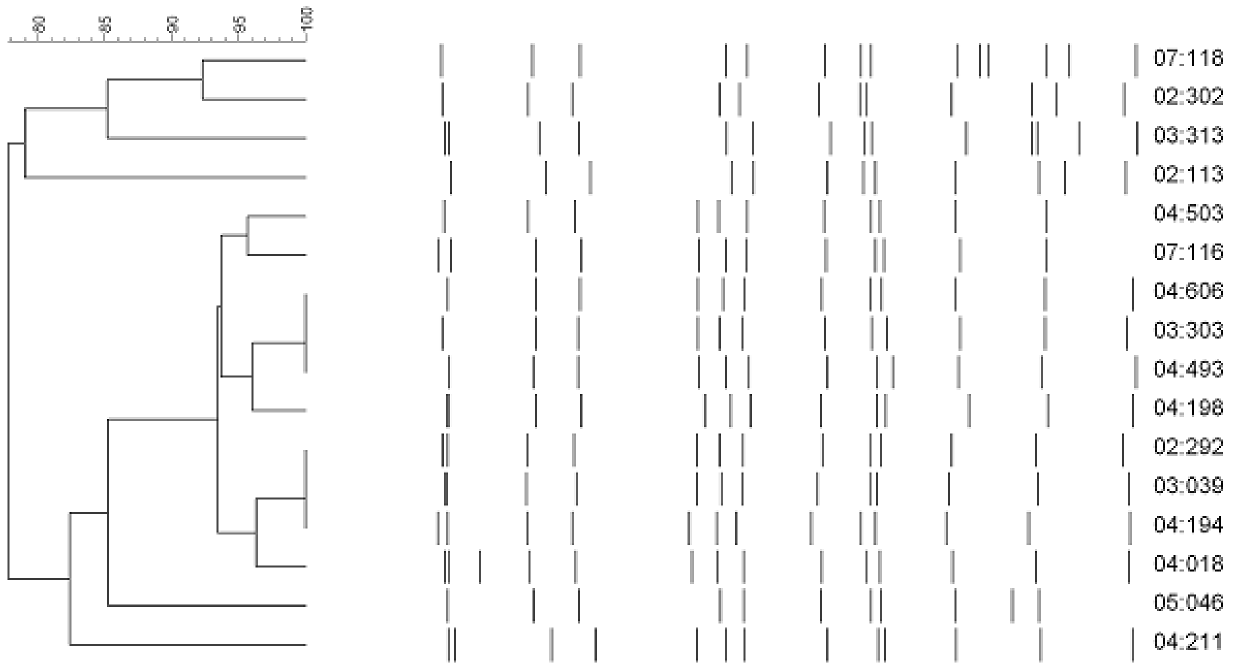


Figure 3.3 Example of IS*6110* RFLP dendrogram. The scale of relatedness is displayed on the left hand side – the vertical lines indicate how related the strains are. For example, 02:292, 03:303 and 04:493 are 100% similar (indistinguishable), whilst 07:118 and 02:302 are approximately 93% similar.

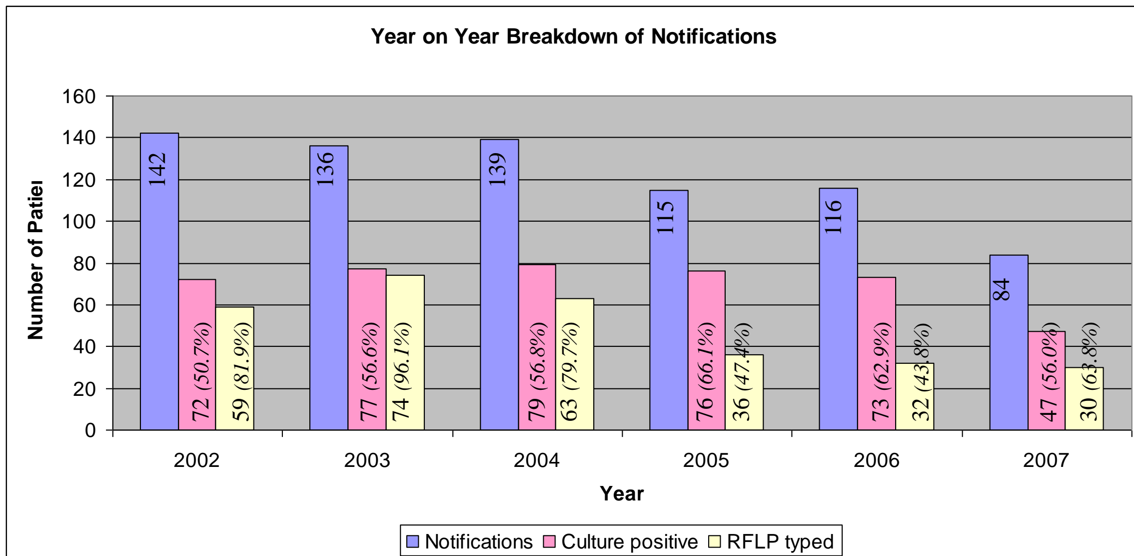


Figure 3.4 Notifications of TB from 2002 – 2007, including the breakdown of culture positive and number (%) of strains typed (Royal Free patients only).

A total of 478 strains were investigated during the years 2002 to 2007. The majority (387, 81.0%) were patients from Royal Free Hospital (RFH). The majority of the remainder were patient samples referred from University College London Hospital (UCLH) (82, 17.2%). The remaining nine isolates were referred from private practice. The notification of these patients was performed by the TB team at UCLH. Full patient demographic details were not available for all of the ULCH patients although if UCLH strains were implicated in significant clusters or lineages, every effort was made to obtain this information.

There was no significant difference in the age, gender or site of disease of the UCLH patients when compared to the RFH. This is shown in table 3.1.

	RFH patients **n = 394** **Number (%)**	**UCLH patients** **n = 84** **Number (%)**	
Patient age			
0-19	20 (5.2%)	6 (7.3%)	p = 0.8157 (Chi-square)
20-34	167 (43.0%)	35 (42.7%)	
35-59	146 (37.6%)	27 (32.9%)	
60+	55 (14.2%)	11 (13.4%)	
Gender			
Female	174 (44.8%)	31 (37.8%)	p = 0.2691 (Fisher's exact test)
Male	210 (54.1%)	50 (61.0%)	
Site of disease			
Extra Pulmonary	122 (31.4%)	26 (32.1%)	p = 0.6651 (Chi-square)
Pulmonary	230 (59.3%)	50 (61.7%)	
Pulmonary & Disseminated	36 (9.3%)	5 (6.2%)	

Table 3.1 The age, gender and site of disease of the Royal Free and University College London Hospital patients. These data show no difference between these patient groups in regard of age, gender or disease site.

3.2.2 Database demographics

The data set demonstrated a male (54%) to female (43%) ratio of 1.3:1 and an age range of <1 year to 97 years. The mean age was 38.7 years, the median 35 years and the mode 25 years (figure 3.5).

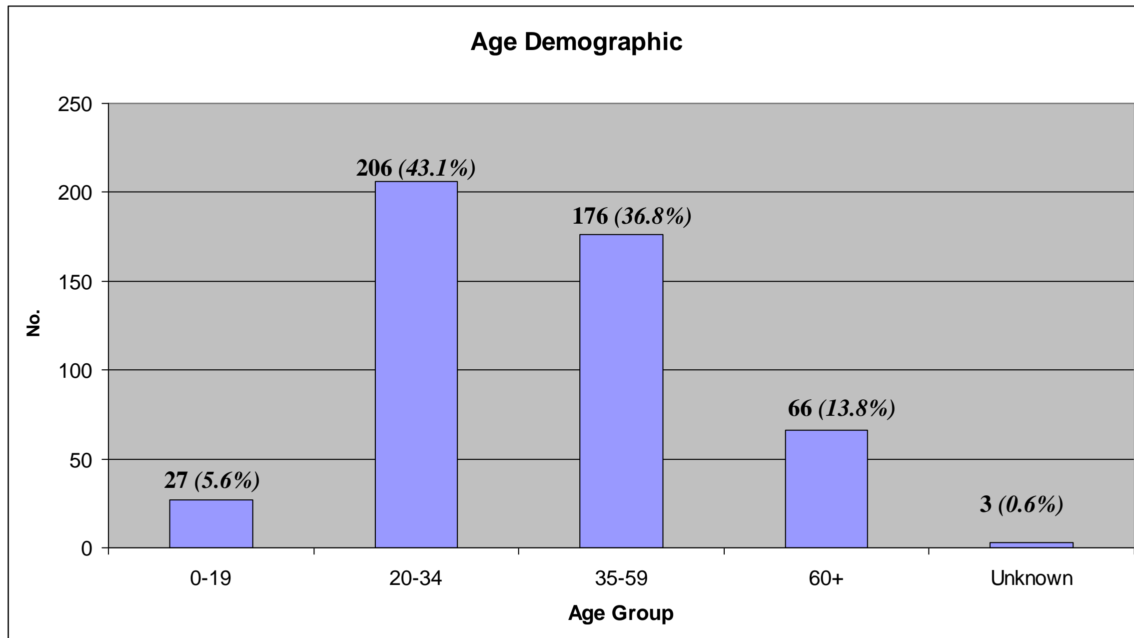


Figure 3.5 Distribution of patients by age. Almost 80% of patients are between 20 and 59.

The place of birth was available for 162 of the 478 patients, the majority of which (118, 72.8%) were born outside of the UK or Ireland. This data was analysed retrospectively and omissions from patient notes and the London TB Register could not always be rectified after the event. The majority of those patients who were born outside the UK were from the African continent (n = 49, 30.2%), with 42 (25.9%) from south Asia and the Middle East. The country of origin for these 162 patients is displayed in table 3.2.

Ethnicity was available for 242 patients. The majority were black African (n = 118, 48.8%), 47 (19.4%) were of European descent whilst 53 (21.9%) were from south Asia, which comprises India, Pakistan, Bangladesh, Sri Lanka, Afghanistan and Nepal. These data are demonstrated in figure 3.6.

Country	**Number**	**%**	**Country**	**Number**	**%**
Afghanistan	5	3.1	Mozambique	1	0.6
Bangladesh	2	1.2	Nigeria	2	1.2
Burma	1	0.6	Pakistan	3	1.9
Cameroon	1	0.6	Philippines	2	1.2
China	2	1.2	Poland	2	1.2
Columbia	1	0.6	Portugal	3	1.9
Dubai	1	0.6	Sierra Leone	2	1.2
Egypt	1	0.6	Slovakia	1	0.6
Ecuador	1	0.6	Somalia	23	14.2
Ethiopia	1	0.6	South Africa	2	1.2
Gambia	1	0.6	Spain	3	1.9
India	26	16.0	Sri Lanka	2	1.2
Iran	1	0.6	Tanzania	1	0.6
Ireland	5	3.1	Thailand	1	0.6
Italy	2	1.2	Tibet	1	0.6
Jamaica	3	1.9	Turkey	3	1.9
Japan	1	0.6	Uganda	1	0.6
Kenya	3	1.9	UK	39	24.1
Kuwait	1	0.6	Vietnam	1	0.6
Malawi	1	0.6	Zambia	2	1.2
Mauritius	1	0.6	Zimbabwe	4	2.5
Morocco	2	1.2			

Table 3.2 Country of birth of 162 TB patients. Sub-Saharan Africa accounts for the majority of patients, followed by the UK and South Asia.

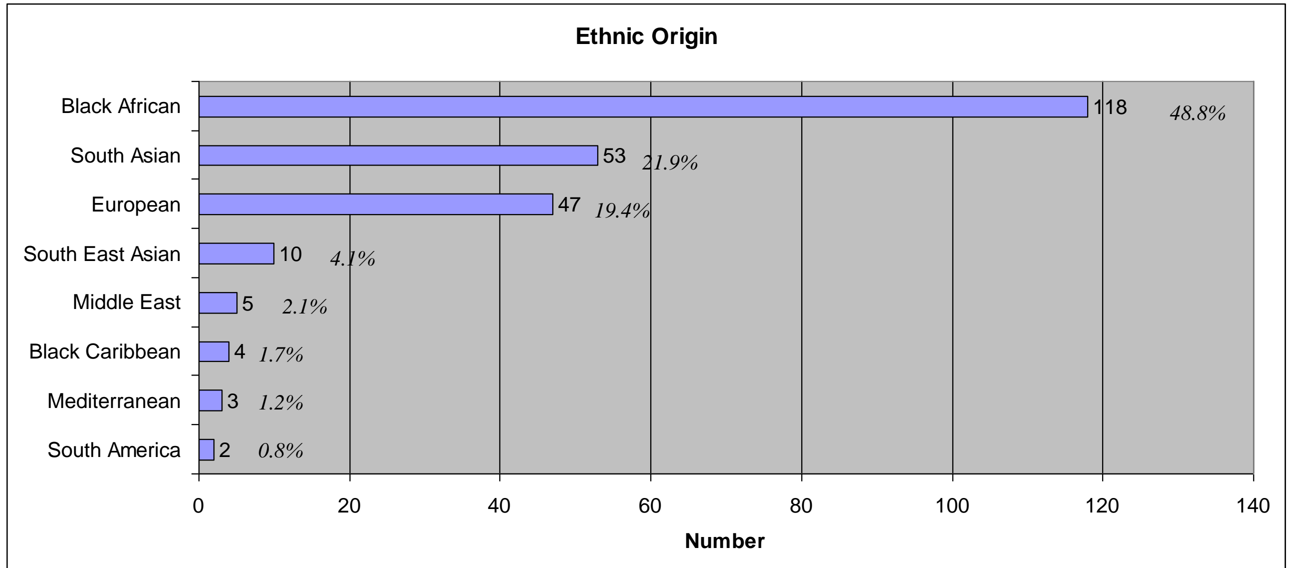


Figure 3.6 Distribution of patients by ethnicity. Black African patients comprise almost 50% of the patients. South Asian and European patients account for approximately 20% each.

The HIV status of 202 patients was known, 59 (29.2%) were HIV positive while 143 (70.8%) were HIV negative. Five patients refused an HIV test. The HIV status of the patients was analysed alongside other demographic details and parameters.

Of the 323 patients with respiratory involvement (figure 3.7), 165 (51.1%) had a respiratory sample that was smear positive. There was no correlation between HIV status and smear positivity (table 3.3).

	Smear positive	Smear negative	
HIV negative	62 (74.7%)	75 (69.4%)	p = 0.5171
HIV positive	21 (25.3%)	33 (30.6%)	(Fisher's exact test)
Total	83	108	

Table 3.3 Comparison of pulmonary smear and HIV status. These data show no association between smear and HIV status (Fisher's exact test).

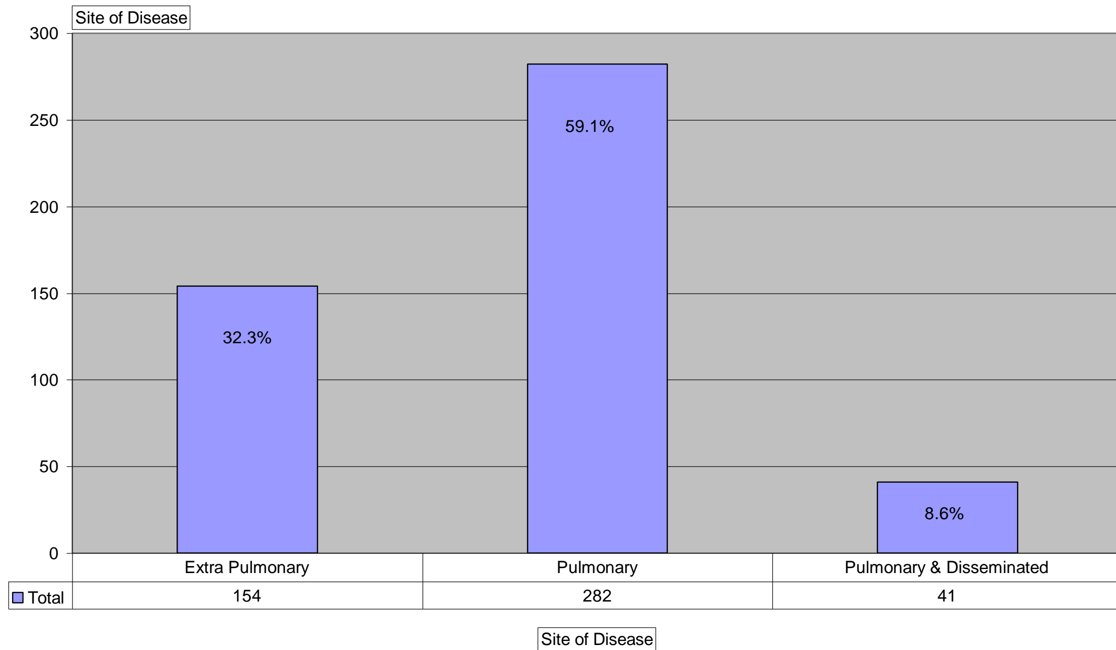


Figure 3.7 Site of disease. Nearly 60% of patients had pulmonary involvement and over 40% had extrapulmonary involvement.

The analysis of HIV status and ethnicity was performed. There were insufficient numbers to statistically examine all ethnicities represented, but there was no significant link between HIV status and black African, European or South Asian ethnicity. These data is shown in table 3.4.

	HIV negative	HIV positive	
Black African	61 (50.8%)	36 (66.7%)	p = 0.3385 (Chi-square)
European	24 (20.0%)	11 (20.4%)	
South Asian	35 (29.2%)	7 (13.0%)	
Total	120	54	

Table 3.4 Comparison of HIV status and ethnicity. These data show no association between HIV status and ethnicity (Chi squared test).

The potential link between smear positivity and ethnicity was investigated. As with HIV status, there were insufficient numbers from all ethnic groups for valid statistical analysis, so the largest represented groups were studied (table 3.5). The data showed that European patients were more likely, whilst south Asian patients were less likely to have smear positive disease.

	Smear negative	Smear positive	
Black African	68 (57.1%)	44 (51.2%)	P = 0.0004** (Chi-square)
European	15 (12.6%)	29 (33.7%)	
South Asian	36 (30.3%)	13 (15.1%)	
Total	119	86	

Table 3.5 Comparison of pulmonary smear status and ethnicity. There is a positive association between smear positivity and European ethnicity.

Additionally, the smear status and country of birth was analysed. Due to small patient numbers, this was broken down to UK or non-UK born (table 3.6). This shows that UK-borne patients are more likely to have smear positive disease.

	Smear negative	Smear positive	
UK born	15 (16.9%)	20 (32.3%)	p = 0.0322*
Non-UK born	74 (83.1%)	42 (67.7%)	(Fisher's exact test)
Total	89	62	

Table 3.6 Comparison of pulmonary smear status and country of birth. There is a positive association between smear positivity and being born in the UK.

3.2.3 Drug susceptibilities

Susceptibility data was available for 93.1% of the database (445/478 isolates). The database contained 370 (83.1%) sensitive isolates of *M. tuberculosis*. Resistance to any drug was present in 75 isolates (16.9%). Of these, isoniazid and streptomycin mono-resistance were the most prevalent, 29 (6.5%) and 18 (4.0%), respectively. Resistance to any of the commonly used first line drugs (rifampicin, isoniazid, ethambutol and pyrazinamide) was seen in 57 isolates (12.8%).

There were 13 (2.9%) isolates that were resistant to at least isoniazid and rifampicin, constituting multidrug resistant TB (MDR TB). The majority of these (10/13), were also resistant to other antimicrobial drugs. No extensively drug resistant (XDR-TB) isolates were identified and the full array of antimicrobial resistances is listed in table 3.7.

Resistance Profile	Number	%
Total number of strains	478	
Sensitivities available	445	93.1
Sensitive	370	83.1
H mono-resistant	29	6.5
S mono-resistant	18	4.0
R mono-resistant	1	0.2
Z mono-resistant	4	0.9
HS resistant	7	1.6
HZ resistant	1	0.2
HSE resistant	1	0.2
H Ethi resistant	1	0.2
HR resistant (MDR)	3	0.7
HRS resistant (MDR)	1	0.2
RHES resistant (MDR)	2	0.4
RHES Ethi resistant (MDR)	1	0.2
HR Cip resistant (MDR)	1	0.2
HSR Clari resistant (MDR)	1	0.2
HR Clari ethi resistant (MDR)	2	0.4
HRE Cip Cap Ethi resistant (MDR)	1	0.2
HRZ Ethi resistant (MDR)	1	0.2
Any resistance	75	16.9
Any first line resistance (not S)	57	12.8
MDR total	13	2.9
MDR plus additional resistance	10	2.2

Table 3.7 Break down of antimicrobial resistance. Key: H = isoniazid, S = streptomycin, R = rifampicin, Z = pyrazinamide, E = ethambutol, Ethi = ethionamide, Cip = ciprofloxacin, Clari = clarithromycin, Cap = capreomycin,

The relationship between drug resistance and other demographic data was analysed. Due to the small numbers of MDR isolates, these were excluded from statistical analyses. There was no link between HIV status and drug resistance (table 3.8). Additionally, there was no correlation between drug resistance and either ethnicity or country of birth (tables 3.9 and 3.10, respectively).

	HIV negative	HIV positive	
Isoniazid resistant	9 (6.8%)	5 (9.3%)	p = 0.7642 (Chi-square)
Resistant to any	8 (6.1%)	3 (5.6%)	
Sensitive	115 (87.1%)	46 (85.2%)	
Total	132	54	

Table 3.8 Comparison of drug susceptibility and HIV status. These data show no association between HIV status and drug resistance.

	Isoniazid resistant	Resistant to any	Sensitive	
Black African	6 (37.5%)	11 (68.8%)	91 (53.5%)	p = 0.3190 (Chi-square)
European	3 (18.8%)	2 (12.5%)	38 (22.4%)	
South Asian	7 (43.8%)	3 (18.8%)	41 (24.1%)	
Total	16	16	170	

Table 3.9 Comparison of drug susceptibility and ethnicity. These data show no association between ethnicity and drug resistance.

	Isoniazid resistant	Resistant to any	Sensitive	
UK born	2 (14.3%)	3 (33.3%)	30 (23.4%)	p = 0.5633 (Chi-square)
Non-UK born	12 (85.7%)	6 (66.7%)	98 (76.6%)	
Total	14	9	128	

Table 3.10 Comparison of drug susceptibility and country of birth. These data show no association between country of birth and drug resistance.

3.2.4 Site of infection

The site of disease was categorised as either pulmonary only, extrapulmonary only or as pulmonary and extrapulmonary. Most patients (323, 67.7%) had pulmonary disease. Of these, 282 (59.1%) patients had pulmonary disease only while 41 (8.6%) had disease in another site also. 154 (32.3%) patients had extrapulmonary disease only. This is summarised in figure 3.7. There was no association between the site of disease and HIV status (table 3.11).

	HIV negative	HIV positive	
Extra Pulmonary	36 (25.2%)	19 (32.2%)	p = 0.3385 (Chi-square)
Pulmonary	91 (63.6%)	31 (52.5%)	
Pulmonary & Disseminated	16 (11.2%)	9 (15.3%)	
Total	143	59	

Table 3.11 Comparison of site of disease and HIV status. These data show no association between HIV status and site of disaese.

3.2.5 Molecular characteristics

When examining the 478 strains in the database, it was found that the mean number of IS*6110* copies was 9.2 and the mode was 12. There were 90 (18.8%) strains which possessed fewer than five copies of IS*6110* and were deemed to be low copy number strains. The largest number of IS*6110* copies possessed by a strain was 21. This is demonstrated in figure 3.8.

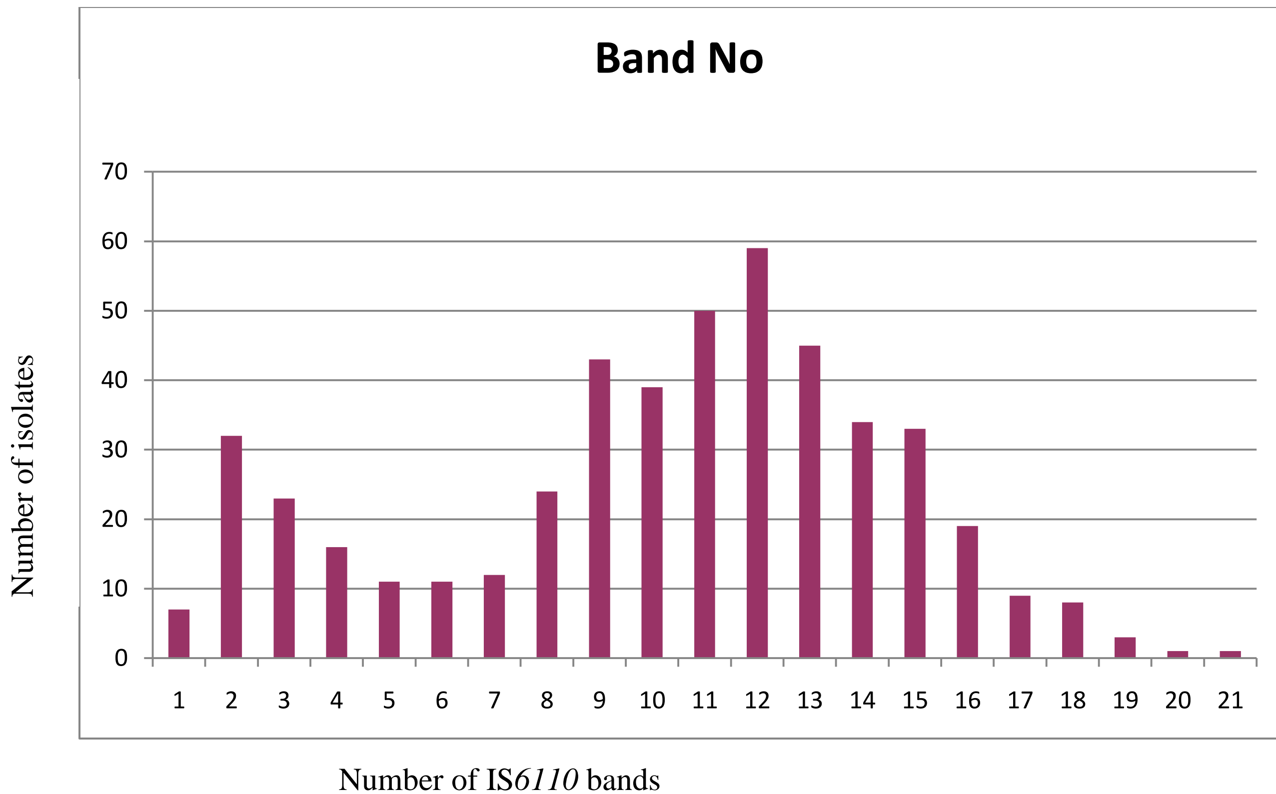


Figure 3.8 Distribution of strains with varying numbers of IS*6110*. These data show a bimodal distribution, comprising approximately 20% of strains with fewer than 5 *IS6110* copies. The mode number of copies was 12 and the maximum 21.

There were 28 clusters of 100% similarity involving 101 strains. However only 20 of these clusters contained strains which possessed five or more copies of IS*6110*. The majority of these clusters contained only two isolates (13 clusters) or three isolates (five clusters). One cluster contained four isolates and one contained eight isolates. A summary of these clusters and IS*6110* fingerprints is shown in table 3.12.

100% Cluster	Number of IS*6110* copies	No. of isolates per cluster	IS*6110* fingerprint	7.0 kb 4.4 kb 3.6 kb 3.0 kb 2.3 kb 2.0 kb 1.8 kb 1.5 kb 1.4kb
			MT14323	
1	4	2	07:182	
2	2	11	07:147	
3	1	3	03:495	
4	3	3	02:204	
5	2	2	03:442	
6	1	7	07:132	
7	8	2	02:230	
8	8	2	04:511	
9	11	2	02:109	
10	12	3	03:303	
11	13	3	02:292	
12	17	3	04:053	
13	9	2	03:188	
14	8	2	03:220	
15	5	3	04:612	
16	7	2	02:011	
17	12	8	03:014	
18	9	2	04:599	
19	8	2	04:468	
20	13	2	04:419	
21	14	3	02:093	
22	15	2	03:138	
23	11	6	04:473	
24	12	3	02:023	
25	13	2	02:129	
26	8	2	03:468	
27	1	3	03:251	
28	1	16	02:017	
Total No. of clustered isolates		101		
No. clustered isolates with ≥5 copies IS*6110*		56		

Table 3.12 Summary of 28 100% clusters. The given cluster number is listed in the first column. The number of IS*6110* copies is listed in column two. The number of isolates in each cluster is shown in column three whilst the fingerprint of a named member of each cluster is shown in columns for and five.

There were 70 lineages of 70% similarity containing 318 isolates. However, three of these lineages (containing a combined 39 isolates) possessed fewer than five copies of IS*6110*. Of the remaining 67 lineages (containing 279 strains), the majority (33) contained only two isolates. The distribution of the remaining lineages is outlined in table 3.13. The six most populated lineages are further displayed, including a representative IS*6110* fingerprint in table 3.14.

Number of isolates in a 70% lineage	**Number of lineages**
2	33
3	9
4	7
5	5
6	6
8	1
10	1
13	2
15	1
16	1
21	1

Table 3.13 Number and size of lineages. Most lineages (33) contained only two isolates. Six lineages contained 10 or more isolates.

70% Lineage	Range of number of IS*6110* copies	No. of isolates per lineage	% Similarity	Example of IS*6110* fingerprint	7.0 kb, 4.4 kb, 3.6 kb, 3.0 kb, 2.3 kb, 2.0 kb, 1.8 kb, 1.5 kb, 1.4 kb
				MT14323	
15	11 - 14	16	80.0%	03:303	
19	14 - 18	15	71.8%	04:053	
41	6 – 11	21	70.5%	02:011	
43	12 - 13	10	77.5%	03:014	
54	11 – 14	13	71.1%	02:093	
61	11 – 14	13	73.9%	02:023	

Table 3.14. Summary of the six most populated 70% lineages. The given cluster number is listed in the first column. The range of IS*6110* copies is listed in column two. The number of isolates in each cluster is shown in column three. The level of relatedness is shown in column four whilst the fingerprint of a named member of each cluster is shown in columns for and five.

Analysis of the 100% clusters was performed to establish two important pieces of information; the level of laboratory contamination and the rate of transmission in the community.

3.2.6 Laboratory Contamination Rate and Transmission Events

Following investigation with the multidisciplinary TB team at the Royal Free Hospital, it was agreed that six clustered isolates were due to laboratory contamination events. Using the total number of strains typed over the period of this study (n = 478), this gives a laboratory contamination rate of 1.3%.

The clustered strains involved in the probable laboratory contamination events (6) were removed from the total number of clustered strains (56) to give a total number of genuinely clustered strains as 50. Using the formula from Small and colleagues (equation 3.1; Small 1994) a rate of transmission rate of 6.3% was calculated.

3.2.7 Secondary Typing: MIRU-VNTR Data

The isolates from the first twelve months of this study were analysed using 12-loci MIRU-VNTR by conventional PCR and agarose gel electrophoresis. The subsequent isolates were genotyped using 24-loci MIRU-VNTR using an automated capillary DNA sequencer (Chapter 2, section 2.2). Isolates were selected for 24 loci MIRU-VNTR typing following the analysis of the IS*6110* genotypes. Isolates that were characterised in one of the six largest 70% lineages and those possessing fewer than five copies of IS*6110* were analysed.

All 12-loci MIRU-VNTR genotyping was optimised, implemented and performed by the author. The optimisation and implementation of the automated 24-loci genotyping method was also performed by the author. Analysis of the majority of the isolates in this chapter by the later method was performed by research assistants Laura Wright and Selina Bannoo.

3.2.7.1 Secondary MIRU-VNTR typing of ≥70% IS*6110* lineages

A MIRU-VNTR genotype was determined for 10 of the 16 isolates in lineage of North London isoniazid resistant strains (lineage 15, table 3.14). All ten were present in a MIRU-VNTR lineage at ≥70% similarity, demonstrating a genuine evolutionary lineage (figure 3.9). The analysis of the low copy number strains using MIRU-VNTR revealed two further isolates that appear to be members of this outbreak (05:373 and 05:344). The IS*6110* genotypes indicated that these isolates were not related to the North London outbreak strains, however the MIRU-VNTR types were indistinguishable from each other

and the outbreak strains. Additionally, 05:373 was an MDR strain, whilst 05:344 was isoniazid mono-resistant. This lineage is displayed in figure 3.9 and is discussed in more detail in Chapter 5.

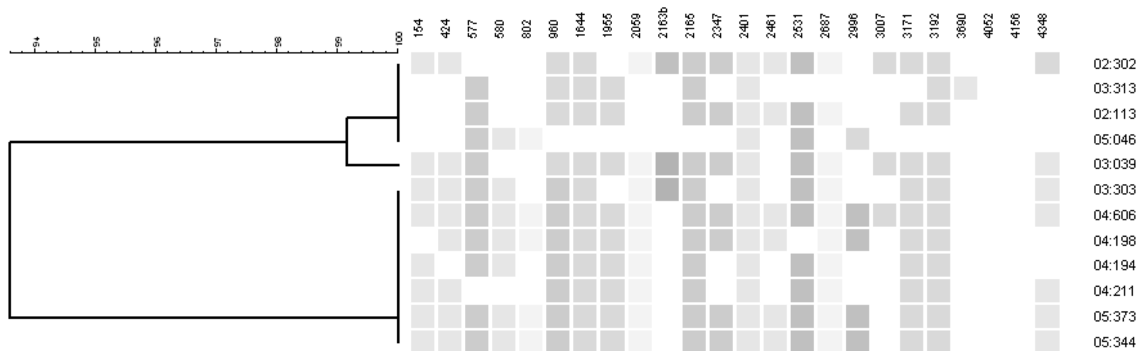


Figure 3.9. MIRU-VNTR dendrogram of IS*6110* RFLP lineage 15. The level of relatedness (%) is on the left hand side. The MIRU-VNTR loci names are listed at the top and isolate numbers are listed on the right. The IS*6110* lineage is split into two groups by MIRU-VNTR that are approximately 94% similar. This confirms that these isolates are closely related.

MIRU-VNTR analysis of 15 isolates in the IS*6110* lineage 19 revealed that two lineages are present. Two smaller lineages, 19a & 19b, consisting of five and six members, respectively, were seen. Isolate 05:085 did not converge with any others in this lineage, but only one of the 24 loci was amplified. Another isolate (06:060) aligned at a similarity of ≥70% with lineage 19b, but as only two loci were amplified, further work is required to confirm this. These results are displayed in figure 3.10.

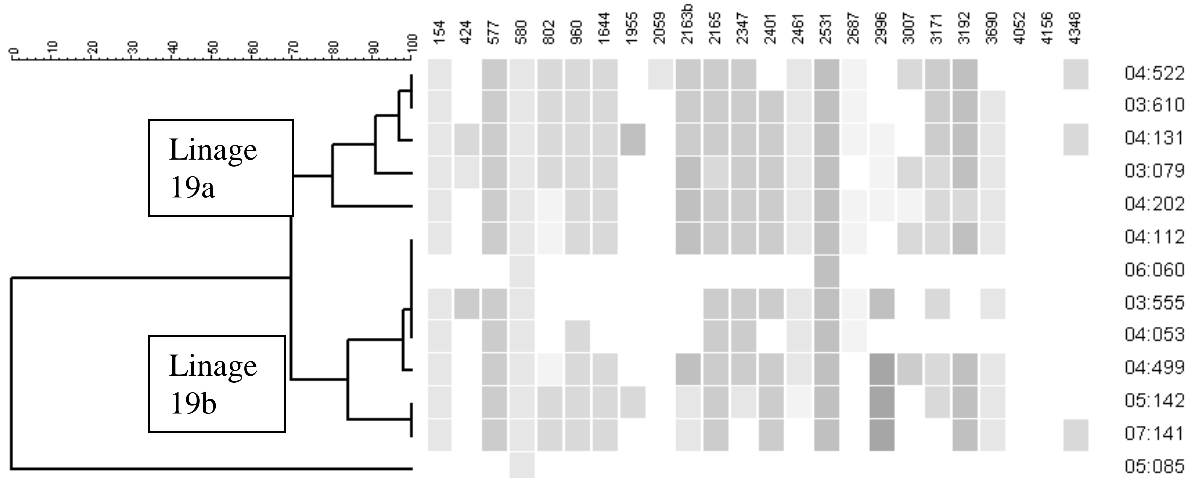


Figure 3.10 MIRU-VNTR dendrogram of IS*6110* RFLP lineage 19. The level of relatedness (%) is on the left hand side. The MIRU-VNTR loci names are listed at the top and isolate numbers are listed on the right. These data show IS*6110* lineage 19 being split into two lineages (19a & 19b) by MIRU-VNTR typing.

The remaining IS*6110* RFLP lineages (41, 43, 54 and 61) were largely confirmed as being evolutionarily related by secondary MIRU-VNTR typing. A MIRU-VNTR genotype was available on 18 of the 21 isolates in IS*6110* lineage 41. Fifteen of these isolates were confirmed at a similarity of ≥70% by MIRU-VNTR. IS*6110* lineage 43 contained ten members. A MIRU-VNTR genotype was available on eight of these and seven were confirmed at a similarity of ≥70% by this secondary typing method. IS*6110* lineages 54 and 61 both contained 13 isolates. A MIRU-VNTR genotype was available on eight and three of the isolates in these lineages and a ≥70% lineage was confirmed for six and three members of these lineages, respectively. The data is displayed in figures 3.11 – 3.14.

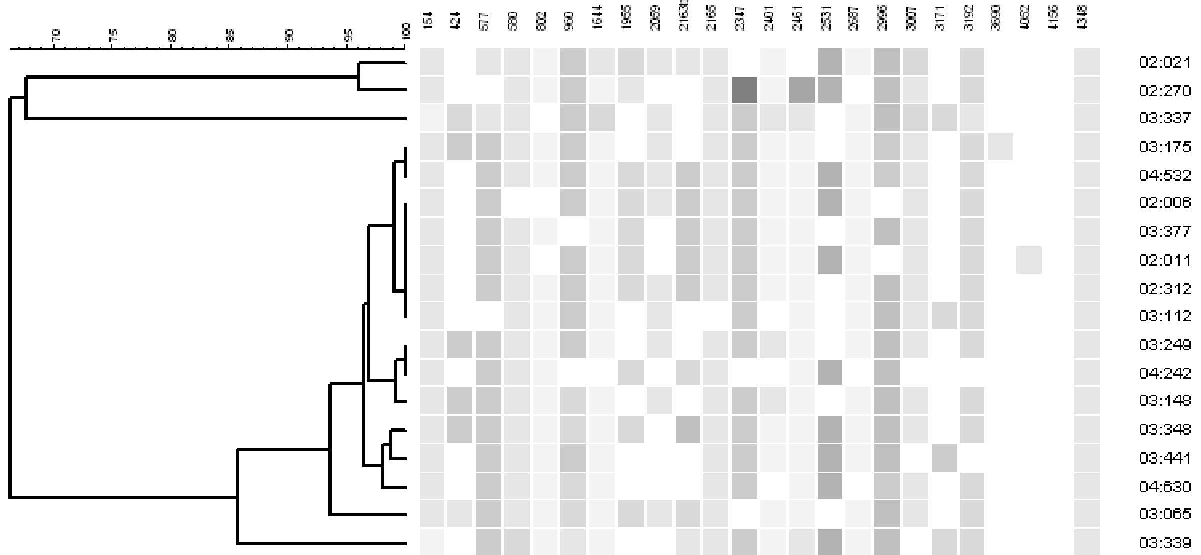


Figure 3.11 MIRU-VNTR dendrogram of IS*6110* RFLP lineage 41. The level of relatedness (%) is on the left hand side. The MIRU-VNTR loci names are listed at the top and isolate numbers are listed on the right.

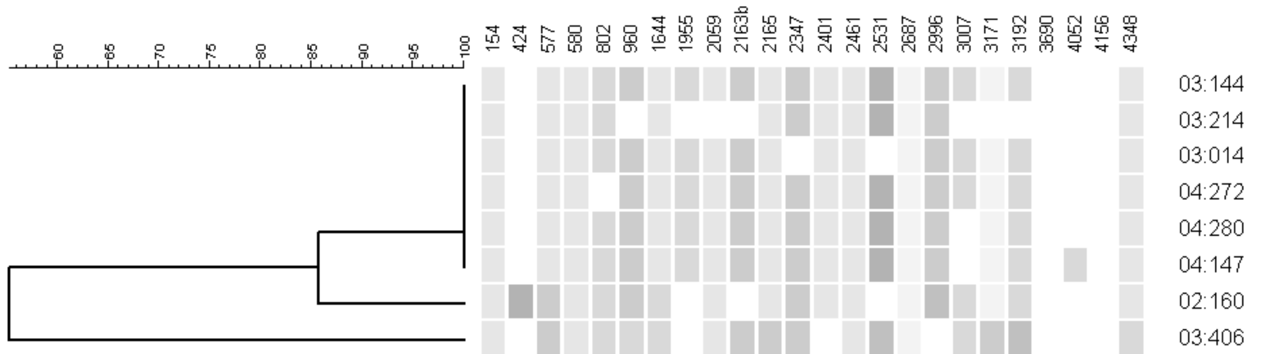


Figure 3.12 MIRU-VNTR dendrogram of IS*6110* RFLP lineage 43. The level of relatedness (%) is on the left hand side. The MIRU-VNTR loci names are listed at the top and isolate numbers are listed on the right.

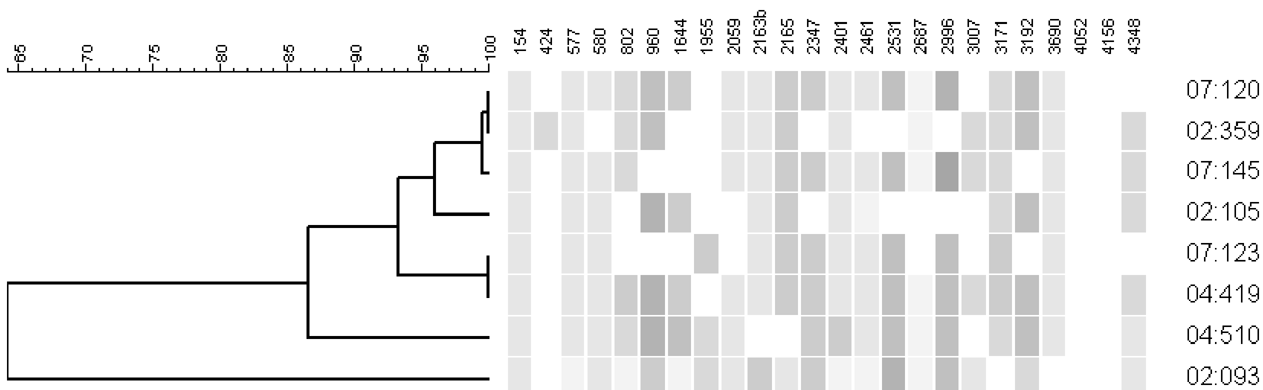


Figure 3.13 MIRU-VNTR dendrogram of IS*6110* RFLP lineage 54. The level of relatedness (%) is on the left hand side. The MIRU-VNTR loci names are listed at the top and isolate numbers are listed on the right.

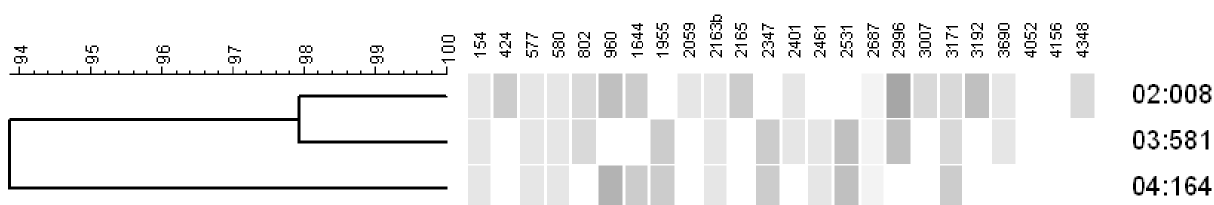


Figure 3.14 MIRU-VNTR dendrogram of IS*6110* RFLP lineage 61. The level of relatedness (%) is on the left hand side. The MIRU-VNTR loci names are listed at the top and isolate numbers are listed on the right.

The MIRU-VNTR genotypes were additionally analysed by Dr. Steve Platt at the Department of Bio Informatics, Health Protection Agency Centre for Infections using BioNumerics version 6.1 to generate a minimum spanning tree (MST). This data is displayed in figure 3.15. The MST demonstrates that the isolates grouped together into lineages by IS*6110* RFLP group closely together when typed with MIRU-VNTR.

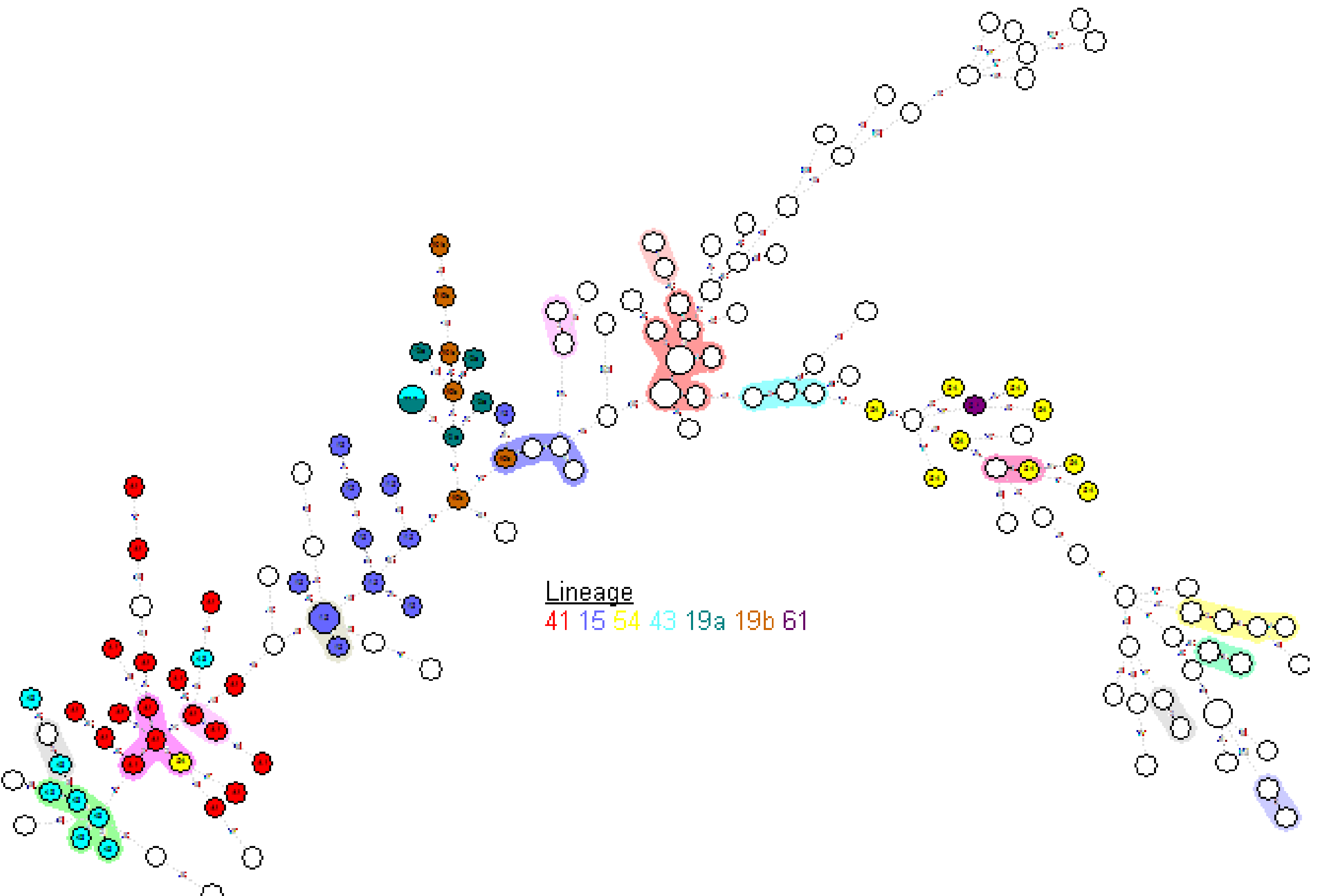


Figure 3.15 Minimum spanning tree (MST) of all isolates typed by MIRU-VNTR. Each node represents a single isolate, except larger nodes, which represent two identical isolates. IS6110 lineages 15, 19, 41, 43, 54 and 61 are highlighted. Numbers in light grey indicate the total number of loci/allele differences between isolates. The distance between nodes (isolates) are proportional to the number of differences on loci copy numbers between isolates. Isolates surrounded by shading are closely related and vary by fewer than two alleles at two loci (Courtesy Dr. Steven Platt).

3.2.7.2 Secondary MIRU-VNTR typing of IS6110 low copy number strains

The use of MIRU-VNTR typing revealed one possible transmission event of a strain with few copies of IS*6110*. The isolates 03:109 and 03:268 shared 18 common MIRU-VNTR loci (table 4.1). The patients were not shown to be epidemiologically linked when their notes were reviewed, however. The MIRU-VNTR dendrogram of all low IS*6110* copy number strains is displayed in figure 4.8.

	154	424	577	580	802	960	1644	1955	2059	2163	2165	2347	2401	2461	2531	2687	2996	3007	3171	3192	3690	4052	4156	4348
03;109	2	1	4		1	2	3	2	2		2	3	4	2	4	2	4	3	3	2	2	2		2
03;268	2	1	4	2	1	2	3		2		2	3	4	2	4	2	4	3	3	2	2	2		2

Table 3.15 MIRU-VNTR genotypes of two low-copy IS*6110* isolates. The MIRU-VNTR loci are listed across the top of the table. The copy number of each of the two isolates are listed below. The isolates are indistinguishable.

Isolates 02:099 and 03:427 were also matched at 100% similarity (figure 4.8). Investigation of the patients' notes showed that they were not epidemiologically linked. However, these isolates only shared 9 MIRU-VNTR loci. Isolates 05:344 and 05:373 were also indistinguishable by MIRU-VNTR genotyping and were identified as being members of the North London isoniazid outbreak and lineage 15 (figure 4.1).

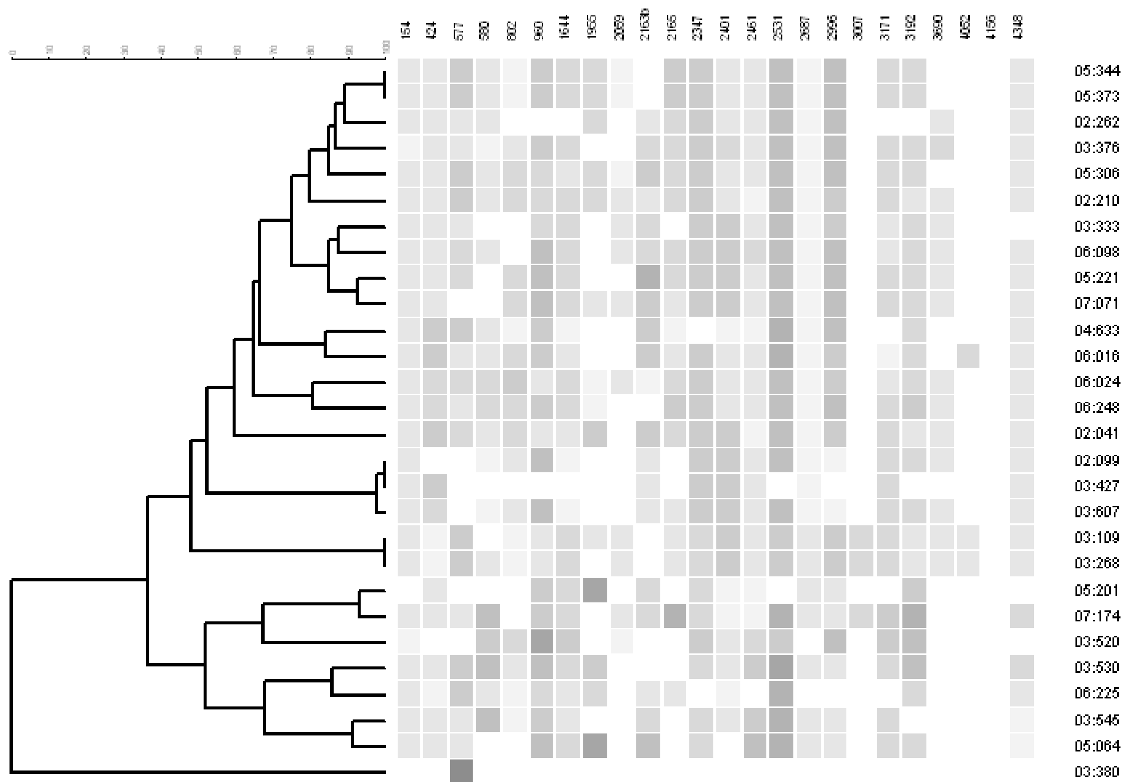


Figure 3.16 MIRU-VNTR dendrogram of all low IS*6110* copy number strains. The level of relatedness (%) is on the left hand side. The MIRU-VNTR loci names are listed at the top and isolate numbers are listed on the right.

3.3 Discussion

Investigations showed that of the patients who were notified to the Health Protection Unit (HPU) during this study, 57.9% were culture positive. This figure is comparable to the national rates in 2007 and 2008 of 57% and 56%, respectively (HPA 2008; HPA 2009). This will vary between centres due to a number of reasons; a proportion of patients remain culture negative despite having symptoms and a clinical history compatible TB. These patients often respond to anti-tuberculosis therapy and are given a clinical diagnosis of TB. Additionally, diagnostic specimens are not submitted from some patients and a proportion of samples become overgrown with contaminating commensal

organisms. Of these culture positive TB patients, 69.3% were genotyped using IS*6110* RFLP. The coverage varied between a low of 30% 2007 and a high of 96% in 2003. The number of patients who are culture positive (58%) is comparable with national data (Health Protection Agency 2009). It is of note that the patients whose isolates were genotyped (69%) were representative of our patient population as a whole with regard to age, gender ethnicity and site of disease. Additionally, this is comparable with the proportion captured in a study in London between 1995 and 1997 (Maguire et al 2002). We know that the patient population in London as a whole is very mobile and transient (Greater London Authority 2006) and these data provide a representative snapshot of the TB in our catchment area between 2002 and 2007.

There were 90 strains (18.8%) which possessed fewer than five copies of IS*6110,* which is comparable to the proportion seen in London between 1995 and 1997 (17.9%). Further work is needed to ascertain whether the strains possessing only one copy if IS*6110* belong to Principal Group IV (Sreevatsan et al. 1997; Baker et al. 2004).

There was no difference in age, gender and site of disease between the patients at the Royal Free and University College London Hospitals and therefore the two patient groups were treated as equal in these regards. The proportion of male patients in the dataset was 54.2% (female 43.1%). This is comparable to the proportions seen in a study in London between 1995 and 1997 (male = 58.6%, female = 41.4%) (Maguire et al. 2002) and the cases seen in the UK as a whole in 2008 (male = 54.2%, female = 45.2%) (HPA 2009).

The age of our patients was comparable to that of a study in London between 1995 and 1997 (Maguire et al. 2002). There was no difference between the ages of the patients in this study and those seen in the data of Maguire (table 3.16).

Age	RFH 2002 – 2007	Maguire 1995 – 1997	
0 – 19	5.6%	7.6%	p value = 0.9187 (two-way Anova)
20 – 34	43.1%	42.3%	
35 – 59	36.8%	33.0%	
60+	13.8%	17.1%	

Table 3.16 Comparison of patient ages between this data set and (Maguire et al.)

There are differences between the ethnicity of our set of patients compared with the UK as a whole. In the UK in 2008 52.2% of TB patients were born in South Asia (India, Pakistan, Bangladesh, Sri Lanka, Afghanistan and Nepal) (HPA 2009). Our data shows that the majority of our TB patients were black African (47.3%), while only 21.9% were from the South Asia. Demographic data, including the country of birth was not available for all patients. This data is collected as part of the routine TB service. Omissions in this data collection may be due to patients being unwilling to divulge potentially sensitive information or due to oversight because of pressures on the clinical service. These gaps are regrettable and cannot always be corrected if this data is collected retrospectively.

The three most commonly represented groups of patients by country of birth were the UK (24.1%), India (16.0%) and Somalia (14.2%). This proportion of TB patients of Black

African ethnicity was higher than those from South Asia (58.8% vs. 21.9%). This is divergent from national data where patients of South Asia ethnicity outnumber those black African (40.6% vs. 21.4%) (HPA 2009). This may be partly explained by the high number of Somalis in the catchment area of the Royal Free. The boroughs of Barnet, Camden, Islington and Westminster contain 13.7% of the nearly 34,000 Somalis in London, whilst only 9.5% of the approximately 172,000 Indians in London (Greater London Authority 2006). The full breakdown of ethnicity between this dataset, recent HPA figures and those seen in London between 1995 and 1997 are demonstrated in table 3.17.

Ethnicity	RFH 2002 – 2007	London 1995 – 1997 (Maguire 2002)	UK 2008 (HPA 2009)	
Black African	47.3%	25.2%	21.4%	p value = 0.0954 (two-way Anova)
Black Caribbean	1.6%	2.9%	2.0%	
South East Asian	8.2%		6.3%*	
European	21.6%	17.9%**	21.3%**	
South Asia	18.0%	22.1	52.2%***	
Middle East	2.0%	Not stated	Not stated	

Table 3.17 comparison of ethnicity in this dataset and that seen nationally (HPA 2009) and in London between 1995 and 1997 (Maguire et al. 2002). The data for these studies has been collected slightly differently. The following caveats apply: * = country of birth East & South East Asia, ** = white ethnicity, *** = country of birth South Asia.

The proportion of TB patients co-infected with HIV was 29.2%. This is significantly higher than the overall proportion in England and Wales in 2003 of 8.3% (Ahmed et al. 2007). This may be because London has a larger proportion of HIV-infected patients than

the rest of the UK (HPA 2008). Additionally, the Royal Free Hospital has one the largest HIV/AIDS centres in the country.

HIV status was not associated with drug resistance, site of disease, smear positivity or ethnicity. These findings are similar to those of Rodger and colleagues who found no association between HIV status and either smear positivity or drug resistance in TB patients in London (Rodger et al. 2010).

Rates of drug resistance in this study were divergent with national data. Resistance to at least one first line drug was seen in 12.8% of isolates. This is a higher rate than that seen in London as a whole in 2008 (8.5%) (HPA 2009). Isoniazid mono-resistance was the most commonly seen resistance profile (6.5% of isolates), which is comparable to national rates of 6.0% in 2008 (HPA 2009). Multidrug resistance was seen in 2.9% of isolates, which is higher than the national average in 2008 (1.1%) (HPA 2009). This may be because complex or drug resistant cases are often transferred from surrounding hospitals for the Royal Free for their care.

There was no evidence linking either ethnicity or country of birth to drug resistance. This may be because the absolute numbers of drug resistant strains is quite small and therefore may not reach statistical significance. Additionally, the dataset includes few patients from areas with high rates of drug resistance, including Eastern Europe, the former Soviet Union and the Far East (WHO 2008).

There was variation between the sites of disease in this dataset compared to national data. HPA data from 2008 revealed that 46% of patients had extrapulmonary disease only, whilst 54% had pulmonary disease (with or without extrapulmonary involvement). These data from the Royal Free show that only 32.3% had extrapulmonary disease only, whilst 67.7 had pulmonary disease (with or without extrapulmonary involvement). These figures may be discrepant because patients at the Royal Free with extrapulmonary disease may be under-represented in this dataset. Extrapulmonary sites, such as lymph nodes, are often only sampled once and may remain culture negative. Alternatively, samples may not be obtained at all and a diagnosis may be made on clinical evidence alone. These situations may mean that an organism will not be available genotyping.

There was an association between ethnicity and smear positivity. European patients were more likely, whilst South Asian patients were less likely to have smear positive disease (p=0.0004). A possible explanation for this data is that given the majority of TB patients in the UK are non-European, other causes may be higher on the list of differential diagnoses. It is possible that these patients go undiagnosed for longer, which allows their disease to progress and develop into a higher bacterial burden. The data collected here does not allow this hypothesis to be examined. A retrospective cohort study of the time difference between onset of symptoms, first presentation and diagnosis would be required for this.

Likewise, there was an association between country of birth and smear positivity. UK born TB patients were more likely to have smear positive disease (p=0.322). This

association is less strong than that demonstrated between ethnicity and smear positivity, possibly because a large proportion of TB patients born in the UK are from ethnic groups with high rates of TB. Therefore, the diagnosis of TB in these patients may be reached more rapidly than in those of white ethnicity born in the UK.

One hundred and one isolates (21.1%) clustered with at least one other isolate. This is significantly lower than the 46% of strains that were in clusters in the Netherlands between 1993 and 1997 (van Soolingen et al. 1999) and 31.8% seen in Madrid between 2002 and 2004 (Iñigo et al. 2008) but more similar to the 22.7% of strains in clusters in London between 1995 and 1997 (Maguire et al. 2002).

The clustered isolates were used to calculate the rate of laboratory contamination and active transmission. Of the isolates typed, 1.3% of positive cultures of *M. tuberculosis* were likely to be laboratory contaminants. This rate is comparable to, or better than, other centres (Anonymous 1997; Burman and Reves 2000; Ruddy et al. 2002; de Boer et al. 2002). Contamination events occur due to a number of reasons. Clerical errors can occur (Nitta et al.), but one study showed that these were less common than laboratory contamination, as the samples that cultured presumed contaminants contained fewer organisms than other positive samples being processed at the same time (Mitchison DA et al. 1980). Additionally, the use of common reagent containers has been shown to be a source of cross contamination between clinical sample (de C Ramos et al. 1999).

The need for standardised procedures and good technique is essential. One study showed that 93% of laboratory contamination events were attributable to a single technician, probably because she was left-handed and the equipment within the biological safety cabinet was not arranged in a suitable manner (Bauer et al. 1997).

For this reason, a number of interventions are in place at our laboratory and are reflected in the Standard Operating Procedures (RFH Microbiology 1; RFH Microbiology 2). Only eight samples are processed in the biological safety cabinet at any one time with only one specimen being opened at a time. There are workflows for right- and left-handed individuals to avoid droplet contamination and the cabinet is cleaned with an appropriate decontaminating agent regularly. Additionally, any disposable equipment, such as Pasteur pipettes or inoculating loops are kept outside the cabinet to avoid contamination, or discarded at the end of the day if they have been inside the cabinet.

Once the probable contaminants in this study were removed from the clusters, it was calculated that 6.3% of typed isolates were from patients involved in active transmission events. This rate is lower than several other studies (1.9.5; Table 1.2). Of note, this rate is lower than a study in London between 1995 and 1997 (14%) (Maguire et al. 2002). This may be due to the smaller geographical area of this study;. Some TB patients in an active chain of transmission may seek medical attention elsewhere and therefore will be excluded from the sample. This will therefore falsely decrease the level of transmission.

The largest cluster, contained eight indistinguishable strains, and included four members of the same Somali family. They were diagnosed over a 13 month period from March 2003 and April 2004 after entering the UK in 1990. Almost 50% of TB patients born outside the UK are diagnosed within five years of entering the country (HPA 2009). 29% of TB patients have been in the UK for ten or more years before TB is diagnosed. The members of this family would be positioned near the top end of the inter-quartile range (2 – 15 years) (HPA 2009). All eight members of the cluster had fully sensitive pulmonary disease. Three of the four family members were sputum smear positive. Epidemiological links between this family and the remaining four members of the cluster were not found following further investigation and contract tracing.

Two isolates from a cohabiting couple were indistinguishable and accounted for one 100% cluster. Both patients were diagnosed with pulmonary TB by bronchoscopy. Transmission of *M. tuberculosis* to household contacts is a common event and the female of the couple was diagnosed following routine contact tracing (Becerra et al. 1986; Wang and Lin 2000; Kilicaslan et al. 2009; Sia et al. 2010).

Two other 100% clusters were of note. Two men in their fourth decades were shown to have indistinguishable strains. They had both spent considerable amounts of time in prison. They were not in the same institution at the time of their diagnoses but records showed that they had both moved around several prisons in the capital during the preceding years. Two further clustered isolates were cultured from one male and one female in their fifth decades. Both had been intermittently homeless and had histories of

misusing alcohol. However, the individuals could not be linked epidemiologically by contact tracing. This does however indicate that strains are circulating within specific communities.

The RFH *IS6110* database shows that 60 of the 67 70% lineages contain six or fewer isolates. There were seven lineages containing eight or more isolates. It is difficult to compare this to other geographic locations as most molecular epidemiological TB studies concentrate on the degree of clustering and active transmission. However, a study in northern Tanzania showed that approximately one third of strains were members of only two 70% lineages as defined by IS*6110* and MIRU typing (McHugh et al. 2005). This may be a result of a more static human population than is seen in London. Indeed, the Greater London Authority estimates that in 2006 almost one in three (31.67%) individuals living in London were born outside of the UK (Greater London Authority 2006). These large lineages were secondarily typed using MIRU-VNTR.

Secondary typing of Lineage 15 as assigned by IS*6110* RFLP was confirmed by MIRU-VNTR typing. Two further isolates shared identical MIRU-VNTR genotypes with other members of this lineage despite being originally categorised as low copy number strains. Review of the original RFLP revealed that not only is RFLP a technically difficult method but it requires subjective interpretation. The additional finding that both of these isolates were isoniazid resistant, and indeed one was multi-drug resistant, adds weight to the argument that these strains should be included in the outbreak lineage despite their putative IS*6110* RFLP genotype. Poor quality images are often difficult to interpret, in

our experience. For example, the RFLP image shown in figure 3.2 shows some variation in DNA quantity between loaded isolates. This is despite procedures put in place to combat this, such as DNA quantification both prior to, and after, restriction. DNA was quantified using serially diluted lambda DNA and agarose gel electrophoresis with ethidium bromide staining. More accurate results may have been obtained if the DNA was quantified using a spectrophotometric method. Strains that show no bands were analysed again to confirm that DNA was successfully extracted and that there were genuinely no copies of IS*6110* present. Lanes that showed large, intense bands were analysed thoroughly by using the contrast and intensity functions on the software to accurately assign the centre of such bands. This lineage (related at 80% by IS*6110* and 94% by MIRU-VNTR) contained 16 isolates. Fourteen of these isolates were resistant to at least isoniazid and genotyping at the HPA MRU revealed they were part of the North London isoniazid-resistant outbreak (Ruddy et al. 2004).

IS*6110* RFLP lineage 19 was subdivided into two smaller lineages by MIRU-VNTR genotyping. The original lineage contained 15 isolates, which were split into one lineage of seven isolates (lineage 19a) and one of four isolates (lineage 19b). Two isolates that were placed in the original lineage by IS*6110* RFLP genotyping were not related at 70% by MIRU-VNTR typing and therefore are not genuine members of this lineage. This finding indicates that 24-loci MIRU-VNTR is more discriminatory in this setting than IS*6110* RFLP.

The remaining four lineages were confirmed by secondary MIRU-VNTR typing. For example, of the isolates for which a MIRU-VNTR type was available, 15 of 18 isolates in lineage 41 were related at 70% by the secondary typing method. Therefore, these 15 isolates are genuinely related within this lineage. All six lineages are summarised in table 3.18. The observation that not all isolates were confirmed at 70% similarity by MIRU-VNTR typing indicates that this method is more discriminatory than IS*6110* RFLP.

Large lineages indicate an over-representation of a group of closely related strains in a population. Successful lineages have been previously documented. The W/Beijing family (van Soolingen et al. 1995), the Kilimanjaro and Meru lineages in northern Tanzania (McHugh et al. 2005) and the Haarlem family (Kremer et al. 1999) are all lineages of strains that are over-represented in a community. The over-representation of a group of strains may indicate their success or 'fitness'. The strains may have a selective advantage over other strains within the same geographic locus.

IS*6110* RFLP Lineage	Number of original members	Number of MIRU-VNTR types available	Number of isolates confirmed at ≥70%	Comments
15	16	10	10 (100%)	Two additional members identified by MIRU-VNTR
19	15	13	11 (85%)	Subdivided into two lineages (19a & 19b)
41	21	18	15 (83%)	
45	10	8	7 (88%)	
54	13	8	6 (75%)	
61	13	3	3 (100%)	

Table 3.18. Summary of the number of isolates in the IS*6110* RFLP lineages that were confirmed as being 70% related by MIRU-VNTR, the secondary typing method.

The Beijing/W family are an example of this is a group of strains that was first seen in the Beijing area of China (van Soolingen et al. 1995). They have a characteristic spoligotype, but are more loosely related by other typing methods. Representatives of the W/Beijing family are now seen worldwide and have been intensively investigated. The success of this group of strains has been hypothesised to be linked to its virulence, fitness (Zhang et al. 1999), mutability (Ebrahimi-Rad et al. 2003), acquisition of drug resistance and immune-modulatory abilities (Lopez et al. 2003; Reed et al. 2007).

Isolates within such lineages may be studied to examine why they are so prevalent and to ascertain any measurable advantage that allows them to spread widely. Fitness assays (Chapter 6) and mutation rate experiments can be used to measure any advantage that is possessed by these strains (O'Sullivan 2008).

The secondary typing of these strains using mycobacterial interspersed repeating units (MIRU), chapter 4, allows a more complete characterisation of the strains found in this location at this particular time. This will allow clusters and chains of transmission to be confirmed or refuted as well as the definitive characterisation of lineages. Additionally, the isolates that possess fewer than five copies of IS*6110* may be studied.

3.3.1 MIRU-VNTR Discussion

Some loci were more variable than others. Locus 154 was amongst the least variable with only two alleles seen (1 and 2 copies). Over 90% of the isolates possessed two copies at this locus. The three most variable loci were 2163, 2165 and 4052. No single allele accounted for more than 40% of the isolates at these loci. These findings are concordant with other published data demonstrating the most and least variable loci (Supply et al. 2006). The summary of the variability of the loci are displayed in table 3.19.

Locus	No. Of Alleles	Copy No.	%	Copy No	%	Copy No	%	Copy No	%	Copy No	%	Copy No	%	Allele	%
154	2	1	**8.0**	2	**92.0**										
424	5	1	**10.8**	2	**50.0**	3	**17.6**	4	**20.3**	6	**1.4**				
577	6	1	**1.8**	2	**38.4**	3	**10.7**	4	**45.5**	5	**2.7**	9	**0.9**		
580	5	1	**7.3**	2	**78.9**	3	**7.3**	4	**0.9**	5	**5.5**				
802	4	1	**30.3**	2	**17.2**	3	**42.4**	4	**10.1**						
960	6	2	**7.2**	3	**22.5**	4	**45.0**	5	**16.2**	6	**4.5**	7	**4.5**		
1644	5	1	**23.4**	2	**14.4**	3	**49.5**	4	**11.7**	5	**0.9**				
1955	7	1	**2.1**	2	**7.2**	3	**57.7**	4	**16.5**	5	**13.4**	6	**1.0**	7	**2.1**
2059	4	0	**17.6**	1	**11.8**	2	**68.1**	4	**2.5**						
2163	6	1	**1.9**	2	**34.3**	3	**21.9**	4	**23.8**	5	**10.5**	6	**7.6**		
2165	5	1	**1.9**	2	**34.3**	3	**21.0**	4	**37.1**	6	**5.7**				
2347	5	1	**3.9**	2	**2.4**	3	**15.0**	4	**78.0**	10	**0.8**				
2401	5	0	**2.5**	1	**18.3**	2	**48.3**	3	**3.3**	4	**27.5**				
2461	6	1	**20.2**	2	**69.0**	3	**3.1**	4	**3.9**	5	**3.1**	7	**0.8**		
2531	6	3	**4.0**	4	**4.0**	5	**60.3**	6	**29.4**	7	**1.6**	10	**0.8**		
2687	4	1	**83.3**	2	**13.8**	4	**1.4**	6	**1.4**						
2996	7	1	**15.2**	2	**11.4**	3	**3.8**	4	**19.0**	5	**44.8**	6	**1.0**	7	**4.8**
3007	4	1	**1.3**	2	**18.8**	3	**77.5**	4	**2.5**						
3171	4	1	**8.5**	2	**10.3**	3	**71.8**	4	**9.4**						
3192	5	2	**6.4**	3	**63.3**	4	**7.3**	5	**21.1**	6	**1.8**				
3690	3	1	**3.3**	2	**83.6**	3	**13.1**								
4052	5	1	**15.8**	2	**21.1**	3	**31.6**	4	**10.5**	5	**21.1**				
4156	3	0	**25.0**	1	**50.0**	3	**25.0**								
4348	4	1	**5.7**	2	**72.1**	3	**20.5**	4	**1.6**						

Table 3.19. Summary of the variability of MIRU-VNTR loci. Loci where a single allele accounts for ≤40% of isolates are the most variable and are highlighted in green. Loci where an allele accounts for 40-50% of isolates are highlighted in yellow. Loci where an allele is seen in 50-60% of isolates are in orange, whilst loci where an allele accounts >60% of isolates (the least variable) are highlighted in red.

It was not possible to amplify all 24 MIRU-VNTR loci for any of the isolates analysed. It is widely reported that some loci are more difficult to amplify than others. Considerable effort was expending to optimise these triplexed PCR reactions. This process involved performing magnesium and primer concentration titrations. Additionally, the injection time of the automated sequencer was altered in an attempt to detect low level amplification of some loci. Failed reactions were often repeated if the level of discrimination was not sufficient. Failed reactions occurred commonly at loci 4052, 4156 and 4348. All three of these loci are grouped together in a similar region of the genome and are amplified using primers labelled with the yellow fluorophore NED.

The consistent failure to amplify these loci warrants further investigation. The complex optimisation performed by Supply and colleagues, particularly related to magnesium concentrations for different master mixes, is testament to this (Supply 2005). Additionally, other groups have noted that some loci are more problematic than others and have raised concerns regarding the conditions used for amplification (Velji et al 2010). This group also highlighted the importance of standardising the amplification conditions, even including the polymerase enzyme, to ensure reproducibility of hypervariable loci. Furthermore, centres performing MIRU-VNTR typing on large numbers of isolates have also experienced difficulties in amplifying some loci (Dr. T Brown, NMRL, personal communication).

In summary, MIRU-VNTR has several advantages over other typing methodologies, for example; it is highly discriminatory when 15 or 24 loci are used, it may be automated, the

genotypes may be analysed objectively between centres. These data suggest that it is more discriminatory than IS*6110* RFLP and can additionally distinguish isolates that possess fewer than five copies of IS*6110* (section 1.9.4.2), which has been described elsewhere (Kremer et al. 2005). The method is less prone to technical difficulties and can provide useful data even if not all 24 loci are amplified. The observation that two potentially linked strains were omitted from an outbreak lineage due to poor RFLP image quality is worrying. The MIRU-VNTR genotypes, together with their phenotypic susceptibilities, suggest that they are indeed related.

Some isolates appear to be indistinguishable using MIRU-VNTR, such as 02:099 and 03:427 (these two isolates were classified as low copy number strains by IS*6110* RFLP). However, due to failed PCR reactions, only nine corresponding loci could be amplified in both isolates. No epidemiological links were found between the patients and considering the gaps in the MIRU-VNTR data, describing these isolates are indistinguishable may be disingenuous.

The clinical usefulness of MIRU-VNTR typing is evident from these data. Small quantities of DNA are required so that MIRU-VNTR typing may be performed on cultures of *M. tuberculosis* at an earlier stage than is required by IS*6110* RFLP. Indeed, given that small quantities of DNA are required, and the degree of automation involved, it is possible to generate a genotype within 24 hours of a positive culture being identified. This advantage, together with its ability to genotype the approximately one in five strains in London that possess fewer than five copies of IS*6110* means that of MIRU-VNTR

should be used as a first line typing technology. MIRU-VNTR is now widely adopted globally as a typing tool due to its ease of use, easily interpretable and transportable data and rapid turnaround. It should therefore be considered as a first-line genotyping methodology and other methods, such as IS*6110* RFLP, may be used as secondary tools.

In summary, the Royal Free genotyping database shows a diverse population of TB patients, which is largely concordant with the conventional epidemiology of TB in London (Maguire et al. 2002; HPA 2009).

Chapter 4: ESAT-6 and the Host Response

4.1 Introduction

The aim of this current study was to test the hypothesis that the antigen specific response seen in the patients' T cells was related to the biology of the infecting organism as compared to that of the host. Here, three hypotheses are examined; firstly, those patients whose T cells do not respond to ESAT-6 stimulation may be infected with related strains of *M. tuberculosis* and therefore the response may be associated with the biology of a variant strain. Secondly, the organisms isolated from patients whose T cells did not respond to ESAT-6 stimulation may lack the *esxA* gene, or may possess the gene, but with mutations. Thirdly, the level of gene expression of *esxA* in the isolates may play a role in the ESAT-6 specific IFN-γ response.

Proteins secreted by *M. tuberculosis* during growth are thought to be important in protective immunity against tuberculosis (Andersen et al. 1991). The study of short-term culture filtrates identified a novel low-molecular-mass protein with a mass of 6kDa that was immunogenic in murine models. The antigen, named early secretory antigen target 6 (ESAT-6), was found to stimulate a dose-dependent IFN-γ response in the *ex-vivo* murine T-cells (Sorensen et al. 1995 ;Andersen et al. 1995 ;Brandt et al. 1996). This antigen was detected in *M. tuberculosis* reference strains $H_{37}Rv$ and $H_{37}Ra$, a clinical isolate of *M. tuberculosis* and *M. africanum*, as well as in *M. kansasii*, *M. szulgai* and *M. marinum*. Interestingly, the antigen was absent from vaccine strains of *M. bovis* BCG (Sorensen et al. 1995; Harboe et al. 1996).

The gene encoding this protein was identified by sequencing, and further analysis showed that ESAT-6 was a polypeptide of 95 amino acids. Identified directly upstream from the gene encoding ESAT-6 was a co-transcribed gene named *lhp* or *esx*B (Berthet et al. 1998). This gene was annotated Rv3874 in the genome sequence of *M. tuberculosis* $H_{37}Rv$ (the gene encoding ESAT-6 was Rv3875 and named *esxA* or *esat*6) (Cole ST et al. 1998). The gene product, isolated from the low-molecular-mass short-term culture filtrate had an apparent molecular mass of 10kDa and was named culture filtrate protein-10 (CFP-10) (Berthet et al. 1998). Renshaw and colleagues demonstrated that ESAT-6 and CFP-10 are found to exist in solution in 1:1 ratio (Renshaw et al. 2005). Additionally, there appeared to be several related proteins in encoded in the *M. tuberculosis* genome and these were termed the *lhp*/*esat*-6 gene family (Berthet et al. 2005).

Comparative genome analysis showed that several regions of difference, including RD1, were present in *M. tuberculosis* $H_{37}Rv$ and *M. bovis*, but absent from the vaccine strain *M. bovis* BCG, and was reasoned to be a contributor to the vaccine strain's attenuation (Mahairas et al. 1996; Gordon et al. 1999; Behr et al. 1999). The genes *esa*t6 and *lhp* were found to be in the *M. tuberculosis* RD1 (spanning the open reading frames Rv3871 – Rv3879) in a cluster with a further 12 genes (figure 4.1). Analysis of the *M. tuberculosis* proteome revealed that there were several proteins similar to ESAT-6. This ESAT-6 family comprises 14 members ranging in size from 90 to 120 amino acids (Tekaia et al. 1999). Additional homologues were also described in *Streptomyces coelicolor* and *Corynebacterium diphtheriae* (Gey Van Pittius et al. 2001).

The roles of ESAT-6, CFP-10 and RD1 in mycobacterial virulence and immunity were investigated. An *esat*6 knockout mutant of *M. bovis* created by Wards and colleagues demonstrated reduced virulence by histopathological examination of lesions in animal models (Wards, de Lisle, and Collins 2000). Additionally, the loss of RD1 was shown, by Pym and colleagues (Pym et al. 2002a) to have contributed to the attenuation in *M. bovis* BCG and the vaccine strain *M. microti. M. bovis* BCG and *M. microtti* were complemented by knocking-in RD1, RD3, RD4, RD5, RD7 and RD9. The strains complemented with RD1 demonstrated altered colonial morphology towards that of virulent tubercle bacilli as well as increased virulence in mice, by the observation of increased growth rates, granuloma formation and splenomegaly, compared to controls. The affect of deleting RD1 from *M. tuberculosis* H_{37}Rv was demonstrated by Lewis and colleagues. The *M. tuberculosis* H_{37}Rv RD1 knock-out showed reduced growth rates (similar to *M. bovis* BCG controls) in the human-macrophage-like cell line THP-1, human monocyte-derived macrophages as well as mice (Lewis et al. 2003).

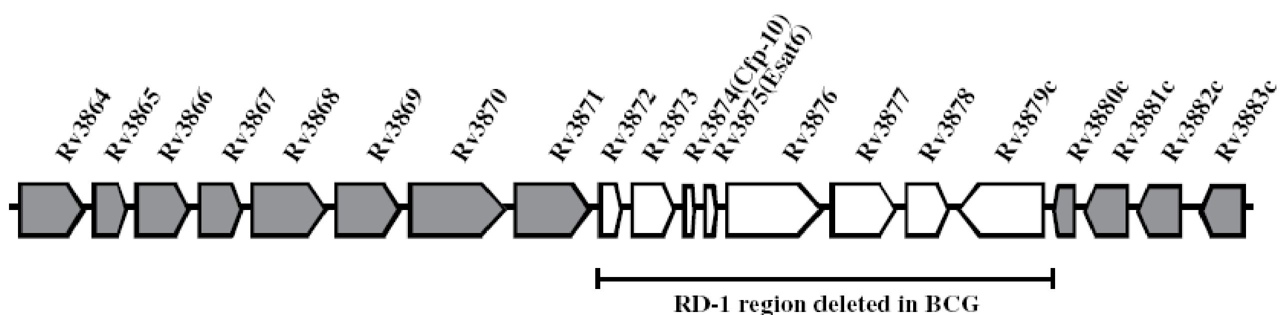


Figure 4.1 The genes present in the RD-1 along with the flanking genes are shown as an extended RD-1 (extRD-1) locus of *M. tuberculosis*. The arrows indicate the direction of transcription. The clear boxes represent genes present in RD-1 region, which is deleted consistently in all strains of BCG (Ganguly, Siddiqui, and Sharma 2008b).

The inoculation of mice with complemented *M. bovis* BCG and *M. microtti* with genes encoding ESAT-6 and CFP-10 was unsuccessful in generating ESAT-6-specific T-cell responses. This was only achieved when the vaccine strains were complemented with the entire RD1 region, suggesting that flanking genes are necessary for ESAT-6 secretion (Pym et al. 2002a). Additionally, reduced lysis of pneumocytes and macrophages, as well as reduced tissue invasion have been shown by RD1 knock-outs of *M. tuberculosis* reference strains $H_{37}Rv$, Erdman and CDC1551 (Hsu et al. 2003). The lytic activity of ESAT-6 in pneumocytes has also been shown in cell culture (Kinhikar et al. 2010).

Involvement of ESAT-6, CFP-10 and surrounding genes in RD1 in virulence and immunomodulation have been described. For example, Ganguly and colleagues (Ganguly et al. 2007) showed that ESAT-6 and CFP-10 down-regulated the bacterial lipopolysaccharide (LPS)-induced reactive oxygen species (ROS) production, which is important for killing intracellular pathogens. Other examples are listed in table 4.1.

Process	**Mechanism implicated**	**Reference(s)**
Pathogenesis/ virulence	Cell lysis	(Hsu et al. 2003; Gao et al. 2004)
	Granuloma formation	(Volkman et al. 2004)
	Cytokine suppression	(Stanley et al. 2003)
	Type I Interferon (IFN) induction	(Stanley et al. 2007)
	Phagosomal maturation arrest	(Tan et al. 2006)
	Cytosolic bacterial translocation	(van der Wel N. et al. 2007)
	Interaction with Toll-like receptor	(Pathak et al. 2007)
Immunomodu -lation	Down-regulation of macrophage cell signaling	(Basu et al. 2006; Ganguly et al. 2007; Ganguly et al. 2008a)
	Down-regulation of dendritic cell and T-cell functions	(Natarajan et al. 2003; Latchumanan et al. 2005; Sinha et al. 2007)

Table 4.1 Summary of the effects of the products of extRD1 (adapted from (Ganguly, Siddiqui, and Sharma 2008b)).

Many secretion systems involved in bacterial virulence have been described (Lee et al. 2001; Finlay et al. 1997), but not in *M. tuberculosis.* The precise mechanism of ESAT-6 and CFP-10 secretion was not clear as the proteins lack a secretory signal sequence (Pallen 2002) and appeared to be sec-independent (Economou 1999). Guinn and colleagues (Guinn et al. 2004) showed that the export of ESAT-6 and CFP-10 were impaired and *M. tuberculosis* attenuated by disrupting five genes within RD1, indicating that these genes play a role in secretion.

Furthermore, a specialised Sec-independent secretion pathway, termed the snm pathway has been described (Stanley et al. 2003). Genes necessary for the secretion of ESAT-6 and CFP-10 were located outside of RD1 (termed the extended RD1 – extRD1) (Gao et al. 2004) and was named the ESAT-6 secretion system (*ESX-1*) (Brodin et al. 2004). Indeed, the individual genes of the extRD1 were investigated by Brodin and colleagues. Inactivation of *pe35* (Rv3872 – upstream of *esxA* and *esxB* and within RD1) impaired expression of CFP-10 and ESAT-6 and therefore may play a role in regulation (Brodin et al. 2006). Further evidence of the regulation of snm secretion from outside the extRD1 has been demonstrated. Two genes located in the gene cluster Rv3616c-Rv3614c, named *snm*9 and *snm*10*,* have been shown to be essential to snm secretion. Mutants of *M. tuberculosis* lacking these genes due to secrete ESAT-6 and CFP-10, and are attenuated in mice and macrophage models (MacGurn et al. 2005). The involvement of *ESX-1* was further elucidated by Champion and colleagues (Champion et al. 2006), who found that the unstructured C terminus of CFP-10 was recognised by Rv3871, which itself interacts with the membrane protein Rv3870. Mutations in this gene impaired CFP-10 secretion via *ESX-1*. Furthermore, the yeast *Saccharomyces cerevisiae* was complemented with

ESX-1 by Teutschbein and colleagues (Teutschbein et al. 2009). This group demonstrated that this sec-independent secretory system involved ATP-dependent chaperones and membrane proteins, encoded in the extRD1. They also showed that both ESAT-6 and CFP-10 exist as hetero- and homo-dimers.

The possibility of utilising these antigens in assays for the diagnosis of TB was explored shortly after the description of the antigen. ESAT-6 and CFP-10 specific immune responses were seen in mice (Brandt et al. 1996), guinea pigs (Elhay, Oettinger, and Andersen 1998; Colangeli et al. 2000) and human lymphocytes (Lalvani et al. 1998; Ulrichs et al. 1998; Mustafa et al. 1998; Ravn et al. 1999).

The tuberculin skin test (TST) is used to measure an individual's exposure to *M. tuberculosis* and therefore identify latent infection. Purified protein derivative (PPD) is inoculated subdermally and the scale of the delayed-type hypersensitivity response, indicating long-term T cell memory, is measured in the size of the induration generated. PPD is a culture filtrate of tubercle bacilli (Seibert and Dufour 1948) that contains over 200 antigens that are shared with *M. bovis* BCG and most non-tuberculous mycobacteria (Huebner, Schein, and Bass, Jr. 1993). Therefore, interpretation of this test is hampered due to exposure to environmental mycobacteria, previous BCG vaccination (Tissot et al. 2005; Wang et al. 2002) and in the immunocompromised. The development of assays measuring the immune memory response to specific *M. tuberculosis* antigens is an advance on this assay.

The potential of the use of these antigens in diagnostic assays was further explored. The enumeration of T cells that secreted IFN-γ in response to ESAT-6 (with or without CFP-10) stimulation was achieved by enzyme linked immunospot (ELISPOT) assay (Pathan et al. 2001; Chapman et al. 2002, Lalvani et al. 2001a; Lalvani et al. 2001b). ESAT-6 specific T cells were detected in patients with culture-confirmed pulmonary TB, presumed previous *M. tuberculosis* infection, TB lymphadenitis and household contacts but not in unexposed controls (Pathan et al.2001). No cross reactivity was seen in unexposed controls who had been BCG vaccinated. This group then used this assay to measure the extent of latent *M. tuberculosis* infection in 100 healthy adults in Mumbai, India compared to 40 BCG-vaccinated UK residents. Of the Mumbai cohort, 80% responded to at least one of ESAT-6 and CFP-10, whilst no UK patients did (Lalvani et al. 2001a). This method was also shown to be useful in HIV-infected individuals. The TST suffers from poor sensitivity and specificity in HIV patients (Huebner, Schein, and Bass, Jr. 1993) but Chapman and colleagues demonstrated that ESAT-6 or CFP-10 specific T cells were detectable in this patient group (Chapman et al. 2002).

ESAT-6, together with CFP-10 is used a stimulatory antigen in these assays. Both assays utilise the antigens ESAT-6 and CFP-10 but the QuantiFERON-TB Gold in-tube assay (QFT-IT) additionally uses the antigen TB7.7, encoded in RD11 (section 1.5.1).

A recent meta-analysis of these assays showed that they are both a huge improvement on the conventional TST (Diel, Loddenkemper, and Nienhaus 2010). The analysis investigated the role of these assays in the detection of active TB as a surrogate for latent

TB infection, as there is no established gold-standard for this. The sensitivity and specificity of TST, QFT-IT and T-SPOT.*TB* are summarised in table 4.2. Additionally, the meta analysis showed low rates of indeterminate results for both IGRA. These rates were considerably higher in immunocomprimised patients (Diel, Loddenkemper, and Nienhaus 2010).

	Sensitivity % (95% CI)	Specificity % (95% CI)
TST	69.9% (67 – 72)	Not measured
QFT-IT	81.0% (78 – 83)	99.2% (98 – 100)
T-SPOT.*TB*	87.5% (85 – 90)	86.3% (81 – 90)

Table 4.2 Summary of the sensitivity and specificity of commercial IGRA and TST in the diagnosis of active TB (Diel, Loddenkemper, and Nienhaus 2010).

The analysis of IGRA in children was less favourable (Bamford et al. 2010). The use of a single IGRA alone was no better at diagnosing active TB than TST. The analysis found that a negative IGRA did not rule out TB, but that the combined use of IGRA and TST increased the sensitivity to over 90% regardless of which IGRA was used.

The use of these assays is recommended in the diagnosis of latent TB infection by NICE. Alongside the tuberculin skin test, or Mantoux, and chest x-ray these assays are used when contract tracing individuals who have been exposed to a confirmed case (NICE 2006).

IGRAs measure the IFN-γ response from circulating lymphocytes. Several studies have investigated the response to cells obtained from the lung, the primary site of *M. tuberculosis* infection (Barry et al. 2003; Wilkinson et al. 2005; Breen et al. 2005). Breen and colleagues showed that IFN-γ responses were demonstrable in such cells from TB patients. However, data obtained by our group has demonstrated that not all tuberculosis patients give a positive IFN-γ response, despite being later confirmed to have active disease by conventional methods (Breen et al. 2005; Breen et al. 2008a). Immunological and clinical data in this chapter was obtained from Dr. Ronan Breen. All other practical work was performed by the author.

The hypothesis of this chapter is that the biology of the infecting organism has an effect on the immunological response of the patients' T cells. One aspect of this is the gene expression of the antigen in question, *esx*A. Gene expression may be measured in a number of ways. As well as allowing the detection of multiple targets, microarrays may be used to measure gene expression see Miller and Tang for review (Miller and Tang 2009). Gene expression in *M. tuberculosis* in response to varying environmental conditions and exposure to antibiotics has been monitored (Kaushal et al. 2002; Betts et al. 2003; Voskuil et al. 2003; Rengarajan et al. 2005). Oligonucleotide arrays have been used to investigate gene expression related to drug resistance (Wilson et al. 1999; Keller et al. 2004; Fu and Shinnick 2007).

Microarrays are incredibly useful for screening for up and down-regulation across an entire genome but they are expensive and can be technically difficult to perform. Additionally, any findings must be confirmed by another method, such as reverse

transcriptase PCR or Western blotting. A more targeted method than microarrays to investigate the expression of individual genes is the use of reverse transcriptase polymerase chain reaction (RT PCR). Gene expression (determined by the detection of mRNA) may be measured absolutely or relatively. Absolute measurement involves using quantitative PCR (qPCR) and the establishment of a standard curve (Schmittgen and Zakrajsek 2000).

The alternative is to measure the relative change in gene expression compared to a reference gene. This may be performed using the ΔΔCt method (Livak and Schmittgen 2001). This method measures the relative change in gene expression in the target gene compared to not only a reference gene, but also to a reference strain. The advantage of the relative quantification is that it is not necessary to quantify the RNA entering the assay.

This reference gene must be selected carefully and should ideally be expressed at a constant rate by all strains. For this study, the housekeeping genes *sig*A and 16S rRNA were chosen as reference genes. *sigA* encodes the principle sigma factor in *M. tuberculosis*, which is a regulatory transcription factor (Gomez, Chen, and Bishai 1997). 16S rRNA encodes the small subunit of the bacterial ribosome.

4.2 Results

4.2.1 IFN-γ release assay

Colleagues in the respiratory and immunology departments of the Royal Free Hospital collected respiratory samples from 17 suspected pulmonary TB patients following ethics approval. IFN-γ responses were measured by flow cytometry and a percentage change compared to a negative control was calculated (Breen et al. 2005).

PPD-specific, IFN-γ producing CD4+ T cells were present in the respiratory samples obtained from all patients. The proportion of CD4+ T cells producing IFN- γ in response to PPD exposure range from 2.69% to 67.11% and are shown in table 4.3. However, analysis of CD4+ T cells exposed to ESAT-6 showed that three patient samples had no IFN-γ producing CD4+ cells. The full range of responses is listed in table 4.3 as well as the ratio of ESAT-6 to PPD responding cells. Samples were also processed for microscopy and mycobacterial culture.

4.2.2 IS*6110* RFLP genotyping

The IS*6110* RFLP profile was available for 14 of the 17 strains studied. Analysis of these profiles demonstrated that none of the strains were closely related - all 14 strains were distinct and none were related at a similarity of greater or equal to 70% (figure 4.2). In addition, none of these strains were members of any of the large lineages described in chapter 3.

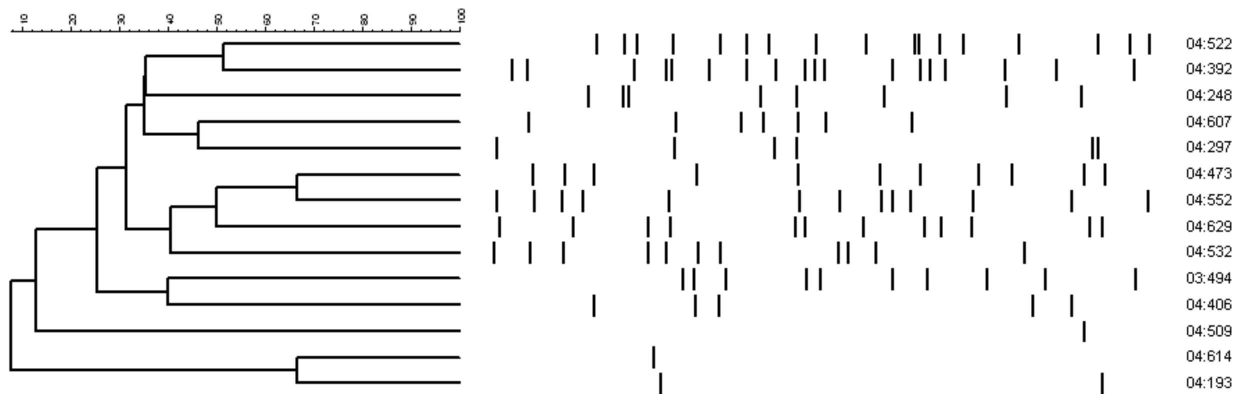


Figure 4.2 Dendrogram showing the relatedness of the patient isolates. The level of relatedness is displayed on the left-hand side. The fingerprints are shown in the centre and the isolate numbers are on the right.

4.2.3 *esx*A PCR and sequencing

The gene encoding ESAT-6, *esxA*, was detected in all strains of *M. tuberculosis* isolated from the 17 patients in the study. Sequence analysis showed no mutations in *esxA* in any of the *M. tuberculosis* strains. A summary of the immunological, PCR and sequencing data is shown in table 4.3.

Patient isolate	**CD4+ IFN- γ response to ESAT-6 (%)**	**CD4+ IFN- γ response to PPD (%)**	**ESAT-6:PPD response ratio**	***esxA* PCR**	**esxA sequence**
04:614	0.00	16.65	0	Positive	Wild Type
04:607	0.00	13.94	0	Positive	Wild Type
04:552	0.00	3.85	0	Positive	Wild Type
04:406	0.03	8.97	0.003	Positive	Wild Type
04:539	0.74	25.58	0.029	Positive	Wild Type
03:497	4.08	67.11	0.061	Positive	Wild Type
04:629	1.90	20.35	0.093	Positive	Wild Type
04:193	6.35	67.11	0.104	Positive	Wild Type
04:473	3.67	31.09	0.118	Positive	Wild Type
04:545	4.37	35.02	0.125	Positive	Wild Type
04:525	10.06	79.32	0.127	Positive	Wild Type
05:218	0.92	6.45	0.143	Positive	Wild Type
04:279	4.22	28.41	0.149	Positive	Wild Type
04:248	3.93	24.40	0.161	Positive	Wild Type
04:393	1.02	4.94	0.206	Positive	Wild Type
03:494	0.68	2.69	0.253	Positive	Wild Type
05:189	12.39	48.00	0.258	Positive	Wild Type

Table 4.3 Summary of INF-γ responses, PCR results and sequencing data. WT = wild type. Immunological data from (Breen et al.)

4.2.4 Reverse transcriptase quantitative polymerase chain reaction (RT qPCR)

A reverse transcriptase quantitative polymerase chain reaction (RT qPCR) was designed to measure the expression of *esxA*. Reference genes were chosen to be *sig*A and 16S rRNA (Chapter 2, section 2.3.3).

Although ESAT-6 was first identified in the filtrate of short-term cultures of *M. tuberculosis* (Sorensen et al. 1995), it was not known during which growth phase the gene was optimally expressed. An experiment was designed to measure *esxA* expression in *M. tuberculosis* $H_{37}Rv$ during lag, log and stationary phases of growth.

Cultures were incubated (in triplicate) for the optimal time (as ascertained from the time-trial experiment). RNA was then extracted, RT qPCR was performed and relative gene expression was calculated using the ΔΔCt method (Livak and Schmittgen 2001).

4.2.4.1 Establishing a reference gene

The use of the 16S rRNA gene as a reference gene generated highly variable results between biological replicates. Work carried out by our group showed that using *sig*A as a reference gene provided more reproducible data (O'Sullivan 2007). The 16S rRNA gene was therefore substituted with *sigA* and these experiments produced more reliable data so this gene was selected as the reference gene for all subsequent analysis. One explanation of highly variable results was the consistently high level of 16S rRNA gene expression which resulted in the inability of the assay to demonstrate small differences in expression. This is demonstrated in three example strains in table 4.4.

	Mean *esx*A CT	**Mean 16S CT**	**Mean *sig*A CT**
04:552	28.55	11.09	21.28
04:607	28.30	11.45	20.64
04:614	30.16	11.61	21.29

Table 4.4 comparison of mean CT values for *esx*A, 16S and *sig*A for three patient isolates. Lower CT values indicate the presence of larger quantities of mRNA and therefore higher expression.

4.2.4.2 *esxA* Expression Time Trial

The RT qPCR data was analysed from the time trial experiment using *M. tuberculosis* $H_{37}Rv$ to ascertain the time of optimal translation of *esxA*. The amplification of *esxA* mRNA from each time point was compared to the reference gene at the same time point. This was then compared to the level of gene expression relative to the cultures at day zero. The relative fold change in gene expression was calculated using the ΔΔCt method (Livak and Schmittgen 2001). Cultures at day 28 (stationary phase) demonstrated the largest relative increase in *esxA* expression. For this reason, day 28 (stationary phase) was selected as the time point for RNA extraction for all of the clinical strains. Relative gene expression for days 4, 14 and 28 are shown in table 4.5.

	Relative fold change ($2^{-\Delta\Delta Ct}$)
Day 0 vs. 7	5.21
Day 0 vs. 14	4.46
Day 0 vs. 28	9.98

Table 4.5 Change in *esxA* expression over time.

4.2.4.3 *esx*A Expression by Q-RT-PCR

RT qPCR data was available for 11 strains. The relative fold change of expression of *esxA* compared to the reference gene *sigA* and the reference strain *M. tuberculosis* $H_{37}Rv$ at day 28 varied from 1.10 to 109467.16-fold. The data for all 11 isolates are shown in table 4.6.

Strain	Relative fold change ($2^{-\Delta\Delta Ct}$)	ESAT6:PPD IFN-γ response
04:552	3.24	0
04:607	1.98	0
04:614	1.28	0
04:539	2.04	0.029
03:497	8.58	0.061
04:629	3.46	0.093
04:545	6.31	0.125
04:525	3.24	0.127
04:279	4.86	0.149
04:248	1.10	0.161
04:391	109467.16	0.206

Table 4.6. Fold change of *esxA* expression in each clinical change relative to *M. tuberculosis* $H_{37}Rv$ compared to ESAT/PPD IFN-γ response.

The correlation between gene expression and ESAT-6:PPD response was plotted using regression analysis and is shown in figure 4.3 and demonstrates a non-linear relationship between the relative expression of *esx*A and ESAT6/PPD response, and therefore, no direct association between the two.

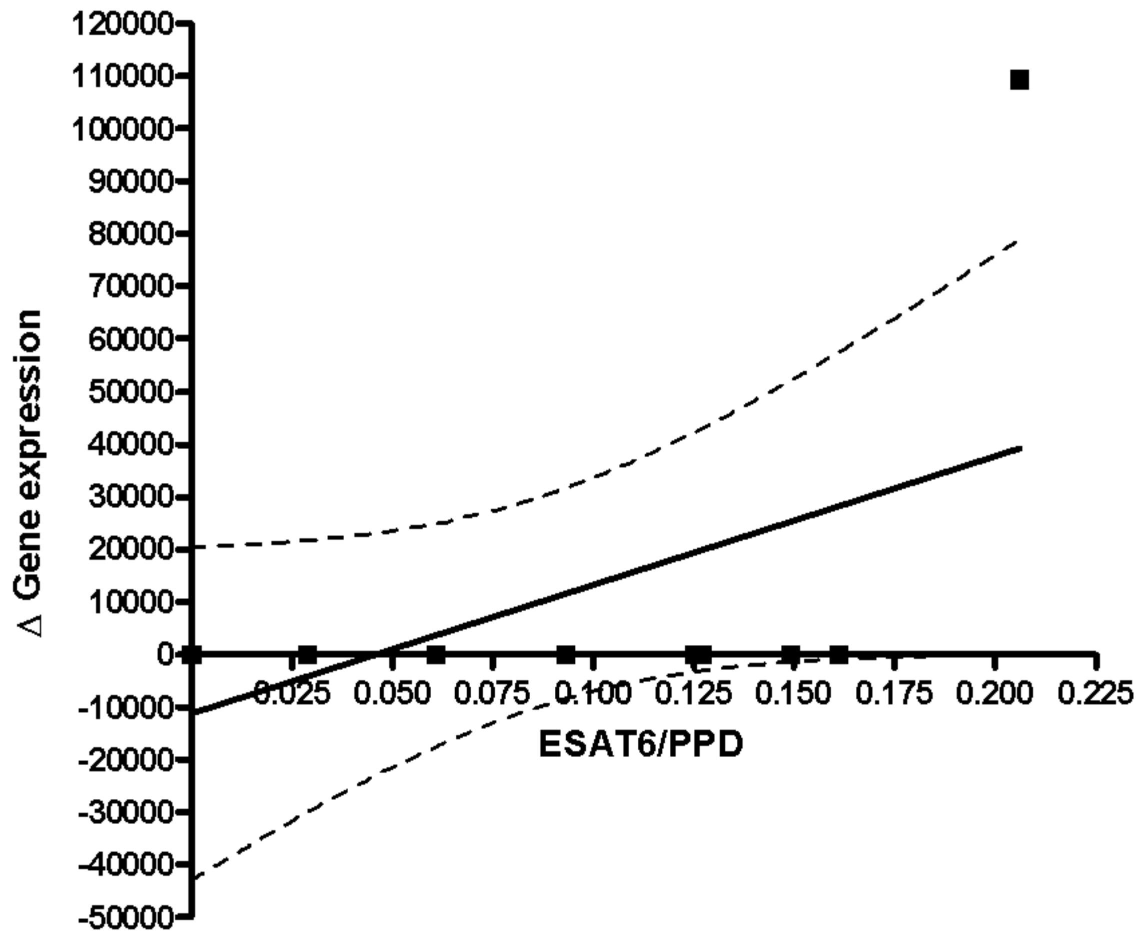


Figure 4.3 Regression analysis plotting relative change in *esxA* expression and ESAT-6:PPD response.

4.3 Discussion

Some strains produced no IFN-γ at all following exposure to ESAT-6. The CD4+ cells did, however, respond to PPD stimulation, indicting that they were capable of mounting a cytokine response. These patients were all shown subsequently to have culture-confirmed TB, so the possibility that this phenomenon was related to the biology of the infecting organism was examined. The genotyping database was interrogated to determine whether this sub-set of patients were infected with a related strain, however the typing

data showed that the differences in IFN-γ responses (in these non-responding cases as well as more widely) were not due to a specific strain type. Those that failed to generate a response following ESAT-6 stimulation were not infected with related strains and there were no relationships between any of the isolates at a level of 70% or greater (figure 5.2). The strains were also shown not to be members of any of the six large lineages discussed in chapters 3 and 4.

The presence of *esxA* was proven in all isolates, regardless of IFN-γ response. Likewise, no mutations were discovered, leading to the conclusion that the failure of patients' CD4+ cells to produce IFN-γ in response to ESAT-6 stimulation is not due to deletion or mutation of the gene causing a modified or inactive protein.

The time-trial experiment was performed to ascertain the time of optimal *esxA* expression in *M. tuberculosis* $H_{37}Rv$. Results indicated that this occurred at 28 days (stationary phase). Stationary phase cultures (day 28) were sampled due to the data obtained from the time trial experiment. No published data was available outlining the time of optimum expression of *esx*A, so this experiment was performed to provide the best opportunity to demonstrate differences between strains.

The exact physiological state of *M. tuberculosis* in the lung is unclear. Therefore, the relevance of stationary state cultures to what is seen *in vivo* is equally uncertain. Indeed, there are two models for persistent bacilli in the lung; slowly replicating organisms that interact with components of the immune system (Ehlers 2009) and non-replicating

organisms that are reawakened with resuscitation promoting factors (Garton et al. 2008). The physiological state of these organisms is not fully understood and therefore the relevance of *in vitro* models is subject to discussion. However, stationary phase cultures were sampled in this work as expression of *esx*A in these isolates was greatest at this time.

Using *sig*A rather than 16S rRNA as a reference gene provided more reproducible data. This may be partly due to the high level of expression of 16S which results in the inability of the assay to demonstrate small differences in expression. The use of *sig*A, which is expressed at a lower level (table 5.4) allows these variations to be seen.

The gene expression data shows that all strains demonstrated an up-regulation of *esxA* when compared to *M. tuberculosis* H_{37}Rv. This may be because ESAT-6 is an immunomodulator and virulence factor that is not required for survival in the laboratory-adapted strain *M. tuberculosis* H_{37}Rv. One isolate showed very high-levels of up-regulation of *esxA* compared to *M. tuberculosis* H_{37}Rv. The expression of *esxA* was increased over 100,000-fold in this isolate, which was cultured from a TB patient whose CD4+ T cells demonstrated the largest ESAT-6:PPD IFN-γ response (0.206 in isolate 04:391, table 5.6; figure 5.3). However, the IFN-γ response in this patient's cells was not statistically significantly higher than the next largest responder (04:248), whose isolate had an *esxA* expression comparable to that of *M. tuberculosis* H_{37}Rv. The increase in gene expression in this isolate was remarkable. This was seen in two biological, and three technical, replicates. It may be possible that the growth rate of this strain is significantly

different to the others, that is, it may have reached stationary phase significantly earlier than day 28. The growth rates of the isolates were not measured, however. This may be necessary to ensure that all of the experiments were performed during stationary phase to allow comparisons between the strains. This warrants further investigation of the expression of other genes associated with this isolate to determine if this level of up-regulation is a common factor. This may be achieved by using RT PCR to look at a range of genes in this isolate, or by utilising a microarray. To investigate whether 04:391 has generally higher levels of expression when compared to its peers, it may be beneficial to investigate housekeeping gene expression whilst in steady state. Additionally, the expression of these genes may also be compared during times of stress, possibly in a chemostat model (Hampshire et al 2004). Microarrays have the benefit of being able to analyse multiple genes in a single assay. However, in the work described, a single RT PCR was designed as a small number of genes were being targeted.

The summation of these data leads us to the conclusion that the failure of an immunological test to diagnose *M. tuberculosis* infection is not a consequence of the modified genotype of the organism. However, there may be post-transcriptional, translational or post-translational factors that may have an effect on the secreted protein, and therefore the T-cell response. The nature of the excreted may be assessed by protein purification and Western blotting.

These findings confirm that the relationship between host and organism in *M. tuberculosis* infection is a complex one. We are led to the conclusion that the failure of

the patients' CD4+ T cells to respond to ESAT-6 stimulation may be due to a function of the hosts' immune response with little contributory effect of the bacterium. It has been shown previously in animal models that an increased susceptibility to disease is seen when there are defects in CD4 T cell function (Saunders et al. 2002; Scanga et al. 2000), or with IFN-γ expression or recognition (Cooper et al. 1993; Flynn et al. 1993). Increased susceptibility has also been seen in such patients (Havlir and Barnes 1999; Ottenhoff et al. 2005). The relationship between immunological response and clinical presentation could be further investigated by the analysis of the patients' HLA types.

Chapter 5: The Relative Fitness of Over-Represented Strains and Acquired Isoniazid Resistance

5.1 Introduction

The measure of the success of a bacterium may be considered by the extent to which it is able to survive, divide and spread between susceptible hosts. In order to achieve this, it must be assumed that the rate of division is paramount in the organism's success. The time taken for a bacterium to divide, or generation time, may therefore be taken as a measure of fitness. The fitness of an organism may be measured in a number of ways; for example, the measurement of growth rates *in vitro*, either in appropriate media (Billington, McHugh, & Gillespie 1999; Bjorkman et al. 1999; Gagneux et al. 2006b) including chemostat (Modi & Adams 1991), in competition assays (Bouma & Lenski 1988; Schrag, Perrot, & Levin 1997) or in animal models (with or without competition) (Bjorkman, Hughes, & Andersson 1998; Heym et al. 1997; Li et al. 1998; Wilson, de Lisle, & Collins 1995). Fitness may also be inferred by the measurement of the size of clusters (van Soolingen et al 2000; van Doorn et al 2006). Conversely, large clusters may not always indicate a strain with enhanced fitness. For example, outbreaks in prisons often occur due to the prolonged exposure of susceptible individuals to infective cases (Sosa et al 2008).

Decreased virulence of isoniazid resistant *M. tuberculosis* in guinea pigs was reported in the 1950s, with the organisms displaying the highest levels of resistance exhibiting the largest loss in virulence (Middlebrook & Cohn 1953; Mitchison DA 1954) .There is

assumed wisdom that the acquisition of antimicrobial resistance will result in a biological cost (Gillespie & McHugh 1997). It is now known that the correlation between acquired resistance mutations and fitness cost is far from linear.

It is true that although a fitness cost is likely as a result of the acquisition of antimicrobial resistance, there are a number of different compensatory mechanisms that enable the organism to persist. These mechanisms often allow the retention of antimicrobial resistance as well as the reversion to fitness to some or all of the degree of the sensitive ancestor organism (Andersson & Levin 1999). Although there are several documented compensatory mechanisms for other bacteria (Moore, Rozen, & Lenski 2000; Reynolds 2000; Maisnier-Patin & Andersson 2004), the only documented compensatory mechanism that has been proposed in *M. tuberculosis* is the hyper-expression of the alkyl hyperperoxide reductase gene, *ahp*C, in isoniazid resistant isolates with *kat*G mutations (Sherman et al. 1996). However, this mutation in the promoter region of *ahp*C has been seen in isolates that do not possess *kat*G mutations, so this mechanism is far from confirmed (Gagneux et al. 2006a).

The degree of fitness cost suffered due to the acquisition of resistance-conferring mutations may be ascertained by *in vitro* or *in vivo* studies, or by the measurement of the prevalence of these mutants in the community. If particular resistance-conferring mutations are common in the community then it may be inferred that the strains are still successful and have not suffered a significant loss. Indeed, studies on strains carrying the most common resistance mutations have confirmed this. Sander and colleagues

demonstrated that the when the most commonly occurring mutations causing streptomycin, amikacin and clarithromycin resistance were induced in *M. smegmatis*, these strains suffered little of no fitness cost in *in vitro* competition assays (Sander et al. 2002). Likewise, the most commonly seen mutations in *rps*L causing streptomycin resistance in clinical isolates of *M. tuberculosis* are the same as those that have been shown experimentally to exhibit no fitness cost in *S. typhimurium* and *E. coli* (Bottger et al. 1998). Additionally, the mutation causing most clinically seen isoniazid resistance (*kat*G Ser-315-Thr) is not associated with any fitness cost in mouse models (Pym, Saint-Joanis, & Cole 2002). Furthermore, the most commonly occurring *rpo*B mutations conferring rifampicin resistance in *M. tuberculosis* are associated with a smaller fitness cost when compared to less frequently occurring mutations (Billington, McHugh, & Gillespie 1999; Mariam et al. 2004: Gagneux et al. 2006b; Huitric et al. 2006;).

If a resistant strain is prevalent in the community, and is therefore said to have suffered a minimal fitness cost, then it stands to reason that the strain must be transmitting successfully between individuals. For this reason, not only is the degree of prevalence of a particular mutation important, but also the size and frequency of clusters of cases associated with the mutated strains. In a Dutch cohort, strains with the most frequently occurring isoniazid-conferring mutation (mutation at *kat*G codon 315) were found in clusters as often as isoniazid susceptible strains (van Soolingen et al. 2000), whilst isoniazid resistant strains with other mutations were less likely to be found in clusters than those with mutations at codon 315 (van Doorn et al. 2006). A study in San Francisco found that isoniazid resistant strains of *M. tuberculosis* with the *kat*G mutation Ser-315-

Thr or a mutation in the *inh*A promoter region were more likely to be found in clusters (Gagneux et al.2006a).

Outbreaks of drug resistant strains are well documented. For example, the outbreak of an MDR strain in New York (Frieden et al. 1996) may be attributable to coincidence or the increased susceptibility of the immunocompromised hosts rather than properties of the organism. Likewise, an MDR TB outbreak in a London hospital was more likely to be due to improper airflows in a patient's room than the increased fitness of the strain itself (Breathnach et al. 1998).

Continued transmission in an ongoing outbreak of isoniazid resistant *M. tuberculosis* in London has been reported (HPA 2009). The first 100 cases in this outbreak were investigated to ascertain the number of secondary cases of active and latent disease (Neely et al. 2010). Analysis revealed that a large number of contacts (11%) were diagnosed with active disease whilst 13% of contacts were said to have latent TB infection. These rates were considerably higher than those previously described. NICE guidance states that the reviewed literature (Ansari et al. 1998; Esmonde & Petheram 1991; Hardinge, Black, & Chamberlain 1999; Hussain et al. 1992; Irish, Jolly, & Baker 1997; Kumar, Innes, & Skinner 1992; Ormerod 1992; Teale, Cundall, & Pearson 1991) indicates that approximately 1% of contacts progress to active disease (NICE 2006).

The phylogeographic background of the bacterium may also play a role in the fitness and drug resistance of *M. tuberculosis*. Differences in immunogenicity and virulence have

been reported and are reviewed (Gagneux & Small 2007). The Beijing family of strains is often cited as being particularly associated with drug resistance. A comparison of twelve studies in South East Asia and countries of the former Soviet Union show a high proportion of MDR strains. These countries have poor TB control strategies, which undoubtedly contribute to the number of cases of MDR-TB, however, the high prevalence of strains belonging to the Beijing lineage may also be a factor (Borrell & Gagneux 2009).

Fitness of drug-resistant strains can be measured in laboratory studies or be shown to be successful by their prevalence in the community. It is useful to quantify this when considering whether acquired drug resistance has caused a fitness cost on a group of strains. If the majority of a group of related strains is drug susceptible, then it may be considered that the acquisition of drug resistance is either relatively recent, or is associated with decreased transmissibility. Conversely, if a group of related strains are largely drug resistant, then it may be proposed that the acquisition of resistance has had little of no effect on the fitness of the organism and is associated with ongoing transmissibility. This may be summarised as the relative rate of transmission (RRT) in the following formula:

$$\mathrm{RRT} = \frac{\text{no. sensitive members of group}}{\text{no. resistant members of group}}$$

Where: $\mathrm{RRT} < 1$ resistance associated with increased transmissibility
$\mathrm{RRT} > 1$ resistance associated with decreased transmissibility

Equation 5.1. Calculation of the relative rate of transmission (RRT).

The lineages identified in the Royal Free database are populated with successful, and evolutionarily-related, strains. It may be argued that these strains possess some evolutionary advantage that allows them to succeed in a particular population. The most densely populated lineages in the Royal Free database were examined to establish whether there were any factors that may contribute to the success of their strains. In addition, the strains isolated from the extended family in North London were investigated to assess whether they possessed a fitness advantage.

5.2 Results

The six most populated lineages from chapter 3 (lineages 15, 19, 41, 43, 54 and 61) were investigated. Demographic data was obtained from WinPath at the Royal Free Hospital. Drug susceptibility tests were performed at the Health Protection Agency's Mycobacterium Reference Unit.

The most outstanding characteristic was the prevalence of drug resistance in lineage 15 and the genotype of the 16 isolates was investigated further. The isoniazid resistance-causing mutation was ascertained for all isolates by PCR and sequence analysis. DNA extraction, PCR amplification and sequencing were performed by Dr. Alastair McGregor as part of his MSc degree project under the supervision of the author. Sequence analysis was performed by the author.

Additionally, the generation time of each strain was measured (chapter 2, section 2.4.3) and compared to the reference strain *M. tuberculosis* $H_{37}Rv$, both fully sensitive strains in

lineage as well as unrelated isoniazid sensitive and resistant strains. The generation times of the isolates were measured by both Dr. Alastair McGregor and the author.

5.2.1 Lineage 15

Analysis of the IS*6110* database revealed six lineages that contained 10 or more isolates (Chapter 3, table 3.14). The drug susceptibilities of these strains were investigated (table 5.1). It was seen that lineage 15 (figure 5.1) contained a higher proportion of drug resistant strains. The relative rate of transmission for each lineage was calculated as follows (table 5.1).

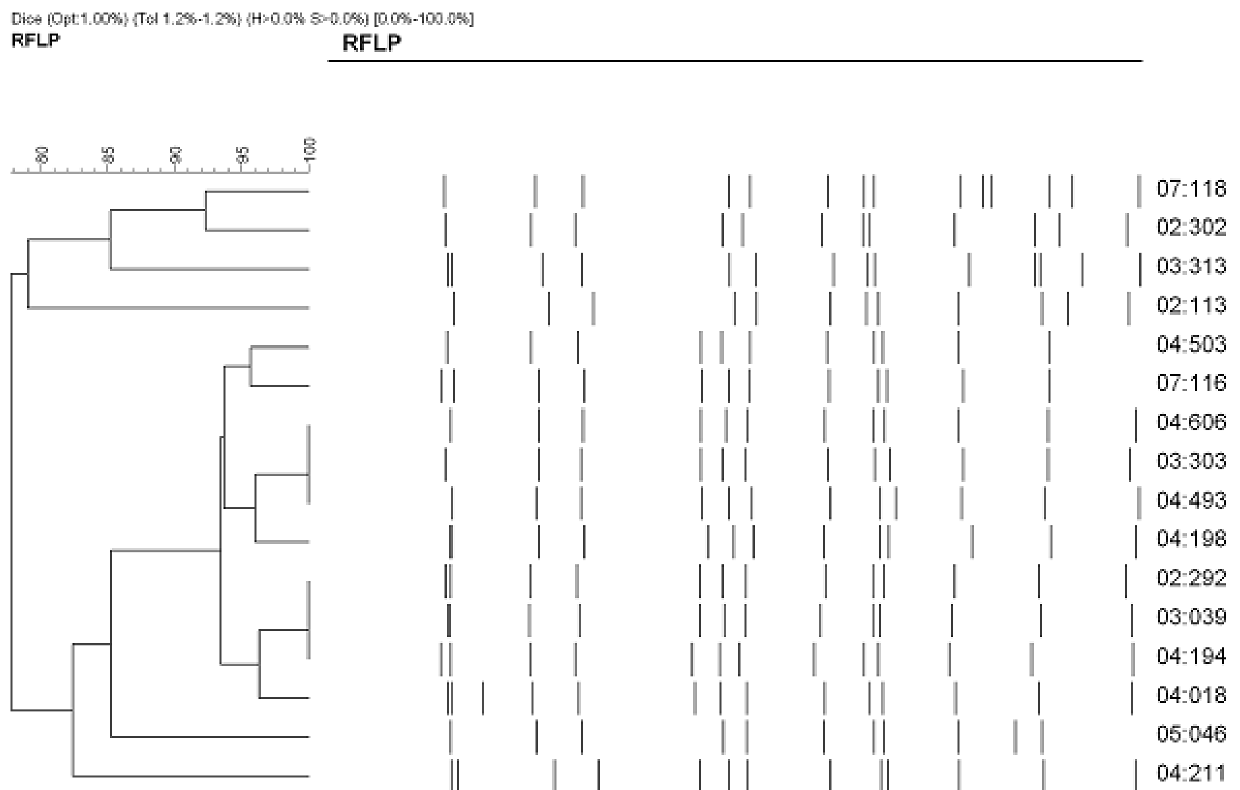


Figure 5.1 Dendrogram of IS*6110* fingerprints of strains within lineage 15.

Lineage No.	Number of isolates	% (number) resistant	Resistance profile(s)	Relative Rate of Transmission (RRT)
15	16	87.5% (14)	10 x INH mono-resistance 1 x INH and eth resistance 1 x INH, RIF, clari & eth resistance 2 x streptomycin mono-resistance	0.14
19	15	6.7% (1) U = 3	INH mono-resistance	11
41	21	9.5% (2) U = 2	INH mono-resistance RIF & INH resistance	8.50
43	10	0% (0)	N/A	N/A
54	13	0.8% (1) U = 1	INH mono-resistance	11
61	13	0.8% (1) U = 1	INH mono-resistance	12

Table 5.1 The six most populated 70% lineages in the RFH database. The lineages were assigned a number consecutively from the start of the dendrogram. Thr relative rate of transmission of each lineage has been calculated. U = unknown (sensitivities not available). INH = isoniazid, RIF = rifampicin, clari = clarithromycin, eth = ethambutol.

All isolates were referred to the MRU for susceptibility testing. Due to their phenotypic drug resistance patterns, the MRU performed genotyping and they were confirmed to be part of the North London isoniazid resistant outbreak (Ruddy et al. 2004). It was noted that the isolates were not identical by IS*6110* RFLP genotyping (performed at the Royal Free Hospital), but closely related at $\geq$ 80% similarity (figure 5.1). One strain (04:606) was not viable and no DNA was available, so no further analysis of this was possible. DNA only was available from 02:292. Amplification and sequencing of *kat*G and the promoter region of *inh*A showed heterogenous polymorphisms, which are summarised in table 5.2. Cytosine was substituted for thymine (C→T) at a locus 15bp upstream from the start codon of the *mab*A gene (which is co-transcribed with *inhA*) in eight of the ten isolates that were isoniazid resistant. The two other isoniazid resistant isolates possessed

mutations that substituted serine for thrionine at codon 315. No isoniazid-encoding mutations were found in these loci of the two sensitive isolates in the lineage. A non-resistance causing mutation resulting in the substitution of arginine with leucine at codon 463 was detected in one of the isoniazid sensitive controls (04:011).

Type	Isolate	Resistant to	Sensitivities	Polymorphism
Lineage 15	02:302	S	HRZE	None detected
	07:118	H	RZES	*kat*G S315T
	03.313	S	HRZE	None detected
	02.113		HRZES	None detected
	04.503	H	RZES	*inh*A C→T
	07.116	H, ethi	RZES	*inh*A C→T
	03.303	H	RZES	*inh*A C→T
	04.493	H	RZES	*inh*A C→T
	04:198	H	RZES	*inh*A C→T
	03.039	H	RZES	*inh*A C→T
	04.194	H	RZES	*kat*G S315T
	04.018	H,R,clari,ethi	ZES	*inh*A C→T
	05.046		HRZES	None detected
	04.211	H	RZES	*inh*A C→T
H resistant control	05.177	H	RZES	*inh*A C→T
H sensitive control	05.094		HRZES	None detected
H sensitive control	04.011		HRZES	*kat*G R463L

Table 5.2 Resistance profiles of strains in lineage RFH-15 and corresponding isoniazid resistance mutations. Key: H = isoniazid, R = rifampicin, S = streptomycin, Z = pyrazinamide, E = ethambutol, ethi = ethionamide, clari = clarithromycin. Polymorphisms: *kat*G S315T = serine to thrionine substitution at codon 315, *inh*A C→T = cytosine to threonine substitution at 15bp upstream of the start codon of *mab*A.

MIRU-VNTR typing was performed on the isolates to confirm their relatedness. All available MIRU-VNTR types are displayed in table 5.3.

	154	424	577	580	802	960	1644	1955	2059	2163	2165	2347	2401	2461	2531	2687	2996	3007	3171	3192	3690	4052	4156	4348
02:113			4			3	3	3			4	4	2	2	5	1			3	3				
02:302	2	2				3	3		1	5	4	4	2	2	5	1		3	3	3				2
03:039	2	2	4			3	3	3	1	6	4	4	2		5	1		3	3	3				2
03:303	2	2	4	2		4	3		1	6	4		2		5	1			3	3				2
03:313			4			3	3	3			4		2							3	2			
04:194	2		4	2		4	3	3	1		4		2		5	1			3	3				
04:198		2	4	2	1	4	3	3	1		4	4	2	2		1	5		3	3				
04:211	2	2				4	3	3	1		4		2		5	1			3	3				2
04:606	2	2	4	2	1	4	3	3	1		4	4	2	2	5	1	5	3	3	3				2
05:046			4	2	1								2		5		3							

Table 5.3. MIRU-VNTR genotypes of the isolates in lineage 15. Blank cells indicate that an amplicon was not generated at the locus.

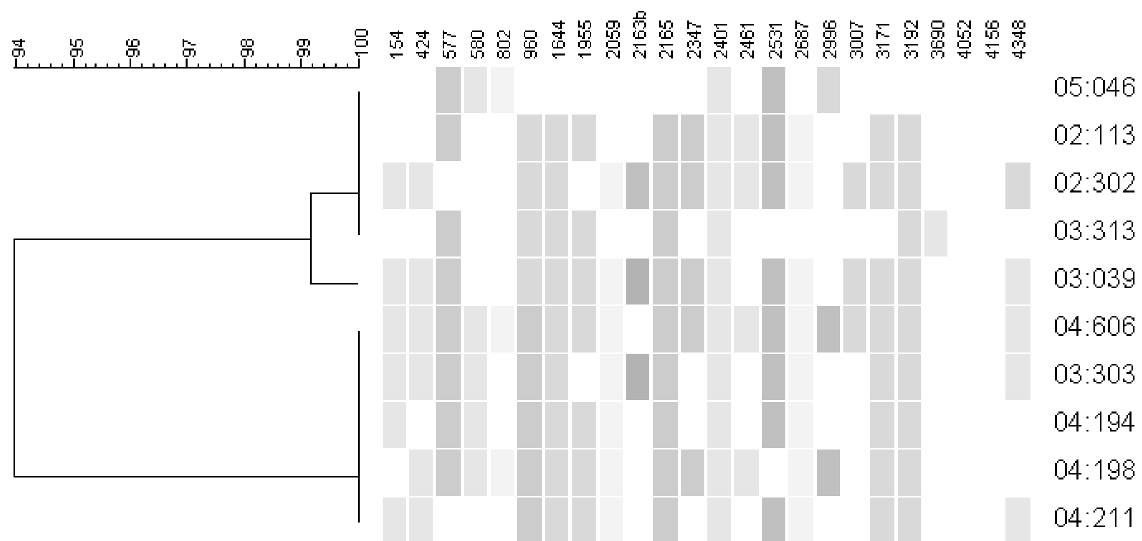


Figure 5.2. MIRU-VNTR dendrogram of lineage 15.

The generation time of all isolates and controls was measured. The median generation time was compared to that of the laboratory strain *M. tuberculosis* $H_{37}Rv$ as well as the susceptible controls.

Type	**Isolate**	**Resistant to**	**Polymorphism**	**Median generation time G (hours)**	**Relative fitness to $H_{37}Rv$**	**Relative fitness to 02:113**	**Relative fitness to 05:046**
Lineage 15 isolates	02:302	s	None detected	16.3	1.05	1.10	1.02
	07:118	h	*kat*G S315T	21.5	0.79	0.83	0.77
	03.313	s	None detected	17.5	0.98	1.02	0.95
	02.113	fully sensitive	None detected	17.8	0.96	1.00	0.94
	04.503	h	*inh*A C→T	19.1	0.89	0.93	0.87
	07.116	h, eth	*inh*A C→T	17.3	0.99	1.03	0.97
	03.303	h	*inh*A C→T	18.5	0.92	0.96	0.90
	04.493	h	*inh*A C→T	16.9	1.01	1.06	0.99
	04:198	h	*inh*A C→T	17.4	0.98	1.03	0.96
	03.039	h	*inh*A C→T	21.8	0.78	0.82	0.76
	04.194	h	*kat*G S315T	18.4	0.93	0.97	0.91
	04.018	h,r,c,eth	*inh*A C→T -767	22.9	0.75	0.78	0.73
	05.046	fully sensitive	None detected	16.7	1.03	1.07	1.00
	04.211	h	*inh*A C→T -767	18.1	0.94	0.98	0.92
H res control	05.177	h	*inh*A C→T -767	18.6	0.92	0.96	0.90
H sens control	05.094	fully sensitive	None detected	19.5	0.88	0.91	0.86
H sens control	04.011	fully sensitive	*kat*G R463L	21.2	0.81	0.84	0.79

Table 5.4 Resistance profiles of strains in lineage RFH-15, corresponding isoniazid resistance mutations and relative fitness. Key: h = isoniazid, r = rifampicin, s = streptomycin, z = pyrazinamide, e = ethambutol, eth = ethionamide, c = clarithromycin

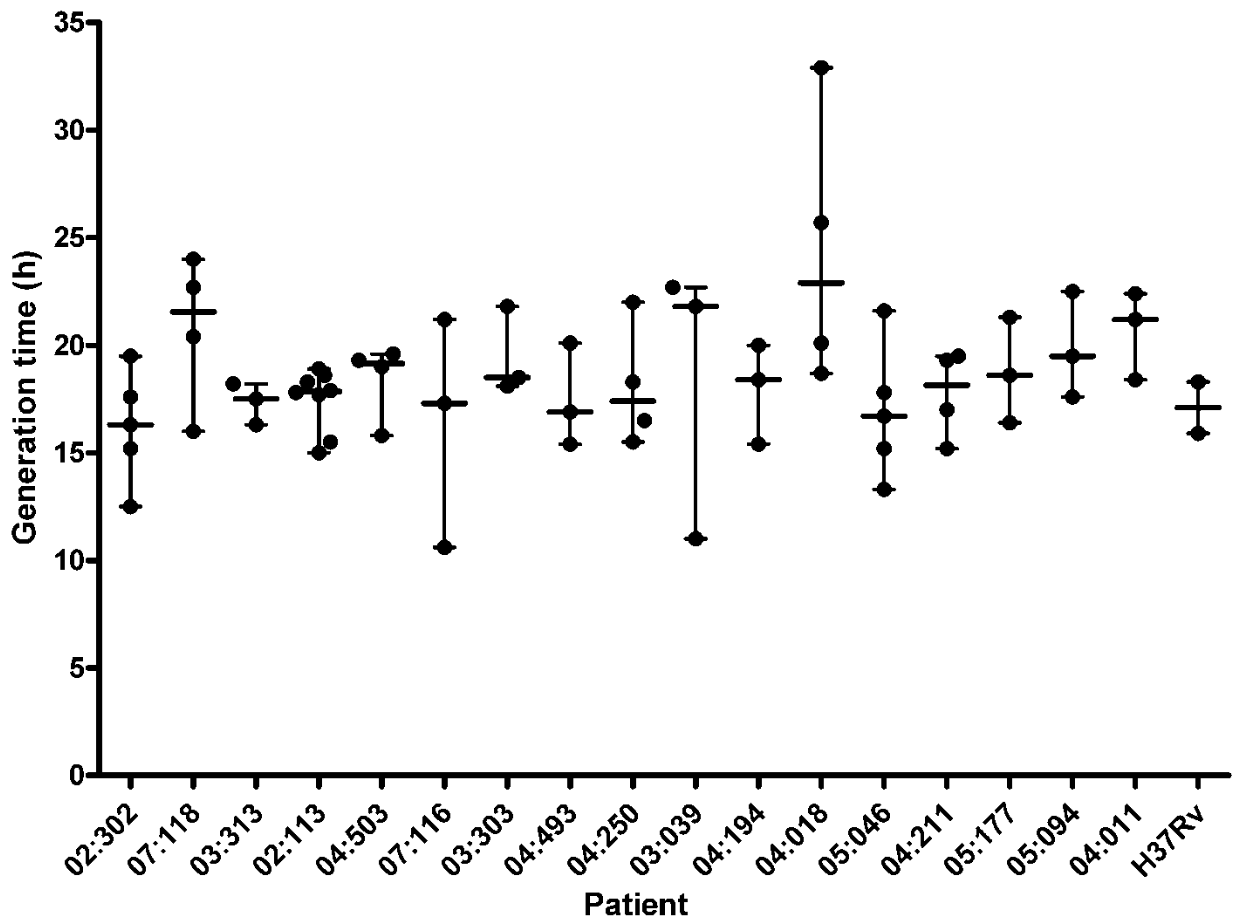


Figure 5.3. Generation times of the isolates in lineage 15. Error bars indicate one standard deviation. The median generation times of the isolates were not significantly different (p = 0.2956 - one-way Anova with Kruskal-Wallis post test).

The IS*6110* RFLP fingerprints of the strains in lineage 15 were further investigated. The fingerprints were not identical, but were closely related at $\geq$80% (figure 5.1). There are two fully susceptible isolates (05:046 and 02:113) that are hypothesised as the precursor strains, and may be used to develop a map of evolution through the alteration of IS*6110* patterns and acquired resistance-conferring mutations from that point on. A hypothetical relationship between the strains, the insertion/deletion of IS*6110* and the acquisition of antimicrobial resistance is shown in figure 5.4.

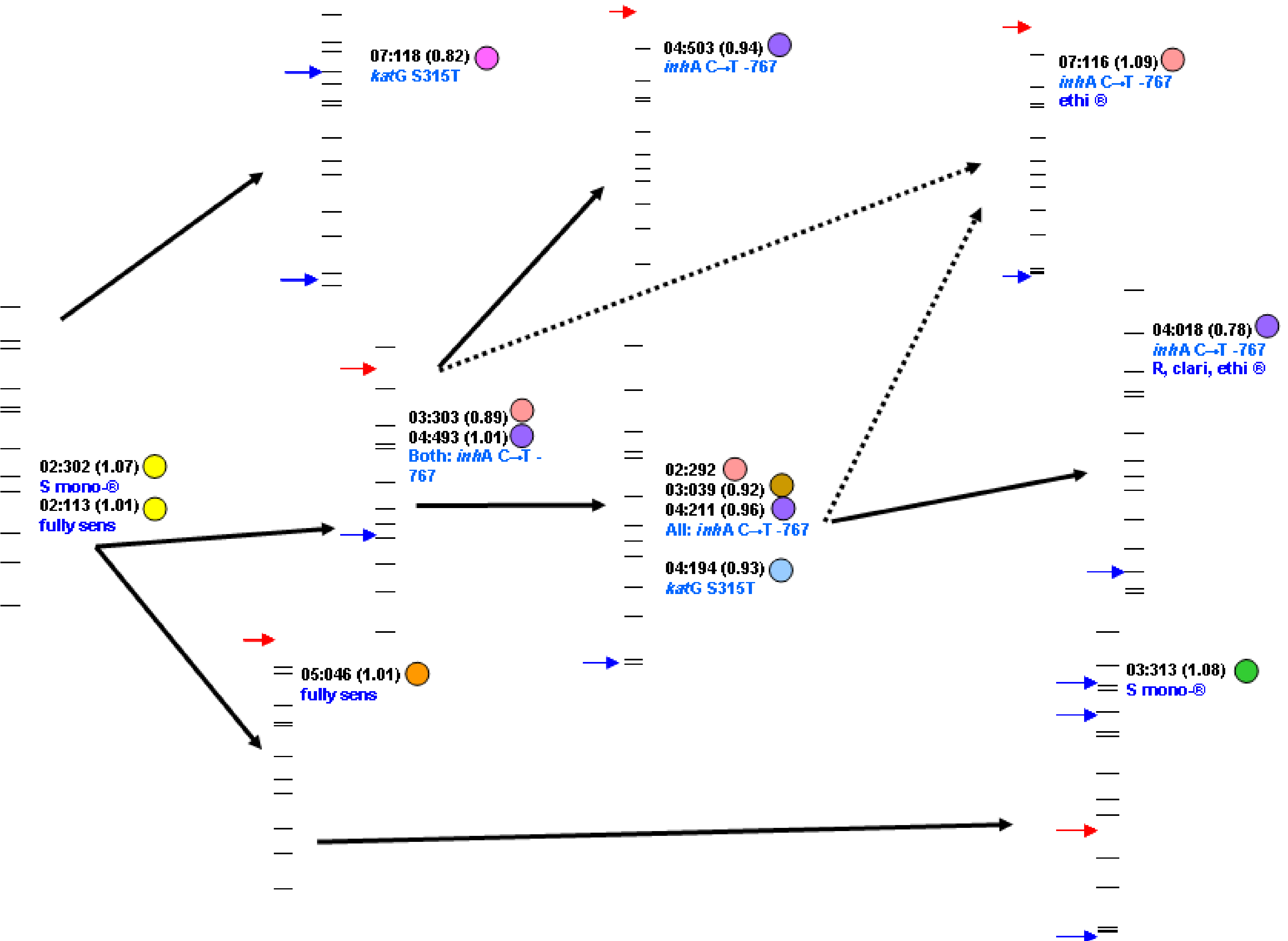


Figure 5.4 Split tree analysis - Relationship of the strains in lineage 15, including insertion (→) and deletion (→) of IS*6110* from the hypothesised progenitor strains (02:302 and 02:113). All strains are isoniazid resistant, except 05:046 & 02:113, which are fully sensitive (fully sens), and 02:302 & 03:313, which are streptomycin mono-resistant (S mono-®). The isoniazid resistance-conferring mutations are highlighted and additional resistances (®) are in blue (R = rifampicin, clari = clarithromycin, ethi = ethionamide. The first two digits in the isolate number indicate the year of isolation, for example, 02:113 was isolated in 2002. Dotted arrows indicate that a strain may have evolved from more than one possible precursor. Coloured circles relate to the MIRU-VNTR genotypes. Isolates with the same colour circle are indistinguishable by MIRU-VNTR genotyping. The relative fitness to *M. tuberculosis* $H_{37}Rv$ is in brackets.

5.2.2 Extended Family with streptomycin-resistant TB

A 26 year old Somalian man presented to the TB team at the Royal Free Hospital with cough, fever, lethargy and headache. His chest X-ray showed cavitations and nodular shadowing. A chest CT scan confirmed this and a diagnosis of active TB was made. He was commenced on standard quadruple therapy and contact tracing was initiated. Twenty members of the extended family were screened (figure 5.5). Screening was performed according to NICE guidance (NICE 2006), and involved an interview, tuberculin skin test, chest X-ray and interferon gamma release assay (IGRA). Screening revealed a further seven cases of active TB and two cases of latent TB. Including the index case, there were 8 cases of active TB and six of these patients were culture positive.

The six strains isolated from the extended family were all highly resistant to streptomycin and were confirmed to be indistinguishable from one another by MIRU typing (15 loci MIRU 42433 23315 14323, performed at HPA MRU). This outbreak occurred prior to the adoption of 24-loci MIRU-VNTR typing. This number of loci were analysed as standard by the HPA Mycobacterium Reference Unit at this time.

The hypothesis is that the strain involved in this outbreak has suffered a negligible fitness cost as a result of acquiring streptomycin resistance because more close contacts had active disease or latent infection than would normally be expected. Fitness studies were performed on those strains that were isolated from the active cases (table 5.5).

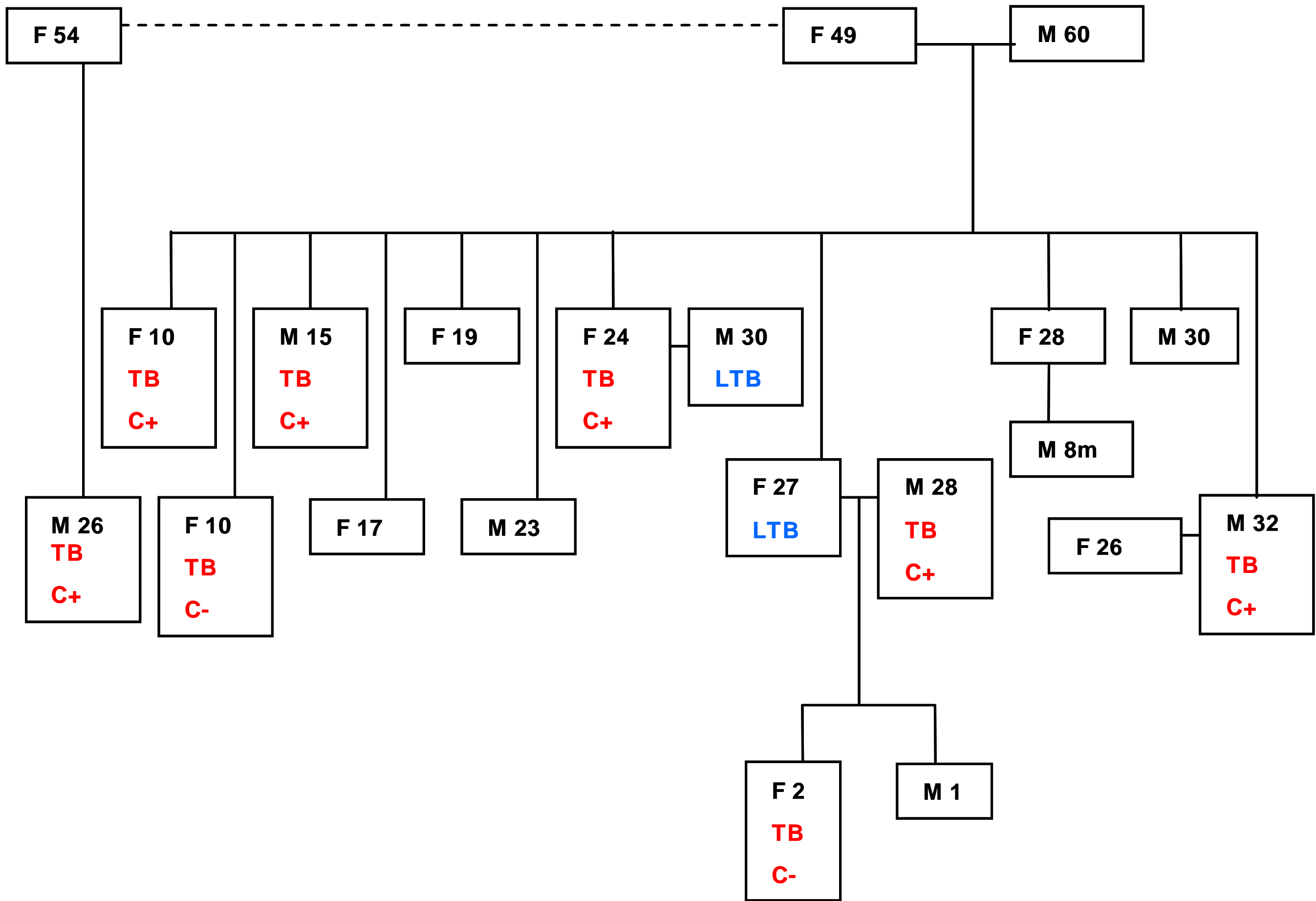


Figure 5.5 Family tree of extended family with a high rate of TB. Dotted line indicates siblings. All individuals in this family tree were contact traced. TB = active disease, LTB = latent TB, C+ = culture positive, C- = culture negative

Patient	**Median generation time (hours)**	**Relative fitness to $H_{37}Rv$**
M 26	23.3	0.73
F 10	19.2	0.89
M 15	20.3	0.84
F 24	19.4	0.88
M 28	22.5	0.76
M 32	17.3	0.99

Table 5.5 Fitness data of the strains isolated from the extended family with a high rate of active TB and latent TB infection.

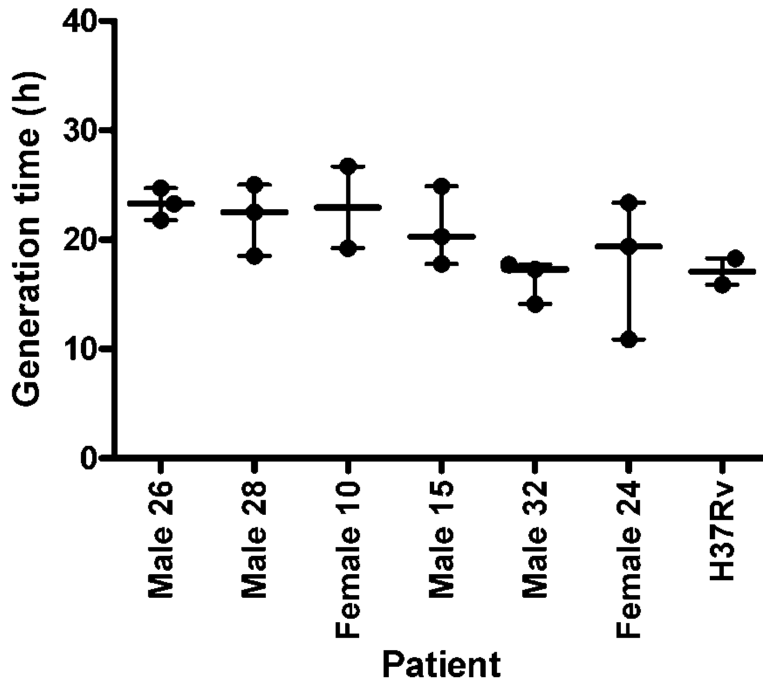


Figure 5.6. Generation times of the isolates from the extended family. Error bars indicate one standard deviation. The median generation times of the isolates were not significantly different (p = 0.1683 - one-way Anova with Kruskal-Wallis post test).

5.3 Discussion

The aim of this work was to investigate a large drug resistant lineage of *M. tuberculosis* and a family-based outbreak of streptomycin resistant *M. tuberculosis*. The hypothesis was that the acquisition of drug resistance confers a fitness deficit and that this may be related to genotype.

5.3.1 Lineage 15

The strains in lineage 15 are closely related, as indicated by their $\geq$80% similarity by IS*6110* fingerprinting. However, although these strains are shown to cluster together with the North London isoniazid resistant outbreak strains, the data shown in this chapter demonstrates that the strains are far from homogeneous. The outbreak strains have evolved as they have been transmitted between individuals, and are no longer identical to the strain at the point source of the outbreak. The isolates were cultured from patients between 2002 and 2007, so there was ample opportunity for the acquisition of drug resistance as well as insertion or deletion is IS*6110* given its half life of two to five years (de Boer et al. 1999).

Firstly, the antibiograms of the strains vary widely and include fully sensitive, isoniazid and streptomycin monoresistant as well as MDR strains. This is concordant with the reports of MDR cases in the North London outbreak (Maguire H, Forrester, & Adam 2006). Furthermore, the point mutations causing isoniazid resistance vary even within clusters of strains that are indistinguishable by IS*6110* RFLP and MIRU-VNTR typing. It may be reasoned that the strain was already successfully circulating in the North London area and the acquisition of antimicrobial resistance occurred secondarily. It has already

been demonstrated that strains in this outbreak have become multi-drug resistant by the independent acquisition of unusual *rpo*B mutations (Jenkins et al. 2005).

The outbreak strains studied in this chapter were isolated between 2002 and 2007, and there are a total of eight insertions and seven deletions of IS*6110* during this time. It appears that the strains have gained and lost the insertion sequence IS*6110* at a higher rate than described elsewhere; an insertion or deletion between every 2-5 years (de Boer et al. 1999; Yeh et al. 1998). The estimation of the time it takes for these insertion sequences to change was largely generated by studying serial isolates from individual patients. One hypothesis for this increased rate of change is the stress caused by transmitting between hosts. It may be possible to serially passage these strains to ascertain whether the rate of insertion and deletion slows in a less competitive environment.

Finally, the fitness of the resistant strains is not significantly different from either the laboratory strain *M. tuberculosis* $H_{37}Rv$, the susceptible strains in lineage 15 or unrelated isoniazid susceptible and resistant strains. This indicates that if there was any fitness cost initially associated with the acquisition of resistance-conferring mutations then this was either very small, or the organisms have compensated for it since. This allows the organisms to successfully transmit between individuals despite the acquisition of drug resistance. Indeed, related strains have been isolated from individuals around the country and cases now exceed 300 (Maguire H, Forrester, & Adam 2006). The investigation of

the mutation rate of these strains with the addition of whole genome analysis may prove insightful in this regard.

Not all isolates had an equal number of data points for the fitness experiments. The experiment was designed with nine replicates per isolate. However, practical constraints, including contamination and lack of growth in the 1:10,000 dilution tubes were experienced. Additionally, some variation was seen between the replicates. These biological experiments are difficult to control and various factors may result in problematic data. For example, *M. tuberculosis* clumps together in liquid culture. Every effort was made to break up the clumps of organisms prior to inoculation, but even small variations may result in variation in generation times. Additionally, although the sampling and inoculation of the cultures was tightly monitored, small variations in the phase of growth of the organisms may result in some variation.

The evolution of the organisms in lineage 15 (the North London isoniazid resistant TB outbreak) are hypothesised in figure 5.4. The proposed progenitor strain (02:113) was isolated in 2002. This organism is fully susceptible to first line anti-tuberculosis agents. A second strain (02:302), also isolated in 2002, is indistinguishable using both IS*6110* and MIRU-VNTR genotyping, but has developed streptomycin mono-resistance. The figure shows hypothesised routes of evolution from these early strains. The insertion and deletion of IS*6110*, alteration in MIRU-VNTR type and the acquisition of antimicrobial resistance is displayed. Additionally, the fitness of the strains relative to the reference

strain $H_{37}Rv$ are displayed. These data demonstrate that the strains have suffered no significant fitness deficit despite the alteration in genotype.

Investigating the broader evolutionary origins of these strains might prove enlightening. Gagneux and colleagues showed that strains in the group that he defined as European-American were three times more likely to give rise to a secondary case compared with other strains (Gagneux et al. 2006c). LSP and SNP analysis may reveal if the origins of these isolates has contributed to their success.

Given that the organisms studied from this diverse outbreak are as fit as the reference strain *M. tuberculosis* $H_{37}Rv$ as well as the fully susceptible outbreak strains, it seems reasonable to conclude that, in the absence of interventions, the outbreak will grow. Data indicating a high rate of infection in close contacts (11%) (Ruddy et al. 2004) suggests that these are highly competitive strains. However, epidemiological factors may play their part; many patients in this outbreak are prisoners and drug users who reside in confined dwellings with limited ventilation.

Reversion to drug susceptibility, even when antibiotic pressure is removed, is rarely seen (Andersson & Hughes 2010). Considering that these strains appear to have already compensated for any acquired fitness cost, reversion seems unlikely. Additionally, patients infected with drug resistant *M. tuberculosis* who are started on empirical treatment will respond less well, remain infectious for longer and risk the development of

further resistance. The additional factor of a fit, successful strain in this setting is ominous.

In conclusion, it appears that this lineage, which is associated with the North London isoniazid resistant outbreak, comprises a heterogenous collection of closely related strains. These strains appear to have thrived in the community, generating large numbers of clinical cases whilst evolving independently via the acquisition of antimicrobial resistance and the alteration of their genotypes.

5.3.2 Extended Family with TB

Contact tracing in the extended family generated a large number of active and latent cases of TB. Of the 20 individuals screened, seven (35%) were considered to have active disease and were treated accordingly, whilst two (10%) were offered prophylaxis for latent TB infection. These rates are considerably higher than those seen elsewhere. NICE guidance states that the reviewed literature (Ansari et al. 1998; Esmonde & Petheram 1991; Hardinge, Black, & Chamberlain 1999; Hussain et al. 1992; Irish, Jolly, & Baker 1997; Kumar, Innes, & Skinner 1992; Ormerod 1992; Teale, Cundall, & Pearson 1991) indicates that approximately 1% of contacts progress to active disease (NICE 2006). Contact tracing interviews conducted by the TB nurses revealed nothing extraordinary in their domestic arrangements that might increase the likelihood of transmission. This makes the high number of secondary cases highly unusual. The finding that the isolates were indistinguishable by 15-loci MIRU-VNTR typing supports their epidemiological link. These strains were isolated after the 2002 – 2007 database was closed and the

genotyping was performed at the National Mycobacteria Reference Laboratory using their standard 15-loci profile. The additional information gained by amplifying further MIRU-VNTR loci is open to debate. In such an outbreak when the individuals are clearly linked over a short period of time, it may be argued that additional loci are unnecessary.

The generation time of the organism isolated from the index case appeared to be longer than that of the organisms isolated from the other family members as well as laboratory strain *M. tuberculosis* $H_{37}Rv$ (25.6 hours vs. 17.0 hours), but this was not statistically significant ($p = 0.4232$ - Kruskal-Wallis test). The organisms isolated from five further cases demonstrated generation times that were far more similar to the reference strain and considerably shorter than the index isolate (table 6.5). The organism appears to have suffered negligible fitness cost despite acquiring streptomycin resistance and is therefore not prohibited from transmitting between various human hosts. This has been reported previously. For example, an isolate of *M. tuberculosis* from a patient in an outbreak of MDR TB was shown to be fitter than other strains in the outbreak (including the strain cultured from the index case). The patient deteriorated rapidly and died and it was therefore proposed that the strain had adapted to the environment of its host (Gillespie et al. 2002).

It would be interesting to investigate whether the strains possess the most commonly seen mutations in *rps*L causing streptomycin resistance in *M. tuberculosis* as these have been shown experimentally to exhibit no fitness cost in *S. typhimurium* and *E. coli* (Bottger et al. 1998).

In conclusion, the strain appears to be highly successful at transmitting between hosts, despite being highly resistant to streptomycin. Additionally, any fitness cost experienced by the organism appears to be ameliorated in subsequent hosts.

In summary, we have described two outbreak situations involving drug resistant strains. In both settings it has been demonstrated that if the organisms experienced any fitness cost as a result of the acquisition of these resistance-conferring mutations, they have since compensated for it. The isoniazid resistant outbreak strains appear to be more heterogenous than the family outbreak strain. This is not unexpected given the time frames involved. The longer time-line involved in the north London outbreak has allowed the organisms to alter their IS*6110* RFLP patterns so that a split tree diagram of evolution may be hypothesised. What is unexpected, however, is the diversity of isoniazid conferring mutations present. This indicates that the successful strains independently acquired resistance whilst propagating in the community. The isolates from the family outbreak were cultured over only a few months, so diversity is far less likely and this is a more classical point-source outbreak.

The relationship between the acquisition of drug resistance and fitness cost is far from straight forward. It is clear from these examples that where a resistance-conferring mutation is associated with little or no fitness cost that these successful strains will thrive and propagate, given the opportunity.

Chapter 6: The Role of Molecular Epidemiology in Clinical Practice

6.1 Introduction

Although the two processes frequently work in unison, it is important to distinguish traditional epidemiology and genotyping. Traditional epidemiology can identify cases of TB that are related in time and space, but this is often not sufficient. Genotyping is required to definitively identify outbreaks and chains of transmission. Genotyping is extremely useful in several clinical settings, for example, in the identification and investigation of outbreaks, the detection of false positive cultures and the differentiation of re-infection and relapse. Prospective, real-time genotyping of *M. tuberculosis* isolates can therefore benefit both the laboratory, and the clinical teams. Our clinical colleagues often identify trends or links between patients that may be related to the infecting organism. It is important therefore, to investigate whether these observed findings are linked via the organisms' genotype. In this chapter, the diagnostic utilisation of genotyping is examined.

6.2 Example 1: Paradoxical reaction and Strain Type

A paradoxical reaction (PR) in a tuberculosis patient is characterised as the transient worsening of symptoms or the appearance of new lesions or symptoms despite appropriate anti-tuberculosis therapy (Breen et al. 2004). It was hypothesised that the appearance of a paradoxical reaction may be related to strain type.

Between January 2002 and December 2003 all adult patients seen at our centre with culture positive MTB had IS*6110* RFLP typing performed on one isolate. A case note review was performed retrospectively for clinical evidence of PR (Breen et al. 2004)

Isolates from 145 patients were typed but 45 were excluded as the notes for 24 were unavailable or incomplete and 21 patients were lost to follow up or care was transferred). This large exclusion rate was due to the fact that this analysis was performed retrospectively and the notes were not available for all patients. Of the 100 (69%) patients' notes reviewed, 52 were male and the age range was 16–81 years. Ethnicity data showed that 48% were black African, 16% Asian, and 19% from the UK. PR occurred in 20 patients (20%) (HIV positive 10/26 (38%); HIV negative or unknown 10/74 (14%)). All patients with PR had distinct IS*6110* RFLP profiles suggesting 20 separate strains (figure 6.1). Only four of the twenty strains (03:390, 03:334, 03:063 and 03:092) were related to another at ≥70%. This number is surprisingly low considering that one third of the isolates (318/478, 66.5%) on the whole database were related to another isolate at this level (section 3.3.5).

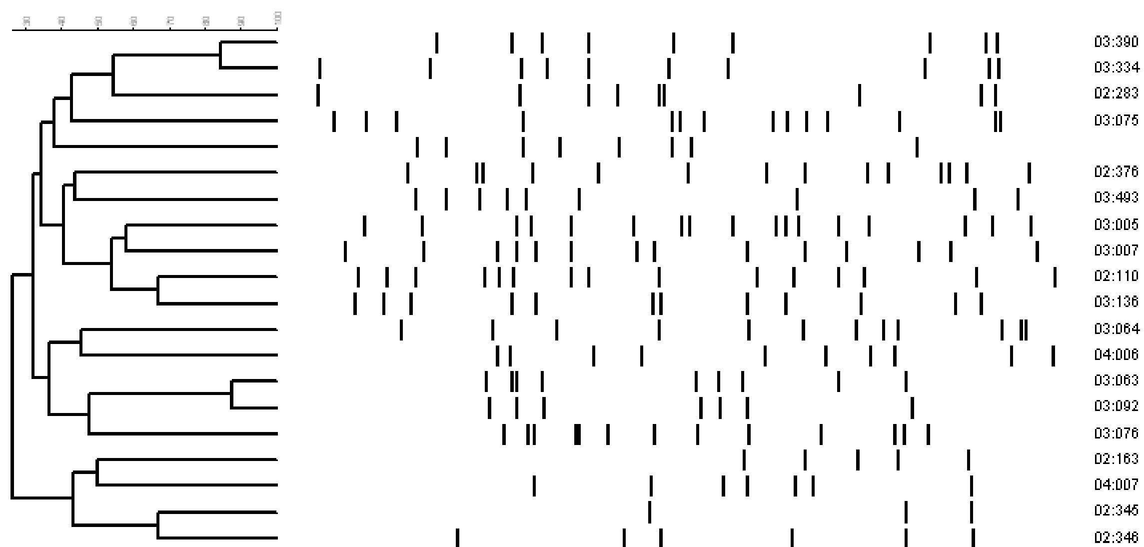


Figure 6.1 Dendrogram of patient strains isolated from patients with paradoxical reactions.

The possibility exists that paradoxical TB reactions may be a consequence of specific host response genes, such as HLA (Perrin et al. 2005). It is noteworthy that not all patients in this group could be included for analysis as this retrospective study did not offer the opportunity re-interview patients who notes were incomplete. However, the data available do indicate that there is no association between genotype and paradoxical reaction in the patients that were studied.

6.3 Example 2: Suspected Contamination due to a Faulty Bronchoscope

In order to allow intervention in real time, genotyping results are often required rapidly. This is evident in the identification of laboratory contamination (chapter 3) and the

investigation of suspected outbreaks (chapter 5). In this example, MIRU-VNTR genotyping was used to investigate a suspected iatrogenic outbreak of *M. tuberculosis* involving a bronchoscope. Such outbreaks are not uncommon and many have been documented in the literature (Cetre et al. 2005; Chroneou et al. 2008; Diaz Granados et al. 2009).

M. tuberculosis was isolated from four patients on the same bronchoscopy list at the Royal Free Hospital. The bronchoscope used was subsequently found to be faulty and could not be thoroughly decontaminated. It was possible that these infections had been transmitted during the procedure or were episodes of contamination of the obtained samples from non-infected patients. A specimen from a further patient was culture negative, but positive by ProbeTec Strand Displacement Amplification (SDA) assay (Becton Dickinson, Oxford, UK). Organisms from the four culture-positive patients, together with the specimen from the SDA-positive patient were strain typed using 12 loci MIRU-VNTR.

The data showed several differences between the isolates (table 6.1), both in loci and allele number. Some PCR reactions failed, but these did not detract from the information provided. Twelve-loci MIRU-VNTR was performed on these isolates as robotic and automated sequencing technologies were not available at this time. Twelve loci were amplified by conventional PCR and examined using agarose gel electrophoresis with ethidium bromide. This allowed a rapid assessment of whether the strains were unique. Had the strains been indistinguishable using these limited loci, then secondary typing using IS*6110* RFLP would have been used.

	MIRU loci and corresponding alleles											
Isolate	**2**	**4**	**10**	**16**	**20**	**23[2]**	**24**	**26[2]**	**27**	**31**	**39**	**40**
03:409 Patient A	2	2	2		2	5	1	7	3	5	4	3
03:420 Patient B	2	2	3	4	1	5	1	1	3	3	2	
03:434 Patient C	2	2	3	1	2	8	1	5	2	3	2	
03:435 Patient C		2	3	1	2	8	1		2	3	2	
03:485 Patient D	2	2	5	1	2	5		1	3	3	2	2

Table 6.1 Summary of MIRU-VNTR type for the four isolates in the suspected bronchoscope outbreak. Blank cells indicate a failed PCR reaction

PCR products could not be generated from the specimen from the SDA-positive patient as the SDA DNA extraction method was unsuitable for use with the MIRU-VNTR PCR.

MIRU-VNTR typing demonstrated that the four isolates were different strains. IS*6110* RFLP typing would have taken a number of weeks whereas these MIRU-VNTR results were available within two days. These results had an important impact clinically; the TB team continued with anti-tuberculosis therapy as the evidence indicated that the four patients were genuinely infected with distinct strains (Shorten et al. 2005). Had genotyping not been available, or delayed in the case of IS*6110* RFLP, then the decision to treat each patient would have had to be made purely on a clinical basis. Indeed, those with other differential diagnoses or those with symptoms inconsistent with TB may have been denied treatment until further specimens were obtained. During this time, their disease may have progressed and they may have infected others.

6.4 Example 3: Differentiation of relapse and re-infection

A 34 year old female patient from the Congo presented at Accident and Emergency complaining of cough, fever, night sweats, shortness of breath at rest, being unsteady on her feet and dizziness. She was admitted to the infectious diseases ward.

Chest x-ray showed changes in the upper right zone. Sputum obtained from the patient was smear positive for acid fast bacilli. SDA performed on the sample was also positive. A fully susceptible isolate of *M. tuberculosis* was subsequently cultured from the sputum. An HIV test performed on the patient was positive. The patient's CD4 count was 11 cells per µL, but no viral load was recorded. Standard anti-TB and anti-retrovirals were started. The patient was recorded to be 100% compliant with her treatment and the CD4 count and viral load were after one month of treatment were 40 cells per µL and 1420 copies per mL, respectively.

The patient returned to the Royal Free hospital two years later following a trip to Congo. She was unwell and complained of fever and sweats. Chest x-ray revealed evidence of old TB in the upper right zone as well as consolidation in the mid left mid zone. She was diagnosed with advanced HIV (CD4 count 18 cells per µL, viral load 164 copies per mL)

The patient was unable to produce any sputum, so underwent a bronchoscopy. An inflammatory lesion was observed in the left main bronchus. The sample obtained during bronchoscopy was microscopy and SDA positive. The lung biopsy showed a granulomatous inflammatory infiltrate composed of inflammatory cells, macrophages and an occasional multinucleate giant cell. Acid fast bacilli were seen and fungal stains were

negative. *M. tuberculosis* that was highly resistant to streptomycin was cultured from the specimen.

The clinical team noted that the patient had been compliant with her previous TB treatment so questioned whether this may be a case of re-infection rather than relapse. IS*6110* RFLP genotyping was performed on the isolates cultured from the samples obtained upon her first presentation and the isolate from two years later. The IS*6110* genotypes are shown in figure 6.2.

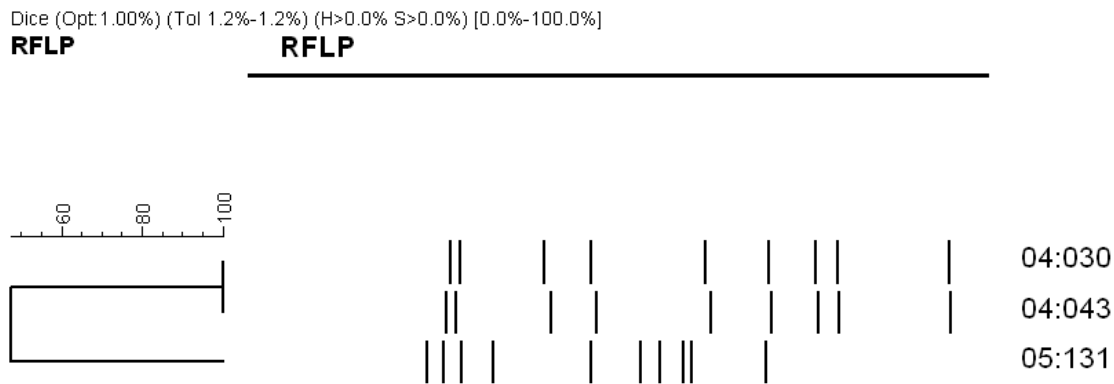


Figure 6.2 The IS*6110* RFLP genotypes of the isolates obtained from the patient on two occasions, approximately 17 months apart. The isolates from December 2003 (04:030 and 04:043) were fully susceptible. The isolate from April 2005 (05:131) was highly resistant to streptomycin.

IS*6110* RFLP genotyping revealed that the isolates were different (50% similarity) and indicated that the patient had been re-infected with another stain, rather than experienced relapsed. She was seen in the TB clinic on 10/08/2005. The clinical team felt that she was poorly compliant with her anti-retroviral treatment as her viral load had risen from less

than 50 to greater than 12,000 copies per mL. Her TB has responded well, however, and was declared cured after six months' treatment.

6.5 Example 4: Genotype and Contract tracing

A 34 year old woman from Indonesia was referred by her general practitioner to Accident and Emergency following a cough and breathlessness that had not responded to erythromycin. She was previously well, but was 7 weeks pregnant and had been in the UK for 20 months. She had been BCG vaccinated as a child. She had a history of cough and sputum production, fevers and weight loss of 2kg in the last 3 months (despite being pregnant). She had become breathless in the previous 2 weeks.

She was admitted and a sputum sample was positive for acid fast bacilli and she was started on standard anti-TB treatment and prednisolone. A chest x-ray and CT taken at this time showed extensive shadowing and a cavity in the upper right lobe (figures 6.3 and 6.4).

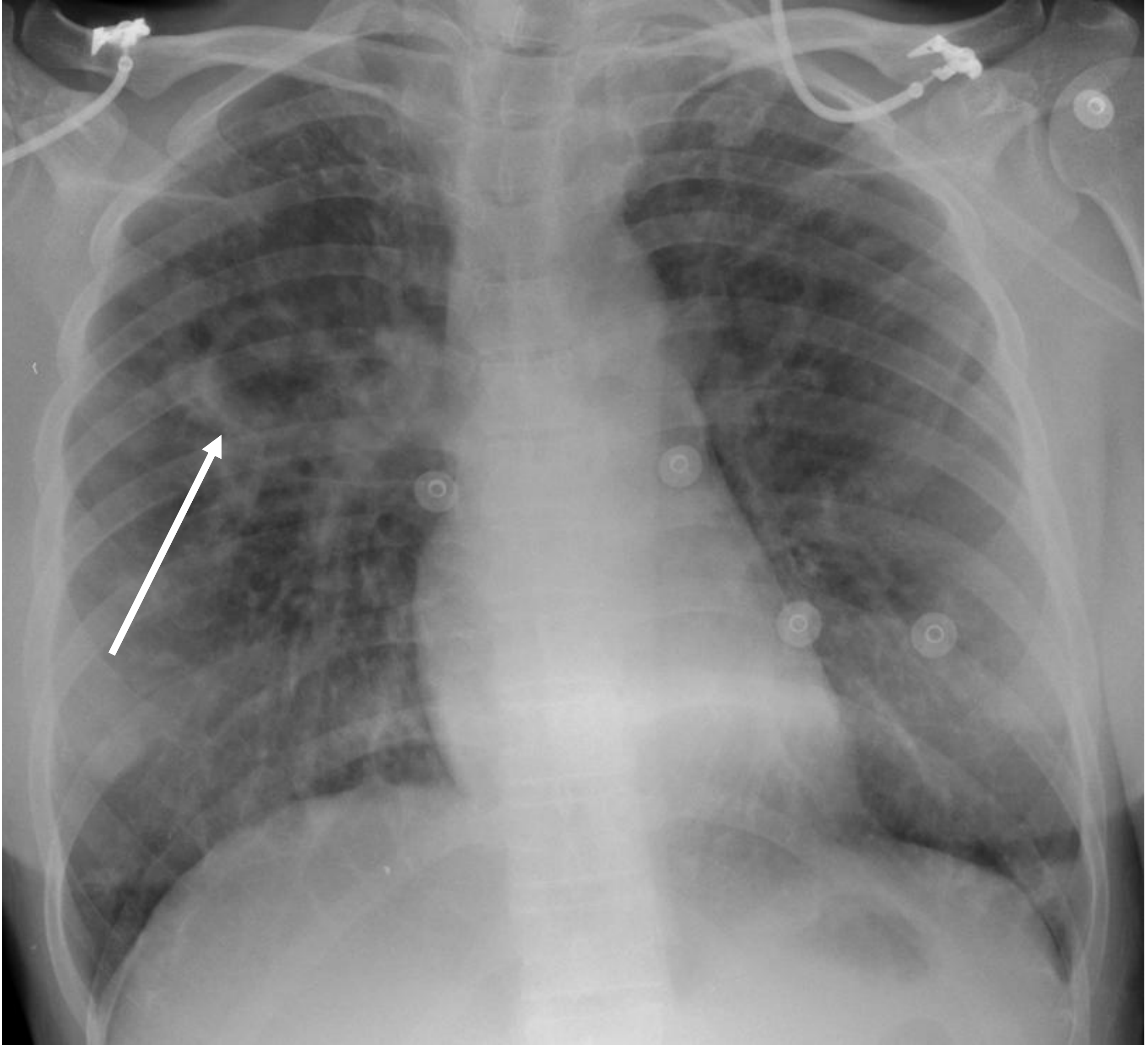


Figure 6.3 Chest x-ray three weeks after presentation. The cavity in the upper right lobe is indicated with an arrow.

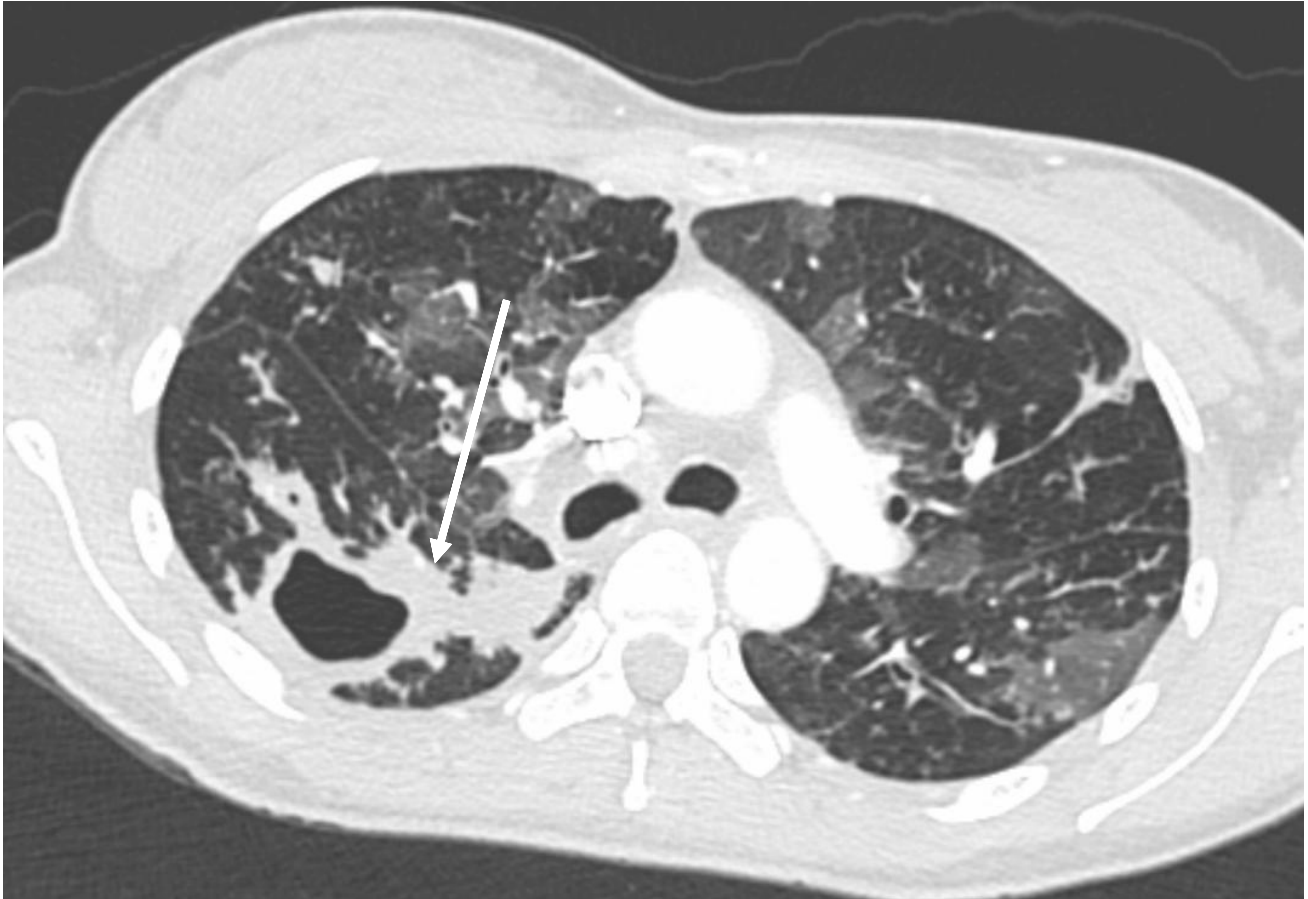


Figure 6.4 Chest CT three weeks after presentation. The cavity in the upper right lobe is indicated with an arrow.

A fully sensitive isolate of *M. tuberculosis* was cultured and the patient progressed well on treatment for nine months, apart from a suspected paradoxical reaction after 21 days of treatment. The patient was HIV negative. The chest x-ray at completion of treatment was indicative of resolution (figure 6.5).

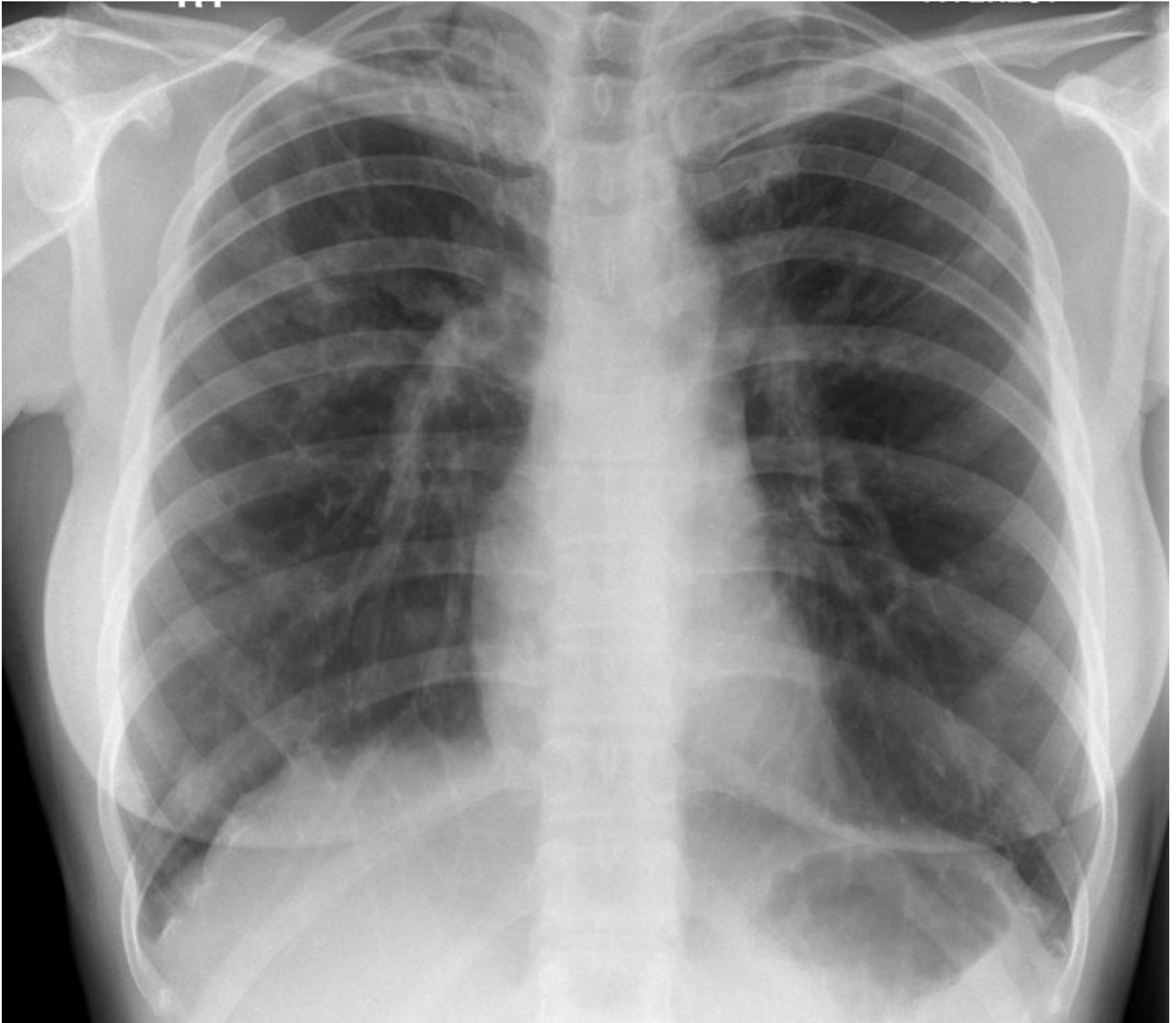


Figure 6.5. Chest x-ray at the completion of treatment.

Once the diagnosis had been made, contact tracing was begun. The patient's partner was a 40 year-old Iranian man who had had an intermittent cough for four weeks. He was a smoker who was otherwise very well. He had no history of TB, had lived in the UK for several years and had been BCG vaccinated as a child.

He received a chest x-ray, which was reported as normal (figure 6.6). He also had a Tuberculin skin test (Heaf), which was positive (grade IV). He had three smear negative (and subsequently culture negative) sputum samples and his CRP was less than 1mg/L.

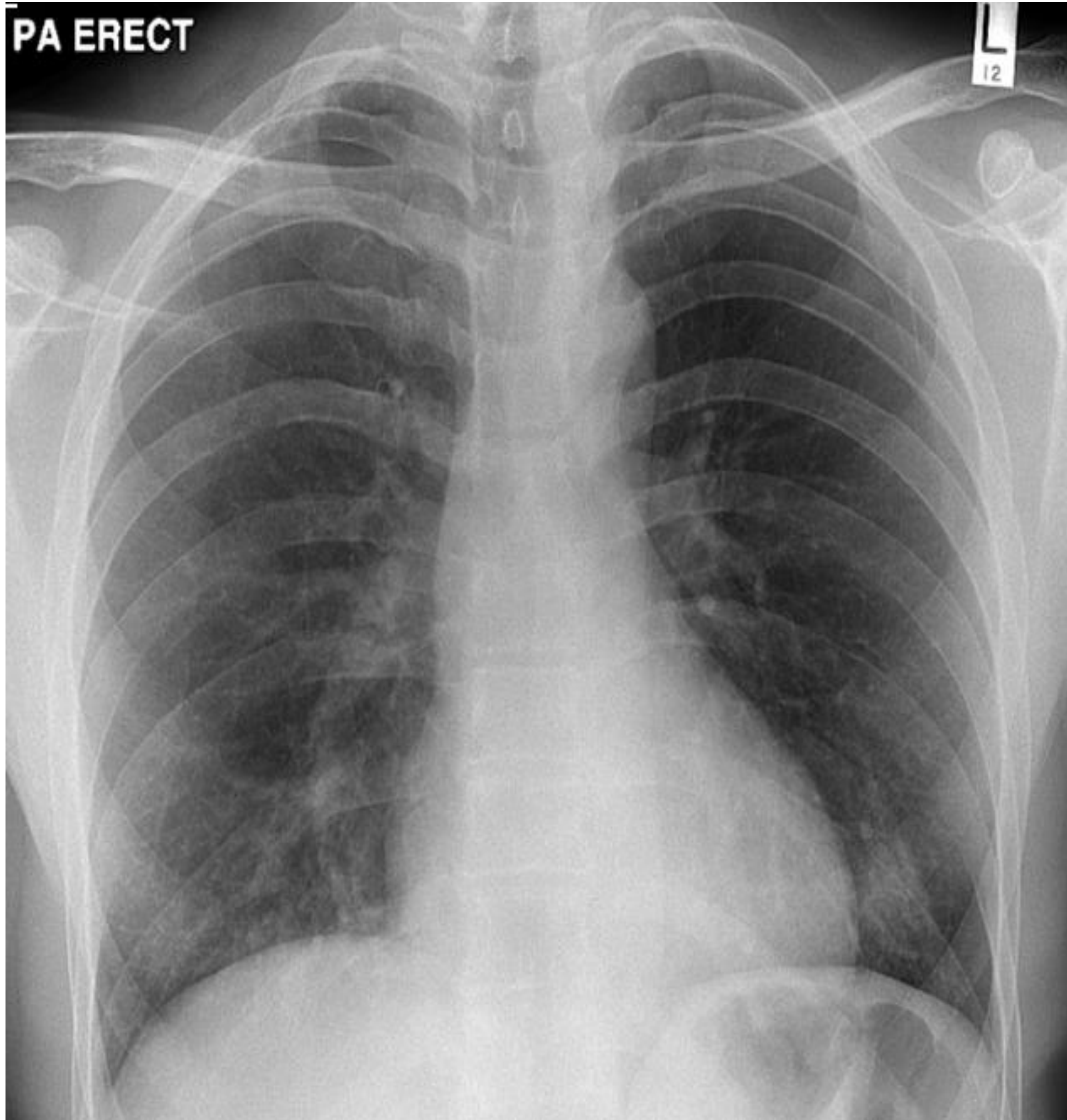


Figure 6.6. Chest x-ray of male contact at day 8 of the patient's treatment.

He was reviewed 10 weeks after his partner had begun anti-TB treatment. He had no new symptoms and his CRP remained less than 1mg/L. His chest x-ray however, indicated the possibility of an abnormality in the left lower zone (figure 6.7).

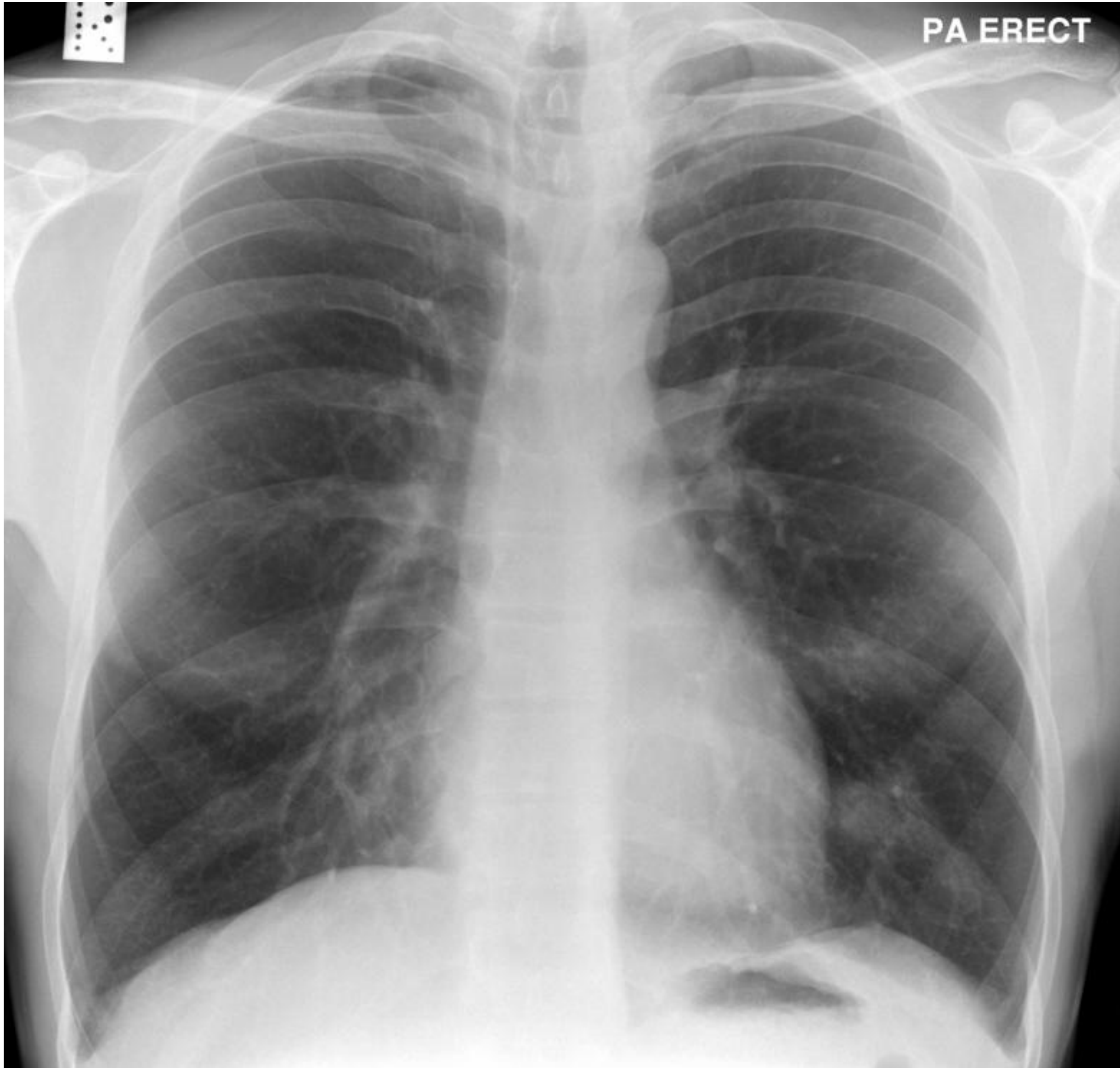


Figure 6.7. Male contact's chest x-ray 10 weeks after his partner had started anti-TB treatment.

He produced an induced sputum sample that was microscopy negative, but *M. tuberculosis* NAAT positive.

The patient remained well, but the team were unsure how to manage him, given that he was a close contact of a smear positive case. His skin test was positive, although he was born outside the UK and was therefore more likely to have been exposed to tuberculous and non-tuberculous mycobacteria. As he had a positive NAAT, had minor respiratory symptoms and the chest x-ray was not clear, he was sent for a chest CT (figure 6.8).

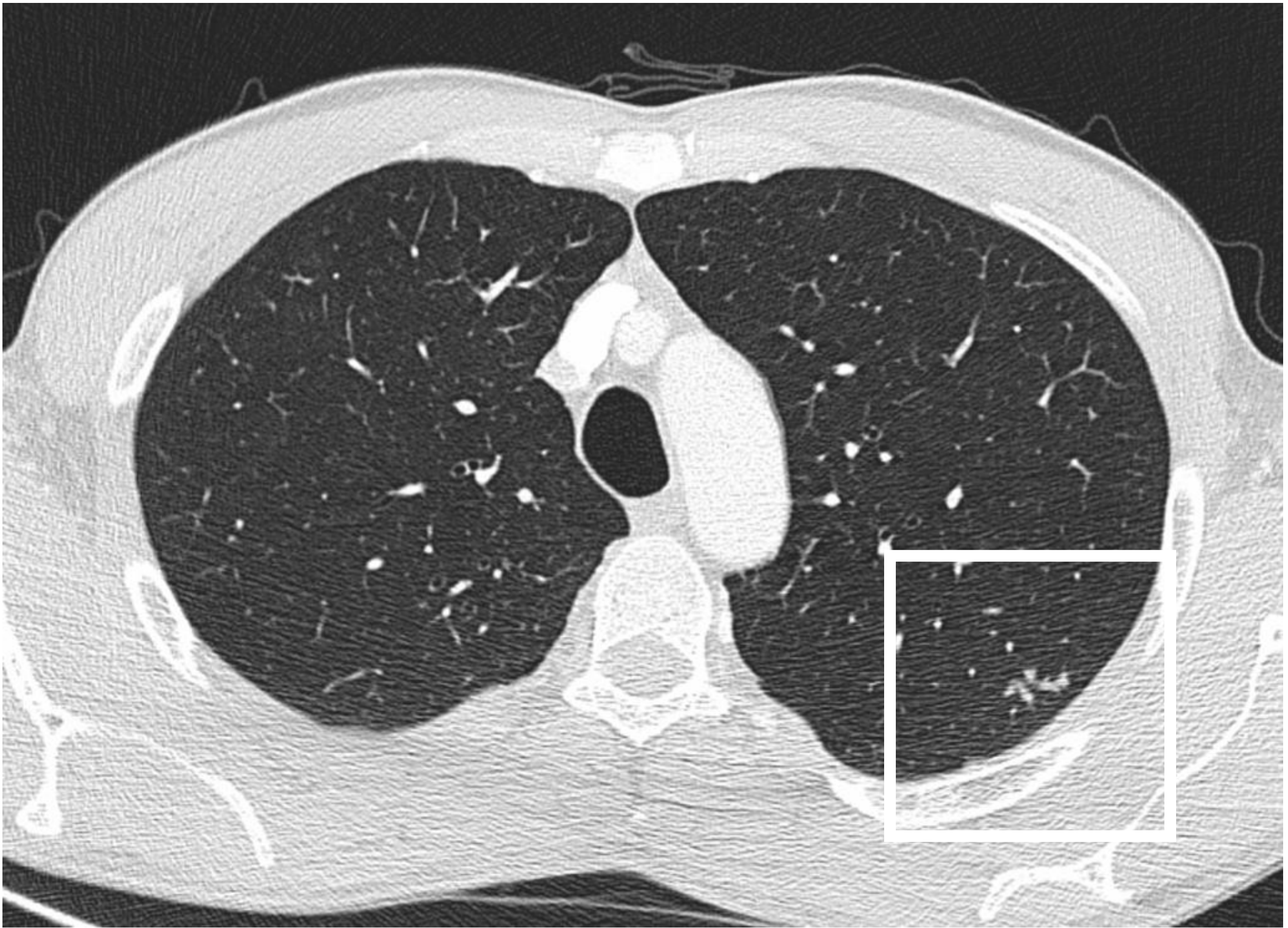


Figure 6.8. Chest CT of male contact at week 12 of the patient's treatment. The highlighted area in the left lower zone is magnified in figure 6.8a.

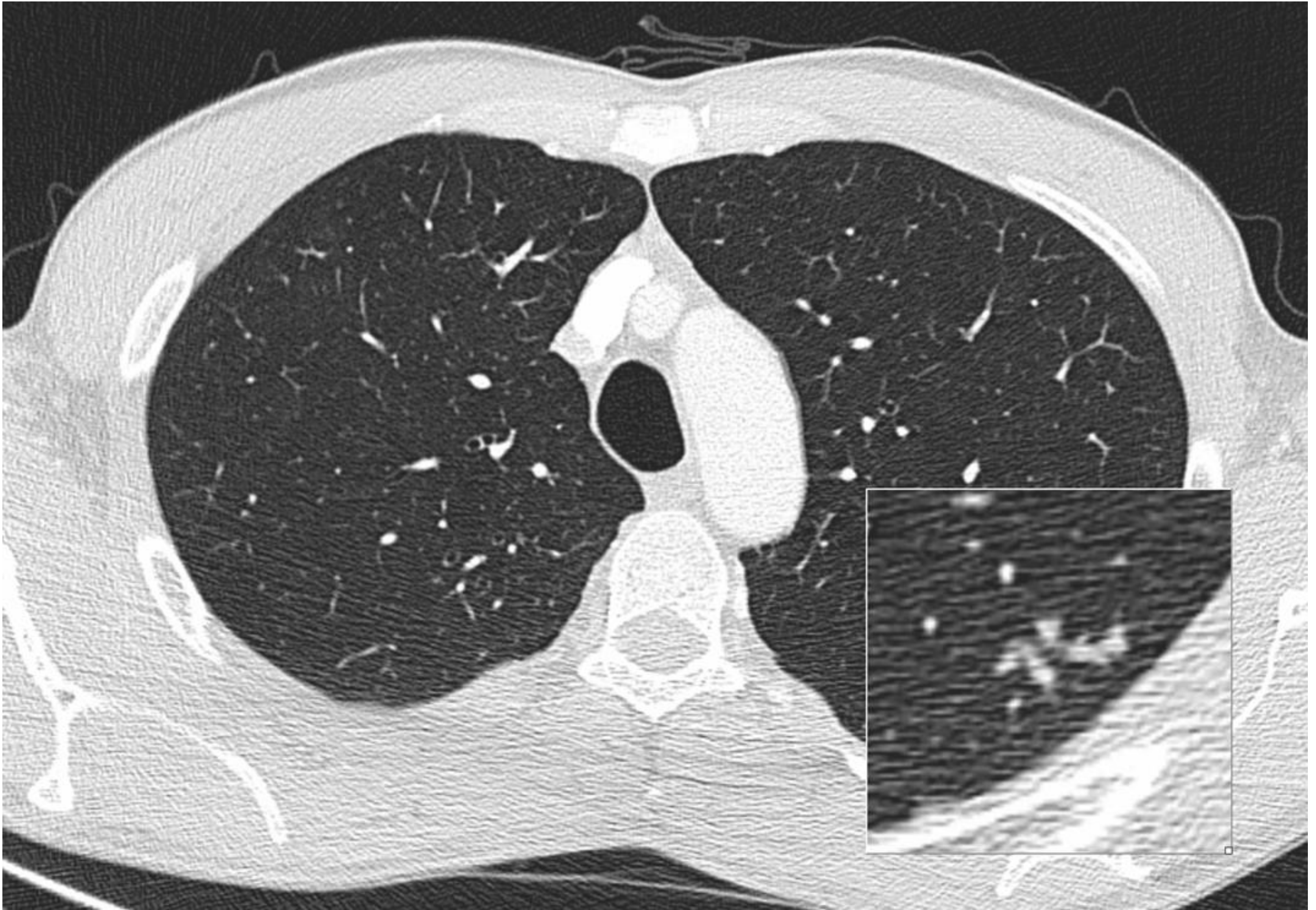


Figure 6.8a Chest CT of male contact at week 12 of the patient's treatment. The highlighted area in the left lower zone is magnified.

It was then decided to treat the male contact for active TB. A fully susceptible strain of *M. tuberculosis* was subsequently cultured from an induced sputum sample. MIRU-VNTR typing of the isolates from the two patients revealed them to be indistinguishable (table 6.2), and therefore indicative of transmission from the female patient to her partner. Twelve-loci MIRU-VNTR was performed on these isolates as robotic and automated sequencing technologies were not available at this time. The isolation of indistinguishable strains by 12-loci MIRU-VNTR from such close household contacts within four months of each other is an extremely strong indicator of transmission. It is unlikely that adding

further MIRU-VNTR loci would produce different results given the short time frame between presentation of the individuals.

	MIRU-VNTR locus														
	2165	**2461**	**577**	**580**	**3192**	**154**	**960**	**1644**	**2059**	**2531**	**2687**	**2996**	**3007**	**4348**	**802**
♀	**3**	**2**	**3**	**3**	**3**	**2**	**5**	**3**	**1**	**5**	**1**	**5**	**3**	**1**	**3**
♂	**3**	**2**	**3**	**3**	**3**	**2**	**5**	**3**	**1**	**5**	**1**	**5**	**3**	**1**	**3**

Table 6.2. 15 loci MIRU-VNTR typing results of the two patients.

In conclusion, MIRU-VNTR demonstrated that the two individuals were infected with an indistinguishable strain of *M. tuberculosis*. This is not entirely surprising considering their relationship and domestic arrangements. However, the two patients presented with different clinical phenotypes and had different radiological findings. Furthermore, if NICE guidance had been followed then the male patient would not have been diagnosed as quickly. The patient's first chest x-ray was normal and as such, no further follow up should have been initiated (figure 6.9). The subsequent clinical investigation and genotyping allowed the diagnosis of active TB that had been contracted from his partner.

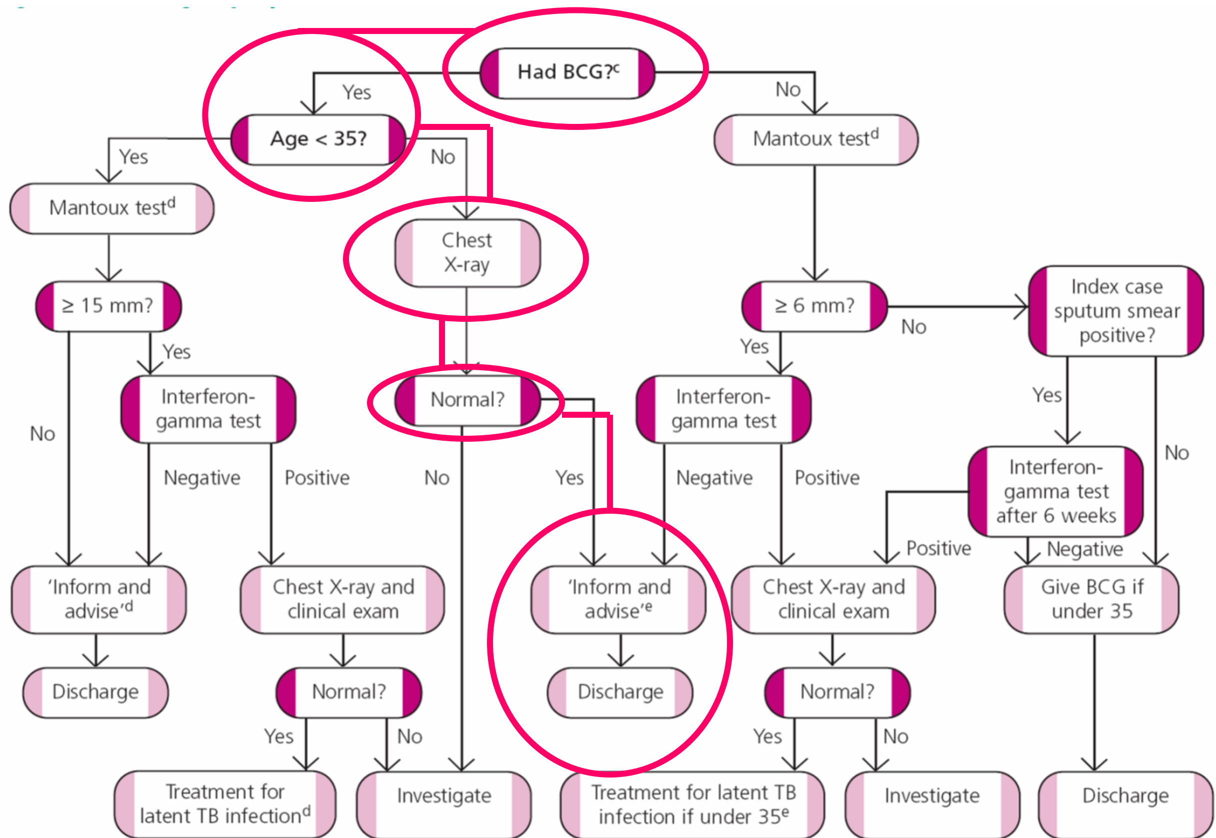


Figure 6.9 NICE guidance for contact tracing. The male patient initially had a normal chest x-ray so should have been discharged, as indicated by the red circles. Adapted from (NICE 2006).

6.6 Discussion

In summary, as well as studying the evolutionary relationships between *M. tuberculosis*, genotyping can provide extremely useful information that is relevant to our clinical colleagues. The real-time investigation of suspected and known outbreaks can be used to rationally guide contact tracing and early detection and intervention, especially in the case of drug resistant strains.

It would be naïve to believe that genotyping *M. tuberculosis* alone can control the spread of infection, but its use, alongside classic epidemiology and contact tracing is important. Additionally, the use of genotyping alongside social network analyses, which focuses contact tracing investigations on locations rather than individuals can accurately identify and define transmission events (Barnes et al. 1997; Cook et al. 2007; Fitzpatrick et al. 2001; Klovdahl et al. 2001). Furthermore, high rates of “false clusters” have been seen when using only 12 MIRU-VNTR loci (Gopaul et al 2006). The use of this laboratory data alone would generate many false alarms that may initiate unnecessary public health interventions. Therefore, the use of an appropriate genotyping tool alongside epidemiological data can ensure that public health resources are used rationally.

The fact that TB often has a long incubation period and disease course, and is not geographically constrained makes the combination of laboratory and clinical epidemiology even more important. Outbreaks can occur over years, rather than weeks or months with other infectious diseases (Fitzpatrick et al. 2001; Ruddy et al. 2004), which means that identifying both cases and contacts as rapidly as possible is imperative. The establishment of a national, prospective MIRU-VNTR genotyping database by the Health Protection Agency will enable the identification of related cases and outbreaks in real-time.

The utilization of automated 24-loci MIRU-VNTR can enable a genotype to be available within days of a positive culture being confirmed. This will bridge the gaps between the laboratory, the clinical TB team and those engaged in public health. It is reasonable to

suggest that some outbreaks may have been identified sooner had such technology and systems been in place. For example, some outbreaks are identified initially due to an unusual drug-susceptibility pattern (Edlin et al. 1992; Ruddy et al. 2004). If real-time genotyping is available for all isolates of *M. tuberculosis*, further resources may be able to be provided at an earlier stage.

Additionally, if real-time genotyping is available to all TB services, then the identification of false positive cases due to contamination can be achieved more quickly (Ruddy et al. 2002). This will prevent patients being inappropriately treated and investigated, which will, in turn, allow resources to be targeted to genuine cases. The monitoring of laboratory cross-contamination will not only provide the clinical teams with confidence in the laboratory service, but it will reassure them that the correct patients are being treated. Furthermore, real-time monitoring of contamination events will help to improve and maintain laboratory standards.

The use of genotyping may be used to bridge the gap between laboratory and clinical investigations. Patients presenting with unusual phenotypes may be investigated to determine whether this may be organism-related. In a period when cases of tuberculosis are once again increasing (HPA 2009), the coordination of laboratory and clinical data can only bring benefits in the understanding of the disease. The inescapable interaction between the biology of the organism and the host must be studied in tandem in the continued search for improved diagnostics, treatments and vaccines.

In conclusion, the advances in the knowledge of the genome of *M. tuberculosis* over the previous two decades have a direct impact on the clinical and public health management of the disease. Enabling the link between the transfer of laboratory data into the clinical realm and vice versa must be continued to identify new cases, contacts and outbreaks in the short term, whilst improving the possibility of eradication in the long term.

Chapter 7: Final Discussion

Tuberculosis is an ongoing global concern. The United Nations has predicted that the world's population will reach in excess of 9 billion by 2050. Almost all of this increase will be seen in the developing world; six of the eight countries predicted to contribute to this increase are in Africa or South Asia (United Nations 2005). The Global Plan to Stop TB 2011-2015 (WHO 2010) has set targets for reducing cases and deaths in the next five years. However, there is concern because the increase in population density that is predicted will affect areas with high rates of TB. Therefore, if the initiatives are not successful, then increases in the population in the areas with high rates of TB may lead to an increase in numbers of cases. Other concerns surround the development and spread of drug resistance. The United Nations has instigated several Millennium Development Goals, including tackling infectious diseases, such as tuberculosis (United Nations 2000). Halting the continued dissemination of disease can only be achieved with concerted action, including increased detection of active disease and latent infection, improved, shortened therapy (including the development of novel agents) and a more effective vaccine.

The hypothesis presented here is that the characterisation of the genotype of *M. tuberculosis* could be used to give an insight into the biology of the organism that relates to its clinical phenotype. The hypothesis was addressed by performing *M. tuberculosis* genotyping alongside the collection of demographic data for all culture positive TB patients at the Royal Free Hospital, Hampstead, London. We have shown that this geographic location is the home of a diverse population of *M. tuberculosis* by the

demonstration of a low rate of transmission (6.3%) compared to other studies (table 1.2). Additionally, IS*6110* RFLP genotyping revealed in excess of 60 70% lineages. This finding, in combination with the high rates of migration in London demonstrates a diverse collection of strains. The identification of over-represented lineages amongst the dataset allowed the association between genotype and phenotype to be studied.

It is clear from the data presented in chapter 3 however, that IS*6110* RFLP used alone is not sufficient to address all of the issues raised, as has been found by other authors (Gillespie, Dickens, & McHugh 2000; van Deutekom et al. 2005). As shown in chapter 3, nearly one in five isolates of *M. tuberculosis* in this data set posses too few copies of this insertion sequence for this method to be utilised for their differentiation. This finding, combined with the deficiencies identified when interpreting poor quality images from this technically demanding tool make it an imperfect genotyping method. MIRU-VNTR was shown to be a valuable secondary genotyping method in this setting. The higher level of discrimination seen in subdividing lineages generated by IS*6110* RFLP, together with the comparative ease of use and transportability of the data makes a compelling case for its adoption as the genotyping method of choice, as is the case for the UK national typing strategy. The advantages, disadvantages and applications of various genotyping methodologies are summarised in table 7.1.

Genotyping *M. tuberculosis* has multiple benefits in the study and control of the disease. Calculation of the rate of transmission as 6.3%, as performed in chapter 3, allows the rational allocation of public health resources to areas with the highest burden. It also

allows outbreaks to be identified more rapidly, such as those in North London and within households described in chapters 3 and 5. The identification of such outbreaks subsequently allows the targeting of public health interventions. The identification of contamination events, whether they be pre-analytical or within the laboratory are essential for good patient care. The realisation that the resilient nature of this organism can cause false positive cultures in even the most specialised of laboratories must drive up standards of practice.

<table>
<tr><th>Genotyping methodology</th><th>Relative discrimination</th><th>Advantages</th><th>Disadvantages</th><th>Uses</th></tr>
<tr><td>IS6110 RFLP</td><td>High</td><td>Highly discriminatory</td><td>Unsuitable for <5 copies of IS6110
Technically difficult</td><td>Outbreak/epidemiological and evolutionary investigation</td></tr>
<tr><td>12-loci MIRU-VNTR</td><td>Moderate</td><td rowspan="3">Easy to perform

Automatable
Data easily transportable</td><td>Moderate discrimination</td><td>Outbreak/epidemiological investigation</td></tr>
<tr><td>15-loci MIRU-VNTR</td><td>High</td><td rowspan="2">Cost associated with automation</td><td>Outbreak/epidemiological investigation</td></tr>
<tr><td>24-loci MIRU-VNTR</td><td>High</td><td>Outbreak/epidemiological and evolutionary investigation</td></tr>
<tr><td>Spoligotyping</td><td>Low</td><td>Data easily transportable</td><td>Low discrimination</td><td>Phylogenetic and low discrimination epidemiological investigations</td></tr>
<tr><td>Whole genome sequencing</td><td>High</td><td>Highly discriminatory</td><td>High cost</td><td>In depth evolutionary analysis</td></tr>
</table>

Table 7.1 The advantages, disadvantages and applications of IS6110, MIRU-VNTR, spoligotyping and whole genome sequencing.

It is imperative to ascertain what question is being asked when genotyping *M. tuberculosis*. The in-depth evolutionary study of the organism may require the use of several methodologies, investigating numerous genomic targets (Gagneux & Small

2007). However, 24-loci MIRU-VNTR typing data has shown concordance with phylogenetic groups in over 90% of cases (Brown et al 2010). Conversely, only 12 loci may be necessary to refute an apparent outbreak in a specific setting (Shorten et al. 2005). It is worth noting however, that the use of 12-loci MIRU-VNTR in isolation in prospective genotyping can generate a large number of “false clusters” resulting in unnecessary public health interventions (Gopaul et all 2006). Therefore, it is important to select the genotyping tool appropriate for the question being asked.

It is worthwhile considering the limitations of this type of data. Studies that use molecular genotyping as their foundation are inherently limited by the fact that one can only genotype the strains that are isolated. Approximately 60% of TB patients in this data set and nationally (HPA 2009) who were diagnosed with TB were culture positive for *M. tuberculosis*. Whilst it is fair to say that a significant proportion of patients who commence anti-tuberculosis treatment in the absence of a positive culture are likely to have active infection, the isolation of an organism is nevertheless reassuring for the clinical team. This demonstrates the need for improved diagnostic tools because the early detection of cases, alongside appropriate treatment and contact tracing, is essential in preventing the spread of TB (WHO 2006). Ultimately however, this means that the conclusions drawn between the relationship between genotype and clinical phenotype may be biased. It is known that the relative fitness of some clinical strains varies (O’Sullivan et al 2010) and that this may therefore affect the chance of their recovery from clinical specimens. There appears to be no evidence, however, that certain lineages

of *M. tuberculosis* are more difficult to culture, perhaps suggesting that the proportion that are cultured and genotyped are representative of the population as a whole.

It would be worthwhile expanding this dataset to include demographic data (for example, country of birth, ethnicity and HIV status) on all TB patients, rather than just those who provided and positive culture. The analysis of the demographic data of all TB patients may provide further insight into the population within our geographic region. It may identify particular groups of patients who present with particular phenotypes of those who are less likely to have their diagnosis confirmed with culture. Investigations into the possible reasons for this may be performed and more resources could be targeted at obtaining positive cultures from these patients.

The differences in the country of birth and ethnicity of patients seen in these data compared to nationally have been described in chapter 3. The differences are likely to be due to the ethnic makeup of the catchment area of the Royal Free Hospital. Data collected by The Greater London Authority demonstrates that large numbers of Somalis live in the London boroughs that are served by the Royal Free. It is logistically important to identify which patient groups are at risk within the surrounding area. Ultimately, the aim must be to encourage patients to engage with a healthcare setting. This may be achieved by providing literature in the appropriate languages, engaging with hard to reach patient groups, such as the homeless and those that misuse drugs or alcohol, together with the use of culturally sensitive practices.

The analysis of country and birth and ethnicity data provided the surprising conclusion that European patients and those born within the UK were more likely to have smear positive pulmonary disease. One hypothesis for this finding is that TB may not always be suspected initially in these patients. Symptomatic patients from sub-Saharan Africa and South Asia may be investigated for TB more rapidly, whilst diagnosis in the European/UK-born group may be delayed. This would allow an increase in bacterial load, leading to the increased probability of smear positivity.

Further investigation of this dataset with regard of phylogeographic groups may be revealing. For example, it would be of interest to ascertain whether isolates in this database with one copy of IS*6110* were more likely to belong to Principle Group IV, similar to those described by Baker and colleagues (Baker et al. 2004). Additionally, the country of birth of the TB patients could be investigated alongside the phylogeographic groups of their isolates as others have done (Dale et al. 2005; Reed et al. 2009).

The rate of transmission in this study was lower than that described in a previous study in London (Maguire et al. 2002). It is likely that this is due to the restricted geographical location and limited time-frame of the study. TB is an infection with a prolonged incubation and disease course. It is also not geographically constrained. It is likely that not all patients within chains of transmission have been identified within this study as symptoms in some patients developed outside this time frame or that they presented to other healthcare centres. For these reasons, the establishment of a prospective nationwide genotyping database at the Health Protection Agency is extremely positive.

Differences in discriminatory power between methods are well documented (Kremer et al. 2005) and false clusters of strains have been seen when using one method alone (Gillespie, Dickens, & McHugh 2000). It has been hypothesised that to establish evolutionary links between isolates, more than one genotyping methodology is required (McHugh et al. 2005). The benefit of this approach can be seen when examining the apparent lineages identified by IS*6110* RFLP in chapter 3. One group of strains (lineage 19) appeared to be closely related at ≥70% by IS*6110* RFLP. Subsequent secondary typing with MIRU-VNTR revealed that the isolates were sub-divided into two smaller lineages. This was particularly well demonstrated in the minimum spanning tree (figure 4.8), in which it can be seen that members of lineage 19b have evolved from members of 19a. It would be fascinating to look deeper into the field epidemiology of the patients in this lineage to ascertain the chain of transmission and the time-line involved, in comparison with the stability of the MIRU-VNTR loci gained and lost.

The identification of a lineage of closely related strains in chapter 3 that were noticeably more drug resistant than other lineages in the dataset was of great interest. It was confirmed that the isolates were part of the ongoing North London isoniazid resistant outbreak (Ruddy et al. 2004) but further analysis revealed findings that were unusual in a normal outbreak setting. It is clear that the strains are closely related, as shown by two genotyping methods, but they were not identical, as one might expect.

The large proportion of drug resistant strains in this lineage indicated that the acquisition of resistance mutations had not had a detrimental effect on their ability to transmit between hosts. This was confirmed by *in vitro* fitness assays that demonstrated that the generation time of the resistant strains was not significantly longer than that of the fully susceptible strains in the lineage as well as the laboratory reference strain *M. tuberculosis* $H_{37}Rv$. In depth analysis of the genotyping, along with sequence analysis of drug resistance genes showed how the strains have evolved independently during the course of the outbreak whilst remaining competitive (figure 6.4). The presence of more than one isoniazid conferring mutation in the lineage, alongside the addition and deletion of the insertion sequence IS*6110*, clearly demonstrates the independent evolution of these organisms. Unusual *rpo*B mutations have already been documented in MDR strains within the outbreak (Jenkins et al. 2005), so it would be interesting to ascertain the nature of all the mutations causing resistance in these isolates to examine how divergent they were.

The North London isoniazid resistant outbreak strains that were studied in this work were isolated between 2002 and 2007, and there are a total of eight insertions and seven deletions of IS*6110* during this time. It appears that the strains have gained and lost the insertion sequence IS*6110* at a higher rate than described elsewhere (de Boer et al. 1999; Yeh et al. 1998). It is possible that these strains have evolved more rapidly, whilst maintaining their competitiveness and acquiring resistance mutations. This may be due to selective pressure during the outbreak, or maybe because of a permissive genotype.

The analysis of mutation rates and DNA repair in this group of strains may shed some light on this situation.

The organisms identified in this large outbreak are clearly successful. These genotypes are associated with the acquisition of drug-resistance conferring mutations with minimal fitness cost. Other genotypes, notably the Beijing/W family, are associated with a particular phenotype, such as progression to active disease (de Jong et al. 2008), pulmonary cavitation (Kato-Maeda et al. 2001), host immune response (Lopez et al. 2003; Theus, Cave, & Eisenach 2005), growth rates in ex-vivo monocytes (Theus, Cave, & Eisenach 2005) and macrophages (Zhang JID 1999), the ability to cause disseminated disease (Caws et al. 2008; Garcia de Viedma et al. 2005; Thwaites et al. 2008) and pulmonary smear negativity (Dale et al. 2005). Additionally, Beijing strains have been associated with radiologically advanced disease (Drobniewski et al 2005). The lineage of strains identified here has an associated phenotype that confers and advantage. They have therefore taken the opportunity to propagate within London and beyond. Similar analyses on related outbreak strains that have been isolated more recently would be fascinating.

The North London isoniazid resistant outbreak consists of large numbers of Caucasian and Black Caribbean patients (Ruddy et al. 2004), which is divergent from national UK data (HPA 2009). Gagneux and colleagues (Gagneux et al. 2006) noted that patients infected with isolates in the European-American group were more likely to give rise to secondary cases than other groups. Baker and colleagues (Baker et al 2004) and Brown and colleagues (Brown et al 2010) additional found that *M. tuberculosis* lineage correlate

with country of birth. Given that is has been shown that phylogeographic groups align with the country of birth of TB patients (Dale et al. 2009) it is possible that these outbreak strains may be European-American in origin.

In summary, it appears that this lineage, which is associated with the North London isoniazid resistant outbreak, comprises a heterogenous collection of closely related strains. These strains appear to have thrived in the community, generating large numbers of clinical cases whilst evolving independently via the acquisition of antimicrobial resistance and the alteration of their genotypes.

The hypothesis that bacterial genotype is related to disease phenotype was further examined in chapter 5. Patients who were subsequently culture positive for *M. tuberculosis* underwent a novel immunological test (Breen et al. 2005), with varying results. The assay measured the immunological response to a single *M. tuberculosis* antigen (ESAT6) *in vitro*. The finding that not all patients gave a positive reaction was intriguing and the possibility that this was a phenomenon was associated to a clonal *M. tuberculosis* genotype was investigated.

The differences in IFNγ production by the patients' T cells were not related to the genotype of the infecting organism. There was no relationship between the isolates in this patient group according the analysis of IS*6110* RFLP (figure 5.2) and all strains possessed a unaltered copy of the gene encoding ESAT6 (*esx*A) (table 5.3). All strains except one exhibited minimal changes in the regulation of this gene when compared to a

reference strain. One isolate showed considerable up-regulation of this gene, but this was not linked to the infected patient's IFNγ response (table 5.6). The IFNγ response in the patient from whom this isolate was cultured was the highest amongst the cohort, but was not statistically significantly higher than those of the other patients. Additional work is needed to further elucidate this finding. It may be possible that the growth rate of this strain is significantly different to the others, that is, it may have reached stationary phase significantly earlier than day 28. This warrants further investigation of the expression of other genes associated with this isolate to determine if this level of up-regulation is a common factor. This may be achieved by using RT PCR to look at a range of genes in this isolate, or by utilising a microarray. Further investigative avenues for these strains include the analysis of the other genes within the RD1, including *lhp* that encodes the co-expressed CFP-10. It may be of interest to examine the expression of the genes encoding the other ESAT-6 homologues (Tekaia et al. 1999). To assess whether the isolate that showed high levels of upregualtion of *esx*A (04:391) has generally high levels of transcription it may be of interest to investigate a number of housekeeping genes, including other sigma factors and genes encoding heat shock proteins.

The interpretation of RT PCR data in this setting prompts some debate regarding the reference genes used and the manor in which the data is analysed. This fascinating group of patients raised a question regarding a single antigen of *M. tuberculosis.* The design of an RT PCR for this gene, along with appropriate controls, allowed this question to be answered directly. Alternatively, it may be beneficial to use a microarray when investigating the regulation of multiple gene targets. In the case of the isolate that was up-

regulated significantly, the utilisation of a microarray would provide an insight to a larger proportion of the genome, which may begin to answer some of the questions raised. It is useful to note that regardless of the method used, RT PCR or microarrays, the careful selection of reference genes is imperative. The utilisation of a gene that is consistently expressed at a level that allows for variation in the target gene to be recorded is vital (table 5.4). As seen in chapter 5, the selection of a gene that is consistently expressed at a high level can make an assay insensitive.

The finding that not all patients in this group demonstrated a positive IGRA result despite having culture confirmed TB is intriguing. This suggests that IGRA must be interpreted as part of the whole picture of TB diagnostics and not alone. As with any diagnostic tool, whether it be laboratory based, radiological or clinical, the sum of these is far greater than the individual parts.

From these data it was concluded that the hosts' genotype may play a more significant role in this setting. Different disease phenotypes have been shown in patients and animal models. For example, increased susceptibility to disease has been shown in animal models when defects in CD4 T cell function (Saunders et al. 2002; Scanga et al. 2000), or with IFN-γ expression or recognition (Cooper et al. 1993; Flynn et al. 1993). Increased susceptibility has also been seen in such patients (Havlir & Barnes 1999; Ottenhoff et al. 2005). These associations between disease phenotype with host genotype indicate that the variations in immune response in these patients may be driven by anomalies within the immune systems of the patients. This warrants further investigation in the areas of patient

HLA typing and immune function tests. HLA typing was not performed on the patients in this dataset as the emphasis was on the biology of the infecting organism. Additionally, ethics approval was not in place to perform this analysis retrospectively on any patient material that was collected for diagnostic purposes.

In attempting to understand the relationship between pathogen genotype and disease phenotype, a number of methodologies were utilised. The technologies used were appropriate at the beginning of this work, but of course technological advances have not ceased in the following years. The development of next generation sequencing makes it possible to sequence an entire bacterial genome in hours, rather than weeks, at a fraction of the cost (Metzker 2010). As this type of technology become more prevalent, it is worth noting that the amount of data generated is vast. Indeed, the techniques involved in sequencing megabases of DNA are relatively simple when compared to the data analysis software and hardware that is required subsequently. The phylogeographic evolution of *M. tuberculosis* has subsequently been described by analysing SNPs and LSPs (Gagneux & Small 2007). These polymorphisms may be considered as surrogates for the evolutionary process, but the technology to study every SNP in an entire genome is very real. The application of such technology to this dataset may answer some of the unanswered questions.

As with all new technologies, it will take some time to establish the real meaning of the generated data. Whole genome sequencing has been used clinically to subdivide clusters of *M. tuberculosis* infection (Schurch et al. 2010). It has also been shown that the genome

of the reference strain *M. tuberculosis* $H_{37}Rv$ varies between different research centres (Ioerger et al. 2010). This finding may have an impact on wide-reaching aspects of tuberculosis research. This reference strain is used globally in the development of vaccines, drugs and diagnostics and it is therefore concerning that the strains held in different centres vary significantly. One may see different vaccine or drug responses with these different strains. It may therefore be necessary to agree on a consensus strain that is distributed globally to all relevant research centres.

It will take some time to establish the degree of difference expected between strains before the technology may be adopted widely. This will be achieved by sequencing the entire genome of multiple strains, which will take time and some existing data may need to be reviewed – indeed analysis of multiple isolates of *M. tuberculosis* $H_{37}Rv$ indicates that some bases were incorrectly sequenced originally and may need to be reissued (Ioerger et al. 2010). Whole genome analysis may play a role in the analysis of successful lineages, such as those identified in this work. Novel virulence factors and potential drug targets may be identified by the large scale adoption of this technology.

Additionally, genotyping can play an important role in the development of both new antimicrobial agents and vaccines. As demonstrated in this thesis, clinical strains can vary widely in many respects. The genomic differences which appear through evolutionary processes and are partially demonstrated with genotyping may alter many crucial aspects of the organisms' physiology. For example, a novel antimicrobial agent may not be considered to be active against *M. tuberculosis* until it has been shown to be effective

against a range of genotypically distinct strains. Furthermore, clinical trials involving novel agents and regimens are using wide-reaching geographical locations partly for this reason. Likewise, multiple unrelated strains should be screened when investigating potential vaccine targets and when assessing the protective effects of novel vaccines.

The utilisation of *M. tuberculosis* genotyping arguably provides the greatest benefit in a clinical setting. In chapter 7, the rapid confirmation of four genuine cases of tuberculosis enabled appropriate treatment to be commenced promptly, which prevents ongoing transmission. Additionally, the instigation of contact tracing allows the identification of actively and latently infected individuals and their subsequent management. The collaboration between the laboratory and clinicians was further demonstrated in the investigation of a group of patients presenting with a paradoxical reactions and also in the case of household transmission. Such continued collaboration will further enhance both the understanding of the biology of the *M. tuberculosis* and the clinical presentation of its hosts. These two disciplines must work together to help to eradicate the disease.

In conclusion, genotyping *M. tuberculosis* plays a pivotal role in the holistic approach to understanding tuberculosis. The relationship between genotype and phenotype is complicated, but the identification of outbreaks and over-represented lineages of strains can further add to the understanding of the evolutionary success of the organism. Prospective, real-time genotyping can enable the monitoring and control of disease as well as its prevention. Bridging the gaps between the laboratory and clinical practice is essential for the eradication of this global pathogen.

References

Aber VR. Allen BW, Mitchison DA, Ayuma P, Edwards EA & Keyes AB. 1980, "Quality control in tuberculosis bacteriology. 1. Laboratory studies on isolated positive cultures and the efficiency of direct smear examination." Tubercle 61(3), 123-33.

Ahmed, A. B., Abubakar, I., Delpech, V., Lipman, M., Boccia, D., Forde, J., Antoine, D & Watson, J. M. 2007, "The growing impact of HIV infection on the epidemiology of tuberculosis in England and Wales: 1999 2003", Thorax, vol. 62, no. 8, pp. 672-676.

Alland, D., Kalkut, G. E., Moss, A. R., McAdam, R. A., Hahn, J. A., Bosworth, W., Drucker, E., & Bloom, B. R. 1994, "Transmission of tuberculosis in New York City. An analysis by DNA fingerprinting and conventional epidemiologic methods", N.Engl.J.Med., vol. 330, no. 24, pp. 1710-1716.

American Thoracic Society 2000, "Targeted tuberculin testing and treatment of latent tuberculosis infection. Am.J.Respir.Crit Care Med., vol. 161, no. 4 Pt 2, p. S221-S247.

Andersen, P, Askgaard D, Ljungqvist L, Bennedsen J, Heron I. 1991, "Proteins released from *Mycobacterium tuberculosis* during growth." Infect.Immun. 59.6: 1905-10.

Andersen P, Andersen AB, Sørensen AL, Nagai S, 1995, "Recall of long-lived immunity to *Mycobacterium tuberculosis* infection in mice." The Journal of Immunology 154.7: 3359-72.

Andersson, D. I. & Levin, B. R. 1999, "The biological cost of antibiotic resistance", Curr.Opin.Microbiol., vol. 2, no. 5, pp. 489-493.

Andersson, D. I. & Hughes, D. 2010, "Antibiotic resistance and its cost: is it possible to reverse resistance?", Nat Rev Microbiol., vol. 8, no. 4, pp. 260-271.

Anonymous 1997. Multiple Misdiagnoses of Tuberculosis Resulting from Laboratory Error -- Wisconsin,.Centers for Disease Control and Prevention. 46[34], 797-801.

Anonymous 2004. Tuberculosis Associated with Blocking Agents Against Tumor Necrosis Factor-Alpha — California, 2002–2003. Centers for Disease Control and Prevention. Morbidity and Mortality Weekly Report 53[30].

Anonymous 2006. Emergence of *Mycobacterium tuberculosis* with Extensive Resistance to Second-Line Drugs - Worldwide, 2000-2004. Centers for Disease Control and Prevention. Morbidity and Mortality Weekly Report 55[11].

Anonymous, 2007. Extensively Drug-Resistant Tuberculosis - United States, 1993-2006. Centers for Disease Control and Prevention. Morbidity and Mortality Weekly Report 56[11].

Ansari, S., Thomas, S., Campbell, I. A., Furness, L., & Evans, M. R. 1998, "Refined tuberculosis contact tracing in a low incidence area", Respir.Med., vol. 92, no. 9, pp. 1127-1131.

Ansorge, W. J. 2009. Next-generation DNA sequencing techniques. New Biotechnology 25, 195-203.

Antoine D. Maguire H and Story A. 2006. Epidemiology and response to the growing problem of tuberculosis in London. Eurosurveillance 3, 25-28.

Antonucci, G., Girardi, E., Raviglione, M. C., & Ippolito, G. 1995, "Risk factors for tuberculosis in HIV-infected persons. A prospective cohort study, JAMA, vol. 274, no. 2, pp. 143-148.

Appelmelk, B. J., van, D., I, van Vliet, S. J., Vandenbroucke-Grauls, C. M., Geijtenbeek, T. B., & van, K. Y. 2003, "Cutting edge: carbohydrate profiling identifies new pathogens that interact with dendritic cell-specific ICAM-3-grabbing nonintegrin on dendritic cells", J.Immunol., vol. 170, no. 4, pp. 1635-1639.

Arend, S. M., Thijsen, S. F., Leyten, E. M., Bouwman, J. J., Franken, W. P., Koster, B. F., Cobelens, F. G., van Houte, A. J., & Bossink, A. W. 2007a, "Comparison of two interferon-gamma assays and tuberculin skin test for tracing tuberculosis contacts", Am.J.Respir.Crit Care Med., vol. 175, no. 6, pp. 618-627.

Arend, S. M., Thijsen, S. F. T., Leyten, E. M. S., Bouwman, J. J. M., Franken, W. P. J., Koster, B. F. P. J., Cobelens, F. G. J., van Houte, A. J., & Bossink, A. W. J. 2007b, "Comparison of Two Interferon-{gamma} Assays and Tuberculin Skin Test for Tracing Tuberculosis Contacts", American Journal of Respiratory and Critical Care Medicine, vol. 175, no. 6, pp. 618-627.

Atun RA, Lebcir R, Drobniewski F, Coker R. 2005. Impact of an effective multidrug-resistant tuberculosis control programme in the setting of an immature HIV epidemic: system dynamics simulation model. International Journal of STD & AIDS 16: 560–570

Aziz, M. A. & Wright, A. 2005, "The World Health Organization/International Union against Tuberculosis and Lung Disease Global Project on Surveillance for Anti-Tuberculosis Drug Resistance: A Model for Other Infectious Diseases", Clinical Infectious Diseases, vol. 41, no. S4, p. S258-S262.

Baker, L., Brown, T., Maiden, M. C., & Drobniewski, F. 2004, "Silent nucleotide polymorphisms and a phylogeny for *Mycobacterium tuberculosis*", Emerg.Infect.Dis., vol. 10, no. 9, pp. 1568-1577.

Bamford, A. R, Crook AM, Clark JE, Nademi Z, Dixon G, Paton JY, Riddell A, Drobniewski F, Riordan A, Anderson ST, Williams A, Walters S, Kampmann B. 2010. "Comparison of interferon-{gamma} release assays and tuberculin skin test in predicting active tuberculosis (TB) in children in the UK: a paediatric TB network study." Arch.Dis.Child 95.3: 180-86.

Bannister B, Gillespie S and Jones J 2006. Infection, Microbiology and Management 3rd Edition. Blackwell Publishing.

Barry, S. M., Lipman MC, Bannister B, Johnson MA, Janossy G. 2003. "Purified protein derivative-activated type 1 cytokine-producing CD4+ T lymphocytes in the lung: a characteristic feature of active pulmonary and nonpulmonary tuberculosis." J.Infect.Dis. 187.2: 243-50.

Barlow, R. E. L., Gascoyne-Binzi, D. M., Gillespie, S. H., Dickens, A., Qamer, S., & Hawkey, P. M. 2001, "Comparison of Variable Number Tandem Repeat and IS6110-Restriction Fragment Length Polymorphism Analyses for Discrimination of High- and Low-Copy-Number IS6110 *Mycobacterium tuberculosis* Isolates", Journal of Clinical Microbiology, vol. 39, no. 7, pp. 2453-2457.

Barnes, P. F., Yang, Z., Preston-Martin, S., Pogoda, J. M., Jones, B. E., Otaya, M., Eisenach, K. D., Knowles, L., Harvey, S., & Cave, M. D. 1997, "Patterns of tuberculosis transmission in Central Los Angeles", JAMA, vol. 278, no. 14, pp. 1159-1163.

Barnes, P. F. & Cave, M. D. 2003, "Molecular epidemiology of tuberculosis", N.Engl.J.Med., vol. 349, no. 12, pp. 1149-1156.

Basu, S. K., Kumar D, Singh DK, Ganguly N, Siddiqui Z, Rao KV, Sharma P. 2006. "*Mycobacterium tuberculosis* secreted antigen (MTSA-10) modulates macrophage function by redox regulation of phosphatases." FEBS J. 273.24 (2006): 5517-34.

Bauer, J., Thomsen VO, Poulsen S, Andersen AB. 1997. "False-positive results from cultures of *Mycobacterium tuberculosis* due to laboratory cross-contamination confirmed by restriction fragment length polymorphism." Journal of Clinical Microbiology 35.4: 988-91.

Becerra, M. C., Pachao-Torreblanca IF, Bayona J, Celi R, Shin SS, Kim JY, Farmer PE, Murray M. 2005. "Expanding tuberculosis case detection by screening household contacts." Public Health Rep. 120.3: 271-77.

Behr, M. A., Wilson, M. A., Gill, W. P., Salamon, H., Schoolnik, G. K., Rane, S., & Small, P. M. 1999, "Comparative Genomics of BCG Vaccines by Whole-Genome DNA Microarray", Science, vol. 284, no. 5419, pp. 1520-1523.

Bemer P, Palicova F, Rusch-Gerdes S, Drugeon, HB, Pfyffer GE. 2002. Multicenter Evaluation of Fully Automated BACTEC Mycobacteria Growth Indicator Tube 960 System for Susceptibility Testing of *Mycobacterium tuberculosis.* Journal of Clinical Microbiology 40(1): 150-154

Bennett, A. F., Dao, K. M., & Lenski, R. E. 1990, "Rapid evolution in response to high-temperature selection", Nature, vol. 346, no. 6279, pp. 79-81.

Bercovier, H., Kafri, O., & Sela, S. 1986, "Mycobacteria possess a surprisingly small number of ribosomal RNA genes in relation to the size of their genome", Biochem.Biophys.Res.Commun., vol. 136, no. 3, pp. 1136-1141.

Berthet, F. X., Rasmussen, P. B., Rosenkrands, I., Andersen, P., & Gicquel, B. 1998, "A *Mycobacterium tuberculosis* operon encoding ESAT-6 and a novel low-molecular-mass culture filtrate protein (CFP-10)", Microbiology, vol. 144 (Pt 11), pp. 3195-3203.

Betts, J. C., McLaren A, Lennon MG, Kelly FM, Lukey PT, Blakemore SJ, Duncan K. 2003. "Signature gene expression profiles discriminate between isoniazid-, thiolactomycin-, and triclosan-treated *Mycobacterium tuberculosis*." Antimicrobial Agents and Chemotherapy 47 (9): 2903-13.

Bifani, P. J., Mathema B, Liu Z, Moghazeh SL, Shopsin B, Tempalski B, Driscol J, Frothingham R, Musser JM, Alcabes P, Kreiswirth BN. 1999. "Identification of a W variant outbreak of *Mycobacterium tuberculosis* via population-based molecular epidemiology." JAMA 282 (24): 2321-27.

Bifani, P. J., Mathema, B., Kurepina, N. E., & Kreiswirth, B. N. 2002, "Global dissemination of the *Mycobacterium tuberculosis* W-Beijing family strains", Trends in Microbiology, vol. 10, no. 1, pp. 45-52.

Billington, O. J., McHugh, T. D., & Gillespie, S. H. 1999, "Physiological cost of rifampin resistance induced in vitro in *Mycobacterium tuberculosis*", Antimicrobial Agents and Chemotherapy, vol. 43, no. 8, pp. 1866-1869.

Bjorkman, J., Hughes, D., & Andersson, D. I. 1998, "Virulence of antibiotic-resistant Salmonella typhimurium", Proc.Natl.Acad.Sci.U.S.A, vol. 95, no. 7, pp. 3949-3953.

Bjorkman, J., Samuelsson, P., Andersson, D. I., & Hughes, D. 1999, "Novel ribosomal mutations affecting translational accuracy, antibiotic resistance and virulence of Salmonella typhimurium", Mol.Microbiol., vol. 31, no. 1, pp. 53-58.

Bocchino, M., Galati, D., Sanduzzi, A., Colizzi, V., Brunetti, E., & Mancino, G. 2005, "Role of mycobacteria-induced monocyte/macrophage apoptosis in the pathogenesis of human tuberculosis", Int.J.Tuberc.Lung Dis., vol. 9, no. 4, pp. 375-383.

Boehme CC, Nabeta P, , Hillemann D, Nicol MP, Shenai S, Krapp F, Allen J, Tahirli R, Blakemore R, Rustomjee R, Milovic A, Jones M, O'Brien SM, Persing DH, Ruesch-Gerdes S, Gotuzzo E, Rodrigues C, Alland D, Perkins D. 2010. Rapid Molecular Detection of Tuberculosis and Rifampin Resistance NEJM 363(11): 1005-1015.

Bonora, S. & Di, P. G. 2008, "Interactions between antiretroviral agents and those used to treat tuberculosis", Curr.Opin.HIV.AIDS, vol. 3, no. 3, pp. 306-312.

Borrell, S. & Gagneux, S. 2009, "Infectiousness, reproductive fitness and evolution of drug-resistant *Mycobacterium tuberculosis*", Int.J.Tuberc.Lung Dis., vol. 13, no. 12, pp. 1456-1466.

Bottger, E. C., Springer, B., Pletschette, M., & Sander, P. 1998, "Fitness of antibiotic-resistant microorganisms and compensatory mutations", Nat Med., vol. 4, no. 12, pp. 1343-1344.

Bouma, J. E. & Lenski, R. E. 1988, "Evolution of a bacteria/plasmid association", Nature, vol. 335, no. 6188, pp. 351-352.

Brandt, L., Oettinger T, Holm A, Andersen AB, Andersen P. 1996. "Key epitopes on the ESAT-6 antigen recognized in mice during the recall of protective immunity to *Mycobacterium tuberculosis*." J.Immunol. 157 (8): 3527-33.

Breathnach, A. S., de, R. A., Holdsworth, G. M., Bateman, N. T., O'Sullivan, D. G., Rees, P. J., Snashall, D., Milburn, H. J., Peters, B. S., Watson, J., Drobniewski, F. A., & French, G. L. 1998, "An outbreak of multi-drug-resistant tuberculosis in a London teaching hospital", J.Hosp.Infect., vol. 39, no. 2, pp. 111-117.

Breen, R. A., Smith, C. J., Bettinson, H., Dart, S., Bannister, B., Johnson, M. A., & Lipman, M. C. 2004, "Paradoxical reactions during tuberculosis treatment in patients with and without HIV co-infection", Thorax, vol. 59, no. 8, pp. 704-707.

Breen, R. A., Hopkins S, Janossy G, Lipman MC, Cropley I. 2005. "Local IFN-{gamma} responses in TB." Thorax 60 (9): 788-89.

Breen, R. A., Miller, R. F., Gorsuch, T., Smith, C. J., Ainsworth, J., Ballinger, J., Swaden, L., Cropley, I., Johnson, M. A., & Lipman, M. C. 2006, "Virological response to highly active antiretroviral therapy is unaffected by antituberculosis therapy", J.Infect.Dis., vol. 193, no. 10, pp. 1437-1440.

Breen, R. A. M., Barry SM, Smith CJ, Shorten RJ, Dilworth JP, Cropley I, McHugh TD, Gillespie SH, Janossy G, Lipman MC. 2008a. "Clinical application of a rapid lung-orientated immunoassay in individuals with possible tuberculosis." Thorax 63 (1): 67-71.

Breen, R. A., Leonard, O., Perrin, F. M., Smith, C. J., Bhagani, S., Cropley, I., & Lipman, M. C. 2008b, "How good are systemic symptoms and blood inflammatory markers at detecting individuals with tuberculosis?", Int.J.Tuberc.Lung Dis., vol. 12, no. 1, pp. 44-49.

Brennan, P. J. 2003, "Structure, function, and biogenesis of the cell wall of *Mycobacterium tuberculosis*", Tuberculosis (Edinb.), vol. 83, no. 1-3, pp. 91-97.

Briken, V., Porcelli, S. A., Besra, G. S., & Kremer, L. 2004, "Mycobacterial lipoarabinomannan and related lipoglycans: from biogenesis to modulation of the immune response", Mol.Microbiol, vol. 53, no. 2, pp. 391-403.

British HIV Association 2009. British HIV Association guidelines for the treatment of TB/ HIV co-infection 2009.

British Medical Research Council 1948, "Streptomycin treatment of pulmonary tuberculosis", Br.Med.J., vol. 2, no. 4582, pp. 769-782.

British Thoracic Society 2005, "BTS recommendations for assessing risk and for managing *Mycobacterium tuberculosis* infection and disease in patients due to start anti-TNF-alpha treatment", Thorax, vol. 60, no. 10, pp. 800-805.

Brodin, P., Majlessi, L., Marsollier, L., de Jonge, M. I., Bottai, D., Demangel, C., Hinds, J., Neyrolles, O., Butcher, P. D., Leclerc, C., Cole, S. T., & Brosch, R. 2006, "Dissection of ESAT-6 system 1 of *Mycobacterium tuberculosis* and impact on immunogenicity and virulence", Infect.Immun., vol. 74, no. 1, pp. 88-98.

Brodin, P., Rosenkrands I, Andersen P, Cole ST, Brosch R. 2004. "ESAT-6 proteins: protective antigens and virulence factors?" Trends Microbiol. 12 (11): 500-08.

Brosch, R., Gordon, S. V., Marmiesse, M., Brodin, P., Buchrieser, C., Eiglmeier, K., Garnier, T., Gutierrez, C., Hewinson, G., Kremer, K., Parsons, L. M., Pym, A. S., Samper, S., van Soolingen, D., & Cole, S. T. 2002, "A new evolutionary scenario for the *Mycobacterium tuberculosis* complex", Proceedings of the National Academy of Sciences of the United States of America, vol. 99, no. 6, pp. 3684-3689.

Brosch, R., Philipp, W. J., Stavropoulos, E., Colston, M. J., Cole, S. T., & Gordon, S. V. 1999, "Genomic analysis reveals variation between *Mycobacterium tuberculosis* H37Rv and the attenuated M. tuberculosis H37Ra strain", Infect.Immun., vol. 67, no. 11, pp. 5768-5774.

Brown T, Nikolayevskyy V, Velji P, and Drobniewski F. 2010. Associations between Mycobacterium tuberculosis Strains and Phenotypes> Emerging Infectious Disaeses 16(2): 272-280

Brudey, K., Driscoll, J. R., Rigouts, L., Prodinger, W. M., Gori, A., Al-Hajoj, S. A., Allix, C., Aristimuno, L., Arora, J., Baumanis, V., Binder, L., Cafrune, P., Cataldi, A., Cheong, S., Diel, R., Ellermeier, C., Evans, J. T., Fauville-Dufaux, M., Ferdinand, S., Garcia, d., V, Garzelli, C., Gazzola, L., Gomes, H. M., Guttierez, M. C., Hawkey, P. M., van Helden, P. D., Kadival, G. V., Kreiswirth, B. N., Kremer, K., Kubin, M., Kulkarni, S. P., Liens, B., Lillebaek, T., Ho, M. L., Martin, C., Martin, C., Mokrousov, I., Narvskaia, O., Ngeow, Y. F., Naumann, L., Niemann, S., Parwati, I., Rahim, Z., Rasolofo-Razanamparany, V., Rasolonavalona, T., Rossetti, M. L., Rusch-Gerdes, S., Sajduda, A., Samper, S., Shemyakin, I. G., Singh, U. B., Somoskovi, A., Skuce, R. A., van, S. D., Streicher, E. M., Suffys, P. N., Tortoli, E., Tracevska, T., Vincent, V., Victor, T. C., Warren, R. M., Yap, S. F., Zaman, K., Portaels, F., Rastogi, N., & Sola, C. 2006, "*Mycobacterium tuberculosis* complex genetic diversity: mining the fourth international spoligotyping database (SpolDB4) for classification, population genetics and epidemiology", BMC Microbiol., vol. 6, p. 23.

Burman, W. J. & Reves, R. R. 2000, "Review of false-positive cultures for *Mycobacterium tuberculosis* and recommendations for avoiding unnecessary treatment", Clin.Infect.Dis., vol. 31, no. 6, pp. 1390-1395.

Burns, J. L., Hedin, L. A., & Lien, D. M. 1989, "Chloramphenicol resistance in Pseudomonas cepacia because of decreased permeability", Antimicrobial Agents and Chemotherapy, vol. 33, no. 2, pp. 136-141.

Caccamo, N., Sireci, G., Meraviglia, S., Dieli, F., Ivanyi, J., & Salerno, A. 2006, "gammadelta T cells condition dendritic cells in vivo for priming pulmonary CD8 T cell responses against *Mycobacterium tuberculosis*", Eur.J.Immunol., vol. 36, no. 10, pp. 2681-2690.

Caley, M, Fowler T, Welch S, Wood A. 2010. "Risk of developing tuberculosis from a school contact: retrospective cohort study, United Kingdom, 2009." Euro.Surveill 15 (11).

Cameron, S. J. 1974, "Tuberculosis and the blood--a special relationship?", Tubercle., vol. 55, no. 1, pp. 55-72.

Camus, J. C., Pryor, M. J., Medigue, C., & Cole, S. T. 2002, "Re-annotation of the genome sequence of *Mycobacterium tuberculosis* H37Rv", Microbiology, vol. 148, no. 10, pp. 2967-2973.

Cave, M. D., Eisenach, K. D., McDermott, P. F., Bates, J. H., & Crawford, J. T. 1991, "IS6110: conservation of sequence in the *Mycobacterium tuberculosis* complex and its utilization in DNA fingerprinting", Mol.Cell Probes, vol. 5, no. 1, pp. 73-80.

Caws, M., Thwaites, G., Dunstan, S., Hawn, T. R., Lan, N. T., Thuong, N. T., Stepniewska, K., Huyen, M. N., Bang, N. D., Loc, T. H., Gagneux, S., van, S. D., Kremer, K., van der, S. M., Small, P., Anh, P. T., Chinh, N. T., Quy, H. T., Duyen, N. T., Tho, D. Q., Hieu, N. T., Torok, E., Hien, T. T., Dung, N. H., Nhu, N. T., Duy, P. M., van Vinh, C. N., & Farrar, J. 2008, "The influence of host and bacterial genotype on the development of disseminated disease with *Mycobacterium tuberculosis*", PLoS.Pathog., vol. 4, no. 3, p. e1000034.

Cetre, J. C., Nicolle, M. C., Salord, H., Perol, M., Tigaud, S., David, G., Bourjault, M., & Vanhems, P. 2005, "Outbreaks of contaminated broncho-alveolar lavage related to intrinsically defective bronchoscopes", J.Hosp.Infect., vol. 61, no. 1, pp. 39-45.

Champion, P. A., Stanley SA, Champion MM, Brown EJ, Cox JS. 2006. "C-terminal signal sequence promotes virulence factor secretion in *Mycobacterium tuberculosis*." Science 313 (5793): 1632-36

Chan, E. D. & Iseman, M. D. 2008, "Multidrug-resistant and extensively drug-resistant tuberculosis: a review", Curr.Opin.Infect.Dis., vol. 21, no. 6, pp. 587-595.

Chan, E. D., Laurel, V., Strand, M. J., Chan, J. F., Huynh, M. L., Goble, M., & Iseman, M. D. 2004, "Treatment and outcome analysis of 205 patients with multidrug-resistant tuberculosis", Am.J.Respir.Crit Care Med., vol. 169, no. 10, pp. 1103-1109.

Chapman, A. L., Munkanta M, Wilkinson KA, Pathan AA, Ewer K, Ayles H, Reece WH, Mwinga A, Godfrey-Faussett P, Lalvani A. 2002. "Rapid detection of active and latent tuberculosis infection in HIV-positive individuals by enumeration of *Mycobacterium tuberculosis*-specific T cells." AIDS 16 (17): 2285-93

Chaves, F., Yang, Z., el, H. H., Alonso, M., Burman, W. J., Eisenach, K. D., Dronda, F., Bates, J. H., & Cave, M. D. 1996, "Usefulness of the secondary probe pTBN12 in DNA fingerprinting of *Mycobacterium tuberculosis*", Journal of Clinical Microbiology, vol. 34, no. 5, pp. 1118-1123.

Chen, F. J. & Lo, H. J. 2003, "Molecular mechanisms of fluoroquinolone resistance", J.Microbiol.Immunol.Infect., vol. 36, no. 1, pp. 1-9.

Chocano-Bedoya, P. & Ronnenberg, A. G. 2009, "Vitamin D and tuberculosis", Nutr.Rev, vol. 67, no. 5, pp. 289-293.

Chroneou, A., Zimmerman, S. K., Cook, S., Willey, S., Eyre-Kelly, J., Zias, N., Shapiro, D. S., Beamis, J. F., Jr., & Craven, D. E. 2008, "Molecular typing of Mycobacterium chelonae isolates from a pseudo-outbreak involving an automated bronchoscope washer", Infect.Control Hosp.Epidemiol., vol. 29, no. 11, pp. 1088-1090.

Chua, J., Vergne, I., Master, S., & Deretic, V. 2004, "A tale of two lipids: *Mycobacterium tuberculosis* phagosome maturation arrest", Curr.Opin.Microbiol., vol. 7, no. 1, pp. 71-77.

Coetzee, N., Liu WK, Astbury N, Williams P, Robinson S, Afza M, Duggal HV. 2009. "Legionnaires' disease cluster linked to a metal product aqueous pre-treatment process, Staffordshire, England, May 2008." Euro.Surveill 14 (40).

Cohn, M. L., Kovitz, C., Oda, U., & Middlebrook, G. 1954, "Studies on isoniazid and tubercle bacilli. II. The growth requirements, catalase activities, and pathogenic properties of isoniazid-resistant mutants", Am.Rev Tuberc., vol. 70, no. 4, pp. 641-664.

Colangeli, R., Spencer JS, Bifani P, Williams A, Lyashchenko K, Keen MA, Hill PJ, Belisle J, Gennaro ML. 2000. "MTSA-10, the product of the Rv3874 gene of *Mycobacterium tuberculosis*, elicits tuberculosis-specific, delayed-type hypersensitivity in guinea pigs." Infect.Immun. 68 (2): 990-93.

Colditz, G. A., Berkey, C. S., Mosteller, F., Brewer, T. F., Wilson, M. E., Burdick, E., & Fineberg, H. V. 1995, "The efficacy of bacillus Calmette-Guerin vaccination of newborns and infants in the prevention of tuberculosis: meta-analyses of the published literature", Pediatrics, vol. 96, no. 1 Pt 1, pp. 29-35.

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, vol. 271, no. 9, pp. 698-702.

Cole ST. Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, Gordon SV, Eiglmeier K, Gas S, Barry CE 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 MA, Rajandream MA, Rogers J, Rutter S, Seeger K, Skelton J, Squares R, Squares S, Sulston JE, Taylor K, Whitehead S, Barrell BG. 1998. "Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence." Nature 393, 537-544.

Cook, V. J., Sun, S. J., Tapia, J., Muth, S. Q., Arguello, D. F., Lewis, B. L., Rothenberg, R. B., & McElroy, P. D. 2007, "Transmission network analysis in tuberculosis contact investigations", J.Infect.Dis., vol. 196, no. 10, pp. 1517-1527

Cooksey, R. C., Morlock, G. P., McQueen, A., Glickman, S. E., & Crawford, J. T. 1996, "Characterization of streptomycin resistance mechanisms among *Mycobacterium tuberculosis* isolates from patients in New York City", Antimicrobial Agents and Chemotherapy, vol. 40, no. 5, pp. 1186-1188.

Cooper, A. M., Dalton, D. K., Stewart, T. A., Griffin, J. P., Russell, D. G., & Orme, I. M. 1993, "Disseminated tuberculosis in interferon gamma gene-disrupted mice", J.Exp.Med., vol. 178, no. 6, pp. 2243-2247.

Cooper, A. M. 2009, "Cell-mediated immune responses in tuberculosis", Annu.Rev Immunol., vol. 27, pp. 393-422.

Crofton, J. & Mitchison, D. A. 1948, "Streptomycin resistance in pulmonary tuberculosis", Br.Med.J., vol. 2, no. 4588, pp. 1009-1015.

Dahle, U. R., Sandven, P., Heldal, E., & Caugant, D. A. 2003, "Continued Low Rates of Transmission of *Mycobacterium tuberculosis* in Norway", Journal of Clinical Microbiology, vol. 41, no. 7, pp. 2968-2973.

Dale, J. W., Brittain, D., Cataldi, A. A., Cousins, D., Crawford, J. T., Driscoll, J., Heersma, H., Lillebaek, T., Quitugua, T., Rastogi, N., Skuce, R. A., Sola, C., van, S. D., & Vincent, V. 2001, "Spacer oligonucleotide typing of bacteria of the *Mycobacterium tuberculosis* complex: recommendations for standardised nomenclature", Int.J.Tuberc.Lung Dis., vol. 5, no. 3, pp. 216-219.

Dale, J. W., Al-Ghusein, H., Al-Hashmi, S., Butcher, P., Dickens, A. L., Drobniewski, F., Forbes, K. J., Gillespie, S. H., Lamprecht, D., McHugh, T. D., Pitman, R., Rastogi, N., Smith, A. T., Sola, C., & Yesilkaya, H. 2003a, "Evolutionary relationships among strains of *Mycobacterium tuberculosis* with few copies of IS6110", The Journal of Bacteriology, vol. 185, no. 8, pp. 2555-2562.

Dale, J. W., Al-Ghusein, H., Al-Hashmi, S., Butcher, P., Dickens, A. L., Drobniewski, F., Forbes, K. J., Gillespie, S. H., Lamprecht, D., McHugh, T. D., Pitman, R., Rastogi, N., Smith, A. T., Sola, C., & Yesilkaya, H. 2003b, "Evolutionary Relationships among

Strains of *Mycobacterium tuberculosis* with Few Copies of IS6110", The Journal of Bacteriology, vol. 185, no. 8, pp. 2555-2562.

Dale, J. W., Bothamley, G. H., Drobniewski, F., Gillespie, S. H., McHugh, T. D., & Pitman, R. 2005a, "Origins and properties of *Mycobacterium tuberculosis* isolates in London", J.Med.Microbiol., vol. 54, no. Pt 6, pp. 575-582.

Dale, J. W., Bothamley, G. H., Drobniewski, F., Gillespie, S. H., McHugh, T. D., & Pitman, R. 2005b, "Origins and properties of *Mycobacterium tuberculosis* isolates in London", Journal of Medical Microbiology, vol. 54, no. 6, pp. 575-582.

Daley, C. L., Small, P. M., Schecter, G. F., Schoolnik, G. K., McAdam, R. A., Jacobs, W. R., Jr., & Hopewell, P. C. 1992, "An outbreak of tuberculosis with accelerated progression among persons infected with the human immunodeficiency virus. An analysis using restriction-fragment-length polymorphisms", N.Engl.J.Med., vol. 326, no. 4, pp. 231-235.

David, H. L. 1970, "Probability distribution of drug-resistant mutants in unselected populations of *Mycobacterium tuberculosis*", Appl.Microbiol., vol. 20, no. 5, pp. 810-814.

Davies, J. 1994, "Inactivation of antibiotics and the dissemination of resistance genes", Science, vol. 264, no. 5157, pp. 375-382.

Davies, P. D. & Cullen, D. 2008, "Drug resistant tuberculosis: Service for drug resistant tuberculosis exists in the UK", BMJ, vol. 336, no. 7657, p. 1324.

de Boer, A. S., Borgdorff, M. W., de Haas, P. E., Nagelkerke, N. J., van Embden, J. D., & van, S. D. 1999, "Analysis of rate of change of IS6110 RFLP patterns of *Mycobacterium tuberculosis* based on serial patient isolates", J.Infect.Dis., vol. 180, no. 4, pp. 1238-1244.

de Boer, Annette S., Blommerde B, de Haas PE, Sebek MM, Lambregts-van Weezenbeek KS, Dessens M, van Soolingen D. 2002. "False-Positive *Mycobacterium tuberculosis* Cultures in 44 Laboratories in The Netherlands (1993 to 2000): Incidence, Risk Factors, and Consequences." Journal of Clinical Microbiology 40 (11): 4004-09.

de C Ramos, M, Soini H, Roscanni GC, Jaques M, Villares MC, Musser JM. 1999. "Extensive cross-contamination of specimens with *Mycobacterium tuberculosis* in a reference laboratory." Journal of Clinical Microbiology 37 (4): 916-19.

de Jong R., Altare, F., Haagen, I. A., Elferink, D. G., Boer, T., van, B., V, Kabel, P. J., Draaisma, J. M., van Dissel, J. T., Kroon, F. P., Casanova, J. L., & Ottenhoff, T. H. 1998, "Severe mycobacterial and Salmonella infections in interleukin-12 receptor-deficient patients", Science, vol. 280, no. 5368, pp. 1435-1438.

de Jong, B. C., Hill, P. C., Aiken, A., Awine, T., Antonio, M., Adetifa, I. M., Jackson-Sillah, D. J., Fox, A., Deriemer, K., Gagneux, S., Borgdorff, M. W., McAdam, K. P., Corrah, T., Small, P. M., & Adegbola, R. A. 2008, "Progression to active tuberculosis,

but not transmission, varies by *Mycobacterium tuberculosis* lineage in The Gambia", J.Infect.Dis., vol. 198, no. 7, pp. 1037-1043.

de Vries G., van Hest RA, Burdo CC, van Soolingen D, Richardus JH. 2009. "A *Mycobacterium tuberculosis* cluster demonstrating the use of genotyping in urban tuberculosis control." BMC Infect.Dis.9: 151

Deretic, V., Singh, S., Master, S., Harris, J., Roberts, E., Kyei, G., Davis, A., de, H. S., Naylor, J., Lee, H. H., & Vergne, I. 2006, "*Mycobacterium tuberculosis* inhibition of phagolysosome biogenesis and autophagy as a host defence mechanism", Cell Microbiol., vol. 8, no. 5, pp. 719-727.

DiazGranados, C. A., Jones, M. Y., Kongphet-Tran, T., White, N., Shapiro, M., Wang, Y. F., Ray, S. M., & Blumberg, H. M. 2009, "Outbreak of Pseudomonas aeruginosa infection associated with contamination of a flexible bronchoscope", Infect.Control Hosp.Epidemiol., vol. 30, no. 6, pp. 550-555.

Diel, R., R. Loddenkemper, and A. Nienhaus. 2010. "Evidence-based comparison of commercial interferon-gamma release assays for detecting active TB: a metaanalysis." Chest 137 (4): 952-68.

Dietrich, J. & Doherty, T. M. 2009, "Interaction of *Mycobacterium tuberculosis* with the host: consequences for vaccine development", APMIS, vol. 117, no. 5-6, pp. 440-457.

Dinnes J, Deeks J, Kunst H, Gibson A, Cummins E, Waugh N, Drobniewski F and Lalvani A. 2007 A systematic review of rapid diagnostic tests for the detection of tuberculosis infection. Health Technology Assessment. 11(3).

Djuretic T, Herbert J, Drobniewski F, Yates M, Smith EG, Magee JG, Williams R, Flanagan P, Watt B, Rayner A, Crowe M, Chadwick MV, Middleton AM, Watson JM. 2002. Antibiotic resistant tuberculosis in the United Kingdom: 1993–1999 Thorax 57:477–482

Donnabella, V., Martiniuk, F., Kinney, D., Bacerdo, M., Bonk, S., Hanna, B., & Rom, W. N. 1994, "Isolation of the gene for the beta subunit of RNA polymerase from rifampicin-resistant *Mycobacterium tuberculosis* and identification of new mutations", Am.J.Respir.Cell Mol.Biol., vol. 11, no. 6, pp. 639-643.

Donoghue, H. D. 2009, "Human tuberculosis--an ancient disease, as elucidated by ancient microbial biomolecules", Microbes.Infect. vol. 11, no. 14-15, pp. 1156-1162.

Donoghue, H. D., Spigelman, M., Greenblatt, C. L., Lev-Maor, G., Bar-Gal, G. K., Matheson, C., Vernon, K., Nerlich, A. G., & Zink, A. R. 2004, "Tuberculosis: from prehistory to Robert Koch, as revealed by ancient DNA", Lancet Infect.Dis., vol. 4, no. 9, pp. 584-592.

Douglass, J. & Steyn, L. M. 1993, "A ribosomal gene mutation in streptomycin-resistant *Mycobacterium tuberculosis* isolates", J.Infect.Dis., vol. 167, no. 6, pp. 1505-1506.

Draper and Daffe 2005. Tuberculosis and the Tubercle Bacillus. 261-273.

Drawz, S. M. & Bonomo, R. A. 2010, "Three decades of beta-lactamase inhibitors", Clin.Microbiol.Rev, vol. 23, no. 1, pp. 160-201.

Drobniewski, F. A., Gibson, A., Ruddy, M., & Yates, M. D. 2003, "Evaluation and Utilization as a Public Health Tool of a National Molecular Epidemiological Tuberculosis Outbreak Database within the United Kingdom from 1997 to 2001", Journal of Clinical Microbiology, vol. 41, no. 5, pp. 1861-1868.

Drobniewski F, Balabanova Y, Nikolayevsky V, Ruddy M, Kuznetzov S, Zakharova S, Melentyev A, Fedorin I. 2005. Drug-Resistant Tuberculosis, Clinical Virulence, and the Dominance of the Beijing Strain Family in Russia. JAMA, 293(22): 2726-2731.

Drobniewski FA, Hoffner S, Rusch-Gerdes S, Skenders G, Thomsen V and the WHO European Laboratory Strengthening Task Force 2006. Recommended standards for modern tuberculosis laboratory services in Europe. Eur Respir J. 28: 903–909

Drobniewski F, Rüsch-Gerdes S, Hoffner S; Subcommittee on Antimicrobial Susceptibility Testing of Mycobacterium tuberculosis of the European Committee for Antimicrobial Susceptibility Testing (EUCAST) of the European Society of Clinical Microbiology and Infectious Diseases (ESCMID). 2007. Antimicrobial susceptibility testing of Mycobacterium tuberculosis (EUCAST document E.DEF 8.1)--report of the Subcommittee on Antimicrobial Susceptibility Testing of Mycobacterium tuberculosis of the European Committee for Antimicrobial Susceptibility Testing (EUCAST) of the European Society of Clinical Microbiology and Infectious Diseases (ESCMID). Clin Microbiol Infect. 13(12):1144-56.

Easterbrook, P. J., Gibson, A., Murad, S., Lamprecht, D., Ives, N., Ferguson, A., Lowe, O., Mason, P., Ndudzo, A., Taziwa, A., Makombe, R., Mbengeranwa, L., Sola, C., Rostogi, N., & Drobniewski, F. 2004, "High Rates of Clustering of Strains Causing Tuberculosis in Harare, Zimbabwe: a Molecular Epidemiological Study", Journal of Clinical Microbiology, vol. 42, no. 10, pp. 4536-4544.

Ebrahimi-Rad, M., Bifani, P., Martin, C., Kremer, K., Samper, S., Rauzier, J., Kreiswirth, B., Blazquez, J., Jouan, M., van, S. D., & Gicquel, B. 2003, "Mutations in putative mutator genes of *Mycobacterium tuberculosis* strains of the W-Beijing family", Emerg.Infect.Dis., vol. 9, no. 7, pp. 838-845.

Economou A. 1999. "Following the leader: bacterial protein export through the Sec pathway." Trends Microbiol. 7 (8): 315-20.

Edlin, B. R., Tokars, J. I., Grieco, M. H., Crawford, J. T., Williams, J., Sordillo, E. M., Ong, K. R., Kilburn, J. O., Dooley, S. W & Castro, K. G. 1992, "An outbreak of

multidrug-resistant tuberculosis among hospitalized patients with the acquired immunodeficiency syndrome", N.Engl.J.Med., vol. 326, no. 23, pp. 1514-1521.

Elhay, M. J., Oettinger T, and Andersen P. 1998. "Delayed-type hypersensitivity responses to ESAT-6 and MPT64 from *Mycobacterium tuberculosis* in the guinea pig." Infect.Immun. 66 (7): 3454-56.

Ehlers, S. 2009, "Lazy, dynamic or minimally recrudescent? On the elusive nature and location of the mycobacterium responsible for latent tuberculosis", Infection, vol. 37, no. 2, pp. 87-95.

Eickhoff, T. C. 1977, "The current status of BCG immunization against tuberculosis", Annu.Rev Med., vol. 28, pp. 411-423.

Eisenach, K. D., Crawford, J. T., & Bates, J. H. 1988, "Repetitive DNA sequences as probes for *Mycobacterium tuberculosis*", Journal of Clinical Microbiology, vol. 26, no. 11, pp. 2240-2245.

Epstein MD, Schluger NW, Davidow AL, Bonk S, Rom WN, Hanna B. 1998. Time to detection of Mycobacterium tuberculosis in sputum culture correlates with outcome in patients receiving treatment for pulmonary tuberculosis. Chest 113(2):379-86.

Ernst, J. D. 1998, "Macrophage receptors for *Mycobacterium tuberculosis*", Infect.Immun., vol. 66, no. 4, pp. 1277-1281.

Esmonde, T. F. & Petheram, I. S. 1991, "Audit of tuberculosis contact tracing procedures in south Gwent", Respir.Med., vol. 85, no. 5, pp. 421-424.

Fair, E., O'Meara, M., Corbally, N., Keogh, B., & Hannan, M. 2006, "Molecular epidemiologic investigation of tuberculosis in an area of increasing incidence in inner-city Dublin", Ir.Med.J., vol. 99, no. 3, pp. 87-90.

Feil, E. J. & Spratt, B. G. 2001, "Recombination and the population structures of bacterial pathogens", Annu.Rev Microbiol., vol. 55, pp. 561-590.

Ferrara, G., Losi, M., D'Amico, R., Roversi, P., Piro, R., Meacci, M., Meccugni, B., Dori, I. M., Andreani, A., Bergamini, B. M., Mussini, C., Rumpianesi, F., Fabbri, L. M., & Richeldi, L. 2006 "Use in routine clinical practice of two commercial blood tests for diagnosis of infection with *Mycobacterium tuberculosis*: a prospective study", The Lancet, vol. 367, no. 9519, pp. 1328-1334.

Filliol, I., Motiwala, A. S., Cavatore, M., Qi, W., Hazbon, M. H., Bobadilla del Valle, M., Fyfe, J., Garcia-Garcia, L., Rastogi, N., Sola, C., Zozio, T., Guerrero, M. I., Leon, C. I., Crabtree, J., Angiuoli, S., Eisenach, K. D., Durmaz, R., Joloba, M. L., Rendon, A., Sifuentes-Osornio, J., Ponce de Leon, A., Cave, M. D., Fleischmann, R., Whittam, T. S., & Alland, D. 2006, "Global Phylogeny of *Mycobacterium tuberculosis* Based on Single Nucleotide Polymorphism (SNP) Analysis: Insights into Tuberculosis Evolution,

Phylogenetic Accuracy of Other DNA Fingerprinting Systems, and Recommendations for a Minimal Standard SNP Set", The Journal of Bacteriology, vol. 188, no. 2, pp. 759-772.

Finken, M., Kirschner, P., Meier, A., Wrede, A., & Bottger, E. C. 1993, "Molecular basis of streptomycin resistance in *Mycobacterium tuberculosis*: alterations of the ribosomal protein S12 gene and point mutations within a functional 16S ribosomal RNA pseudoknot", Mol.Microbiol, vol. 9, no. 6, pp. 1239-1246.

Fitzpatrick, L. K., Hardacker, J. A., Heirendt, W., Agerton, T., Streicher, A., Melnyk, H., Ridzon, R., Valway, S., & Onorato, I. 2001, "A preventable outbreak of tuberculosis investigated through an intricate social network", Clin.Infect.Dis., vol. 33, no. 11, pp. 1801-1806.

Fleischmann, R. D., Alland, D., Eisen, J. A., Carpenter, L., White, O., Peterson, J., DeBoy, R., Dodson, R., Gwinn, M., Haft, D., Hickey, E., Kolonay, J. F., Nelson, W. C., Umayam, L. A., Ermolaeva, M., Salzberg, S. L., Delcher, A., Utterback, T., Weidman, J., Khouri, H., Gill, J., Mikula, A., Bishai, W., Jacobs, J. W., Jr., Venter, J. C., & Fraser, C. M. 2002, "Whole-genome comparison of *Mycobacterium tuberculosis* clinical and laboratory strains", The Journal of Bacteriology, vol. 184, no. 19, pp. 5479-5490.

Flynn, J. L., Chan, J., Triebold, K. J., Dalton, D. K., Stewart, T. A., & Bloom, B. R. 1993, "An essential role for interferon gamma in resistance to *Mycobacterium tuberculosis* infection", J.Exp.Med., vol. 178, no. 6, pp. 2249-2254.

Fomukong, N. G., Dale, J. W., Osborn, T. W., & Grange, J. M. 1992, "Use of gene probes based on the insertion sequence IS986 to differentiate between BCG vaccine strains", J.Appl.Bacteriol., vol. 72, no. 2, pp. 126-133.

Fortune, S. M., Solache, A., Jaeger, A., Hill, P. J., Belisle, J. T., Bloom, B. R., Rubin, E. J., & Ernst, J. D. 2004, "*Mycobacterium tuberculosis* inhibits macrophage responses to IFN-gamma through myeloid differentiation factor 88-dependent and -independent mechanisms", J.Immunol., vol. 172, no. 10, pp. 6272-6280.

French, C. E., Kruijshaar, M. E., Jones, J. A., & Abubakar, I. 2009, "The influence of socio-economic deprivation on tuberculosis treatment delays in England, 2000-2005", Epidemiol.Infect., vol. 137, no. 4, pp. 591-596.

French, M. A., Price, P., & Stone, S. F. 2004, "Immune restoration disease after antiretroviral therapy", AIDS, vol. 18, no. 12, pp. 1615-1627.

Frieden, T. R., Sherman, L. F., Maw, K. L., Fujiwara, P. I., Crawford, J. T., Nivin, B., Sharp, V., Hewlett, D., Jr., Brudney, K., Alland, D., & Kreisworth, B. N. 1996, "A multi-institutional outbreak of highly drug-resistant tuberculosis: epidemiology and clinical outcomes", JAMA, vol. 276, no. 15, pp. 1229-1235.

Frothingham, R., Hills, H. G., & Wilson, K. H. 1994, "Extensive DNA sequence conservation throughout the *Mycobacterium tuberculosis* complex", Journal of Clinical Microbiology, vol. 32, no. 7, pp. 1639-1643.

Frothingham, R. & Meeker-O'Connell, W. A. 1998, "Genetic diversity in the *Mycobacterium tuberculosis* complex based on variable numbers of tandem DNA repeats", Microbiology, vol. 144, no. 5, pp. 1189-1196.

Frothingham, R. 1999, "Evolutionary bottlenecks in the agents of tuberculosis, leprosy, and paratuberculosis", Medical Hypotheses, vol. 52, no. 2, pp. 95-99.

Fu, L. M. and Shinnick T M. 2007. "Understanding the action of INH on a highly INH-resistant *Mycobacterium tuberculosis* strain using Genechips." Tuberculosis.(Edinb.) 87 (1): 63-70.

Gaber, K. A. Maggs A, Thould G, Goldman JM. 2005. "An outbreak of tuberculosis in the South West of England related to a public house." Prim.Care Respir.J. 14 (1): 51-55.

Gagneux, S., DeRiemer, K., Van, T., Kato-Maeda, M., de Jong, B. C., Narayanan, S., Nicol, M., Niemann, S., Kremer, K., Gutierrez, M. C., Hilty, M., Hopewell, P. C., & Small, P. M. 2006, "Variable host-pathogen compatibility in *Mycobacterium tuberculosis*", Proceedings of the National Academy of Sciences of the United States of America, vol. 103, no. 8, pp. 2869-2873.

Gagneux, S., Burgos, M. V., Deriemer, K., Encisco, A., Munoz, S., Hopewell, P. C., Small, P. M., & Pym, A. S. 2006a, "Impact of bacterial genetics on the transmission of isoniazid-resistant *Mycobacterium tuberculosis*", PLoS.Pathog., vol. 2, no. 6, p. e61.

Gagneux, S., Long, C. D., Small, P. M., Van, T., Schoolnik, G. K., & Bohannan, B. J. 2006b, "The competitive cost of antibiotic resistance in *Mycobacterium tuberculosis*", Science, vol. 312, no. 5782, pp. 1944-1946

Gagneux, S. & Small, P. M. 2007, "Global phylogeography of *Mycobacterium tuberculosis* and implications for tuberculosis product development", Lancet Infect.Dis., vol. 7, no. 5, pp. 328-337.

Gandhi, N. R., Moll, A., Sturm, A. W., Pawinski, R., Govender, T., Lalloo, U., Zeller, K., Andrews, J., & Friedland, G. 2006, "Extensively drug-resistant tuberculosis as a cause of death in patients co-infected with tuberculosis and HIV in a rural area of South Africa", Lancet, vol. 368, no. 9547, pp. 1575-1580.

Ganguly, N., Giang PH, Basu SK, Mir FA, Siddiqui I, Sharma P. 2007. "*Mycobacterium tuberculosis* 6-kDa early secreted antigenic target (ESAT-6) protein downregulates lipopolysaccharide induced c-myc expression by modulating the extracellular signal regulated kinases 1/2." BMC Immunol. 8: 24.

Ganguly, N., Giang PH, Gupta C, Basu SK, Siddiqui I, Salunke DM, Sharma P. 2008a. "*Mycobacterium tuberculosis* secretory proteins CFP-10, ESAT-6 and the CFP10:ESAT6 complex inhibit lipopolysaccharide-induced NF-kappaB transactivation by downregulation of reactive oxidative species (ROS) production." Immunol.Cell Biol. 86 (1): 98-106.

Ganguly, N., I. Siddiqui, and P. Sharma. 2008b. "Role of M. tuberculosis RD-1 region encoded secretory proteins in protective response and virulence." Tuberculosis.(Edinb.) 88 (6): 510-17.

Gao, L. Y., Guo S, McLaughlin B, Morisaki H, Engel JN, Brown EJ. 2004. "A mycobacterial virulence gene cluster extending RD1 is required for cytolysis, bacterial spreading and ESAT-6 secretion." Mol.Microbiol. 53 (6): 1677-93.

Garcia de Viedma, D., Lorenzo, G., Cardona, P. J., Rodriguez, N. A., Gordillo, S., Serrano, M. J., & Bouza, E. 2005, "Association between the infectivity of *Mycobacterium tuberculosis* strains and their efficiency for extrarespiratory infection", J.Infect.Dis., vol. 192, no. 12, pp. 2059-2065.

Garnier, T., Eiglmeier, K., Camus, J. C., Medina, N., Mansoor, H., Pryor, M., Duthoy, S., Grondin, S., Lacroix, C., Monsempe, C., Simon, S., Harris, B., Atkin, R., Doggett, J., Mayes, R., Keating, L., Wheeler, P. R., Parkhill, J., Barrell, B. G., Cole, S. T., Gordon, S. V., & Hewinson, R. G. 2003, "The complete genome sequence of Mycobacterium bovis", Proceedings of the National Academy of Sciences of the United States of America, vol. 100, no. 13, pp. 7877-7882.

Garton, N. J., Waddell, S. J., Sherratt, A. L., Lee, S. M., Smith, R. J., Senner, C., Hinds, J., Rajakumar, K., Adegbola, R. A., Besra, G. S., Butcher, P. D., & Barer, M. R. 2008, "Cytological and transcript analyses reveal fat and lazy persister-like bacilli in tuberculous sputum", PLoS.Med., vol. 5, no. 4, p. e75.

Gey Van Pittius, N. C., Gamieldien J, Hide W, Brown GD, Siezen RJ, Beyers AD. 2001. "The ESAT-6 gene cluster of *Mycobacterium tuberculosis* and other high G+C Gram-positive bacteria." Genome Biol. 2001;2(10):RESEARCH0044. Epub 2001 Sep 19.

Gibson A, Brown T, Baker L, Drobniewski F. 2005. Can 15-Locus Mycobacterial Interspersed Repetitive Unit–Variable-Number Tandem Repeat Analysis Provide Insight into the Evolution of *Mycobacterium tuberculosis*? Applied and Environmental Microbiology 71(12): 8207–8213

Gilbert, R. L., Antoine, D., French, C. E., Abubakar, I., Watson, J. M., & Jones, J. A. 2009, "The impact of immigration on tuberculosis rates in the United Kingdom compared with other European countries", Int.J.Tuberc.Lung Dis., vol. 13, no. 5, pp. 645-651.

Gillespie, S. H. & McHugh, T. D. 1997, "The biological cost of antimicrobial resistance", Trends Microbiol., vol. 5, no. 9, pp. 337-339.

Gillespie, S. H., Dickens, A., & McHugh, T. D. 2000, "False molecular clusters due to nonrandom association of IS6110 with *Mycobacterium tuberculosis*", Journal of Clinical Microbiology, vol. 38, no. 6, pp. 2081-2086.

Gillespie, S. H., Billington, O. J., Breathnach, A., & McHugh, T. D. 2002, "Multiple drug-resistant *Mycobacterium tuberculosis*: evidence for changing fitness following passage through human hosts", Microb.Drug Resist., vol. 8, no. 4, pp. 273-279.

Girardi E, Raviglione MC, Antonucci G, Godfrey-Faussett P, & Ippolito G. 2000. "Impact of the HIV epidemic on the spread of other diseases: the case of tuberculosis." AIDS 14: S47-S56.

Glynn, J. R., Whiteley, J., Bifani, P. J., Kremer, K., & van, S. D. 2002, "Worldwide occurrence of Beijing/W strains of *Mycobacterium tuberculosis*: a systematic review", Emerg.Infect.Dis., vol. 8, no. 8, pp. 843-849.

Go, M. F., Kapur, V., Graham, D. Y., & Musser, J. M. 1996, "Population genetic analysis of Helicobacter pylori by multilocus enzyme electrophoresis: extensive allelic diversity and recombinational population structure", The Journal of Bacteriology, vol. 178, no. 13, pp. 3934-3938.

Goble, M., Iseman, M. D., Madsen, L. A., Waite, D., Ackerson, L., & Horsburgh, C. R., Jr. 1993, "Treatment of 171 patients with pulmonary tuberculosis resistant to isoniazid and rifampin", N.Engl.J.Med., vol. 328, no. 8, pp. 527-532.

Golub, J. E., Bur, S., Cronin, W. A., Gange, S., Baruch, N., Comstock, G. W., & Chaisson, R. E. 2006, "Delayed tuberculosis diagnosis and tuberculosis transmission", Int.J.Tuberc.Lung Dis., vol. 10, no. 1, pp. 24-30.

Gomez, J. E., J. M. Chen, and W. R. Bishai. 1997. "Sigma factors of *Mycobacterium tuberculosis*." Tuber.Lung Dis. 78 (3-4): 175-83.

Gomez-Reino, J. J., Carmona, L., Valverde, V. R., Mola, E. M., & Montero, M. D. 2003, "Treatment of rheumatoid arthritis with tumor necrosis factor inhibitors may predispose to significant increase in tuberculosis risk: a multicenter active-surveillance report", Arthritis Rheum., vol. 48, no. 8, pp. 2122-2127.

Gonzalez-Juarrero, M., Turner, O. C., Turner, J., Marietta, P., Brooks, J. V., & Orme, I. M. 2001, "Temporal and spatial arrangement of lymphocytes within lung granulomas induced by aerosol infection with *Mycobacterium tuberculosis*", Infect.Immun., vol. 69, no. 3, pp. 1722-1728.

Krishna K. Gopaul, Timothy J. Brown, Andrea L. Gibson, Malcolm D. Yates, and Francis A. Drobniewski. 2006. Progression Toward an Improved DNA Amplification-Based Typing Technique in the Study of *Mycobacterium tuberculosis* Epidemiology. Journal of Clinical Microbiology 44(7): 2492–2498

Gordon, S. V., Brosch, R., Billault, A., Garnier, T., Eiglmeier, K., & Cole, S. T. 1999, "Identification of variable regions in the genomes of tubercle bacilli using bacterial artificial chromosome arrays", Mol.Microbiol., vol. 32, no. 3, pp. 643-655.

Gorge, O., Lopez, S., Hilaire, V., Lisanti, O., Ramisse, V., & Vergnaud, G. 2008, "Selection and validation of a multilocus variable-number tandem-repeat analysis panel for typing Shigella spp", Journal of Clinical Microbiology, vol. 46, no. 3, pp. 1026-1036.

Goyal, M., Young, D., Zhang, Y., Jenkins, P. A., & Shaw, R. J. 1994, "PCR amplification of variable sequence upstream of katG gene to subdivide strains of *Mycobacterium tuberculosis* complex", Journal of Clinical Microbiology, vol. 32, no. 12, pp. 3070-3071.

Granger, C. 2006, "The specificity of interferon-{gamma}-based blood tests in the identification of latent tuberculosis infection", European Respiratory Journal, vol. 28, no. 6, p. 1283.

Greater London Authority. 2006. London borough residents by country of birth. An analysis of 2001 Census data. Data Management and Analysis Group.

Greater London Authority. 2008. A Profile of Londoners by Country of Birth. Estimates from the 2006 Annual Population Survey. Data Management and Analysis Group.

Groenen, P. M., Bunschoten, A. E., van, S. D., & van Embden, J. D. 1993, "Nature of DNA polymorphism in the direct repeat cluster of *Mycobacterium tuberculosis*; application for strain differentiation by a novel typing method", Mol.Microbiol., vol. 10, no. 5, pp. 1057-1065.

Guiney, D. G., Jr. 1984, "Promiscuous transfer of drug resistance in gram-negative bacteria", J.Infect.Dis., vol. 149, no. 3, pp. 320-329.

Guinn, K. M., Hickey, M. J., Mathur, S. K., Zakel, K. L., Grotzke, J. E., Lewinsohn, D. M., Smith, S., & Sherman, D. R. 2004, "Individual RD1-region genes are required for export of ESAT-6/CFP-10 and for virulence of *Mycobacterium tuberculosis*", Mol.Microbiol., vol. 51, no. 2, pp. 359-370.

Gutacker, M. M., Mathema, B., Soini, H., Shashkina, E., Kreiswirth, B. N., Graviss, E. A., & Musser, J. M. 2006, "Single-nucleotide polymorphism-based population genetic analysis of *Mycobacterium tuberculosis* strains from 4 geographic sites", J.Infect.Dis., vol. 193, no. 1, pp. 121-128.

Gutierrez, M. C., Vincent, V., Aubert, D., Bizet, J., Gaillot, O., Lebrun, L., Le, P. C., Le Pennec, M. P., Mathieu, D., Offredo, C., Pangon, B., & Pierre-Audigier, C. 1998, "Molecular fingerprinting of *Mycobacterium tuberculosis* and risk factors for tuberculosis transmission in Paris, France, and surrounding area", Journal of Clinical Microbiology, vol. 36, no. 2, pp. 486-492.

Gutierrez, M. C., Brisse, S., Brosch, R., Fabre, M., Omais, B., Marmiesse, M., Supply, P., & Vincent, V. 2005, "Ancient origin and gene mosaicism of the progenitor of *Mycobacterium tuberculosis*", PLoS.Pathog., vol. 1, no. 1, p. e5.

Gutierrez, M. G., Master, S. S., Singh, S. B., Taylor, G. A., Colombo, M. I., & Deretic, V. 2004, "Autophagy is a defense mechanism inhibiting BCG and *Mycobacterium tuberculosis* survival in infected macrophages", Cell, vol. 119, no. 6, pp. 753-766.

Hampshire T, Soneji S, Bacon J, James BW, Hinds J, Laing K, Stabler RA, Marsh PD, Butcher PD. 2004. "Stationary phase gene expression of *Mycobacterium tuberculosis* following a progressive nutrient depletion: a model for persistent organisms?" Tuberculosis (Edinb.) 84 (3-4): 228-38.

Harboe, M., Oettinger, T., Wiker, H. G., Rosenkrands, I., & Andersen, P. 1996, "Evidence for occurrence of the ESAT-6 protein in *Mycobacterium tuberculosis* and virulent Mycobacterium bovis and for its absence in Mycobacterium bovis BCG", Infect.Immun., vol. 64, no. 1, pp. 16-22.

Hardinge, F. M., Black, M., & Chamberlain, P. 1999. TB contact tracing in South Buckinghamshire from 1994 to mid 1998. Am.J.Respir.Crit Care Med. 159(Suppl), a303.

Havlir, D. V. & Barnes, P. F. 1999, "Tuberculosis in patients with human immunodeficiency virus infection", N.Engl.J.Med., vol. 340, no. 5, pp. 367-373.

Health Protection Agency. HPA TB Report 2008. 2008.

Health Protection Agency. Tuberculosis in the UK. Annual report on tuberculosis surveillance in the UK 2009. 2009.

Health Protection Agency. HIV in the United Kingdom: 2008 Report. 2008.

Heathcock, R. and C. Watts. 2008. "Measles outbreaks in London, United Kingdom - a preliminary report." Euro.Surveill 13 (15).

Hedgecock, L. W. & Faucher, I. O. 1957, "Relation of pyrogallol-peroxidative activity to isoniazid resistance in *Mycobacterium tuberculosis*", Am.Rev Tuberc., vol. 75, no. 4, pp. 670-674.

Heifets LB, and Cangelosi GA. 1999. Drug susceptibility testing of *Mycobacterium tuberculosis:* a neglected problem at the turn of the century. INT J TUBERC LUNG DIS 3(7):564–581

Helke, K. L., Mankowski, J. L., & Manabe, Y. C. 2006, "Animal models of cavitation in pulmonary tuberculosis", Tuberculosis.(Edinb.), vol. 86, no. 5, pp. 337-348.

Henao-Tamayo, M., Junqueira-Kipnis, A. P., Ordway, D., Gonzales-Juarrero, M., Stewart, G. R., Young, D. B., Wilkinson, R. J., Basaraba, R. J., & Orme, I. M. 2007, "A mutant of *Mycobacterium tuberculosis* lacking the 19-kDa lipoprotein Rv3763 is highly attenuated in vivo but retains potent vaccinogenic properties", Vaccine, vol. 25, no. 41, pp. 7153-7159.

Hermans, P. W., van Soolingen, D., Dale, J. W., Schuitema, A. R., McAdam, R. A., Catty, D., & van Embden, J. D. 1990, "Insertion element IS986 from *Mycobacterium tuberculosis*: a useful tool for diagnosis and epidemiology of tuberculosis", Journal of Clinical Microbiology, vol. 28, no. 9, pp. 2051-2058.

Hershberg, R., Lipatov, M., Small, P. M., Sheffer, H., Niemann, S., Homolka, S., Roach, J. C., Kremer, K., Petrov, D. A., Feldman, M. W., & Gagneux, S. 2008, "High functional diversity in *Mycobacterium tuberculosis* driven by genetic drift and human demography", PLoS.Biol., vol. 6, no. 12, p. e311.

Heym, B., Honore, N., Truffot-Pernot, C., Banerjee, A., Schurra, C., Jacobs, W. R., Jr., van Embden, J. D., Grosset, J. H., & Cole, S. T. 1994, "Implications of multidrug resistance for the future of short-course chemotherapy of tuberculosis: a molecular study", Lancet, vol. 344, no. 8918, pp. 293-298.

Heym, B., Stavropoulos, E., Honore, N., Domenech, P., Saint-Joanis, B., Wilson, T. M., Collins, D. M., Colston, M. J., & Cole, S. T. 1997, "Effects of overexpression of the alkyl hydroperoxide reductase AhpC on the virulence and isoniazid resistance of *Mycobacterium tuberculosis*", Infect.Immun., vol. 65, no. 4, pp. 1395-1401.

Higgins, C. F., McLaren, R. S., & Newbury, S. F. 1988, "Repetitive extragenic palindromic sequences, mRNA stability and gene expression: evolution by gene conversion? A review", Gene, vol. 72, no. 1-2, pp. 3-14.

Hillemann D, Richter E, Rüsch-Gerdes S. 2006. Use of the BACTEC Mycobacteria Growth Indicator Tube 960 automated system for recovery of Mycobacteria from 9,558 extrapulmonary specimens, including urine samples. J Clin Microbiol. 44(11):4014-7.

Hillemann, D., Rusch-Gerdes, S., & Richter, E. 2007, "Evaluation of the GenoType MTBDRplus assay for rifampin and isoniazid susceptibility testing of *Mycobacterium tuberculosis* strains and clinical specimens", Journal of Clinical Microbiology, vol. 45, no. 8, pp. 2635-2640.

Hinchey, J., Lee, S., Jeon, B. Y., Basaraba, R. J., Venkataswamy, M. M., Chen, B., Chan, J., Braunstein, M., Orme, I. M., Derrick, S. C., Morris, S. L., Jacobs, W. R., Jr., & Porcelli, S. A. 2007, "Enhanced priming of adaptive immunity by a proapoptotic mutant of *Mycobacterium tuberculosis*", J.Clin.Invest, vol. 117, no. 8, pp. 2279-2288.

Hirsh, A. E., Tsolaki, A. G., DeRiemer, K., Feldman, M. W., & Small, P. M. 2004, "Stable association between strains of *Mycobacterium tuberculosis* and their human host populations", Proceedings of the National Academy of Sciences of the United States of America, vol. 101, no. 14, pp. 4871-4876.

Honore, N. & Cole, S. T. 1994, "Streptomycin resistance in mycobacteria", Antimicrobial Agents and Chemotherapy, vol. 38, no. 2, pp. 238-242.

Hou, J. M., D'Lima, N. G., Rigel, N. W., Gibbons, H. S., McCann, J. R., Braunstein, M., & Teschke, C. M. 2008, "ATPase activity of *Mycobacterium tuberculosis* SecA1 and SecA2 proteins and its importance for SecA2 function in macrophages", The Journal of Bacteriology, vol. 190, no. 14, pp. 4880-4887.

Hovav, A. H. & Bercovier, H. 2006, "Pseudo-rationale design of efficient TB vaccines: lesson from the mycobacterial 27-kDa lipoprotein", Tuberculosis.(Edinb.), vol. 86, no. 3-4, pp. 225-235.

Hovav, A. H., Mullerad, J., Davidovitch, L., Fishman, Y., Bigi, F., Cataldi, A., & Bercovier, H. 2003, "The *Mycobacterium tuberculosis* recombinant 27-kilodalton lipoprotein induces a strong Th1-type immune response deleterious to protection", Infect.Immun., vol. 71, no. 6, pp. 3146-3154.

Hsu T, Hingley-Wilson SM, Chen B, Chen M, Dai AZ, Morin PM, Marks CB, Padiyar J, Goulding C, Gingery M, Eisenberg D, Russell RG, Derrick SC, Collins FM, Morris SL, King CH, Jacobs WR Jr. 2003. "The primary mechanism of attenuation of bacillus Calmette-Guerin is a loss of secreted lytic function required for invasion of lung interstitial tissue." Proceedings of the National Academy of Sciences of the United States of America 100 (21): 12420-25

Huebner, R. E., Schein, M. F., & Bass, J. B., Jr. 1993, "The tuberculin skin test", Clin.Infect.Dis., vol. 17, no. 6, pp. 968-975.

Hughes AL, Friedman R, & Murray M. 2002. Genomewide pattern of synonymous nucleotide substitution in two complete genomes of *Mycobacterium tuberculosis*. Emerging Infectious Diseases 8 (11): 1342-1346.

Huitric, E., Werngren, J., Jureen, P., & Hoffner, S. 2006, "Resistance levels and rpoB gene mutations among in vitro-selected rifampin-resistant *Mycobacterium tuberculosis* mutants", Antimicrobial Agents and Chemotherapy, vol. 50, no. 8, pp. 2860-2862.

Hunter, P. R. & Gaston, M. A. 1988, "Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity", Journal of Clinical Microbiology, vol. 26, no. 11, pp. 2465-2466.

Hunter, R. L., Olsen, M. R., Jagannath, C., & Actor, J. K. 2006, "Multiple roles of cord factor in the pathogenesis of primary, secondary, and cavitary tuberculosis, including a revised description of the pathology of secondary disease", Ann.Clin.Lab Sci., vol. 36, no. 4, pp. 371-386.

Hussain, S. F., Watura, R., Cashman, B., Campbell, I. A., & Evans, M. R. 1992, "Tuberculosis contact tracing: are the British Thoracic Society guidelines still appropriate?", Thorax, vol. 47, no. 11, pp. 984-985.

Hussain, R., Talat, N., Shahid, F., & Dawood, G. 2007, "Longitudinal tracking of cytokines after acute exposure to tuberculosis: association of distinct cytokine patterns with protection and disease development", Clin.Vaccine Immunol., vol. 14, no. 12, pp. 1578-1586.

Iñigo J, Arce A, Palenque E, García de Viedma D, Chaves F. 2008. "Decreased tuberculosis incidence and declining clustered case rates, Madrid." Emerg.Infect.Dis. 14 (10): 1641-43.

Ioerger, T. R., Feng, Y., Ganesula, K., Chen, X., Dobos, K. M., Fortune, S., Jacobs, W. R., Jr., Mizrahi, V., Parish, T., Rubin, E., Sassetti, C., & Sacchettini, J. C. 2010, "Variation among genome sequences of H37Rv strains of *Mycobacterium tuberculosis* from multiple laboratories", The Journal of Bacteriology, vol. 192, no. 14, pp. 3645-3653.

Irish, C., Jolly, E., & Baker, T. 1997. "Contact tracing smear negative and non-pulmonary tuberculosis in a high incidence area." Thorax 52 (Suppl 6), a34.

Jacobs, M., Togbe, D., Fremond, C., Samarina, A., Allie, N., Botha, T., Carlos, D., Parida, S. K., Grivennikov, S., Nedospasov, S., Monteiro, A., Le, B. M., Quesniaux, V., & Ryffel, B. 2007, "Tumor necrosis factor is critical to control tuberculosis infection", Microbes.Infect., vol. 9, no. 5, pp. 623-628.

Jalapathy, K. V., Prabha, C., & Das, S. D. 2004, "Correlates of protective immune response in tuberculous pleuritis", FEMS Immunol.Med.Microbiol., vol. 40, no. 2, pp. 139-145.

Jang, S., Uzelac, A., & Salgame, P. 2008, "Distinct chemokine and cytokine gene expression pattern of murine dendritic cells and macrophages in response to *Mycobacterium tuberculosis* infection", J.Leukoc.Biol., vol. 84, no. 5, pp. 1264-1270.

Jenkins, C., Claxton, A. P., Shorten, R. J., McHugh, T. D., & Gillespie, S. H. 2005, "Rifampicin resistance in tuberculosis outbreak, London, England", Emerg.Infect.Dis., vol. 11, no. 6, pp. 931-934.

Jimenez-Martinez, M. C., Linares, M., Baez, R., Montano, L. F., Martinez-Cairo, S., Gorocica, P., Chavez, R., Zenteno, E., & Lascurain, R. 2004, "Intracellular expression of interleukin-4 and interferon-gamma by a *Mycobacterium tuberculosis* antigen-stimulated CD4+ CD57+ T-cell subpopulation with memory phenotype in tuberculosis patients", Immunology, vol. 111, no. 1, pp. 100-106.

Johnsborg, O., Eldholm, V., & Havarstein, L. S. 2007, "Natural genetic transformation: prevalence, mechanisms and function", Res.Microbiol., vol. 158, no. 10, pp. 767-778.

Jones, D., Metzger, H. J., Schatz, A., & Waksman, S. A. 1944, "Control of Gram negative bacteria in experimental animals by streptomycin", Science, vol. 100, no. 2588, pp. 103-105.

Kam, K. M., Yip, C. W., Tse, L. W., Leung, K. L., Wong, K. L., Ko, W. M., & Wong, W. S. 2006, "Optimization of variable number tandem repeat typing set for differentiating *Mycobacterium tuberculosis* strains in the Beijing family", FEMS Microbiol.Lett., vol. 256, no. 2, pp. 258-265.

Kamerbeek, J., Schouls, L., Kolk, A., van Agterveld, M., van Soolingen, D., Kuijper, S., Bunschoten, A., Molhuizen, H., Shaw, R., Goyal, M., & van Embden, J. 1997, "Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for

diagnosis and epidemiology", Journal of Clinical Microbiology, vol. 35, no. 4, pp. 907-914.

Kanduma, E., McHugh, T. D., & Gillespie, S. H. 2003, "Molecular methods for *Mycobacterium tuberculosis* strain typing: a users guide", J.Appl.Microbiol., vol. 94, no. 5, pp. 781-791.

Kappelman, J., Alcicek, M. C., Kazanci, N., Schultz, M., Ozkul, M., & Sen, S. 2008, "First Homo erectus from Turkey and implications for migrations into temperate Eurasia", Am.J.Phys.Anthropol., vol. 135, no. 1, pp. 110-116.

Kapur, V., Li, L. L., Hamrick, M. R., Plikaytis, B. B., Shinnick, T. M., Telenti, A., Jacobs, W. R., Jr., Banerjee, A., Cole, S., Yuen, K. Y., & . 1995, "Rapid Mycobacterium species assignment and unambiguous identification of mutations associated with antimicrobial resistance in *Mycobacterium tuberculosis* by automated DNA sequencing", Arch.Pathol.Lab Med., vol. 119, no. 2, pp. 131-138.

Kapur, V., Li, L. L., Iordanescu, S., Hamrick, M. R., Wanger, A., Kreiswirth, B. N., & Musser, J. M. 1994, "Characterization by automated DNA sequencing of mutations in the gene (rpoB) encoding the RNA polymerase beta subunit in rifampin-resistant *Mycobacterium tuberculosis* strains from New York City and Texas", Journal of Clinical Microbiology, vol. 32, no. 4, pp. 1095-1098.

Kapur, V., Whittam, T. S., & Musser, J. M., "Is *Mycobacterium tuberculosis* 15,000 Years Old", The Journal of Infectious Diseases, vol. 170, no. 5, pp. 1348-1349.

Karboul, A., van Pittius, N., Namouchi, A., Vincent, V., Sola, C., Rastogi, N., Suffys, P., Fabre, M., Cataldi, A., Huard, R., Kurepina, N., Kreiswirth, B., Ho, J., Gutierrez, M., & Mardassi, H. 2006. "Insights into the evolutionary history of tubercle bacilli as disclosed by genetic rearrangements within a PE_PGRS duplicated gene pair." BMC Evolutionary Biology 6 (107)

Kato-Maeda, M., Rhee, J. T., Gingeras, T. R., Salamon, H., Drenkow, J., Smittipat, N., & Small, P. M. 2001, "Comparing genomes within the species *Mycobacterium tuberculosis*", Genome Res., vol. 11, no. 4, pp. 547-554.

Kaushal D, Schroeder BG, Tyagi S, Yoshimatsu T, Scott C, Ko C, Carpenter L, Mehrotra J, Manabe YC, Fleischmann RD, Bishai WR. 2002. "Reduced immunopathology and mortality despite tissue persistence in a *Mycobacterium tuberculosis* mutant lacking alternative -â factor, SigH." Proceedings of the National Academy of Sciences of the United States of America 99 (12): 8330-35

Keim, P., Klevytska, A. M., Price, L. B., Schupp, J. M., Zinser, G., Smith, K. L., Hugh-Jones, M. E., Okinaka, R., Hill, K. K., & Jackson, P. J. 1999, "Molecular diversity in Bacillus anthracis", J.Appl.Microbiol., vol. 87, no. 2, pp. 215-217.

Keller C, Lauber J, Blumenthal A, Buer J, Ehlers S. 2004. "Resistance and susceptibility to tuberculosis analysed at the transcriptome level: lessons from mouse macrophages." Tuberculosis.(Edinb.) 84 (3-4): 144-58.

Kettaneh, A., Seng, L., Tiev, K. P., Toledano, C., Fabre, B., & Cabane, J. 2006, "Human leukocyte antigens and susceptibility to tuberculosis: a meta-analysis of case-control studies", Int.J.Tuberc.Lung Dis., vol. 10, no. 7, pp. 717-725.

Kik SV, Verver S, van Soolingen D, de Haas PE, Cobelens FG, Kremer K, van Deutekom H, Borgdorff MW. 2008. "Tuberculosis Outbreaks Predicted by Characteristics of First Patients in a DNA Fingerprint Cluster." American Journal of Respiratory and Critical Care Medicine 178 (1): 96-104.

Kilicaslan Z, Kiyan E, Kucuk C, Kumbetli S, Sarimurat N, Ozturk F, Yapici D, Al S, Erboran T, Iliksu N. 2009. "Risk of active tuberculosis in adult household contacts of smear-positive pulmonary tuberculosis cases." Int.J.Tuberc.Lung Dis. 13(1): 93-98.

Kincaid, E. Z. & Ernst, J. D. 2003, "*Mycobacterium tuberculosis* exerts gene-selective inhibition of transcriptional responses to IFN-gamma without inhibiting STAT1 function", J.Immunol., vol. 171, no. 4, pp. 2042-2049.

Kinhikar AG, Verma I, Chandra D, Singh KK, Weldingh K, Andersen P, Hsu T, Jacobs WR Jr, Laal S. 2010. "Potential role for ESAT6 in dissemination of M. tuberculosis via human lung epithelial cells." Mol.Microbiol. 75 (1): 92-106.

Klevytska, A. M., Price, L. B., Schupp, J. M., Worsham, P. L., Wong, J., & Keim, P. 2001, "Identification and characterization of variable-number tandem repeats in the Yersinia pestis genome", Journal of Clinical Microbiology, vol. 39, no. 9, pp. 3179-3185.

Klovdahl, A. S., Graviss, E. A., Yaganehdoost, A., Ross, M. W., Wanger, A., Adams, G. J., & Musser, J. M. 2001, "Networks and tuberculosis: an undetected community outbreak involving public places", Soc.Sci.Med., vol. 52, no. 5, pp. 681-694.

Korf, J., Stoltz, A., Verschoor, J., De, B. P., & Grooten, J. 2005, "The *Mycobacterium tuberculosis* cell wall component mycolic acid elicits pathogen-associated host innate immune responses", Eur.J.Immunol., vol. 35, no. 3, pp. 890-900.

Kourbatova, E. V., Borodulin, B. E., Borodulina, E. A., del, R. C., Blumberg, H. M., & Leonard, M. K., Jr. 2006, "Risk factors for mortality among adult patients with newly diagnosed tuberculosis in Samara, Russia", Int.J.Tuberc.Lung Dis., vol. 10, no. 11, pp. 1224-1230.

Kremer, K., Au, B. K., Yip, P. C., Skuce, R., Supply, P., Kam, K. M., & van, S. D. 2005, "Use of variable-number tandem-repeat typing to differentiate *Mycobacterium tuberculosis* Beijing family isolates from Hong Kong and comparison with IS6110 restriction fragment length polymorphism typing and spoligotyping", Journal of Clinical Microbiology, vol. 43, no. 1, pp. 314-320.

Kremer, K., van Soolingen, D., Frothingham, R., Haas, W. H., Hermans, P. W. M., Martin, C., Palittapongarnpim, P., Plikaytis, B. B., Riley, L. W., Yakrus, M. A., Musser, J. M., & van Embden, J. D. A. 1999, "Comparison of Methods Based on Different Molecular Epidemiological Markers for Typing of *Mycobacterium tuberculosis* Complex Strains: Interlaboratory Study of Discriminatory Power and Reproducibility", Journal of Clinical Microbiology, vol. 37, no. 8, pp. 2607-2618.

Kremer, K., Arnold, C., Cataldi, A., Gutierrez, M. C., Haas, W. H., Panaiotov, S., Skuce, R. A., Supply, P., van der Zanden, A. G., & van, S. D. 2005, "Discriminatory power and reproducibility of novel DNA typing methods for *Mycobacterium tuberculosis* complex strains", Journal of Clinical Microbiology, vol. 43, no. 11, pp. 5628-5638.

Kruijshaar, M. E., Watson, J. M., Drobniewski, F., Anderson, C., Brown, T. J., Magee, J. G., Smith, E. G., Story, A., & Abubakar, I. 2008, "Increasing antituberculosis drug resistance in the United Kingdom: analysis of National Surveillance Data", BMJ, vol. 336, no. 7655, pp. 1231-1234.

Kugelberg, E., Lofmark, S., Wretlind, B., & Andersson, D. I. 2005, "Reduction of the fitness burden of quinolone resistance in Pseudomonas aeruginosa", J.Antimicrob.Chemother., vol. 55, no. 1, pp. 22-30.

Kumar, S., Innes, J. A., & Skinner, C. 1992. "Yield from tuberculosis contact tracing in Birmingham." Thorax 47; 875.

Kwara, A., Schiro, R., Cowan, L. S., Hyslop, N. E., Wiser, M. F., Roahen Harrison, S., Kissinger, P., Diem, L., & Crawford, J. T. 2003, "Evaluation of the Epidemiologic Utility of Secondary Typing Methods for Differentiation of *Mycobacterium tuberculosis* Isolates", Journal of Clinical Microbiology, vol. 41, no. 6, pp. 2683-2685.

Lacey, R. W. 1984, "Evolution of microorganisms and antibiotic resistance", Lancet, vol. 2, no. 8410, pp. 1022-1025.

Ladel, C. H., Blum, C., Dreher, A., Reifenberg, K., & Kaufmann, S. H. 1995a, "Protective role of gamma/delta T cells and alpha/beta T cells in tuberculosis", Eur.J.Immunol., vol. 25, no. 10, pp. 2877-2881.

Ladel, C. H., Hess, J., Daugelat, S., Mombaerts, P., Tonegawa, S., & Kaufmann, S. H. 1995b, "Contribution of alpha/beta and gamma/delta T lymphocytes to immunity against Mycobacterium bovis bacillus Calmette Guerin: studies with T cell receptor-deficient mutant mice", Eur.J.Immunol., vol. 25, no. 3, pp. 838-846.

Lalvani A, Brookes R, Wilkinson RJ, Malin AS, Pathan AA, Andersen P, Dockrell H, Pasvol G, Hill AV. 1998. "Human cytolytic and interferon gamma-secreting CD8+ T lymphocytes specific for *Mycobacterium tuberculosis*." Proceedings of the National Academy of Sciences of the United States of America 95 (1): 270-75.

Lalvani, A., Nagvenkar P, Udwadia Z, Pathan AA, Wilkinson KA, Shastri JS, Ewer K, Hill AV, Mehta A, Rodrigues C. 2001a. "Enumeration of T cells specific for RD1-

encoded antigens suggests a high prevalence of latent *Mycobacterium tuberculosis* infection in healthy urban Indians." J.Infect.Dis 183 (3): 469-77.

Lalvani A, Pathan AA, McShane H, Wilkinson RJ, Latif M, Conlon CP, Pasvol G, and Hill AVS. 2001b "Rapid Detection of *Mycobacterium tuberculosis* Infection by Enumeration of Antigen-specific T Cells." Am J Respir Crit Care Med. Vol 163. pp 824–828

Lalvani, A. 2007, "Diagnosing tuberculosis infection in the 21st century: new tools to tackle an old enemy", Chest, vol. 131, no. 6, pp. 1898-1906.

Latchumanan VK, Balkhi MY, Sinha A, Singh B, Sharma P, Natarajan K. 2005. "Regulation of immune responses to *Mycobacterium tuberculosis* secretory antigens by dendritic cells." Tuberculosis (Edinb.) 85 (5-6): 377-83.

Le Fleche P., Hauck, Y., Onteniente, L., Prieur, A., Denoeud, F., Ramisse, V., Sylvestre, P., Benson, G., Ramisse, F., & Vergnaud, G. 2001, "A tandem repeats database for bacterial genomes: application to the genotyping of Yersinia pestis and Bacillus anthracis", BMC Microbiol., vol. 1, p. 2.

Le Fleche P., Fabre, M., Denoeud, F., Koeck, J. L., & Vergnaud, G. 2002, "High resolution, on-line identification of strains from the *Mycobacterium tuberculosis* complex based on tandem repeat typing", BMC Microbiol., vol. 2, p. 37.

Lehman, J. 1949, "On the effect of isomers of para-aminosalicylic acid and related substances on the tuberculostatic effect of PAS", Experientia, vol. 5, no. 9, pp. 365-367.

Leung, C. C., Tam, C. M., Chan, S. L., Chan-Yeung, M., Chan, C. K., & Chang, K. C. 2001, "Efficacy of the BCG revaccination programme in a cohort given BCG vaccination at birth in Hong Kong", Int.J.Tuberc.Lung Dis., vol. 5, no. 8, pp. 717-723.

Lewis, K. N., Liao, R., Guinn, K. M., Hickey, M. J., Smith, S., Behr, M. A., & Sherman, D. R. 2003, "Deletion of RD1 from *Mycobacterium tuberculosis* mimics bacille Calmette-Guerin attenuation", J.Infect.Dis., vol. 187, no. 1, pp. 117-123.

Li, Z., Kelley, C., Collins, F., Rouse, D., & Morris, S. 1998, "Expression of katG in *Mycobacterium tuberculosis* is associated with its growth and persistence in mice and guinea pigs", J.Infect.Dis., vol. 177, no. 4, pp. 1030-1035.

Lindstedt, B. A., Tham, W., nielsson-Tham, M. L., Vardund, T., Helmersson, S., & Kapperud, G. 2008, "Multiple-locus variable-number tandem-repeats analysis of Listeria monocytogenes using multicolour capillary electrophoresis and comparison with pulsed-field gel electrophoresis typing", J.Microbiol.Methods, vol. 72, no. 2, pp. 141-148.

Linz, B. & Schuster, S. C. 2007, "Genomic diversity in Helicobacter and related organisms", Research in Microbiology, vol. 158, no. 10, pp. 737-744.

Lipman, M. & Breen, R. 2006, "Immune reconstitution inflammatory syndrome in HIV", Curr.Opin.Infect.Dis., vol. 19, no. 1, pp. 20-25.

Liu, X., Gutacker, M. M., Musser, J. M., & Fu, Y. X. 2006, "Evidence for Recombination in *Mycobacterium tuberculosis*", The Journal of Bacteriology, vol. 188, no. 23, pp. 8169-8177.

Livak, K. J. and T. D. Schmittgen. 2001. "Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method." Methods 25 (4): 402-08.

Lopez, B., Aguilar, D., Orozco, H., Burger, M., Espitia, C., Ritacco, V., Barrera, L., Kremer, K., Hernandez-Pando, R., Huygen, K., & van, S. D. 2003, "A marked difference in pathogenesis and immune response induced by different *Mycobacterium tuberculosis* genotypes", Clin.Exp.Immunol., vol. 133, no. 1, pp. 30-37.

López-Calleja AI, Gavín P, Lezcano MA, Vitoria MA, Iglesias MJ, Guimbao J, Lázaro MA, Rastogi N, Revillo MJ, Martín C, Samper S. 2009. "Unsuspected and extensive transmission of a drug-susceptible *Mycobacterium tuberculosis* strain." BMC Pulm.Med. 9: 3.

Lupski, J. R. & Weinstock, G. M. 1992, "Short, interspersed repetitive DNA sequences in prokaryotic genomes", The Journal of Bacteriology, vol. 174, no. 14, pp. 4525-4529.

Luria, S. E. & Delbruck, M. 1943, " Mutations of Bacteria from Virus Sensitivity to Virus Resistance", Genetics, vol. 28, no. 6, pp. 491-511.

MacGurn JA, Raghavan S, Stanley SA, Cox JS. 2005. "A non-RD1 gene cluster is required for Snm secretion in *Mycobacterium tuberculosis*." Mol.Microbiol. 57 (6): 1653-63.

Magdalena, J., Supply, P., & Locht, C. 1998, "Specific differentiation between Mycobacterium bovis BCG and virulent strains of the *Mycobacterium tuberculosis* complex", Journal of Clinical Microbiology, vol. 36, no. 9, pp. 2471-2476.

Magdalena, J., Vachee, A., Supply, P., & Locht, C. 1998, "Identification of a new DNA region specific for members of *Mycobacterium tuberculosis* complex", Journal of Clinical Microbiology, vol. 36, no. 4, pp. 937-943.

Maguire, H., Dale, J. W., McHugh, T. D., Butcher, P. D., Gillespie, S. H., Costetsos, A., Al-Ghusein, H., Holland, R., Dickens, A., Marston, L., Wilson, P., Pitman, R., Strachan, D., Drobniewski, F. A., & Banerjee, D. K. 2002, "Molecular epidemiology of tuberculosis in London 1995-7 showing low rate of active transmission", Thorax, vol. 57, no. 7, pp. 617-622.

Maguire H, Forrester, S., & Adam, S. 2006. Isoniazid Resistant TB outbreak in London 2000 to 2006. Health Protection Agency. Progress report and updated recommendations of the London Outbreak Control Committee.

Mahairas GG, Sabo PJ, Hickey MJ, Singh DC, Stover CK. 1996. "Molecular analysis of genetic differences between Mycobacterium bovis BCG and virulent *M. bovis*." The Journal of Bacteriology 178 (5): 1274-82.

Mahmoudi, A. & Iseman, M. D. 1993, "Pitfalls in the care of patients with tuberculosis. Common errors and their association with the acquisition of drug resistance", JAMA, vol. 270, no. 1, pp. 65-68.

Maiden, M. C. 1998, "Horizontal genetic exchange, evolution, and spread of antibiotic resistance in bacteria", Clin.Infect.Dis., vol. 27 Suppl 1, p. S12-S20.

Maisnier-Patin, S. & Andersson, D. I. 2004, "Adaptation to the deleterious effects of antimicrobial drug resistance mutations by compensatory evolution", Research in Microbiology, vol. 155, no. 5, pp. 360-369.

Manca, C., Reed, M. B., Freeman, S., Mathema, B., Kreiswirth, B., Barry, C. E., III, & Kaplan, G. 2004, "Differential monocyte activation underlies strain-specific *Mycobacterium tuberculosis* pathogenesis", Infect.Immun., vol. 72, no. 9, pp. 5511-5514.

Manca, C., Tsenova, L., Freeman, S., Barczak, A. K., Tovey, M., Murray, P. J., Barry, C., & Kaplan, G. 2005, "Hypervirulent M. tuberculosis W/Beijing strains upregulate type I IFNs and increase expression of negative regulators of the Jak-Stat pathway", J.Interferon Cytokine Res., vol. 25, no. 11, pp. 694-701.

Manges, A. R., Tellis, P. A., Vincent, C., Lifeso, K., Geneau, G., Reid-Smith, R. J., & Boerlin, P. 2009, "Multi-locus variable number tandem repeat analysis for Escherichia coli causing extraintestinal infections", J.Microbiol.Methods, vol. 79, no. 2, pp. 211-213.

Mariam, D. H., Mengistu, Y., Hoffner, S. E., & Andersson, D. I. 2004, "Effect of rpoB mutations conferring rifampin resistance on fitness of *Mycobacterium tuberculosis*", Antimicrobial Agents and Chemotherapy, vol. 48, no. 4, pp. 1289-1294.

Marsh, J. W., O'Leary, M. M., Shutt, K. A., Pasculle, A. W., Johnson, S., Gerding, D. N., Muto, C. A., & Harrison, L. H. 2006, "Multilocus variable-number tandem-repeat analysis for investigation of Clostridium difficile transmission in Hospitals", Journal of Clinical Microbiology, vol. 44, no. 7, pp. 2558-2566.

Marshall BG, Mitchell DM, Shaw RJ, Marais F, Watkins RM, Coker RJ. 1999. HIV and tuberculosis co-infection in an inner London hospital--a prospective anonymized seroprevalence study. Journal Infection 38(3):162-6.

Martineau AR,Timms PM, Bothamley GH, Hanifa Y, Islam K, Claxton AP, Packe GE, Moore-Gillon JC, Darmalingam M, Davidson RN, Milburn HJ, Baker LV, Barker RD, Woodward NJ, Venton TR, Barnes KE, Mullett CJ, Coussens AK, Rutterford CM, Mein CA, Davies GR, Wilkinson RJ, Nikolayevskyy V, Drobniewski FA, Eldridge SM, Griffiths CJ. 2011. High-dose vitamin D3 during intensive-phase antimicrobial treatment of pulmonary tuberculosis: a double-blind randomised controlled trial. Lancet; 377: 242–50

Marttila, H. J., Soini, H., Huovinen, P., & Viljanen, M. K. 1996, "katG mutations in isoniazid-resistant *Mycobacterium tuberculosis* isolates recovered from Finnish patients", Antimicrobial Agents and Chemotherapy, vol. 40, no. 9, pp. 2187-2189.

Mathema, B., Kurepina, N. E., Bifani, P. J., & Kreiswirth, B. N. 2006, "Molecular epidemiology of tuberculosis: current insights", Clin.Microbiol.Rev, vol. 19, no. 4, pp. 658-685.

Matteelli, A., Migliori, G. B., Cirillo, D., Centis, R., Girard, E., & Raviglion, M. 2007, "Multidrug-resistant and extensively drug-resistant *Mycobacterium tuberculosis*: epidemiology and control", Expert.Rev Anti.Infect.Ther., vol. 5, no. 5, pp. 857-871.

Mazars, E., Lesjean, S., Banuls, A. L., Gilbert, M., Vincent, V., Gicquel, B., Tibayrenc, M., Locht, C., & Supply, P. 2001, "High-resolution minisatellite-based typing as a portable approach to global analysis of *Mycobacterium tuberculosis* molecular epidemiology", Proc.Natl.Acad.Sci.U.S.A, vol. 98, no. 4, pp. 1901-1906.

Mazurek, G. H., Cave, M. D., Eisenach, K. D., Wallace, R. J., Jr., Bates, J. H., & Crawford, J. T. 1991, "Chromosomal DNA fingerprint patterns produced with IS6110 as strain-specific markers for epidemiologic study of tuberculosis", Journal of Clinical Microbiology, vol. 29, no. 9, pp. 2030-2033.

McAdam, R. A., Hermans, P. W., van, S. D., Zainuddin, Z. F., Catty, D., van Embden, J. D., & Dale, J. W. 1990, "Characterization of a *Mycobacterium tuberculosis* insertion sequence belonging to the IS3 family", Mol.Microbiol., vol. 4, no. 9, pp. 1607-1613.

McClure, W. R. & Cech, C. L. 1978, "On the mechanism of rifampicin inhibition of RNA synthesis", J.Biol.Chem., vol. 253, no. 24, pp. 8949-8956.

McHugh, T. D., Newport, L. E., & Gillespie, S. H. 1997, "IS6110 homologs are present in multiple copies in mycobacteria other than tuberculosis-causing mycobacteria", Journal of Clinical Microbiology, vol. 35, no. 7, pp. 1769-1771.

McHugh, T. D. & Gillespie, S. H. 1998, "Nonrandom association of IS6110 and *Mycobacterium tuberculosis*: implications for molecular epidemiological studies", Journal of Clinical Microbiology, vol. 36, no. 5, pp. 1410-1413.

McHugh, T. D., Pope, C. F., Ling, C. L., Patel, S., Billington, O. J., Gosling, R. D., Lipman, M. C., & Gillespie, S. H. 2004, "Prospective evaluation of BDProbeTec strand displacement amplification (SDA) system for diagnosis of tuberculosis in non-respiratory and respiratory samples", J.Med.Microbiol., vol. 53, no. Pt 12, pp. 1215-1219.

McHugh, T. D., Batt, S. L., Shorten, R. J., Gosling, R. D., Uiso, L., & Gillespie, S. H. 2005, "*Mycobacterium tuberculosis* lineage: A naming of the parts", Tuberculosis, vol. 85, no. 3, pp. 127-136.

Means, T. K., Lien, E., Yoshimura, A., Wang, S., Golenbock, D. T., & Fenton, M. J. 1999a, "The CD14 ligands lipoarabinomannan and lipopolysaccharide differ in their requirement for Toll-like receptors", J.Immunol., vol. 163, no. 12, pp. 6748-6755.

Means, T. K., Wang, S., Lien, E., Yoshimura, A., Golenbock, D. T., & Fenton, M. J. 1999b, "Human toll-like receptors mediate cellular activation by *Mycobacterium tuberculosis*", J.Immunol., vol. 163, no. 7, pp. 3920-3927.

Meier, A., Kirschner, P., Bange, F. C., Vogel, U., & Bottger, E. C. 1994, "Genetic alterations in streptomycin-resistant *Mycobacterium tuberculosis*: mapping of mutations conferring resistance", Antimicrobial Agents and Chemotherapy, vol. 38, no. 2, pp. 228-233.

Meier, A., Sander, P., Schaper, K. J., Scholz, M., & Bottger, E. C. 1996, "Correlation of molecular resistance mechanisms and phenotypic resistance levels in streptomycin-resistant *Mycobacterium tuberculosis*", Antimicrobial Agents and Chemotherapy, vol. 40, no. 11, pp. 2452-2454.

Metzker, M. L. 2010, "Sequencing technologies - the next generation", Nat Rev Genet., vol. 11, no. 1, pp. 31-46.

Middlebrook, G. & Cohn, M. L. 1953, "Some observations on the pathogenicity of isoniazid-resistant variants of tubercle bacilli", Science, vol. 118, no. 3063, pp. 297-299.

Middlebrook, G. 1954, "Isoniazid-resistance and catalase activity of tubercle bacilli; a preliminary report", Am.Rev Tuberc., vol. 69, no. 3, pp. 471-472.

Middlebrook, G., Cohn, M. L., & Schaefer, W. B. 1954, "Studies on isoniazid and tubercle bacilli. III. The isolation, drug-susceptibility, and catalase-testing of tubercle bacilli from isoniazid-treated patients", Am.Rev Tuberc., vol. 70, no. 5, pp. 852-872.

Miller, L. P., Crawford, J. T., & Shinnick, T. M. 1994, "The rpoB gene of *Mycobacterium tuberculosis*", Antimicrobial Agents and Chemotherapy, vol. 38, no. 4, pp. 805-811.

Miller, M. B. and Y. W. Tang. 2009. "Basic concepts of microarrays and potential applications in clinical microbiology." Clin.Microbiol.Rev 22 (4): 611-33.

Mitchison DA 1954, "Tubercle bacilli resistant to isoniazid; virulence and response to treatment with isoniazid in guinea-pigs", Br.Med.J..vol. 1, no. 4854, pp. 128-130.

Mitchison DA, Keyes AB, Edwards EA, Ayuma P, Byfield SP, Nunn AJ. 1980. "Quality control in tuberculosis bacteriology. 2. The origin of isolated positive cultures from the sputum of patients in four studies of short course chemotherapy in Africa." Tubercle. 61 (3): 135-44.

Mizuno, K., Chikamatsu, K., Aono, A., Azuma, Y., Yamada, H., & Mitarai, S. 2009, "Clinical evaluation of acid-fast smear examination with light emitting diode fluorescent microscopy", Kekkaku, vol. 84, no. 9, pp. 627-629.

Moazed, D. & Noller, H. F. 1987, "Interaction of antibiotics with functional sites in 16S ribosomal RNA", Nature, vol. 327, no. 6121, pp. 389-394.

Modi, R. I. & Adams, J. 1991. "Coevolution in Bacterial-Plasmid Populations." Evolution 45 (3): 656-667.

Mokrousov, I., Ly, H. M., Otten, T., Lan, N. N., Vyshnevskyi, B., Hoffner, S., & Narvskaya, O. 2005, "Origin and primary dispersal of the *Mycobacterium tuberculosis* Beijing genotype: clues from human phylogeography", Genome Res., vol. 15, no. 10, pp. 1357-1364.

Moore, D. A., Evans, C. A., Gilman, R. H., Caviedes, L., Coronel, J., Vivar, A., Sanchez, E., Pinedo, Y., Saravia, J. C., Salazar, C., Oberhelman, R., Hollm-Delgado, M. G., LaChira, D., Escombe, A. R., & Friedland, J. S. 2006, "Microscopic-observation drug-susceptibility assay for the diagnosis of TB", N.Engl.J.Med., vol. 355, no. 15, pp. 1539-1550.

Moore, F. B., Rozen, D. E., & Lenski, R. E. 2000, "Pervasive compensatory adaptation in Escherichia coli", Proc.Biol.Sci., vol. 267, no. 1442, pp. 515-522.

Morris, S., Bai, G. H., Suffys, P., Portillo-Gomez, L., Fairchok, M., & Rouse, D. 1995, "Molecular mechanisms of multiple drug resistance in clinical isolates of *Mycobacterium tuberculosis*", J.Infect.Dis., vol. 171, no. 4, pp. 954-960.

Moss, A. R., Alland, D., Telzak, E., Hewlett, D., Jr., Sharp, V., Chiliade, P., LaBombardi, V., Kabus, D., Hanna, B., Palumbo, L., Brudney, K., Weltman, A., Stoeckle, K., Chirgwin, K., Simberkoff, M., Moghazeh, S., Eisner, W., Lutfey, M., & Kreiswirth, B. 1997, "A city-wide outbreak of a multiple-drug-resistant strain of *Mycobacterium tuberculosis* in New York", Int.J.Tuberc.Lung Dis., vol. 1, no. 2, pp. 115-121.

Mostowy, S., Cousins, D., Brinkman, J., Aranaz, A., & Behr, M. A. 2002, "Genomic deletions suggest a phylogeny for the *Mycobacterium tuberculosis* complex", J.Infect.Dis., vol. 186, no. 1, pp. 74-80.

Motulsky, H. 1995. Intuitive Biostatistics. Oxford University Press.

Musser, J. M. 1995, "Antimicrobial agent resistance in mycobacteria: molecular genetic insights", Clin.Microbiol.Rev, vol. 8, no. 4, pp. 496-514.

Musser, J. M., Kapur, V., Williams, D. L., Kreiswirth, B. N., van, S. D., & van Embden, J. D. 1996, "Characterization of the catalase-peroxidase gene (katG) and inhA locus in isoniazid-resistant and -susceptible strains of *Mycobacterium tuberculosis* by automated DNA sequencing: restricted array of mutations associated with drug resistance", J.Infect.Dis., vol. 173, no. 1, pp. 196-202.

Musser, J. M., Amin, A., & Ramaswamy, S. 2000, "Negligible Genetic Diversity of *Mycobacterium tuberculosis* Host Immune System Protein Targets: Evidence of Limited Selective Pressure", Genetics, vol. 155, no. 1, pp. 7-16.

Mustafa AS, Amoudy HA, Wiker HG, Abal AT, Ravn P, Oftung F, Andersen P. 1998. "Comparison of antigen-specific T-cell responses of tuberculosis patients using complex or single antigens of *Mycobacterium tuberculosis*." Scand.J.Immunol. 48 (5): 535-43.

Muyoyeta, M., Schaap, J. A., De, H. P., Mwanza, W., Muvwimi, M. W., Godfrey-Faussett, P., & Ayles, H. 2009, "Comparison of four culture systems for *Mycobacterium tuberculosis* in the Zambian National Reference Laboratory", Int.J.Tuberc.Lung Dis., vol. 13, no. 4, pp. 460-465.

Nagvenkar P, Udwadia Z, Pathan AA, Wilkinson KA, Shastri JS, Ewer K, Hill AV, Mehta A, Rodrigues C. 2001. "Enumeration of T cells specific for RD1-encoded antigens suggests a high prevalence of latent *Mycobacterium tuberculosis* infection in healthy urban Indians." J.Infect.Dis. 183 (3): 469-77.

Nair, J., Rouse, D. A., Bai, G. H., & Morris, S. L. 1993, "The rpsL gene and streptomycin resistance in single and multiple drug-resistant strains of *Mycobacterium tuberculosis*", Mol.Microbiol., vol. 10, no. 3, pp. 521-527.

Nakamura, Y., Leppert, M., O'Connell, P., Wolff, R., Holm, T., Culver, M., Martin, C., Fujimoto, E., Hoff, M., Kumlin, E., 1987, "Variable number of tandem repeat (VNTR) markers for human gene mapping", Science, vol. 235, no. 4796, pp. 1616-1622.

Natarajan K, Latchumanan VK, Singh B, Singh S, Sharma P. 2003. "Down-regulation of T helper 1 responses to mycobacterial antigens due to maturation of dendritic cells by 10-kDa *Mycobacterium tuberculosis* secretory antigen." J.Infect.Dis. 187 (6): 914-28.

Naya, H., Romero, H. +., Zavala, A., Alvarez, B., & Musto, H. +. 2002, "Aerobiosis Increases the Genomic Guanine Plus Cytosine Content (GC%) in Prokaryotes", Journal of Molecular Evolution, vol. 55, no. 3, pp. 260-264.

Neely, F., Maguire, H., Le, B. F., Davies, A., Gelb, D., & Yates, S. 2010, "High rate of transmission among contacts in large London outbreak of isoniazid mono-resistant tuberculosis", J.Public Health (Oxf), vol. 32, no. 1, pp. 44-51.

NICE. NICE Tuberculosis Guidelines 2006. National Institute for Health and Clinical Excellence. Clinical diagnosis and management of tuberculosis, and measures for its prevention and control.

Niemann, S., Richter, E., & Rusch-Gerdes, S. 1999, "Stability of *Mycobacterium tuberculosis* IS6110 restriction fragment length polymorphism patterns and spoligotypes determined by analyzing serial isolates from patients with drug-resistant tuberculosis", Journal of Clinical Microbiology, vol. 37, no. 2, pp. 409-412.

Nitta AT, Davidson PT, de Koning ML, Kilman RJ. 1996. "Misdiagnosis of multidrug-resistant tuberculosis possibly due to laboratory-related errors." JAMA 276 (24): 1980-83.

Nnoaham, K. E. & Clarke, A. 2008, "Low serum vitamin D levels and tuberculosis: a systematic review and meta-analysis", Int.J.Epidemiol., vol. 37, no. 1, pp. 113-119.

Noller, H. F. 1984, "Structure of ribosomal RNA", Annu.Rev Biochem., vol. 53, pp. 119-162.

Northrup JM, Miller AC, Nardell E, Sharnprapai S, Etkind S, Driscoll J, McGarry M, Taber HW, Elvin P, Qualls NL, Braden CR. 2002. "Estimated costs of false laboratory diagnoses of tuberculosis in three patients." Emerg.Infect.Dis. 8 (11): 1264-70.

Noss, E. H., Pai, R. K., Sellati, T. J., Radolf, J. D., Belisle, J., Golenbock, D. T., Boom, W. H., & Harding, C. V. 2001, "Toll-like receptor 2-dependent inhibition of macrophage class II MHC expression and antigen processing by 19-kDa lipoprotein of *Mycobacterium tuberculosis*", J.Immunol., vol. 167, no. 2, pp. 910-918.

O'Rourke, M. & Stevens, E. 1993, "Genetic structure of Neisseria gonorrhoeae populations: a non-clonal pathogen", J.Gen.Microbiol., vol. 139, no. 11, pp. 2603-2611.

O'Rourke, M. & Spratt, B. G. 1994, "Further evidence for the non-clonal population structure of Neisseria gonorrhoeae: extensive genetic diversity within isolates of the same electrophoretic type", Microbiology, vol. 140 (Pt 6), pp. 1285-1290.

O'Sullivan, D. M 2007. "Evolution of the mechanisms of drug resistance in *M. tuberculosis*". PhD thesis. University College London.

O'Sullivan, D. M., McHugh, T. D., & Gillespie, S. H. 2008, "The effect of oxidative stress on the mutation rate of *Mycobacterium tuberculosis* with impaired catalase/peroxidase function", J.Antimicrob.Chemother., vol. 62, no. 4, pp. 709-712.

O'Sullivan DM, McHugh TD, Gillespie SH. 2010. Mapping the fitness of Mycobacterium tuberculosis strains: a complex picture. J Med Microbiol. 59(12):1533-5

Ochman, H., Lawrence, J. G., & Groisman, E. A. 2000, "Lateral gene transfer and the nature of bacterial innovation", Nature, vol. 405, no. 6784, pp. 299-304.

Olobo, J. O., Geletu, M., Demissie, A., Eguale, T., Hiwot, K., Aderaye, G., & Britton, S. 2001, "Circulating TNF-alpha, TGF-beta, and IL-10 in tuberculosis patients and healthy contacts", Scand.J.Immunol., vol. 53, no. 1, pp. 85-91.

Ordway, D. J., Costa, L., Martins, M., Silveira, H., Amaral, L., Arroz, M. J., Ventura, F. A., & Dockrell, H. M. 2004, "Increased Interleukin-4 production by CD8 and gammadelta T cells in health-care workers is associated with the subsequent development of active tuberculosis", J.Infect.Dis., vol. 190, no. 4, pp. 756-766.

Orme, I. M. & Cooper, A. M. 1999, "Cytokine/chemokine cascades in immunity to tuberculosis", Immunol.Today, vol. 20, no. 7, pp. 307-312.

Ormerod, L. P. 1992. "Tuberculosis contact tracing. Blackburn 1982–90." Respir.Med. 87(2);,127-131.

Otal, I., Martin, C., Vincent-Levy-Frebault, V., Thierry, D., & Gicquel, B. 1991, "Restriction fragment length polymorphism analysis using IS6110 as an epidemiological marker in tuberculosis", Journal of Clinical Microbiology, vol. 29, no. 6, pp. 1252-1254.

Ottenhoff, T. H., Verreck, F. A., Hoeve, M. A., & van de Vosse 2005, "Control of human host immunity to mycobacteria", Tuberculosis.(Edinb.), vol. 85, no. 1-2, pp. 53-64.

Pai, R. K., Pennini, M. E., Tobian, A. A., Canaday, D. H., Boom, W. H., & Harding, C. V. 2004, "Prolonged toll-like receptor signaling by *Mycobacterium tuberculosis* and its 19-kilodalton lipoprotein inhibits gamma interferon-induced regulation of selected genes in macrophages", Infect.Immun., vol. 72, no. 11, pp. 6603-6614.

Pallen, M. J. 2002, "The ESAT-6/WXG100 superfamily -- and a new Gram-positive secretion system?", Trends Microbiol., vol. 10, no. 5, pp. 209-212.

Palomino, J. C., Martin, A., Von, G. A., & Portaels, F. 2008, "Rapid culture-based methods for drug-resistance detection in *Mycobacterium tuberculosis*", J.Microbiol.Methods, vol. 75, no. 2, pp. 161-166.

Park, S. H. & Bendelac, A. 2000, "CD1-restricted T-cell responses and microbial infection", Nature, vol. 406, no. 6797, pp. 788-792.

Parrish, N., Dionne, K., Sweeney, A., Hedgepeth, A., & Carroll, K. 2009, "Differences in time to detection and recovery of Mycobacterium spp. between the MGIT 960 and the BacT/ALERT MB automated culture systems", Diagn.Microbiol.Infect.Dis., vol. 63, no. 3, pp. 342-345.

Parwati, I., van Crevel R and van Soolingen D. 2010. "Possible underlying mechanisms for successful emergence of the *Mycobacterium tuberculosis* Beijing genotype strains." Lancet Infect.Dis. 10 (2): 103-11.

Patel, N. R., Zhu, J., Tachado, S. D., Zhang, J., Wan, Z., Saukkonen, J., & Koziel, H. 2007, "HIV impairs TNF-alpha mediated macrophage apoptotic response to *Mycobacterium tuberculosis*", J.Immunol., vol. 179, no. 10, pp. 6973-6980.

Pathak, S. K., Basu, S., Basu, K. K., Banerjee, A., Pathak, S., Bhattacharyya, A., Kaisho, T., Kundu, M., & Basu, J. 2007, "Direct extracellular interaction between the early secreted antigen ESAT-6 of *Mycobacterium tuberculosis* and TLR2 inhibits TLR signaling in macrophages", Nat Immunol., vol. 8, no. 6, pp. 610-618.

Pathan AA, Wilkinson KA, Klenerman P, McShane H, Davidson RN, Pasvol G, Hill AV, Lalvani AJ. 2001. Direct ex vivo analysis of antigen-specific IFN-gamma-secreting CD4

T cells in *Mycobacterium tuberculosis*-infected individuals: associations with clinical disease state and effect of treatment. J Immunol. 167(9): 5217-25.

Pepper, T., Joseph, P., Mwenya, C., McKee, G. S., Haushalter, A., Carter, A., Warkentin, J., Haas, D. W., & Sterling, T. R. 2008, "Normal chest radiography in pulmonary tuberculosis: implications for obtaining respiratory specimen cultures", Int.J.Tuberc.Lung Dis., vol. 12, no. 4, pp. 397-403.

Perna, N. T., Plunkett, G., Burland, V., Mau, B., Glasner, J. D., Rose, D. J., Mayhew, G. F., Evans, P. S., Gregor, J., Kirkpatrick, H. A., Posfai, G., Hackett, J., Klink, S., Boutin, A., Shao, Y., Miller, L., Grotbeck, E. J., Davis, N. W., Lim, A., Dimalanta, E. T., Potamousis, K. D., Apodaca, J., Anantharaman, T. S., Lin, J., Yen, G., Schwartz, D. C., Welch, R. A., & Blattner, F. R. 2001, "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7", Nature, vol. 409, no. 6819, pp. 529-533.

Perneger, T. V., Sudre, P., Lundgren, J. D., & Hirschel, B. 1995, "Does the onset of tuberculosis in AIDS predict shorter survival? Results of a cohort study in 17 European countries over 13 years. AIDS in Europe Study Group", BMJ, vol. 311, no. 7018, pp. 1468-1471.

Perrin, F. M., Breen, R. A., Lipman, M. C., Shorten, R. J., Gillespie, S. H., & McHugh, T. D. 2005, "Is there a relationship between *Mycobacterium tuberculosis* strain type and TB paradoxical reaction?", Thorax, vol. 60, no. 8, pp. 706-707.

Pitout, J. D. & Laupland, K. B. 2008, "Extended-spectrum beta-lactamase-producing Enterobacteriaceae: an emerging public-health concern", Lancet Infect.Dis., vol. 8, no. 3, pp. 159-166.

Poulet, S. & Cole ST. 1995a. Characterization of the highly abundant polymorphic GC-rich-repetitive sequence (PGRS) present in *Mycobacterium tuberculosis*. Archives of Microbiology 163 (2); 87-95.

Poulet, S. & Cole ST. 1995b. Repeated DNA sequences in mycobacteria. Archives of Microbiology 163 (2); 79-86.

Pym, A. S., Brodin, P., Brosch, R., Huerre, M., & Cole, S. T. 2002a, "Loss of RD1 contributed to the attenuation of the live tuberculosis vaccines Mycobacterium bovis BCG and Mycobacterium microti", Mol.Microbiol., vol. 46, no. 3, pp. 709-717.

Pym, A. S., Saint-Joanis, B., & Cole, S. T. 2002b, "Effect of katG mutations on the virulence of *Mycobacterium tuberculosis* and the implication for transmission in humans", Infect.Immun., vol. 70, no. 9, pp. 4955-4960.

Pym AS, Brodin P, Majlessi L, Brosch R, Demangel C, Williams A, Griffiths KE, Marchal G, Leclerc C, Cole ST. 2003. "Recombinant BCG exporting ESAT-6 confers enhanced protection against tuberculosis." Nat Med. 9 (5): 533-39.

Quesniaux, V., Fremond, C., Jacobs, M., Parida, S., Nicolle, D., Yeremeev, V., Bihl, F., Erard, F., Botha, T., Drennan, M., Soler, M. N., Le, B. M., Schnyder, B., & Ryffel, B. 2004, "Toll-like receptor pathways in the immune responses to mycobacteria", Microbes.Infect., vol. 6, no. 10, pp. 946-959.

Ramaswamy, S. & Musser, J. M. 1998, "Molecular genetic basis of antimicrobial agent resistance in *Mycobacterium tuberculosis*: 1998 update", Tuber.Lung Dis., vol. 79, no. 1, pp. 3-29.

Ravn P, Demissie A, Eguale T, Wondwosson H, Lein D, Amoudy HA, Mustafa AS, Jensen AK, Holm A, Rosenkrands I, Oftung F, Olobo J, von Reyn F, Andersen P. 1999. "Human T cell responses to the ESAT-6 antigen from *Mycobacterium tuberculosis*." J.Infect.Dis. 179 (3): 637-45.

Redpath, S., Ghazal, P., & Gascoigne, N. R. 2001, "Hijacking and exploitation of IL-10 by intracellular pathogens", Trends Microbiol., vol. 9, no. 2, pp. 86-92.

Reed, M. B., Domenech, P., Manca, C., Su, H., Barczak, A. K., Kreiswirth, B. N., Kaplan, G., & Barry, C. E., III 2004a, "A glycolipid of hypervirulent tuberculosis strains that inhibits the innate immune response", Nature, vol. 431, no. 7004, pp. 84-87.

Reed, M. B., Gagneux, S., DeRiemer, K., Small, P. M., & Barry, C. E., III 2007, "The W-Beijing Lineage of *Mycobacterium tuberculosis* Overproduces Triglycerides and Has the DosR Dormancy Regulon Constitutively Upregulated", The Journal of Bacteriology, vol. 189, no. 7, pp. 2583-2589.

Reed, M. B., Pichler, V. K., McIntosh, F., Mattia, A., Fallow, A., Masala, S., Domenech, P., Zwerling, A., Thibert, L., Menzies, D., Schwartzman, K., & Behr, M. A. 2009, "Major *Mycobacterium tuberculosis* Lineages Associate with Patient Country of Origin", Journal of Clinical Microbiology, vol. 47, no. 4, pp. 1119-1128.

Rengarajan, Jyothi, Barry R. Bloom, and Eric J. Rubin. 2005. "Genome-wide requirements for *Mycobacterium tuberculosis* adaptation and survival in macrophages." Proceedings of the National Academy of Sciences of the United States of America 102 (23): 8327-32.

Renshaw PS, Lightbody KL, Veverka V, Muskett FW, Kelly G, Frenkiel TA, Gordon SV, Hewinson RG, Burke B, Norman J, Williamson RA, Carr MD. 2005 "Structure and function of the complex formed by the tuberculosis virulence factors CFP-10 and ESAT-6." EMBO J. 24 (14): 2491-98.

Reynolds, M. G. 2000, "Compensatory evolution in rifampin-resistant Escherichia coli", Genetics, vol. 156, no. 4, pp. 1471-1481.

Rhee, J. T., Tanaka, M. M., Behr, M. A., Agasino, C. B., Paz, E. A., Hopewell, P. C., & Small, P. M. 2000, "Use of multiple markers in population-based molecular epidemiologic studies of tuberculosis", Int.J.Tuberc.Lung Dis., vol. 4, no. 12, pp. 1111-1119.

Richeldi, L. 2006, "An update on the diagnosis of tuberculosis infection", Am.J.Respir.Crit Care Med., vol. 174, no. 7, pp. 736-742.

Roberts, T., Beyers, N., Aguirre, A., & Walzl, G. 2007, "Immunosuppression during active tuberculosis is characterized by decreased interferon- gamma production and CD25 expression with elevated forkhead box P3, transforming growth factor- beta , and interleukin-4 mRNA levels", J.Infect.Dis., vol. 195, no. 6, pp. 870-878.

Rodger, A. J., Story, A., Fox, Z., & Hayward, A. 2010, "HIV prevalence and testing practices among tuberculosis cases in London: a missed opportunity for HIV diagnosis?", Thorax, vol. 65, no. 1, pp. 63-69.

Rodrigues, L. C., Pereira, S. M., Cunha, S. S., Genser, B., Ichihara, M. Y., de Brito, S. C., Hijjar, M. A., Dourado, I., Cruz, A. A., Sant'Anna, C., Bierrenbach, A. L., & Barreto, M. L. 2005, "Effect of BCG revaccination on incidence of tuberculosis in school-aged children in Brazil: the BCG-REVAC cluster-randomised trial", Lancet, vol. 366, no. 9493, pp. 1290-1295.

Roring, S., Scott, A., Brittain, D., Walker, I., Hewinson, G., Neill, S., & Skuce, R. 2002, "Development of variable-number tandem repeat typing of Mycobacterium bovis: comparison of results with those obtained by using existing exact tandem repeats and spoligotyping", Journal of Clinical Microbiology, vol. 40, no. 6, pp. 2126-2133.

Roring, S., Scott, A. N., Glyn, H. R., Neill, S. D., & Skuce, R. A. 2004, "Evaluation of variable number tandem repeat (VNTR) loci in molecular typing of Mycobacterium bovis isolates from Ireland", Vet.Microbiol., vol. 101, no. 1, pp. 65-73.

Rosenberg, N. A., Tsolaki, A. G., & Tanaka, M. M. 2003, "Estimating change rates of genetic markers using serial samples: applications to the transposon IS6110 in *Mycobacterium tuberculosis*", Theor.Popul.Biol., vol. 63, no. 4, pp. 347-363.

Ross, B. C., Raios, K., Jackson, K., & Dwyer, B. 1992, "Molecular cloning of a highly repeated DNA element from *Mycobacterium tuberculosis* and its use as an epidemiological tool", Journal of Clinical Microbiology, vol. 30, no. 4, pp. 942-946.

Rossau, R., Traore, H., De Beenhouwer, H., Mijs, W., Jannes, G., De Rijk, P., & Portaels, F. 1997, "Evaluation of the INNO-LiPA Rif. TB assay, a reverse hybridization assay for the simultaneous detection of *Mycobacterium tuberculosis* complex and its resistance to rifampin", Antimicrobial Agents and Chemotherapy, vol. 41, no. 10, pp. 2093-2098.

Rouse, D. A., Li, Z., Bai, G. H., & Morris, S. L. 1995, "Characterization of the katG and inhA genes of isoniazid-resistant clinical isolates of *Mycobacterium tuberculosis*", Antimicrobial Agents and Chemotherapy, vol. 39, no. 11, pp. 2472-2477.

Ruddy, M., McHugh, T. D., Dale, J. W., Banerjee, D., Maguire, H., Wilson, P., Drobniewski, F., Butcher, P., & Gillespie, S. H. 2002, "Estimation of the Rate of Unrecognized Cross-Contamination with *Mycobacterium tuberculosis* in London

Microbiology Laboratories", Journal of Clinical Microbiology, vol. 40, no. 11, pp. 4100-4104.

Ruddy, M. C., Davies, A. P., Yates, M. D., Yates, S., Balasegaram, S., Drabu, Y., Patel, B., Lozewicz, S., Sen, S., Bahl, M., James, E., Lipman, M., Duckworth, G., Watson, J. M., Piper, M., Drobniewski, F. A., & Maguire, H. 2004, "Outbreak of isoniazid resistant tuberculosis in north London", Thorax, vol. 59, no. 4, pp. 279-285.

Russell, D. G. 2001, "*Mycobacterium tuberculosis*: here today, and here tomorrow", Nat Rev Mol.Cell Biol., vol. 2, no. 8, pp. 569-577.

Ryll, R., Kumazawa, Y., & Yano, I. 2001, "Immunological properties of trehalose dimycolate (cord factor) and other mycolic acid-containing glycolipids--a review", Microbiol.Immunol., vol. 45, no. 12, pp. 801-811.

Safi, H., Aznar, J., & Palomares, J. C. 1997, "Molecular epidemiology of *Mycobacterium tuberculosis* strains isolated during a 3-year period (1993 to 1995) in Seville, Spain", Journal of Clinical Microbiology, vol. 35, no. 10, pp. 2472-2476.

Safi H, Barnes PF, Lakey DL, Shams H, Samten B, Vankayalapati R and Howard ST. 2004. IS*6110* functions as a mobile, monocyte-activated promoter in *Mycobacterium tuberculosis*Molecular Microbiology. 52(4), 999–1012

Sahiratmadja, E., Alisjahbana, B., de, B. T., Adnan, I., Maya, A., Danusantoso, H., Nelwan, R. H., Marzuki, S., van der Meer, J. W., van, C. R., van, d., V, & Ottenhoff, T. H. 2007, "Dynamic changes in pro- and anti-inflammatory cytokine profiles and gamma interferon receptor signaling integrity correlate with tuberculosis disease activity and response to curative treatment", Infect.Immun., vol. 75, no. 2, pp. 820-829.

Sander, P., Springer, B., Prammananan, T., Sturmfels, A., Kappler, M., Pletschette, M., & Bottger, E. C. 2002, "Fitness cost of chromosomal drug resistance-conferring mutations", Antimicrobial Agents and Chemotherapy, vol. 46, no. 5, pp. 1204-1211.

Saukkonen JJ, Cohn DL, Jasmer RM, Schenker S, Jereb JA, Nolan CM, Peloquin CA, Gordin FM, Nunes D, Strader DB, Bernardo J, Venkataramanan R, Sterling TR 2006. ATS (American Thoracic Society) Hepatotoxicity of Antituberculosis Therapy Subcommittee. "An Official ATS Statement: Hepatotoxicity of Antituberculosis Therapy." American Journal of Respiratory and Critical Care Medicine 174 (8): 935-52.

Saunders, B. M., Frank, A. A., Orme, I. M., & Cooper, A. M. 2002, "CD4 is required for the development of a protective granulomatous response to pulmonary tuberculosis", Cell Immunol., vol. 216, no. 1-2, pp. 65-72.

Savine, E., Warren, R. M., Van Der Spuy, G. D., Beyers, N., van Helden, P. D., Locht, C., & Supply, P. 2002, "Stability of variable-number tandem repeats of mycobacterial interspersed repetitive units from 12 loci in serial isolates of *Mycobacterium tuberculosis*", Journal of Clinical Microbiology, vol. 40, no. 12, pp. 4561-4566.

Scanga, C. A., Mohan, V. P., Yu, K., Joseph, H., Tanaka, K., Chan, J., & Flynn, J. L. 2000, "Depletion of CD4(+) T cells causes reactivation of murine persistent tuberculosis despite continued expression of interferon gamma and nitric oxide synthase 2", J.Exp.Med., vol. 192, no. 3, pp. 347-358.

Schaible, U. E. & Kaufmann, S. H. 2000, "CD1 molecules and CD1-dependent T cells in bacterial infections: a link from innate to acquired immunity?", Semin.Immunol., vol. 12, no. 6, pp. 527-535.

Schlesinger, L. S. 1993, "Macrophage phagocytosis of virulent but not attenuated strains of *Mycobacterium tuberculosis* is mediated by mannose receptors in addition to complement receptors", J.Immunol., vol. 150, no. 7, pp. 2920-2930.

Schmittgen, T. D. and B. A. Zakrajsek. 2000. "Effect of experimental treatment on housekeeping gene expression: validation by real-time, quantitative RT-PCR." J.Biochem.Biophys.Methods 46 (1-2): 69-81

Schrag, S. J., Perrot, V., & Levin, B. R. 1997, "Adaptation to the fitness costs of antibiotic resistance in Escherichia coli", Proc.Biol.Sci., vol. 264, no. 1386, pp. 1287-1291.

Schurch, A. C., Kremer, K., Daviena.O, Kiers, A., Boeree, M. J., Siezen, R. J., & van Soolingen D. 2010. "High resolution typing be integration of genome sequencing data in a large tuberculosis cluster." Journal of Clinical Microbiology. 48(9): 3403-6.

Seibert, F. B. & Dufour, E. 1948, "A study of certain problems in the use of standard tuberculin; fraction of PPD, standardization of tuberculins, and the question of sensitization", Am.Rev Tuberc., vol. 58, no. 3, pp. 363-374.

Selander, R. K., Caugant, D. A., Ochman, H., Musser, J. M., Gilmour, M. N., & Whittam, T. S. 1986, "Methods of multilocus enzyme electrophoresis for bacterial population genetics and systematics", Appl.Environ.Microbiol., vol. 51, no. 5, pp. 873-884.

Shelburne, S. A., Visnegarwala, F., Darcourt, J., Graviss, E. A., Giordano, T. P., White, A. C., Jr., & Hamill, R. J. 2005, "Incidence and risk factors for immune reconstitution inflammatory syndrome during highly active antiretroviral therapy", AIDS, vol. 19, no. 4, pp. 399-406.

Sherman, D. R., Mdluli, K., Hickey, M. J., Arain, T. M., Morris, S. L., Barry, C. E., III, & Stover, C. K. 1996, "Compensatory ahpC gene expression in isoniazid-resistant *Mycobacterium tuberculosis*", Science, vol. 272, no. 5268, pp. 1641-1643.

Shorten, R. J., Gillespie, S. H., Sule, O., Lipman, M., & McHugh, T. D. 2005, "Molecular strain typing of M. tuberculosis isolates from a suspected outbreak involving a faulty bronchoscope", J.Hosp.Infect., vol. 61, no. 1, pp. 86-87.

Sia IG, Orillaza RB, St Sauver JL, Quelapio ID, Lahr BD, Alcañeses RS, Wilson WR, Cockerill FR, Balane GI, Mangubat NV, Wengenack NL, Tupasi TE. 2010. "Tuberculosis attributed to household contacts in the Philippines." Int.J.Tuberc.Lung Dis. 14 (1): 122-25.

Sinha A, Salam N, Gupta S, Natarajan K. 2007. "*Mycobacterium tuberculosis* and dendritic cells: recognition, activation and functional implications." Indian J.Biochem.Biophys. 44 (5): 279-88.

Simpson. 1949. "Measurement of diversity." Nature 163; 688.

Singh, P. P., Parra, M., Cadieux, N., & Brennan, M. J. 2008, "A comparative study of host response to three *Mycobacterium tuberculosis* PE_PGRS proteins", Microbiology, vol. 154, no. Pt 11, pp. 3469-3479.

Skuce, R. A., McCorry, T. P., McCarroll, J. F., Roring, S. M., Scott, A. N., Brittain, D., Hughes, S. L., Hewinson, R. G., & Neill, S. D. 2002, "Discrimination of *Mycobacterium tuberculosis* complex bacteria using novel VNTR-PCR targets", Microbiology, vol. 148, no. Pt 2, pp. 519-528.

Small, P. M., Hopewell, P. C., Singh, S. P., Paz, A., Parsonnet, J., Ruston, D. C., Schecter, G. F., Daley, C. L., & Schoolnik, G. K. 1994, "The epidemiology of tuberculosis in San Francisco. A population-based study using conventional and molecular methods", N.Engl.J.Med., vol. 330, no. 24, pp. 1703-1709.

Smith, N. H., Dale, J., Inwald, J., Palmer, S., Gordon, S. V., Hewinson, R. G., & Smith, J. M. 2003, "The population structure of Mycobacterium bovis in Great Britain: Clonal expansion", Proceedings of the National Academy of Sciences of the United States of America, vol. 100, no. 25, pp. 15271-15275.

Smith, N. H., Gordon, S. V., de la Rua-Domenech, R., Clifton-Hadley, R. S., & Hewinson, R. G. 2006, "Bottlenecks and broomsticks: the molecular evolution of Mycobacterium bovis", Nat Rev Micro, vol. 4, no. 9, pp. 670-681.

Smittipat, N., Billamas, P., Palittapongarnpim, M., Thong-On, A., Temu, M. M., Thanakijcharoen, P., Karnkawinpong, O., & Palittapongarnpim, P. 2005, "Polymorphism of variable-number tandem repeats at multiple loci in *Mycobacterium tuberculosis*", Journal of Clinical Microbiology, vol. 43, no. 10, pp. 5034-5043.

Sodhi, A., Gong, J., Silva, C., Qian, D., & Barnes, P. F. 1997, "Clinical correlates of interferon gamma production in patients with tuberculosis", Clin.Infect.Dis., vol. 25, no. 3, pp. 617-620.

Soini, H., Pan, X., Teeter, L., Musser, J. M., & Graviss, E. A. 2001, "Transmission Dynamics and Molecular Characterization of *Mycobacterium tuberculosis* Isolates with Low Copy Numbers of IS6110", Journal of Clinical Microbiology, vol. 39, no. 1, pp. 217-221.

Somoskovi, A., Kodmon, C., Lantos, A., Bartfai, Z., Tamasi, L., Fuzy, J., & Magyar, P. 2000, "Comparison of Recoveries of *Mycobacterium tuberculosis* Using the Automated BACTEC MGIT 960 System, the BACTEC 460 TB System, and Lowenstein-Jensen Medium", Journal of Clinical Microbiology, vol. 38, no. 6, pp. 2395-2397.

Sorensen, A. L., Nagai, S., Houen, G., Andersen, P., & Andersen, A. B. 1995, "Purification and characterization of a low-molecular-mass T-cell antigen secreted by *Mycobacterium tuberculosis*", Infect.Immun., vol. 63, no. 5, pp. 1710-1717.

Sorlozano, A., Soria, I., Roman, J., Huertas, P., Soto, M. J., Piedrola, G., & Gutierrez, J. 2009, "Comparative evaluation of three culture methods for the isolation of mycobacteria from clinical samples", J.Microbiol.Biotechnol., vol. 19, no. 10, pp. 1259-1264.

Sosa LE, Lobato MN, Condren T, Williams MN, Hadler JL. 2008. Outbreak of tuberculosis in a correctional facility:consequences of missed opportunities. INT J TUBERC LUNG DIS 12(6):689–691

Sreevatsan, S., Pan, X., Stockbauer, K. E., Williams, D. L., Kreiswirth, B. N., & Musser, J. M. 1996, "Characterization of rpsL and rrs mutations in streptomycin-resistant *Mycobacterium tuberculosis* isolates from diverse geographic localities", Antimicrobial Agents and Chemotherapy, vol. 40, no. 4, pp. 1024-1026.

Sreevatsan, S., Pan, X., Stockbauer, K. E., Connell, N. D., Kreiswirth, B. N., Whittam, T. S., & Musser, J. M. 1997, "Restricted structural gene polymorphism in the *Mycobacterium tuberculosis* complex indicates evolutionarily recent global dissemination", Proceedings of the National Academy of Sciences of the United States of America, vol. 94, no. 18, pp. 9869-9874.

Stanley SA, Raghavan S, Hwang WW, Cox JS. 2003. "Acute infection and macrophage subversion by *Mycobacterium tuberculosis* require a specialized secretion system." Proceedings of the National Academy of Sciences of the United States of America 100 (22): 13001-06

Stanley SA, Johndrow JE, Manzanillo P, Cox JS. 2007. "The Type I IFN response to infection with *Mycobacterium tuberculosis* requires ESX-1-mediated secretion and contributes to pathogenesis." J.Immunol. 178 (5): 3143-52.

Steinberg, B. E. & Grinstein, S. 2008, "Pathogen destruction versus intracellular survival: the role of lipids as phagosomal fate determinants", J.Clin.Invest, vol. 118, no. 6, pp. 2002-2011.

Stenger, S. 2005, "Immunological control of tuberculosis: role of tumour necrosis factor and more", Ann.Rheum.Dis., vol. 64 Suppl 4, p. iv24-iv28.

Stenger, S. & Modlin, R. L. 2002, "Control of *Mycobacterium tuberculosis* through mammalian Toll-like receptors", Curr.Opin.Immunol., vol. 14, no. 4, pp. 452-457.

Sterne, J. A., Rodrigues, L. C., & Guedes, I. N. 1998, "Does the efficacy of BCG decline with time since vaccination?", Int.J.Tuberc.Lung Dis., vol. 2, no. 3, pp. 200-207.

Stokes, R. W., Norris-Jones, R., Brooks, D. E., Beveridge, T. J., Doxsee, D., & Thorson, L. M. 2004, "The glycan-rich outer layer of the cell wall of *Mycobacterium tuberculosis* acts as an antiphagocytic capsule limiting the association of the bacterium with macrophages", Infect.Immun., vol. 72, no. 10, pp. 5676-5686.

Story, A., Murad, S., Roberts, W., Verheyen, M., & Hayward, A. C. 2007, "Tuberculosis in London: the importance of homelessness, problem drug use and prison", Thorax, vol. 62, no. 8, pp. 667-671.

Story, A., Bothamley, G., & Hayward, A. 2008, "Crack cocaine and infectious tuberculosis", Emerg.Infect.Dis., vol. 14, no. 9, pp. 1466-1469.

Suerbaum, S., Smith, J. M., Bapumia, K., Morelli, G., Smith, N. H., Kunstmann, E., Dyrek, I., & Achtman, M. 1998, "Free recombination within Helicobacter pylori", Proceedings of the National Academy of Sciences of the United States of America, vol. 95, no. 21, pp. 12619-12624.

Supply, P., Magdalena, J., Himpens, S., & Locht, C. 1997, "Identification of novel intergenic repetitive units in a mycobacterial two-component system operon", Mol.Microbiol., vol. 26, no. 5, pp. 991-1003.

Supply, P., Mazars, E., Lesjean, S., Vincent, V., Gicquel, B., & Locht, C. 2000, "Variable human minisatellite-like regions in the *Mycobacterium tuberculosis* genome", Mol.Microbiol., vol. 36, no. 3, pp. 762-771.

Supply, P., Lesjean, S., Savine, E., Kremer, K., van Soolingen, D., & Locht, C. 2001, "Automated High-Throughput Genotyping for Study of Global Epidemiology of *Mycobacterium tuberculosis* Based on Mycobacterial Interspersed Repetitive Units", Journal of Clinical Microbiology, vol. 39, no. 10, pp. 3563-3571.

Supply, P., Warren, R. M., Banuls, A. L., Lesjean, S., Van Der Spuy, G. D., Lewis, L. A., Tibayrenc, M., van Helden, P. D., & Locht, C. 2003, "Linkage disequilibrium between minisatellite loci supports clonal evolution of *Mycobacterium tuberculosis* in a high tuberculosis incidence area", Mol.Microbiol., vol. 47, no. 2, pp. 529-538.

Supply, P. Multilocus Variable Number Tandem Repeat Genotyping of *Mycobacterium tuberculosis* Technical Guide. Institut de Biologie/Institut Pasteur de Lille. 2005.

Supply, P., Allix, C., Lesjean, S., Cardoso-Oelemann, M., Rusch-Gerdes, S., Willery, E., Savine, E., de Haas, P., van Deutekom, H., Roring, S., Bifani, P., Kurepina, N., Kreiswirth, B., Sola, C., Rastogi, N., Vatin, V., Gutierrez, M. C., Fauville, M., Niemann, S., Skuce, R., Kremer, K., Locht, C., & van Soolingen, D. 2006, "Proposal for Standardization of Optimized Mycobacterial Interspersed Repetitive Unit-Variable-Number Tandem Repeat Typing of *Mycobacterium tuberculosis*", Journal of Clinical Microbiology, vol. 44, no. 12, pp. 4498-4510.

Surikova, O. V., Voitech, D. S., Kuzmicheva, G., Tatkov, S. I., Mokrousov, I. V., Narvskaya, O. V., Rot, M. A., van, S. D., & Filipenko, M. L. 2005, "Efficient differentiation of *Mycobacterium tuberculosis* strains of the W-Beijing family from Russia using highly polymorphic VNTR loci", Eur.J.Epidemiol., vol. 20, no. 11, pp. 963-974.

Suzuki, Y., Yoshinaga, K., Ono, Y., Nagata, A., & Yamada, T. 1987, "Organization of rRNA genes in Mycobacterium bovis BCG", The Journal of Bacteriology, vol. 169, no. 2, pp. 839-843.

Takashima, T., Ueta, C., Tsuyuguchi, I., & Kishimoto, S. 1990, "Production of tumor necrosis factor alpha by monocytes from patients with pulmonary tuberculosis", Infect.Immun., vol. 58, no. 10, pp. 3286-3292.

Tan, T., Lee, W. L., Alexander, D. C., Grinstein, S., & Liu, J. 2006, "The ESAT-6/CFP-10 secretion system of Mycobacterium marinum modulates phagosome maturation", Cell Microbiol., vol. 8, no. 9, pp. 1417-1429.

Taylor, G. M., Young, D. B., & Mays, S. A. 2005, "Genotypic Analysis of the Earliest Known Prehistoric Case of Tuberculosis in Britain", Journal of Clinical Microbiology, vol. 43, no. 5, pp. 2236-2240.

Teale, C., Cundall, D. B., & Pearson, S. B. 1991, "Time of development of tuberculosis in contacts", Respir.Med., vol. 85, no. 6, pp. 475-477.

Telenti, A., Imboden, P., Marchesi, F., Lowrie, D., Cole, S., Colston, M. J., Matter, L., Schopfer, K., & Bodmer, T. 1993, "Detection of rifampicin-resistance mutations in *Mycobacterium tuberculosis*", Lancet, vol. 341, no. 8846, pp. 647-650.

Telenti, A.,Honoré N, Bernasconi C, March J, Ortega A, Heym B, Takiff HE, Cole ST 1997. "Genotypic assessment of isoniazid and rifampin resistance in *Mycobacterium tuberculosis*: a blind study at reference laboratory level." Journal of Clinical Microbiology 35(3): 719-23.

Teutschbein J, Schumann G, Möllmann U, Grabley S, Cole ST, Munder T. 2009. "A protein linkage map of the ESAT-6 secretion system 1 (ESX-1) of *Mycobacterium tuberculosis*." Microbiol.Res. 164(3): 253-59.

Then, R. L. 1982, "Mechanisms of resistance to trimethoprim, the sulfonamides, and trimethoprim-sulfamethoxazole", Rev Infect.Dis., vol. 4, no. 2, pp. 261-269.

Theus, S. A., Cave, M. D., & Eisenach, K. D. 2005, "Intracellular macrophage growth rates and cytokine profiles of *Mycobacterium tuberculosis* strains with different transmission dynamics", J.Infect.Dis., vol. 191, no. 3, pp. 453-460.

Thierry, D., Cave, M. D., Eisenach, K. D., Crawford, J. T., Bates, J. H., Gicquel, B., & Guesdon, J. L. 1990, "IS6110, an IS-like element of *Mycobacterium tuberculosis* complex", Nucleic Acids Research, vol. 18, no. 1, p. 188.

Thwaites, G., Caws, M., Chau, T. T. H., D'Sa, A., Lan, N. T. N., Huyen, M. N. T., Gagneux, S., Anh, P. T. H., Tho, D. Q., Torok, E., Nhu, N. T. Q., Duyen, N. T. H., Duy, P. M., Richenberg, J., Simmons, C., Hien, T. T., & Farrar, J. 2008, "Relationship between *Mycobacterium tuberculosis* Genotype and the Clinical Phenotype of Pulmonary and Meningeal Tuberculosis", Journal of Clinical Microbiology, vol. 46, no. 4, pp. 1363-1368.

Ting, L. M., Kim, A. C., Cattamanchi, A., & Ernst, J. D. 1999, "*Mycobacterium tuberculosis* inhibits IFN-gamma transcriptional responses without inhibiting activation of STAT1", J.Immunol., vol. 163, no. 7, pp. 3898-3906.

Tissot, F., Zanetti, G., Francioli, P., Zellweger, J. P., & Zysset, F. 2005, "Influence of bacille Calmette-Guerin vaccination on size of tuberculin skin test reaction: to what size?", Clin.Infect.Dis., vol. 40, no. 2, pp. 211-217.

Tsenova, L., Ellison, E., Harbacheuski, R., Moreira, A. L., Kurepina, N., Reed, M. B., Mathema, B., Barry, C. E., III, & Kaplan, G. 2005, "Virulence of selected *Mycobacterium tuberculosis* clinical isolates in the rabbit model of meningitis is dependent on phenolic glycolipid produced by the bacilli", J.Infect.Dis., vol. 192, no. 1, pp. 98-106.

Tso, H. W., Lau, Y. L., Tam, C. M., Wong, H. S., & Chiang, A. K. 2004, "Associations between IL12B polymorphisms and tuberculosis in the Hong Kong Chinese population", J.Infect.Dis., vol. 190, no. 5, pp. 913-919.

Tsolaki, A. G., Gagneux, S., Pym, A. S., Goguet de la Salmoniere, Y.-O., Kreiswirth, B. N., van Soolingen, D., & Small, P. M. 2005, "Genomic Deletions Classify the Beijing/W Strains as a Distinct Genetic Lineage of *Mycobacterium tuberculosis*", Journal of Clinical Microbiology, vol. 43, no. 7, pp. 3185-3191.

Turnbull, F. W., Wallace, A. T., Stewart, S., & Crofton, J. W. 1953, "Streptomycin resistance in patients with pulmonary tuberculosis previously treated with P.A.S. alone", Br.Med.J., vol. 1, no. 4822, pp. 1244-1246.

Ulrichs T, Munk ME, Mollenkopf H, Behr-Perst S, Colangeli R, Gennaro ML, Kaufmann SH. 1998. "Differential T cell responses to *Mycobacterium tuberculosis* ESAT6 in tuberculosis patients and healthy donors." Eur.J.Immunol. 28(12): 3949-58

Ulukanligil, M., Aslan, G., & Tasci, S. 2000, "A comparative study on the different staining methods and number of specimens for the detection of acid fast bacilli", Mem.Inst.Oswaldo Cruz, vol. 95, no. 6, pp. 855-858.

United Nations 2000. United Nations - Millenium Development Goals. http://www.un.org/millenniumgoals/ .

United Nations 2005. World population to reach 9.1 billion in 2050, UN projects. http://www.un.org/apps/news/story.asp?NewsID=13451&Cr=population&Cr1

Valone, S. E., Rich, E. A., Wallis, R. S., & Ellner, J. J. 1988, "Expression of tumor necrosis factor in vitro by human mononuclear phagocytes stimulated with whole Mycobacterium bovis BCG and mycobacterial antigens", Infect.Immun., vol. 56, no. 12, pp. 3313-3315.

Van Bambeke F., Balzi, E., & Tulkens, P. M. 2000, "Antibiotic efflux pumps", Biochem.Pharmacol., vol. 60, no. 4, pp. 457-470.

van den Berg, R. J., Schaap, I., Templeton, K. E., Klaassen, C. H., & Kuijper, E. J. 2007, "Typing and subtyping of Clostridium difficile isolates by using multiple-locus variable-number tandem-repeat analysis", Journal of Clinical Microbiology, vol. 45, no. 3, pp. 1024-1028.

van der Wel N, Hava D, Houben D, Fluitsma D, van Zon M, Pierson J, Brenner M, Peters PJ. 2007. "M. tuberculosis and M. leprae translocate from the phagolysosome to the cytosol in myeloid cells." Cell 129(7): 1287-98.

van der Zanden, A. G., Hoentjen, A. H., Heilmann, F. G., Weltevreden, E. F., Schouls, L. M., & van Embden, J. D. 1998, "Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* complex in paraffin wax embedded tissues and in stained microscopic preparations", Mol.Pathol., vol. 51, no. 4, pp. 209-214.

van Deutekom H., Gerritsen, J. J., van, S. D., van Ameijden, E. J., van Embden, J. D., & Coutinho, R. A. 1997, "A molecular epidemiological approach to studying the transmission of tuberculosis in Amsterdam", Clin.Infect.Dis., vol. 25, no. 5, pp. 1071-1077.

van Deutekom, H., Supply, P., de Haas, P. E. W., Willery, E., Hoijng, S. P., Locht, C., Coutinho, R. A., & van Soolingen, D. 2005, "Molecular Typing of *Mycobacterium tuberculosis* by Mycobacterial Interspersed Repetitive Unit-Variable-Number Tandem Repeat Analysis, a More Accurate Method for Identifying Epidemiological Links between Patients with Tuberculosis", Journal of Clinical Microbiology, vol. 43, no. 9, pp. 4473-4479.

van Doorn, H. R., de Haas, P. E., Kremer, K., Vandenbroucke-Grauls, C. M., Borgdorff, M. W., & van, S. D. 2006, "Public health impact of isoniazid-resistant *Mycobacterium tuberculosis* strains with a mutation at amino-acid position 315 of katG: a decade of experience in The Netherlands", Clin.Microbiol.Infect., vol. 12, no. 8, pp. 769-775.

van Embden, J. D., Cave, M. D., Crawford, J. T., Dale, J. W., Eisenach, K. D., Gicquel, B., Hermans, P., Martin, C., McAdam, R., & Shinnick, T. M. 1993, "Strain identification of *Mycobacterium tuberculosis* by DNA fingerprinting: recommendations for a standardized methodology", Journal of Clinical Microbiology, vol. 31, no. 2, pp. 406-409.

van Embden, J. D., van, G. T., Kremer, K., Jansen, R., van Der Zeijst, B. A., & Schouls, L. M. 2000, "Genetic variation and evolutionary origin of the direct repeat locus of

Mycobacterium tuberculosis complex bacteria", The Journal of Bacteriology, vol. 182, no. 9, pp. 2393-2401.

van Kooyk Y. & Geijtenbeek, T. B. 2003, "DC-SIGN: escape mechanism for pathogens", Nat Rev Immunol., vol. 3, no. 9, pp. 697-709.

van Soolingen, D., Hermans, P. W., de Haas, P. E., Soll, D. R., & van Embden, J. D. 1991, "Occurrence and stability of insertion sequences in *Mycobacterium tuberculosis* complex strains: evaluation of an insertion sequence-dependent DNA polymorphism as a tool in the epidemiology of tuberculosis", Journal of Clinical Microbiology, vol. 29, no. 11, pp. 2578-2586.

van Soolingen, D., de Haas, P. E., Hermans, P. W., Groenen, P. M., & van Embden, J. D. 1993, "Comparison of various repetitive DNA elements as genetic markers for strain differentiation and epidemiology of *Mycobacterium tuberculosis*", Journal of Clinical Microbiology, vol. 31, no. 8, pp. 1987-1995.

van Soolingen, D., Qian, L., de Haas, P. E., Douglas, J. T., Traore, H., Portaels, F., Qing, H. Z., Enkhsaikan, D., Nymadawa, P., & van Embden, J. D. 1995, "Predominance of a single genotype of *Mycobacterium tuberculosis* in countries of east Asia", Journal of Clinical Microbiology, vol. 33, no. 12, pp. 3234-3238.

van Soolingen D., Borgdorff, M. W., de Haas, P. E., Sebek, M. M., Veen, J., Dessens, M., Kremer, K., & van Embden, J. D. 1999, "Molecular epidemiology of tuberculosis in the Netherlands: a nationwide study from 1993 through 1997", J.Infect.Dis., vol. 180, no. 3, pp. 726-736.

van Soolingen D., de Haas, P. E., van Doorn, H. R., Kuijper, E., Rinder, H., & Borgdorff, M. W. 2000, "Mutations at amino acid position 315 of the katG gene are associated with high-level resistance to isoniazid, other drug resistance, and successful transmission of *Mycobacterium tuberculosis* in the Netherlands", J.Infect.Dis., vol. 182, no. 6, pp. 1788-1790.

Velji P, Nikolayevskyy V, Brown T, Drobniewski F. 2009. Discriminatory Ability of Hypervariable Variable Number Tandem Repeat Loci in Populationbased Analysis of Mycobacterium tuberculosis Strains, London, UK Emerging Infectious Diseases . 15(10); 1609-1616

Volkman HE, Clay H, Beery D, Chang JC, Sherman DR, Ramakrishnan L. 2004. "Tuberculous granuloma formation is enhanced by a mycobacterium virulence determinant." PLoS.Biol. 2(11): e367.

Voskuil MI, Schnappinger D, Visconti KC, Harrell MI, Dolganov GM, Sherman DR, Schoolnik GK. 2003. "Inhibition of respiration by nitric oxide induces a *Mycobacterium tuberculosis* dormancy program." J.Exp.Med. 198(5): 705-13.

Waksman, S. A., Reilly, H. C., & Schatz, A. 1945, "Strain Specificity and Production of Antibiotic Substances: V. Strain Resistance of Bacteria to Antibiotic Substances, Especially to Streptomycin", Proc.Natl.Acad.Sci.U.S.A, vol. 31, no. 6, pp. 157-164.

Wards, B. J., G. W. de Lisle, and D. M. Collins. 2000. "An esat6 knockout mutant of Mycobacterium bovis produced by homologous recombination will contribute to the development of a live tuberculosis vaccine." Tuber.Lung Dis. 80(4-5): 185-89.

Walker NF, Turner D, Kliner M, Bhagani S, Cropley I, Hopkins S, Lipman M 2008. "Hepatotoxicity due to anti-tuberculosis therapy: guidance and practice - Clinical tuberculosis." Thorax 63.Suppl 7: A144-A149.

Wallis, R. S., Vjecha, M., mir-Tahmasseb, M., Okwera, A., Byekwaso, F., Nyole, S., Kabengera, S., Mugerwa, R. D., & Ellner, J. J. 1993, "Influence of tuberculosis on human immunodeficiency virus (HIV-1): enhanced cytokine expression and elevated beta 2-microglobulin in HIV-1-associated tuberculosis", J.Infect.Dis., vol. 167, no. 1, pp. 43-48.

Wang, C. X. & Mi, Z. H. 2006, "Imipenem-resistant Pseudomonas aeruginosa producing IMP-1 metallo-beta-lactamases and lacking the outer-membrane protein OprD", J.Med.Microbiol., vol. 55, no. Pt 3, pp. 353-354.

Wang, L., Turner, M. O., Elwood, R. K., Schulzer, M., & FitzGerald, J. M. 2002, "A meta-analysis of the effect of Bacille Calmette Guerin vaccination on tuberculin skin test measurements", Thorax, vol. 57, no. 9, pp. 804-809.

Wang, P. D. and R. S. Lin. 2000. "Tuberculosis transmission in the family." J.Infect. 41(3): 249-51.

Warren, R. M., Van Der Spuy, G. D., Richardson, M., Beyers, N., Borgdorff, M. W., Behr, M. A., & van Helden, P. D. 2002, "Calculation of the stability of the IS6110 banding pattern in patients with persistent *Mycobacterium tuberculosis* disease", Journal of Clinical Microbiology, vol. 40, no. 5, pp. 1705-1708.

Wassie, L., Demissie, A., Aseffa, A., Abebe, M., Yamuah, L., Tilahun, H., Petros, B., Rook, G., Zumla, A., Andersen, P., & Doherty, T. M. 2008, "Ex vivo cytokine mRNA levels correlate with changing clinical status of ethiopian TB patients and their contacts over time", PLoS.One., vol. 3, no. 1, p. e1522.

Werner, G., Coque, T. M., Hammerum, A. M., Hope, R., Hryniewicz, W., Johnson, A., Klare, I., Kristinsson, K. G., Leclercq, R., Lester, C. H., Lillie, M., Novais, C., Olsson-Liljequist, B., Peixe, L. V., Sadowy, E., Simonsen, G. S., Top, J., Vuopio-Varkila, J., Willems, R. J., Witte, W., & Woodford, N. 2008, "Emergence and spread of vancomycin resistance among enterococci in Europe", Euro.Surveill, vol. 13, no. 47.

Whalen, C., Horsburgh, C. R., Hom, D., Lahart, C., Simberkoff, M., & Ellner, J. 1995, "Accelerated course of human immunodeficiency virus infection after tuberculosis", Am.J.Respir.Crit Care Med., vol. 151, no. 1, pp. 129-135.

White, V. L. C. & Moore-Gillon, J. 2000, "Resource implications of patients with multidrug resistant tuberculosis", Thorax, vol. 55, no. 11, pp. 962-963.

WHO 2006. The Stop TB Strategy. Building on and enhancing DOTS to meet the TB-related Millennium Development Goals.

WHO 2007. WHO tuberculosis fact sheet No. 104. World Health Organisation.

WHO 2008. Anti-tuberculosis Drug Resistance in the World. Report No. 4.

WHO 2009. Global Tuberculosis Control. Short Report.

WHO 2010 The Global Plan to Stop TB 2011-2015.

Wilkinson, K. A., Wilkinson, R. J., Pathan, A., Ewer, K., Prakash, M., Klenerman, P., Maskell, N., Davies, R., Pasvol, G., & Lalvani, A. 2005, "Ex vivo characterization of early secretory antigenic target 6-specific T cells at sites of active disease in pleural tuberculosis", Clin.Infect.Dis., vol. 40, no. 1, pp. 184-187.

Wilson M, DeRisi J, Kristensen HH, Imboden P, Rane S, Brown PO, Schoolnik GK. 1999. "Exploring drug-induced alterations in gene expression in *Mycobacterium tuberculosis* by microarray hybridization." Proceedings of the National Academy of Sciences of the United States of America 96(22): 12833-38.

Wilson, T. M., de Lisle, G. W., & Collins, D. M. 1995, "Effect of inhA and katG on isoniazid resistance and virulence of Mycobacterium bovis", Mol.Microbiol., vol. 15, no. 6, pp. 1009-1015.

Wirth, T., Hildebrand, F., lix-Beguec, C., Wolbeling, F., Kubica, T., Kremer, K., van, S. D., Rusch-Gerdes, S., Locht, C., Brisse, S., Meyer, A., Supply, P., & Niemann, S. 2008, "Origin, spread and demography of the *Mycobacterium tuberculosis* complex", PLoS.Pathog., vol. 4, no. 9, p. e1000160.

Wisnivesky, J. P., Kaplan, J., Henschke, C., McGinn, T. G., & Crystal, R. G. 2000, "Evaluation of clinical parameters to predict *Mycobacterium tuberculosis* in inpatients", Arch.Intern.Med., vol. 160, no. 16, pp. 2471-2476.

Wright A, Zignol M, Van Deun A, Falzon D, Ruesch-Gerdes S, Feldman K, Hoffner S, Drobniewski F. Barrera L, van Soolingen D, Boulabhal F, Paramasivan CN, Man Kam K, Mitarai S, Nunn P, Raviglione M. 2009 Epidemiology of antituberculosis drug resistance 2002–07:an updated analysis of the Global Project on Anti-Tuberculosis Drug Resistance Surveillance. Lancet. 373: 1861–73

Yang, Z., Chaves, F., Barnes, P. F., Burman, W. J., Koehler, J., Eisenach, K. D., Bates, J. H., & Cave, M. D. 1996, "Evaluation of method for secondary DNA typing of *Mycobacterium tuberculosis* with pTBN12 in epidemiologic study of tuberculosis", Journal of Clinical Microbiology, vol. 34, no. 12, pp. 3044-3048.

Yeager, H., Jr., Lacy, J., Smith, L. R., & LeMaistre, C. A. 1967, "Quantitative studies of mycobacterial populations in sputum and saliva", Am.Rev Respir.Dis., vol. 95, no. 6, pp. 998-1004.

Yee D, Valiquette C, Pelletier M, Parisien I, Rocher I, Menzies D. 2003. "Incidence of Serious Side Effects from First-Line Antituberculosis Drugs among Patients Treated for Active Tuberculosis." American Journal of Respiratory and Critical Care Medicine 167(11): 1472-77.

Yeh, R. W., Ponce de, L. A., Agasino, C. B., Hahn, J. A., Daley, C. L., Hopewell, P. C., & Small, P. M. 1998, "Stability of *Mycobacterium tuberculosis* DNA genotypes", J.Infect.Dis., vol. 177, no. 4, pp. 1107-1111.

Yeremeev, V. V., Lyadova, I. V., Nikonenko, B. V., Apt, A. S., bou-Zeid, C., Inwald, J., & Young, D. B. 2000, "The 19-kD antigen and protective immunity in a murine model of tuberculosis", Clin.Exp.Immunol., vol. 120, no. 2, pp. 274-279.

Youmans, G. P and A. S Youmans. 1949. "A method for the determination of the rate of growth of tubercle bacilli by the use of small inocula." The Journal of Bacteriology 58(2): 247-55.

Zager, E. M. & McNerney, R. 2008, "Multidrug-resistant tuberculosis", BMC Infect.Dis., vol. 8, p. 10.

Zainuddin, Z. F. & Dale, J. W. 1989, "Polymorphic repetitive DNA sequences in *Mycobacterium tuberculosis* detected with a gene probe from a Mycobacterium fortuitum plasmid", J.Gen.Microbiol., vol. 135, no. 9, pp. 2347-2355.

Zhang, M., Gong, J., Yang, Z., Samten, B., Cave, M. D., & Barnes, P. F. 1999, "Enhanced capacity of a widespread strain of *Mycobacterium tuberculosis* to grow in human macrophages", J.Infect.Dis., vol. 179, no. 5, pp. 1213-1217.

Zheng, H., Lu, L., Wang, B., Pu, S., Zhang, X., Zhu, G., Shi, W., Zhang, L., Wang, H., Wang, S., Zhao, G., & Zhang, Y. 2008, "Genetic basis of virulence attenuation revealed by comparative genomic analysis of *Mycobacterium tuberculosis* strain H37Ra versus H37Rv", PLoS.One., vol. 3, no. 6, p. e2375.