

**HUMAN GENETIC SUSCEPTIBILITY TO TUBERCULOSIS:
THE INVESTIGATION OF CANDIDATE GENES
INFLUENCING INTERFERON GAMMA LEVELS AND OTHER
CANDIDATE GENES AFFECTING IMMUNOLOGICAL
PATHWAYS.**

Marlo Möller



Dissertation presented for the degree of Doctor of Philosophy in Medical Biochemistry
at Stellenbosch University

Promoter: Prof Eileen Hoal

Co-promoter: Prof Paul van Helden

December 2007

DECLARATION

I, the undersigned, hereby declare that the work contained in this dissertation is my own original work and that I have not previously in its entirety or in part submitted it at any university for a degree.

Signature: _____

M. Möller

Date: _____

SUMMARY

The infectious disease tuberculosis (TB) is one of the leading causes of death worldwide. The idea that infectious diseases are the most important driving force in natural selection and that they sustain frequent polymorphisms in the human genome was formally suggested by Haldane in 1949. This hypothesis implicated the human genetic component in the response to infectious disease. Today the involvement of host genetics in TB has been proven unequivocally and, together with environmental factors (e.g. nutrition and crowding) and the causative bacterium, *Mycobacterium tuberculosis* (*M.tuberculosis*), may influence the outcome of disease. As is evident, TB is a complex disease and the implication for studying genetic susceptibility is that a number of genes will be involved.

Interferon gamma (IFN- γ) is the major macrophage-activating cytokine during infection with *M.tuberculosis* and its role has been well established in animal models and in humans. This cytokine is produced by activated T helper 1 (Th1) cells. These Th1 responses can best deal with intracellular pathogens such as *M.tuberculosis*. We selected twelve candidate genes based on the hypothesis that genes which regulate the production of IFN- γ may influence TB susceptibility. We also selected polymorphisms from 27 other candidate genes, which may affect immunological pathways involved in TB, to investigate as susceptibility factors based on the following hypotheses: 1) granulomatous diseases can share susceptibility genes; 2) gene expression studies done by DNA-array analysis experiments may reveal TB susceptibility genes; 3) genome-wide linkage studies in TB can determine susceptibility loci and genes in this region are possibly susceptibility factors; and 4) functional susceptibility polymorphisms in genes involved in immune-mediated diseases other than TB may contribute to susceptibility to TB.

This research tested the association of 136 genetic polymorphisms in 39 potentially important genes with TB in the South African Coloured population. Well-designed case-control association studies were used and we attempted to replicate these findings in an independent sample set using family-based case-control designs (transmission disequilibrium tests (TDTs)). In addition, haplotypes and linkage disequilibrium (LD) in the candidate genes were also investigated.

During the case-control analyses we found significant associations for 6 single nucleotide polymorphisms (SNPs) in the following genes: SH2 domain protein 1A, toll-like receptor 2, class II major histocompatibility complex transactivator, interleukin 1 receptor antagonist, runt-related transcription factor 1 and tumour necrosis factor superfamily, member 1B. Discrepant results were obtained during the TDT analyses. The number of families available was small and for this reason we cannot conclude that the case-control results were spurious. We also tested the association of haplotypes with TB. Haplotypes in the interleukin 12, beta (*IL12B*) and toll-like receptor 4 genes were nominally associated with TB in both the case-control and TDT analyses. We observed strong LD for the genes in the South African Coloured population. In total 17 novel SNPs were identified and one novel allele was found for a microsatellite in *IL12B*.

This research contributes to the increasing amount of information available on genes involved in TB susceptibility, which in the future may help to predict high risk individuals.

OPSOMMING

Die infektiewe siekte tuberkulose (TB) is wêreldwyd een van die hooforsake van menslike sterftes. Die idee dat hierdie siektes die belangrikste dryfveer van natuurlike seleksie is en dat dit die rede is waarom polimorfismes in die menslike genoom voortbestaan, was in 1949 formeel voorgestel deur Haldane. Volgens hierdie hipotese speel die mens se genetiese samestelling 'n rol in die reaksie op infektiewe siektes. Die betrokkenheid van die gasheer se genoom in TB is onaanvegbaar en saam met omgewingsfaktore (bv. voeding en oorbevolking) en die veroorsakende bakterium *Mycobacterium tuberculosis* (*M.tuberculosis*), kan dit die uitkoms van die siekte beïnvloed. Die bydrae van al hierdie faktore maak TB 'n komplekse siekte wat impliseer dat verskeie gene betrokke sal wees in die bestudering van genetiese vatbaarheid vir TB.

Interferon gamma (IFN- γ) is die hoof makrofaag-aktiverende sitokien tydens infeksie met *M.tuberculosis* en die rol daarvan in dieremodelle en in die mens is bekend. Die sitokien word geproduseer deur geaktiveerde T helper 1 (Th1) selle. Hierdie Th1 reaksies is die beste verdedigingsmeganisme teen intrasellulêre patogene soos *M.tuberculosis*. Ons het twaalf kandidaatgene gekies gebaseer op die hipotese dat gene wat IFN- γ produksie reguleer TB vatbaarheid mag beïnvloed. Ons het ook polimorfismes van 27 ander kandidaatgene, wat immunologiese paaie betrokke in TB kan affekteer, gekies gebaseer op die volgende hipoteses: 1) granulomateuse siektes kan vatbaarheidsgene deel; 2) geenuitdrukkingstudies kan TB vatbaarheidsgene voorstel; 3) genoomwye koppelingsanalise in TB kan vatbaarheidslokusse voorstel en gene in hierdie areas is moontlike vatbaarheidsfaktore; en 4) funksionele polimorfismes in gene betrokke in ander immuun-mediërende siektes kan bydrae tot vatbaarheid vir TB.

Hierdie studie het die assosiasie van 136 genetiese polimorfismes in 39 kandidaatgene met TB in die Suid-Afrikaanse Kleurlingbevolking getoets. Goedbeplande pasiënt-kontrole assosiasiestudies is gedoen en familie-gebaseerde assosiasiestudies (transmissie disekwilibriumtoetse (TDT)) is gebruik om die resultate van die pasiënt-kontrole studie te dupliseer in 'n onafhanklike groep. Ons het ook haplotipes en koppelings disekwilibrium (KD) in die kandidaatgene ondersoek.

Tydens die pasiënt-kontrole assosiasiestudies het ons beduidende positiewe assosiasies vir 6 enkel nukleotied polimorfismes (ENPs) in die volgende gene gevind: *SH2 domain protein 1A*, *toll-like receptor 2*, *class II major histocompatibility complex transactivator*, *interleukin 1 receptor antagonist*, *runt-related transcription factor 1* en *tumour necrosis factor superfamily, member 1B*. Verskillende resultate is verkry tydens die TDT analises. Die aantal families wat beskikbaar was vir studie, was relatief min en daarom kan ons nie aflei dat die resultate van die pasiënt-kontrole assosiasiestudies vals was nie. Ons het ook die assosiasie van haplotipes met TB getoets. Haplotipes in die *interleukin 12, beta (IL12B)* en *toll-like receptor 4* gene was geassosieer met TB in beide die pasiënt-kontrole en TDT analises. Ons het sterk KD vir die gene in die Suid-Afrikaanse Kleurlingbevolking waargeneem. In totaal is 17 nuwe ENPs geïdentifiseer en een nuwe alleel is gevind vir 'n mikrosatelliet in *IL12B*.

Hierdie navorsing dra by tot die toenemende hoeveelheid inligting beskikbaar oor gene betrokke in TB vatbaarheid, wat in die toekoms mag help om individue met 'n hoë risiko om die siekte te ontwikkel te voorspel.

*This thesis is dedicated to my parents, Erik and Trix Möller, and to my grandmothers,
Susan Möller and Max Kleynhans.*

ACKNOWLEDGEMENTS

Success is not final, failure is not fatal: it is the courage to continue that counts. Winston Churchill

I would like to thank the following people and institutions for their contribution to this work:

My promotor prof Eileen Hoal-van Helden and co-promoter prof Paul van Helden. Thank you for believing in me, for all your support and encouragement, and for the excellent help you gave me during the preparation of my thesis. I have learnt invaluable scientific and life lessons from both of you.

Our collaborators at the Institute for Clinical Molecular Biology in Kiel, Germany. Prof Stefan Schreiber for giving me the opportunity to visit the Institute, dr Ruta Kwiatkowski and dr Almut Nebel for assisting me with experimental design and manuscript writing, dr Andre Franke for designing the SNPlex pool for immune-mediated diseases, dr Christiane Wolf-Schwerin for organising my accommodation and helping with other administrative tasks, Lena Bosson and Tanya Wesse for genotyping, Rainer Vogler for database support, the DNA intake platform for plate-making and especially Meike Davids for teaching me the WGA protocol, Anita Dietsch and Melanie Friskovec for sequencing and my friends in the main office Friederike Flachsbart, Annegret Fischer, Abdou El Sharawy and Weiyue Zheng. I will remember the eight months that I spent in Kiel with fondness.

Prof Rob Warren, Don Hayward, Nelita du Plessis and Ruben van der Merwe for assistance with the *SH2DIA* cloning. Thank you for the great tips and for the use of your protocols and reagents.

All my friends in the Division of Molecular Biology and Human Genetics, especially Elizabeth Hanekom Keet, Kim Stanley, Erika de Wit, André Loxton, Laurianne van der Merwe and my “PhD partner” Chantal Babb. Our tea-time and lunch trips to the TSS saved me! Thank you for the useful discussions about research which we had on our way there. I would also like to thank Elizabeth and Chantal for doing sequencing reactions and capillary electrophoresis.

The Harry Crossley foundation, the South African Medical Research Council and Stellenbosch University for financial support during this project.

The individuals who consented to take part in this research.

My sister Chrizanne and brother Heinrich. Dankie dat julle geluister het as ek gepraat het oor my navorsing, al moes dit soms na Grieks geklink het! Ek waardeer dit baie.

My parents Erik and Trix Möller. Pappa en Mamma, baie dankie vir al die opofferinge wat julle vir my gemaak het, ek weet dit was nie altyd maklik nie en ek sal julle nooit daarvoor kan vergoed nie. Ek is werklik bevoorreg om ouers soos julle te hê. Ek is baie lief vir julle!

TABLE OF CONTENTS

List of abbreviations	i
List of figures	vi
List of tables	ix
List of publications	xiii
CHAPTER 1 INTRODUCTION	
1.1. Tuberculosis	1
1.1.1. The history of TB	1
1.1.2. The pathogen	3
1.1.3. Pathophysiology	4
1.1.4. Current treatment	4
1.1.5. Drug-resistant TB	5
1.1.6. HIV and TB	6
1.2. Host genetic susceptibility to TB	7
1.2.1. Evidence for a genetic influence in susceptibility to <i>M.tuberculosis</i>	7
1.2.2. Approaches and study designs used in TB host genetics	9
1.2.3. Previous genetic susceptibility studies done in TB	17
1.2.4. Genes investigated in this study	19
CHAPTER 2 STUDY OVERVIEW	
2.1. Candidate gene selection	31
2.2. Main hypotheses	31
2.3. Aims of this study	31
CHAPTER 3 MATERIALS AND METHODS	
3.1. Reagents and equipment	32
3.2. Laboratory Information Management System	32
3.3. Study participants	32
3.3.1. Study population and stratification	32
3.3.2. TB case-control sample set	33
3.3.3. Family-based study sample set	34
3.4. DNA samples	34
3.4.1. DNA extraction from blood and quality testing	34
3.4.2. Plate design	35
3.4.3. Whole genome amplification (WGA)	36
3.5. Mutation detection and polymorphism selection	37
3.5.1. Butyrophilin-like 2 gene	38
3.5.2. Interleukin 12B gene	40
3.5.3. Interleukin 12 receptor, beta 2 gene	40
3.5.4. Signaling lymphocytic activation molecule-associated protein	41
3.5.5. Testis expressed sequence 264	41
3.5.6. Toll-like receptor 2	42
3.5.7. Gel electrophoresis, purification of PCR products and sequencing	43
3.5.8. Selection of remaining polymorphisms	44
3.6. Amplification refraction mutation system-polymerase chain reaction	44
3.6.1. ARMS protocol	46
3.6.2. Tetraprimer ARMS-PCR protocol	46
3.7. Capillary electrophoresis of fluorescently labelled PCR products	47
3.7.1. Capillary electrophoresis protocol	47
3.8. TaqMan [®] genotyping	49

3.8.1. TaqMan [®] protocol.....	51
3.9. SNPLex [™] genotyping system	51
3.9.1. SNPLex pool design.....	54
3.9.2. SNPLex protocol.....	55
3.10. Cloning of the <i>SH2D1A</i> gene promoter.....	57
3.10.1. In silico promoter analysis.....	57
3.10.2. Cloning into pGEM [®] -T Easy vector system.....	57
3.10.3. Subcloning into pGL4 luciferase reporter vector.....	60
3.11. Statistical analysis.....	62
3.11.1. Quality checking of the genotype data	62
3.11.2. Genotype distributions of SNPs.....	62
3.11.3. Transmission disequilibrium test.....	63
3.11.4. Haplotypes and linkage disequilibrium	63
3.11.5. Microsatellite analysis	64
3.11.6. Power calculations.	64
CHAPTER 4 CANDIDATE GENES WHICH INFLUENCE INTERFERON GAMMA LEVELS	
4.1. Introduction.....	65
4.2. C-C chemokine ligand-2.....	68
4.2.1. Results.....	69
4.2.2. Discussion.....	73
4.3. Interleukin-4.....	74
4.3.1. Results.....	75
4.3.2. Discussion.....	78
4.4. Interleukin-10.....	79
4.4.1. Results.....	80
4.4.2. Discussion.....	85
4.5. Members of the interleukin-12 family and their receptors	86
4.5.1. Results.....	88
4.5.2. Discussion.....	99
4.6. Interleukin-18.....	102
4.6.1. Results.....	103
4.6.2. Discussion.....	106
4.7. SH2 domain protein 1A	107
4.7.1. Results.....	108
4.7.2. Discussion.....	110
4.8. Toll-like receptors.....	112
4.8.1. Results.....	114
4.8.2. Discussion.....	123
4.9. The wingless-type MMTV integration site family, member 5A and Frizzled homolog 5 (<i>Drosophila</i>).....	125
4.9.1. Results.....	127
4.9.2. Discussion.....	134
CHAPTER 5 CANDIDATE GENES FROM OTHER GRANULOMATOUS DISEASES	
5.1. Introduction.....	136
5.2. Butyrophilin-like 2 gene	138
5.2.1. Results.....	138
5.2.2. Discussion.....	146
5.3. Caspase recruitment domain-containing protein 15 gene.....	147

5.3.1. Results.....	147
5.3.2. Discussion.....	148
CHAPTER 6 CANDIDATE GENES WHICH ARE DIFFERENTIALLY EXPRESSED BETWEEN TUBERCULOSIS CASES AND CONTROLS	
6.1. Introduction.....	151
6.1.1 Testis expressed 264	152
6.1.2. Suppressor of cytokine signalling 3	153
6.2. Results.....	154
6.2.1. <i>TEX264</i>	154
6.2.2. <i>SOCS3</i>	158
6.3. Discussion.....	160
CHAPTER 7 A CANDIDATE GENE FROM A GENOME-WIDE LINKAGE SCREEN	
7.1. Introduction.....	161
7.2. Cathepsin Z.....	162
7.2.1. Results.....	163
7.2.2. Discussion.....	165
CHAPTER 8 FUNCTIONAL POLYMORPHISMS THAT INFLUENCE SUSCEPTIBILITY TO IMMUNE-MEDIATED DISEASES OTHER THAN TUBERCULOSIS	
8.1. Introduction.....	167
8.2. Candidate genes from immune-mediated diseases other than TB	168
8.2.1. Results.....	170
8.2.2. Discussion.....	172
CHAPTER 9 CONCLUDING REMARKS	175
REFERENCES	184
ADDENDUM	
A.1. Buffers and solutions	225
A.2. Reagents	226
A.3. Equipment	229
A.4. Software	231
A.5. Online resources.....	231
A.6. Polymorphisms detected during <i>BTNL2</i> sequencing	232
A.7. Publications.....	236

LIST OF ABBREVIATIONS

A	adenine
<i>ABCB1</i>	ATP-binding cassette, subfamily B (MDR/TAP), member 1 gene
AIDS	acquired immune deficiency syndrome
APC	antigen-presenting cell
Arg	arginine
ARMS	amplification refraction mutation system
<i>ASNA1</i>	ArsA arsenite transporter, ATP-binding, homolog 1 (bacterial) gene
ASO	allele-specific oligonucleotide
Asp	aspartic acid
<i>ATG16L1</i>	ATG16 autophagy related 16-like 1 (<i>S. cerevisiae</i>) gene
ATP	adenosine triphosphate
<i>ATP5G1</i>	ATP synthase, H ⁺ transporting, mitochondrial F ₀ complex, subunit c (subunit 9) gene
BCG	bacillus Calmette-Guérin
bp	base pair
BSA	bovine serum albumin
<i>BTNL2</i>	butyrophilin-like 2 gene
BTNL2	butyrophilin-like 2 protein
C	cytosine
C/EBP- β	CCAAT enhancer binding protein, beta
<i>C14orf2</i>	Chromosome 14 open reading frame 2 gene
<i>CARD15</i>	caspase recruitment domain-containing protein 15 gene
<i>CCL2</i>	C-C chemokine ligand-2 gene
CD	Crohn's disease
CEPH	Centre d'Etude du Polymorphisme Humain
CI	confidence interval
<i>CIITA</i>	Class II, major histocompatibility complex, transactivator gene
CIS	cytokine-inducible SH2-containing protein
<i>CTLA4</i>	Cytotoxic T-lymphocyte-associated protein 4 gene
CTS _Z	cathepsin Z protein
<i>CTS_Z</i>	cathepsin Z gene
D'	Lewontin's standardised disequilibrium coefficient D'
dATP	deoxyadenosine triphosphate
dH ₂ O	distilled water
DNA	deoxyribonucleic acid
dNTPs	deoxynucleoside triphosphate
ECM	extracellular matrix remodelling
<i>Exo I</i>	Exonuclease I
<i>FCRL3</i>	Fc receptor-like 3 gene
FZD5	frizzled homolog 5 protein

<i>FZD6</i> frizzled homolog 5 gene
g gram
G guanine
Gln glutamine
Gly glycine
GWA genome-wide association
H ₂ O water
HIV human immunodeficiency virus
HLA human leukocyte antigen
HWE Hardy-Weinberg equilibrium
IBD inflammatory bowel disease
<i>IFNGR1</i> IFN- γ receptor-1 gene
<i>IFNGR2</i> IFN- γ receptor-2 gene
IFN- γ interferon gamma
IKMB Institute for Clinical Molecular Biology
<i>IL10</i> interleukin-10 gene
IL-10 interleukin-10
IL-12 interleukin-12
IL-12(p40) ₂ interleukin-12 p 40 homodimer
<i>IL12B</i> interleukin-12, beta gene
IL-12p40 interleukin-12 p40 subunit
IL-12R interleukin-12 receptor complex
<i>IL12RB1</i> interleukin-12 receptor, beta 1 gene
<i>IL12RB2</i> interleukin-12 receptor, beta 2 gene
IL-12R β 1 interleukin-12 receptor, beta 1
IL-12R β 2 interleukin-12 receptor, beta 2
<i>IL18</i> interleukin-18 gene
IL-18 interleukin-18
<i>IL1RN</i> Interleukin 1 receptor antagonist gene
IL-23 interleukin-23
<i>IL23R</i> Interleukin23 receptor gene
IL-4 interleukin-4
<i>IL4</i> interleukin-4 gene
<i>IL6ST</i> Interleukin 6 signal transducer (gp130, oncostatin M reporter) gene
Ile isoleucine
indel insertion/deletion polymorphism
<i>INSIG2</i> Insulin induced gene 2
IPTG isopropyl β -D-1-thiogalactopyranoside
IS6110 insertion sequence 6110
JAK/STAT Janus kinase/signal transducer and activators of transcription
<i>KIAA2013</i> Hypothetical protein KIAA2013 gene

LB	Luria Broth
LD	linkage disequilibrium
LIMS	Laboratory Information Management System
LPS	lipopolysaccharides
LRR	leucine-rich repeats
LSO	locus-specific oligonucleotide
<i>LY6G6D</i>	Lymphocyte antigen 6 complex, locus G6D gene
<i>M.tuberculosis</i>	<i>Mycobacterium tuberculosis</i>
MCP-1	monocyte chemoattractant protein-1
MDA	multiple displacement amplification
MDR	multidrug-resistant
µg	microgram
mg	milligram
MgCl ₂	magnesium chloride
µl	microlitre
ml	millilitre
µM	micro molar
mM	millimolar
mRNA	messenger RNA
<i>MS4A2</i>	Membrane-spanning 4 domains, subfamily F, member 2 gene
MSMD	Mendelian susceptibility to mycobacterial disease
MyD88	myeloid differentiation factor 88
NaClO	Sodium hypochlorite
NaOH	sodium hydroxide
NCBI	National Centre for Biotechnology Information
<i>NELL1</i>	Nel-like 1 gene
<i>NEMO</i>	NFκB essential modulator gene
NFκB	nuclear factor κB
ng	nanogram
NK	natural killer
NOD2	nucleotide-binding oligomerization domain 2 protein
<i>NOLA3</i>	Nucleolar protein family A, member 3 gene
<i>NRAMP1</i>	natural resistance associated macrophage protein 1 gene
ns	not significant
OLA	oligonucleotide ligation assay
OR	odds ratio
p	probability value
<i>PADI4</i>	Peptidyl arginine deiminase, type IV gene
PAMP	pathogen-associated molecular patterns
PCR	polymerase chain reaction
<i>PPARG</i>	Peroxisome proliferator-activated receptor gamma gene
PPD	purified protein derivate

pTB	pulmonary tuberculosis
<i>PTGER4</i>	Prostaglandin E receptor 4 (subtype EP4) gene
<i>PTPN22</i>	Protein tyrosine phosphatase, non-receptor type 22 gene
PXR	pregnane X receptor
qRT-PCR	quantitative reverse-transcriptase PCR
r^2	square of the correlation coefficient between two loci
RE	response elements
<i>RIN3</i>	Ras and Rab interactor 3 gene
RNA	ribonucleic acid
rpm	revolutions per minute
<i>RUNX1</i>	Runt-related transcription factor 1 gene
SAP	Shrimp Alkaline Phosphatase
SAP	signaling lymphocytic activation molecule-associated protein
SB	sodium boride
<i>SH2D1A</i>	SH2 domain protein 1A, Duncan's disease (lymphoproliferative syndrome)
<i>SLC22A4</i>	Solute carrier family 22, (organic cation transporter), member 4 gene
<i>SLC22A5</i>	Solute carrier family 22, (organic cation transporter), member 5 gene
SNP	single nucleotide polymorphism
SOB		Super Optimal Broth
SOC	Super Optimal Broth with “B” changed to “C” for catabolite repression, reflective of the added glucose.
SOCS	suppressor of cytokine signalling proteins
SOCS-1	Suppressor of cytokine signalling 1 protein
<i>SOCS3</i>	Suppressor of cytokine signalling 3 gene
SOCS-3	Suppressor of cytokine signalling 3 protein
<i>SP110</i>	SP110 nuclear body protein gene
<i>STAT1</i>	signal transducer and activator of transcription-1 gene
STAT4	signal transducer and activator of transcription-4 protein
STR	short tandem repeat
T	thiamine
TB	tuberculosis
TBE	tris-boric acid-ethylene diamine tetra-acetic acid buffer
TDT	transmission disequilibrium test
TE	tris-ethylene diamine tetracetic acid buffer
<i>TEX264</i>	Testis expressed 264 gene
Th	T helper
<i>TH1</i>	TH1-like (<i>Drosophila</i>) gene
Thr	threonine
TIR	Toll/interleukin-1 receptor
TLR	toll-like receptor

<i>TLR2</i> toll-like receptor 2 gene
<i>TLR4</i> toll-like receptor 4 gene
T_m melting temperature
TNF tumour necrosis factor
<i>TNF</i> Tumour necrosis factor (TNF superfamily, member 2) gene
TNFR2 TNF receptor 2
<i>TNFRSF1A</i> Tumour necrosis factor superfamily, member 1A gene
<i>TNFRSF1B</i> Tumour necrosis factor superfamily, member 1B gene
<i>TNFSF15</i> Tumour necrosis factor (ligand) superfamily, member 15 gene
Trp tryptophan
<i>TUBB1</i> tubulin, beta 1 gene
<i>UBE3A</i> human papilloma virus E6-associated protein, Angelman syndrome gene
UTR untranslated region
UV ultra violet
V volt
<i>VDR</i> vitamin D receptor gene
WGA whole genome amplification
WHO World Health Organization
WNT/FZD Wntless /Frizzled
WNT5A wntless-type MMTV integration site family, member 5A protein
<i>WNT5A</i> wntless-type MMTV integration site family, member 5A gene
X times
XDR extensively drug-resistant
X-gal 5-bromo-4-chloro-3-indolyl- beta-D-galactopyranoside
XLP X-linked lymphoproliferative disease
% percent
/ per
°C degrees celcius
5' 5 prime end
3' 3 prime end

LIST OF FIGURES

Figure 1. Estimated incidence of new tuberculosis cases in 2005.....	1
Figure 2. Scanning electron micrograph of <i>M.tuberculosis</i>	4
Figure 3. <i>M.tuberculosis</i> enters the host through the respiratory system.	5
Figure 4. Estimated HIV prevalence in new TB cases, 2005.	6
Figure 5. Methods and strategies for identifying susceptibility genes in complex diseases.....	9
Figure 6. Illustration of LD blocks, recombination hotspots, haplotypes and haplotype tagging SNPs.....	16
Figure 7. Agarose gels used for quality testing of the DNA samples.....	35
Figure 8. Schematic representation of plate merging.	35
Figure 9. The multiple displacement amplification method used for whole genome amplification.....	36
Figure 10. The concept of the ARMS.....	44
Figure 11. The concept behind the tetraprimer ARMS-PCR.....	45
Figure 12. Graphical representation of capillary electrophoresis of fluorescently labelled primers.....	48
Figure 13. The TaqMan [®] platform.	49
Figure 14. Graphical representation of the TaqMan [®] SNP Genotyping method.....	50
Figure 15. Cluster plot used during genotype calling.....	50
Figure 16. Graphical representation of the SNPLex [™] genotyping system.	52
Figure 17. Graphical representation of a ZipChute probe.	53
Figure 18. Cluster plots used during the analysis of SNPLex [™] results.	54
Figure 19. The SNPLex [™] automated high-throughput pipeline for designing a multiplex pool.....	54
Figure 20. Simplified overview of the interleukin-12/interleukin-23/IFN- γ axis.....	66
Figure 21. Plot of LD between all markers of <i>CCL2</i> analysed in control individuals of South African Coloured descent, generated by Haploview v3.3.	70
Figure 22. Comparison of LD between the South African Coloured control population (n = 482) and the Framingham Offspring Study (n = 1797).	72
Figure 23. Plot of LD between <i>IL4</i> markers analysed in control individuals of South African Coloured descent generated by Haploview v3.3.....	76
Figure 24. Plot of LD between <i>IL10</i> markers analysed in control individuals of South African Coloured descent generated by Haploview v3.3.....	81
Figure 25. Selected members of the IL-12 family and their receptors.	86
Figure 26A. Sequencing from <i>IL12RB2</i> revealed the presence of the rs3762317 SNP, indicated with an arrow on the chromatogram.....	90
Figure 26B. A 2% agarose gel with the amplification products generated by the ARMS-PCR.	90
Figure 27. D5S2941 trinucleotide repeat analysis with Genemapper version 3.7.....	91

Figure 28. Sequencing of D5S2941 to confirm the presence of the (ATT) ₇ and (ATT) ₁₀ alleles.....	92
Figure 29. Distribution graph of (ATT) _n repeats of the D5S2941 microsatellite.	93
Figure 30. Plot of LD between <i>IL12B</i> markers analysed in control individuals of South African Coloured descent generated by Haploview v3.3.....	94
Figure 31. Plot of LD between <i>IL12RB1</i> markers analysed in control individuals of South African Coloured descent generated by Haploview v3.3.	95
Figure 32. Plot of LD between <i>IL12RB2</i> markers analysed in control individuals of South African Coloured descent generated by Haploview v3.3.	97
Figure 33. Plot of LD between <i>IL18</i> markers analysed in control individuals of South African Coloured descent generated by Haploview v3.3.....	104
Figure 34A. Sequencing from <i>SH2D1A</i> revealed the presence of the rs990545 polymorphism, indicated with an arrow on the chromatogram	108
Figure 34B. A 2% agarose gel with the amplification products generated by the tetraprimer ARMS-PCR.....	108
Figure 35. Sequencing of the pGEM-T and pGL4 constructs were aligned with the expected <i>SH2D1A</i> insert.....	110
Figure 36. Illustration of the role of TLRs in initiating the immune system. The extracellular leucine rich repeat (LRR) and intracellular Toll/interleukin-1 receptor (TIR) domains are indicated.....	112
Figure 37. (GT) _n dinucleotide repeat analysis with Genemapper version 3.7.	117
Figure 38. Distribution graph of (GT) _n repeats of the TLR2_GT microsatellite.....	118
Figure 39. Plot of LD between <i>TLR2</i> markers analysed in control individuals of South African Coloured descent generated by Haploview v3.3.....	120
Figure 40. Plot of LD between <i>TLR4</i> markers analysed in control individuals of South African Coloured descent generated by Haploview v3.3.....	121
Figure 41. Possible model of WNT5A and FZD5 as modulators of the relationship between antigen-presenting cells and T lymphocytes.....	125
Figure 42. Plot of LD between <i>WNT5A</i> markers analysed in control individuals of South African Coloured descent generated by Haploview v3.3.....	129
Figure 43. Plot of LD between <i>FZD5</i> markers analysed in control individuals of South African Coloured descent generated by Haploview v3.3.....	131
Figure 44. Amplification of the <i>BTNL2</i> products for sequencing.	139
Figure 45. Alignment of sequence data with the expected product using the Sequencher program.	139
Figure 46. SGCcaller was used to call genotypes in the sequenced samples.....	140
Figure 47. Plot of LD between all markers analysed in control individuals of South African Coloured descent generated by Haploview v3.3.....	143
Figure 48. Comparison of LD between the South African Coloured and German populations for 6 common SNPs.....	144
Figure 49. Expression profile of TEX264.	152
Figure 50. Expression profile of SOCS3.	154
Figure 51. Amplification of <i>TEX264</i> products for sequencing.....	155

Figure 52. Sequencing of the coding regions of <i>TEX264</i> and the detection of novel polymorphisms, highlighted in grey.....	155
Figure 53. Plot of linkage disequilibrium (LD) between <i>TEX264</i> markers analysed in control individuals of South African Coloured descent.....	157
Figure 54. Plot of LD between <i>SOCS3</i> markers analysed in control individuals of South African Coloured descent generated by Haploview v3.3.....	159
Figure 55. Plot of LD between <i>CTSZ</i> markers analysed in control individuals of South African Coloured descent generated by Haploview v3.3.....	164
Figure 56. Bladder biopsy showing chronic granulomatous inflammation.....	167

LIST OF TABLES

Table 1. Selected twin studies of TB.	8
Table 2. Genetic nomenclature.	10
Table 3. Types of population association studies.	13
Table 4. Publications reporting significant linkage or association with TB.	20
Table 5. Publications reporting no significant linkage or association with TB.	26
Table 6. Characteristics of the family-based and case-control samples.	33
Table 7. The names, sequences and product sizes of the primers for <i>BTNL2</i>	38
Table 8. The names, sequences and product sizes of the primers for <i>IL12B</i>	40
Table 9. The names, sequences and product sizes of the primers for <i>IL12RB2</i>	40
Table 10. The names, sequences and product sizes of the primers for <i>SH2D1A</i>	41
Table 11. The names, sequences and product sizes of the primers for <i>TEX264</i>	42
Table 12. The names, sequences and product sizes of the primers for <i>TLR2</i>	43
Table 13. Primers used for genotyping the rs3762317 SNP in <i>IL12RB2</i>	46
Table 14. Primers used to genotype the rs990545 SNP in <i>SH2D1A</i>	47
Table 15. Primers used during the multiplex PCR prior to capillary electrophoresis.	48
Table 16. TaqMan [®] Genotyping Assays used during this project.	51
Table 17. The names, sequences and product sizes of the primers used for <i>SH2D1A</i> promoter cloning.	58
Table 18. The names and sequences of the primers used for sequencing of the pGEM(SAP G) and pGEM(SAP A) constructs.	59
Table 19. Polymorphisms genotyped in the <i>CCL2</i> gene.	69
Table 20. Single-point statistical analysis in the <i>CCL2</i> case-control study.	69
Table 21. Single-point statistical analysis in the <i>CCL2</i> TDT.	70
Table 22. <i>CCL2</i> case-control haplotype analysis by Haploview.	71
Table 23. <i>CCL2</i> TDT haplotype analysis by Haploview.	71
Table 24. Comparison of allele frequencies of rs1024611 between populations.	72
Table 25. Polymorphisms genotyped in the <i>IL4</i> gene in the South African Coloured population.	75
Table 26. Single-point statistical analysis in the <i>IL4</i> case-control study.	75
Table 27. Single-point statistical analysis in the <i>IL4</i> TDT.	76
Table 28. <i>IL4</i> case-control haplotype analysis by Haploview.	77
Table 29. <i>IL4</i> TDT haplotype analysis by Haploview.	77
Table 30. Comparison of allele frequencies of rs2243250 between populations.	78
Table 31. Polymorphisms genotyped in <i>IL10</i> in this study.	79
Table 32. Single-point statistical analysis in the <i>IL10</i> case-control study.	80
Table 33. Single-point statistical analysis in the <i>IL10</i> TDT.	80
Table 34. <i>IL10</i> case-control haplotype analysis by Haploview.	82
Table 35. <i>IL10</i> TDT haplotype analysis by Haploview.	82
Table 36. <i>IL10</i> three SNP haplotype analysis by Haploview.	82

Table 37. <i>IL10</i> four SNP haplotype analysis by Haploview.....	83
Table 38. Comparison of allele frequencies of <i>IL10</i> markers between populations.....	84
Table 39. Comparison of 3 SNP haplotype frequencies of <i>IL10</i> between populations..	84
Table 40. Polymorphisms genotyped in the <i>IL12B</i> , <i>IL12RB1</i> and <i>IL12RB2</i> genes in the South African Coloured population.	87
Table 41. Single-point statistical analysis in the <i>IL12B</i> case-control study.	88
Table 42. Single-point statistical analysis in the <i>IL12B</i> TDT.....	89
Table 43. Single-point statistical analysis in the <i>IL12RB1</i> case-control study.....	89
Table 44. Single-point statistical analysis in the <i>IL12RB1</i> TDT.	89
Table 45. Single-point statistical analysis in the <i>IL12RB2</i> case-control study.....	90
Table 46. Single-point statistical analysis in the <i>IL12RB2</i> TDT.	91
Table 47. Frequencies of the respective alleles of D5S2941 in the TB cases and controls.	92
Table 48. Microsatellite analysis of D5S2941 in the case-control study.....	93
Table 49. D5S2941 S and L allele TDT analysis.	93
Table 50. <i>IL12B</i> case-control haplotype analysis by Haploview.....	94
Table 51. <i>IL12B</i> TDT haplotype analysis by Haploview.	95
Table 52. <i>IL12RB1</i> case-control haplotype analysis by Haploview.	96
Table 53. <i>IL12RB1</i> TDT haplotype analysis by Haploview.....	96
Table 54. <i>IL12RB2</i> case-control haplotype analysis by Haploview.	96
Table 55. <i>IL12RB2</i> TDT haplotype analysis by Haploview.....	97
Table 56. Comparison of allele frequencies of <i>IL12B</i> SNPs between populations.	98
Table 57. Comparison of allele frequencies of <i>IL12RB1</i> SNPs between populations....	98
Table 58. Comparison of allele frequencies of <i>IL12RB2</i> SNPs between populations....	99
Table 59. Polymorphisms genotyped in <i>IL18</i> in this study.	102
Table 60. Single-point statistical analysis in the <i>IL18</i> case-control study.....	103
Table 61. Single-point statistical analysis in the <i>IL18</i> TDT.	103
Table 62. <i>IL18</i> case-control haplotype analysis by Haploview.	104
Table 63. <i>IL18</i> TDT haplotype analysis by Haploview.....	105
Table 64. Comparison of allele frequencies of <i>IL18</i> between populations.	105
Table 65. Genotype frequencies and association result of the <i>SH2D1A</i> variant investigated in the South African Coloured female population using a case-control study.....	109
Table 66. Allele frequencies and association results of the <i>SH2D1A</i> variant investigated in the South African Coloured population using a case-control study.....	109
Table 67. Polymorphisms genotyped in the <i>TLR2</i> and <i>TLR4</i> genes.	114
Table 68. Single-point statistical analysis in the <i>TLR2</i> case-control study.	115
Table 69. Single-point statistical analysis in the <i>TLR2</i> TDT.....	115
Table 70. Single-point statistical analysis in the <i>TLR4</i> case-control study.	116
Table 71. Single-point statistical analysis in the <i>TLR4</i> TDT.....	116

Table 72. Frequencies of the respective alleles of TLR2_GT in the TB cases and controls.....	117
Table 73. Microsatellite analysis of TLR_GT in the case-control study.....	118
Table 74. TLR2_GT TDT analysis.....	119
Table 75. <i>TLR2</i> case-control haplotype analysis by Haploview.....	119
Table 76. <i>TLR2</i> TDT haplotype analysis by Haploview.....	120
Table 77. <i>TLR4</i> case-control haplotype analysis by Haploview.....	121
Table 78. <i>TLR4</i> TDT haplotype analysis by Haploview.....	121
Table 79. Comparison of allele frequencies of rs4696480 and rs5743708 between populations.....	122
Table 80. Comparison of allele frequencies of rs4986790 and rs4986791 between populations.....	122
Table 81. Tagging SNPs genotyped in <i>WNT5A</i> and <i>FZD5</i> in this study.....	126
Table 82. Single-point statistical analysis in the <i>WNT5A</i> case-control study.....	127
Table 83. Single-point statistical analysis in the <i>WNT5A</i> TDT.....	128
Table 84. Single-point statistical analysis in the <i>FZD5</i> case-control study.....	128
Table 85. Single-point statistical analysis in the <i>FZD5</i> TDT.....	129
Table 86. <i>WNT5A</i> case-control haplotype analysis by Haploview.....	130
Table 87. <i>WNT5A</i> TDT haplotype analysis by Haploview.....	130
Table 88. <i>FZD5</i> case-control haplotype analysis by Haploview.....	132
Table 89. <i>FZD5</i> TDT haplotype analysis by Haploview.....	132
Table 90. Comparison of <i>WNT5A</i> allele frequencies between the South African Coloured population and the four HapMap populations.....	133
Table 91. Comparison of <i>FZD5</i> allele frequencies between the South African Coloured population and the four HapMap populations.....	134
Table 92. Comparison of TB, sarcoidosis and Crohn's disease.....	136
Table 93. Single-point statistical analysis in the <i>BTNL2</i> case-control study design.....	140
Table 94. Single-point statistical analysis in the <i>BTNL2</i> TDT.....	141
Table 95. <i>BTNL2</i> case-control haplotype analysis by Cocaphase.....	142
Table 96. <i>BTNL2</i> TDT haplotype analysis by TDTphase.....	143
Table 97. Comparisons of the allele frequencies in the African-American, white American, South African Coloured and German control populations.....	145
Table 98. Genotype frequencies of the <i>CARD15</i> variants investigated in the South African Coloured population using a case-control study design.....	148
Table 99. Allele frequencies of the <i>CARD15</i> variants investigated in the South African Coloured population is a case-control study design.....	148
Table 100. TDT results for <i>CARD15</i>	148
Table 101. Genes differentially expressed between the four TB groups.....	151
Table 102. Genes differentially expressed between the latent and recurrent groups....	152
Table 103. Polymorphisms genotyped in <i>SOCS3</i>	154
Table 104. Polymorphisms genotyped in the <i>TEX264</i> gene.....	156

Table 105. Single-point statistical analysis in the <i>TEX264</i> case-control study.	156
Table 106. Single-point statistical analysis in the <i>TEX264</i> TDT.	156
Table 107. <i>TEX264</i> case-control haplotype analysis by Haploview.	157
Table 108. <i>TEX264</i> TDT haplotype analysis by Haploview.	157
Table 109. Single-point statistical analysis in the <i>SOCS3</i> case-control study.	158
Table 110. Single-point statistical analysis in the <i>SOCS3</i> TDT.	158
Table 111. <i>SOCS3</i> case-control haplotype analysis by Haploview.	159
Table 112. <i>SOCS3</i> TDT haplotype analysis by Haploview.	160
Table 113. Polymorphisms genotyped in <i>CTSZ</i> in this study.	162
Table 114. Single-point statistical analysis in the <i>CTSZ</i> case-control study design.	163
Table 115. Single-point statistical analysis in the <i>CTSZ</i> TDT.	164
Table 116. <i>CTSZ</i> case-control haplotype analysis by Haploview.	165
Table 117. <i>CTSZ</i> TDT haplotype analysis by Haploview.	165
Table 118. Functional polymorphisms previously associated with other immune-mediated diseases.	168
Table 119. Single-point statistical analysis in the case-control study.	171
Table 120. Single-point statistical analysis in the TDT.	172
Table 121. Polymorphisms genotyped and their single-point association results.	175
Table A1. SNPs detected during <i>BTNL2</i> sequencing.	232

LIST OF PUBLICATIONS

Möller M, Nebel A, Kwiatkowski R, van Helden P, Hoal E, Schreiber S : Host susceptibility to tuberculosis: *CARD15* polymorphisms in a South African population *Molecular and Cellular probes* 2007. 21: 148-151

Möller M, Kwiatkowski R, Nebel A, van Helden P, Hoal E, Schreiber S: Allelic variation in *BTNL2* and susceptibility to tuberculosis in a South African population *Microbes and infection* 2007. 9: 522-528

Van Helden PD, **Möller M**, Babb C, Warren R, Walzl G, Uys P, Hoal E. TB epidemiology and human genetics. *Innate Immunity to Pulmonary Infection, Novartis Foundation Symposium* 2006. 279:17-31

CHAPTER 1

INTRODUCTION

1.1. TUBERCULOSIS

Tuberculosis (TB) is a common infectious disease in the developing world and can be fatal when left untreated. The disease most frequently affects the respiratory system, but it can also affect bones, joints, the central nervous system, the lymphatic system, the circulatory system and the genito-urinary system. Even though TB has been studied for centuries, it is still responsible for more human deaths than any other single infectious agent. The causative bacterium in humans is *Mycobacterium tuberculosis* (*M.tuberculosis*) and although the development of effective antibiotics during the twentieth century has reduced mortality rates in developed countries [1], TB continues to be a major threat to the world's population. In developing countries, such as South Africa, TB has been a public health problem of long standing.

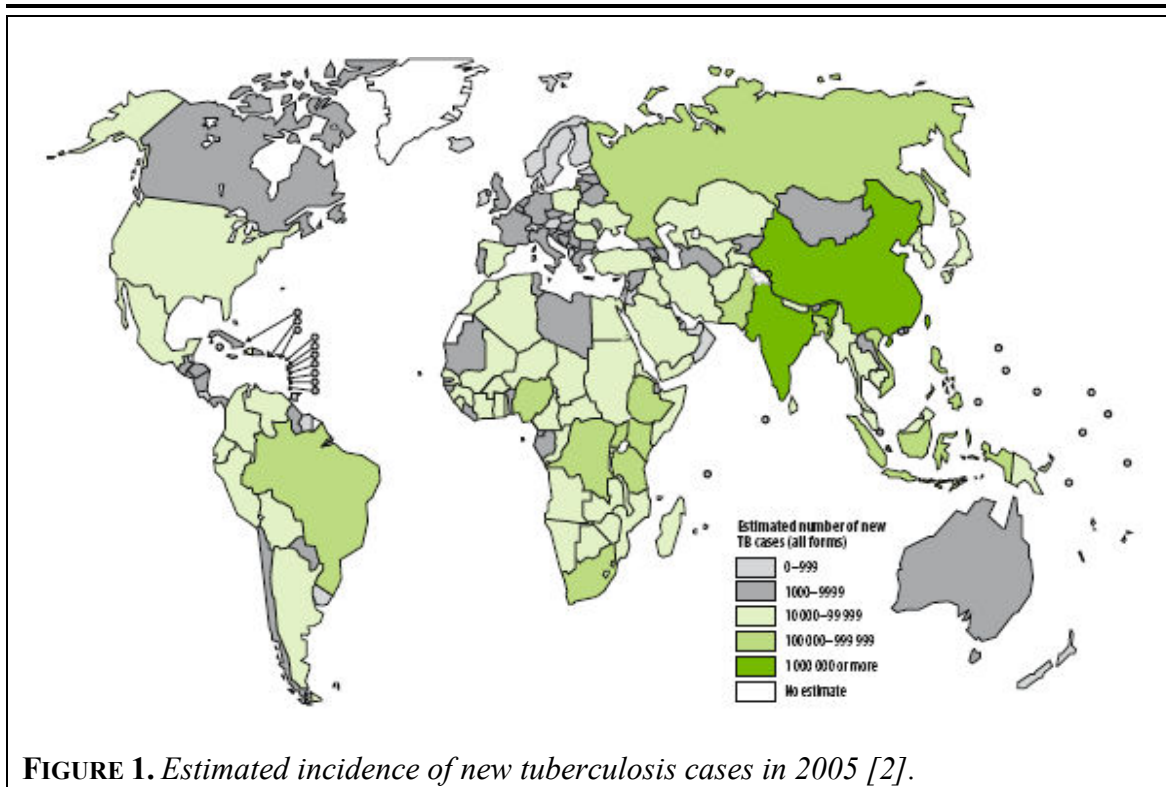


FIGURE 1. *Estimated incidence of new tuberculosis cases in 2005 [2].*

Worldwide, *M.tuberculosis* is thought to infect one new case every second [3] and one third of the global population is latently infected. Only 10% of those infected will ever develop the active disease [4], while 90% will remain healthy. It is estimated that in 2005, 1.6 million people died of TB and approximately 8.8 million new cases (Figure 1) developed active disease [2]. The prevalence of TB is still increasing globally, especially in Africa, where the incidence of the disease parallels the spread of the human immunodeficiency virus (HIV) [4].

1.1.1. The history of TB

TB has been a human burden throughout the history of mankind and evidence of the bacterium has been found in several archaeological sites, ranging from the Greek and Roman empires to locations in Europe and ancient Egypt [5]. The disease was probably described in early written accounts from many cultures [6] and was documented as

phthisis by Hippocrates. TB had several other attention-grabbing names, such as consumption, white plague, king's evil and lupus vulgaris [5] and well known victims of the disease include John Keats, Emily and Charlotte Brontë, Lord Byron, Eleanor Roosevelt and Vivian Leigh [1]. The typical symptoms of TB, such as a thin and pale face, became signs of beauty during the nineteenth century as the disease was romanticised by artists and writers [1]. Aristotle (384-322 BC) was probably the first to suggest that TB was contagious, stating that "one takes the disease because there is in the air something disease-producing" [6]. Several other individuals also postulated the contagiousness of the disease, such as the Florentine physician Fracastorius (1546) [6], the Italian anatomist Valsalva (1666-1723) [6] and the British doctor Benjamin Martin (1790) [7]. Others considered TB to be hereditary, since the disease was often observed in many members of the same family. However, the infectious nature of the disease was only convincingly proven in 1865 by Jean-Antoine Villemin [7].

Robert Koch made the important discovery that demystified the cause of TB in 1882, namely, by isolating and culturing the infectious organism [5]. This discovery diverted attention to the significance of the pathogen and the importance of the host was ignored for many years thereafter. In 1890 Koch produced a glycerine extract of the bacterium, named "old tuberculin", which he presented as a cure to the disease [1]. This announcement proved to be false, since tuberculin did not protect against TB, but led to severe inflammatory reactions in individuals with active disease [8]. However, it was found that tuberculin could be used to distinguish between *M.tuberculosis*-infected and uninfected animals based on their reaction when inoculated with it [1]. Clemens Freiherr von Pirquet further investigated tuberculin and determined that children who did not have active TB, but had a positive tuberculin reaction, were latently infected with the bacterium [7]. Old tuberculin resulted in non-specific reactions and Florence Seibert developed purified protein derivate (PPD) [7]. This is still widely used as a diagnostic test. Koch also developed staining methods for the identification of the bacterium. These techniques were improved on by Paul Ehrlich, whose detection method was used to develop the Ziehl-Neelsen staining method, which is still an important TB diagnostic tool [1].

Treatment before the advent of antibiotics was based on rest, fresh air and a healthy diet. These measures were implemented at sanatoriums dedicated to the treatment of TB [7]. South Africa, with a climate that was proclaimed as being beneficial to TB patients, was therefore promoted as health resort during the nineteenth century and numerous sanatoria were established [6]. Several therapies were employed in these institutions. This included the collapse of lungs by removing part of the ribs (thoracoplasty) or by collecting air in the pleural space (pneumothorax). The goal of these practices was the closure of TB cavities, which would hopefully result in sputum sterilisation and a decrease in the contagiousness of the TB patient [7].

TB was one of the first diseases for which a vaccine was developed [8]. The first success in immunising against the disease was achieved by Albert Calmette and Camille Guérin [7]. They attenuated *Mycobacterium bovis* (*M.bovis*) for use as a vaccine, which was named bacillus Calmette-Guérin (BCG) [1]. Unfortunately, the BCG vaccine was not as effective in the developing world as in the developed countries, possibly at least partly due to the connection between *M.tuberculosis*, BCG and nonpathogenic

mycobacteria [8]. The development of an effective vaccine is still a priority today, more than 85 years after the development of BCG [9].

TB mortality rates started to decline, probably due to the improvement of socio-economic circumstances, which includes diet, of the general population. The isolation of streptomycin in 1944 [7], the first antibiotic and a chemical compound effective against *M.tuberculosis*, contributed to the lore that antibiotics would assist in reducing TB mortality rates. The development of other antibiotics effective against the bacterium, such as isoniazid and rifampicin, fuelled the belief that TB would be eradicated. However, drug-resistant isolates of *M.tuberculosis* were observed soon after the introduction of antibiotics [10]. This, together with the HIV pandemic (*Section 1.1.6.*), the global population increase, poverty and the failure of many TB programmes [1] has led to an increase in the global incidence of TB. The resurgence of TB was declared a global health emergency by the World Health Organization (WHO) in 2003, the first disease to be classified in this manner [1].

In the past, most of the understanding of *M.tuberculosis* resulted from animal studies and clinical and descriptive epidemiologic observations. The development of molecular biology techniques enhanced the understanding of the biology of the bacterium [11]. This included the complete genomic sequence [12,13] of the H37Rv *M.tuberculosis* laboratory strain, which facilitated the discovery of unique mycobacterial characteristics. Several genotyping methods used to study the molecular epidemiology of TB have made their appearance in the literature, such as insertion sequence 6110 (IS6110) restriction fragment length polymorphism analysis, spoligotyping and mycobacterial interspersed repetitive unit typing. New diagnostic techniques for rapid detection of the bacterium have also been developed, for example polymerase chain reaction (PCR) methods and real-time PCR techniques.

1.1.2. The pathogen

The name of the genus of the TB pathogen, *Mycobacterium*, meaning fungus-bacterium, is a reference to the mould-like pellicles which are formed when the bacteria are grown in liquid media [14]. Infection by several members from the *M.tuberculosis* complex, which also includes *M.bovis*, *M.microti* and *M.africanum*, may result in TB [14]. *M.tuberculosis* (Figure 2) however, is the main TB disease-causing bacterium in humans [15]. The mycobacteria are not only responsible for TB, but also for leprosy, where the causative bacterium is *M.leprae*. The secret of the survival of *M.tuberculosis* in the host probably lies in the fact that it has evolved specialised strategies to persist and replicate within macrophages. It is these characteristics that distinguish pathogenic and non-pathogenic mycobacteria from each other [16]. Firstly, the bacterium has the ability to block the delivery of the phagosome to the lysosome and can thus avoid instant destruction [17]. Secondly, the bacilli are able to inhibit the local immune responses that activate macrophages [18]. The cell wall of the tubercle bacillus contains high levels of lipids and the microbe is therefore not easily stained by Gram staining. However, it is considered to be Gram-positive [14]. After staining, dyes cannot be readily removed with acid-alcohol, and the bacteria are therefore also named acid-fast bacilli [19]. In sputum, or other samples, the bacilli can occur individually or in small clumps [15].



FIGURE 2. Scanning electron micrograph of *M.tuberculosis* (Courtesy: CDC).

1.1.3. Pathophysiology

Infection by *M.tuberculosis* is a complex, multistage process progressing from the first encounter with the bacterium. For this reason a multistep course of disease has to be imagined (Figure 3) [20]. The consequences of mycobacterial infection are determined by two factors, namely the virulence of the infecting strain, as well as the resistance of the host [21]. TB is spread through the air by droplet nuclei containing *M.tuberculosis* from the lungs of individuals with active disease [5] to the respiratory tract of uninfected individuals. These particles are 1-5 μm in size and can remain in the air for minutes to hours after release from diseased lungs [22]. The droplet nuclei are inhaled and move to the alveoli where the bacteria are phagocytosed by alveolar macrophages and also probably by dendritic cells [23], which have the activity to destroy most potential infectious microbes [16]. The phagocytosis of the bacterium invokes a strong host cellular immune response and a cascade of events is triggered that involves cytokines and chemokines [5]. In some cases the bacilli are immediately destroyed.

Alternatively, the bacteria will begin to replicate in the intracellular environment and migrate to lymph nodes in the lung through the lymphatic system [24]. In the first 2-8 weeks after infection, cell-mediated immunity will develop [22] and conversion to tuberculin reactivity takes place [24]. To limit the spread and replication of the bacilli, granulomas are formed by activated T lymphocytes and macrophages. The centres of these granulomas consist of dormant bacteria. At this stage the immune system can contain the infection, but if this does not occur, the infection may progress to active disease [22]. In these cases, the bacteria continue to replicate and disease symptoms will start to appear. Common symptoms of TB include persistent coughing, fever, coughing of blood, night sweats, weight loss and chest pain [22].

1.1.4. Current treatment

The production of antibiotics, such as isoniazid and streptomycin played a role in the considerable decrease in TB mortality rates seen during the TB chemotherapy revolution [1]. Before the introduction of these medicines, 50% of patients with active TB died within two years [25]. The main goal of treatment is to ensure the elimination of the bacteria from the host, while preventing transmission of the disease, drug resistance and death [22]. Unfortunately, long-term treatment with a combination of drugs is needed to cure TB. Most treatment regimens have two phases, namely

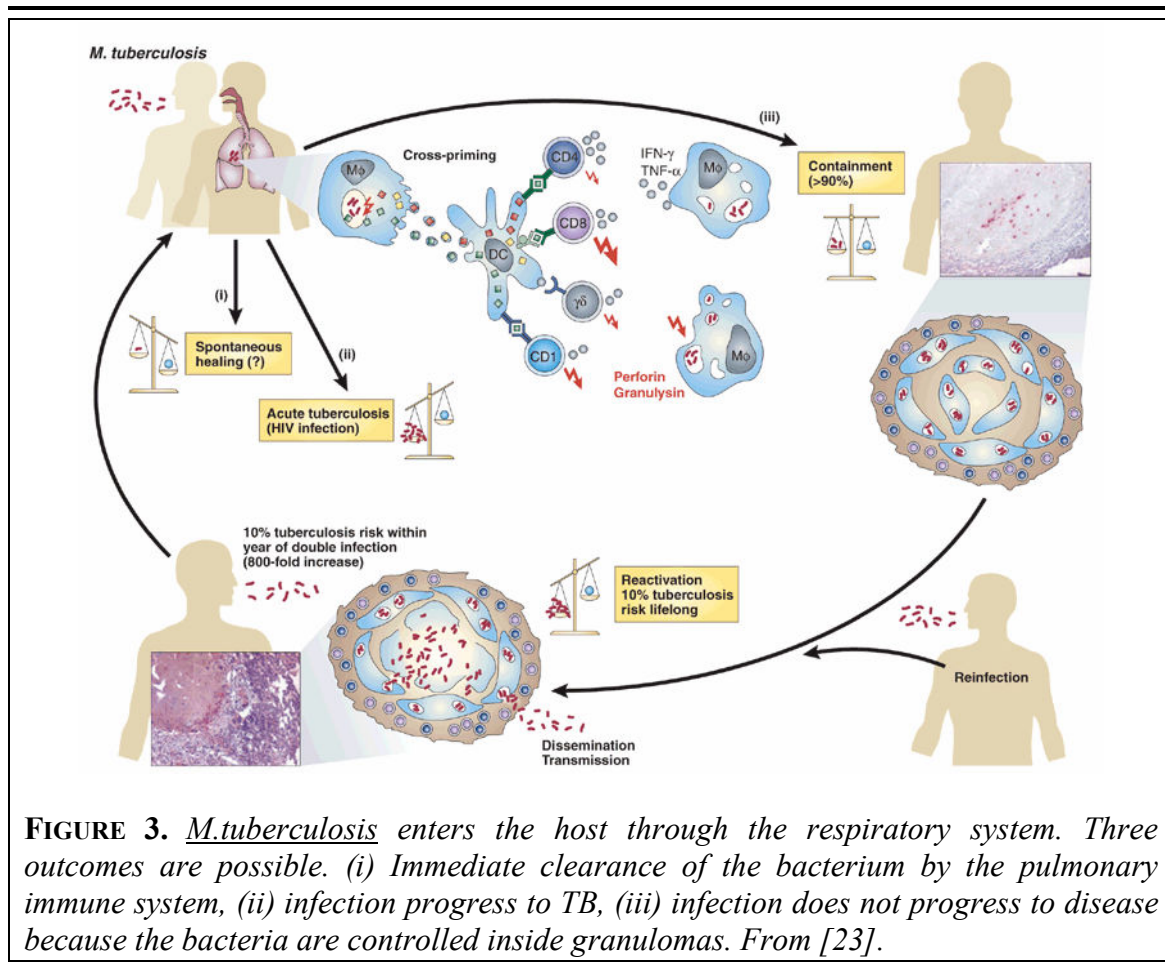


FIGURE 3. *M.tuberculosis* enters the host through the respiratory system. Three outcomes are possible. (i) Immediate clearance of the bacterium by the pulmonary immune system, (ii) infection progress to TB, (iii) infection does not progress to disease because the bacteria are controlled inside granulomas. From [23].

intensive and continuation [25]. A typical intensive phase lasts for two months and most patients receive first-line drugs [1], such as isoniazid, rifampicin, pyrazinamide and ethambutol. After two weeks of treatment, the patient is no longer considered to be contagious. During the continuation phase of four months, the therapy is continued with isoniazid and rifampicin [25]. In cases with multidrug-resistant TB (MDR-TB), the treatment period is extended and second-line drugs, such as ethionamide and kanamycin [26], are incorporated into the regimen. Extreme cases with MDR or extensively drug-resistant (XDR) infections may be treated by surgical means, which involves the removal of granulomas [1].

1.1.5. Drug-resistant TB

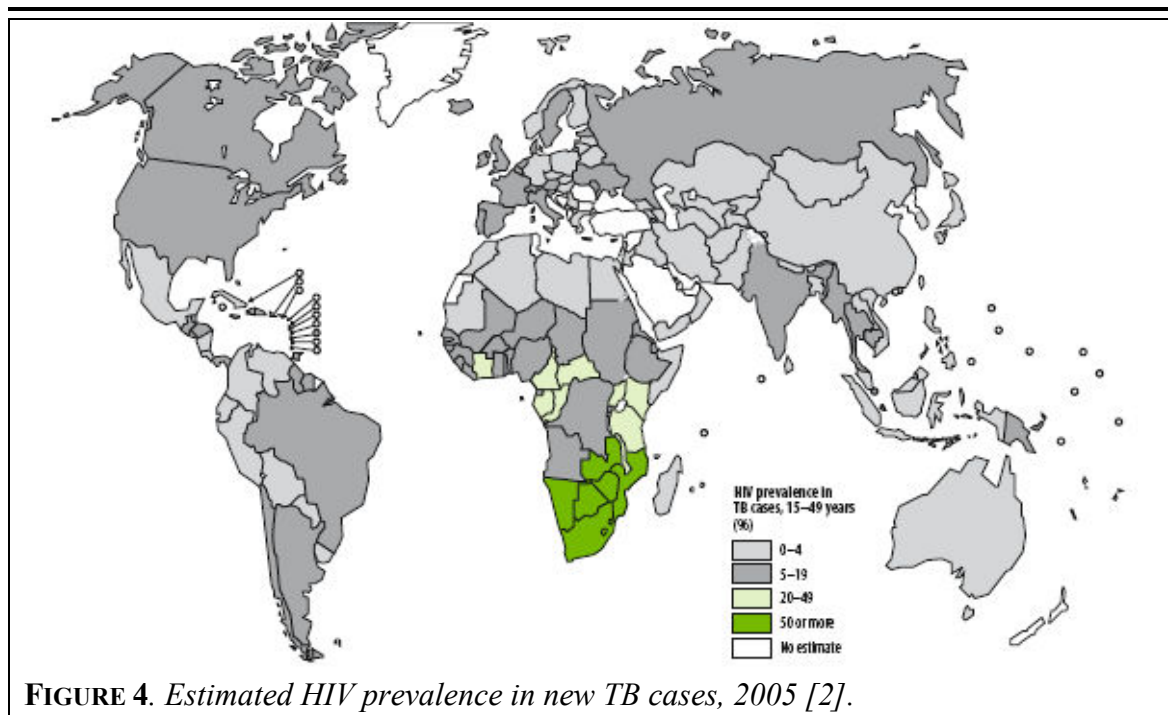
The appearance of the first drug-resistant strains of TB was noted soon after the introduction of antibiotics [10]. While anti-TB drugs kill susceptible *M.tuberculosis* they also select for drug-resistant bacteria, since these bacteria are resistant to the drugs [26]. One of the other contributing factors to drug resistance is the exceptional length of treatment (Section 1.1.4) which leads to the interruption of therapy by some patients. Drug-resistant TB is defined as resistance to at least one first-line drug [27]. During the early 1990's an increased incidence of MDR-TB was observed [5]. These are *M.tuberculosis* strains that are resistant to at least isoniazid and rifampicin [26], two of the most effective first-line drugs [5]. The MDR-TB strains can be transmitted and the

most extensive MDR-TB outbreak reported to date was in the United States [28]. Infections with MDR-TB are more difficult and expensive to treat and the rate of mortality ranges from 40 to 60%, which is similar to that of patients with untreated TB [29].

Recently a new class of MDR-TB was defined and named XDR-TB. These isolates are resistant to isoniazid, rifampicin, any fluoroquinolone and at least one of three injectable second-line drugs (capreomycin, kanamycin, and amikacin) [30]. Treatment success in XDR-TB is very difficult to achieve as few active drugs, if any at all, remain to combat the disease.

1.1.6. HIV and TB

The worldwide resurgence of TB is linked to the HIV pandemic and almost one-third of the 40 million HIV infected individuals in the world are co-infected with TB [31]. Currently, HIV infection is the major cause of progression from latent TB to active disease.



The production of stimulatory cytokines and the decrease in CD4 cells during TB infection also accelerates the progression of HIV to acquired immune deficiency syndrome (AIDS) [32]. The highest prevalence of HIV in new TB cases is found in Africa (Figure 4), with more than 50% [4] of new TB cases in South Africa estimated to be co-infected with both organisms. TB also differs from other opportunistic infections observed in HIV positive patients, since the bacterium can be transmitted to HIV negative patients through the air [31].

1.2. HOST GENETIC SUSCEPTIBILITY TO TB

Although infection with *M.tuberculosis* is required, it is not sufficient to cause TB in most humans. This is substantiated by the observation that only 10% of infected people who are immunocompetent will ever develop the active disease, while the majority of the population control the bacterium effectively. The genetic make-up of the host could determine the outcome after infection and several studies have proven that genetic factors contribute to the outcome of TB (*Section 1.2.1.*). Before this finding was widely accepted, the misconception existed that other diseases, such as cardiovascular disease, cancer and diabetes were influenced by genetic factors, but that death from infection was due to unfavourable conditions or misfortune [33]. A classical epidemiological study on the premature death of adoptees in Denmark suggested that the genetic contribution to infectious disease is greater than for cancer or cardiovascular disease [34]. The explanation of the genetic control of susceptibility to TB is expected to provide new and more effective tools for prevention and control of this problematic disease. However, up to the present day this field has achieved limited success and some of the genes and mechanisms which determine susceptibility to TB still remain unidentified.

TB differs from Mendelian diseases, since disease in a family does not follow a Mendelian pattern and is polygenic and multifactorial. Genetic susceptibility studies in TB are exceptionally complicated because of the presence of two different genomes and the influence their interaction can have on the disease. Although several genes have been identified as susceptibility genes for TB (*Section 1.2.3.*), it is necessary to keep in mind that other genes, the environment (including socio-economic conditions) and *M.tuberculosis* itself can have an influence on the disease, this being the reason that no single major susceptibility gene has been identified for TB [35].

1.2.1. Evidence for a genetic influence in susceptibility to *M.tuberculosis*

The idea that a genetic influence is present in susceptibility or resistance to TB has long been established and differences in susceptibility to TB in diverse populations were recognised even 100 years ago [33]. Before the discovery of *M.tuberculosis*, it was observed that TB frequently occurred in several members of the same family. This convinced some scholars that the disease was hereditary. However, the discovery of the bacterium led to the interpretation that host genetic factors were excluded from the risk of contracting TB and turned attention to the importance of the pathogen, while the host genome was largely ignored.

The role of the host genome in TB has been suggested by several observations. The inhabitants of the Qu'Appelle Indian Reservation serve as a first example. When they were originally exposed to TB, almost 10% of the population died of the disease per annum. The high prevalence of TB in this population was, in part, a consequence of the lack of "natural" resistance through the exposure to tubercle bacilli for many generations. After 40 years more than half of the families were eradicated, but the death rate had fallen to 0.2%, a decrease that could be attributed to the strong selection against susceptibility genes for TB [36]. This is a very plausible explanation, since it has also been observed that Europeans have greater resistance to TB than populations from sub-Saharan African descent, most likely due to the longer time that European populations

have been in contact with the bacterium [37]. These population differences are not only due to social factors, since a study in a nursing home in the USA determined that individuals from African descent were twice as likely (compared to individuals from European descent) to be infected with *M.tuberculosis* [38]. Another observation, one of the oldest and arguably the most important, is the wide range of responses seen in individuals exposed to *M.tuberculosis*. Some individuals are never infected, some are infected but never develop clinical disease and others are infected and present with active TB. A tragic example of this variable outcome of disease was provided in 1926 by the inadvertent immunisation of children with a virulent strain of *M.tuberculosis* in Lübeck, Germany. The same dose of bacteria was given to 251 children. Of these children, 77 died, 127 had radiological signs of disease and 47 showed no evidence of TB [39].

Some of the most convincing evidence that host genes are important in TB susceptibility was provided by classical twin studies. Monozygous and dizygous twins were compared to see how often both members develop the disease (concordance, Table 1). It was found that monozygous twins, who are considered to be identical in their genetic make-up, are more likely to show concordance for the development of TB than dizygous twins [40-42]. This indicated that genetic factors are involved in the progression of the disease, since twins share a similar environment.

TABLE 1. *Selected twin studies of TB, adapted from Cooke et al [43].*

Country	Concordance in monozygotes (%)	Concordance in dizygotes (%)	Reference
UK	32	25	[40]
USA	62	18	[41]

Mendelian susceptibility to mycobacterial infection is a rare human syndrome and affected individuals are exceptionally susceptible to otherwise non-pathogenic mycobacteria, including BCG, and *Salmonella*. Numerous studies have identified the mutations in these cases, mostly in genes that are vital for immunity against intracellular pathogens in the interleukin-12/interleukin-23/interferon- γ axis [44-50]. The existence of these individuals implies that the normal human genome could contribute to susceptibility to TB and the findings suggested candidate genes to investigate in the general population (*Chapter 4*).

Studies in animal models of mycobacterial infections have also found evidence that a genetic component is involved in susceptibility to TB. Several animal models for TB exist, such as the mouse, rabbit, guinea pig, fish and non-human primates [51]. However, mice and rabbits are most often used in studies assessing the genetic component of TB susceptibility. Rabbits are relatively resistant to TB [52], but Lurie et al [53] used the rabbit as an animal model and developed inbred resistant and susceptible strains (which were subsequently lost, unfortunately). These strains were infected with aerosols of human and bovine TB. More viable bacteria were present in the susceptible rabbits, and this was primarily due to the bactericidal effectiveness of the alveolar macrophages of the resistant strain [52]. Inbred strains of mice showed

different patterns of susceptibility after infection with *M.tuberculosis*. Resistant mice can control bacterial replication, reduce lung injury and survive longer, while susceptible mice show more severe symptoms and die prematurely [54]. The use of the mouse model led to the identification of the first TB susceptibility gene, discussed later (Section 1.2.3.2).

1.2.2. Approaches and study designs used in TB host genetics

The ideal study design for evaluating genetic susceptibility in complex diseases such as TB would be based on the complete resequencing of the human genome in a large collection of cases and controls, since this would allow genotyping of all the variation, rare or common, in an individual’s genome [55]. However, at this stage, applying such a method would be time-consuming, expensive and impractical. Given this, several study designs and approaches have been devised and successfully used in complex disease genetics (Figure 5). These include genome-wide linkage and candidate gene association analyses, which have been used in the majority of genetic susceptibility studies in TB. All the susceptibility genes involved in TB will not be identified by applying a single study design, since all study approaches have limitations. Two broad strategies exist to identify genes involved in a complex disease: 1) the experimental and 2) hypothesis-driven approaches [56]. Experimental approaches, which include linkage analysis and genome-wide association analysis, can identify a region containing an unknown potential susceptibility gene. Candidate genes can also be selected by hypotheses which are based on the functions of putative susceptibility genes in animals or humans. This approach utilises association studies and screens for mutations or polymorphisms in the gene of interest. Genetic nomenclature, also used in TB host genetics, are defined in Table 2.

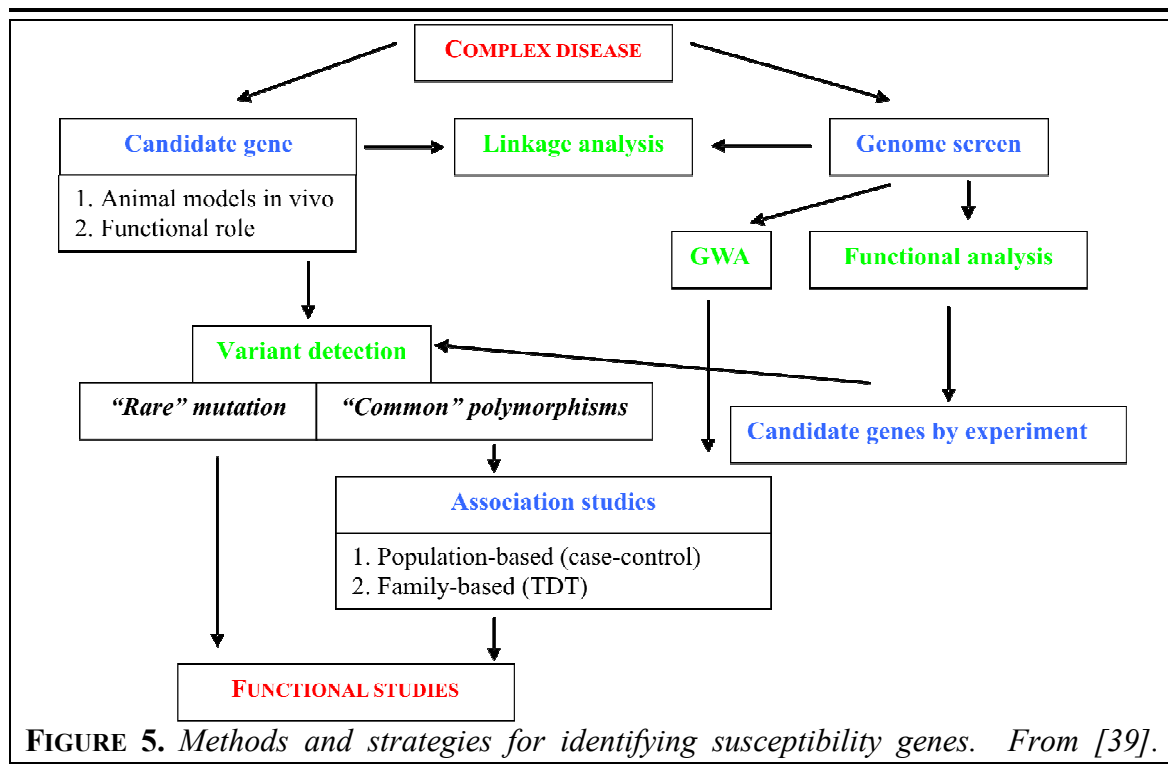


FIGURE 5. Methods and strategies for identifying susceptibility genes. From [39].

TABLE 2. *Genetic nomenclature. Adapted from [57].*

Term	Definition
Locus	A specific position in DNA.
Polymorphism	Variations in DNA that originated during evolution as a result of mutations. Indicates that a locus has more than one form in the population. Should occur with a frequency of greater than 1% in the population to be classified as such.
Allele	Refers to a particular form occurring at a locus. In humans, each individual will have a maximum of two forms, since it is possible that their parents differed from each other at this locus and that the individual inherited a different form from each parent.
Genotype	A description of the two alleles at a given locus.
Single nucleotide polymorphisms	A variation type in DNA where a single nucleotide has more than one allele in a population. Coding SNPs are found in genes and non-coding SNPs are present in promoters, introns, or intergenic regions. Synonymous SNPs will not result in an amino acid change in the protein, while non-synonymous SNPs will lead to a change of the amino acid at that position. (SNPs)
Microsatellite repeat	A more complex form of variation. Consists of short nucleotide sequences (di-, tri-, tetra- or pentanucleotides) that are repeated several times in DNA.
Insertion/deletion	A polymorphism that consists of insertions or deletions. (indel)
Haplotype	A section of DNA on a single chromosome where certain alleles of different markers tend to be inherited as a unit.
Linkage disequilibrium	The non-random association of alleles at different loci. (LD)
Haplotype tagging SNP	A SNP that acts as an identifying marker for other SNPs through LD. Using haplotype tags reduces genotyping costs and redundancy.
True association	Association with the true causal variant. Also known as direct association.
Association by LD	Association with a marker in strong LD with the causal locus. Also known as indirect association.
False positive association	Association due to confounding by stratification between cases and controls. Also known as confounded association.

Polymorphisms used in genetic studies can be selected from resequencing data, single nucleotide polymorphism (SNP) databases (such as dbSNP [58]) or from the International HapMap project [59]. Resequencing of genes, made possible by the elucidation of the complete human genome sequence [60,61], allows the identification of novel polymorphisms. These novel polymorphisms may well be involved in genetic susceptibility to disease in particular populations and could therefore be used in association studies. The dbSNP database, started in September 1998, is a public repository to deposit genetic variation. The database is available freely [58] and contains more than 12 million sequence variations, as well as data from the International HapMap project and the Human Genome Project.

The HapMap project attempts to determine the common arrangement of DNA variation in the human genome [59]. The goal of the first phase of the study was to genotype one common SNP (with an allele frequency greater than 5%) every 5 kilobases in 269 DNA samples from 4 populations [62]. Phase 2 will extend the project by typing another 4.6 million SNPs in the HapMap samples. The project proved the redundancy among SNPs located closely together (which enabled researchers to select haplotype tagging SNPs which represent the common variation in a genomic region, thereby optimising association studies) and made it possible to extract genomic variation information without resequencing.

1.2.2.1. Linkage studies

Linkage studies are used to trace chromosomal regions containing putative susceptibility genes, either by a genome-wide scan, which ensures that all major genomic regions involved in disease susceptibility are identified, or by concentrating on a candidate region. This provides the opportunity to find new genes and pathways that might not previously have been thought to contribute to the disease studied. This approach can be very successful in monogenic diseases, since it allows the fine-mapping of the gene of interest. However, this is not the case in complex disease studies, where (in general) several regions containing many genes may be identified. Linkage studies assume that chromosomal regions segregate non-randomly with the disease of interest in large affected families and aim to identify these regions [39]. Model-based linkage analysis by the logarithm of odds (lod) score method needs a defined model that specifies the relationship between the phenotype and factors that may influence its expression. The model is provided by segregation analysis. Model-free linkage studies are used when little is known about the relationship between the phenotype and the gene, which is generally the case with complex diseases. Once evidence for linkage has been found, fine genetic and physical maps are constructed to narrow down the interval on the chromosome and allow gene identification by candidate gene selection (when the function of the gene is known) or by positional cloning (when the function is unknown) [63].

A TB linkage study focused on the chromosome 17q11-q21 candidate region. This region is syntenic to mouse chromosome 11 which was previously identified as a susceptibility region for another intracellular disease [64]. The results showed evidence that the chromosome 17q11-q21 region contained TB susceptibility genes and further evaluation of this region suggested that four genes contributed separately to susceptibility [65].

Genome-wide scans have been employed in TB genetics. The first of these studies in TB was done using sibpairs from The Gambia and South Africa [66] and identified chromosome 15q and Xq as containing possible susceptibility genes. Fine-mapping of chromosome 15q11-13 suggested that ubiquitin protein ligase E3A (human papilloma virus E6-associated protein, Angelman syndrome) (*UBE3A*) or a nearby gene may be involved in TB pathogenesis [67]. A linkage study in an extended family of aboriginal Canadians indicated chromosome 2q35 to be involved in disease [68]. Another study in Brazilians, which examined TB and leprosy families, suggested three regions (10q26.13, 11q12.3 and 20p12.1) [69]. A recent genome scan in Moroccan families recognized chromosome 8q12-q13 as containing a major dominant acting TB susceptibility gene [70]. A fifth study, using South African and Malawian populations, identified areas on chromosome 6 (Cooke et al, unpublished data). The little overlap between susceptibility regions observed in these studies is probably because linkage of a genotype for a genetic marker to infectious disease may be unique to a specific family or population and thus impossible to identify in other studies [71]. These studies are also exquisitely sensitive to the phenotype definition.

1.2.2.2. Association studies

Association studies attempt to investigate the role of polymorphisms in candidate genes selected on the basis of their function or location [39] in humans or animals. These studies have greater power than linkage analyses and can detect genes of smaller effect [72]. Traditionally association studies have a population-based case-control study design, but family-based association studies are also used and recently whole genome association studies have been successfully done. Types of population-based association studies are given in Table 3.

Association studies can be used to examine the influence of single polymorphisms, multiple polymorphisms in a candidate gene or the most common SNPs representing all variation in the complete genome. When choosing a candidate gene to investigate in an association study, the influence of the gene on the development of TB and the functional effects of the particular polymorphism should ideally be taken into consideration to increase the likelihood that only true associations will be detected. In the past, polymorphisms studied were frequently genetic markers and not necessarily polymorphisms that had a direct effect on the function of the gene product. Due to technological advancements, such as the availability of the human genome sequence, it is now usually possible to choose a functionally important polymorphism in the gene. A genetic polymorphism is more likely to affect a function if it is associated with an amino acid substitution in the gene product, if it results in a frameshift mutation in the coding area or if it affects gene transcription, RNA splicing, mRNA stability or mRNA translation [71].

A problem with population association studies using a single polymorphism or candidate gene approach is that not all polymorphisms in the candidate gene are examined for susceptibility or resistance. Finding no association with an allele does not necessarily mean that the whole candidate gene is excluded as a susceptibility gene, only that a specific allele is disqualified [73]. Whole genome association studies promise to address this issue. Association of an allele with TB is only a statistical finding and is not automatically a reflection of genetic linkage, since the associated

allele may be in linkage disequilibrium (LD) with the true susceptibility allele nearby, or can be a false positive.

TABLE 3. *Types of population association studies. Adapted from [74].*

Association study type	Description
Candidate polymorphism	Individual polymorphisms that are likely to be involved in disease are investigated.
Candidate gene	Several polymorphisms (usually in the putative promoter, coding exons and possible splice sites) in a candidate gene are typed. The gene is selected from a linkage study or on the basis of its homology with a gene of known function in a model organism.
Haplotype tag SNPs	A small subset of the SNPs which are sufficient to capture the full haplotype information of a gene (“tag” the gene) are genotyped. This reduces genotyping costs.
Fine mapping	A study done in a candidate region, identified by a linkage study, genome-wide association study or admixture mapping. This region may contain many genes.
Genome-wide	Thousands of common SNPs are genotyped throughout the whole genome to identify variants in LD with the causal variant. In Europeans, 300 000 SNPs should be sufficient to cover the whole genome, while more SNPs would have to be genotyped in African populations.

Another factor to consider when doing association studies is that multiple comparison testing may result in false positives, since random statistical fluctuations can occur [75]. There are various ways to correct for multiple testing [76-80], but none have been deemed to be completely satisfactory (*Sections 3.11.2 and 3.11.4*) [75].

a) Population-based case-control studies for polymorphisms and candidate genes

In our case-control studies, allele frequencies of polymorphisms in candidate genes for TB susceptibility are compared in unrelated individuals with the disease, and healthy controls without TB. In other words these studies identify an association between a particular polymorphism in a gene and the disease. This study design is the most commonly used approach in association studies, since large study groups can be collected without the need to recruit the family members of the participants. A well-designed case-control study should have the cases and controls matched as closely as possible in terms of age, gender, ethnicity and geographical location i.e. they should ideally have an equal chance of contracting the disease.

Usually allele frequencies are considered as a first approach in the analysis of case-control studies. There will be twice as many alleles as participants, since each individual has two alleles at any autosomal locus. For this analysis to be valid, the cases and controls combined should be in Hardy-Weinberg equilibrium (HWE, *section*

3.11.1.). However, the allelic analysis is questionable when interpreting the relative risk of a disease. For this reason, genotypes are considered in the second analysis and this approach is preferred whenever possible. It allows the comparison of individuals who do not carry a certain marker allele (AA) with individuals who carry a single copy of the gene (AB) and with those who are homozygous for the marker allele (BB). HWE is only checked in the controls and the test provides the chance to test dominant or recessive allele effects [39].

Like all study designs, the case-control study design has some disadvantages. Many studies have reported positive associations that could not be replicated in other studies. In some cases the original study reported a false positive, but in other cases replication studies were underpowered to detect the original association [75]. Technical problems, such as methods used or genotyping cases and controls separately or on different days could also lead to false positive associations. The biggest concern in case-control studies however is population stratification. Bias may arise in these studies when there is admixture in the study population or when multiple subgroups that differ in disease prevalence are present in a sample set. This may lead to overrepresentation of one or more of these groups in the diseased cases of the association study. In this situation, a genetic marker with different allele frequencies in the various subgroups may show a false positive association with the disease [55]. However, there are methods available to correct for population stratification [81].

b) Family-based studies for polymorphisms and candidate genes

The problem of population stratification in association studies is addressed by the use of a family-based study design. This was the reason for the development of the transmission disequilibrium test (TDT) [82]. It was first used to control for population stratification while testing for linkage in the presence of association between a marker and a disease susceptibility locus. The TDT essentially evaluates if a marker allele is transmitted from heterozygous parents to their affected children more often than expected, in other words the number of parents who transmit the marker to an affected child is compared with the number of parents who transmit the other allele. This means that the transmission deviates from 50%, the expected Mendelian frequency when there is no linkage [83].

Although the TDT was introduced to test for linkage in the presence of association, it can also be used to test for association in the presence of linkage, but only if the data consists of nuclear families with one affected child [84]. The TDT has been extended to accommodate several study designs. These include the S-TDT (used when the genotypes from both parents are unavailable) [85], the 1-TDT (used when only one parent is available) [86] and a robust TDT that can handle incomplete genotypes for both parents and children without making assumptions about missing data [87]. Acquiring sample sets for these tests can be difficult and time-consuming, because DNA samples from the parents and the affected children are needed. This can be a problem if the parents are deceased, which is likely if the disease is late-onset [71].

The TDT is often used as a second study approach to test for association with a disease, as all association findings need to be replicated in another population or sample set. Association of an allele with TB is only a statistical finding and is not necessarily a

reflection of genetic linkage, but as stated above, the TDT can be used to test for linkage. For a TDT to have the same power as a case-control study, the number of families investigated should be equal to that of the unrelated cases [88]. A TDT is also more susceptible to technical errors than a case-control study [55].

c) Genome-wide association studies

Genome-wide association (GWA) studies allow the genotyping of the most frequent genetic polymorphisms in the genome without making assumptions about the genomic location of the causal variants. Since most of the genome is surveyed, it eliminates the disadvantages of the single polymorphism or candidate gene approach where only a few polymorphisms are investigated [55]. The completion of the human genome sequence, the deposition of SNPs into public databases, the rapid improvements in SNP genotyping methods (such as the development of microarray platforms) and the International HapMap Project have allowed the genetic association field to progress to this study design. Previous studies [89-91] and the International HapMap project have shown that most common variation in the genome can be represented by approximately 300 000 SNPs [74] in white populations. African and other populations with greater variation and less LD will need more SNPs [55] to ensure coverage of the entire genome.

GWA studies have the potential to identify many false associations and therefore replication in independent populations is essential [57]. However, these studies are very expensive. In addition, the best approach to adjust for multiple comparisons in these studies has not been determined. The proposed strategies to eliminate false-positive results are to use a two-step study design and to adopt strict rules for declaring significant associations. A few individuals are genotyped genome-wide in the first stage of the study. In the second stage, promising SNPs are genotyped in the remainder of the study population [74]. In cases where a significant association is determined, the results are verified in another population, preferably using another genotyping method to exclude technical artefacts [92]. These studies need a large sample size to ensure adequate power.

Smaller scale GWA study designs involving extensive marker coverage have been employed in other complex diseases. A study in Crohn's disease tested 19 799 coding SNPs in 735 cases and 368 controls. In the subsequent analysis, markers with $p \leq 0.01$ in the first analysis were evaluated in 380 Crohn's disease family trios, 498 single cases and 1032 unrelated controls. This study identified one novel association and three previously reported susceptibility variants for Crohn's disease. The novel finding was replicated in samples from the UK using another genotyping method [92]. Similar studies were done in myocardial infarction [89,93-95] and age-related macular degeneration [96]. A bigger GWA study in inflammatory bowel disease was done using more than 300 000 SNPs on the Illumina HumanHap300 Genotyping BeadChip and identified the interleukin-23 receptor gene as a susceptibility factor [97]. There are currently no reported GWA studies in TB or other infectious diseases, but such studies will probably be done in the future. The Wellcome Trust Case-Control Consortium has initiated a study that will consider up to 2000 cases and 3000 controls for eight complex diseases such as TB and malaria using a GWA study design [57].

1.2.2.3. Haplotypes

A haplotype is a section of DNA sequence on a single chromosome where certain alleles of different markers tend to be inherited as a unit. Therefore, an individual's genotypes at several strongly linked SNPs contribute to two haplotypes, each consisting of alleles from one parent [74]. If a susceptibility haplotype can be identified in a group of individuals with a disease, such as TB, typing alleles within the haplotype allows the identification of a conserved region, which may pinpoint the polymorphism contributing to the disease. Haplotypes are mostly inferred by the use of statistical methods, such as Phase [98,99], Fastphase [100] and Cocophase [101], since laboratory-based haplotyping or typing of family members to determine the unknown phase is expensive. However, true haplotypes are more informative than inferred haplotypes [74] and statistical methods can result in incorrect haplotype classification [102].

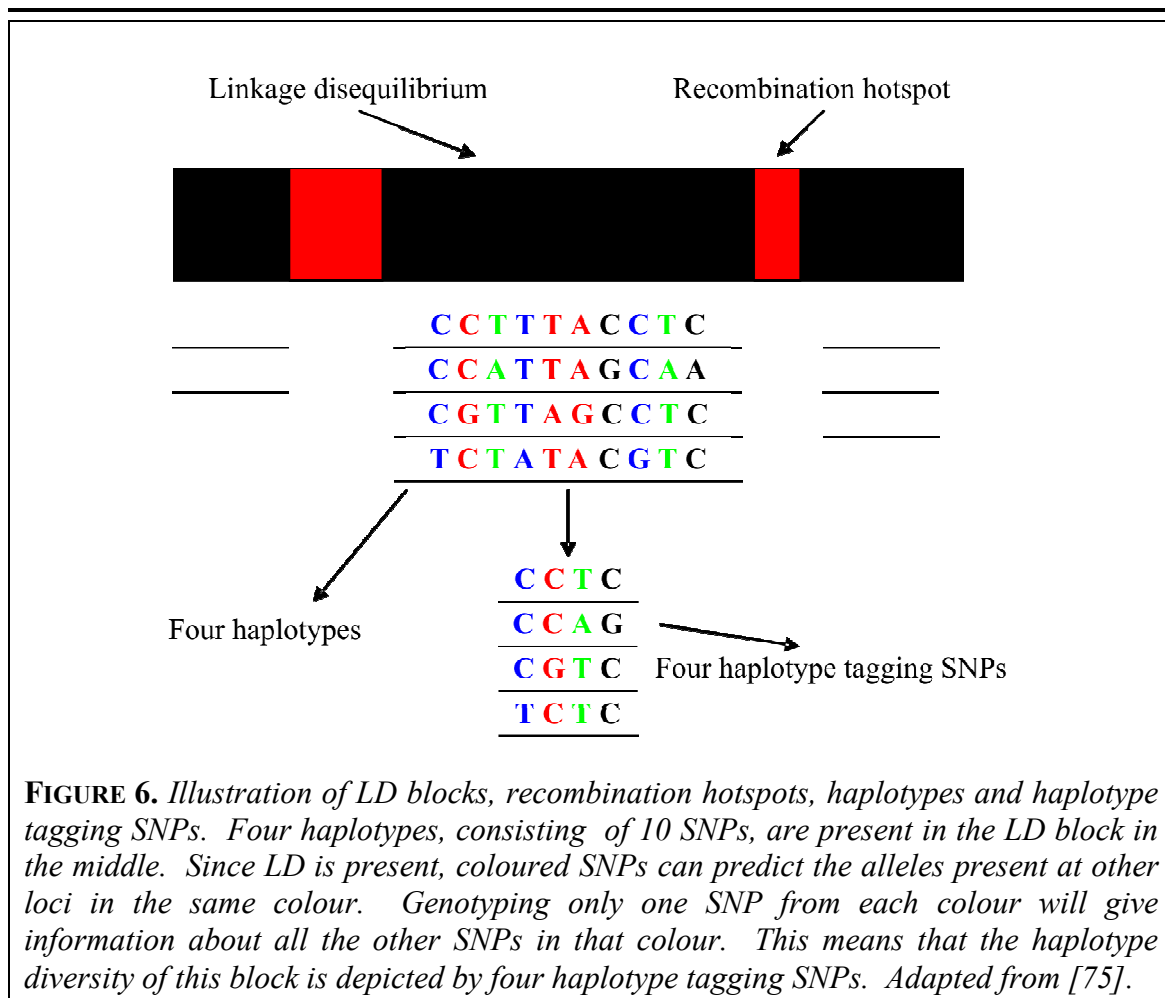


FIGURE 6. Illustration of LD blocks, recombination hotspots, haplotypes and haplotype tagging SNPs. Four haplotypes, consisting of 10 SNPs, are present in the LD block in the middle. Since LD is present, coloured SNPs can predict the alleles present at other loci in the same colour. Genotyping only one SNP from each colour will give information about all the other SNPs in that colour. This means that the haplotype diversity of this block is depicted by four haplotype tagging SNPs. Adapted from [75].

The diversity of haplotypes in the human genome is low and can lead to considerable associations of SNPs with neighbouring variants. Therefore the genome can be divided into distinct blocks separated by recombination hotspots (Figure 6) [89,91,103]. This observation led to the generation of a haplotype map (HapMap) [62] of the human genome which offers the potential to extract extensive information without resequencing and allows the selection of haplotype tagging SNPs to optimise association studies, since redundant SNPs are excluded from genotyping. However,

haplotype block structure can differ significantly between populations [91,104-107] if the block structure resulted due to selection, bottlenecks or admixture in the population history [91,108,109]. Consequently, when using information from the HapMap, it is important to determine that the data is applicable to the study population [110].

1.2.2.4. Linkage disequilibrium

LD can determine disease susceptibility, reveal the history of human populations and supply the tools for understanding fundamental biological processes [110]. LD is the non-random statistical association of sequence variants at different positions along the chromosome as it occurs in gametes. The fact that DNA mutations occur sporadically leads to linkage of SNPs along the chromosome. This means that the presence of one allele at a certain SNP can give information about the presence of alleles at other variants [62]. LD can complicate association studies, because an apparently disease-associated SNP may actually be in LD with the true disease-causing SNP (indirect association, *table 2*). Therefore it is important to consider haplotypes (*Section 1.2.2.3.*) when investigating association with disease [111]. However, LD can also aid in the discovery of susceptibility variants, since the effect can be detected by genotyping SNPs. This has been exploited in the HapMap and can assist in the selection of polymorphisms.

LD between genetic markers can be used to identify functional genetic variants that influence human diseases. Several measures of LD exist, but Lewontin's standardised disequilibrium coefficient D' [112] and the square of the correlation coefficient between two loci, r^2 [111,113], are the two most important. The D' value ranges from 0 to 1, where 0 indicates no LD and 1 indicates complete LD [113]. For example, consider two alleles at each of two loci, with A1 and A2 at locus A and B1 and B2 at locus B. If only two out of four possible haplotypes are found in the population (A1B1 and A2B2), then these loci are in complete LD and $D' = 1$ [111]. D' is useful to assess the probability of historical recombination [114]. A disadvantage of D' is that the estimate of this value is highly inflated in small samples [113]. This means that the value can indicate high LD, especially when one allele is very rare, even when the markers are in linkage equilibrium [113]. The r^2 -value reflects the statistical power to detect LD and is correlated to the D' value. In other words a low r^2 will mean that a large sample size, n , is required to detect LD between the markers. Thus, $r^2 = 1$ when two SNPs evolved together and remained associated during recombination [62], in other words when only two haplotypes are present [115]. Since intermediate values of r^2 can be easily interpreted [113] and r^2 takes allele frequency differences into account [113], this value is more useful in association studies [114]. For example, assume two nearby loci where one is functionally associated with disease, while the other is a marker in LD with the susceptibility locus. In order to have the same power to detect the association between the marker and disease, the sample size has to be increased by $1/r^2$ in comparison with the sample size needed to detect association with the true disease-causing locus [113].

1.2.3. Previous genetic susceptibility studies done in TB

Several susceptibility genes for TB have been identified (Table 4), but some were not associated with every population examined (Table 5) or could not be replicated at all (*Section 1.2.3.4.*). This emphasizes the complexities of characterising host

susceptibility in different ethnic populations living in diverse environments [113]. Even so, certain genes have consistently been associated with TB in diverse populations (Section 1.2.3.1). From the susceptibility genes identified thus far, it seems that each gene makes a small contribution to the development of the disease, which can be expected since TB is a complex disease [113]. A few of the candidate genes previously investigated in a number of TB genetic susceptibility studies are discussed in sections 1.2.3.1.-1.2.3.4.

1.2.3.1. Human leukocyte antigen genes

The human leukocyte antigen (HLA) region consists of approximately 200 genes, many of which are involved in antigen presentation. Genes that are involved in protective immunity show greater variance than other genes [116]. This is also the case for the HLA region, which varies between populations and is highly polymorphic. This is thought to be the effect of different selection pressures, such as infectious disease [117]. The HLA class I and class II genes are involved in antigen presentation to T cells and each protein binds a different range of peptides [118]. There are three class I α -chain genes in humans, namely *HLA-A*, *-B*, and *-C*. There are four subclasses of genes in the HLA class II region, named *HLA-DR*, *-DP*, *-DM* and *-DQ* respectively. These sets of genes can give rise to four types of HLA class II molecules. Genes in the HLA class III region encode, amongst others, heat shock proteins, complement proteins, tumour necrosis factor- α (TNF- α) and TNF- β [118]. HLA genes have been examined in several TB susceptibility studies (Tables 4 and 5) and were some of the first genes to be associated with the disease. The HLA-DR subtypes, in particular, were consistently associated with TB and other mycobacterial diseases.

1.2.3.2. Natural resistance associated macrophage protein 1

The natural resistance associated macrophage protein 1 (*NRAMP1*, renamed solute carrier family 11A member 1 (*SLC11A1*)) was the first susceptibility gene to be identified from a mouse model of mycobacterial disease. Resistance to several intracellular mycobacteria, including *M.bovis* (BCG) and *M.lepraemurium* [39], in inbred strains of mice is controlled by a single dominant gene, which was designated *Bcg* [119]. This gene, *Nramp1*, was isolated from the mouse genome by positional cloning and it was determined that a non-synonymous nucleotide substitution in codon 169 led to an amino acid change (glycine to aspartic acid) which caused susceptibility to mycobacteria in mice [119,120]. Additional proof of the effect of *Nramp1* was acquired when an *Nramp1* knockout mouse strain and a mouse strain with the non-synonymous nucleotide substitution had identical phenotypes [121]. *Nramp1* does not seem to affect susceptibility to *M.tuberculosis* in mice [122]. The human homologue of this candidate gene, derived from the mouse model, was named *NRAMP1* and mapped to chromosome 2q35 [123]. The protein product of *NRAMP1* is a divalent cation transporter, which is recruited to the phagolysosomal membrane when the macrophage is activated [124]. Several TB association studies in different populations have considered *NRAMP1* as a candidate gene (Tables 4 and 5). Combined, these studies suggest that *NRAMP1* is involved in *M.tuberculosis* susceptibility in humans. However, the effect of this gene is not strong enough to suggest that it is the major gene involved in TB.

1.2.3.3. Vitamin D Receptor

The vitamin D receptor (*VDR*) gene mediates the effects of the active metabolite of vitamin D, 1,25-dihydroxyvitamin D₃, which suppresses the growth of *M.tuberculosis* in vitro [125,126] by stimulating cell-mediated immunity and activating monocytes [127]. Conflicting results have been found in association studies of *VDR* in TB. The rare allele of codon 352 (“t”) was associated with protection against pulmonary TB in the Gambia [128], only associated with female TB patients from India [129] and associated with TB in the Gujarati population only when taking deficiency of serum vitamin D into consideration [130]. This polymorphism was not associated with TB in the Chinese-Han population, but the *FokI* polymorphism was associated with disease [131], while a case-control and family-based study in West African populations indicated a role for *VDR* haplotypes in susceptibility to TB [132]. Subsequent studies did not find any association between *VDR* polymorphisms and TB [133,134]. A meta-analysis of studies indicated that results were inconclusive and that the studies were underpowered [135]. A recent study in South Africans [136], which attempted to replicate a previous study [137], determined that the *Apal* “AA” genotype and “T”-containing *TaqI* genotypes predicted a faster response to TB treatment, but did not detect an association with TB in a case-control analysis.

1.2.3.4. SP110 nuclear body protein

Recently, a mouse strain extremely predisposed to TB was used to identify a susceptibility region on chromosome 1 of the mouse. This region was named susceptibility to TB 1 (*sst1*) and a candidate gene from this region, Intracellular pathogen resistance 1 (*Ipr1*), was found to mediate resistance to TB in mice [138]. The closest human homologue of *Ipr1* is SP110b, a protein encoded by the SP110 nuclear body protein gene (*SP110*). The initial association study in West Africa [139] identified three polymorphisms in the gene that possibly influenced genetic susceptibility to TB. However three large subsequent studies in Ghanian [140], Russian [141] and South African [142] populations failed to replicate this association. Since a number of independent studies could not replicate the original finding, it is likely that this gene does not contribute to genetic susceptibility to TB in humans [142].

1.2.4. Genes investigated in this study

Thirty-nine genes were chosen as candidate genes for this study according to their role in TB or other granulomatous diseases, or from DNA-array analyses, genome-wide linkage studies or genes known to be involved in other immune-mediated diseases (*Section 2.2*). These genes will be discussed in more detail in the chapters to follow.

TABLE 4. Publications reporting significant linkage or association with TB. Adapted from [57] and updated (complete as on 15-06-2007).

Candidate Gene	Population	Phenotype	Sample Size	Reported Results	Ref
MHC Region (Class I, II & III)					
A1 Supertype	Indian	pTB ^a & Miliary TB	TB = 235; Con ^b = 289	OR = 0.43; pc = 0.001	[143]
A10; B8	Indian (South)	Smear -ve pTB	TB = 152; Con = 404	p < 0.01	[144]
A2; B5	Egyptian	pTB	TB = 42; Con = 156	N/A	[145]
A26; B17; B27; DR14	Iranian	pTB	TB = 44; Con = 108	p < 0.05 (pc > 0.05)	[146]
B12	Thai	pTB	TB = 35; Con = 35?	Decreased frequency in cases	[147]
B14	Italian	Cavitary TB	TB = 54; Con = 1089	RR = 3.9; p = 0.001	[148]
Bw15	African American	pTB with Cavitation	TB = 60; Con = 100	Increased frequency in cases	[149]
Bw46	Thai	pTB	TB = 35; Con = 35?	Increased frequency in cases	[147]
Cw Allotype 1	Indian	pTB & Miliary TB	TB = 235; Con = 289	OR = 1.69; p = 0.005	[143]
Cw Allotype 2	Indian	pTB & Miliary TB	TB = 235; Con = 289	OR = 2.31; p = 0.000004	[143]
DQ B57 (Asp/Asp)	Cambodian	pTB	TB = 436; Con = 107	OR = 3.05; p = 0.001	[150]
DQA1*0101; DQB1*0501; DRB1*1501	Mexican	pTB	TB = 65; Con = 95	OR = 6.16 - 7.92	[151]
DQA1*0301; DQA1*0501	Iranian	pTB	TB = 40; Con = 100	OR = 0.25; OR = 0.53	[152]
DQA1*0601, DQB1*0301	Thai	pTB	TB = 82; Con = 160	OR = 0.4, p < 0.02	[153]
DQB1*02	Polish	pTB	TB = 38; Con = 58	OR = 0.39; p = 0.01	[154]
DQB1*0301-0304	South African Venda	TB	TB = 95; Con = 117	OR = 2.58; p = 0.001	[155]
DQB1*0402; DR4; DR8	Mexican	pTB	TB = 65; Con = 95	Decreased frequency in cases	[151]
DQB1*05	Polish	pTB	TB = 38; Con = 58	OR = 2.84; pc = 0.002	[154]
DQB1*0502	Thai	pTB	TB = 82; Con = 160	OR = 2.06, p = 0.01 (pc = 0.13)	[153]
DQB1*0503	Cambodian	pTB	TB = 126; Con = 88	p = 0.005	[156]
DR2	Indian	pTB	25 Families	p = 0.001	[157]
DR2	Russian	pTB	N/A	Increased frequency in cases	[158]

Candidate Gene	Population	Phenotype	Sample Size	Reported Results	Ref
DR2	Indian (South)	Smear +ve pTB	TB = 204; Con = 404	Attributable risk = 0.29; p = 0.01	[144]
DR2	Indian (North)	pTB	TB = 153; Con = 289	RR = 1.8; pc = 0.029	[159]
DR2; DQ1	Indian	pTB	TB = 209; Con = 122	RR = 2.3; RR = 2.8	[160]
DR2; DQw1	Indonesian	pTB	TB = 101; Con = 65	Attributable risk = 36% & 39%	[161]
DR2; DRw53	Tuvonian Russian	pTB	N/A	N/A	[162]
DR3	Mexican	pTB	TB = 51; Con = 54	Decreased frequency in cases	[163]
DR3	Russian	pTB	N/A	Decreased frequency in cases	[158]
DR4	Thai	pTB	TB = 35; Con = 35?	Increased frequency in cases	[147]
DR4	Italian	Cavitary TB	TB = 54; Con = 1089	RR = 2.7; p = 0.001	[148]
DRB1 *13 & DRB1 *14	Tuvonian Russian	pTB	14 Pedigrees	Transmitted more frequently	[164]
DRB1*07; DQA1*0101	Iranian	pTB	TB = 40; Con = 100	OR = 2.7; OR = 2.66	[152]
DRB1*0803	Korean	Drug Resistant pTB	TB = 81; Con = 200	OR = 2.63; p = 0.047	[165]
DRB1*11	Chinese	pTB	TB = 74; Con = 90	RR = 0.12; p < 0.05	[166]
DRB1*13	Polish	pTB	TB = 31; Con = 58	RR = 0.04; p < 0.001	[167]
DRB1*1302	South African Venda	TB	TB = 92; Con = 117	OR = 5.05; p < 0.001	[155]
DRB1*15	Chinese	pTB	TB = 74; Con = 90	RR = 2.91; p < 0.05	[166]
DRB1*1501	Indian	pTB	TB = 22; Con = 36	p < 0.05	[168]
DRB1*1501; DQB1*0601	Southern Indian	Sputum +ve pTB	TB = 126; Con = 87	OR = 2.68; OR = 2.32	[169]
DRB1*16	Polish	pTB	TB = 31; Con = 58	RR = 9.7; p < 0.01	[167]
HSP70-1A	Indian (North)	DR15 -ve pTB	N/A	RR = 12.6; p = 0.02	[170]
TAP-A/F	Indian (North)	pTB	TB = 57; Con - 40	RR = 4.3; pc = 0.01	[159]
TNF	Sicilian	pTB	TB = 45; Con = 100	p = 0.05	[171]
TNF (-308G; -238A)	Colombian	pTB	TB = 135; Con = 430	OR = 1.8; OR = 2.2	[172]
TNF -238	Gambian	pTB	TB = 206; Con = 229	OR = 2.54; p = 0.00001	[173]
Other Candidates:					
BCHE (cholinesterase)	Russian	pTB	N/A	RR = 6.92	[174]
CCL18 (rs2015086; rs14304)	Brazilian	pTB	92 Pedigrees (627 Ind)	RR = 0.4; RR = 0.38	[65]

Candidate Gene	Population	Phenotype	Sample Size	Reported Results	Ref
CCL2 (-2518G)	Mexican	pTB	TB = 445; Con = 334	OR = 2.43; p = 0.0003	[175]
CCL2 (-2518G)	Korean	pTB	TB = 129; Con = 162	OR = 2.63; p = 0.0001	[175]
CCL4 (rs1719144)	Brazilian	pTB	92 Pedigrees (627 Ind)	RR = 0.35; p = 0.002	[65]
CR1 (Q1022H)	Malawian	HIV-ve pTB	TB = 196; Con = 670	OR = 3.12; p = 0.03	[133]
DC-SIGN	South African	pTB	TB = 351; Con = 360	OR = 1.85; p = 8.2 x 10 ⁻⁴	[176]
ESD (Esterase)	Tuvonian Russian	pTB	TB = 73; Con = 251	N/A	[177]
Haptoglobin (2:2)	Russian	pTB	TB = 223; Con = 567	Significantly increased in cases	[178]
IFNG	Caucasian	Smear +ve pTB	TB = 113; Con = 207	OR = 3.75; 0.0017	[179]
IFNG	South African	pTB	TB = 313; Con = 235	OR = 1.64; p = 0.0055	[180]
IFNG (874 TT)	Sicilian	pTB	TB = 45; Con = 97	p = 0.02	[181]
IFNG (874T)	Turkish	pTB	TB = 319; Con = 115	OR = 0.7; p = 0.024	[182]
IFNG (874T/A)	Hong Kong Chinese	pTB	TB = 385; Con = 451	OR = 2.24; p < 0.001	[183]
IFNG (874T/A)	Colombian	TB	TB = 190; Con = 135	p = 0.01	[184]
IFNG (T/T 874)	Croatian	Microscopy +ve TB	TB = 54; Con = 175	OR = 3.12; p = 0.012	[185]
IFNGR1	Croatian	pTB	TB = 120; Con = 87	p = 0.02	[186]
IFNGR1 (-611A, -56C)	Caucasian	DNMT & pTB	TB = 55; Con = 86	Increased in Cases; p = 0.004 & 0.03	[187]
IFNGR1 (FA1 (CA)n)	Croatian	pTB	TB = 244; Con = 521	OR = 0.24; p = 0.0023	[188]
IL10	Turkish	TB	TB = 81; Con = 50	p = 0.014	[189]
IL10 (-1082)	Cambodian	pTB	TB = 358; Con = 106	OR = 1.84; p = 0.01	[190]
IL10 (-1082)	Malawian	HIV+ve pTB	TB = 155; Con = 541	OR = 0.37; p = 0.007	[133]
IL10 (-1082A)	Sicilian	pTB	TB = 45; Con = 100	p = 0.05	[171]
IL12RBI (641A/G)	Japanese	TB	TB = 87; Con = 265	OR = 2.53; p = 0.0078	[191]
IL12RBI (1094T/C)	Japanese	TB	TB = 87; Con = 265	OR = 1.59; p = 0.0087	[191]
IL12RBI (1132C/G)	Japanese	TB	TB = 87; Con = 265	OR = 2.83; p = 0.0032	[191]
IL1B	Japanese	TB	TB = 87; Con = 265	p = 0.002	[191]
IL1B (+3953 TT/TC)	Columbian	pTB	TB = 122; Con = 166	OR = 0.3; p = 0.001	[192]
IL1B (-511C)	Gambian	pTB	TB = 335; Con = 298	OR = 0.58; p = 0.015	[193]

Candidate Gene	Population	Phenotype	Sample Size	Reported Results	Ref
IL1RA	Gambian	pTB	TB = 404; Con = 417	OR = 0.46; p = 0.032	[194]
IL4 (-590C/T)	South Indian	pTB	TB = 129; Con = 127	p < 0.01	[195]
IL8 (-251A)	Caucasian	pTB	TB = 106; Con = 107	OR = 3.41; p < 0.006	[196]
IL8 (-251A)	African American	pTB	TB = 180; Con = 167	OR = 3.46; p < 0.01	[196]
MBP (B)	South African	TB Meningitis	TB = 91; Con = 79	p = 0.017	[197]
MBP (B)	African American	pTB	TB = 176; Con = 71	OR = 0.34; p < 0.01	[198]
MBP (B; C)	West African	TB Incidence	626 Individuals	r = 0.565; t = 2.273	[199]
MBP (B; C; D & X)	Danish	pTB	TB = 59; Con = 250	p = 0.03	[200]
MBP (B; C; D)	Indian	pTB	TB = 202; Con = 109	OR = 6.5; p = 0.008	[201]
MBP (C)	Gambian	pTB	TB = 397; Con = 422	OR = 0.79; p = 0.037	[202]
MBP (C)	Malawian	HIV+ve pTB	TB = 154; Con = 546	OR = 1.69; p = 0.034	[133]
NOS2A (-1026)	Brazilian	pTB	92 Pedigrees (627 Ind)	RR = 3.25; p = 0.021	[65]
P2RX7	Gambian	pTB	TB = 646; Con = 694	OR = 0.70; p = 0.003	[203]
P2X7	Southeast Asians	TB	TB = 119; Con = 554	OR = 1.9; p < 0.05	[204]
PGM1	Indian	pTB	N/A	Significant difference	[205]
PGM1 (*2+ allele)	South Indian	pTB	TB = 204; Con = ?	N/A	[206]
SLC11A1 GT(n) & D543N	Japanese	pTB	TB = 202; Con = 267	OR = 2.07; p = 0.0003	[207]
SLC11A1 (TGTG+/del)	Chinese Han	pTB	TB = 147; Con = 145	$\chi^2 = 7.79$; p < 0.01	[208]
SLC11A1 (D543N & TGTG+/del)	Chinese Han	pTB	TB = 110; Con = 180	OR = 1.93; OR = 2.22	[209]
SLC11A1 (CAAA+/del)	Malawian	HIV -ve/+ve pTB	TB = 239/259; Con = 762	OR = 0.65; OR = 0.70	[133]
SLC11A1 (GT(9) & TGTG+/del)	South African	pTB	TB = 265; Con = 224	p = 0.002 & p = 0.013	[210]
SLC11A1 (D543N; TGTG+/del)	Chinese Han	pTB	TB = 120; Con = 240	OR = 2.59; OR = 1.89	[131]
SLC11A1 (INT4 + D543N)	Korean	NTM Lung Disease	TB = 41; Con = 50	OR = 10.88; p = 0.04	[211]
SLC11A1 (INT4 & D543N)	Chinese	Severe TB	TB = 127; Con = 91	OR = 2.29; OR = 2.27	[212]

Candidate Gene	Population	Phenotype	Sample Size	Reported Results	Ref
SLC11A1 (GT(n) & 274C/T)	Houston (USA)	Pediatric TB	184 Nuclear Families	p = 0.04 & OR = 1.75; p = 0.01	[213]
SLC11A1	Chinese	TB	TB = 278; Con = 282	OR = 1.59; p = 0.0163	[214]
SLC11A1 (D2S424)	Canadian Indian	pTB	1 Family (81 Ind)	LOD = 3.81; p = 0.00001	[68]
SLC11A1 (D543N & TGTG+/del)	Cambodian	pTB	TB = 358; Con = 106	OR = 0.59; p = 0.02	[190]
SLC11A1 (INT4 C/C, D543N G/A)	Peruvian	pTB	TB = 507; Con = 513	OR = 1.72 & 1.40; p < 0.05	[215]
SLC11A1 (3' UTR)	Korean	pTB	TB = 192; Con = 192	OR = 1.85; p = 0.02	[216]
SLC11A1 (5' CA)	Tanzanian	TB	TB = 443; Con = 426	OR = 1.45; p = 0.014	[134]
SLC11A1 (D543N)	Japanese	Cavitary Lesion in TB	TB = 95; Con = 90	OR = 5.16	[217]
SLC11A1 (GT(n))	Gambian	pTB	TB = 329; Con = 324	OR = 1.40; p = 0.024	[218]
SLC11A1 (GT(m))	Caucasian US	pTB	TB = 135; Con = 108	OR = 2.02	[219]
SLC11A1 (GT(m); INT4; D543N & TGTG+/d)	Gambian	Smear +ve pTB	TB = 410; Con = 417	OR = 4.07; p < 0.001	[220]
SLC11A1 (INT4)	Guinea-Conakry	pTB	44 Families (160 Ind)	$\chi^2 = 4.14$; p = 0.036	[221]
SLC11A1 (INT4)	Danish	Microscopy +ve TB	TB = 104; Con = 176	RR = 1.9; p = 0.013	[222]
SLC11A1 (INT4)	Chinese exposed to silica dust	pTB	TB = 61; Con = 122	OR = 2.73	[223]
SLC11A1 (region)	Brazilian	pTB	37 Pedigrees (287 Ind)	LOD = 0.51; p = 0.025	[224]
SP110	Republic of Guinea	pTB	99 Families	p < 0.015	[139]
SP110	Guinea-Bissau	pTB	102 Families	p = 0.002	[139]
SP110	Gambian	pTB	219 Families	p < 0.02	[139]
SP-A1	Mexican	pTB	TB = 107; Con = 101	OR = 4.51; p = 0.008	[225]
SP-A1 (307A, 776T)	Ethiopian	pTB	181 Pedigrees (226 TBses)	p < 0.019	[226]
SP-A2	Mexican	pTB	TB = 107; Con = 102	OR = 9.57; p = 0.38	[225]
SP-A2 (A1660G; G1649C)	Indian	pTB	TB = 17; Con = 19	OR = 16.3; p < 0.001	[227]
SP-A2 (355C, 751C)	Ethiopian	pTB	181 Pedigrees (226 TBses)	p < 0.042	[226]

Candidate Gene	Population	Phenotype	Sample Size	Reported Results	Ref
SP-B (B1012_A, AAGG_1)	Mexican	pTB	TB = 107; Con = 103	OR = 2.36; OR = 0.12	[225]
STAT5B (rs2230097)	Brazilian	pTB	92 Pedigrees (627 Ind)	RR = 0.36; p = 0.038	[65]
TIRAP (558)	Vietnamese	TB	TB = 358; Con = 392	OR = 2.25; p = 0.001	[228]
TLR2 (Arg753Gln)	Turkish	pTB	TB = 151; Con = 116	1.60 - 6.04 fold increased risk	[229]
UBE3A	African	pTB	180 Pedigrees	$\chi^2 = 4.17$; p = 0.03	[67]
VDR (FokI; BsmI; ApaI; TaqI)	West African	pTB	382 Trios	$\chi^2 = 22.11$; p = 0.009	[132]
VDR (BsmI; FokI)	Indian	Spinal TB	TB = 64; Con = 103	OR = 2.2; OR = 2.4	[230]
VDR (ff)	Gujarati Asian	Extra-pulmonary TB	TB = 52; Con = 116	OR = 2.8	[130]
VDR (ff)	Chinese Han	pTB	TB = 76; Con = 171	OR = 3.67	[231]
VDR (-ff)	Chinese Han	pTB	TB = 120; Con = 240	OR = 2.35; p = 0.03	[131]
VDR (TaqI; FokI)	Peruvian	TB Treatment	TB = 103; Con = 206	RR = 5.6; RR = 9.6	[137]
VDR (tt)	Gambian	pTB	TB = 408; Con = 414	OR = 0.53; p = 0.01	[128]

^a pTB = pulmonary TB,

^b Con = control

TABLE 5. Publications reporting no significant linkage or association with TB. Adapted from [57] and updated (updated on 15-06-2007).

Candidate Gene	Population	Phenotype	Sample Size	Reported Results	Ref
MHC Region (Class I, II & III)					
Aw30; Aw33; B7; B15; B17	Mexican American	pTB ^a	TB = 100; Con ^b = 100	Ns ^c	[232]
DQB1*0501	Iranian	pTB	TB = 40; Con = 100	ns	[152]
DR2	Northern Indian	pTB	TB = 124; Con = 109	ns (after correction)	[157]
DRB; DQA, DQB	South Indian	pTB	TB = 38; Con = 36; 12 families	No association; No linkage	[233]
HLA	Indian	pTB	21 Families	No Linkage	[234]
HLA	Brazilian	pTB	98 Pedigrees (704 Ind)	ns	[235]
HLA -A; -B; -C	Italian	Current TB	TB = 68; Con = 1089	ns	[148]
HLA-A & -B	Hong Kong Chinese	pTB	TB = 256; Con = 100	ns	[236]
HLA-A, -B, -C	Northern Indian	pTB	TB = 124; Con = 109	ns	[157]
LT	Indian	pTB	TB = 210; Con = 120	ns	[237]
LTA (5' UTR m/sat)	Malawian	HIV -ve/+ve pTB	TB = 198/237; Con = 707	ns	[133]
TAP1	Korean	pTB	TB = 219; Con = 210	ns	[238]
TAP1*0101; TAP2*0101	Columbian	pTB	TB = 122; Con = 166	ns	[192]
TNF	Brazilian	pTB	37 pedigrees (287 Ind)	LOD = 0.01	[224]
TNF	Cambodian	pTB	TB = 126; Con = 88	ns	[156]
TNF -1030; -862; -856; -307	Cambodian	pTB	TB = 358; Con = 106	p > 0.05	[190]
TNF (-238; -308; -376; -893)	Malawian	HIV -ve/+ve pTB	TB ~ 181/144; Con = 417	ns	[133]
TNF (-308)	Japanese	TB	TB = 87; Con = 265	ns	[191]
TNF (-238; -308)	Colombian	TB	TB = 190; Con = 135	ns	[184]
	Indian	pTB	TB = 210; Con = 120	ns	[237]
Other Candidates:					
BTNL2	South African Coloured	TB	TB = 432; Con = 482	ns	[239]
CARD15	Gambian	pTB	TB = 320; Con = 320	ns	[240]
CARD15	South African Coloured	TB	TB = 432; Con = 482	ns	[241]
CCL3 (-906 promoter m/sat)	Malawian	HIV -ve/+ve pTB	TB = 147/206; Con = 580	ns	[133]

Candidate Gene	Population	Phenotype	Sample Size	Reported Results	Ref
CCL3 (-459)	Mexican	pTB	TB = 445; Con = 518	ns	[175]
CCL5 (-471)	Mexican	pTB	TB = 445; Con = 518	ns	[175]
CD14 (-159)	Polish	pTB	Not Given - Abstract Only	ns	[242]
CD14 (-159CT)	Columbian	pTB	TB = 267; Con = 112	ns	[243]
CXCR1	Caucasian US	pTB	TB = 106; Con = 107	ns	[196]
CXCR1	African American	pTB	TB = 180; Con = 167	ns	[196]
CXCR2	Caucasian US	pTB	TB = 106; Con = 107	ns	[196]
CXCR2	African American	pTB	TB = 180; Con = 167	ns	[196]
DC-SIGN	Colombian	TB	TB = 110; Con = 299	p = 0.34	[244]
DC-SIGN and L-SIGN (neck-region)	South African Coloured	TB	TB = 351; Con = 360	ns	[245]
Haptoglobin (HP)	Zimbabwean	pTB	TB = 98; Con = 98	p = 0.5	[246]
ICAM1 (179A/T)	Malawian	HIV -ve/+ve pTB	TB = 209/217; Con = 596	ns	[133]
IFNA17 (551T/G)	Japanese	pTB	TB = 114; Con = 110	p = 0.155	[247]
IFNB (153C/T)	Japanese	pTB	TB = 114; Con = 110	p = 0.137	[247]
IFNG (T874A, G2109A)	Croatian	pTB	TB = 253; Con = 519	ns	[185]
IFNG	Japanese	TB	TB = 87; Con = 265	ns	[191]
IFNG (+874T/A)	African American	TB	TB = 240; Con = 174	ns	[248]
IFNG (+874T/A)	Caucasian	TB	TB = 161; Con = 64	ns	[248]
IFNG (+874T/A)	Hispanic	TB	TB = 319; Con = 98	ns	[248]
IFNG (+874T/A)	South Indian	TB	TB = 129; Con = 127	ns	[195]
IFNG (1348T/A)	Japanese	pTB	TB = 114; Con = 110	p = 0.55	[247]
IFNG (874T/A)	Malawian	HIV -ve/+ve pTB	TB = 213/238; Con = 703	ns	[133]
IFNGR1	Gambian	pTB	TB = 297; Con = 285	ns	[249]
IFNGR1	Japanese	TB	TB = 87; Con = 265	ns	[191]
IFNGR1 (167T/C)	Japanese	pTB	TB = 114; Con = 110	p = 0.213	[247]
IFNGR1 (395)	Iranian	pTB	TB = 50; Con = 54	ns	[250]
IFNGR1 (-611; -56)	African American	pTB	TB = 76; Con = 114	ns	[187]
IFNGR1 (-611; -56)	Caucasian	pTB	TB = 70; Con = 128	ns	[187]
IFNGR2	Japanese	TB	TB = 87; Con = 265	ns	[191]

Candidate Gene	Population	Phenotype	Sample Size	Reported Results	Ref
IFNGR2 (839G/A)	Japanese	pTB	TB = 114, Con = 110	p = 0.498	[247]
IkBL (T738C)	Colombian	pTB	TB = 122; Con = 166	ns	[192]
IL10	Gambian	pTB	TB = 792; Con = 816	ns	[194]
IL10	Caucasian	pTB	TB = 113; Con = 207	ns	[179]
IL12 - 3' UTR	Russian	pTB	TB = 58; Con = 127	ns	[251]
IL12A	Japanese	TB	TB = 87; Con = 265	ns	[191]
IL12B	Japanese	TB	TB = 87; Con = 265	ns	[191]
IL12p40	Caucasian US	pTB	TB = 106; Con = 107	ns	[196]
IL12p40	African American	pTB	TB = 180; Con = 167	ns	[196]
IL12RB1 (+705, +1158, +1196, +1637, +1664)	Koreans	pTB	TB = 115; Con = 151	ns	[252]
IL12RB1 (641A/G)	Japanese	pTB	TB = 114, Con = 110	p = 0.610	[247]
IL12RB2	Japanese	TB	TB = 87; Con = 265	ns	[191]
IL12RB2 (365C/T)	Japanese	pTB	TB = 114, Con = 110	p = 0.59	[247]
IL18R	Japanese	TB	TB = 87; Con = 265	ns	[191]
IL19	Japanese	TB	TB = 87; Con = 265	ns	[191]
IL1beta	Gujarati Asian	pTB	TB = 89; Con = 114	ns	[253]
IL1beta -511; +3953	Cambodian	pTB	TB = 358; Con = 106	p = 0.32 & 0.78	[253]
IL1RA	Gujarati Asian	pTB	TB = 89; Con = 114	ns	[253]
IL1RA	Indian	pTB	TB = 202; Con = 109	ns	[129]
IL1RA	Cambodian	pTB	TB = 358; Con = 106	$\chi^2 = 3.27$; p = 0.19	[190]
IL1RN	Gambian	pTB	TB = 35; Con = 298	ns	[193]
IL23A	Japanese	TB	TB = 87; Con = 265	ns	[191]
IL23R	Japanese	TB	TB = 87; Con = 265	ns	[191]
IL27	Japanese	TB	TB = 87; Con = 265	ns	[191]
IL27RA	Japanese	TB	TB = 87; Con = 265	ns	[191]
IL6 (-174G/C)	Colombian	TB	TB = 190; Con = 135	ns	[184]
IL8 -251 & +781	Gambian	pTB	TB = 284; Con = 245	p = .50; p = .42	[254]
MBL2	Tanzanian	TB	TB = 443; Con = 426	ns	[134]
MBP (B; C; D)	Hispanic	pTB	TB = 198; Con = 46	ns	[198]

Candidate Gene	Population	Phenotype	Sample Size	Reported Results	Ref
MBP (B; C; D)	Caucasian	pTB	TB = 113; Con = 69	ns	[198]
MBP (C; D)	African American	pTB	TB = 176; Con = 71	ns	[198]
MBP (codons 52, 54 & 57)	Polish	pTB	Not Given - Abstract Only	ns	[242]
MMP-1 1G/2G	Japanese	pTB	TB = 105; Con = 106	ns	[255]
NLI-IF	Caucasian US	pTB	TB = 94; Con = 145	p > 0.05	[256]
NOS2A (-954)	Mexican	pTB	TB = 445; Con = 518	ns	[175]
NRAMP1	Thai	TB	TB = 149; Con = 147	ns	[257]
SLC11A1	Russian (Slavic)	pTB	TB = 58; Con = 127	ns	[251]
SLC11A1	Taiwanese	pTB	TB = 49; Con = 48	ns	[258]
SLC11A1	Moroccan	pTB	116 Pedigrees	ns	[259]
SLC11A1	Polish	pTB	TB = 85; Con = 93	ns	[260]
SLC11A1	Thai	TB	TB = 149; Con = 147	ns	[257]
SLC11A1	Japanese	TB	TB = 87; Con = 265	ns	[191]
SLC11A1 (1703G/A)	Japanese	pTB	TB = 114; Con = 110	P = 0.144	[247]
SLC11A1 (ATA(n))	Gambian	pTB	TB = 318; Con = 146	ns	[261]
SLC11A1 (D543N)	Chinese exposed to silica dust	pTB	TB = 61; Con = 122	ns	[223]
SLC11A1 (INT4)	Polish	pTB	Not Given - Abstract Only	ns	[242]
SLC11A2	South African	pTB	TB = 265; Con = 224	ns	[210]
SP110	South African Coloured	TB	TB = 381; Con = 417	ns	[142]
SPP1 (2514C/T)	Japanese	pTB	TB = 114; Con = 110	p = 0.643	[247]
STAT1	Japanese	TB	TB = 87; Con = 265	ns	[191]
TGFbeta1	Colombian	TB	TB = 190; Con = 135	ns	[184]
TGFbeta1 (T869C)	Japanese	pTB	TB = 101; Con = 110	ns	[262]
TLR2 (Arg677Trp; Arg753Gln)	Korean	NTM Lung Disease	TB = 80; Con = 84	ns	[263]
TLR2 (Int 2 m/sat)	Malawian	HIV -ve/+ve pTB	TB = 215/249; Con = 742	ns	[133]
TLR4	Gambian	pTB	TB = 307; Con = 298	$\chi^2 = 0.19$; p = 1.00	[264]
TLR4 (896A/G)	Malawian	HIV -ve/+ve pTB	TB = 162/120; Con = 427	ns	[133]
TNFRSF1B	Japanese	TB	TB = 87; Con = 265	ns	[191]
TNFSF5 (CD40)	West African	pTB	121 Trios	ns	[265]

Candidate Gene	Population	Phenotype	Sample Size	Reported Results	Ref
UBE3A	Japanese	TB	TB = 87; Con = 265	ns	[191]
VDR	Cambodian	pTB	TB = 358; Con = 106	$\chi^2 = 0.99$; p = 0.60	[253]
VDR (FokI, BsmI, ApaI, TaqI)	South African Venda	TB	TB ~ 85; Con ~ 88	ns	[155]
VDR	Japanese	TB	TB = 87; Con = 265	ns	[191]
VDR	Tanzanian	TB	TB = 443; Con = 426	ns	[134]
VDR (TaqI; ApaI; BsmI)	Malawian	HIV -ve/+ve pTB	TB ~212/225; Con = 672	ns	[133]

^a pTB = pulmonary TB

^b Con = control

^c ns = not significant

CHAPTER 2

STUDY OVERVIEW

2.1. CANDIDATE GENE SELECTION

Selection of a candidate gene for a genetic study can be based on two approaches as previously discussed (*Section 1.2.2*). In this project both experimental means and hypotheses were used.

2.2. MAIN HYPOTHESES

1. Interferon gamma (IFN- γ) is the major macrophage-activating cytokine during infection with *Mycobacterium tuberculosis* (*M.tuberculosis*) and its role has been well established in animal models [266,267] and in humans [180,268-270]. Genes regulating the expression of interferon gamma (IFN- γ) may influence resistance or susceptibility to tuberculosis (TB) (*Chapter 4*).

2. TB, sarcoidosis and Crohn's disease (CD) are all granulomatous diseases. Because of this similarity, susceptibility genes in sarcoidosis and CD could also be candidate genes in TB susceptibility (*Chapter 5*).

3. Gene expression studies previously performed by DNA-array analysis experiments [271] in our study population identified genes differentially expressed between TB cases and healthy controls. These genes could determine the outcome of *M.tuberculosis* infection (*Chapter 6*).

4. Genome-wide linkage studies can lead to the discovery of genes that influence disease susceptibility even where a function in disease might not have been suspected. Such a study, as yet unpublished, in the South African Coloured population has led to the identification of a novel susceptibility locus. Genes in this region could therefore influence TB susceptibility (*Chapter 7*).

5. Functional polymorphisms that influence susceptibility to immune-mediated diseases other than TB, such as arthritis, diabetes and Alzheimer's disease, could influence susceptibility to TB (*Chapter 8*).

2.3. AIMS OF THIS STUDY

1. To identify novel polymorphisms in the South African Coloured population.
2. To conduct case-control association studies in the South African Coloured population to test association of genetic markers with TB. These markers can be novel (as derived from resequencing results using DNA from the South African Coloured population) or selected from previous studies.
3. To test the findings from the case-control association studies by transmission disequilibrium test (TDT) analyses using families from the South African Coloured population.
4. To evaluate haplotypes and linkage disequilibrium in the candidate genes.

CHAPTER 3

MATERIALS AND METHODS

3.1. REAGENTS AND EQUIPMENT

The buffers and solutions, reagents, equipment, software and online resources used in this study are listed in the addendum.

3.2. LABORATORY INFORMATION MANAGEMENT SYSTEM

The Laboratory Information Management System (LIMS) used during this project is a public domain database system that allows the effective management of the laboratory workflow and the analysis of single nucleotide polymorphism (SNP) genotyping results [272,273]. Large-scale SNP studies require and generate huge amounts of information, such as patient data, sample data, plate data, assay information, experiments done and genotyping results. It also requires the control of robots. The LIMS of the Institute for Clinical Molecular Biology (IKMB, www.ikmb.uni-kiel.de), used in this study, is composed of a Standard Query Language Server Database and several Visual Basic and web-based tools.

Patient data is entered and edited with the PatientTool program. This data includes the population, pedigrees and disease status of individuals. Each individual receives a unique number. The SampleTool program is used to enter and edit sample data such as the concentration and amount of DNA available. It is also used to design and name 96 and 384-well plates (to track the DNA samples in specific plates). All plates are labelled using a barcode system. The Pipettor program is used to control the robots that sets up the plates. The GenoTool program imports genotyping data into the database, checks the Mendelian inheritance in families and evaluates the quality of the genotyping by testing for Hardy-Weinberg equilibrium. The program can also be used to export data into linkage files [272] for use in analysis programs.

Assay information is stored in the database and comprises the genotyping method used, primers, assay conditions and variants typed. Other programs from the LIMS used in this study are ProjectManager (used to track the experiment process) and AssayManager (used to manage the barcodes of the assay components) [273]. All the information for a specific sample from the different programs in the database is linked, enabling the integration of the data.

3.3. STUDY PARTICIPANTS

Samples for this study were collected in the Western Cape Province of South Africa, where the incidence of tuberculosis (TB) is greater than 500 per 100 000 population per year [274]; chiefly from Ravensmead/Uitsig and surrounding suburbs in the Cape Town metropolitan area. In some of these suburbs the incidence of TB was 1340 per 100 000 population in 1996 [275] and a low prevalence of HIV exists [276].

3.3.1. Study population and stratification

The population studied here, known as the South African Coloured population, has a mixed ancestry, dating back many generations, that includes San, Khoi, Malaysian, African black and European genetic contributions [277]. Since the study population was established through the mixing of racial ethnicities many years ago (even centuries), there was a possibility that population stratification existed which could bias

any association results found. Population stratification results when a case-control sample contains subgroups with allele-frequency differences. These differences might then contribute to any association found and lead to false positive or negative results [75]. To test for population stratification in the South African Coloured population, 25 unlinked SNP markers were genotyped in a previous investigation [176] using the population of the present study. No significant stratification was observed ($p = 0.26$). Using admixed populations, such as the South African Coloured population, in disease association studies has advantages. This group is a present-day homogenous population with a large degree of genetic diversity, since genetic input from several populations contributed to the group. A high number of alleles and genotypes are therefore available to genotype. Some of these polymorphisms may be present in certain populations (for example Europeans and Asians) and absent from other populations (for example African blacks). For this reason, the South African Coloured population offers an unequalled opportunity to evaluate the association of unique markers with disease in a single population [278].

3.3.2. TB case-control sample set

Large case-control sample sets are easier to collect than family-based samples and thought to be the most powerful method to identify modest risk variants in common diseases [75]. The samples from the case-control group were unrelated as reported by the participants. The controls were healthy people older than 15 years with no history of TB who lived in the same high incidence community as the TB patients and were likely therefore to have been exposed to TB. Recent surveys have shown a PPD conversion of $\geq 8\%$ of individuals aged ≥ 5 [279]. Cases consisted of two groups: active TB patients (who had clinical symptoms and bacteriologically confirmed TB) and previous TB patients (who had received treatment and had no clinical symptoms at the time of sampling). All TB cases were human immunodeficiency virus (HIV)-negative. The characteristics of the case-control samples used in this study are listed in Table 6.

TABLE 6. *Characteristics of the family-based and case-control samples.*

Characteristics	Families	Case-control
Diseased individuals (n)	385	431
<i>% male</i>	48	53
<i>Average age (years)</i>	31	34
Healthy individuals (n)	167	483
<i>% male</i>	29	23
<i>Average age (years)</i>	44	27
Total individuals (n)	552	914
Both parents and 2 or more children	30	-
Trios (complete)	54	-
Single parent and 1 or more children	101	-
Total families (n)	185	-

3.3.3. Family-based study sample set

Family-based studies have the advantage that population stratification is circumvented, since there is a common genetic background among the family members. These studies can therefore ensure that an association found with a marker in a case-control study is not due to population stratification. The samples for the family-based study were collected from the same community as the case-control group and were used to confirm associations observed in the case-control study in a smaller and independent sample set [280]. The group consisted of 185 families with 522 individuals. Ideally the amount of families for a transmission disequilibrium test (TDT) should be equal to the number of cases in the case-control cohort, in order to achieve the same power [280], but it is more difficult to collect families than unrelated cases. Both parents and one or more children were sampled in 84 families and single parents and one or more children were genotyped in 101 families. Characteristics for the family-based samples used in this study are listed in Table 6.

3.4. DNA SAMPLES

3.4.1. DNA extraction from blood and quality testing

Blood samples were collected with permission from the Ethics Committee of the Faculty of Health Sciences (Stellenbosch University, South Africa, Project registration number 95/072) and the DNA was purified using the Nucleon BACC3 Kit for blood and cell cultures (Amersham Biosciences, Buckinghamshire, UK).

The concentrations of DNA samples were determined with a NanoDrop[®] ND-1000 Spectrophotometer and the NanoDrop[®] v3.0.1 Software (Inqaba Biotechnology, Pretoria, RSA). All genomic DNA samples were diluted to a final concentration of 30ng/μl with autoclaved distilled water and stored in 0.5 ml tubes. Samples were quality tested in order to determine suitability for whole genome amplification (WGA). All the samples from the family cohort, 64 randomly selected cases and 64 randomly selected controls were tested for quality on agarose gels (Figure 7).

In order to ensure a high quality product after WGA, samples with a clear high molecular weight band and minimum degradation were used. Five μl of each genomic DNA sample at 30 ng/μl, together with 5 μl of 2X loading buffer (0.25% bromophenol blue, 0.25% xylene cyanol FF, 30% glycerol in water) was loaded onto a 0.8% agarose gel (2.4 g of agarose, 300 ml TBE, 2.4 μl of ethidium bromide (10mg/ml, Invitrogen, Karlsruhe, Germany)). Two μl of Hyperladder I (Bioline, Luckenwalde, Germany) was loaded as a molecular weight marker.

The gel was electrophoresed at 110 V for 1 hour, analysed under UV light and photographed with the BioDoc Analyze Video Documentation System (Whatman Biometra, Goettingen, Germany). In the case-control group only 2% of the tested samples were deemed to be unsuitable for WGA, and therefore no further testing was done.

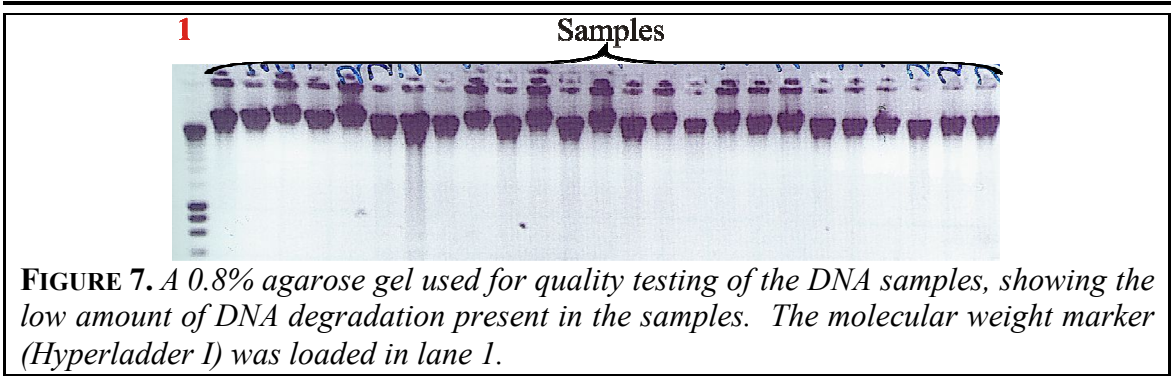


FIGURE 7. A 0.8% agarose gel used for quality testing of the DNA samples, showing the low amount of DNA degradation present in the samples. The molecular weight marker (Hyperladder I) was loaded in lane 1.

3.4.2. Plate design

The 96 well plate layouts were planned using the SampleTool program in the LIMS and contained 92 DNA samples. These designs were used during WGA and TaqMan® genotyping of the samples. Each plate received a unique name. Well C3 always contained a positive control with DNA from the Centre d’Etude du Polymorphisme Humain (CEPH), Paris, France and wells D3, D10 and H12 served as negative controls. For the SNPlex™ genotyping method and some TaqMan® SNP Genotyping Assays, 384 well plates were needed. The 384 well plates were designed with the SampleTool program from the design of the 96 well plates for WGA and each received an unique name. Each 384 well plate layout was prepared by merging four 96 well plates (Figure 8). Sixteen 96 well plates (plate names TBSA01 – TBSA16), and four 384 well plates (plate names XTBSA01-XTBSA04), were designed with the SampleTool program in the LIMS.

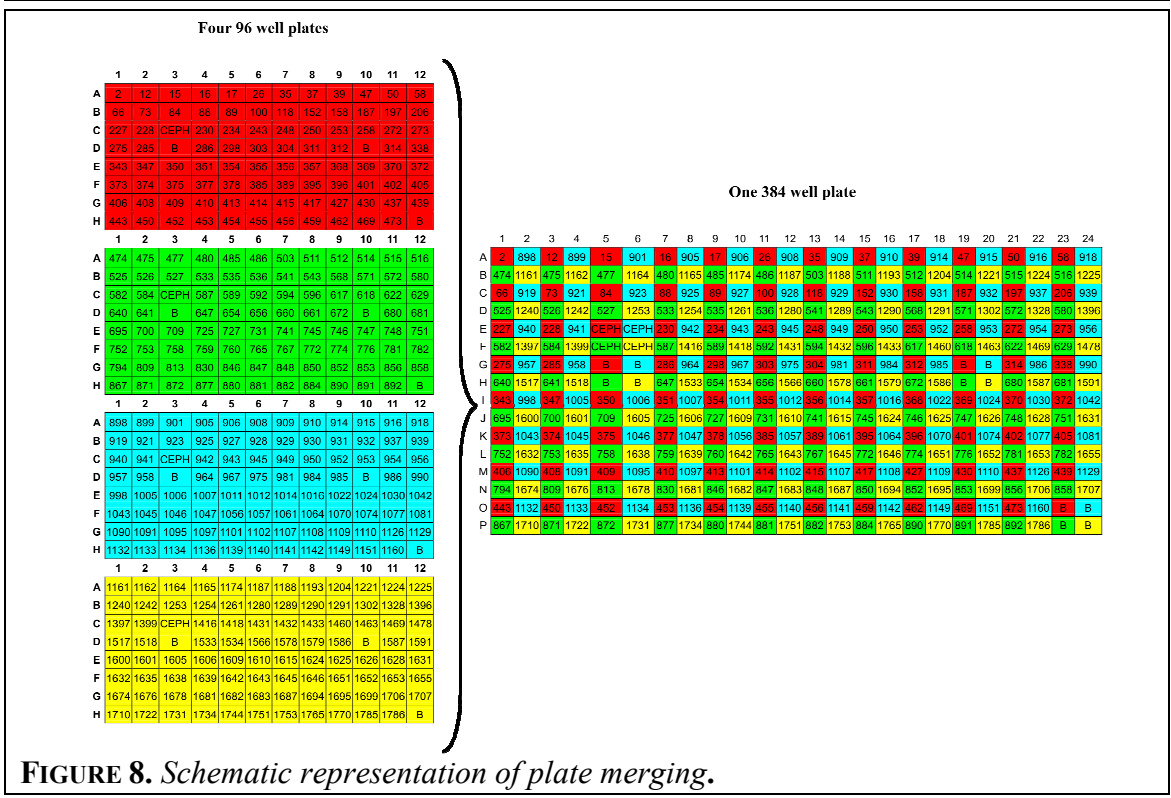


FIGURE 8. Schematic representation of plate merging.

3.4.3. Whole genome amplification (WGA)

The quantity of genomic DNA available for genetic studies is usually limited [281] and acquiring samples for these studies can be difficult [282]. This places restrictions on the type and number of genotyping methods that can be used [283]. The availability of high throughput genotyping methods has increased the demand for DNA, because of the increase in the number of potential genotyping experiments [284]. WGA can generate large amounts of DNA that are suitable for downstream applications [282]. Approaches to WGA can be either PCR-based or non-PCR based. The use of PCR-based methods for WGA has been successful in some studies, but several disadvantages have been identified [283]. These include nonspecific amplification, the overamplification of certain regions and the production of short products [285], which may cause errors during SNP genotyping. Non-PCR based methods are therefore more popular, because they preserve sequence representation, will not cause the preferential amplification of certain alleles and provide excellent coverage of the genome [283].

In this study, multiple displacement amplification (MDA), a non-PCR based method, was used for WGA (Figure 9). This method uses the bacteriophage Phi29 DNA polymerase and random hexamer primers. Phi29 DNA polymerase consists of a single subunit, has proofreading ability, will displace DNA strands and is very stable [286]. The hexamer primers anneal to the denatured genomic DNA at several sites and Phi29 starts replication at these sites. As replication continues, the downstream DNA strands will be displaced by the polymerase and form new single-stranded DNA. The hexamer primers bind to the new DNA, continuing the processes of priming and strand displacement to produce high molecular weight, double-stranded DNA.

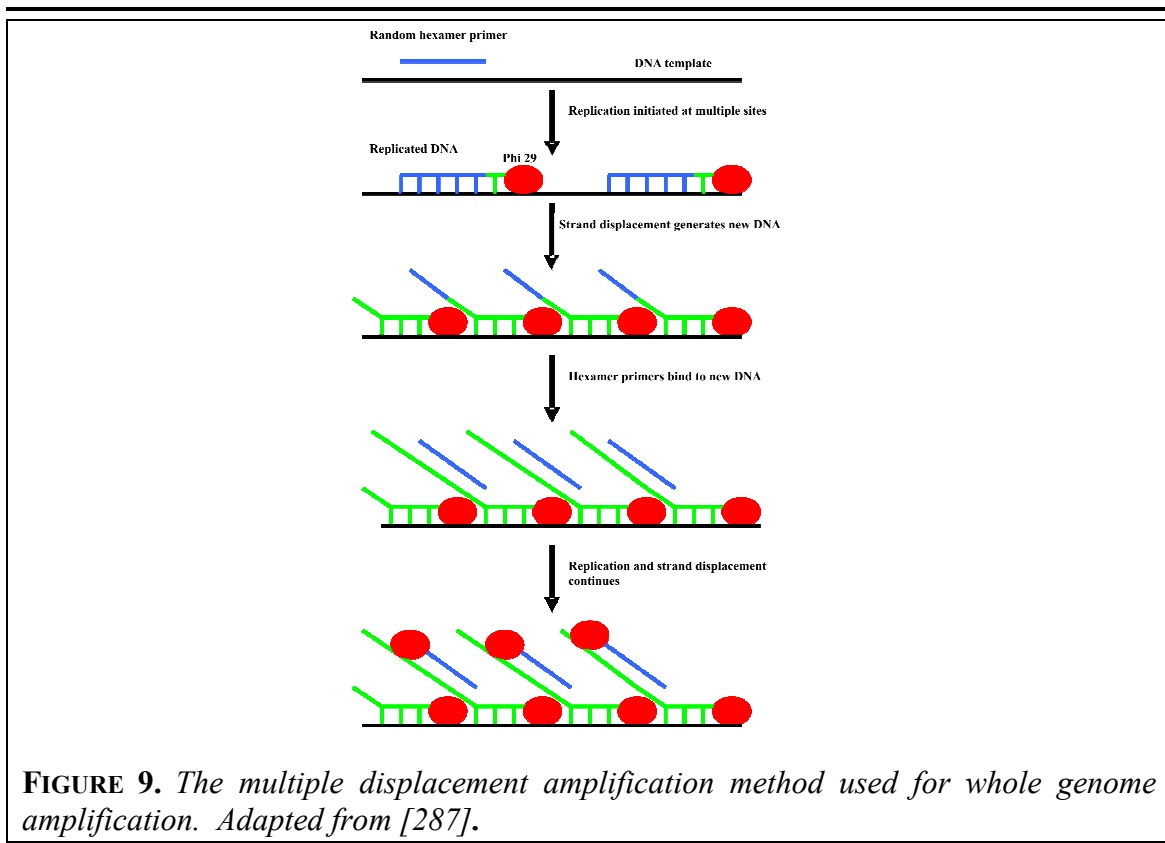


FIGURE 9. *The multiple displacement amplification method used for whole genome amplification. Adapted from [287].*

3.4.3.1. WGA Protocol

Sixteen 96 well plates were used for WGA. The GenomiPhi DNA amplification kit (Amersham Biosciences, Freiburg, Germany) was used for the process after a comparative test done by the host institution (IKMB). This kit can produce 4-7 µg of product DNA from 10 ng of starting DNA.

The workbench was cleaned with a 5% NaClO solution. A 96-well plate was marked with the platenam from the LIMS and the date. The sample buffer from the WGA kit was vortexed and 9 µl was added to each well. One µl of genomic DNA at 30ng/µl was added to the sample buffer. In well C3, 1 µl of CEPH DNA was added to each 96 well plate to serve as positive control. This sample was used to control for the possibility of preferential allele amplification during WGA. No DNA was added to wells D3, D10 and H12 and these blank wells were used to control for contamination during WGA. The plate was sealed with adhesive PCR film (Abgene, Epsom, UK). The template DNA was heat denatured at 95°C for 3 minutes in a GeneAmp PCR system 9700 and cooled to 4°C on ice. For each sample, 9 µl of reaction buffer and 1 µl of enzyme were combined on ice. This enzyme mix was added to the plate, but not to D3, D10 and H12. All the other wells contained a volume of 20 µl. The plate was sealed with an ALPS 300 Platesealer (Abgene, Epsom, UK) and incubated at 30°C overnight (16-18 hours) for the amplification process. After amplification, the enzyme was heat-inactivated at 65°C for 10 minutes in a GeneAmp PCR system 9700 and the plate was cooled to 4°C.

The WGA products were diluted with 1X TE in a 1:10 dilution using the 96-channel Robbins Scientific Hydra system. The syringes of the Hydra were washed with 580 µl of 5% NaClO and rinsed with dH₂O before adding the TE, to remove any contaminants. For the 1:10 dilution, 180 µl of 1X TE was added to the 20 µl WGA product and mixed. The 200 µl was divided into two 100 µl (per well) plates with the Hydra system and these were fragmented for 10 minutes at 95°C in a GeneAmp PCR system 9700 for use in the SNplex™ genotyping system. The plates were marked with a *f* to indicate that fragmentation was completed. The sixteen 96-well plates were merged into four 384-well SNplex plates (as depicted in Figure 8) with the Tecan Genesis RSP 150 robot (Tecan Deutschland GmbH, Crailsheim Germany) for use in the SNplex™ genotyping method (*Section 3.9*). During the merging process, 5 µl of each sample was transferred to the 384 well plates and dried at 60°C.

3.5. MUTATION DETECTION AND POLYMORPHISM SELECTION

In order to determine the frequency of known polymorphisms and the presence of novel variants in some candidate genes, the coding exons and flanking regions of these genes were amplified and sequenced in at least 10 individuals. Half of these individuals were healthy controls and the other half were tuberculosis cases. The reason for selecting the samples for sequencing in this manner was to ensure that possible disease-causing polymorphisms (which may occur at a low frequency in healthy controls) were not missed. Criteria that determined the selection of a SNP or repeat marker for genotyping from sequencing data were as follows: a minor allele frequency of at least 5% in the sequencing data and presence in a functional region of the gene i.e. in an exon, promoter or splice site.

Sequence data was aligned to the expected products with Sequencher v4.2 (Gene Codes Corporation, Michigan, USA). Variations noted in this setup were automatically typed in all the sequenced samples with SGcaller [288]. In order to determine if polymorphisms were known or novel, the sequence containing the variation was aligned to the human genome with Blat (<http://genome.ucsc.edu/cgi-bin/hgBlat>). Blat indicates the presence of a known SNP in a region by listing the rs-number for the SNP. The rs-number of a SNP is a unique record identifier for the Single Nucleotide Polymorphism Database (dbSNP), which was established by the National Human Genome Research Institute and The National Center for Biotechnology Information (NCBI). These records can be found at <http://www.ncbi.nlm.nih.gov/projects/SNP/>. If no rs-number could be found in dbSNP through Blat or in the literature, it was considered to be novel and submitted to the NCBI.

3.5.1. Butyrophilin-like 2 gene

The butyrophilin-like 2 gene (*BTNL2*) has eight coding exons. These were amplified together with the putative promoter region in 48 individuals from the South African Coloured population.

3.5.1.1 PCR primers

Primers were designed as previously described [289]. The names, sequences and the expected product sizes for each primer pair are listed in Table 7.

TABLE 7. *The names, sequences and product sizes of the primers for BTNL2 [289].*

Exon	Primer name	Sequence	Size (bp)
1	BTNL2_e1_F	5'-TGT CCT CTA AGG GCA AAA CAG C-3'	693
	BTNL2_e1_R	5'-GTT TCC CCC TTT CCA AGA AAC TA-3'	
2	BTNL2_e2_F	5'-CAA TGG AAG CCA TGG AGT GTG-3'	489
	BTNL2_e2_R	5'-AAC GTG GTG CAA CCA CTG ATT-3'	
3	BTNL2_e3_F	5'-CAT TTT GAC AGG CTG GAC ACC-3'	511
	BTNL2_e3_R	5'-CCT CAA CTG TCA CAA AGC TTA CCA-3'	
4	BTNL2_e4_F	5'-GTG ACA GCA TTT TTG TTG TGG AG-3'	474
	BTNL2_e4_R	5'-GGT GTC CAT CTG ATG CTC ACT G-3'	
5	BTNL2_e5_F	5'-GGT TTC TAA ACT CCA ATG GAG CTGTT-3'	527
	BTNL2_e5_R	5'-CAA ATG TCA GAG AAA TTG TCC AGG A-3'	
6	BTNL2_e6_F	5'-CCT GGC CGA AGT TAT TTT GT-3'	738
	BTNL2_e6_R	5'-TAG AAT CCC TGG GTG TCC TG-3'	
7	BTNL2_e7_F	5'-AGC AGG CCA ACT GCT TCC T-3'	402
	BTNL2_e7_R	5'-ACT GAG CCT GGA TTG CAT GAT-3'	
8	BTNL2_e8_F	5'-CTT GTC TTC CTG TTT GGC TGT TAA-3'	497
	BTNL2_e8_R	5'-GGG AAG GTG TAG ATA GGG CAC TT-3'	

3.5.1.2. PCR amplification

PCR reactions were carried out in a total volume of 25 μ l. Each reaction contained 10 ng of genomic DNA, 2.5 μ l of 10X enzyme buffer (Applied Biosystems, Darmstadt, Germany), 0.5 μ l of dNTPs (10 mM stock, Amersham Biosciences, Freiburg, Germany), 3 μ l of $MgCl_2$ (25 mM), 0.2 μ l forward and reverse primers (20 μ M stock) and 0.15 μ l of AmpliTaq Gold DNA polymerase (Applied Biosystems, Darmstadt, Germany). A GeneAmp PCR system 9700 was used with the following touchdown cycling program: 10 minutes of denaturation at 95°C followed by 16 cycles of 30 seconds at 95°C, 45 seconds at 66°C (dropping 0.5°C per cycle) and 1 minute at 72°C. This was followed by 20 cycles of 30 seconds at 95°C, 45 seconds at 58°C and 1 minute at 72°C. The reaction was terminated at 72°C for 10 minutes. Control for contamination was done by the inclusion of water blanks in every batch of samples amplified.

3.5.1.3. Gel electrophoresis of *BTNL2* PCR products

To verify the success of each *BTNL2* PCR reaction, agarose gel electrophoresis of the samples was performed. A 2% agarose TBE gel was prepared with 2g of agarose, 100ml of 0.5X TBE buffer and 0.8 μ l of 10mg/ml ethidium bromide (Invitrogen, Karlsruhe, Germany). Five μ l of PCR product together with 5 μ l of 2X loading buffer (0.25% bromophenol blue, 0.25% xylene cyanol FF, 30% glycerol in water) were loaded in the gel. Two μ l of a 100 bp DNA ladder (Invitrogen, Karlsruhe, Germany) was electrophoresed with the samples as a molecular weight marker. The samples were run for 1 hour at 110V and the gel was photographed with a BioDoc Analyze Video documentation system (Whatman Biometra, Germany) using UV light.

3.5.1.4. Purification of *BTNL2* PCR products

PCR products were digested with Exonuclease I (to remove primers, Exo I, Amersham Biosciences, Freiburg, Germany) and Shrimp Alkaline Phosphatase (to remove unincorporated dNTPs, SAP, Amersham Biosciences, Freiburg, Germany). Each reaction was carried out in a total volume of 10 μ l with 0.3 μ l of SAP, 0.15 μ l of Exo I, 1.55 μ l of dH_2O and 8 μ l of PCR product. The reactions were incubated in a GeneAmp PCR system 9700 at 37°C for 15 minutes, followed by 72°C for 15 minutes to deactivate the enzymes.

3.5.1.5. Sequencing of *BTNL2* PCR products

Sequencing reactions were performed with BigDye Terminator v1.1 Cycle Sequencing Kit (Applied Biosystems, Darmstadt, Germany). Each reaction was carried out in a total volume of 10 μ l with 1 μ l of Big Dye v1.1, 1 μ l of primer (3.2 μ M), 0.5 μ l of 5X SB buffer, 5.5 μ l of dH_2O and 2 μ l of purified PCR product. The sequencing cycle was done in a GeneAmp PCR system 9700 with the following cycling program: 1 minute at 96°C followed by 25 cycles of 10 seconds at 96°C, 5 seconds at 55°C and 4 minutes at 60°C. The sequencing products were purified with Sephadex columns and analysed on a 96-capillary 3730xl DNA analyzer (Applied Biosystems, Darmstadt, Germany). To confirm the presence of SNPs, both the forward and reverse strands were sequenced.

3.5.2. Interleukin 12B gene

The microsatellite D5S2941, in intron 2 of the interleukin 12B gene (*IL12B*), was amplified and sequenced to confirm the presence of a novel allele (Section 4.5.1.2.).

3.5.2.1. PCR primers

Primers from a previous study [290] were modified for sequencing purposes (Table 8).

TABLE 8. The names, sequences and product sizes of the primers for *IL12B*.

Region	Primer name	Sequence	Size (bp)
Intron 2	IL12msF	5'-GTT GGC CTC AGT ACG CTT CTT-3'	133
	IL12msRu	5'-TCA CCA GTG GAG ATT TTC ATT C-3'	

3.5.2.2. PCR amplification

PCR reactions were carried out in a total volume of 25 µl. Each reaction contained 100 ng of genomic DNA, 2.5 µl of 10X enzyme buffer with 1.5mM MgCl₂, 2 µl of dNTPs (1.25 mM stock, Bioline, Celtic Diagnostics, South Africa), 1.5 µl forward and reverse primers (10 µM stock) and 0.075 µl of Super-therm GOLD HotStart TAQ DNA polymerase (Southern Cross Biotechnology, South Africa). An Eppendorf Mastercycler (Merck, South Africa) was used with the following cycling program: 10 minutes of denaturation at 95°C followed by 35 cycles of 45 seconds at 94°C, 45 seconds at 60°C and 1 minute at 72°C. The reaction was terminated at 72°C for 10 minutes and cooled at 4°C for 2 minutes. Control for contamination was done by the inclusion of water blanks in every batch of samples amplified. Gel electrophoresis, purification of PCR products and sequencing reactions were done as described in section 3.5.7.

3.5.3. Interleukin 12 receptor, beta 2 gene

A region from the putative promoter of the interleukin 12 receptor, beta 2 (*IL12RB2*) gene was sequenced to determine the presence of the known functional SNP rs3762317.

3.5.3.1. Primer design

Primers were designed from the genomic sequence of the *IL12RB2* gene to amplify a 723 bp product (Table 9).

TABLE 9. The names, sequences and product sizes of the primers for *IL12RB2*.

Region	Primer name	Sequence	Size (bp)
Promoter	IL12RB2proF	5'-CTC AGC AAA CCT GTA CTG GGC-3'	723
	IL12RB2proR	5'-CCC AGC TCT TAC CAC ACC TGA-3'	

3.5.3.2. PCR amplification

PCR reactions were carried out in a total volume of 25 µl. Each reaction contained 100 ng of genomic DNA, 2.5 µl of 10X enzyme buffer with 1.5mM MgCl₂, 2 µl of dNTPs (1.25 mM stock), 1.5 µl IL12RB2proF and IL12RB2proR primers (10 µM stock) and 0.075 µl of Super-therm GOLD HotStart TAQ DNA polymerase.

An Eppendorf Mastercycler was used. The PCR program used was as follows: 10 minutes of denaturation at 95°C followed by 35 cycles of 1 minute at 94°C, 1 minute at 60°C and 2 minutes at 72°C. Control for contamination was done by the inclusion of water blanks in every batch of samples amplified. Gel electrophoresis, purification of PCR products and sequencing reactions were done as described in section 3.5.7.

3.5.4. Signaling lymphocytic activation molecule-associated protein

A region from the putative promoter of the signaling lymphocytic activation molecule-associated protein (*SH2D1A*) gene was sequenced in 10 individuals to determine the frequency of 3 promoter polymorphisms (rs990545, rs7357894, rs12164382) from the dbSNP database in the South African Coloured population.

3.5.4.1. Primer design

Primers were designed from the genomic sequence of the *SH2D1A* gene to amplify a 695 bp product (Table 10).

3.5.4.2. PCR amplification

PCR reactions were carried out in a total volume of 25 µl. Each reaction contained 100 ng of genomic DNA, 2.5 µl of 10X enzyme buffer with 1.5mM MgCl₂, 2 µl of dNTPs (1.25 mM stock), 1.5 µl of SAPproF and SAPproR primers (10 µM stock) and 0.075 µl of Super-therm GOLD HotStart TAQ DNA polymerase. An Eppendorf Mastercycler was used with the following cycling conditions: 10 minutes at 95°C followed by 35 cycles of 1 minute at 94°C, 1 minute at 60°C and 2 minutes at 72°C. The reaction was terminated at 72°C for 10 minutes and cooled at 4°C for 2 minutes. Control for contamination was done by the inclusion of water blanks in every batch of samples amplified. Gel electrophoresis, purification of PCR products and sequencing reactions were done as described in section 3.5.7.

TABLE 10. *The names, sequences and product sizes of the primers for SH2D1A.*

Region	Primer name	Sequence	Size (bp)
Promoter	SAPproF	5'-GAA AGG TGG GCC CTG TAC AC-3'	695
	SAPproR	5'-TAG TAG CTT GGT TCG ATC GAG C-3'	

3.5.5. Testis expressed sequence 264

The five coding exons of the testis expressed sequence 264 (*TEX264*) gene were sequenced in 14 individuals to determine the presence of new SNPs and the frequencies of known SNPs.

3.5.5.1. Primer design

Primers were designed from the genomic sequence of *TEX264* to amplify regions for sequencing (Table 11).

3.5.5.2. PCR amplification

PCR reactions were carried out in a total volume of 25 μ l. Each reaction contained 100 ng of genomic DNA, 2.5 μ l of 10X enzyme buffer with 1.5mM MgCl₂, 2 μ l of dNTPs (1.25 mM stock), 1.5 μ l of forward and reverse primers (10 μ M stock) and 0.075 μ l of

Super-therm GOLD HotStart TAQ DNA polymerase. An Eppendorf Mastercycler was used with the following cycling conditions: 10 minutes at 95°C followed by 30 cycles of 45 seconds at 94°C, 45 seconds at the T_m of the primer and 1 minute at 72°C. The reaction was terminated at 72°C for 10 minutes and cooled at 4°C for 2 minutes. The T_m used for each exon is listed in Table 11. Control for contamination was done by the inclusion of water blanks in every batch of samples amplified. Gel electrophoresis, purification of PCR products and sequencing reactions were done as described in section 3.5.7.

TABLE 11. *The names, sequences and product sizes of the primers for TEX264.*

Exon	Primer name	Sequence	Size (bp)	T _m (°C)
1	TEXEx1.1F	5'-AGG CTC CGT AAC CGA ACC CT-3'	559	64.0
	TEXEx1.1R	5'-GAG GGC GGA AGG ACT ATG GAG-3'		
	TEXEx1.2F	5'-GCT AAA TCA ACC TCA GCA ACC G-3'	316	61.1
	TEXEx1.2R	5'-CCA GTG GGT ATG CCT CTC CT-3'		
2	TEXEx2F	5'-GGC ACA TTG TCA GAG TAG GCT AC-3'	508	60
	TEXEx2R	5'-CAG GCT AAA TCT GGA AGG TGC-3'		
3	TEXEx3F	5'-CCT GAT CTG GCA CAG TCC CA-3'	522	63.3
	TEXEx3R	5'-GGA TAT GCT CCG GCT TCC A-3'		
4	TEXEx4F	5'-GCT CTG TGA GAC CAG TGC CTT-3'	618	62.2
	TEXEx4R	5'-GCA TAG CCT GAC ACC CAC AGT-3'		
5	TEXEx5.1F	5'-AGT GCA CTG CGT GCT GAG TT-3'	392	62.2
	TEXEx5.1R	5'-GTG CAG GCC ATG GGT TAC TC-3'		
	TEXEx5.2F	5'-AGC ACA GCT ACA GCG AGT CAG-3'	541	61.1
	TEXEx5.2R	5'-CCA GCC TAT CTG TAC AGC CAT C-3'		

3.5.6. Toll-like receptor 2

The GT repeat in intron 2 of the toll-like receptor 2 (*TLR2*) gene was sequenced in 14 individuals to determine the spectrum of repeats present in the South African Coloured population.

3.5.6.1. Primer design

The forward primer was designed according to a previous study [291], but was modified for sequencing. The reverse primer was designed from the genomic sequence of *TLR2* (Table 12).

3.5.6.2. PCR amplification

PCR reactions were carried out in a total volume of 25 µl (100 ng of genomic DNA, 2.5 µl of 10X enzyme buffer with 1.5mM MgCl₂, 2 µl of dNTPs (1.25 mM stock), 1.5 µl TLR2msF and TLR2msRu primers (10 µM stock) and 0.075 µl of Super-therm GOLD HotStart TAQ DNA polymerase amplified with the following cycling program: 10 minutes at 95°C followed by 35 cycles of 1 minute at 94°C, 1 minute at 62°C and 1 minute at 72°C in an Eppendorf Mastercycler. The reaction was terminated at 72°C for 10 minutes and cooled at 4°C for 2 minutes). Control for contamination was done by the inclusion of water blanks in every batch of samples amplified. Gel electrophoresis, purification of PCR products and sequencing reactions were done as described in section 3.5.7.

TABLE 12. The names, sequences and product sizes of the primers for *TLR2*.

Region	Primer name	Sequence	Product size
<i>TLR2</i>	TLR2msF	5'-GCA TTG CTG AAT GTA TCA GGG A -3'	292 bp
	TLR2msRu	5'-GAT TCT TCC TTG GAG AGG CTG-3'	

3.5.7. Gel electrophoresis, purification of PCR products and sequencing reactions

All PCR products from *IL12B*, *IL12RB2*, *SH2D1A*, *TEX264* and *TLR2* were electrophoresed, purified and sequenced using the following protocol:

A 2% agarose sodium boride (SB) gel was prepared with 2g of agarose, 100ml of 0.5X SB (*Addendum A.1*) buffer and 5 µl of 10mg/ml ethidium bromide (Sigma-Aldrich, South Africa). Five µl of PCR product together with 5 µl of 1X Orange Loading Dye Solution (Fermentas, Inqaba Biotechnology, South Africa) were loaded on the gel. Five µl of a 100 bp DNA ladder (Promega, Whitehead Scientific, South Africa) was electrophoresed with the samples as a molecular weight marker. The samples were run for 30 minutes at 150V and the gel was photographed with a Syngene G box (Vacutec, South Africa) using UV light.

Purification of the PCR products prior to sequencing was done with ExoSAPit (AEC Amersham, South Africa). Each reaction was carried out in a total volume of 19.2 µl with 7.2ul of enzyme and 12 µl of PCR product. The reactions were incubated in an Eppendorf Mastercycler at 37°C for 15 minutes followed by 72°C for 15 minutes to deactivate the enzymes.

Sequencing reactions were performed with BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, South Africa). Each reaction was carried out with 1.5 µl of Big Dye v3.1, 0.5 µl of DMSO, 5 µl of the primers used for the PCR (1.1 µM), and 5 µl

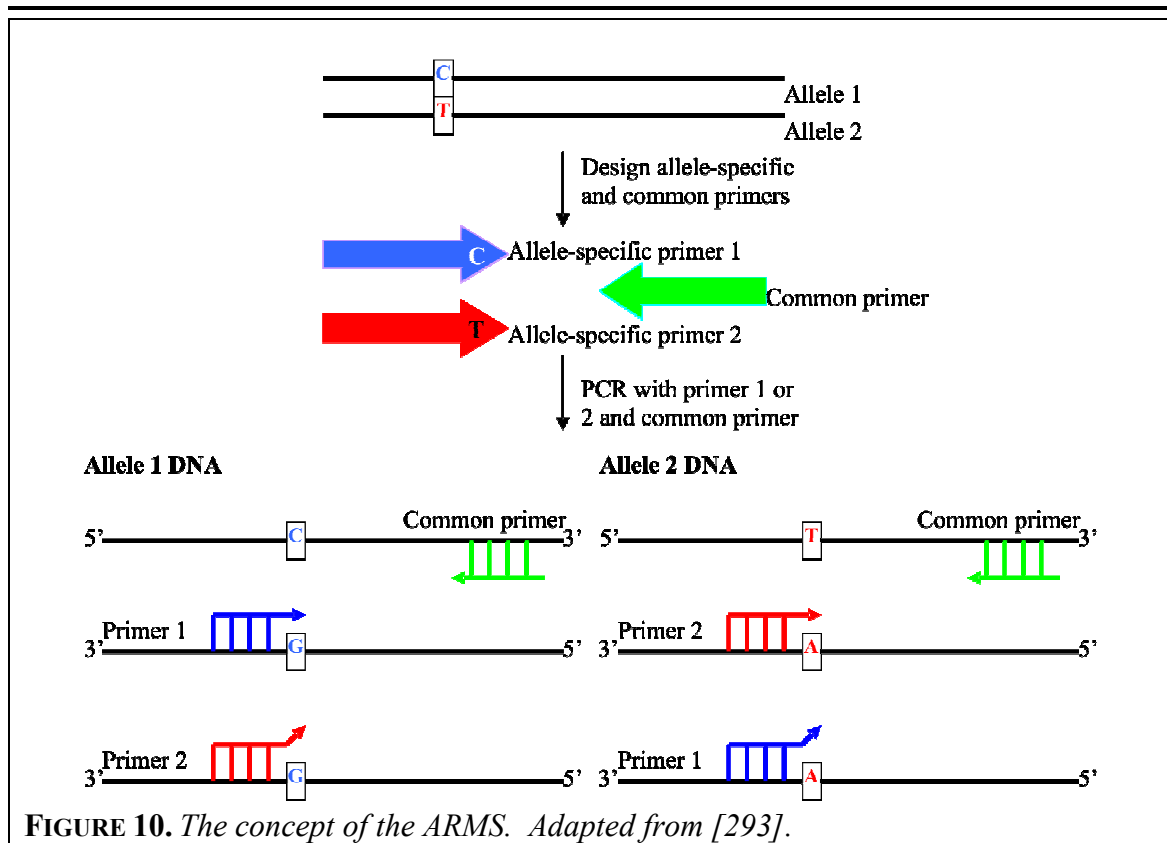
of purified PCR product. A GeneAmp PCR system 9700 was used with the following cycling program: 5 minutes at 95°C followed by 35 cycles of 45 seconds at 95°C and 4 minutes at 58°C. The sequencing products were purified with Centri-Sep 96 well plates (Princeton Separations, Adelphia, New Jersey) and analysed on a 16-capillary 3130xl DNA analyzer (Applied Biosystems, South Africa). To confirm the presence of SNPs, both the forward and reverse strands were sequenced with the respective forward and reverse primers.

3.5.8. Selection of remaining polymorphisms

Remaining polymorphisms were selected from the literature and the Hapmap project (Section 1.2.2., <http://www.hapmap.org/>). The markers chosen were either previously found to be functional in mycobacterial or related diseases or had a minor allele frequency greater than 5%.

3.6. AMPLIFICATION REFRACTION MUTATION SYSTEM-POLYMERASE CHAIN REACTION

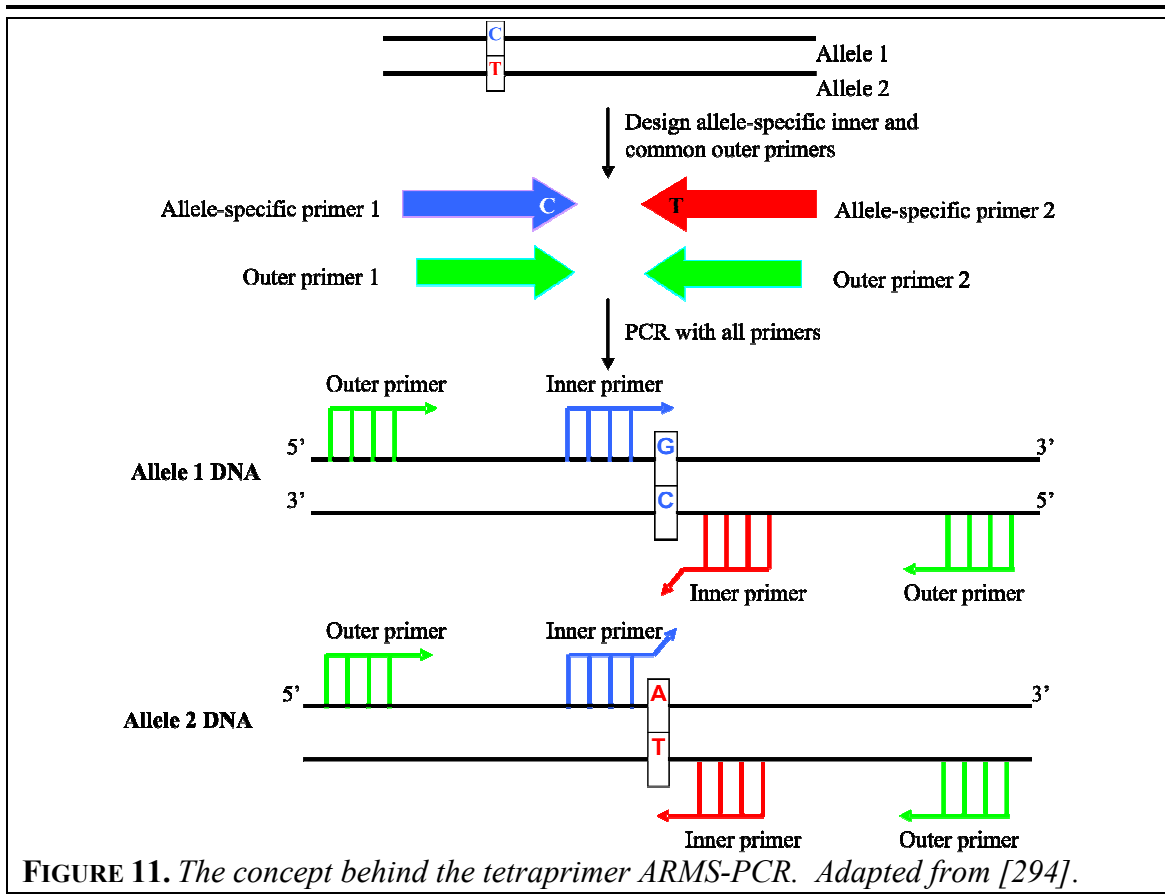
The amplification refraction mutation system-polymerase chain reaction (ARMS-PCR, Figure 10) method allows the genotyping of samples by PCR and agarose gel electrophoresis only. It is a simple and reliable system which makes it possible to distinguish between heterozygotes and homozygotes at a certain allele.



Two PCR reactions are carried out for each sample. The one primer will be the same in both reactions, while the second primer will be unique to each reaction. The only difference between these unique primers will be that their 3' ends correspond to one of

the two alleles of the polymorphism of interest. In theory these allele-specific primers will anneal only to DNA with a complementary base at that allele. To enhance allele-specificity a deliberate mismatch at position -3 bases from the 3' end is incorporated into these primers. The presence of a DNA fragment on an agarose gel will indicate which base is present at the locus. Control primers are included to amplify another sequence in the genome and are used to check that the PCR reaction has worked [292].

An adaptation of the ARMS method, the tetraprimer ARMS-PCR [294] (Figure 11), simplifies this technique even further by reducing the number of PCR reactions per sample to one. This minimises costs and is time-saving. Two sets of allele-specific primers (four primers) are designed for each reaction. One allele-specific (or inner) primer is designed in the sense direction, while the other allele-specific primer is designed in the anti-sense direction. The counterparts (or outer primers) are designed to amplify different sized allele-specific PCR products when combined with the inner primers. The PCR product sizes must be different enough to be distinguished by gel electrophoresis. The PCR product from the outer primers also serves as an internal control to verify the success of the PCR reaction. When using this technique, the shorter and allele-specific products will have a lower amplification efficiency than the largest product and non-specific amplification of the shorter products might take place [295]. These problems are resolved by lowering the concentration of the outer primers and by adjusting the PCR program [294].



3.6.1. ARMS protocol

The ARMS technique was used to genotype the rs3762317 SNP in the promoter of the *IL12RB2* gene. Allele-specific and common primers were designed as previously described [296] and a mismatch was incorporated three bases from the 3' end of the allele-specific primers. SAPproF and SAPproR (Table 13) were used as internal control primers to amplify a 695 bp product from the promoter in *SH2D1A*. The allele-specific primers IL12-G and IL12-A together with the common IL12-R primer amplified the region surrounding the SNP of interest.

PCR reactions for the ARMS-PCR were carried out in a total volume of 25 µl. Each reaction contained 60 ng of genomic DNA, 2.5 µl of 10X enzyme buffer with 1.5mM MgCl₂, 2 µl of dNTPs (1.25 mM stock), 0.3 µl of SAPproF and SAPproR primers (the control primers, 10 µM stock), 1.5 µl of the common primer IL12-R, 1.5 µl of the allele-specific IL12-G or IL12-A primers (10 µM stock) and 0.075 µl of Super-therm GOLD HotStart TAQ DNA polymerase. An Eppendorf Mastercycler and touchdown PCR program was used as follows: 10 minutes of denaturation at 95°C followed by 10 cycles of 45 seconds at 94°C, 1 minute at 70.3°C (dropping 1°C per cycle) and 1 minute at 72°C. This was followed by 35 cycles of 45 seconds at 94°C, 45 seconds at 60.3°C and 1 minute at 72°C. The reaction was terminated at 72°C for 10 minutes and cooled at 4°C for 2 minutes. Control for contamination was done by the inclusion of water blanks in every batch of samples amplified. To verify the success of each PCR reaction agarose gel electrophoresis of the samples was performed on a 2% SB gel as described in section 3.5.7.

TABLE 13. Primers used for genotyping the rs3762317 SNP in *IL12RB2*.

Primer name	Sequence	Product size
SAPproF	5'-GAA AGG TGG GCC CTG TAC AC-3'	695 bp
SAPproR	5'-TAG TAG CTT GGT TCG ATC GAG C-3'	
IL12-G	5'-ACT AGC GCA TAA ATA ATG GTG TG-3'	494bp
IL12-A	5'-ACT AGC GCA TAA ATA ATG GTG TA-3'	
IL12-R	5'-CTC TGC ACC GCC CAG CTC TTA-3'	

3.6.2. Tetraprimer ARMS-PCR protocol

The tetraprimer ARMS-PCR technique was used to genotype a promoter SNP, G-631A (rs990545), in the *SH2D1A* gene. Outer and inner primers (Table 14) for the reaction were designed from the genomic sequence of the gene. The sizes of the resulting PCR products were selected in such a way that the differences in size could be easily visualised on an agarose gel. Homozygotes would have the 695 bp product and either a 450 or 296 bp product, while heterozygotes would have all three products.

PCR reactions for the tetraprimer ARMS-PCR were carried out in a total volume of 25 µl. Each reaction contained 100 ng of genomic DNA, 2.5 µl of 10X enzyme buffer with 1.5mM MgCl₂, 2 µl of dNTPs (1.25 mM stock), 0.5 µl of SAPproF and SAPproR primers (the outer primers, 10 µM stock), 5 µl of SAP-631G and SAP-631A primers

(the inner primers, 10 μ M stock) and 0.075 μ l of Super-therm GOLD HotStart TAQ DNA polymerase.

An Eppendorf Mastercycler and touchdown PCR program was used as follows: 10 minutes of denaturation at 95°C followed by 12 cycles of 45 seconds at 95°C, 1 minute at 72°C (dropping 1°C per cycle) and 1 minute at 72°C. This was followed by 22 cycles of 45 seconds at 95°C, 1 minute at 60°C and 1 minute at 72°C. The reaction was terminated at 72°C for 10 minutes and cooled at 4°C. Control for contamination was done by the inclusion of water blanks in every batch of samples amplified. Gel electrophoresis was done as described in section 3.5.7.

TABLE 14. Primers used to genotype the rs990545 SNP in *SH2D1A*.

Primer name	Sequence	Size (bp)
SAPproF	5'-GAA AGG TGG GCC CTG TAC AC-3'	695
SAPproR	5'-TAG TAG CTT GGT TCG ATC GAG C-3'	
SAP-631G	5'-GGA GAA GAA GAA CTT GCC TAG-3'	450
SAP-631A	5'-AAT CAG ACA TAG TAG ATG AAC AGT GAA TTT-3'	296

3.7. CAPILLARY ELECTROPHORESIS OF FLUORESCENTLY LABELLED PCR PRODUCTS

Roughly 94-95% of the human genome is made up of repeated sequences. Eukaryotic and prokaryotic genomes contain many repetitive DNA sequences and these are present in all sizes [297]. Short tandem repeats (STRs), or microsatellites, are defined as tandemly repeated regions of DNA that consist of 1-6 base pair long motifs [298]. These markers, that can be polymorphic, are primarily used for genetic mapping [299] and have been reported to be associated with several diseases, such as a (CA)_n repeat in the interferon- γ gene that was correlated with the production of interferon- γ and associated with pulmonary TB [300,301]. However, the manner by which these STRs influence disease are still not clear [302].

STRs can be genotyped using capillary electrophoresis and fluorescently labelled PCR products (Figure 12). The fluorescent dye is linked to a PCR primer and this is incorporated into the amplified region during PCR. Multiple regions can be simultaneously amplified using a multiplex PCR and uniquely labelled primers that amplify differently sized regions. These products are separated according to size using capillary electrophoresis and the results are represented by peaks on an electropherogram.

3.7.1. Capillary electrophoresis protocol

The *TLR2* (GT)_n and *IL12B* (ATT)_n (D5S2941) microsatellites were amplified using a multiplex PCR and fluorescently labelled primers (Table 15). Primers were designed by modifying previously described primers [290,291] or from the genomic sequence of the genes. Each reaction was done in a total volume of 10 μ l and contained 1 μ l of WGA-DNA, 1 μ l of 10X enzyme buffer with 1.5mM MgCl₂, 0.8 μ l of dNTPs (1.25 mM

stock), 0.2 µl of each of the TLR2msF, TLR2msRl, IL12msF and IL12msRl primers and 0.03 µl of Super-therm GOLD HotStart TAQ DNA polymerase.

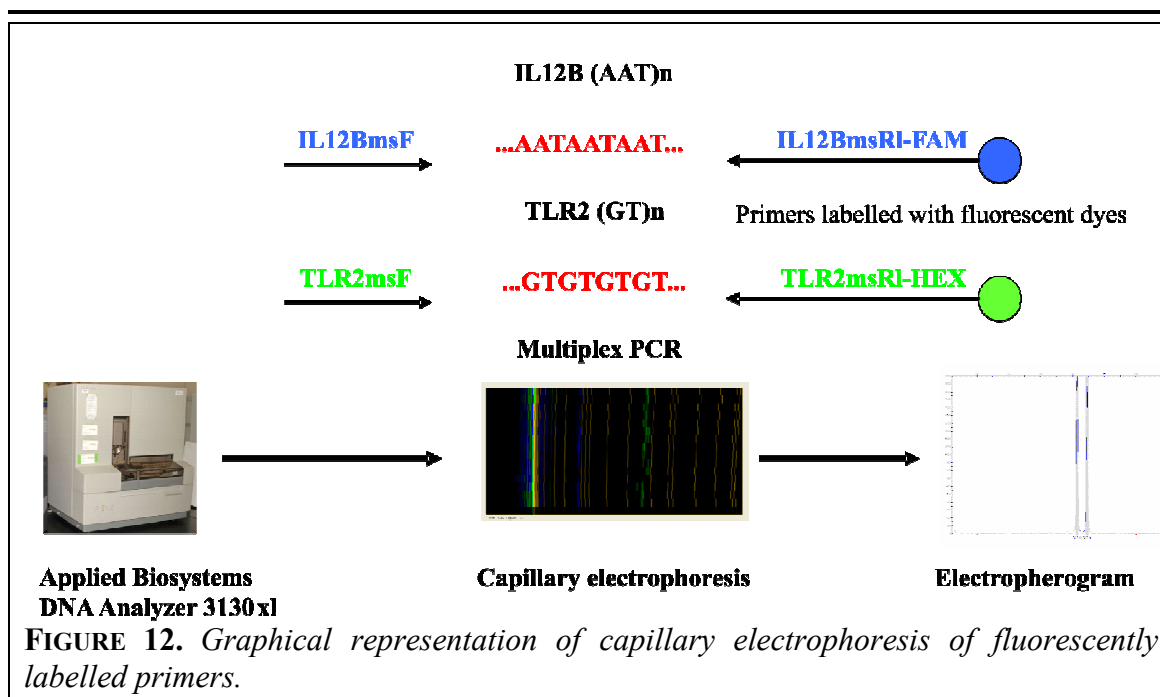


TABLE 15. Primers used during the multiplex PCR prior to capillary electrophoresis.

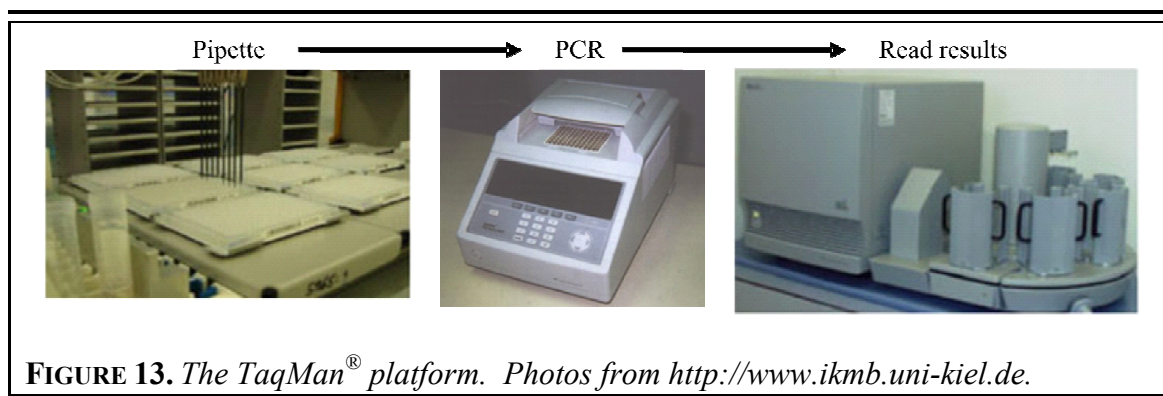
Region	Primer name	Sequence	Size (bp)
TLR2	TLR2msF	5'-GCA TTG CTG AAT GTA TCA GGG A -3'	292
	TLR2msRl	5' HEX-GAT TCT TCC TTG GAG AGG CTG-3'	
IL12B	IL12msF	5'-GTT GGC CTC AGT ACG CTT CTT-3'	133
	IL12msRl	5' FAM-TCA CCA GTG GAG ATT TTC ATT C-3'	

An Eppendorf Mastercycler was used for amplification and the following cycling conditions were used: 10 minutes at 95°C followed by 25 cycles of 1 minute at 94°C, 1 minute at 60°C and 1.5 minutes at 72°C. The reaction was terminated with 10 minutes at 72°C and cooled at 4°C for 2 minutes.

After amplification, samples were prepared for capillary electrophoresis. One µl of fluorescently labelled PCR product, 10 µl of HiDi™ Formamide (Applied Biosystems, South Africa) and 0.5 µl of the Genescan™ – 500 LIZ™ size standard (Applied Biosystems, South Africa) were mixed. The samples were denatured for 5 minutes at 94°C in a GeneAmp PCR System 9700 and placed on ice to preserve the single-stranded DNA conformation. Samples were electrophoresed in a 36 cm capillary array using the 3130 Performance Optimized Polymer 7 (Applied Biosystems, South Africa) on a 3130xl DNA analyzer. Following data collection, samples were analysed with the Genemapper version 3.7 software program (Applied Biosystems, South Africa).

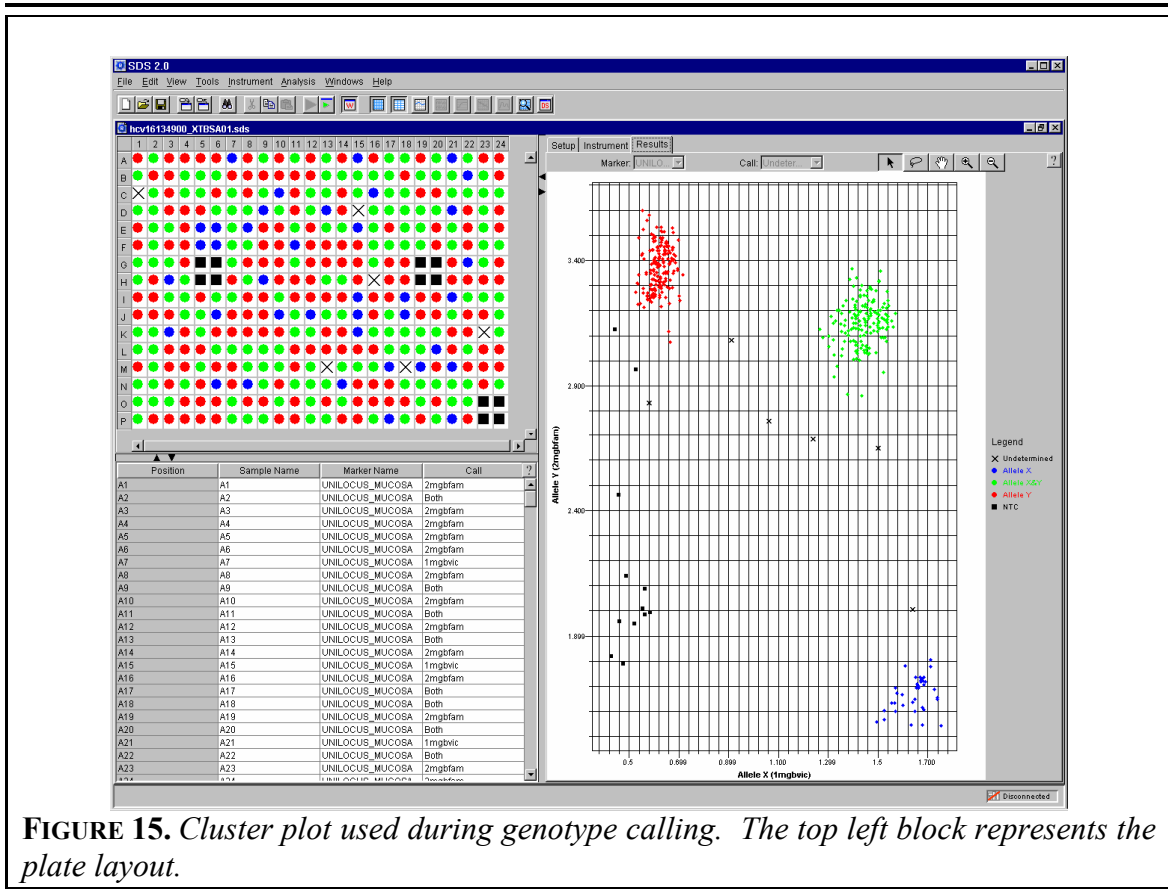
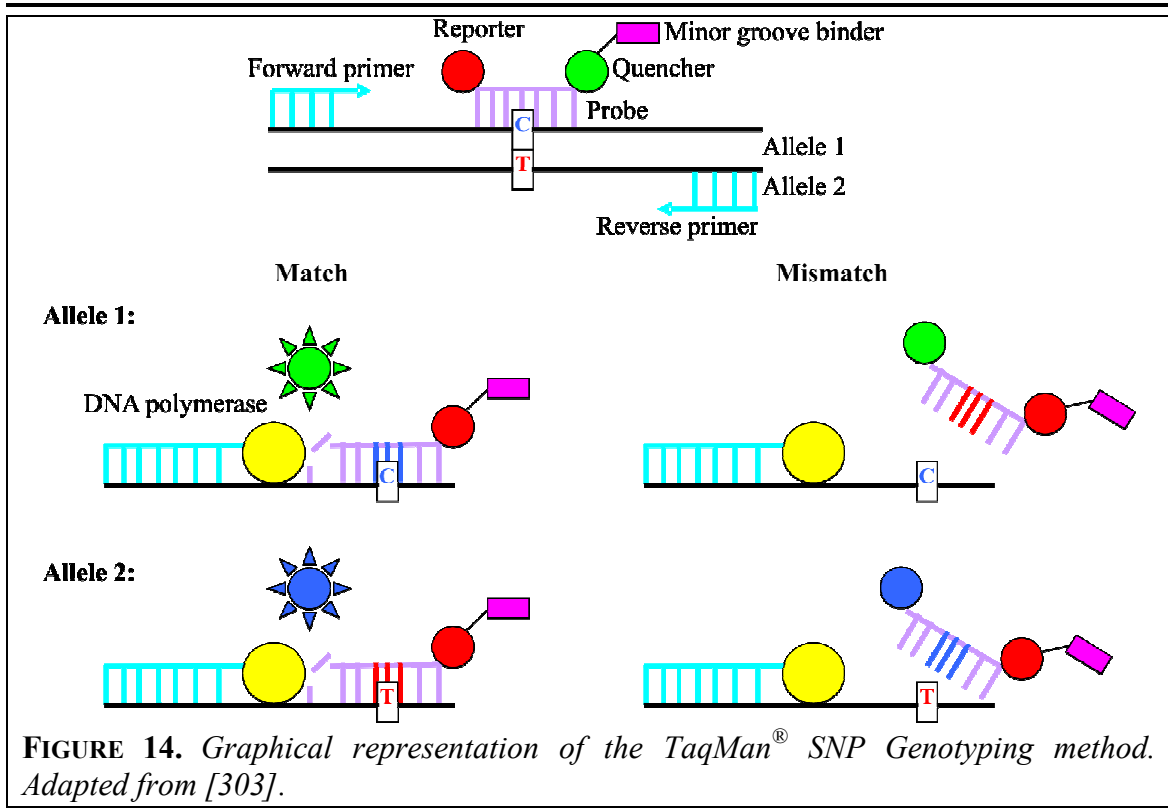
3.8. TAQMAN[®] GENOTYPING

TaqMan[®] SNP Genotyping Assays [303], done in a single-step PCR reaction (Figure 13), use the 5' to 3' exonuclease activity of Taq DNA polymerase [304]. Each assay contains 2 locus-specific PCR primers that surround the SNP of interest and two TaqMan[®] probes (Figure 14). Every probe consists of an allele-specific oligonucleotide labelled with a fluorescent reporter dye at the 5' end. The 3' end of the probe is unable to extend [304] and is labelled with a non-fluorescent quencher and a minor groove binder. The minor groove binder stabilises the double-stranded probe structure and therefore increases the melting temperature (T_m) of the probe without increasing the length of the probe. This enhances allelic discrimination. The quencher diminishes the fluorescence emitted by the fluorescent dye [303].



The two probes are labelled with different fluorescent reporter dyes to discriminate between the two alleles of a SNP in a single reaction [305]. During the PCR, the probes and primers are allowed to hybridise to the target DNA (Figure 14). If no hybridisation occurs, the probe will remain intact and will not emit a detectable fluorescent signal when excited, due to the presence of the quencher molecule. If hybridisation does occur, the 5' nuclease activity during each cycle of the PCR reaction cleaves the probe thus leading to the release of the reporter dye from the quencher and a fluorescent signal can be detected [306].

The results of a SNP TaqMan[®] Genotyping Assay are read at the PCR endpoint, as opposed to TaqMan[®] assays designed for gene expression which are read in realtime [303]. Genotype calls are done by plotting the normalised intensity of the fluorescent reporter dyes on a cluster plot (Figure 15) using the data analysis software. For a diallelic SNP, a maximum of three clusters would be expected, with two of these representing the homozygotes of each allele and the third representing heterozygotes. An additional cluster in the vicinity of the cross of the x- and y-axis would represent the control samples without any DNA. Several pre-designed and validated assays [307] are available from Applied Biosystems and these were used during this project. The TaqMan[®] Genotyping Assays offer the following advantages: a single enzymatic reaction is required followed by PCR and endpoint reading using the closed tube, universal reactions and PCR conditions are used, the location of the primers and probes around the SNP of interest is flexible and insertion or deletion polymorphisms can be typed [303].



3.8.1. TaqMan[®] protocol

TaqMan[®] Genotyping Assays were used to genotype the following SNPs: rs2857654 from *CCL2*, rs2243250 from *IL4* and rs557077 from *WNT5A* (Table 16). R702W (rs2066844), G908R (rs2066845) and 1007fs (rs2066847) from *CARD15* were genotyped using the TaqMan assays and protocols of Hampe et al [308,309]. Both 96-well and 384-well plates were used for genotyping SNPs with the TaqMan[®] Genotyping Assays. Pipetting was done with the Tecan Genesis Workstation 150 (Tecan Deutschland GmbH, Crailsheim, Germany). For each 5 µl reaction, 2.5 µl of TaqMan[®] Universal PCR Master Mix (Applied Biosystems, Darmstadt, Germany), 0.25 µl of the 20X TaqMan[®] SNP Genotyping Assay Mix (Applied Biosystems, Darmstadt, Germany) and 2.5 ng of dried DNA were used. PCR reactions were completed in either a Biometra T1 Thermocycler (Whatman Biometra, Goettingen, Germany) or in a GeneAmp PCR system 9700 with the following cycling program: 10 minutes at 95°C followed by 40 cycles of 15 seconds at 92°C and 1 minute at 60°C. The plates were read in the 7900 HT Sequence Detection System (Applied Biosystems, Darmstadt, Germany) and the fluorescence analysed with the SDS 2.0 software.

TABLE 16. *TaqMan[®] Genotyping Assays used during this project.*

Gene	SNP	Applied Biosystems Assay ID
<i>CCL2</i>	rs2857654	C_16134900_10
<i>IL4</i>	rs2243250	C_16176216_10
<i>WNT5A</i>	rs557077	C_3214537_10

3.9. SNPLEX[™] GENOTYPING SYSTEM

The SNplex[™] Genotyping System (Applied Biosystems, Darmstadt, Germany) allows high-throughput genotyping and up to 48 SNPs can be genotyped in one reaction [303] (Figure 16). The method relies on the oligonucleotide ligation assay (OLA) and multiplex PCR to discriminate between alleles and to amplify the target [310]. The products are hybridised with universal fluorescently labeled Zipchute probes and separated and detected with Applied Biosystems 3130xl or 3730xl DNA analysers [303]. The traditional OLA makes use of two adjacent oligonucleotides. The 5' probe is biotinylated and the 3' probe is the reporter sequence. The two probes will anneal to the target DNA sequence and the DNA ligase will covalently bind it if there is perfect complementarity. The 5' probe is captured on streptavidin and analysed for the 3' reporter [311].

In the SNplex Genotyping System, three unlabeled oligonucleotides are used to detect a SNP during OLA. Two of these are allele-specific oligonucleotides (ASOs) and will correspond to the 2 alleles of the SNP. They contain unique ZipCode sequences. Each of the ASOs have a linker that contains an internal spacer, a partial Zipcode sequence, the forward primer site and a hairpin structure that is complimentary to the Zipcode sequence of its ASO. In a 48-plex reaction there are 96 ASO linkers. The third oligonucleotide is the locus-specific oligonucleotide (LSO), which contains the reverse primer site. The LSO linker is universal for the reaction, i.e. there is only one LSO linker in a 48-plex [310].

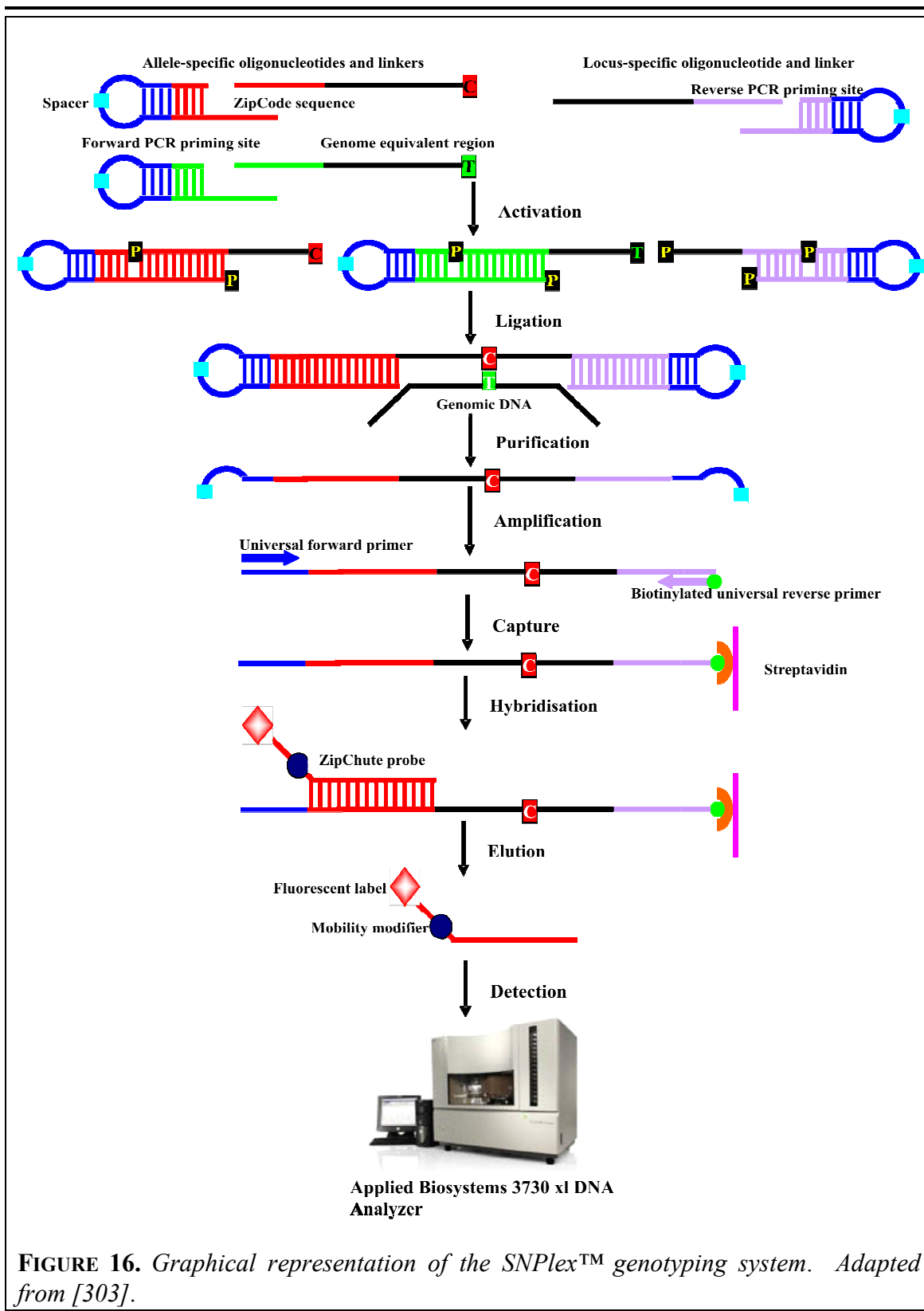
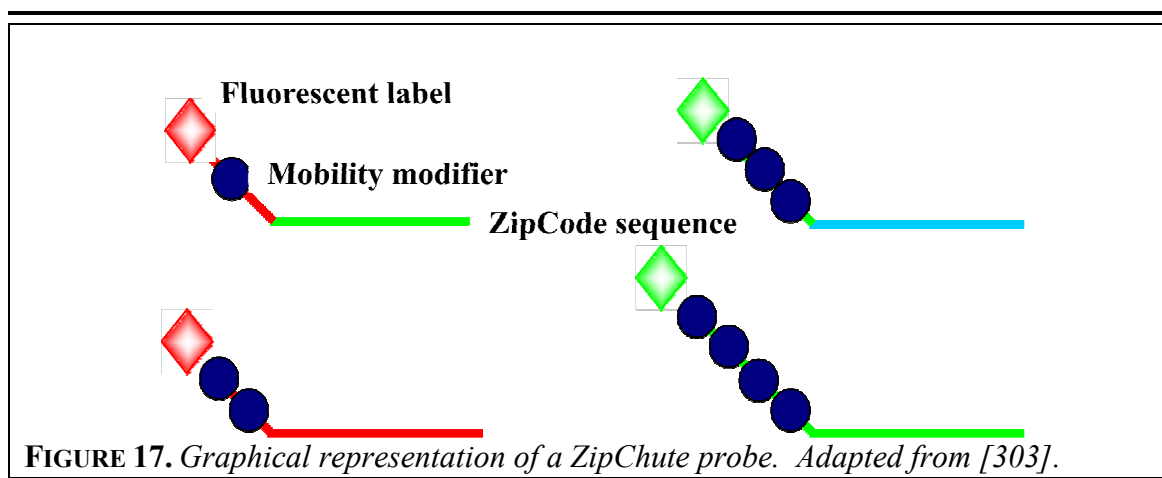


FIGURE 16. Graphical representation of the SNPLex™ genotyping system. Adapted from [303].

The linkers and probes are activated by phosphorylation before the OLA step. During the SNplex OLA process, the ASO that corresponds to the base present at the SNP will anneal to the DNA together with the LSO. The ASOs and LSO at a particular SNP will be ligated to their specific linkers and to each other. After ligation, the unused probes, linkers and genomic DNA are removed by digestion with the exonuclease I and λ -exonuclease enzymes [303]. The products are diluted before the multiplex PCR reaction with universal primers. One of the PCR primers is labelled with biotin and after PCR this will lead to biotinylated amplification products. The labelled amplicons are captured on streptavidin-coated microtiter plates and non-biotinylated amplicons are removed [310].

ZipChute probes (Figure 17) are used to report the genotypes during SNplex genotyping. Each probe corresponds to one of the unique ZipCode sequences of the ASOs, which is incorporated into the amplification product. The ZipChute probes also contain varying numbers of mobility modifiers, so that size separation of the different probes is possible during capillary electrophoresis. The fluorescent label is detected by the 3730x1 DNA analyser. Each SNP has two ZipChute probes, one for each allele of the SNP and each probe pair is labelled with the same fluorescent marker. Both the size and the colour of the ZipChute probes are therefore used to discriminate between SNPs and their alleles [303]. The universal ZipChute probes are added to the wells of the microtiter plates where the biotinylated amplicons are captured. After hybridisation and washing steps, the ZipChute probes are released from the amplification products by incubating the plate at 37°C. The plates are loaded onto the 3730x1 DNA analyser.



The resulting sample files are analysed with the GeneMapper Analysis Software v3.5.1 [310]. This software can be used for automatic allele calling of the SNPs genotyped. Electrophoretic mobility bins are assigned to each marker, two bins for each ZipChute pair. In order to perform the allele calling, the peaks in the mobility bins are compared relative to an internal size standard. To correct for variability that might occur during different runs, an allelic ladder that contains all the ZipChute probes are also analysed. The signal intensities of the peaks are used by a clustering algorithm to determine the genotype of a sample and the results are presented as cluster (Figure 18) or polar plots [303]. The resulting genotypes can be exported as comma-separated value files or as text delimited files for use in other programs such as Microsoft Excel or a LIMS.

3.9.1. SNplex pool design

The SNPs selected for genotyping were submitted online at the myScience Environment of Applied Biosystems (<http://myscience.appliedbiosystems.com>) for assay design. A list of the selected SNPs was compiled and uploaded and either specified the SNP IDs from NCBI's dbSNP or the SNP-containing sequence as prescribed [312]. The list was submitted to the automated high-throughput pipeline (Figure 19) that designs multiplex pools from the submitted SNP information. Every genotyping method has some limitations and in the case of the SNplex Genotyping System, it is not possible to design assays for all SNPs. There are several reasons for this: 1) The SNP may be present in a region that is homologous to another in the genome, for example in duplications or repeats, and this could lead to non-specificity of the assay, 2) other SNPs in the region may influence probe binding and 3) certain oligonucleotides in the same pool

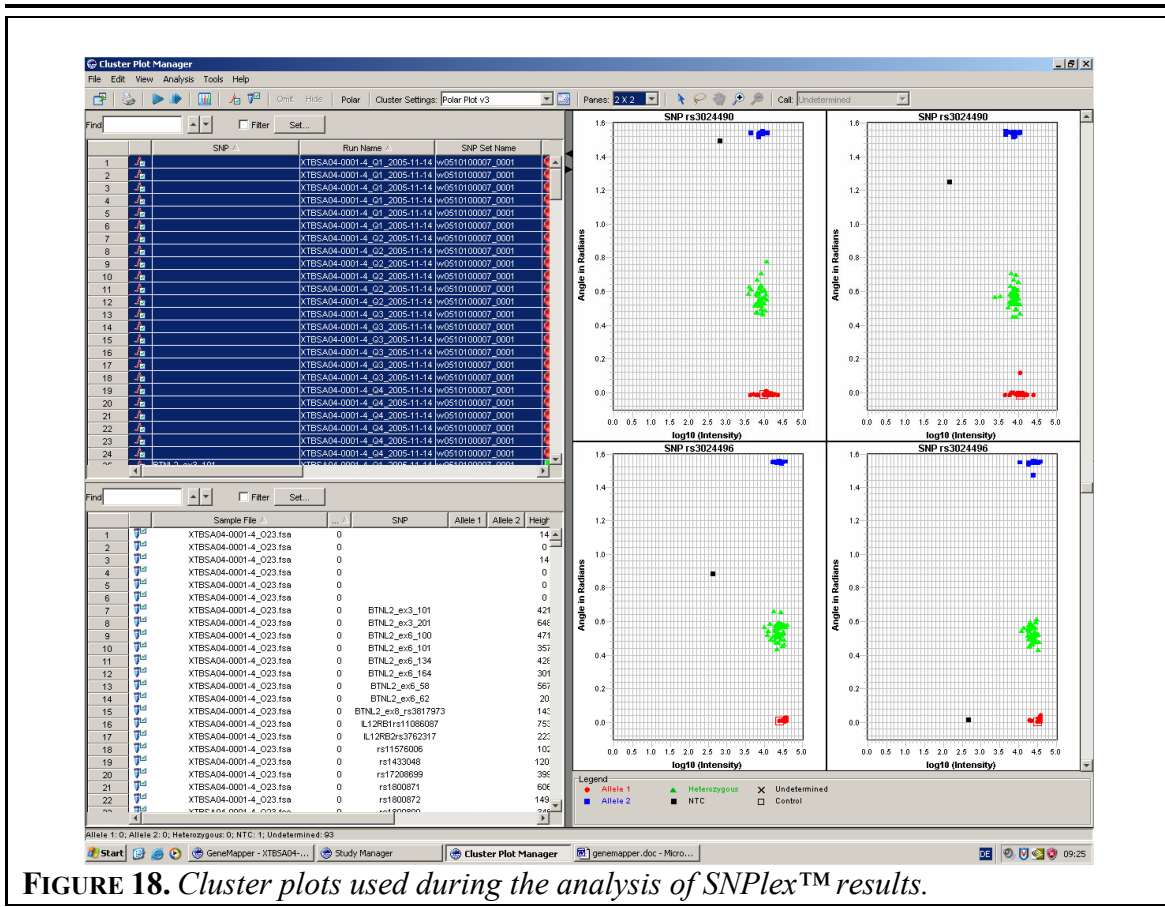


FIGURE 18. Cluster plots used during the analysis of SNplex™ results.

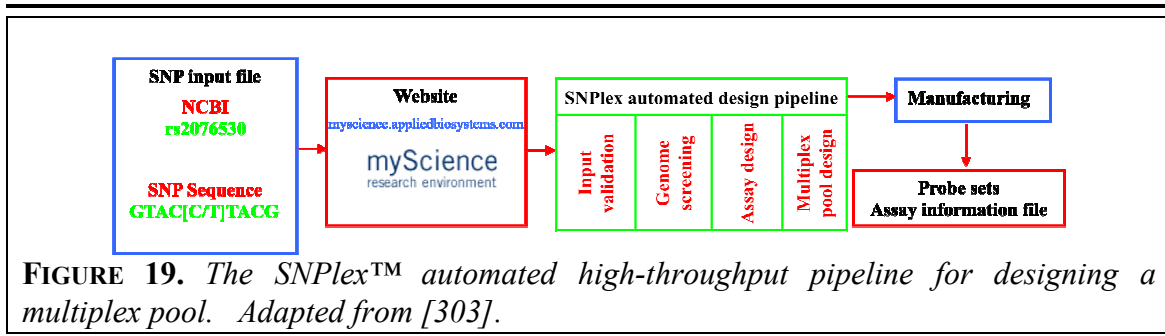


FIGURE 19. The SNplex™ automated high-throughput pipeline for designing a multiplex pool. Adapted from [303].

might interact with each other [312]. The pipeline controls for all these issues and will reject any SNP that does not fulfill these criteria. For this reason it is desirable to target more than 48 SNPs per pool during the design phase, so that an optimised design can be made.

3.9.2. SNPlex protocol

3.9.2.1. Phosphorylation and activation

The SNPlex System Phosphorylation Kit was used to phosphorylate the 5' ends of the SNPlex ligation probes and linkers. The kit consists of SNPlex kinase, SNPlex kinase buffer, SNPlex enhancer, 10X deoxyadenosine triphosphate (dATP) and SNPlex universal linkers. The pooled SNPlex ligation probes were ordered separately. The phosphorylation master mix was prepared with 0.1 μ l of pooled ligation probes, 0.05 μ l of universal linkers, 0.125 μ l of nuclease free water, 0.05 μ l of kinase buffer, 0.025 μ l of kinase, 0.1 μ l of enhancer and 0.05 μ l of dATP for a total volume of 0.5 μ l per reaction. The mix was incubated at 37°C for 1 hour. After incubation the activated ligation probe pool was diluted 1:1 with 0.1X TE buffer, pH 8.

3.9.2.2. OLA

The SNPlex System Ligation Kit contains SNPlex ligase and SNPlex ligase buffer. The kit was used to prepare the OLA master mix with 3.475 μ l of nuclease free water, 0.5 μ l of ligation buffer and 0.025 μ l of ligase per reaction. A total of 4 μ l of the OLA master mix was pipetted with the Tecan Genesis RSP 150 (Tecan Deutschland GmbH, Crailsheim, Germany) to each well of the 384-well plate at 4°C, followed by 1 μ l of the activated SNPlex ligation probe pool. The plates were sealed and transferred from the ice to a GeneAmp PCR system 9700 at 90°C. The cycling consisted of 3 minutes at 90°C followed by 30 cycles of 90°C for 15 seconds, 60°C for 30 seconds and 51°C for 30 seconds with a 2% ramp. This was followed by 10 minutes at 99°C and incubation at 4°C. The OLA plates were centrifuged after thermal cycling.

3.9.2.3. Purification

The SNPlex System Purification Kit was used to create the exonuclease reaction mix. The mixture for one 5 μ l reaction contained 4.2 μ l of nuclease-free water, 0.5 μ l of SNPlex exonuclease buffer, 0.2 μ l of SNPlex lambda exonuclease and 0.1 μ l of SNPlex exonuclease I and was pipetted to the 384-well plates with the Tecan Genesis RSP 150 (Tecan Deutschland GmbH, Crailsheim, Germany). The plates were vortexed, centrifuged and sealed with a heat-seal cover. The thermal cycler was pre-heated to 37°C and the program consisted of 90 minutes at 37°C followed by 10 minutes at 80°C and incubation at 4°C. The plates were centrifuged and 15 μ l of nuclease free water was added to each well.

3.9.2.4. Amplification

The SNPlex System Amplification kit containing the SNPlex amplification master mix (2X) and SNPlex amplification primers (20X) were used to perform the amplification reaction. Each reaction contained 2.42 μ l of nuclease free water, 5 μ l of the amplification master mix, 0.5 of the amplification primers and 2.08 μ l diluted OLA products. The plates were sealed and transferred to the thermal cycler. The PCR

program consisted of 10 minutes at 95°C followed by 30 cycles of 15 seconds at 95 °C, 1 minute at 63°C and incubation at 4°C.

3.9.2.5. Hybridisation

The components of the SNPlex System Hybridisation kit are SNPlex hybridisation plates, SNPlex hybridisation wash buffer, SNPlex hybridisation binding buffer, SNPlex ZipChute dilution buffer and the SNPlex System ZipChute kit, 48-plex (SNPlex denaturant, positive hybridisation control and SNPlex ZipChute mix, 48-plex).

The hybridisation plates containing the streptavidin were washed with 100 µl of hybridisation wash buffer (diluted 1:10). This was repeated three times and was performed with the Power Washer 384 (Tecan Deutschland GmbH, Crailsheim, Germany). The plates were shaken upside down on a paper towel to remove the excess water. The PCR plates from the amplification reaction were vortexed and centrifuged. For each reaction, 17.491 µl hybridisation binding buffer and 0.009 µl of the positive hybridisation control were combined in a master mix. Of this, 17.5 µl was pipetted to each well of the PCR plate with the Tecan robot. The mixture was pipetted to the hybridisation plates and the plates were sealed with Thermowell sealing plates (Corning Incorporated Costar, Corning, USA). The plates were placed on the TiMix control shaker (Edmund Bühler Labortechnik, Tübingen, Germany) and incubated for 1 hour at 600 rpm after which they were centrifuged to collect the contents in the bottom of the wells. The contents of the wells were aspirated and the plates were washed with 100 µl of SNPlex hybridisation buffer in the Power Washer 384 and this was repeated three times. The excess water was shaken from the plates by hand and 50 µl of NaOH was added to each well. The plates were incubated without a seal on the TiMix control rotary shaker (Edmund Bühler Labortechnik, Tübingen, Germany) for 10 minutes at 800 rpm. The supernatant was removed from the wells, each well was washed five times and the excess water was removed by draining the plates upside down on paper towels.

3.9.2.6. Elution

The ZipChute hybridisation mix was prepared with 0.05 µl SNPlex ZipChute mix, 11.25 µl SNPlex denaturant and 13.7 µl SNPlex ZipChute dilution buffer per reaction. The ZipChute mix is light-sensitive and was therefore stored in the dark. A total of 25 µl of ZipChute hybridisation mix was pipetted to each well with the Tecan Te-Mo robot. The plates were incubated for 1 hour at 37°C at 800 rpm on the TiMix control rotary shaker with a TH15 incubator hood (Edmund Bühler Labortechnik, Tübingen, Germany) after which the supernatant was aspirated. The plates were washed with 100 µl of SNPlex hybridisation buffer in the Power Washer 384 and this was repeated four times. The plates were centrifuged upside down at 1000 rpm for 1 minute to remove all buffer solution.

3.9.2.7. Detection

The Sample loading master mix was prepared with the SNPlex System Assay Standards Kit. The total volume of 17.5 µl per reaction consisted of 0.54 µl of SNPlex size standard and 16.96 µl of SNPlex sample loading reagent. This master mix was pipetted to the plates with the Tecan Te-Mo robot and the plates were sealed with foil. The plates

were incubated at 37°C for 30 minutes at 800 rpm on a TiMix control rotary shaker with a TH15 incubator hood. The plates were centrifuged and 7.5 µl of the eluted ZipChutes were pipetted to new 384 well PCR plates. One µl of allelic ladder was pipetted to the control wells. The plates were loaded onto the 3730xl DNA analyser for analysis of the ZipChute fluorescence.

3.9.2.8. Automated allele calling

After the analysis of the ZipChute fluorescence, the .fsa output files were evaluated with the GeneMapper Analysis Software v3.5.1 (Applied Biosystems Darmstadt, Germany). Alleles were automatically called. The results were verified by looking at the distribution of the “clouds” for each SNP in the cluster plots and by checking that the genotypes for the positive CEPH controls were the same on all the plates for the same assay. A maximum of three clustered clouds had to be visible on the plot. In order to ensure genotyping of high quality, the logarithm of the intensity of the fluorescence for a sample had to be greater than 3. The genotypes were exported from Genemapper in tab delimited files. These files were imported into the LIMS for further analysis.

3.10. CLONING OF THE *SH2D1A* GENE PROMOTER

3.10.1. In silico promoter analysis

Sequencing of the putative promoter of *SH2D1A* did not reveal any novel SNPs, but 3 known variants were present in this region, namely G-631A (rs990545), G-494A (rs7357894) and C-347T (rs12164382). The polymorphisms were in complete linkage disequilibrium (LD, $D' = 1$ and $r^2 = 1$) as assessed by Haploview (*Section 3.11.4*) [313]. The SNPInspector program [314] (www.genomatix.de/products/SNPInspector) was used to determine the potential effects of the SNPs on the *SH2D1A* sequence. Only G-631A created 2 putative transcription factor binding sites (for CCAAT enhancer binding protein, beta (C/EBP-β) and pregnane X receptor (PXR), respectively).

3.10.2. Cloning into pGEM[®]-T Easy vector system

3.10.2.1. Amplification and purification of promoter fragment

The 731bp promoter fragments from two male DNA samples, one containing the A-allele (SAP A) and the other containing the G-allele (SAP G) of the G-631A SNP, were PCR amplified. Primers (Table 17) were designed to incorporate endonuclease recognition sites for the *KpnI* and *XhoI* restriction enzymes in the amplification product with three base overhangs to facilitate binding of the specific restriction enzymes during later stages of the experiment.

PCR reactions were carried out in a total volume of 25 µl. Each reaction contained 100 ng of genomic DNA, 2.5 µl of 10X enzyme buffer with 1.5mM MgCl₂, 2 µl of dNTPs (1.25 mM stock), 1 µl forward and reverse primers (10 µM stock) and 0.075 µl of Super-therm GOLD HotStart TAQ DNA polymerase. An Eppendorf Mastercycler was used with the following cycling program: 10 minutes of denaturation at 95°C followed by 35 cycles of 45 seconds at 94°C, 45 seconds at 61°C and 1 minute at 72°C. The reaction was terminated at 72°C for 10 minutes and cooled at 4°C for 2 minutes. Control for contamination was done by the inclusion of water blanks in every batch of samples amplified.

TABLE 17. The names, sequences and product sizes of the primers used for *SH2D1A* promoter cloning. The endonuclease recognition site of *KpnI* (5'-GGTAC↓C-3') and *XhoI* (5'-C↓TCGAG-3') are indicated in red bold type. The three base overhangs, which were added to facilitate restriction enzyme binding, are underlined.

Primer name	Sequence	Size
SAP_ <i>KpnI</i>	5'- <u>AGG</u> GGT ACC AAG ATG ACT GCG TGA GGT AGA-3'	731 bp
SAP_ <i>XhoI</i>	5'- <u>CCG</u> CTC GAG CTC GCC GGT TTC CCT GCT-3'	

The PCR products were purified by vacuum using the Wizard SV Gel and PCR Clean-up System (Promega, Whitehead Scientific, South Africa). From each sample, ten PCR reactions were pooled (250 µl) and mixed with an equal volume of membrane binding solution. The SV minicolumn was placed into the adapter of the vacuum manifold. The prepared PCR product was transferred to the minicolumn and incubated for 1 minute. The vacuum was applied to pull the liquid through the minicolumn after which 600 µl of Membrane Wash Solution was added to the minicolumn. The vacuum was applied and the wash step was repeated. The minicolumn was transferred to a collection tube and centrifuged (Eppendorf centrifuge 5415 D, Merck, South Africa) at full speed for 5 minutes. The minicolumn was transferred to a clean 1.5 ml microcentrifuge tube and 50 µl of nuclease-free water was added to the minicolumn. After a 1 minute incubation step the minicolumn and microcentrifuge tube was centrifuged for 1 minute at full speed to elute the DNA. The DNA concentration was determined with the NanoDrop[®] ND-1000 Spectrophotometer and the NanoDrop[®] v3.0.1 Software.

3.10.2.2. Ligation to pGEM[®]-T Easy vector

The amplified promoter fragments were each ligated to the pGEM[®]-T Easy vector using the T4 DNA ligase and 10X buffer (Promega, Whitehead Scientific, South Africa). This is possible, because the Super-therm GOLD HotStart TAQ DNA polymerase adds adenosine bases to the 3' ends of the amplified fragments, which enables ligation to the 3' tyrosine overhangs of the pGEM[®]-T Easy vector. The ligation was done in a 10 µl reaction with 1 µl of pGEM[®]-T Easy vector, 1 µl of 10X ligation buffer, 3 µl of PCR product and 1 µl of T4 DNA ligase. The ligation reactions were incubated overnight at 4°C. Purification of the ligation reactions were done by ethanol precipitation. Briefly, 90 µl of dH₂O were added to the ligations. The diluted ligation reactions were added to 100 µl of 24:1 chloroform: isoamylalcohol, mixed and centrifuged (Eppendorf Centrifuge 5415, Merck, South Africa) at 14 000 rpm for 30 minutes at 4°C. The top layer was aspirated and added to 2 volumes of ethanol, after which the centrifugation step was repeated. The ethanol was aspirated and the tubes were incubated at 65°C until all ethanol had evaporated. The DNA was dissolved in 10 µl of dH₂O.

3.10.2.3. Transformation

The purified ligation products were used for transformations with competent *Escherichia coli* (strain DH5-α) cells. For each transformation, 2 µl of ligation product and 50 µl of competent cells were transferred to a cuvette and electroporated (Bio-Rad Gene pulser, capacitance = 125 µF, resistance = 1000 Ω, 250 V). After

electroporation, 1 µl of SOC (*Addendum A.1.*) medium were added to the cells. This was incubated in a shaker (Orbital shaker incubator LM-530, Yih Der, Japan) at 37°C for 1 hour. From each transformation reaction 50, 100 and 150 µl were plated onto LB agar plates supplemented with 100 mM ampicillin, 40 µg /ml X-gal and 0.2mM IPTG. The plates were incubated overnight at 37°C to allow colony formation.

3.10.2.4. Clone selection

After overnight incubation, blue and white colonies were visible on the plates. White colonies, which indicated vectors containing inserts, were selected with a pipette tip and swirled in a 25 µl PCR reaction. This screening PCR (*Section 3.10.2.1.*), with the SAP_KpnI and SAP_XhoI primers, were used to determine if the selected colonies contained the insert. The remainder of the single colony on the tip was inoculated in 20 µl of LB and placed on a shaker at 37°C until the results from the screening PCR were available. One µl of this 20 µl culture was used to inoculate 11ml of LB medium with 11 µl of ampicillin if the colony showed amplification of the desired insert. The 50ml tubes were incubated overnight at 37°C in a shaker.

3.10.2.5. Small-scale plasmid extraction (Miniprep)

After overnight growth, glycerol stocks were prepared by collecting 850 µl of the bacterial culture and mixing it with 150 µl of 100% glycerol. This was stored at -80°C. The Wizard Plus SV Miniprep Kit (Promega, Whitehead Scientific, South Africa) was used for plasmid extraction. The remaining 10ml of bacterial culture was centrifuged for 10 minutes at 3000rpm and the supernatant was discarded. The cell pellet was resuspended in 300 µl of cell resuspension solution by pipetting vigorously. The resuspended cells were transferred to a 2ml tube to which 350 µl of cell lysis solution was added and was gently rotated. The cells were resuspended by adding 10 µl of alkaline protease and incubating for 5 minutes. After incubation, 450 µl of neutralisation solution was added and the contents of the tube were flicked to the top of the tube to create a homogenous solution. This was centrifuged at 13.2 rpm for 10 minutes. The clear top layer was transferred to the minicolumn and centrifuged for 1 minute. The columns were washed with 600 µl of column wash solution and centrifuged for 5 minutes at 13.2 rpm. This step was repeated. The plasmid DNA was eluted with 25 µl of nuclease-free water at 65°C and the concentration was determined with the NanoDrop® ND-1000 Spectrophotometer and the NanoDrop® v3.0.1 Software. The resultant constructs were named pGEM(SAP G) and pGEM(SAP A) and were sequenced with the Sp6 and T7 primers (Table 18) to verify the nucleotide sequence and orientation of the inserts.

TABLE 18. *The names and sequences of the primers used for sequencing of the pGEM(SAP G) and pGEM(SAP A) constructs.*

Primer name	Sequence
Sp6	5'-TAT TTA GGT GAC ACT ATA G-3'
T7	5'-TAA TAC GAC TCA CTA TAG GG -3'

3.10.3. Subcloning into pGL4 luciferase reporter vector

3.10.3.1. Digestion and purification

The pGEM(SAP_G) and pGEM(SAP_A) constructs and the luciferase-reporter vector, the pGL4.10[*luc2*] plasmid (Promega, Whitehead Scientific, South Africa), were double digested using the *KpnI* and *XhoI* restriction enzymes. The digestion of the pGEM constructs were done in a total volume of 40 μ l using 8 μ g of plasmid each, 20 μ l of 10X buffer (SuRE/cut buffer L, Roche, South Africa), 5 μ l of *KpnI* (10 U/ μ l, Roche, South Africa), 0.5 μ l of bovine serum albumin (BSA) and 5 μ l of *XhoI* (10 U/ μ l, Roche, South Africa). The pGL4.10[*luc2*] plasmid was digested in a total volume of 30 μ l using 4 μ g of plasmid, 10 μ l of 10X buffer (SuRE/cut buffer L), 2 μ l of *KpnI*, 0.3 μ l of BSA and 2 μ l of *XhoI* (10 U/ μ l, Roche, South Africa). The reactions were incubated overnight at 37°C. To prevent recircularisation of the digested pGL4.10[*luc2*] plasmid during ligation, the 5' ends were dephosphorylated using 1 μ l of SAP. This was incubated at 37°C for 1 hour and the enzyme was inactivated by incubation at 65°C for 15 min.

The reactions were resolved on a 1% agarose gel at 150 V for 1 hour. The fragments were excised from the gel under UV light with a sterile scalpel. Membrane binding solution was added to the gel slices as determined by its weight (10 μ l of solution for every 10 mg of gel) and placed at 65°C to melt before proceeding with the Wizard SV Gel and PCR Clean-up System (Promega) as described in section 3.10.2.1. The DNA was eluted with 20 μ l of nuclease-free water. The DNA concentration was determined with the NanoDrop[®] ND-1000 Spectrophotometer and the NanoDrop[®] v3.0.1 Software.

3.10.3.2. Ligation and ethanol precipitation

The digested promoter fragments were each ligated to the pGL4.10[*luc2*] luciferase reporter vector using T4 DNA ligase and 10X buffer (Promega, Whitehead Scientific, South Africa). The ligation was done according to the manufacturer's protocol which recommends that a 3:1 vector: insert DNA ratio should be used. The digested SAP A, SAP G and pGL4.10[*luc2*] products had concentrations of 16.8, 12.1 and 61.6 ng/ μ l respectively. To achieve the 3:1 ratio, 3 and 4 μ l of each of SAP A and SAP G were used in a 10 μ l reaction with 1.6 μ l of digested pGL4.10[*luc2*], 1 μ l of 10 ligation buffer and 1 μ l of T4 DNA ligase. The ligation reactions were incubated overnight at 4°C.

Purification of the ligation reactions were done by ethanol precipitation. Briefly, 90 μ l of dH₂O were added to the ligations. The diluted ligation reactions were added to 100 μ l of 24:1 chloroform: isoamylalcohol and centrifuged (Eppendorf Centrifuge 5415) at 14 000 rpm for 30 minutes at 4°C. The top layer was aspirated and added to 2 volumes of ethanol, after which the centrifugation step was repeated. The ethanol was aspirated and the tubes were incubated at 65°C until all ethanol had evaporated. The DNA was dissolved in 10 μ l of dH₂O.

3.10.3.3. Transformation

The purified ligation products were used for transformations with competent DH5- α *Escherichia coli* cells. This was done as described in section 3.10.2.3. From each transformation reaction 100, 150 and 300 μ l were plated onto LB agar plates

supplemented with 100 mM ampicillin and 40µg/ml X-gal. The plates were incubated overnight at 37°C to allow colony formation.

3.10.3.4. Clone selection

After overnight incubation, white colonies which contained the vector, were visible on the plates. Clones were screened by PCR as described in section 3.10.2.4. Positive reactions for each of the SAP A and SAP G fragments were selected and 1 µl of these reactions were inoculated into 50 ml LB medium with 50 µl of ampicillin. The cultures were incubated overnight at 37°C in a shaker for use in an endotoxin-free plasmid extraction.

3.10.3.5. Endotoxin-free plasmid extraction (Midiprep)

After overnight incubation, glycerol stocks were prepared as described in section 3.10.2.5. The remainder of the bacterial cultures were used for plasmid extractions with the endotoxin-free PureYield™ Plasmid Midiprep system (Promega, Whitehead Scientific, South Africa) using an adjusted protocol of the manufacturer. The cells were pelleted using centrifugation (Eppendorf Centrifuge 5810 R) at 2000 x g for 15 minutes and the supernatants were discarded. The pellets were resuspended in 3 ml of Cell Resuspension Solution until dissolved. After resuspension, 3 ml of Cell Lysis Solution were added to each reaction and the tubes were gently inverted 5 times. This solution was incubated for 3 minutes at room temperature and 5 ml of Neutralization Solution were added to the lysed cells. This solution was mixed by inverting the tubes 5 times. The lysates were incubated in an upright position for 3 minutes to allow a white precipitate to form.

Two blue lysate clearing PureYield™ Clearing Columns were placed into 50 ml Falcon plastic tubes. The lysates were added to the column and incubated for 2 minutes to allow the cellular debris to rise to the top. The columns were centrifuged at 1500 x g for 5 minutes. The filtered lysates were transferred to the white DNA binding PureYield™ Binding Columns in new 50 ml Falcon tubes and centrifuged at 1500 x g for 3 minutes. After the first centrifugation step, 5 ml of Endotoxin Removal wash solution were added to the columns and centrifuged at 1500 x g for 3 minutes. The flowthroughs were discarded and 20 ml of Column Wash Solution were added. The columns were centrifuged at 1500 x g for 5 minutes. The flowthroughs were discarded and the columns centrifuged for an additional 10 minutes to ensure the removal of the Column Wash Solution. The tips of the columns were tapped on a paper towel to ensure that all the Column Wash Solution had been removed. The columns were transferred to clean 50 ml tubes and the DNA was eluted with 600 µl of nuclease-free water by centrifuging at 1500 x g for 5 minutes. The filtrates were collected and transferred to 1.5 ml tubes.

The DNA concentrations were determined with the NanoDrop® ND-1000 Spectrophotometer and the NanoDrop® v3.0.1 Software. The correct nucleotide sequences of the inserts were verified by DNA sequence analysis with the Rvprimer3 (Promega, Whitehead Scientific, South Africa) and SAP_ *XhoI* primers. The resultant pGL4 luciferase constructs were named pGL4(SAP A) and pGL4(SAP G).

3.11. STATISTICAL ANALYSIS

3.11.1. Quality checking of the genotype data

Allele frequencies can be used to predict genotype frequencies in a stable population if there is no mutation, migration, natural selection or non-random mating present. This is the Hardy-Weinberg model [315,316] and it is valid for most human populations [317]. If the assumptions of this model are true for a population, the genotype and allele frequencies will remain constant over consecutive generations, and the population is considered to be in Hardy-Weinberg equilibrium (HWE). The model allows the comparison of a population's actual genetic structure with the genetic structure expected if the population were not evolving. If genotype frequencies differ from those expected under equilibrium, it can be assumed that one or more of the model's assumptions are violated. In case-control association studies a departure from HWE can indicate genotyping error [318], population stratification or, sometimes if found in disease cases, an association of the marker with the disease [317].

The model consists of two equations: one that calculates allele frequencies and one that calculates genotype frequencies. The $p + q = 1$ equation describes allele frequencies for two alleles of a marker and if the frequency for one allele is known, the other can be calculated. The $1 = p^2 + 2pq + q^2$ equation is used to calculate the expected genotype frequencies from known allele frequencies, where p and q are the proportions of alleles. p^2 and q^2 are the proportions of the homozygotes pp and qq , while pq is the proportion of the heterozygotes. To calculate the expected number of pp genotypes, the equation $p^2 \times n$ was used, where p refers to the allele frequency of p and n is the total number of samples examined in the population. The same calculation would be used to determine the expected number of qq and pq genotypes. Once the expected values have been calculated, the chi-square test (χ^2 , *section 3.10.2.1.*) is performed to determine if the population is in HWE. This value must be less than 3.84 for the population not to be significantly out of HWE, because the test has one degree of freedom.

The basic principles of inheritance were described by Gregor Mendel in 1866. According to Mendel's law of segregation, alleles separate during meiosis and one allele of each parent is inherited by their descendant [319]. These principles can be applied in pedigrees, since the 4 parental alleles of a bi-allelic marker can only result in certain combinations of alleles in the offspring. Deviations from these expected genotypes, called Mendel errors, could indicate non-paternity or errors in the genotyping method.

The quality of genotyping was verified by checking for conformance with HWE in the control population and by Mendel checking in families. Mendel checking was done with the Genotool program in the IKMB LIMS [272,273].

3.11.2. Genotype distributions of SNPs

Contingency tables were drawn up to determine significant differences in the distribution of genotypes (2 x 3 table) between cases and controls. Only genotypes were considered, since allelic tests are more prone to results in false positive associations. The two-tailed chi-square test or the Fisher's exact test (depending on the data analysed, *sections 3.10.2.1* and *3.10.2.2.*) was done to establish whether the results were significantly different between the genotype data. Prism version 4.02 was used to

calculate the odds ratios and p-values as reported. When a significant result was found, Bonferroni corrections for multiple testing were done by considering the number of independent linkage disequilibrium (LD) blocks (*Section 1.2.2.4.*) and the number of independent SNPs [76] in each gene.

3.11.2.1. The chi-square test

The chi-square (χ^2) test is a statistical test that determines if observed frequencies are significantly different from expected frequencies. A null hypothesis (H_0 : there is no significant difference between the observed and expected frequencies) and an alternative hypothesis (H_1 : there is a significant difference between the observed and expected frequencies) is stated. [320] The value for χ^2 is calculated by the following equation:

$$\chi^2 = \sum \frac{(\text{observed number} - \text{expected number})^2}{\text{expected number}}$$

The chi-square tests were done with Prism version 4.02.

3.11.2.2. The Fisher's exact test

This test must be used for small sample sizes and when the expected number in any cell of a contingency table is less than six [321] and calculates the exact p-value. In this study the test was done using Prism version 4.02. Prism always uses this exact test for 2 x 2 contingency tables. For genotype data, which consists out of 2 x 3 tables, the two smallest genotype numbers in a group are added together and then compared to the remaining data.

3.11.3. Transmission disequilibrium test

In this study, a transmission disequilibrium test (TDT, *discussed in section 1.2.2.2.*) was done to confirm results found in the case-control studies. If TDT results suggested that there was proof of linkage, it could be deduced that LD exists between the marker allele and the disease allele. The TDTs in the family sample set were analysed with Haploview v3.3 [313]. Single markers and haplotypes were evaluated. Power calculations for TDTs were done with the Genetic Power Calculator [322].

3.11.4. Haplotypes and linkage disequilibrium

Haplotypes and LD (*discussed in sections 1.2.2.3. and 1.2.2.4.*) in the case-control cohort were analysed with Haploview v3.3 [313]. Pairwise LD was analysed in terms of D' . The D' value ranges from 0 to 1, where 0 indicates no LD and 1 indicates complete LD. Haploview automatically selects haplotype blocks according to a commonly used block definition from Gabriel et al [89]. This block definition calculated the 95% confidence levels for the D' values and the comparisons are labelled as strong LD, uncertain or strong recombination. If 95% of the comparisons are labelled as strong LD, a block will be created. Markers with a minor allele frequency of less than 0.05 are ignored in this algorithm. With this definition, several overlapping blocks can be found, therefore the program sorts the list of possible blocks, starts with the largest and keeps adding the blocks until they overlap with an already declared block [89,313]. The chi-square test was used by Haploview to determine any significant

differences between the two groups. In addition, we tested global significance for these haplotypes using the Cocaphase program in the Unphased suite [101] and permutation testing in Haploview, since haplotypes in a haplotype block is not independent and Bonferroni correction is too conservative in this instance. Using the Cocaphase program, 10 000 permutation replicates were done for each block. In Haploview, 10 000 or 100 000 permutation replicates were done, depending on the distribution graph generated by the program. The “global” (empirical) p value from Cocaphase and Haploview represents the total significance when the observed versus expected values of all of the haplotypes are analysed together.

3.11.5. Microsatellite analysis

The numbers of repeats from the microsatellite markers were determined by direct counting and plotted as a distribution graph. The distribution of the peaks on this graph (either bi- or trimodal) determined the number of subclasses (two (S or L) or three (S, M or L)) into which the alleles were divided. This methodology was previously described by studies in the analysis of microsatellite polymorphisms [291,323-325]. To evaluate the genotypic effect on the development of TB, the genotype classes (three (S/S, S/L, L/L) or six (S/S, S/M, S/L, M/M, M/L, L/L)) resulting from this classification were divided into two subgroups depending on the presence or absence of the S allele. Possible associations between the disease groups and classes of alleles were analysed for significance by the chi-square test. Distribution graphs were prepared with Microsoft Excel and the chi-square tests were done with Prism version 4.02.

3.11.6. Power calculations

Our case-control studies had the power to detect an odds ratio of 2.15 with 95% confidence given an allele frequency of at least 5% based on the sample size available (Epi Info 2000, CDC). The TDTs had 68% power to detect the same odds ratio as the case-control studies given an allele frequency of at least 5% [322]. In order to have 95% power 409 family trios would be needed for the TDT analysis.

CHAPTER 4

CANDIDATE GENES WHICH INFLUENCE INTERFERON GAMMA LEVELS

4.1. INTRODUCTION

Interferon gamma (IFN- γ) is a member of the interferon family. These cytokines play a crucial role in resistance to pathogens [326] and were first identified as agents that interfere with viral replication [327]. IFNs are classified into type 1 and type 2 depending on receptor specificity and sequence homology. IFN- γ is the only type 2 IFN [328] and was first identified in lymphocyte supernatants as a unique lymphocyte antiviral activity. The cytokine is produced not only by natural killer (NK) cells [329], CD4+ and CD8+ T cells [330,331], but also by $\gamma\delta$ T cells [332], NK T cells [333], B cells, myeloid cells [334] and type II alveolar epithelial cells [335].

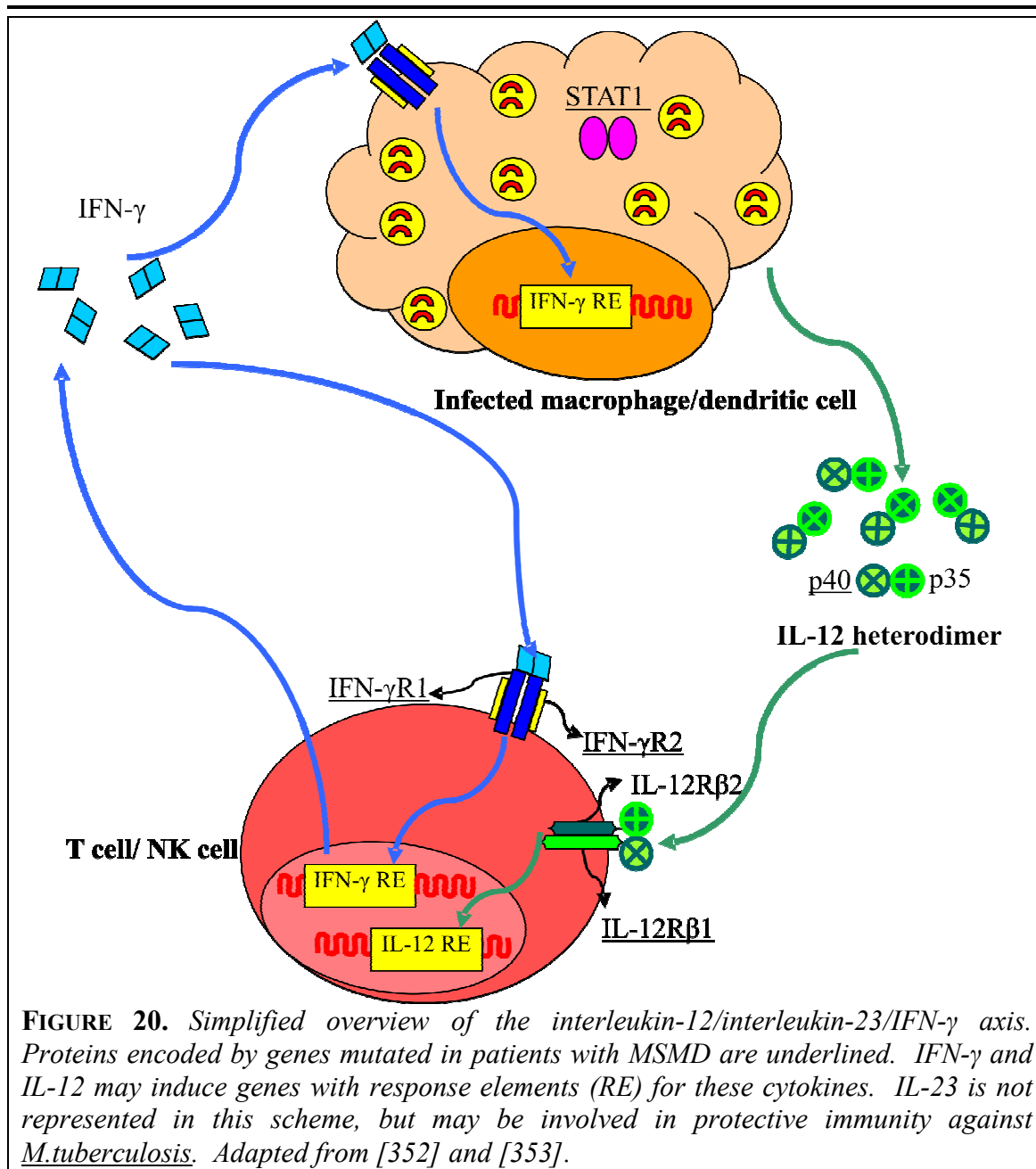
There has been a great deal of interest in IFN- γ since its discovery, because the macrophage, an important target cell of IFN- γ , plays a central role in the immune system. The IFN- γ /macrophage system is vital in natural and acquired resistance to infection and defects in the system may lead to inflammatory and autoimmune disease [336].

The T helper 1 (Th1) cell response, which is required to contain *Mycobacterium tuberculosis* (*M.tuberculosis*) infection, is largely characterised by IFN- γ production. Production of IFN- γ is essential for the effectiveness of the innate immune system against *M.tuberculosis*. However production of this cytokine alone is not sufficient to protect against disease [337]. Even so, convincing evidence for its importance in the control of mycobacterial infections have been found in both experimental and clinical studies. Mice with a disrupted IFN- γ gene show increased susceptibility to tuberculosis (TB) [338-340] and replacement of the gene into the lung confers resistance [341].

Humans with mutations in genes of the interleukin-12 (IL-12)/interleukin-23 (IL-23)/IFN- γ axis (Figure 20), such as interleukin-12, beta (*IL12B*) [50,342,343], interleukin-12 receptor, beta 1 (*IL12RB1*) [45,46,344-346], IFN- γ receptor-1 (*IFNGR1*) [268,347-349], IFN- γ receptor-2 (*IFNGR2*) [350] or signal transducer and activator of transcription-1 (*STAT1*) [351], have an increased susceptibility to even non-pathogenic mycobacteria. These mutations are all associated with the rare human syndrome known as Mendelian susceptibility to mycobacterial disease (MSMD, Online Mendelian Inheritance in Man #209950). The affected patients are also susceptible to *M.tuberculosis* and *Salmonella*, but not to other infections.

In addition to these experimental and clinical studies, genetic association studies have also suggested that this cytokine is important in protecting against mycobacterial infection. The functional 874A \rightarrow T polymorphism, a common variant in the IFN- γ gene, has been associated with TB in the South African Coloured population [180]. This polymorphism is also associated with TB in Spanish [179], Sicilian [181], Turkish [182], Colombian [354] and Hong Kong Chinese [183] populations, but not in Malawians [133], Croatians [185] and Americans from Caucasian, African or Hispanic descent [248].

In most TB patients the production of *M.tuberculosis*-induced IFN- γ by peripheral blood mononuclear cells is reduced at the time of diagnosis [355]. During and after successful treatment, these levels increase significantly [356]. It is also known that the



IFN- γ concentrations in sputum and bronchoalveolar lavage fluid can be used as an estimate of disease activity [357,358]. Currently, experimental data propose that IFN- γ is the best correlate of protective immunity against TB [359,360]. However, it is known that *M.tuberculosis* modulates the anti-mycobacterial action of IFN- γ by adopting various mechanisms, such as inhibiting the response of macrophages to IFN- γ [361,362]. A recent study indicated that the bacterium accomplished this by downregulating the stimulating protein 1, which in turn was responsible for downregulating IFN- γ R1 and altering the cells sensitivity of the cells to IFN- γ [363].

IFN- γ has been considered as an immunotherapy to treat TB [364-367] and certain studies using aerosolized IFN- γ as a treatment have reported favourable results when

using this approach [368,369]. However administering IFN- γ subcutaneously to patients with chronic TB and advanced multidrug-resistant (MDR)-TB found that this treatment did not affect the disease parameters [360]. The authors concluded that treatment with aerosolised IFN- γ might be useful in treating MDR-TB in patients that respond poorly to treatment.

Given the clinical and experimental evidence showing a crucial role of IFN- γ in host defence against TB, we investigated polymorphisms in candidate genes which could potentially modulate expression levels of this vital cytokine.

4.2. C-C CHEMOKINE LIGAND-2

Chemokines play an important role in the development of immune responses against TB. The C-C chemokine ligand-2 gene (*CCL2*) encodes monocyte chemoattractant protein-1 (MCP-1), a protein which is essential for the recruitment of monocytes, T lymphocytes [370] and NK cells [371] to the site of mycobacterial infection. It may also take part in the localisation of TB in the lungs by contributing to granuloma formation [370] and possibly have a role in T cell differentiation [372]. Mice deficient in *CCL2* are more susceptible to TB during the early stages of the infection [373].

It is known that MCP-1 production is promoted by IFN- γ [374]. In turn, MCP-1 attracts antigen-specific IFN- γ -secreting T cells to the site of infection [373]. Since MCP-1 is responsible for the recruitment of immune cells to the infection site, it is possible that levels of this chemokine will determine the initial cellular immune response at the site of disease and will therefore indirectly influence IFN- γ levels. MCP-1 also inhibits production of IL-12p40, a key player in the IL-12/IL-23/IFN- γ pathway [175]. For this reason it is likely that the clinical presentation of TB and its outcome will be partially regulated by MCP-1, which could conceivably have a major effect on disease progression.

It is well documented that polymorphisms in the promoter region of a gene can affect its expression levels. This is also true for *CCL2*, where the expression of the gene seems to be controlled by two sections in the 5' regulatory region. The proximal regulatory region, located approximately 150 bp upstream [375] of the transcriptional start site, responds to various cytokines and contains several putative transcription factor binding sites [376-378]. In addition, a distal regulatory region, which contains two binding sequences for nuclear factor- $\kappa\beta$ and other unknown cis-acting regulatory elements [376,379,380], is located 1.8 to 2.7 kb upstream [375] of the transcriptional start site.

The functional [381-383] *CCL2* promoter polymorphism rs1024611 (named MCP-1-2518 [381] or MCP-1-2578 [384] in the respective studies) was associated with increased susceptibility to pulmonary TB in Mexicans and the finding was replicated in Koreans [175]. Patients with a GG genotype had a higher risk of progression to disease. Monocytes from these individuals with a GG genotype were stimulated with *M.tuberculosis* antigens and produced higher concentrations of MCP-1 and lower concentrations of IL-12p40 than monocytes from individuals with the AA genotype [175]. This polymorphism was not associated with TB in a smaller study of Brazilians [65]. In addition, the T allele of the rs1024610 single nucleotide polymorphism (SNP, named MCP-2136 [384] or MCP-1-2076 [385]) may also influence MCP-1 levels [384]. The rs1024610 and rs1024611 SNPs were in complete linkage disequilibrium (LD) in previous studies [384,385], but both polymorphisms contributed separately to increased MCP-1 levels in patients [384].

The rs1024611 *CCL2* polymorphism (Table 19) was investigated in the South African Coloured population to test the possible association of the gene with TB based on a previous association study which found an association between it and disease [175]. In addition six other SNPs from this gene were genotyped to elucidate the LD structure of the region in this unique South African population. Both a case-control study design and a transmission disequilibrium test (TDT) were used.

TABLE 19. Polymorphisms genotyped in the *CCL2* gene.

Polymorphism	Class	Location	Genotyping method
rs2857654	SNP	Promoter	TaqMan
rs1024611	SNP	Promoter	SNPlex
rs1024610	SNP	Promoter	SNPlex
rs3760399	SNP	Promoter	SNPlex
rs3760396	SNP	Promoter	SNPlex
rs4586	SNP	Exon 2, synonymous	SNPlex
rs2530797	SNP	3' UTR ^a	SNPlex

^a untranslated region (UTR)

4.2.1. Results

4.2.1.1. Single-point statistical analysis

All of the SNPs genotyped in *CCL2* were in Hardy-Weinberg equilibrium (HWE) in the control population. The threshold for significance was set at $p = 0.025$ after Bonferroni correction, since 6 of the 7 genotyped SNPs were in LD with each other. Under this criterion, none of the analysed SNPs reached statistical significance. The functional [381-383] promoter polymorphism rs1024611 was not associated with TB in our case-control study ($p = 0.24$, Table 20) or TDT ($p = 0.59$, Table 21), a finding similar to a previous study considering Brazilian TB patients [65]. However, it was previously associated with increased susceptibility to pulmonary TB in Mexicans and in Koreans [175]. These contradictory results suggest an ethnic-specific role of *CCL2* polymorphisms in TB. The other SNP with a known functional effect, rs1024610, was also not associated with TB in the South African Coloured population. None of the five remaining SNPs, which included the synonymous SNP rs4586 and the functional distal promoter polymorphism rs2857654 [375], were associated with disease in the case-control or TDT analyses.

TABLE 20. Single-point statistical analysis in the *CCL2* case-control study.

Polymorphism	Allele		Genotype frequencies						Cases vs Controls p value ^c
			Cases (n ^a =432)			Controls (n=482)			
	1 ^b	2	11	12	22	11	12	22	
rs2857654	A	C	0.10	0.43	0.47	0.11	0.43	0.46	0.90
rs1024611	G	A	0.06	0.33	0.61	0.08	0.36	0.56	0.24
rs1024610	A	T	0.01	0.16	0.83	0.02	0.19	0.79	0.17
rs3760399	C	T	0.01	0.11	0.88	0.01	0.11	0.88	0.62
rs3760396	G	C	0.01	0.11	0.88	0.01	0.11	0.88	0.89
rs4586	T	C	0.20	0.52	0.28	0.24	0.48	0.28	0.26
rs2530797	G	A	0.06	0.37	0.57	0.05	0.4	0.55	0.82

^a number of investigated individuals.

^b Allele 1 is the minor allele.

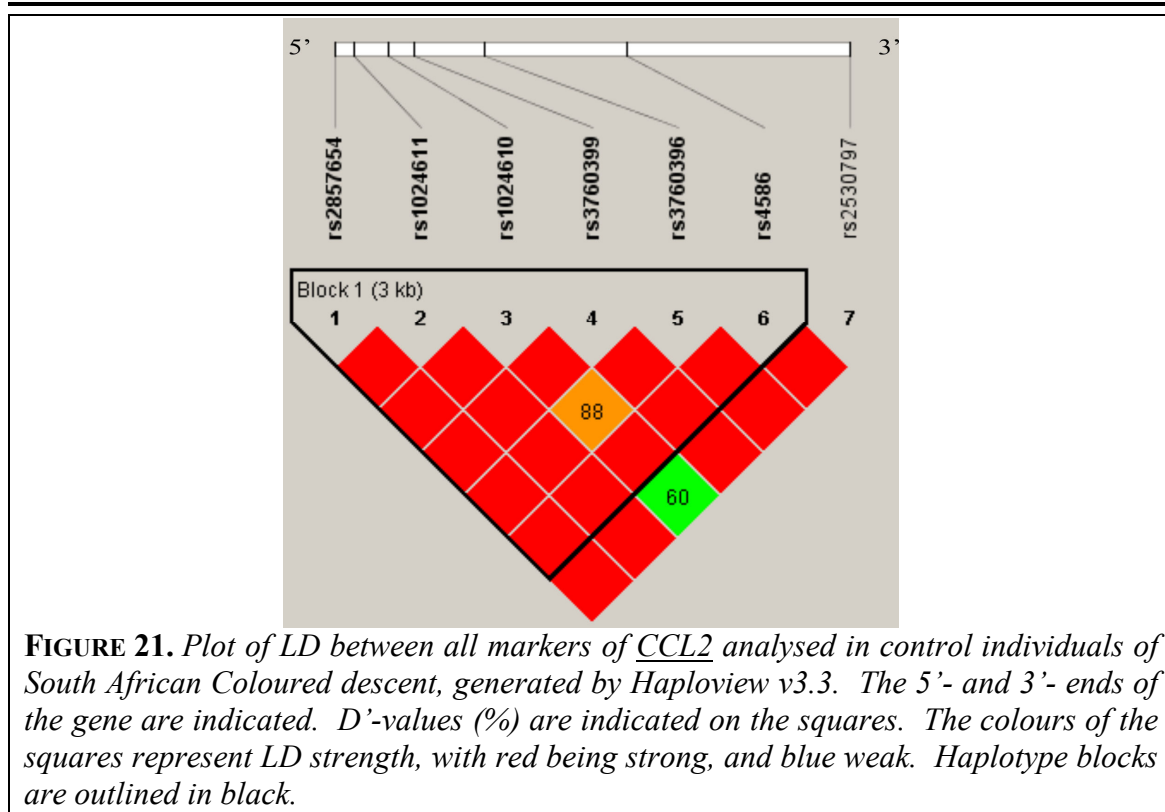
^c p value from a genotype-based χ^2 test

TABLE 21. Single-point statistical analysis in the *CCL2* TDT.

SNP	Overtransmitted	Transmitted:Untransmitted	χ^2	p value
rs2857654	C	53 : 47	0.36	0.55
rs1024611	G	45 : 40	0.29	0.59
rs1024610	A	20 : 12	2	0.16
rs3760399	C	8 : 6	0.29	0.59
rs3760396	G	17 : 13	0.53	0.47
rs4586	T	57 : 53	0.15	0.70
rs2530797	G	42 : 41	0.01	0.91

4.2.1.2. Haplotype analysis

The Haploview program [313] was used to identify haplotype blocks (Figure 21) for the markers genotyped in *CCL2* in the control samples. Strong LD exists across the *CCL2* gene in the South African population. The promoter polymorphisms and exonic SNP were part of a single haplotype block, while the 3'UTR SNP (rs2530797) was not included in this block. Since no additional information on the haplotype block containing the 3'UTR SNP was available, it was not considered in the statistical analysis of haplotypes with TB.



Seven haplotypes were determined for the case-control samples in the haplotype block (Table 22). Haplotype 1 was the most frequent haplotype in controls, while haplotype 2 had the highest frequency in cases. Haplotype 5 had a nominally significant difference

between cases and controls ($p = 0.02$). To test the global significance of the haplotypes, we did 10 000 permutations using Haploview [313]. According to the calculation of the program, 1154 out of these 10 000 permutations exceeded the best observed χ^2 (haplotype 5, $\chi^2 = 5.79$, $p = 0.02$). This means that there is a 12% chance that the significant nominal p value of haplotype 5 was observed due to chance alone. Haploview also identified the same single haplotype block for the markers genotyped in *CCL2* for the samples from the TB families. None of the estimated haplotypes were significantly associated with TB before or after permutation testing using the haplotype TDT analysis in Haploview (Table 23).

TABLE 22. *CCL2* case-control haplotype analysis by Haploview [313].

Block 1 ^a	Freq _{Cases}	Freq _{Controls}	χ^2	p value	p value _{Permutation} ^b
1 A-G-T-T-C-C	0.22	0.26	2.99	0.08	0.45
2 C-A-T-T-C-T	0.24	0.24	0.01	0.93	1
3 C-A-T-T-C-C	0.22	0.20	1.78	0.18	0.73
4 C-A-A-T-C-T	0.09	0.11	2.78	0.10	0.50
5 A-A-T-T-C-C	0.09	0.06	5.78	0.02	0.12
6 C-A-T-T-G-C	0.07	0.07	0.16	0.69	1
7 C-A-T-C-C-T	0.07	0.06	0.23	0.63	1

^a The order of the SNPs in each block corresponds to Figure 21.

^b Permutation test p values were calculated from 10 000 permutations in Haploview.

TABLE 23. *CCL2* TDT haplotype analysis by Haploview [313].

Block 1 ^a	Transmitted:untransmitted	χ^2	p value	p value _{Permutation} ^b
1 C-A-T-T-C-T	25.0 : 33.0	1.11	0.29	0.80
2 A-G-T-T-C-C	30.0 : 26.6	0.20	0.65	1
3 C-A-T-T-C-C	21.6 : 20.7	0.02	0.90	1
4 C-A-A-T-C-T	13.0 : 10.5	0.27	0.60	1
5 C-A-T-T-G-T	11.0 : 9.3	0.13	0.71	1
6 A-A-T-T-C-C	6.0 : 8.4	0.40	0.53	0.99
7 C-A-T-C-C-T	6.0 : 4.0	0.40	0.53	0.98

^a The order of the SNPs in each block corresponds to Figure 21.

^b Permutation test p values were calculated from 100 000 permutations in Haploview.

4.2.1.3. Comparison of allele frequency and LD between different populations

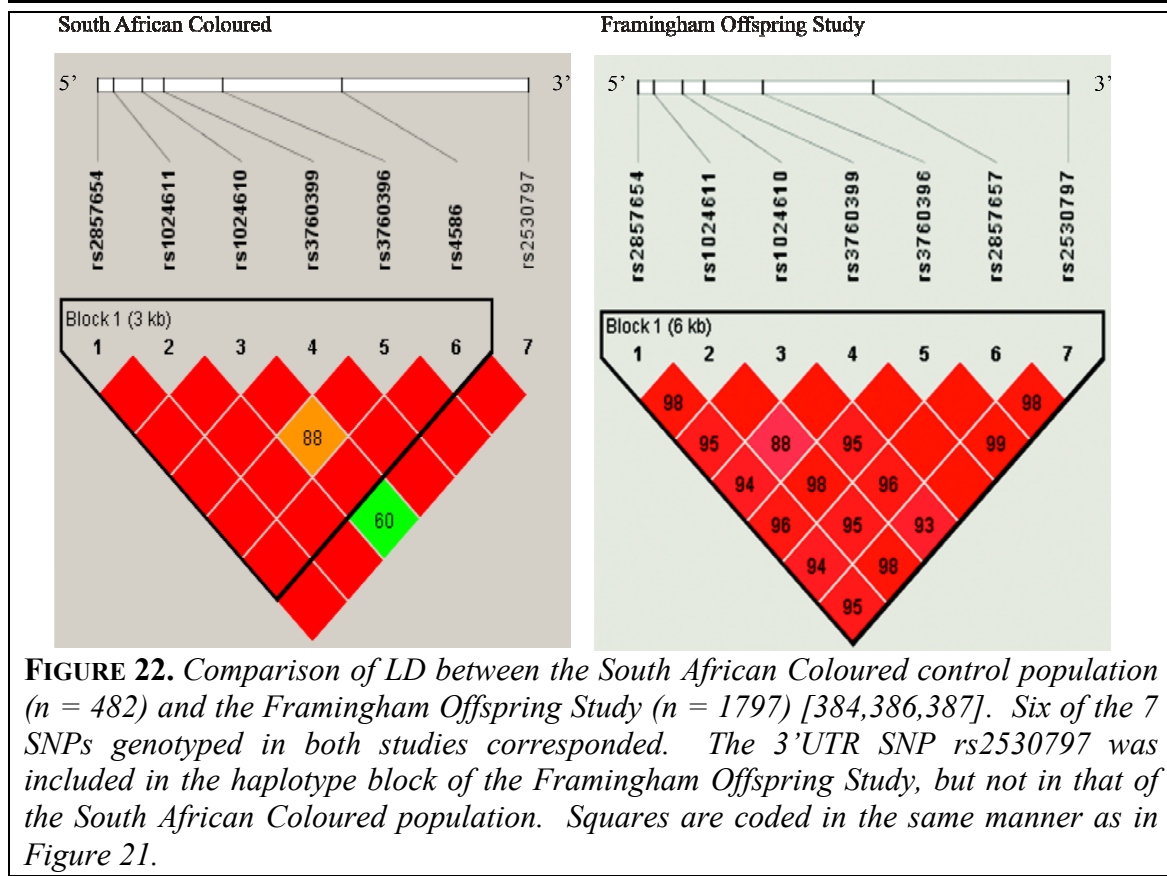
We compared the allele frequencies of rs1024611 between populations to evaluate the distribution of the SNP in 4 different populations, namely South African Coloureds, Mexicans [175], Koreans [175], African Americans [375] and Caucasians [375] (Table 24). Allele frequencies of control samples from each study were considered. The allele frequencies of the South African Coloureds, African Americans and Caucasians were similar, and the G allele was the minor allele in these 3 populations. The G allele was also the minor allele in the Koreans, but it was the major allele in Mexicans. Spanish, Amerindian and African populations contributed to the admixture in the Mexican

population [175]. The A allele is the major allele in African Americans, suggesting that the G allele was probably the major allele in the Spanish or Amerindian ancestors of the Mexican population.

TABLE 24. Comparison of allele frequencies of rs1024611 between populations.

Allele	South African	Mexican [175]	Korean [175]	African American [375]	Caucasian [375]
A	0.78	0.49	0.64	0.78	0.71
G	0.22	0.51	0.36	0.22	0.29

We also compared the LD structure (Figure 22) of *CCL2* between the South African Coloured control population (n = 482) and a community-based cohort from Framingham, New England, namely the Framingham Offspring Study (n = 1797) [386,387], in which an association between *CCL2* polymorphisms and MCP-1 levels and myocardial infarction was found [384]. Both the present study and the Framingham Offspring Study genotyped 7 SNPs from *CCL2*, of which 6 corresponded. McDermott et al [384] included the intronic rs2857657 and we included the synonymous rs4586. In both studies strong LD was found across the *CCL2* region. The 3'UTR SNP was not part of the haplotype block determined by Haploview in the South African Coloured population, but it was part of the single haplotype block determined by McDermott et al [384].



4.2.2. Discussion

CCL2 encodes the MCP-1 protein, a chemokine which recruits immune cells to the site of TB infection. MCP-1 levels are influenced by IFN- γ production and it could possibly also regulate IFN- γ levels in return since it determines the migration of IFN- γ producing cells to the disease site. This study in the South African Coloured population did not find an association between TB susceptibility and polymorphisms in *CCL2*. This is similar to the finding of Jamieson et al [65] in a Brazilian population, but differs from the association result of Flores-Villaneuva et al in Mexicans and Koreans [175].

We did not observe any association between the genotypes of the three functional SNPs (rs2857654, rs1024610 and rs1024611) and TB susceptibility in the case-control or TDT studies. None of the four remaining SNPs showed any indication of association either. We also tested for an association between TB and the estimated haplotypes. One haplotype in the case-control group had a nominally significant p value, but this was lost after correcting for multiple testing and is unlikely to be a true reflection of association with this region.

In addition, we compared allele frequencies of the rs1024611 SNP between populations. The South African Coloureds, African Americans [375] and Caucasians [375] had similar frequencies and shared the same major allele (A). The alternate allele (G) was the major allele in Mexicans [175]. Since this is the allele previously associated with TB susceptibility in Mexicans and Koreans [175], the lower frequency of the allele in the South African Coloured, African American and Caucasian populations suggest that it is not the cause of TB susceptibility in these populations, but probably an ethnic-specific susceptibility factor.

LD analysis of the *CCL2* region in the South African Coloured population confirmed the strong LD observed in other populations for this region. However, where another study found one haplotype block in this region [384], the results from this study suggest that two haplotype blocks are present in the South African Coloured population, dividing the gene in 5' and 3' segments. The rs1024610 and rs1024611 SNPs were in complete LD in the South African Coloured population, as in previous studies [384,385].

In summary, polymorphisms in *CCL2* are not associated with TB in the South African population and the gene is probably an ethnic-specific susceptibility factor.

4.3. INTERLEUKIN-4

Interleukin-4 (IL-4) is traditionally a T helper 2 (Th2) cytokine and can be produced by a variety of cell types, such as T cells, NK cells, eosinophils, basophils and some antigen-presenting cells (APCs) [388]. During the early stages of the Th1 response, IL-4 can promote the associated reactions, but it may downregulate the Th1 response during later phases [388]. In addition, IL-4 has a regulatory influence on the alternative activation of mouse macrophages [389], a process where an increased expression of, and phagocytosis by, the mannose receptor is observed. *M.tuberculosis* makes use of the mannose receptors to enter the macrophages without hindrance. IL-4 enhances this process due to its anti-inflammatory properties [390].

Contradictory results as to the importance of IL-4 in TB have been observed in genetically deficient *IL4*^{-/-} mice infected with mycobacteria [390-395]. It has been argued that IL-4 does not play an important role in all mouse models of TB [391] as these studies differed in the mouse strain used and in the dosage of mycobacteria administered [390]. In two studies, where mice with C57BL/6 [395] and B6129 [393] genetic backgrounds were used, inactivation of *IL4* did not affect the disease. In contrast, *M.tuberculosis*-infected *IL4/IL13*^{-/-} mice on a C57BL/6 background generated significantly higher levels of IFN- γ mRNA than wild type mice on day 50 after infection [394]. BALB/c mice normally have an increased expression of IL-4 during TB, as do humans. When *IL4* was disrupted, IFN- γ levels remained high [391], as opposed to *IL4*^{+/+} BALB/c mice, where IL-4 also had unfavourable effects on the Th1 response.

Contrasting effects of IL-4 are also observed during TB disease in humans. IFN- γ production is favoured during mild disease, while IL-4 levels increase in advanced stages of TB with a simultaneous decrease in IFN- γ production and the Th1 immune response [195]. These seemingly contradictory effects of IL-4 could be explained by the presence of the competitive IL-4 antagonist, IL-4 δ 2. IL-4 δ 2 is a splice variant of *IL4*, lacking exon 2 of the gene, and displays Th1-like properties by inhibiting the actions of IL-4 [396,397]. Individuals with latent TB infection, but who remain healthy, have high levels of this variant mRNA [398]. During TB chemotherapy IL-4 δ 2, but not IL-4 levels, increased in HIV positive and negative patients [399]. In addition, there is a difference in the stability of the two *IL4* mRNA products in TB patients, with IL-4 being more stable than IL-4 δ 2 [400]. Therefore the ratio of IL-4 and IL-4 δ 2 may play a role in disease progression and outcome [400].

Polymorphisms in the *IL4* promoter could influence the transcription levels of the gene. Several such polymorphisms in *IL4* have been identified previously. Specifically, a functional SNP (rs2243250, IL-4-C589T), located 589 bp upstream of the translational site, has been associated with increased promoter strength, stronger binding of transcription factors and with different levels of IL-4 activity [401,402]. The CC genotype of this polymorphism was previously associated with protection against pulmonary TB in a South Indian population [195], but not in Gambians [403]. The rs2243250 SNP, together with rs2070874, was also associated with asthma and total IgE levels in white Americans [404]. No linkage was found between the chromosome 5q region, which contains *IL4*, and TB in Brazilian families [405,406].

TABLE 25. Polymorphisms genotyped in the *IL4* gene in the South African Coloured population.

Polymorphism	Class	Location	Genotyping method
rs2243248	SNP	Promoter	SNPlex
rs2243250 (-589)	SNP	Promoter	TaqMan
rs2070874 (-33)	SNP	Promoter	SNPlex
rs2243251	SNP	Exon 1, synonymous	SNPlex
rs2243291	SNP	3' UTR	SNPlex

IL4 polymorphisms (Table 25) were investigated in the South African Coloured population to test the possible association of the gene with TB and to attempt replication of the previous association found with rs2243250 in the South Indian TB population [195]. In addition, several other SNPs from this gene were genotyped to elucidate the LD structure of the region in this unique South African population. Case-control and family-based association studies were employed.

4.3.1. Results

4.3.1.1. Single-point statistical analysis

All SNPs were in HWE for the control population. The threshold for significance was set at $p = 0.025$ after Bonferroni correction, using the methodology of Nicodemus et al [76], since 4 of the 5 genotyped SNPs were in strong LD with each other. Under this criterion, none of the analysed SNPs (Tables 26 and 27) were associated with TB.

TABLE 26. Single-point statistical analysis in the *IL4* case-control study.

Polymorphism	Allele		Genotype frequencies						Cases vs Controls p value ^c
			Cases (n ^a =432)			Controls (n=482)			
	1 ^b	2	11	12	22	11	12	22	
rs2243248	G	T	0.13	0.47	0.40	0.11	0.41	0.48	0.08
rs2243250	C	T	0.28	0.49	0.23	0.24	0.53	0.23	0.37
rs2070874	C	T	0.35	0.45	0.20	0.36	0.50	0.14	0.05
rs2243251	A	G	0.54	0.38	0.08	0.61	0.31	0.08	0.08
rs2243291	C	G	0.32	0.47	0.21	0.28	0.51	0.20	0.37

^a number of investigated individuals.

^b Allele 1 is the minor allele.

^c p value from a genotype-based χ^2 test

The rs2243250 SNP has a known functional effect on *IL4* transcription, but was not associated with TB in the South African Coloured population in the case-control ($p = 0.37$) or TDT ($p = 0.088$) analyses. This SNP was also not associated with TB in a Gambian population [403], but was associated with disease in a South Indian population [195]. The rs2070874 SNP had the smallest p value ($p = 0.05$) of the polymorphisms studied, but this was not considered to be statistically significant, because of the

adjusted threshold ($p = 0.025$) for significance. None of the other SNPs were associated with disease either.

TABLE 27. Single-point statistical analysis in the *IL4* TDT.

SNP	Overtransmitted	Transmitted:Untransmitted	χ^2	p value
rs2243248	G	51 : 48	0.091	0.76
rs2243250	T	58 : 41	2.9	0.088
rs2070874	T	64 : 54	0.85	0.36
rs2243251	G	41 : 33	0.87	0.35
rs2243291	C	54 : 42	1.5	0.22

4.3.1.2. Haplotype analysis

Haplotype and LD analysis for *IL4* was done with the Haploview program. The program identified one haplotype block consisting of 4 SNPs. Strong LD exists in this block (Figure 23). The rs2243248 SNP, which was not included in this block, is in the distal promoter region.

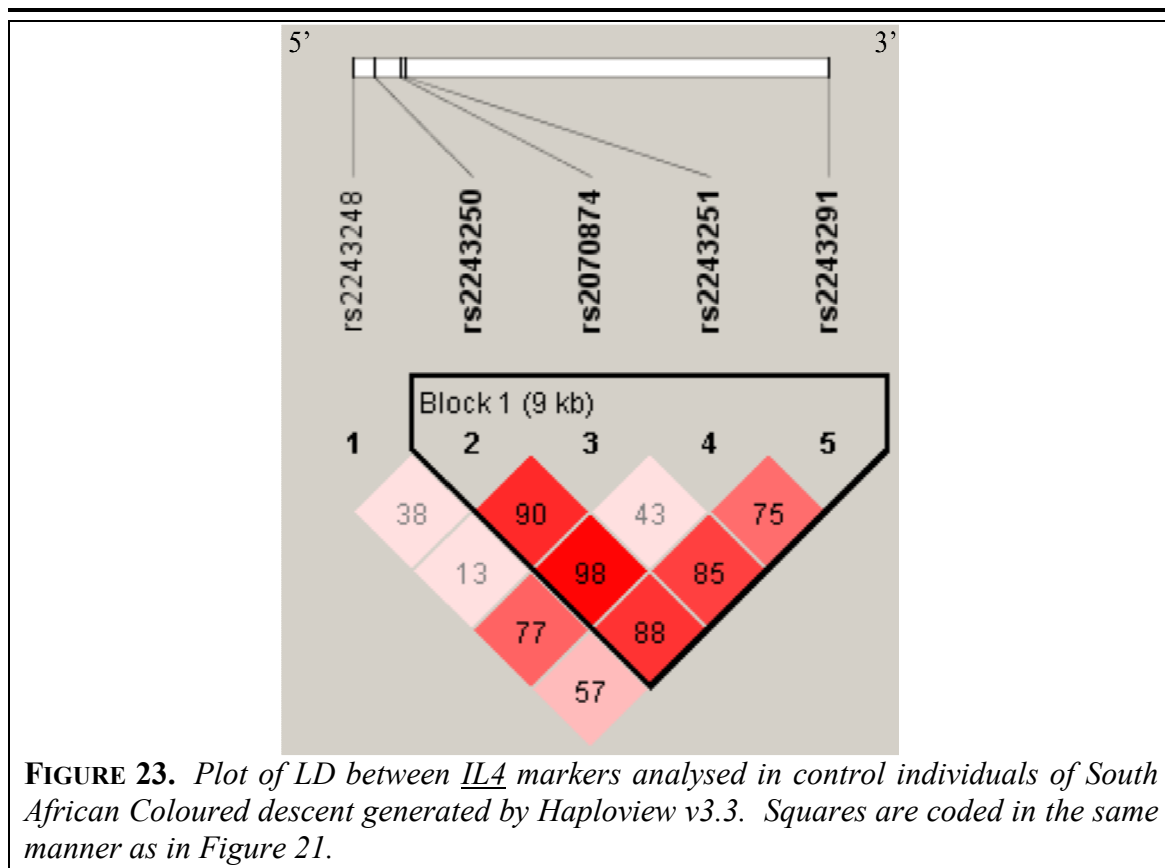


FIGURE 23. Plot of LD between *IL4* markers analysed in control individuals of South African Coloured descent generated by Haploview v3.3. Squares are coded in the same manner as in Figure 21.

In both the case-control (Table 28) and TDT (Table 29) samples eight haplotypes were present in the single haplotype block identified. Haplotype 7 had the highest frequency in both the cases and controls. Haplotype 1 had a nominally significant p value of 0.02 in the case-control study. The empirical p value ($p_{\text{value}_{\text{permutation}}}$) represents the

number of times the observed data would be expected to occur in the data set analysed. During permutation testing in the case-control samples, p values less than 0.02 were observed on 1107 out of 10 000 occasions (11%). These permutation results imply that there is an 11% chance that the significant difference between cases and controls with the rare haplotype 1 is likely to be due to random variation in the selection of the SNPs.

Similarly, permutation testing in the TDT resulted in the observation of p values less than 0.05 on 28933 out of 100 000 occasions (29%). These permutation results suggest that the significant difference in the transmission of the rare haplotype 5 from parents to their affected children is likely to be the result of random variation in the selection of the SNPs and not a reflection of a true association.

TABLE 28. *IL4 case-control haplotype analysis by Haploview [313].*

Block 1 ^a	Freq _{Cases}	Freq _{Controls}	χ^2	p value	p value _{Permutation} ^b
1 T-T-G-G	0.03	0.02	5.40	0.02	0.11
2 T-T-G-C	0.16	0.13	4.10	0.04	0.22
3 T-T-A-C	0.20	0.22	1.17	0.28	0.91
4 T-C-G-C	0.07	0.09	2.04	0.15	0.65
5 T-C-A-C	0.06	0.05	0.62	0.43	0.98
6 C-T-A-C	0.02	0.02	0.43	0.51	0.99
7 C-C-A-G	0.41	0.44	2.06	0.15	0.64
8 C-C-A-C	0.04	0.03	0.64	0.42	0.98

^a The order of the SNPs in each block corresponds to Figure 23.

^b Permutation test p values were calculated from 10 000 permutations in Haploview.

TABLE 29. *IL4 TDT haplotype analysis by Haploview [313].*

Block 1 ^a	Transmitted:untransmitted	χ^2	p value	p value _{Permutation} ^b
1 T-T-G-G	3 : 5	0.58	0.45	1
2 T-T-G-C	18 : 16	0.09	0.77	1
3 T-T-A-C	27 : 21	0.87	0.35	1
4 T-C-G-C	13 : 7	1.80	0.18	0.73
5 T-C-A-C	11 : 11	0.01	0.94	1
6 C-T-A-C	7 : 5	0.54	0.46	1
7 C-C-A-G	29 : 39	1.46	0.23	0.82
8 C-C-A-C	1 : 6	3.75	0.05	0.29

^a The order of the SNPs in each block corresponds to Figure 23.

^b Permutation test p values were calculated from 100 000 permutations in Haploview.

4.3.1.3. Comparison of allele frequency of rs2243250 between different populations

We compared the allele frequency of rs2243250 (Table 30) between five different populations, namely the South African Coloureds from this study, Indians [195] and the populations from the HapMap [59] (Section 4.9), namely Han Chinese from Beijing, Japanese from Tokyo, Yoruba Nigerians from Ibadan and individuals of European

descent from Utah. The T allele was the major allele in the South African Coloured, Asian and Yoruba Nigerian populations, while the C allele was the major allele in the Indians and Europeans, who also shared similar allele frequencies. The intermediate frequency of the C allele (0.47) in the South African Coloureds controls reflects the genetic contribution of a number of parental populations.

TABLE 30. Comparison of allele frequencies of rs2243250 between populations.

Allele	South African	Indian [195]	Asian, Beijing and Tokyo [59]	Yoruba Nigerians [59]	Europeans, Utah [59]
C	0.47	0.89	0.27	0.17	0.84
T	0.53	0.11	0.73	0.83	0.16

4.3.2. Discussion

In this study, polymorphisms in *IL4* were tested as susceptibility factors in TB. Promoter polymorphisms of the gene have previously been shown to influence transcription of *IL4*. We did not find any association with the SNPs studied and TB in cases and controls or families from the South African Coloured population.

The rs2243250 SNP has a known functional effect on *IL4* transcription [401,402] and the CC genotype was previously associated with TB in a South Indian population by Vidyarani et al [195]. We, and a study in The Gambia [403], did not find an association with this polymorphism and TB. We also did not see any association with the other SNPs or haplotypes and TB susceptibility. We have determined that the allele frequency of rs2243250 differed significantly between the South African Coloureds and other populations.

The discrepancies reported in association studies with TB for polymorphisms in *IL4* could be the result of differences in allele frequencies and the extent of LD in the various populations studied. However, it is also possible that previous studies reporting an association with *IL4* may reflect the common problem of an initial association which proves to be a false positive. Another possibility is that alternative splicing, and not increased expression of the *IL4* gene, is the actual regulatory mechanism in the production of IL-4, since the product of alternative splicing of *IL4*, namely IL-4 δ 2, is a competitive antagonist of IL-4. The apparent exclusion of a significant genetic effect of this gene in TB susceptibility is therefore not an indication that the cytokine is not important in TB disease.

4.4. INTERLEUKIN-10

Interleukin-10 (IL-10) is a powerful Th2 regulatory cytokine and plays an essential role during the latent TB stage, where increased production of this cytokine promotes reactivation of disease [407] and suppression of cell mediated immunity against the intracellular infection. This is mediated via inhibition of MHC-restricted cytotoxicity against infected macrophages [408]. In mice, this cytokine inhibits the production of IFN- γ [409,410]. IL-10 may also downregulate interleukin-12 (IL-12) responses [411-413] and thereby influence IFN- γ levels [413,414] (*Section 4.5.*). The multifunctional IL-10 also inhibits the replication of macrophages and lymphocytes and the secretion of other inflammatory cytokines [415]. Individuals innately susceptible to TB after infection with *M.tuberculosis* produce more IL-10 in response to lipopolysaccharides (LPS) than those who are resistant [218].

The three most frequently investigated polymorphic variants (rs1800872, rs1800871 and rs1800896) in *IL10*, studied in numerous diseases and disorders such as brucellosis [416], psoriasis [417] and TB [171,179,194,418], are located in the proximal promoter. These SNPs are linked and in most cases three haplotypes (CCG, ATA and CCA) are observed [419]. The haplotypes and individual SNPs have been shown to correlate with IL-10 production, transcriptional activity and nuclear-binding activity [419-422]. The rs1800896 SNP has previously been associated with TB in Cambodians [190], Sicilians [171] and Turkish [189], but not in Korean [418], Gambian [194] or Spanish populations [179]. The rs1800872 SNP is in complete LD with rs1800871 and these SNPs were associated with TB in Koreans [418], but not in Gambians [194]. A 4 SNP haplotype of *IL10*, which consisted of rs1800896, rs1800871, rs1800872 and rs3024496, were also associated with TB in the Koreans [418]. Polymorphisms in the distal region of the *IL10* promoter, such as rs1800895, may also affect IL-10 production [423]. *IL10* variants (Table 31) were investigated to test the possible association of the gene with TB and to attempt replication of previous association studies in different population groups. The selected polymorphisms included three frequently studied functional markers, namely rs1800896, rs1800871 and rs1800872, as well as a SNP in the distal promoter region (rs1800890). In addition other SNPs from this gene were genotyped to elucidate the LD structure of the region in the South African Coloureds.

TABLE 31. *Polymorphisms genotyped in IL10 in this study.*

Polymorphism	Class	Location	Genotyping method
rs1800890 (-3575)	SNP	Promoter	SNPlex
rs1800893 (-1330)	SNP	Promoter	SNPlex
rs1800896 (-1082)	SNP	Promoter	SNPlex
rs1800871 (-819)	SNP	Promoter	SNPlex
rs1800872 (-592)	SNP	Promoter	SNPlex
rs3024490	SNP	Intron 1 (HapMap)	SNPlex
rs3790622	SNP	Intron 1 (HapMap)	SNPlex
rs3024496 (+117)	SNP	3' UTR	SNPlex
rs3024498	SNP	3' UTR	SNPlex

4.4.1. Results

4.4.1.1. Single-point statistical analysis

All polymorphisms were analysed in the case-control and family samples and were in HWE for the control samples. Since all nine genotyped SNPs were in strong LD with each other, the threshold for significance was set at $p = 0.05$. We found no association between the functional promoter rs1800896 SNP and TB in the case-control (Table 32) or TDT (Table 33) analysis. The frequently studied rs1800872 and rs1800871 promoter polymorphisms were not associated with TB either. The rs1800890 SNP in the distal region of the promoter was also not associated with disease. None of the other SNPs genotyped in *IL10* were associated with TB susceptibility.

TABLE 32. Single-point statistical analysis in the *IL10* case-control study.

Polymorphism	Allele	Genotype frequencies							Cases vs Controls p value ^c
		Cases (n ^a =432)			Controls (n=482)				
		1 ^b	2	11	12	22	11		
rs1800890	T A	0.05	0.38	0.58	0.06	0.34	0.60	0.34	
rs1800893	A G	0.16	0.53	0.31	0.16	0.46	0.37	0.12	
rs1800896	G A	0.09	0.46	0.45	0.11	0.42	0.47	0.43	
rs1800871	T C	0.09	0.43	0.48	0.11	0.48	0.42	0.15	
rs1800872	A C	0.09	0.43	0.48	0.11	0.48	0.42	0.16	
rs3024490	T G	0.11	0.41	0.48	0.13	0.46	0.42	0.14	
rs3790622	T C	0	0	1	0	0.01	0.99	0.17	
rs3024496	C T	0.23	0.53	0.24	0.24	0.48	0.29	0.21	
rs3024498	G A	0.01	0.24	0.75	0.01	0.19	0.80	0.17	

^a number of investigated individuals.

^b Allele 1 is the minor allele.

^c p value from a genotype-based χ^2 test

TABLE 33. Single-point statistical analysis in the *IL10* TDT.

SNP	Overtransmitted	Transmitted:Untransmitted	χ^2	p value
rs1800890	A	37 : 36	0.01	0.91
rs1800893	G	58 : 48	0.94	0.33
rs1800896	A	48 : 41	0.55	0.46
rs1800871	C	47 : 39	0.74	0.39
rs1800872	C	47 : 40	0.56	0.45
rs3024490	G	46 : 34	1.80	0.18
rs3790622	-	-	-	-
rs3024496	C	51 : 41	1.09	0.30
rs3024498	A	18 : 11	1.69	0.19

4.4.1.2. Haplotype analysis

Haplotypes in *IL10* have previously been investigated in TB susceptibility. We analysed LD and haplotypes with the Haploview program, which identified a single haplotype block, consisting of all 9 SNPs genotyped in this gene region (Figure 24). The SNPs were mostly in strong LD. Only the rs3790622 and rs3024498 variants displayed weak LD (D' value = 5%). This could be explained by the extremely low allele frequency of rs3790622, which was hardly polymorphic in the South African Coloured population.

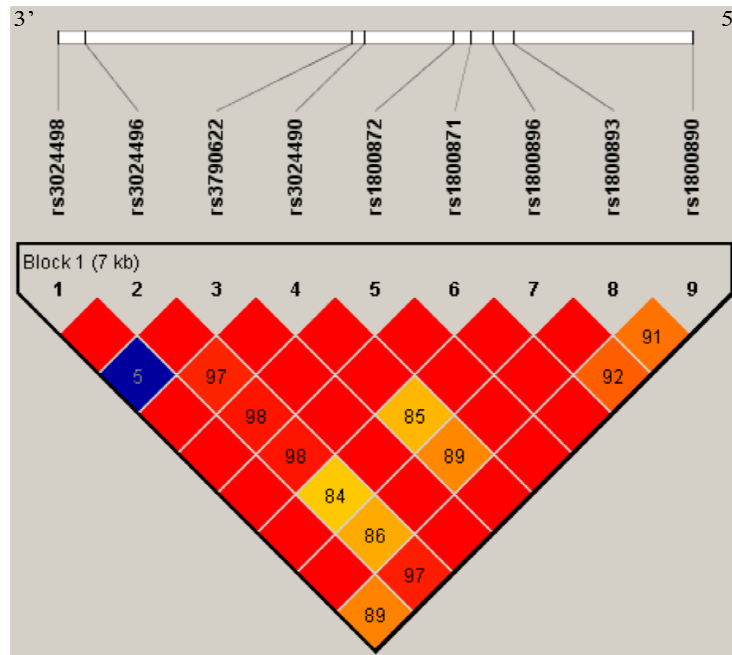


FIGURE 24. Plot of LD between *IL10* markers analysed in control individuals of South African Coloured descent generated by Haploview v3.3. Squares are coded in the same manner as in Figure 21.

Eight haplotypes (Table 34) were estimated in the case-control samples for the single haplotype block of *IL10*, while ten haplotypes were estimated in the family samples (Table 35). Haplotypes 3 and 9 (Table 35) were present in the family samples, but not in the case-control group, possibly due to their low frequencies. The most frequent haplotype in the cases and controls was haplotype 8 (Table 34). This haplotype was also the most frequent in the TDT analysis. Haplotype 1 was nominally associated ($p = 0.05$) in the case-control group, but permutation results implied that this was due to random variation in the selection of the SNPs (Table 34). Haplotype 2 (Table 35) in the family samples had a nominally significant p value of 0.0048 and was transmitted more frequently to affected offspring than expected. After permutation calculations, this association remained significant, with an empirical p value of 0.01, implying that the result is true. However, this association was not observed in the case-control analysis, where the haplotype had a low frequency (Table 34, haplotype 2). None of the other haplotypes were associated with TB.

TABLE 34. *IL10 case-control haplotype analysis by Haploview [313].*

Block 1 ^a	Freq _{Cases}	Freq _{Controls}	χ^2	p value	p value _{Permutation} ^{b, c}
1 A-C-C-G-C-C-A-A-A	0.10	0.08	3.71	0.05	0.30
2 A-C-C-G-C-C-A-G-A	0.09	0.10	0.16	0.69	1
4 A-C-C-G-C-C-G-A-A	0.07	0.07	0.01	0.93	1
5 A-C-C-G-C-C-G-A-T	0.10	0.12	2.23	0.14	0.61
6 A-T-C-G-C-C-A-G-A	0.17	0.15	0.97	0.32	0.96
7 A-T-C-G-C-C-G-A-A	0.02	0.02	0.06	0.80	1
8 A-T-C-T-A-T-A-G-A	0.31	0.34	2.27	0.13	0.60
10 G-C-C-G-C-C-G-A-T	0.12	0.10	3.43	0.06	0.40

^a The order of the SNPs in each block corresponds to Figure 24.

^b Permutation test p values were calculated from 100 000 permutations in Haploview.

^c p values less than 0.05 were observed on 30 018 out of 100 000 occasions (30%)

TABLE 35. *IL10 TDT haplotype analysis by Haploview [313].*

Block 1 ^a	Transmitted:untransmitted	χ^2	p	p _{Permutation} ^{b, c}
1 A-C-C-G-C-C-A-A-A	13 : 18	0.78	0.38	1
2 A-C-C-G-C-C-A-G-A	17 : 3	9.75	0.0048	0.01
3 A-C-C-G-C-C-A-G-T	2 : 1	0.33	0.56	1
4 A-C-C-G-C-C-G-A-A	9 : 5	1.14	0.29	0.93
5 A-C-C-G-C-C-G-A-T	12 : 15	0.33	0.56	1
6 A-T-C-G-C-C-A-G-A	21 : 15	1	0.32	1
7 A-T-C-G-C-C-G-A-A	1 : 5	3.54	0.061	0.34
8 A-T-C-T-A-T-A-G-A	29 : 33	0.26	0.61	1
9 A-T-C-T-C-C-G-A-A	1 : 4	2.86	0.088	0.58
10 G-C-C-G-C-C-G-A-T	9 : 13	0.73	0.39	1

^a The order of the SNPs in each block corresponds to Figure 24.

^b Permutation test p values were calculated from 100 000 permutations in Haploview.

^c p values less than 0.048 were observed on 624 out of 100 000 occasions (0.6%)

TABLE 36. *IL10 three SNP haplotype analysis by Haploview [313].*

Block 1 ^a	Freq _{Cases}	Freq _{Controls}	χ^2	p value	p value _{Permutation} ^{b, c}
1 C-C-A	0.37	0.34	2.71	0.10	0.22
2 A-T-A	0.31	0.34	3.26	0.07	0.16
3 C-C-G	0.32	0.32	0.02	0.90	0.99

^a The order of the SNPs is rs1800872, rs1800871 and rs1800896.

^b Permutation test p values were calculated from 10 000 permutations in Haploview.

^c p values less than 0.16 were observed on 624 out of 10 000 occasions (16%)

We also considered a 3 SNP haplotype, previously associated with TB in Gambians, Spanish, Sicilians and Koreans [171,179,194,418] (Table 36, rs1800872, rs1800871 and

rs1800896), but it was not significantly associated with disease in the South African Coloured case-control study population. As in previous studies of the 3 SNP haplotype, rs1800896, rs1800871 and rs1800872 were linked and only 3 haplotypes, namely CCA, ATA and CCG were observed [419].

The estimation of a 4 SNP haplotype, previously associated with TB in Koreans [418] (Table 37, rs3024496, rs1800872, rs1800871 and rs1800896), indicated that 5 haplotypes were present in the South African Coloured population. The Korean study inferred 4 haplotypes containing these SNPs [418] and found that the second most frequent haplotype in their study (TCCA, haplotype 4 in Table 37) was associated with disease. This was not the case in our study, where the haplotype was not as frequently observed (0.16 in controls) and also not associated with TB.

TABLE 37. *IL10* four SNP haplotype analysis by Haploview [313].

Block 1 ^a	Freq _{Cases}	Freq _{Controls}	χ^2	p value	p value _{Permutation} ^{b, c}
1 T-A-T-A	0.30	0.34	2.93	0.09	0.27
2 C-C-C-G	0.30	0.29	0.03	0.85	1
3 C-C-C-A	0.20	0.18	1.61	0.20	0.66
4 T-C-C-A	0.17	0.16	0.62	0.43	0.94
5 T-C-C-G	0.03	0.03	0.03	0.88	1

^a The order of the SNPs is rs30224496, rs1800872, rs1800871, rs1800896.

^b Permutation test p values were calculated from 10 000 permutations in Haploview.

^c p values less than 0.09 were observed on 2707 out of 10 000 occasions (27%).

4.4.1.3. Comparison of allele frequencies and haplotypes of *IL10* between populations

Three promoter polymorphisms and a 3'UTR SNP have been studied extensively in diverse populations. We collected allele frequency information for rs1800896, rs1800871, rs1800872 and rs3024496 from previously published studies [171,179,183,189,190,194,418,424,425,425,426] and compared the results with that of the South African Coloured population (Table 38). In the South African Coloured population the major A allele of rs1800896 had similar allele frequencies to the Cambodian and Gambian populations (66-68%). The respective Chinese studies, as well as the Korean study indicated that rs1800896 was not a highly polymorphic SNP in those populations, with the A allele being the major allele and having frequencies higher than 90%. The UK white, Spanish and Sicilian populations also had the A and G alleles at approximately equal frequencies (A allele frequency = 0.52, 0.56 and 0.45 respectively). The G allele was the major allele only in Brazilians.

Four populations, namely the South African Coloureds, Brazilians, Gambians and Turkish had allele frequency data for rs1800871. The T allele of this SNP was the major allele only in the South African Coloured population. The Brazilian and Turkish allele frequencies were the inverse of the South African Coloureds, with the C allele as major allele. Allele frequencies of rs1800872 were available in five populations. The South African Coloureds shared the same minor allele (A) and similar allele frequencies with the Turkish. The Koreans were the only population with the A allele as major

allele. Only two studies had allelic information for rs3024496. This SNP was highly polymorphic in the South African Coloured population, but not in Koreans.

TABLE 38. Comparison of allele frequencies of *IL10* markers between populations.

Population	rs1800896		rs1800871		rs1800872		rs3024496	
	A	G	T	C	A	C	C	T
South African Coloured	0.68	0.32	0.66	0.34	0.34	0.66	0.48	0.52
Cambodian [190]	0.67	0.33	-	-	-	-	-	-
Brazilian [424]	0.31	0.69	0.34	0.66	-	-	-	-
Chinese [425]	0.96	0.04	-	-	-	-	-	-
Gambian [194]	0.66	0.34	0.47	0.53	0.47	0.53	-	-
Hong Kong Chinese [183]	0.96	0.04	-	-	-	-	-	-
Korean [418]	0.92	0.08	-	-	0.67	0.33	0.04	0.96
Sicilian [171]	0.45	0.55	-	-	-	-	-	-
Spanish [179]	0.56	0.44	-	-	-	-	-	-
Turkish [189]	0.77	0.23	0.33	0.67	0.33	0.67	-	-
Ugandan [426]	-	-	-	-	0.42	0.58	-	-
UK white [425]	0.52	0.48	-	-	-	-	-	-

We also compared frequencies of the 3 SNP haplotype consisting of rs1800872, rs1800871 and rs1800896 (Table 39). This haplotype has been studied repeatedly in numerous populations and in most cases only 3 haplotypes, namely CCA, ATA and CCG were observed [419]. This is also true for the South African Coloured population. All three of these haplotypes were present at approximately the same frequencies in the South African Coloureds (0.32-0.34). The minor haplotype in Koreans and Hong Kong Chinese was CCG, two populations which also differ markedly in the allele frequencies of the other polymorphisms in *IL10* when compared to the rest. The GTA haplotype is rarely identified in populations and was present in Estonian whites [417].

TABLE 39. Comparison of 3 SNP haplotype frequencies of *IL10* between populations.

Population	Ref	1	2	3	4
		C-C-A ^a	A-T-A	C-C-G	G-T-A
South African Coloured		0.34	0.34	0.32	-
Estonian whites	[417]	0.26	0.28	0.45	0.01
Hong Kong Chinese	[183]		0.92 ^b	0.08	-
Koreans	[418]	0.23	0.70	0.07	-
Kwazulu-Natal	[427]	0.28	0.34	0.38	-
Spanish	[416]	0.37	0.27	0.36	-
Turkish	[189]	0.44	0.33	0.23	-

^a The SNPs are listed in the following order: rs1800872, rs1800871 and rs1800896.

^b Frequency of combined CCA and ATA.

4.4.2. Discussion

Polymorphisms in *IL10* have been studied in various diseases, including TB. However, association results from these TB studies have been conflicting. We attempted to replicate previous association studies in other populations and to elucidate the LD and haplotype structure of this gene in the South African Coloured population.

We found no association between the rs1800896 SNP, which has previously been shown to influence IL-10 production, and TB susceptibility in the case-control or TDT analysis. This result is similar to findings in Korean [418], Gambian [194] and Spanish populations [179], even though the polymorphism was previously associated with TB in Cambodians [190], Sicilians [171] and Turkish [189]. The rs1800871 and rs1800872 SNPs, also previously associated with TB disease, were not susceptibility factors in this study either. However, these SNPs were in complete LD with each other in the South African Coloured population, as was found in other populations [194,418]. None of the remaining SNPs were associated with disease in the case-control or TDT analysis.

Haplotype analysis of the SNPs genotyped in *IL10* showed a single association with one haplotype in the TDT analysis. This association remained significant after permutation testing. However, this haplotype was not associated with disease in the case-control study. This seemingly conflicting result could be explained by the smaller sample size available for the family-based association test. Since the case-control association study had a bigger sample size, and therefore higher power than the TDT, it is probable that this association is not true and that, given the availability of more TB families to genotype, it would disappear.

Since a three SNP haplotype of *IL10* was previously associated with TB, we also tested this in the South African Coloured population. In contrast to studies done in Gambians, Spanish, Sicilians and Koreans [171,179,194,418], there was no association evident in the South African Coloured population. The CCG haplotype has previously been associated with higher, and the ATA haplotype with lower, IL-10 production [428]. A four SNP haplotype, previously associated with TB in Koreans [418], was also not associated with disease in this study. Our comparison of allele frequencies and haplotypes of *IL-10* in different populations highlights the genetic diversity that exists between populations.

In conclusion, our results suggest that *IL10* polymorphisms and haplotypes are not susceptibility determinants for TB in the South African Coloured population.

4.5. MEMBERS OF THE INTERLEUKIN-12 FAMILY AND THEIR RECEPTORS

Interleukin-12 (IL-12, IL12-p70), IL-12(p40)₂ (IL12p80) and interleukin-23 (IL-23) are all dimeric members of the IL-12 family [429]. The heterodimers IL-12 and IL-23 share the p40 chain (encoded by the interleukin-12, beta (*IL12B*) gene [430]), but each cytokine also contains a unique subunit, namely p35 [431] and p19 [432] respectively, while the homodimer IL-12(p40)₂ is composed of two p40 chains [433,434] (Figure 25). In order for these cytokines to have any biological activity, they need to bind to their specific receptors. The high affinity interleukin-12 receptor complex (IL-12R) [435] consists of two subunits, namely interleukin-12 receptor, beta 1 (IL-12Rβ1) and interleukin-12 receptor, beta 2 (IL-12Rβ2) [435], encoded by the *IL12RB1* [436,437] and *IL12RB2* [438] genes respectively. IL-23 also binds to IL-12Rβ1, but not measurably to IL-12Rβ2 [432], and attaches to the IL-23R subunit instead [439]. The respective heterodimeric receptors of IL-12 and IL-23 share the IL-12Rβ1 chain and it is likely that IL-12(p40)₂ may also bind to it [440]. Even though IL-12, IL-12(p40)₂ and IL-23 all utilise these cytokine and receptor subunits [441] (Figure 25), each cytokine has specific and diverse functions.

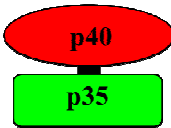
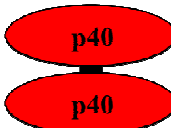
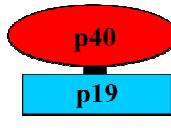
Cytokine	IL-12	IL-12(p40) ₂	IL-23
			
Receptor			
Binding chain	IL-12Rβ1	IL-12Rβ1	IL-12Rβ1
Signalling chain	IL-12Rβ2	?	IL-23R

FIGURE 25. Selected members of the IL-12 family and their receptors. Adapted from [429].

IL-12 is a key promoter of the Th1 immune responses and plays a vital role in the protective cell-mediated immunity against *M.tuberculosis* as part of the IL-12/IL-23/IFN-γ pathway [442]. This cytokine is released from macrophages after phagocytosis of the bacterium and induces IFN-γ production by Th1 lymphocytes, which in turn will activate macrophages (Figure 20). A positive feedback loop is therefore created which will enhance the antimicrobial activity of the macrophages and increase killing of the mycobacterium [443,444]. IFN-γ production is reduced in the absence of IL-12 [445]. It has been shown in mice that deficiency of the p35 chain is less detrimental to host response against TB than the lack of p40 [446,447]. IL-12(p40)₂ can not induce IFN-γ production in T cells, is an antagonist of IL-12 that binds to IL-12Rβ1 and blocks the IL-12p70 activation of T cells [443]. It is also a chemoattractant for macrophages [448], and may induce tumor necrosis factor-α [449] and nitric oxide synthase [450]. IL-12(p40)₂ binds to the IL-12R, but this interaction does not result in a biological response [433,434,451]. It has been shown that IL-12(p40)₂ is sufficient to ensure the *M.tuberculosis*-induced maturation of pulmonary dendritic cells to that of a T cell activating phenotype, but it is not known through which receptor it acts [452].

IL-23 was first thought to have similar functions to IL-12, since it induced the production of IFN- γ in human T cells and activated signal transducer and activator of transcription 4 (STAT4) [432,439]. However IL-23 can induce interleukin-17 production in T cells, something which IL-12 cannot [453]. The IL-23 p19 subunit has a negligible effect on TB progression in both early and chronic *M.tuberculosis* infection [453]. It can however reduce the severity of TB infection and may promote granuloma formation when IL-12 is absent [441].

MSMD (*as discussed in 4.1.*) is a rare syndrome and affected individuals are usually predisposed to clinical disease which is caused by normally nonpathogenic mycobacteria, but they are also susceptible to *M.tuberculosis* and *Salmonella*. Mutations in *IL12B* [50,342,343] and *IL12RB1* [45,46,344-346] of these patients are often the cause of this disease. More common polymorphisms in these genes may have an effect on the genetic control of *M.tuberculosis* in the general population [454]. Considering this hypothesis, the importance of the p40 and IL-12R β 1 subunits in TB, their effects on IFN- γ production and the conflicting findings of previous association studies done in other populations [191,444,454-457], *IL12B* and *IL12RB1* were selected to study as susceptibility factors for TB.

TABLE 40. Polymorphisms genotyped in the *IL12B*, *IL12RB1* and *IL12RB2* genes in the South African Coloured population.

Polymorphism	ClassLocation	Genotyping method
<i>IL12B</i>		
rs730691	SNP 5' UTR	SNPlex
rs3212217	SNP 5' UTR	SNPlex
rs3212220	SNP 5' UTR	SNPlex
D5S2941 ((ATT) _n)	STR ^a Intron 1	Capillary electrophoresis
rs3213096	SNP Exon 3, non-synonymous	SNPlex
rs2288831	SNP Intron 3, HapMap	SNPlex
rs2853696	SNP Intron 5, HapMap	SNPlex
rs3212227 (+1188/1159)	SNP 3' UTR	SNPlex
<i>IL12RB1</i>		
rs393548 (-111)	SNP Promoter	SNPlex
rs2305743	SNP Intron 2, HapMap	SNPlex
rs11086087 (5970)	SNP Exon 4, non-synonymous	SNPlex
rs429774	SNP Intron 6, HapMap	SNPlex
rs375947 (365/1094/1158)	SNP Exon 10, non-synonymous	SNPlex
<i>IL12RB2</i>		
rs11576006 (-1033)	SNP Promoter	SNPlex
rs3762317 (-465)	SNP Promoter	ARMS-PCR

^a short tandem repeat (STR)

We investigated polymorphisms from these genes (Table 40) to determine their involvement in TB susceptibility in the South African Coloured population. Promoter

polymorphisms in *IL12RB2* (Table 40) were also investigated, since this gene was previously associated with leprosy [458], an infectious disease caused by *Mycobacterium leprae*. Coding SNPs in this gene were previously shown to have no influence on mycobacterial infection [252,455], but the degree of expression of this gene, possibly regulated by promoter polymorphisms, could determine the intensity of the cell-mediated immune response to mycobacteria [458]. For this reason we included rs3762317, which disrupts a GATA transcription factor binding site [296], and rs11576006, which participates in the creation of another GATA site [296], in the genotyping experiments.

4.5.1. Results

4.5.1.1. Single-point statistical analysis

All polymorphisms were in HWE in controls. Bonferroni corrections were applied using the methodology of Nicodemus et al [76].

a) *IL12B*

Seven SNPs in *IL12B* were studied as susceptibility markers for TB in the South African Coloured population. After Bonferroni correction [76] the level of significance required was set at 0.025. Under this criterion, none of the analysed SNPs (Tables 41 and 42) reached statistical significance, even though rs2853696 had a nominally significant p value ($p = 0.04$). The 3'UTR SNP (rs3212227) [459,460], may influence gene expression levels and has previously been associated with TB in certain populations [247,456,461], but not in others [444]. This polymorphism was not associated with TB in this study either ($p_{\text{Case-control}} = 0.27$, $p_{\text{TDT}} = 1$). None of the remaining SNPs, which included 3 promoter variants, 1 nonsynonymous SNP and 2 intronic SNPs, were associated with TB in the case-control or TDT analyses. The rs3213096 and rs2853696 SNPs had extremely low minor allele frequencies in the South African Coloured population.

TABLE 41. Single-point statistical analysis in the *IL12B* case-control study.

Polymorphism	Allele		Genotype frequencies						Cases vs Controls p value ^c
			Cases (n ^a =432)			Controls (n=482)			
	1 ^b	2	11	12	22	11	12	22	
rs730691	T	C	0.24	0.49	0.27	0.26	0.50	0.24	0.51
rs3212217	C	G	0.09	0.36	0.55	0.06	0.38	0.56	0.18
rs3212220	T	G	0.10	0.40	0.50	0.06	0.44	0.49	0.11
rs3213096	A	G	0	0.003	0.997	0	0.004	0.996	0.82
rs2288831	C	T	0.10	0.40	0.50	0.06	0.45	0.49	0.12
rs2853696	A	G	0.003	0.114	0.883	0.01	0.16	0.83	0.04
rs3212227	C	A	0.09	0.40	0.51	0.06	0.44	0.50	0.23

^a number of investigated individuals.

^b Allele 1 is the minor allele.

^c p value from a genotype-based χ^2 test

TABLE 42. Single-point statistical analysis in the *IL12B* TDT.

SNP	Overtransmitted	Transmitted:Untransmitted	χ^2	p value
rs730691	C	52 : 41	1.30	0.25
rs3212217	G	36 : 32	0.24	0.63
rs3212220	T	43 : 40	0.11	0.74
rs3213096	- ^a	-	-	-
rs2288831	C	43 : 40	0.11	0.74
rs2853696	G	19 : 10	2.79	0.09
rs3212227	-	40 : 40	0	1

^a this SNP was not present in the family-based samples.

b) *IL12RB1*

The association of *IL12RB1* polymorphisms with TB susceptibility in the South African Coloured population were tested by genotyping 5 polymorphisms from the gene. The threshold for significance was set at $p = 0.013$ after Bonferroni correction, since 3 of the 5 genotyped SNPs were independent of the others. None of the SNPs were associated with TB in the case-control analysis (Table 43). The rs393548 SNP was significantly associated with TB in the TDT analysis (Table 44, $p = 0.013$). This polymorphism was previously associated with TB [454] in 101 Moroccan families, but was not associated with TB in the case-control study of the South African Coloured population. No other SNP attained significance in the TDT analysis.

TABLE 43. Single-point statistical analysis in the *IL12RB1* case-control study.

Polymorphism	Allele		Genotype frequencies						Cases vs Controls p value ^c
			Cases (n ^a =432)			Controls (n=482)			
	1 ^b	2	11	12	22	11	12	22	
rs393548	C	G	0.04	0.32	0.64	0.03	0.32	0.65	0.52
rs2305743	A	G	0.03	0.33	0.65	0.04	0.29	0.66	0.32
rs11086087	C	G	0.02	0.19	0.78	0.01	0.22	0.77	0.50
rs429774	C	T	0.12	0.44	0.44	0.12	0.44	0.44	0.95
rs375947	G	A	0.15	0.44	0.42	0.16	0.43	0.42	0.89

^a number of investigated individuals.

^b Allele 1 is the minor allele.

^c p value from a genotype-based χ^2 test.

TABLE 44. Single-point statistical analysis in the *IL12RB1* TDT.

SNP	Overtransmitted	Transmitted:Untransmitted	χ^2	p value
rs393548	G	39 : 20	6.12	0.013
rs2305743	G	25 : 20	0.56	0.46
rs11086087	G	23 : 22	0.02	0.88
rs429774	C	51 : 48	0.09	0.76
rs375947	A	49 : 44	0.27	0.60

c) *IL12RB2*

Two promoter polymorphisms from *IL12RB2* were genotyped in this study, one with an ARMS-PCR genotyping method (Figure 26, section 3.6.1). There was no association between either of these SNPs and TB in the case-control (Table 45) or TDT (Table 46) analysis, since the threshold for significance was set at 0.025 after Bonferroni correction using the method of Nicodemus et al [76]. The rs3762317 SNP was previously associated with leprosy and determined the clinical type of the disease [458]. The G allele of this SNP disrupts a GATA site and leads to a threefold increase in promoter activity [296]. The rs11576006 SNP creates a GATA transcription factor binding site [296] together with the rs3762315 SNP, which could not be genotyped in this study due to technical issues.

TABLE 45. Single-point statistical analysis in the *IL12RB2* case-control study.

Polymorphism	Allele		Genotype frequencies						Cases vs Controls p value ^c
	1 ^b	2	Cases (n ^a =432)			Controls (n=482)			
			11	12	22	11	12	22	
rs11576006	C	T	0.12	0.50	0.38	0.11	0.44	0.45	0.07
rs3762317	G	A	0.18	0.57	0.24	0.17	0.51	0.31	0.08

^a number of investigated individuals.

^b Allele 1 is the minor allele.

^c p value from a genotype-based χ^2 test

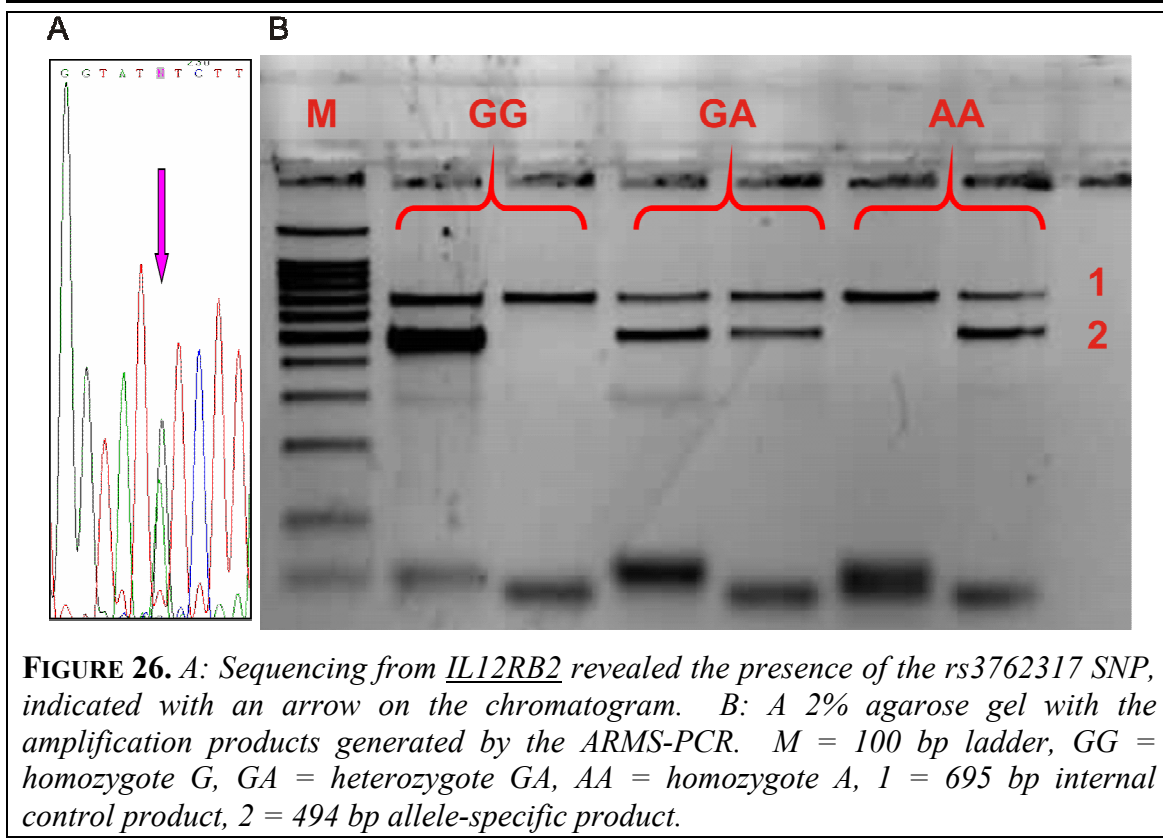


TABLE 46. Single-point statistical analysis in the *IL12RB2* TDT.

SNP	Overtransmitted	Transmitted:Untransmitted	χ^2	p value
rs11576006	T	44 : 40	0.19	0.66
rs3762317	A	22 : 20	0.095	0.76

4.5.1.2. *IL12B* microsatellite analysis

The D5S291 microsatellite, an (ATT)_n repeat, was previously associated with TB in the Hong Kong Chinese population [456]. In that study, and several other studies, only (ATT)₈ and (ATT)₉ alleles were observed. However, in the South African Coloured population we also detected (ATT)₇ and (ATT)₁₀ alleles (Figure 27) with the Genemapper version 3.7 software program (Applied Biosystems). The presence of an (ATT)₇ allele in two Swedish families was mentioned in only one other study [290], while the (ATT)₁₀ allele was novel. We verified the presence of these alleles by sequencing the microsatellite (Figure 28, section 3.5.2.).

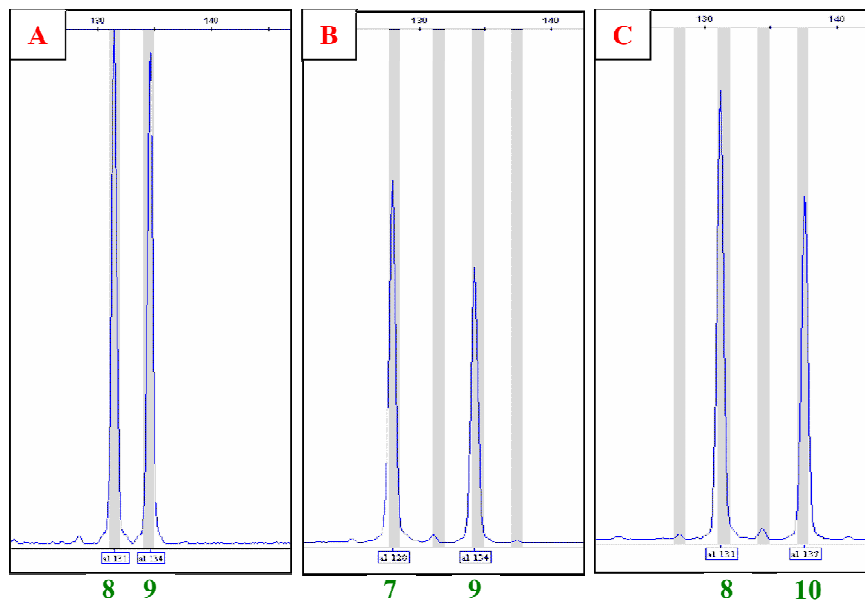


FIGURE 27. D5S291 trinucleotide repeat analysis with Genemapper version 3.7. The common (ATT)₈ and (ATT)₉ alleles can be seen in A. The (ATT)₇ allele is visible in B and the novel (ATT)₁₀ allele is present in C.

The number of ATT repeats from the D5S291 microsatellite marker was determined by direct counting (Table 47) and plotted as a distribution graph (Figure 29). Since this graph was bimodal, the alleles were divided into two subclasses, as previously described [323-325]. The shorter repeats, with (ATT)₇ and (ATT)₈, were designated as S alleles and the longer repeats, with (ATT)₉ and (ATT)₁₀, were designated as L alleles. The genotypic effect of the repeats on susceptibility to TB was evaluated by dividing the three genotypes (S/S, S/L and L/L) into two subgroups according to the presence or absence of the S allele: genotypes containing the S allele (S/S and S/L) and genotypes

without the S allele (L/L). Possible associations between the disease groups and classes of alleles were analysed for significance by the χ^2 test (Table 48). We did not find any association with the D5S2941 microsatellite and TB susceptibility using a genotype-based test ($p = 0.27$) or comparing genotypes with and without the S allele ($p = 0.17$) in the case-control samples. In the TDT the L allele was transmitted more frequently than expected to affected children, but this did not reach statistical significance (Table 49).

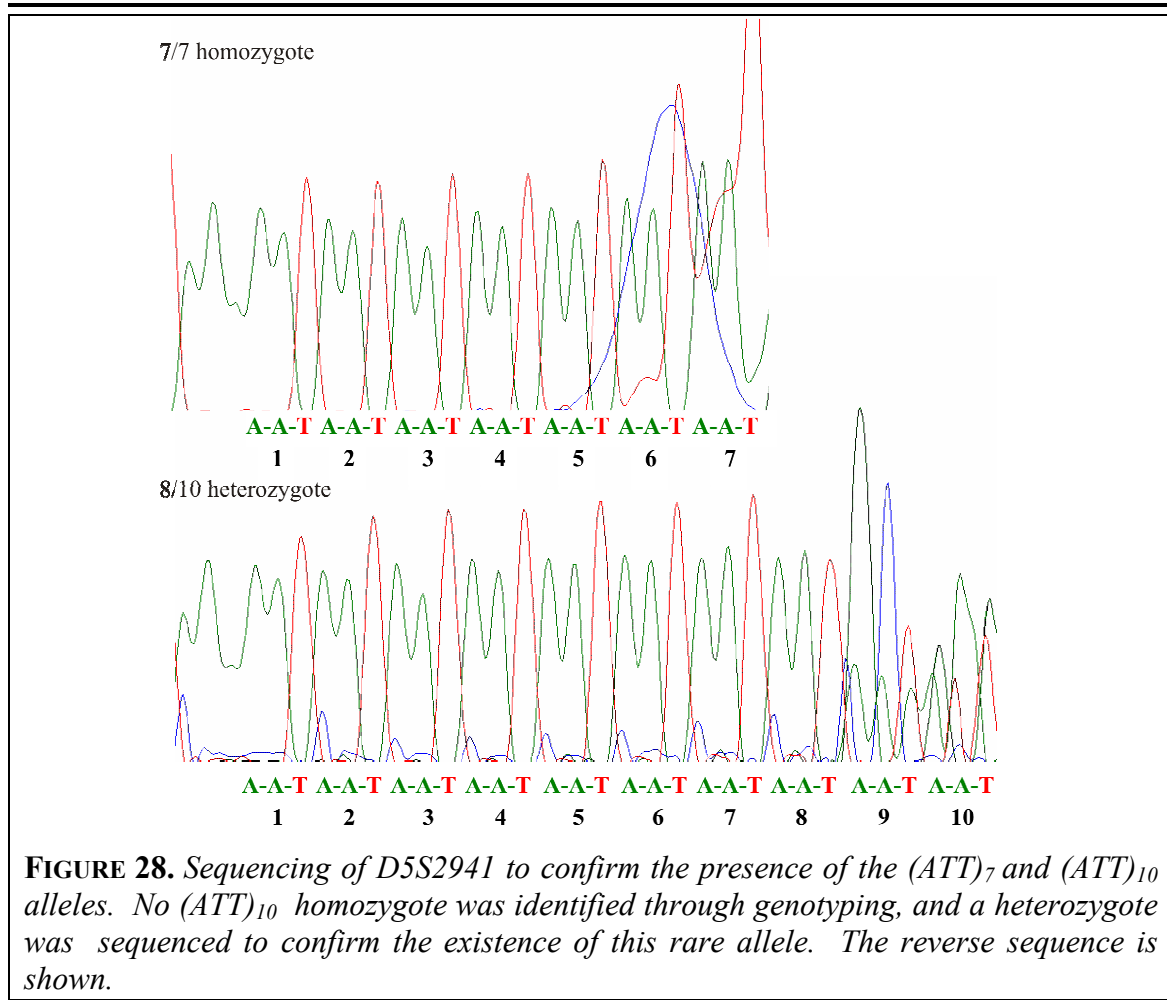


TABLE 47. Frequencies of the respective alleles of D5S2941 in the TB cases and controls.

Allele	Cases _{(n) alleles}	Cases _{Frequency}	Controls _{(n) alleles}	Controls _{Frequency}
(ATT) ₇	18	0.02	13	0.01
(ATT) ₈	581	0.68	633	0.69
(ATT) ₉	239	0.28	253	0.28
(ATT) ₁₀	16	0.02	13	0.01

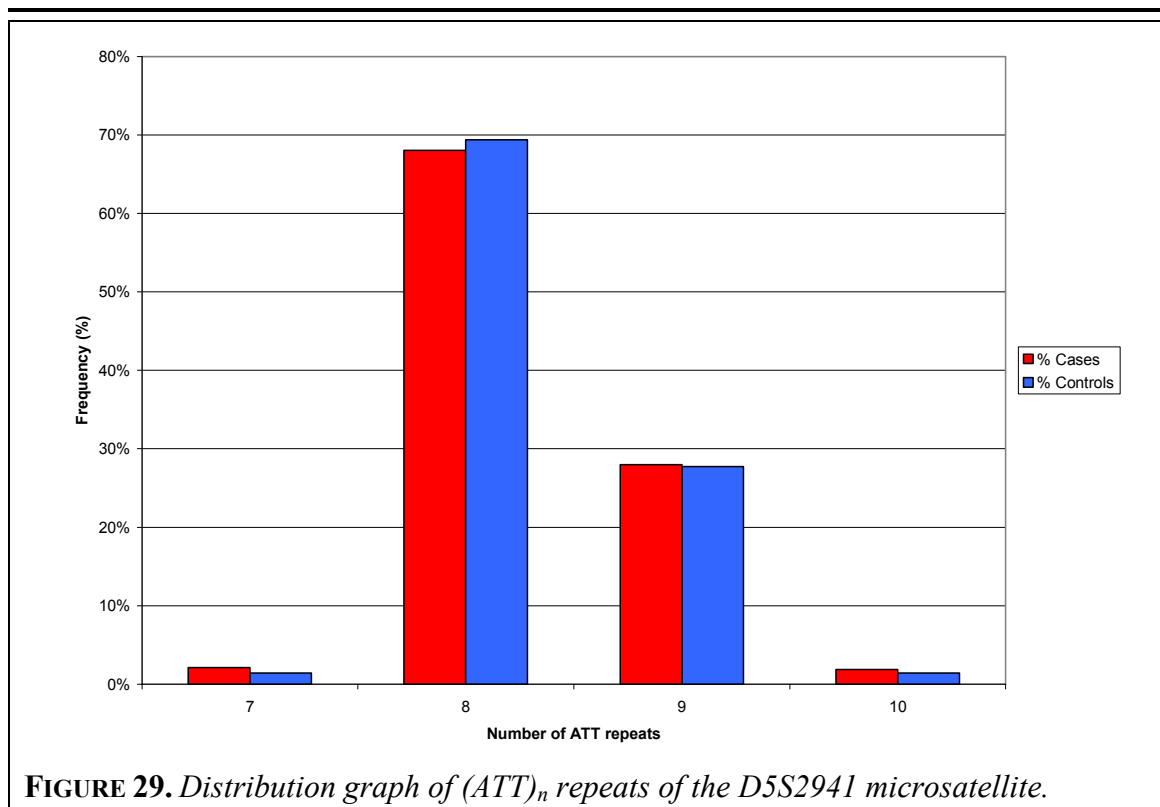


TABLE 48. Microsatellite analysis of D5S2941 in the case-control study.

D5S2941	Allele		Genotype frequencies						Cases vs Controls p value
			Cases (n ^a =432)			Controls (n=482)			
	1 ^b	2	11	12	22	11	12	22	
	L	S	0.09	0.41	0.50	0.08	0.45	0.47	0.27 ^c ; 0.17 ^d

^a number of investigated individuals.

^b Allele 1 is the minor allele.

^c p value from genotype-based χ^2 test.

^d p value of comparison of genotypes with and without S allele.

TABLE 49. D5S2941 S and L allele TDT analysis.

D5S2941	Overtransmitted	Transmitted:Untransmitted	χ^2	p value
	L	43 : 35	0.082	0.37

4.5.1.3. Haplotype analysis

Haplotypes in the respective genes were analysed with Haploview v3.3.

a) IL12B

D' values between markers in the IL12B gene were usually greater than 50%, indicating strong LD in this region (Figure 30). One haplotype block, which excluded rs3213096

due to its low allele frequency, was identified in *IL12B* (Figure 30). The rs730691 5'UTR SNP was also not part of this block in the South African Coloured population. Four haplotypes were estimated in the haplotype block. Haplotype 3 (Tables 50 and 51) was nominally associated with TB in both the case-control ($p = 0.02$) and TDT ($p = 0.02$) analyses. After 10 000 permutations, these associations became less significant ($p = 0.06$ and $p = 0.03$ respectively). According to the calculation of the program, 570 out of these 10 000 permutations exceeded the best observed χ^2 (haplotype 3, $\chi^2 = 5.76$, $p = 0.02$) in the case-control analysis. This means that there is a 6% chance that the significant nominal p value of haplotype 3 was observed due to chance alone. In permutation calculations using the family-based TDT data, 334 out of these 10 000 permutations exceeded the best observed χ^2 (haplotype 3, $\chi^2 = 5.76$, $p = 0.02$), which means that there is a 3% chance that this association is due to chance alone in the TDT analysis. Haplotype 3 was tagged by the A allele of rs2853696, a SNP with a low allele frequency (0.09) in the South African Coloured population.

TABLE 50. *IL12B* case-control haplotype analysis by Haploview [313].

Block 1 ^a		Freq _{Cases}	Freq _{Controls}	χ^2	p value	p value _{Permutation} ^{b, c}
1	A-G-T-S-G-G	0.63	0.62	0.49	0.49	0.89
2	C-G-C-L-T-C	0.25	0.24	0.21	0.64	0.99
3	A-A-T-S-G-G	0.06	0.09	5.29	0.02	0.06
4	C-G-C-L-T-G	0.03	0.04	1.12	0.29	0.73

^a The order of the SNPs in each block corresponds to Figure 30, with D5S2941 included and rs3213096 excluded.

^b Permutation test p values were calculated from 10 000 permutations in Haploview.

^c p values less than 0.02 were observed on 570 out of 10 000 occasions (6%).

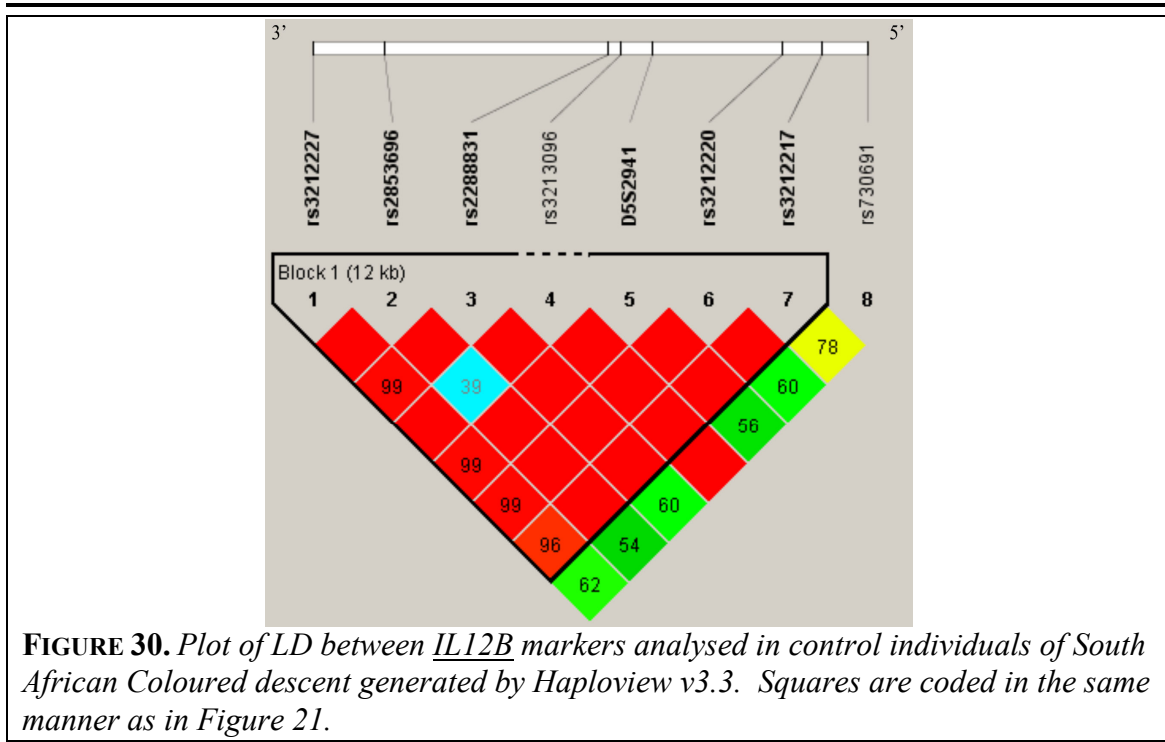


TABLE 51. *IL12B* TDT haplotype analysis by Haploview [313].

Block 1 ^a	Transmitted:untransmitted	χ^2	p	p _{Permutation} ^{b, c}	
1	A-G-T-S-G-G	33 : 25	1.11	0.29	0.51
2	C-G-C-L-T-C	24 : 25	0.02	0.90	1
3	A-A-T-S-G-G	5 : 16	5.76	0.02	0.03
4	C-G-C-L-T-G	5 : 3	0.43	0.51	1

^a The order of the SNPs in each block corresponds to Figure 30.

^b Permutation test p values were calculated from 10 000 permutations in Haploview.

^c p values less than 0.02 were observed on 334 out of 10 000 occasions (3%)

b) *IL12RB1*

The confidence interval definition of haplotype blocks according to Gabriel et al [89] indicated only one haplotype block consisting of two polymorphisms (rs429774 and rs11086087) of the five *IL12RB1* SNPs genotyped in this study (Figure 31). This is due to the varying degrees of LD between the SNPs of *IL12RB2* (Figure 31). Three haplotypes were estimated in this single block in both the case-control (Table 52) and TDT (Table 53) analyses. Haplotype 1 (Table 52) was the most frequent haplotype, while haplotype 3 had the lowest frequency in both cases and controls. Haplotype 1 was transmitted more frequently than the others in the TDT study (Table 53), but none of the haplotypes were associated with TB susceptibility in the South African Coloured population.

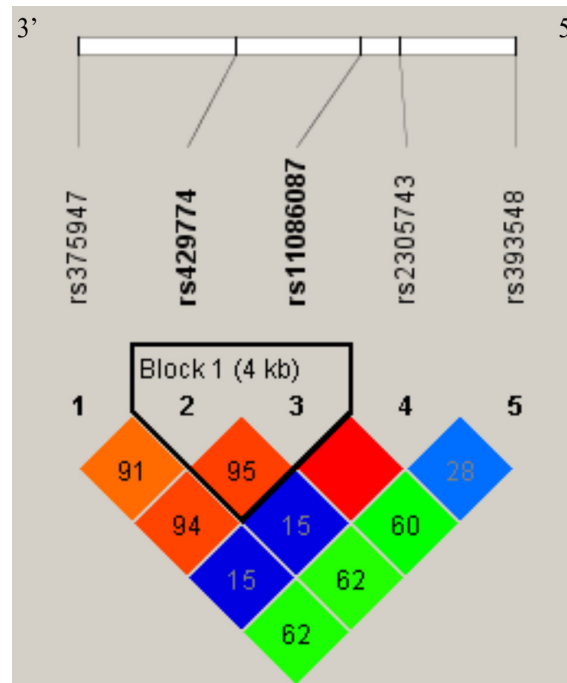


FIGURE 31. Plot of LD between *IL12RB1* markers analysed in control individuals of South African Coloured descent generated by Haploview v3.3. Squares are coded in the same manner as in Figure 21.

TABLE 52. *IL12RB1* case-control haplotype analysis by Haploview [313].

Block 1 ^a		Freq _{Cases}	Freq _{Controls}	χ^2	p value	p value _{Permutation} ^{b, c}
1	T-G	0.66	0.66	0.07	0.79	0.99
2	C-G	0.24	0.22	1.10	0.29	0.59
3	C-C	0.11	0.12	0.87	0.35	0.69

^a The order of the SNPs in each block corresponds to Figure 31.

^b Permutation test p values were calculated from 10 000 permutations in Haploview.

^c p values less than 0.29 were observed on 5910 out of 10 000 occasions (59%).

TABLE 53. *IL12RB1* TDT haplotype analysis by Haploview [313].

Block 1 ^a		Transmitted:untransmitted	χ^2	p	p _{Permutation} ^{b, c}
1	T-G	39 : 31	0.92	0.34	1
2	C-G	19 : 26	1.25	0.26	0.42
3	C-C	18 : 18	0.01	0.94	1

^a The order of the SNPs in each block corresponds to Figure 31.

^b Permutation test p values were calculated from 10 000 permutations in Haploview.

^c p values less than 0.26 were observed on 4230 out of 10 000 occasions (42%).

c) *IL12RB2*

The 2 promoter polymorphisms genotyped in *IL12RB2* were in weak LD with each other, and this is why Haploview did not create a haplotype block for them (Figure 32). However, we did evaluate the combined effect of these SNPs in a haplotype in both the case-control (Table 54) and TDT (Table 55) analyses. Haplotype 2 (Table 54) in the case-control haplotype analysis did show a nominal association with TB (p = 0.05), but after permutation testing this association was no longer significant (p = 0.08). Haplotype 2 was most frequently transmitted to affected offspring in the TDT (Table 55), while haplotype 1 had the highest frequency in cases and controls (Table 54). We did not find any association between TB susceptibility and the remaining haplotypes in the case-control or family samples.

TABLE 54. *IL12RB2* case-control haplotype analysis by Haploview [313].

Block 1 ^a		Freq _{Cases}	Freq _{Controls}	χ^2	p value	p value _{Permutation} ^{b, c}
1	T-G	0.33	0.34	0.11	0.75	0.97
2	T-A	0.29	0.33	3.88	0.05	0.08
3	C-A	0.25	0.22	1.27	0.26	0.49
4	C-G	0.13	0.11	3.35	0.07	0.11

^a The order of the SNPs in each block corresponds to Figure 32.

^b Permutation test p values were calculated from 10 000 permutations in Haploview.

^c p values less than 0.05 were observed on 764 out of 10 000 occasions (8%).

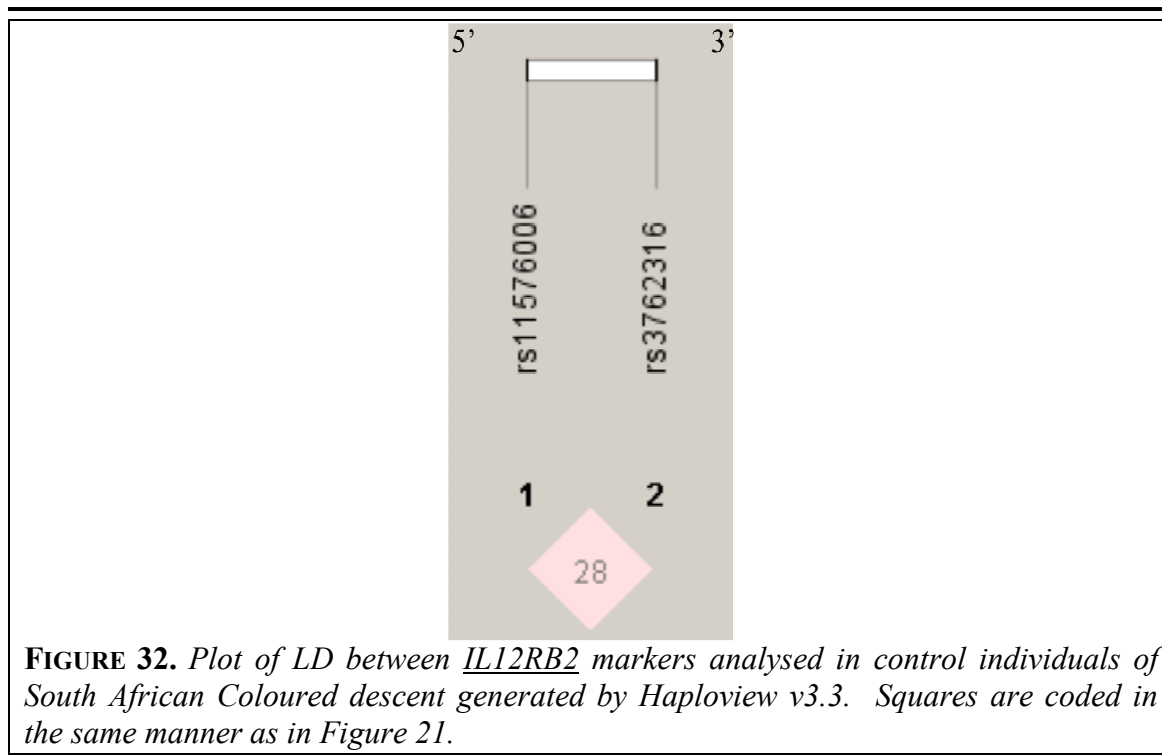


TABLE 55. *IL12RB2* TDT haplotype analysis by Haploview [313].

Block 1 ^a	Transmitted:untransmitted	χ^2	p	p _{Permutation} ^{b, c}	
1	T-G	29 : 25	0.36	0.56	0.85
2	T-A	40 : 31	1.28	0.26	0.41
3	C-A	21 : 31	2.27	0.13	0.18
4	C-G	10 : 13	0.40	0.53	0.82

^a The order of the SNPs corresponds to Figure 32.

^b Permutation test p values were calculated from 10 000 permutations in Haploview.

^c p values less than 0.13 were observed on 1811 out of 10 000 occasions (18%).

4.5.1.4. Comparison of allele frequencies between different populations

We compared allele frequencies of *IL12B* (Table 56), *IL12RB1* (Table 57) and *IL12RB2* (Table 58) polymorphisms between different populations using previously published data [191,252,296,444,454-456,458-465].

The 3'UTR rs3212227 polymorphism from *IL12B* (Table 56) has been genotyped in diverse populations. The C allele was the minor allele in practically all populations, although it had a 50% allele frequency in Japanese, which was similar to the Hong Kong Chinese frequency. The populations from African descent (South African Coloured, Cameroon, African American) also had higher frequencies of the A allele in comparison with populations from European descent. The European populations, e.g. the British, German, Greek and Spanish populations, had similar frequencies as well.

The South African Coloured and Moroccan studies showed similar frequencies for the rs393548 promoter SNP and for rs11086087 in *IL12RB1*. The frequency of the non-synonymous rs375947 polymorphism did not differ widely between populations, with the G allele being the minor allele in all population listed (Table 57).

The rs11576006 and rs3762317 SNPs (Table 58) of *IL12RB2* had lower allele frequencies in the European population than in the South African Coloured population. The frequency of rs3762317 was also slightly lower in the Japanese population (0.39 vs 0.43).

TABLE 56. Comparison of allele frequencies of *IL12B* SNPs between populations.

Population	Ref	rs3212227	
		C	A
South African Coloured		0.28	0.72
African American	[444]	0.33	0.67
British	[462]	0.21	0.79
Cameroon	[459]	0.38	0.62
European	[460]	0.18	0.82
German	[463]	0.21	0.79
Greek	[459]	0.21	0.79
Hong Kong Chinese	[456]	0.45	0.55
Indian	[461]	0.42	0.58
Japanese	[464]	0.50	0.50
Spanish	[465]	0.21	0.79
White American	[444]	0.15	0.85

TABLE 57. Comparison of allele frequencies of *IL12RB1* SNPs between populations.

Population	Ref	rs393548		rs2305743		rs11086087		rs429774		rs375947	
		C	G	A	G	C	G	C	T	G	A
South African Coloured		0.19	0.81	0.19	0.81	0.12	0.88	0.34	0.66	0.37	0.63
Japanese	[191]	-	-	-	-	-	-	-	-	0.39	0.61
Japanese	[455]	-	-	-	-	-	-	-	-	0.36	0.64
Korean	[252]	-	-	-	-	-	-	-	-	0.34	0.66
Moroccan	[454]	0.17	0.83	-	-	0.13	0.87	-	-	0.26	0.74

TABLE 58. Comparison of allele frequencies of *IL12RB2* SNPs between populations.

Population	Ref	rs11576006		rs3762317	
		C	T	G	A
South African Coloured		0.33	0.67	0.43	0.57
European	[296]	0.21	0.79	0.21	0.79
Japanese	[458]	-	-	0.39	0.61

4.5.2. Discussion

Members of the interleukin-12 cytokine family are known to affect IFN- γ production and are, together with their related receptors, necessary for the regulation of T cell responses. We tested the association of common polymorphisms in genes of this cytokine family and their receptors with TB, and evaluated the known allele frequencies of SNPs in these genes between populations.

4.5.2.1. *IL12B*

IL12B encodes the p40 subunit, a chain which is shared by all members of the IL-12 cytokine family. These p40-dependent molecules regulate Th1 responses [429] and are either involved in the host control of *M.tuberculosis* infection [466] (IL-12 and IL-23) or may contribute to the disease by acting as an antagonist (IL-12(p40)₂) [443]. Several variant sites in *IL12B*, including an insertion/deletion polymorphism [467], an intronic microsatellite (D5S2941) [468] and a 3'UTR SNP (rs3212227) [459,460], may influence gene expression levels. These polymorphisms have previously been associated with susceptibility to a variety of diseases [461], such as diabetes [290,462,468], malaria [469], TB [456,461], psoriasis [470] and hepatitis C virus infection [471].

We found a nominally significant association between a haplotype in *IL12B* and resistance to TB in the South African Coloured population. The haplotype occurred more in controls than in cases ($p = 0.02$, OR = 1.53, 95%CI [1.07-2.02]) and was significantly untransmitted to affected offspring in the family-based study ($p = 0.02$). Permutation testing decreased the significance of these associations ($p = 0.06$ and $p = 0.03$ respectively), but revealed that there were very slim chances of these associations existing due to chance alone. The haplotype was tagged by the A allele of the rs2853696 SNP, which was not associated with TB on its own. This suggests that this SNP is in LD with the true susceptibility variant.

The 3'UTR SNP rs3212227 was not associated with TB in the South African Coloured population. It was previously associated with TB in Japanese [247], Hong Kong Chinese [456] and Indian [461] populations, but not in white Americans or African Americans [444]. Kusuhara et al [191] considered 1 intronic SNP (rs2288831) and two SNPs downstream of *IL12B* (rs11135058 and rs6870828) in Japanese, but did not detect any association between them and TB. The non-synonymous rs3213096 was found at an extremely low frequency of 0.002 in the South African Coloured population. This variant, which changes a serine to asparagine at position 226 of the IL-12 protein, was

first identified in a Japanese population [464], but in that study it was not associated with asthma and allergic rhinitis. In the Japanese study an allele frequency of 0.04 was detected. Since our study had the power to detect an odds ratio of 2.15 with 95% confidence given an allele frequency of 0.05 based on the sample size available, we can only exclude the possibility that the nonsynonymous rs3213096 polymorphism plays a substantial role in TB susceptibility.

We identified the novel (ATT)₁₀ allele (which was confirmed by sequencing) as well as the (ATT)₇, (ATT)₈ and (ATT)₉ alleles of the D5S2941 microsatellite. However, none of these were associated with TB in the South African Coloured population. This microsatellite, in intron 2 of *IL12B*, was previously investigated in TB [456] and diabetes [468]. In these studies, which were done in Caucasian-Americans and Hong Kong Chinese, only the (ATT)₈ and (ATT)₉ alleles of this STR were ever observed, although the presence of the (ATT)₇ allele in two Swedish families was mentioned in another diabetes study [290]. There are no reports in the literature which studied this marker in African populations. Since the (ATT)₇, (ATT)₈ and (ATT)₉ alleles were identified in individuals from European and Asian descent, we speculate that the (ATT)₁₀ allele is a genetic contribution from the African parental population of the South African Coloureds.

When comparing the association results found between TB and *IL12B* polymorphisms, conflicting results are apparent. This could be explained by population differences, where *IL12B* may contribute to TB susceptibility to a larger extent in Asian populations, than in European or African populations.

4.5.2.2. *IL12RB1*

Mutations in *IL12RB1* are the most common genetic cause of MSMD with more than 61 MSMD patients from 18 different countries presenting with defects in this gene [454]. IL-12Rβ1, encoded by *IL12RB1*, is the binding chain in the IL-12R and expression of this protein is necessary for interaction of IL-12 with IL-12R [443].

We tested the association of five SNPs in *IL12RB1* with TB in the South African Coloured population using both case-control and family-based association studies. The G allele of the rs393548 SNP was preferentially transferred to affected children in the single-point TDT analysis, although no haplotype was associated with TB in the TDT. This polymorphism was previously associated with TB in 101 Moroccan families [454]. In our case-control study, none of the SNPs evaluated or haplotypes estimated in this gene was associated with TB disease. The apparently contradictory results obtained from the case-control and TDT analysis in the South African Coloureds could be explained by the smaller sample size of the family-based samples. This would mean that the case-control association study had a higher power to detect potential associations than the TDT analysis and that the TDT result is likely to be a false positive.

Akahoshi et al [455] reported the first case-control association study assessing *IL12RB1* in TB susceptibility in the general population. This was done in a Japanese population with 98 TB cases and 197 healthy controls. Two common haplotypes (consisting of 4 SNPs, of which rs375947 was genotyped in our study) were identified, and

homozygosity for the allele 2 haplotype was significantly associated with TB (OR 2.4, 95% CI 1.20-4.99, $p = 0.013$). Healthy subjects with this haplotype had lower levels of IL-12-induced signalling. Remus et al [454] investigated *IL12RB1* in 101 Moroccan families where two promoter polymorphisms in strong LD with each other (rs436857 and rs393548) were associated with disease. However, no association was detected in their study with the haplotype reported by Akahoshi et al [455] in the Japanese population. Lee et al [252] did a case-control study (115 TB patients, 151 controls) in Koreans and investigated 5 SNPs, but no significant differences were observed between cases and controls in terms of the genotypes or haplotypes. A study to investigate the effect of *IL12RB1* polymorphisms on susceptibility to and severity of TB was done in 87 Japanese TB patients and 265 controls [191]. Two intronic SNPs were associated with disease, and a haplotype consisting of a different combination of SNPs to the one identified by Akahoshi et al [455], was associated with resistance to TB. This study did not replicate the association found with the promoter polymorphisms in the Moroccan families [191]. The most recent study to consider polymorphisms and haplotypes in *IL12RB1* was done in an Indonesian population with 314 TB patients and 316 controls [457], but no association was observed for the SNPs or haplotypes tested with TB susceptibility.

This study in the case-control samples of the South African Coloured population, as well as the Korean [252] and Indonesian [457] studies which did not detect any association with *IL12RB1*, considered larger sample sizes than the two positive reports from Japanese studies [191,455]. The association found in the Japanese could be population-specific, but it could also be indicative of a false positive result. For this reason, replication of those results should be attempted in an independent, larger Japanese population [457].

4.5.2.3. *IL12RB2*

IL12RB2 encodes the signalling chain of the IL-12R and is essential for IL-12 responsiveness [435,472]. Ohyama et al [458] determined that SNPs in the 5' flanking region of *IL12RB2* partially determined the clinical types of leprosy and that haplotypes in this region which carried more than one SNP showed a lower transcriptional activity than the wild type. We tested two promoter polymorphisms to determine their possible involvement with TB susceptibility in the South African Coloured population, since the promoter region of *IL12RB2* has previously been shown to affect the transcriptional activity of the gene [296,458].

We did not detect a significant association between TB and the rs11576006 or rs3762317 SNPs. These specific promoter polymorphisms have not previously been studied in any TB association study, but were associated with leprosy. No haplotypes were associated with disease after correcting for multiple tests. Akahoshi et al [247] genotyped a single SNP (rs12142823) in *IL12RB2* in Japanese TB cases and controls, but did not find any association between it and TB. Kusahara et al [191] considered 5 intronic SNPs of *IL12RB2* in a Japanese population, but did not detect any association of this region with TB either.

In conclusion, from our results it is apparent that promoter polymorphisms from *IL12RB2* are not susceptibility factors for TB in the South African Coloured population.

4.6. INTERLEUKIN-18

Interleukin-18 (IL-18) is a proinflammatory cytokine and, together with IL-12, one of the primary inducers of IFN- γ production by T cells [473,474]. The cytokine plays a critical role in the cell-mediated immunity against several pathogens, including intracellular organisms such as *Leishmania*, *Salmonella* and *M.tuberculosis* [475]. In mice IL-18 may influence the course of mycobacterial infection and other cytokines, such as IL-12, can not atone for a deficiency of this cytokine [476]. IL-18 can, in turn, not act independently of the IL-12 system [477], even though they have their own receptors and stimulate different signalling pathways [478]. Peripheral blood mononuclear cells from human TB patients produce less IL-18 than those from healthy individuals and alveolar macrophages secrete IL-18 when exposed to *M.tuberculosis* [473,479-481]. However, the production of IL-18 needs to be tightly regulated, since an excess of this cytokine may lead to topical or systemic injuries in the host [473,480-482]. IL-18 is also involved in the production of interleukin-32 in response to *M.tuberculosis* [483].

Despite intensive resequencing of *IL18*, the gene encoding IL-18, to date no non-synonymous SNPs have been identified. Initially a synonymous SNP in exon 4 of the gene (rs549908) was the main focus of investigation in association studies [484] and studies using monocytes from healthy individuals showed that this SNP may influence *IL18* transcription *in vitro* [485]. However, several polymorphisms in the promoter region of the gene have also been implicated in various diseases, such as sarcoidosis [486], allergic rhinitis [487], arthritis [488] and HIV [489]. The rs1946518 promoter SNP was previously associated with sarcoidosis in Japanese patients [486], but this finding was not verified in a Dutch [490] or independent Japanese population [491]. Other *IL18* polymorphisms (rs187238 and rs189667) were also investigated in Dutch sarcoidosis samples, but no association was observed with disease. Zhou et al showed that the promoter activity of a common haplotype was significantly higher than that of other haplotypes in sarcoidosis patients [491]. In fact, the functionality of the promoter polymorphisms has been proven previously in a multiple sclerosis study [492]. However, the variants studied were not associated with multiple sclerosis [492].

TABLE 59. *Polymorphisms genotyped in IL18 in this study.*

Polymorphism	Class	Location	Genotyping method
rs1946519	SNP	Promoter	SNPlex
rs1946518 (-607)	SNP	Promoter	SNPlex
rs187238 (-137)	SNP	Promoter	SNPlex
rs5744229	SNP	Exon/intron boundary 1	SNPlex
rs189667 (1248)	SNP	Intron 1	SNPlex
rs549908 (105)	SNP	Exon 4, synonymous	SNPlex

To date only one other association study between polymorphisms in *IL18* and susceptibility to TB has been published, namely a recent study by Kusuvara et al which considered 21 candidate genes including *IL18* [191]. Amongst others, they genotyped 6 SNPs spread throughout *IL18*, but did not consider the functional promoter or

synonymous polymorphisms, and found no association between this gene and TB susceptibility.

We selected functional promoter polymorphisms, the synonymous SNP and other variants (Table 59) from the gene to genotype in TB patients and controls from the South African Coloured population.

4.6.1. Results

4.6.1.1. Single-point statistical analysis

All the SNPs were in HWE for the controls. The threshold for significance was set at 0.05, since only one LD block was determined in Haploview (Figure 33). However, none of the analysed SNPs were significantly associated with TB in either the case-control (Table 60) or TDT (Table 61) study.

The rs1946518 and rs187238 SNPs were not associated with TB in the South African Coloured population. The synonymous rs549908 was also not associated with TB in our study, even though this SNP may influence *IL18* transcription *in vitro* [485].

TABLE 60. Single-point statistical analysis in the *IL18* case-control study.

Polymorphism	Allele		Genotype frequencies						Cases vs Controls p value ^c
			Cases (n ^a =432)			Controls (n=482)			
	1 ^b	2	11	12	22	11	12	22	
rs1946519	A	C	0.11	.50	0.39	0.14	0.48	0.38	0.46
rs1946518	T	G	0.11	0.49	0.39	0.14	0.48	0.38	0.43
rs187238	C	G	0.02	0.23	0.75	0.01	0.24	0.75	0.71
rs5744229	A	G	0.01	0.06	0.93	0.01	0.07	0.93	0.85
rs189667	G	A	0.02	0.23	0.75	0.01	0.24	0.75	0.48
rs549908	G	T	0.02	0.22	0.76	0.01	0.23	0.76	0.88

^a number of investigated individuals.

^b Allele 1 is the minor allele.

^c p value from a genotype-based χ^2 test

TABLE 61. Single-point statistical analysis in the *IL18* TDT.

SNP	Overtransmitted	Transmitted:Untransmitted	χ^2	p value
rs1946519	T	45 : 44	0.01	0.92
rs1946518	T	49 : 45	0.17	0.68
rs187238	C	28 : 21	1.00	0.32
rs5744229	G	5 : 1	2.67	0.10
rs189667	G	25 : 16	1.98	0.16
rs549908	G	27 : 20	1.04	0.31

4.6.1.2. Haplotype analysis

Only one haplotype block was identified for the polymorphisms genotyped in *IL18* (Figure 33). The rs5744229 SNP was excluded, since its allele frequency was less than 0.05. There was no association between any of the polymorphisms studied and susceptibility to TB in the case-control (Table 62) or TDT studies (Table 63). Haplotype 1 was the most frequent haplotype in cases, controls and samples from the families.

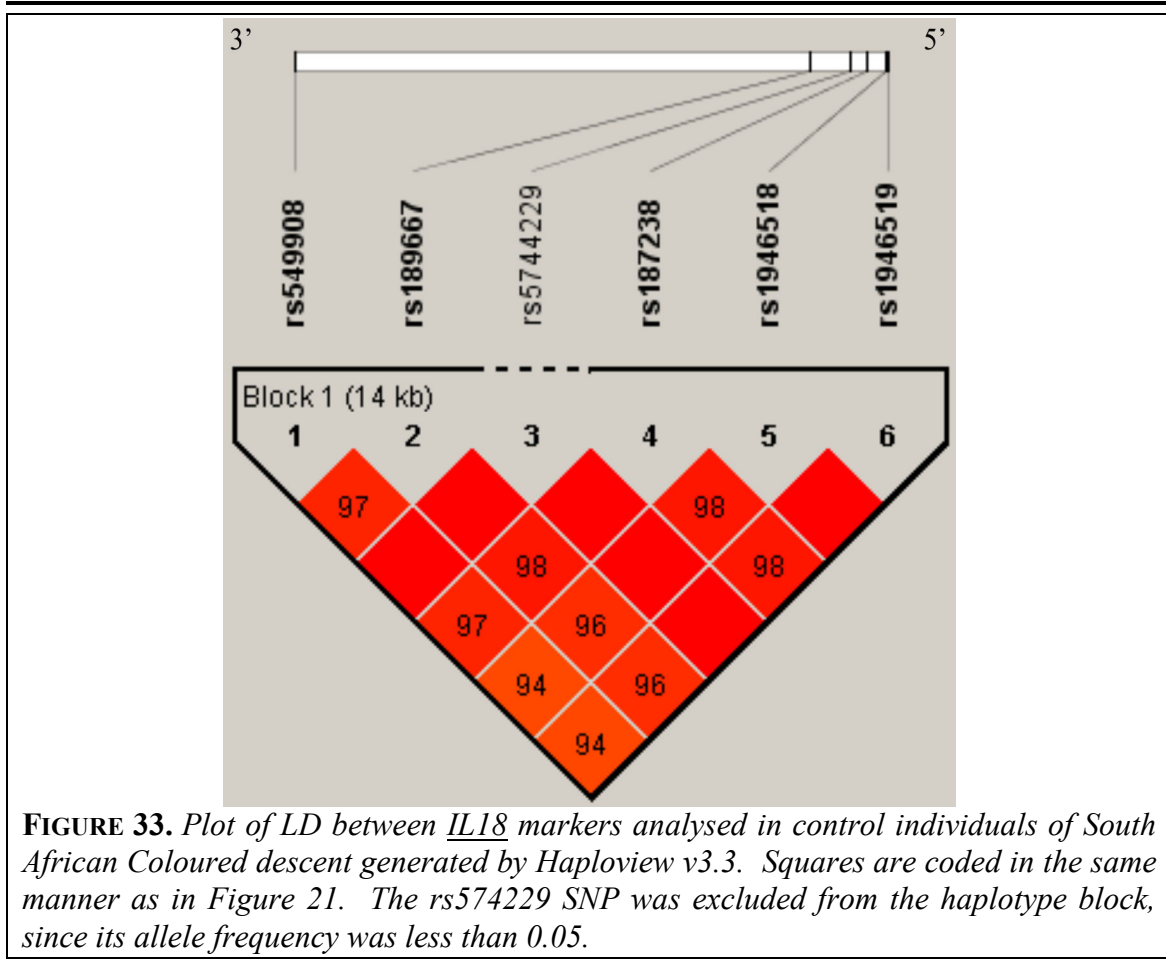


FIGURE 33. Plot of LD between *IL18* markers analysed in control individuals of South African Coloured descent generated by Haploview v3.3. Squares are coded in the same manner as in Figure 21. The rs574229 SNP was excluded from the haplotype block, since its allele frequency was less than 0.05.

TABLE 62. *IL18* case-control haplotype analysis by Haploview [313].

Block 1 ^a		Freq _{Cases}	Freq _{Controls}	χ^2	p value	p value _{Permutation} ^{b, c}
1	T-A-G-G-T	0.63	0.61	0.72	0.39	0.74
2	T-A-G-T-A	0.23	0.25	1.15	0.28	0.58
3	G-G-C-T-A	0.12	0.12	0.006	0.94	1

^a The order of the SNPs in each block corresponds to Figure 33.

^b Permutation test p values were calculated from 10 000 permutations in Haploview.

^c p values less than 0.28 were observed on 5823 out of 10 000 occasions (58%).

TABLE 63. *IL18 TDT haplotype analysis by Haploview [313].*

Block 1 ^a	Transmitted:untransmitted	χ^2	p	p _{Permutation} ^{b, c}	
1	T-A-G-G-T	37 : 27	1.54	0.21	0.28
2	T-A-G-T-A	21 : 33	3.12	0.08	0.07
3	G-G-C-T-A	19 : 14	0.76	0.38	1

^a The order of the SNPs in each block corresponds to Figure 33.

^b Permutation test p values were calculated from 10 000 permutations in Haploview.

^c p values less than 0.08 were observed on 726 out of 10 000 occasions (7%)

4.6.1.3. Comparison of allele frequencies between populations

Two promoter SNPs (rs1946518 and rs187238) and a synonymous SNP (rs5549908) have often been studied in various populations. We compared the allele frequencies for these SNPs in the South African Coloured population with that of other populations (Table 64).

Allele frequencies for the promoter SNP rs1946518 were available for six populations. The South African Coloured population had the same allele frequency for rs1946518 as the Czechoslovakian population (T allele = 0.38), which was similar to the allele frequency of white Americans (T allele = 0.34). The G allele was the major allele of rs1946518 in almost all populations, but not in the Koreans. African Americans had the lowest allele frequency (T allele = 0.28) for this SNP.

The rs187238 allele frequency of the South African Coloured population was compared to that of seven other populations. In this instance, a similar allele frequency was detected in the Korean population (0.13), while the rest of the populations had higher frequencies for this SNP. The synonymous rs5549908 SNP used to be the main focus of investigation in association studies [484]. In the South African Coloured population this SNP was in complete LD with rs187238. Allele frequencies of rs5549908 were similar to those of Japanese and Korean populations, but lower than those of Chinese.

TABLE 64. *Comparison of allele frequencies of IL18 between populations.*

Population	Ref	rs1946518		rs187238		rs5549908	
		T	G	C	G	G	T
South African Coloured		0.38	0.62	0.13	0.87	0.13	0.87
African Americans	[493]	0.28	0.72	0.23	0.77	-	-
Chinese	[488]	-	-	-	-	0.23	0.77
Czechoslovakians	[487]	0.38	0.62	0.28	0.72	-	-
Japanese	[494]	0.43	0.57	0.19	0.81	0.12	0.88
Koreans	[495]	0.54	0.46	0.13	0.87	0.15	0.85
Swiss	[496]	-	-	0.27	0.73	-	-
White Americans	[493]	0.34	0.66	0.27	0.73	-	-

4.6.2. Discussion

IL-18 can induce production of IFN- γ by T cells [473,474] and plays an important role in the cell-mediated immunity against several pathogens such as *M.tuberculosis* [475]. Expression of the *IL18* gene, which encodes this cytokine, may be regulated by promoter polymorphisms. Since no other study that examined these specific polymorphisms of *IL18* in TB has been published to date, we genotyped the SNPs of interest in cases, controls and family samples from the South African Coloured population.

We did not detect any associations between six single SNPs, or 3 haplotypes consisting of these SNPs, in the case-control or TDT analyses. The rs1946518 and rs187238 polymorphisms were previously associated with sarcoidosis [486], asthma [495-497], coronary artery disease [498] and type 1 diabetes [499], while the rs549908 SNP may influence *IL18* transcription [485]. Only one recent TB genetic association study has considered *IL18* polymorphisms [191], but did not detect any association with disease either. These polymorphisms differed from those genotyped in our study.

Allele frequencies of commonly genotyped polymorphisms were compared between different populations. The South African Coloured population shared similar allele frequencies for rs1946518 with Czechoslovakians, and resembled Asian populations for rs187238 and rs549908.

In conclusion, *IL18* polymorphisms are not susceptibility factors in the South African Coloured population.

4.7. SH2 DOMAIN PROTEIN 1A

More male than female TB patients are observed worldwide [117] and even though other factors (societal, environmental or hormonal [66]) may account for this observation, the presence of a TB susceptibility gene on chromosome X cannot be discounted. A previous linkage study [66] using samples from the South African Coloured population indicated that chromosome Xq26 may contain a susceptibility gene for TB. Campbell et al [265] tested the association of variants in the CD40 ligand gene located on chromosome Xq26 with TB, but found no association with the disease.

MSMD (*as discussed in section 4.1.*) is a rare syndrome and affected individuals are typically predisposed to clinical disease which is caused by normally nonpathogenic mycobacteria. Five disease-causing autosomal genes have been found to contain mutations in these individuals. However, these genes could not explain the X-linked recessive inheritance pattern seen in some families [500]. Recently it was shown that X-linked recessive MSMD is caused by mutations in the NF κ B essential modulator gene (*NEMO*) located on chromosome Xq28 [352]. Another study by Bustamante et al [501] examining patients from a large family with X-linked recessive MSMD did not find any mutations in *NEMO*. This family was used in a linkage study which indicated that the Xq25-26.3 and the Xp11.2-p21.2 could contain the causative mutation. Since MSMD is caused by mutations in X-linked genes in some patients, it is possible that common polymorphisms in these genes could contribute to TB susceptibility in the general population. This hypothesis has already been proven true for common variants in certain MSMD-causing genes [454-456,502].

Pasquinelli et al [503] recently found that expression of the signaling lymphocytic activation molecule-associated protein (SAP) abolished the production of IFN- γ in T cells from TB patients and that disease severity correlated with SAP expression. SAP is encoded by the gene known as SH2 domain protein 1A, Duncan's disease (lymphoproliferative syndrome) (*SH2D1A*) and is predominantly expressed by T, NK and some B cells [504]. *SH2D1A*, which is mutated or absent in patients with X-linked lymphoproliferative disease (XLP), is present on chromosome Xq25-q26. This region (Xq26) showed evidence of containing TB susceptibility genes during a genome-wide scan in South African Coloureds and Gambians [66], as mentioned above.

SH2D1A is known to be regulated at both the transcriptional and the post-transcriptional stage and has a functionally confirmed Ets binding site present near the transcription start site of the gene [505]. The human basal promoter (from position -167 to -134) of *SH2D1A* has been identified [505], but it is very likely that other factors are also involved in promoter activity of the gene. Such factors could include differential methylation (which is present in the *SH2D1A* promoter [506]), that may promote or prevent the binding of transcription factors, or SNPs that create or destroy transcription factor binding sites.

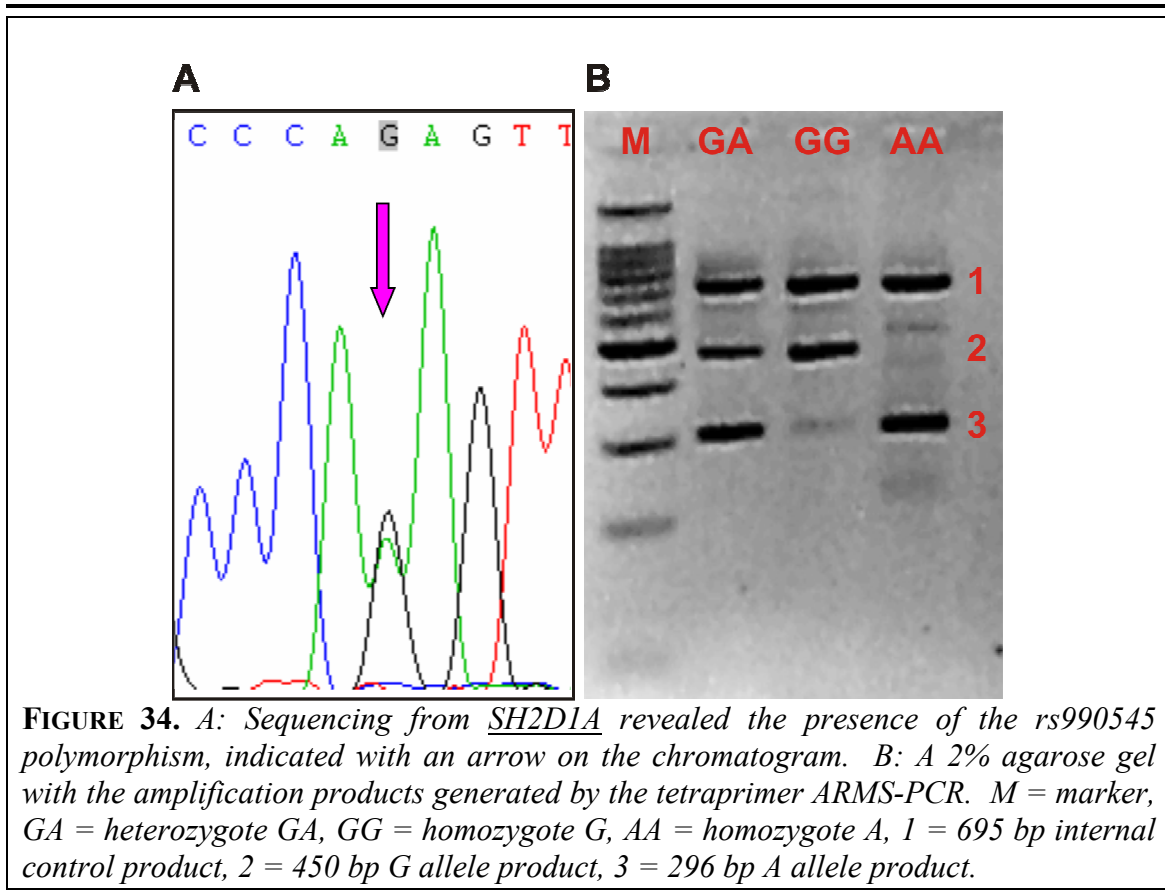
This part of the study further investigates the linkage found with chromosome Xq26 in a previous genome-wide scan [66] by selecting a candidate gene from the region identified. We hypothesise that high levels of SAP may interfere with Th1 cytokine responses, and therefore IFN- γ levels, and could cause an increased susceptibility to TB. To test this, we sequenced the region upstream of *SH2D1A* to identify polymorphisms

that could influence transcription of the gene and genotyped a SNP that creates putative transcription factor binding sites.

4.7.1. Results

4.7.1.1. Sequencing and in silico promoter analysis

Sequencing (Figure 34A) of the putative promoter of *SH2D1A* in 10 individuals did not reveal any novel SNPs, but 3 known variants were present in this region, namely G-631A (rs990545), G-494A (rs7357894) and C-347T (rs12164382). The polymorphisms were in complete linkage disequilibrium ($D' = 1$ and $r^2 = 1$, data not shown). We then used the SNPInspector program [314] to determine the potential effects of the SNPs on the *SH2D1A* sequence. Only rs990545 created 2 putative transcription factor binding sites (for CCAAT enhancer binding protein, beta (C/EBP- β) and pregnane X receptor (PXR), respectively) and we therefore selected this polymorphism to genotype in all the samples from the female and male DNA collections. A tetraprimer ARMS-PCR method (Figure 34B, section 3.6.2.) was used for genotyping.



4.7.1.2. Single-point statistical analysis

The female dataset was in HWE for both cases and controls, indicating the probability of correct genotyping. Since males have one X chromosome only, they were all hemizygous for rs990545 and not tested for HWE.

Genotype and allele frequencies for female subjects and allele data for male individuals are listed in Tables 65 and 66. The G allele of the rs990545 SNP was associated with susceptibility to TB in the female dataset ($p_{\text{genotypes}} = 0.016$; $p_{\text{alleles}} = 0.021$, OR = 1.3 [1.04-1.65]) which was analysed initially as it provided both hetero- and homozygotes. To replicate this finding, we analysed the allelic distribution of this polymorphism in males unrelated to the females and again found a significant association ($p = 0.014$, OR = 1.6 95% CI [1.11-2.33]). Since X-inactivation takes place in females, we compared

TABLE 65. Genotype frequencies and association result of the *SH2D1A* variant investigated in the South African Coloured female population using a case-control study.

rs990545	GG	GA	AA	p value
TB (n^a=269)	137 (0.51) ^b	101 (0.38)	31 (0.11)	0.016
Controls (n=449)	180 (0.40)	212 (0.47)	57 (0.13)	

^a number of samples

^bThe frequency of each genotype is indicated in brackets.

females homozygous for either of the alleles of the SNP separately [507]. We observed a non-significant trend ($p = 0.058$) for this analysis, with the GG genotype occurring more frequently in cases than controls. An analysis of the combined female and male allelic data, revealed a strong association ($p = 0.0009$, OR = 1.39 [1.14-1.68]) between TB and the G allele of rs990545.

TABLE 66. Allele frequencies and association results of the *SH2D1A* variant investigated in the South African Coloured population using a case-control study.

rs990545	G	A	p value	OR ^a [95% CI ^b]
Female heterozygotes and homozygotes				
TB (n^c=269)	375 (0.70) ^d	163 (0.30)	0.021	1.31 [1.04-1.65]
Controls (n=449)	572 (0.64)	326 (0.36)		
Male hemizygotes				
TB (n=270)	194 (0.72)	76 (0.28)	0.014	1.61 [1.11-2.33]
Controls (n=238)	146 (0.61)	92 (0.39)		
Female and male subjects				
TB (n=539)	569 (0.70)	239 (0.30)	0.0009	1.39 [1.14-1.68]
Controls (n=687)	718 (0.63)	418 (0.37)		

^a odds ratio

^b confidence interval

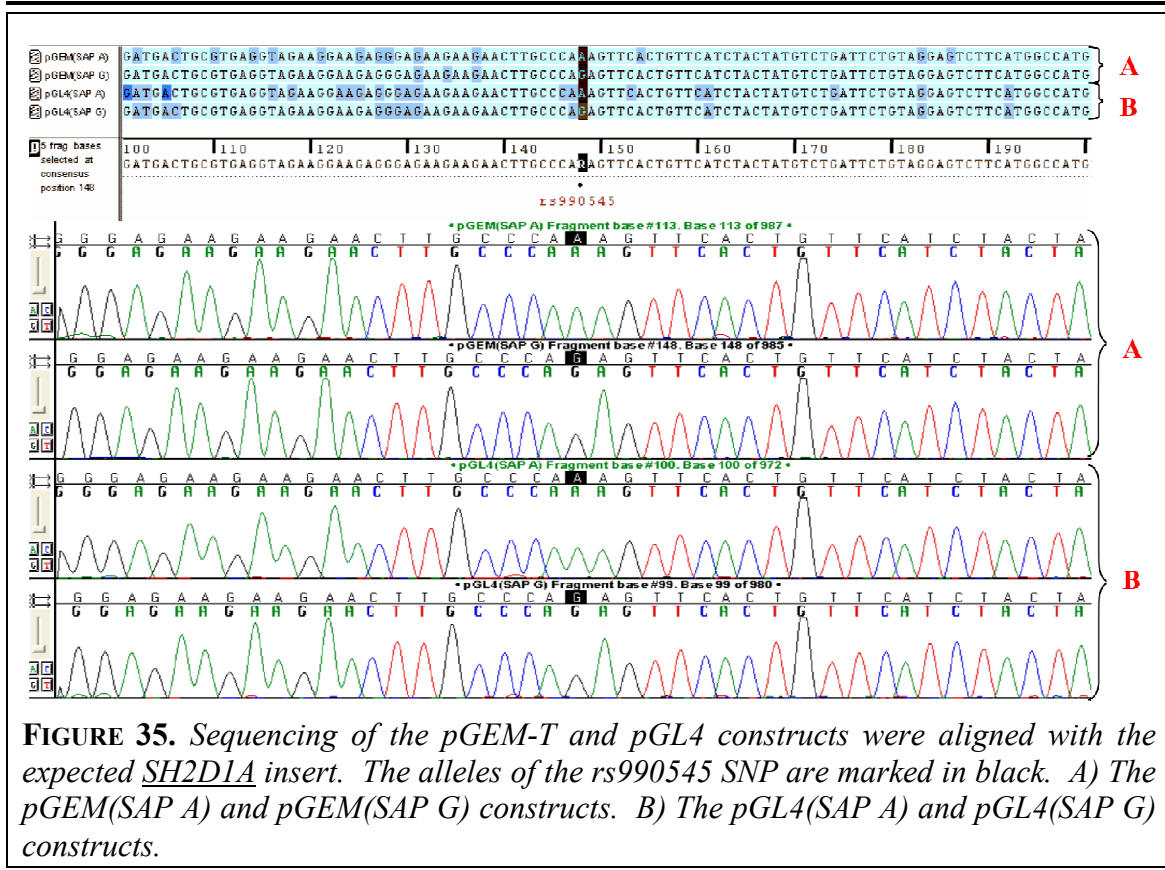
^c number of samples

^d The frequency of each genotype is indicated in brackets.

4.7.1.3. Cloning of the *SH2D1A* gene promoter

Cloning of the *SH2D1A* promoter region into the pGEM[®]-T Easy vector system resulted in two constructs, namely pGEM(SAP G) and pGEM(SAP A). Each contained one

allele (G or A) of the rs990545 SNP. Sequencing of the pGEM constructs confirmed that they contained the correct *SH2D1A* insert (Figure 35A). The insert was subcloned from the pGEM constructs into a pGL4 luciferase reporter vector, namely pGL4.10[*luc2*]. The luciferase constructs were named pGL4(SAP G) and pGL4(SAP A) and the nucleotide sequences of the inserts were verified by sequencing (Figure 35B).



4.7.2. Discussion

Chromosome Xq26 was indicated as a region possibly containing a TB susceptibility gene during the first genome-wide scan [66], but no X-linked gene has been associated with TB to date [265]. Globally, fewer female than male TB patients are detected [117] and even though other factors may contribute to this discrepancy, it is possible that a TB susceptibility gene on the X chromosome may be the underlying cause.

Here we report the association of a polymorphism (rs990545) in the promoter of the X-linked *SH2D1A* gene with susceptibility to TB in a South African population. Expression of SAP, encoded by *SH2D1A*, has previously been shown to be the greatest in TB patients with low immune responses, suggesting that the SAP protein may be involved in the outcome of TB infection [503]. In addition, reduction of SAP expression caused *M.tuberculosis* antigen-stimulated T cells to produce high levels of IFN- γ [503].

We found an association between the G allele of the rs990545 *SH2D1A* SNP and susceptibility to TB in female and male case-control sample sets from the South African Coloured population (Tables 65 and 66). Since *SH2D1A* is X-linked, the data was separated by gender [508]. This association was significant for the female sample set on both the genotypic and allelic level and the finding was replicated in the allelic frequency in the male sample. A combined analysis of the female and male data also showed a highly significant association.

The A allele of rs990545 in *SH2D1A* creates two putative transcription factor binding sites, namely for C/EBP- β and PXR, and this polymorphism could therefore contribute to the regulation of the transcriptional activity of *SH2D1A*. C/EBP- β is stimulated by IFN- γ [509] and has three different isoforms, of which two are activators and one is a repressor [510]. This transcription factor can act as a mycobacterial or bacterial response element [511] and regulates the acute phase-responsive genes under the control of interleukin-6 [509]. PXR is a nuclear receptor and an activator of several cytochrome P450 genes [512]. Interestingly, PXR is activated by rifampicin [513], a drug used extensively to treat TB. Both C/EBP- β and PXR could be involved in regulating *SH2D1A* transcription, since transcription factors can bind to overlapping binding sites without competition [514]. To test this hypothesis in the future two luciferase expression vectors controlled by the *SH2D1A* promoter fragment, containing either the A or G allele of rs990545, were constructed in this study.

Since the A allele of rs990545 is more frequent in controls than TB cases, it is feasible that the creation of the C/EBP- β and PXR transcription factor binding sites could lead to the binding of their respective transcription factors and inhibition of *SH2D1A* transcription. This would cause lower levels of SAP and could lead to a stronger Th1 response, which would be advantageous during infection with *M.tuberculosis*.

The association observed with the rs990545 polymorphism could be true, but could also be due to LD with another polymorphism in the same or another gene on the X chromosome. Recently, X-linked recessive MSMD has been found to be caused by mutations in *NEMO* [352], located in the chromosome Xq28 region. Therefore future studies of the importance of the X-linked *SH2D1A* gene in TB should evaluate the functionality of the rs990545 promoter polymorphism in TB to confirm the hypothesis of this study.

4.8. TOLL-LIKE RECEPTORS

The family of mammalian Toll-like receptors (TLRs) consists of at least 12 proteins, each with a distinct function, which initiate the innate immune response, an essential component in the host defence against infection with micro-organisms. This initiation is done through either interferon-regulatory factor-dependent or nuclear factor- κ B-dependent signalling pathways [515]. This chain of events is set off by the recognition of pathogen-associated molecular patterns (PAMPs), which are microbial components that are crucial for the survival of the micro-organism and therefore difficult for it to change [516]. The TLR-induced mechanism eventually allows the modulation of the adaptive immune system in response to infections. TLRs are type I transmembrane receptors and are characterised by an intracellular Toll/interleukin-1 receptor (TIR) domain and an ectodomain of leucine-rich repeats (LRR) [517-519] (Figure 36).

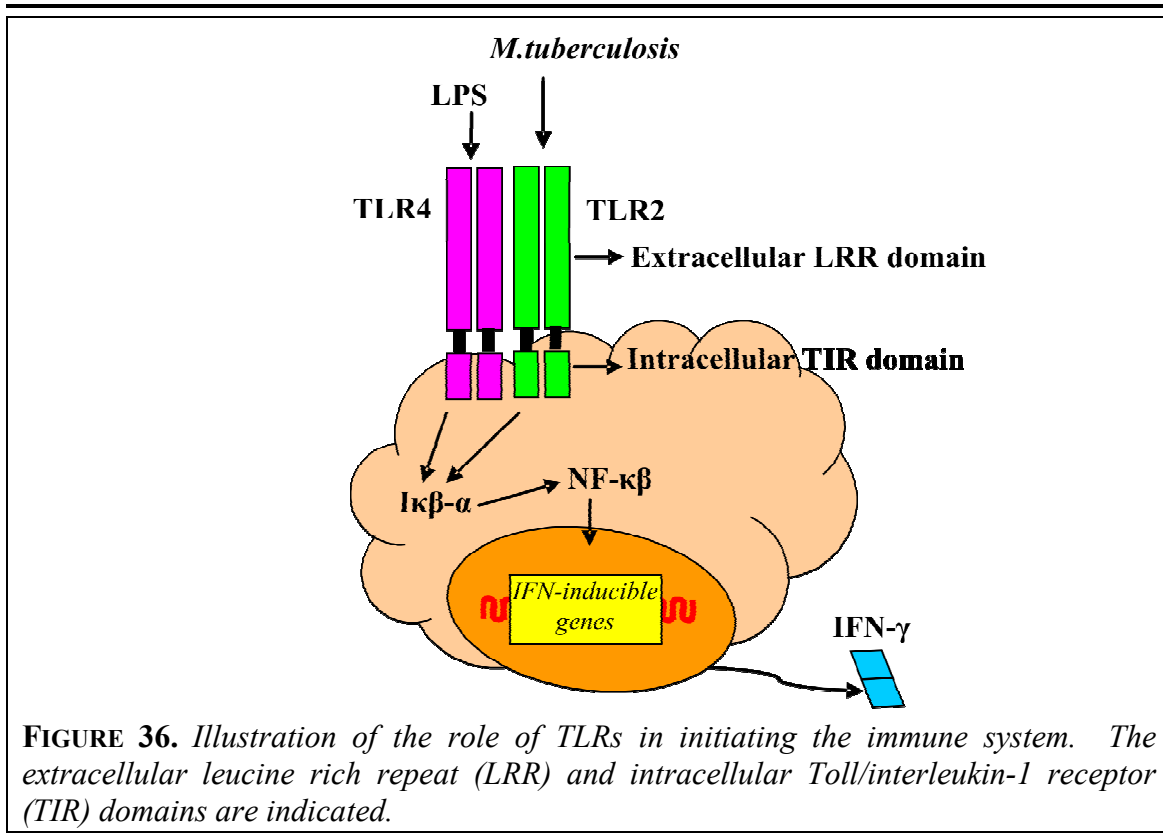


FIGURE 36. Illustration of the role of TLRs in initiating the immune system. The extracellular leucine rich repeat (LRR) and intracellular Toll/interleukin-1 receptor (TIR) domains are indicated.

TLRs are central components of the innate immune response to mycobacterial infection [520,521] and act as part of the pattern recognition system to signal the presence of *M.tuberculosis* in the host [522]. TLR2 and TLR4 were found to be nonredundant recognition systems of the bacterium [522]. The importance of TLR pathways in the sensing of mycobacteria was further shown in a mouse model deficient for myeloid differentiation factor 88 (MyD88), an adaptor molecule utilized by nearly all receptors of the TLR family [523]. These mice were extremely susceptible to infection with *M.tuberculosis* [523]. TLRs are also expressed on T cells and may modulate T cell activation by TLR ligands [524].

TLR2 has a vital role in the responsiveness to several bacterial products [525-529], such as leptospiral LPS [528] and mycobacterial lipoproteins [527]. In Th1 cells from mice, TLR2 stimulation induced IFN- γ production, cell proliferation and cell survival [524]. IL-12 and IL-2 enhanced these effects, indicating that TLR2 is a specific activator of Th1 responses, while other TLRs investigated in the study did not alter Th2 responses [524]. Two recent studies showed that *M.tuberculosis* may inhibit macrophage responses to IFN- γ by a TLR2-dependent pathway [530,531], one of several mechanisms used by the bacterium to inhibit the host immune system. Another popular hypothesis argues that TLR2 favours Th2 responses during *M.tuberculosis* infection and contributes to the disease in this manner [532].

TLR2 knockout mice are highly susceptible to *M.tuberculosis* infection [533,534] and have a short survival time. These mice also have a higher burden of mycobacteria when compared to mice with the functional gene [535]. For this reason, polymorphisms in *TLR2* have been investigated with regards to TB in a variety of populations. An association between TB and *TLR2*, with a SNP named Arg677Trp [536] which appeared to be functional in mycobacterial disease [537], proved to be false, after it was established that Arg677Trp was an artifact due to the presence of a pseudogene region [538]. The rs5743708 SNP (Arg753Gln) was investigated in TB cases and controls from Turkey and it was suggested that this SNP may contribute to the risk of developing TB [229]. There was a slight deviation from Hardy-Weinberg in the control population and the authors speculated that this could be due to the high incidence of consanguinity in the Turkish population [229]. This possibility necessitates the replication of the association of rs5743708 with TB in another population. Yim et al [302] studied the 5'UTR region of the *TLR2* gene and identified a microsatellite polymorphism in intron 2 of the gene. The frequency of this GT repeat polymorphism varied significantly between populations, and modified promoter activity. This microsatellite was subsequently associated with rheumatoid arthritis [323], colorectal cancer [539] and TB [291]. Shorter GT repeats were present more often among TB patients and in validation samples from Korea, were associated with weaker promoter activity and reduced expression of TLR2 on CD14+ peripheral blood monocytes [291]. Variants in *TLR2* have also been associated with other diseases, such as asthma where a promoter polymorphism, rs4696480, was associated with disease [540] and staphylococcal infection [536].

TLR4 was the first mammalian toll receptor to be characterised [518] and was identified as the mediator of LPS inflammatory responses [264,541-543]. This receptor can also respond to heat-labile soluble mycobacterial components as well as viable *M.tuberculosis* and TLR4 will initiate the innate immune system when exposed to these [544,545]. In addition, mice deficient in TLR4 developed chronic lung infection when exposed to *M.tuberculosis* and produced less IL-12p40 and MCP-1 [546]. Two missense mutations in *TLR4*, namely rs4986790 (Asp299Gly) and rs4986791 (Thr399Ile), affect the extracellular domain of the TLR4 protein and were associated with hyporesponsiveness to LPS [547], an increased incidence of septic shock [548] and with Gram-negative and haematogenous osteomyelitis [549]. The rs4986790 polymorphism was previously investigated in Gambian TB patients, but no influence in LPS responsiveness or susceptibility with the disease was determined [264]. The rs4986791 SNP was not tested, since it was not present in Gambians [550].

The results of current studies indicate that the activation of several TLRs contribute to an efficient innate response against mycobacteria and support the idea that both the TLR2 and TLR4 proteins play a role in sensing mycobacteria and initiating an antimycobacterial immune response [551]. Due to the importance of the pattern-recognition receptors in recognising invading microorganisms, we chose *TLR2* and *TLR4* (encoding the receptors TLR2 and TLR4 respectively), as candidate genes. We selected polymorphisms (Table 67) in the *TLR2* and *TLR4* genes to genotype in the South African Coloured population, based on their possible function in host defence against TB.

TABLE 67. Polymorphisms genotyped in the *TLR2* and *TLR4* genes.

Polymorphism	Class	Location	Genotyping method
<i>TLR2</i>			
rs4696480	SNP	Promoter	SNPlex
TLR2_GT	STR	Intron 2	Capillary electrophoresis
rs3804099	SNP	Exon 3, synonymous	SNPlex
rs3804100	SNP	Exon 3, synonymous	SNPlex
rs5743704	SNP	Exon 3, non-synonymous	SNPlex
rs5743708 (Arg753Gln)	SNP	Exon 3, non-synonymous	SNPlex
<i>TLR4</i>			
rs1927914	SNP	Promoter	SNPlex
rs10759932	SNP	Promoter	SNPlex
rs2770148	SNP	Exon 3/intron 2 boundary	SNPlex
rs4986790 (Asp299Gly)	SNP	Exon 3, non-synonymous	SNPlex
rs4986791 (Thr399Ile)	SNP	Exon 3, non-synonymous	TaqMan
rs11536889	SNP	3' UTR	SNPlex
rs11536891	SNP	3' UTR	SNPlex

4.8.1. Results

4.8.1.1. Single-point statistical analysis

Single-point statistical evaluations of *TLR2* and *TLR4* polymorphisms were done with case-control and TDT analyses. All SNPs were in HWE in the control samples. Correction for multiple testing was done using the definition in Haploview based on the study of Nicodemus et al [76].

a) *TLR2*

Five SNPs were studied as TB susceptibility factors in the South African Coloured population using case-control (Table 68) and TDT analyses (Table 69). After correcting for multiple tests, the cut-off for significance was set at 0.01. According to this criterion only the synonymous rs3804100 SNP was significantly associated with TB ($p = 0.01$, OR = 1.67 95%CI [1.1-2.5]). However, this association was only detected in the case-control samples (Table 68). The non-synonymous rs5743708 SNP was previously

associated with TB in a Turkish population [229], but this finding was not replicated in our study, which had a larger sample size and therefore more power. None of the three remaining SNPs were associated with TB.

TABLE 68. Single-point statistical analysis in the *TLR2* case-control study.

Polymorphism	Allele		Genotype frequencies						Cases vs Controls p value ^c
			Cases (n ^a =432)			Controls (n=482)			
	1 ^b	2	11	12	22	11	12	22	
rs4696480	A	T	0.14	0.43	0.43	0.12	0.47	0.41	0.51
rs3804099	T	C	0.17	0.48	0.35	0.20	0.49	0.31	0.29
rs3804100	C	T	0	0.08	0.92	0	0.14	0.86	0.01
rs5743704	A	C	0	0.01	0.99	0	0.01	0.99	0.88
rs5743708	A	G	0	0.03	0.97	0	0.01	0.99	0.38

^a number of investigated individuals.

^b Allele 1 is the minor allele.

^c p value from a genotype-based χ^2 test

TABLE 69. Single-point statistical analysis in the *TLR2* TDT.

SNP	Overtransmitted	Transmitted:Untransmitted	χ^2	p value
rs4696480	T	50 : 46	0.17	0.68
rs3804099	T	61 : 48	1.55	0.21
rs3804100	C	15 : 8	2.13	0.14
rs5743704	C	4 : 2	0.67	0.41
rs5743708	G	1 : 0	1	0.32

b) *TLR4*

Seven *TLR4* SNPs were genotyped to test their possible association with TB in the case-control (Table 70) and family-based (Table 71) studies. The level for significance was set at 0.01 after correcting for multiple tests. With this stringent requirement for significance, none of the SNPs were associated with TB in the case-control (Table 70) or TDT (Table 71) analyses.

The nonsynonymous rs4986790 and rs4986791 SNPs were previously associated with hyporesponsiveness to LPS [547] and affect the extracellular domain of the *TLR4* protein. The rs4986790 SNP was previously investigated in a Gambian TB association study who found no association between this SNP and disease [264] and rs4986791 was not present in samples from the Gambians. For this reason we included these two polymorphisms in our study. However, we did not detect any association between these coding SNPs and TB (Tables 70 and 71).

TABLE 70. Single-point statistical analysis in the *TLR4* case-control study.

Polymorphism	Allele		Genotype frequencies						Cases vs Controls p value ^c
			Cases (n ^a =432)			Controls (n=482)			
	1 ^b	2	11	12	22	11	12	22	
rs1927914	T	C	0.10	0.44	0.46	0.14	0.45	0.41	0.09
rs10759932	C	T	0.01	0.22	0.77	0.01	0.27	0.72	0.20
rs2770148	T	C	0.04	0.34	0.63	0.03	0.26	0.71	0.04
rs4986790	G	A	0	0.12	0.88	0	0.09	0.91	0.22
rs4986791	T	C	0	0.05	0.95	0	0.04	0.96	0.63
rs11536889	C	G	0.01	0.12	0.87	0.01	0.17	0.82	0.09
rs11536891	C	T	0.06	0.39	0.55	0.08	0.34	0.58	0.31

^a number of investigated individuals.

^b Allele 1 is the minor allele.

^c p value from a genotype-based χ^2 test

TABLE 71. Single-point statistical analysis in the *TLR4* TDT.

SNP	Overtransmitted	Transmitted:Untransmitted	χ^2	p value
rs1927914	C	43 : 29	2.72	0.10
rs10759932	T	28 : 24	0.31	0.58
rs2770148	C	33 : 23	1.79	0.18
rs4986790	-	9 : 9	0	1
rs4986791	C	11 : 8	0.47	0.49
rs11536889	G	12 : 8	0.80	0.37
rs11536891	C	31 : 28	0.15	0.70

4.8.1.2. *TLR2* microsatellite analysis

The *TLR2* microsatellite (*TLR2*_GT), present in intron 2 of the gene and associated with weaker promoter activity [291], was genotyped in the South African Coloured population to determine its association with TB. This marker was previously associated with TB in Koreans [291]. In that study a number of repeat alleles, ranging from (GT)₁₂ to (GT)₂₈, were observed.

To genotype this microsatellite, we used capillary electrophoresis of fluorescently labelled PCR products (*Section 3.7*). The Genemapper version 3.7 program was used in the analysis of the results and represented the repeats visually on electropherograms (Figure 37). The repeat alleles were confirmed by sequencing (*Section 3.5.6*). We detected alleles ranging from (GT)₁₂ to (GT)₂₇ (Table 72). We did not detect any (GT)₁₄ or (GT)₂₈ alleles in the South African Coloured population, which differs from the findings in the Korean study [291].

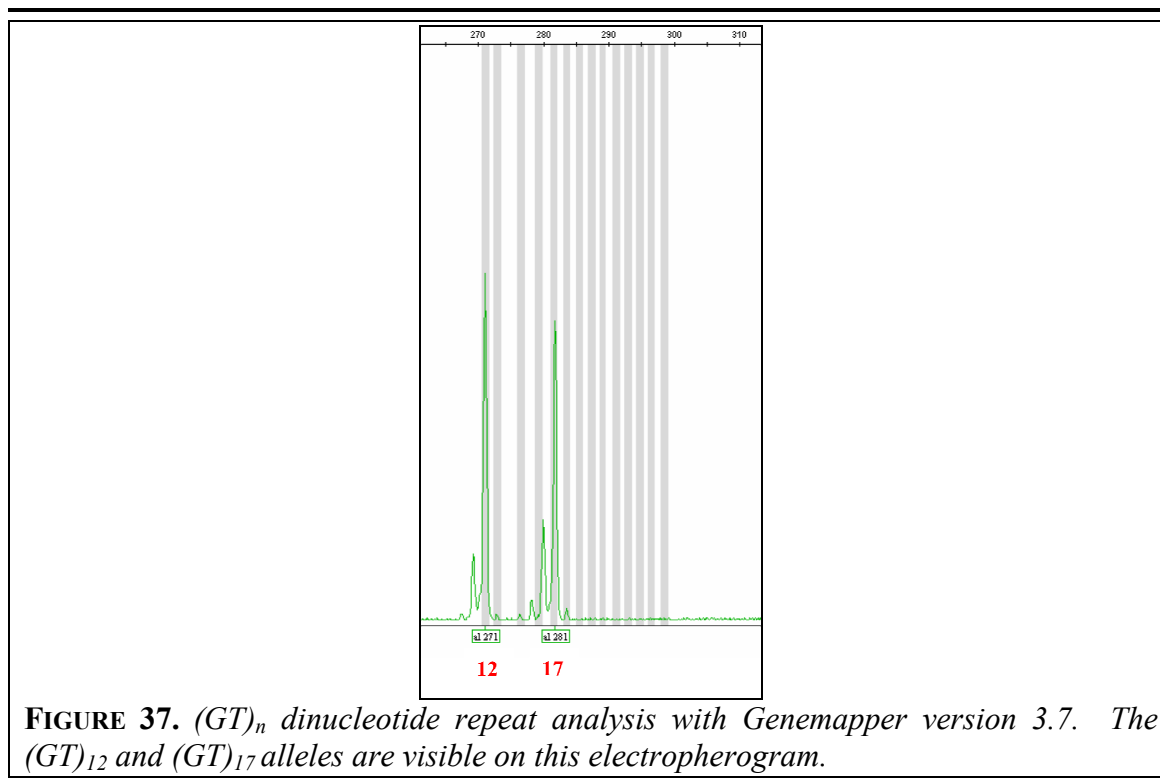


TABLE 72. Frequencies of the respective alleles of TLR2_GT in the TB cases and controls.

Allele	Cases(n) alleles	CasesFrequency	Controls(n) alleles	ControlsFrequency
(GT)₁₂	39	0.05	36	0.04
(GT)₁₃	34	0.04	55	0.06
(GT)₁₄	0	0.00	0	0.00
(GT)₁₅	31	0.04	24	0.03
(GT)₁₆	17	0.02	27	0.03
(GT)₁₇	52	0.06	59	0.07
(GT)₁₈	181	0.21	163	0.19
(GT)₁₉	118	0.14	119	0.14
(GT)₂₀	104	0.12	106	0.12
(GT)₂₁	47	0.06	34	0.04
(GT)₂₂	58	0.07	67	0.08
(GT)₂₃	101	0.12	110	0.13
(GT)₂₄	14	0.02	23	0.03
(GT)₂₅	38	0.05	42	0.05
(GT)₂₆	7	0.01	6	0.01
(GT)₂₇	1	0.00 ^a	1	0.00 [*]

^a This allele occurred in less than 1% of the samples

The numbers of GT repeats from the TLR2 microsatellite marker were determined by direct counting and plotted as a distribution graph (Figure 38). Since this graph was trimodal, the alleles were divided into three subclasses, as described by studies in the analysis of microsatellite polymorphisms [291,323-325]. Alleles with 16 or less GT repeats were designated as S alleles, those with (GT)₁₇ to (GT)₂₂ as M alleles and those with 23 or more GT repeats as L alleles. To evaluate the genotypic effect on the development of TB, the six genotypes (S/S, S/M, S/L, M/M, M/L, L/L) resulting from this classification were divided into two subgroups depending on the presence of the S allele [291]. Genotypes including the S allele were S/S, M/S and L/S, while genotypes without the S allele were M/M, M/L and L/L. Association between the TLR2 microsatellite and TB was tested using a two-tailed chi-square test. For the TDT, the M and L alleles were combined in a single group.

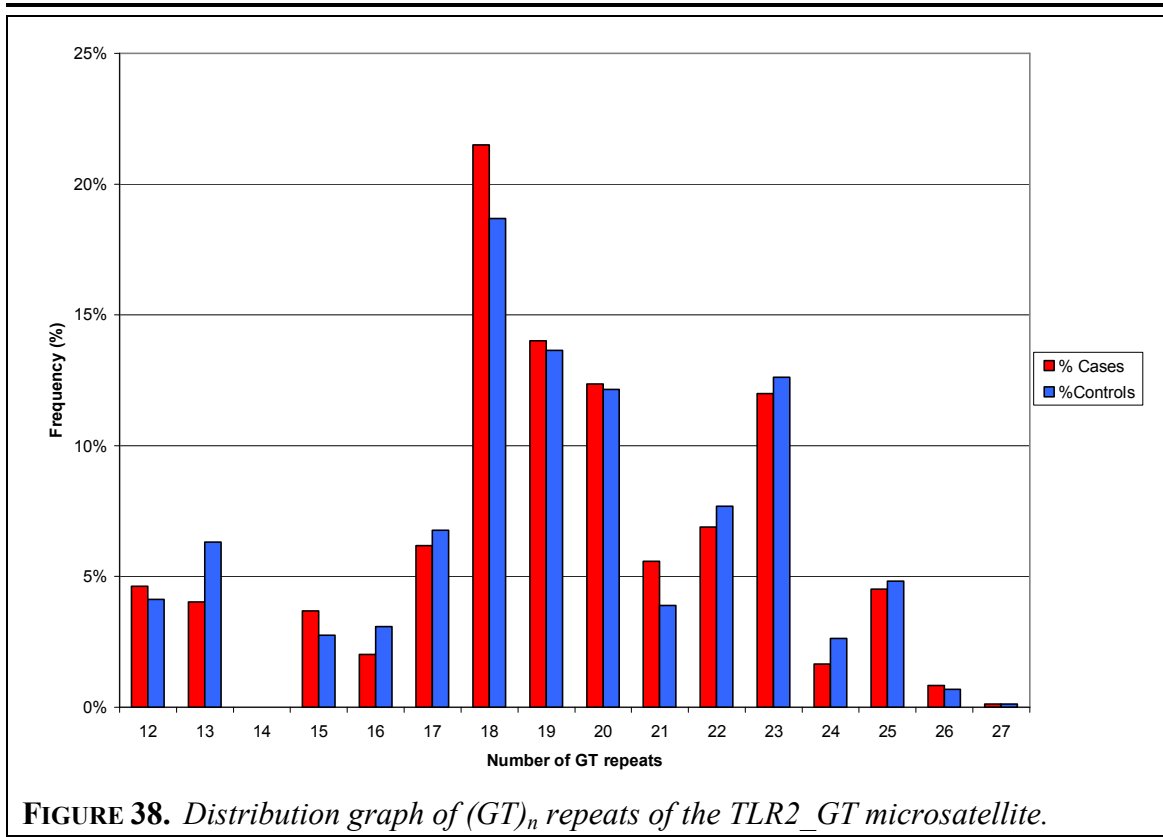


TABLE 73. Microsatellite analysis of TLR GT in the case-control study.

TLR2_GT	Allele		Genotype frequencies					Cases vs Controls	
			Cases (n ^a =432)			Controls (n=482)			
	1 ^b	2	11	12	22	11	12	22	p value
	S	M/L	0.02	0.24	0.73	0.03	0.62	0.71	0.49 ^c ; 0.36 ^d

^a number of investigated individuals.

^b Allele 1 is the minor allele.

^c p value from genotype-based χ^2 test.

^d p value of comparison of genotypes with and without S allele.

We did not find any association with the TLR2_GT microsatellite and TB susceptibility (Table 73) using a genotype-based test ($p = 0.49$) or comparing genotypes with and without the S allele ($p = 0.36$). A TDT in the family-based samples established that the S allele was transmitted to affected children more often than expected, but this was not significantly so (Table 74).

TABLE 74. *TLR2 GT TDT analysis.*

TLR2_GT	Overtransmitted	Transmitted:Untransmitted	χ^2	p value
S	S	25 : 15	2.5	0.11

4.8.1.3. Haplotype analysis

Haplotype and LD analysis for the *TLR2* and *TLR4* genes were done with the Haploview program.

a) *TLR2*

LD across the *TLR2* gene region was not strong (Figure 39). However, the rs3804099 and rs3804100 SNPs were in complete LD. These two SNPs were also part of a single haplotype block identified in this region, which did not include rs4696480, TLR2_GT, rs5743704 and rs5743708. Three haplotypes were identified in the haplotype block (Tables 75 and 76).

Haplotype 1 was only associated with disease in the case-control analysis ($p = 0.02$) and remained significant after permutation testing ($p = 0.04$). Haplotype 2 (Tables 75 and 76) was nominally associated with TB in both the case-control and TDT studies. This association was still significant in the case-control group after permutation testing, but became less significant in the TDT analysis. According to calculation of the Haploview program 2% of permutations exceeded the best p value, meaning that there is a 2% chance that the nominally significant p value was a random statistical fluctuation. However, in the TDT haplotype analysis (Table 76) there was a 9% chance of this result not being significant.

TABLE 75. *TLR2 case-control haplotype analysis by Haploview [313].*

Block 1 ^a	Freq _{Cases}	Freq _{Controls}	χ^2	p value	p value _{Permutation} ^{b, c}
1 C-C	0.04	0.07	5.86	0.02	0.04
2 C-T	0.55	0.49	7.20	0.01	0.02
3 T-T	0.41	0.45	2.50	0.11	0.25

^a The order of the SNPs in the block corresponds to Figure 39.

^b Permutation test p values were calculated from 10 000 permutations in Haploview.

^c p values less than 0.01 were observed on 186 out of 10 000 occasions (2%).

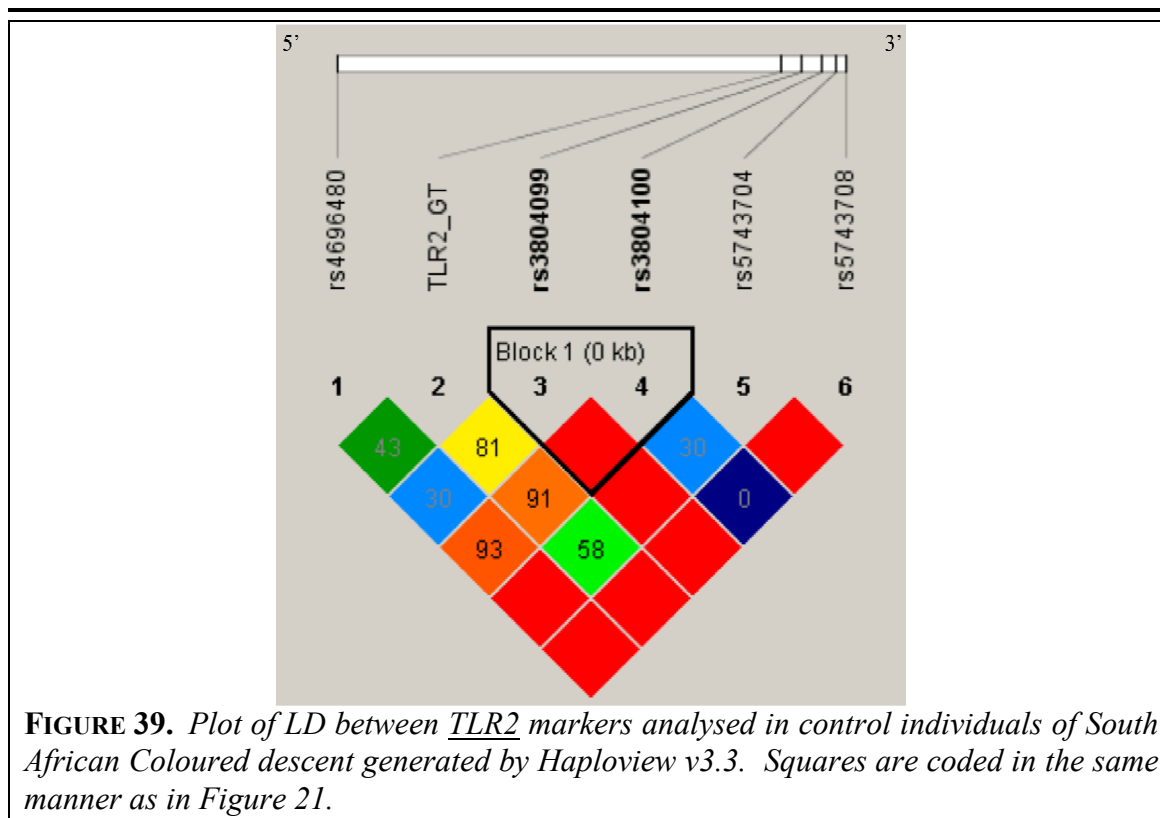


FIGURE 39. Plot of LD between *TLR2* markers analysed in control individuals of South African Coloured descent generated by Haploview v3.3. Squares are coded in the same manner as in Figure 21.

TABLE 76. *TLR2* TDT haplotype analysis by Haploview [313].

Block 1 ^a	Transmitted:untransmitted	χ^2	p value	p value _{Permutation} ^{b, c}
1 C-C	10 : 6	0.89	0.35	0.67
2 C-T	33 : 51	4.21	0.04	0.09
3 T-T	50 : 35	2.64	0.10	0.22

^a The order of the SNPs in each block corresponds to Figure 39.

^b Permutation test p values were calculated from 10 000 permutations in Haploview.

^c p values less than 0.04 were observed on 859 out of 10 000 occasions (9%).

b) *TLR4*

LD across the *TLR4* gene region varied (Figure 40). The rs1927914, rs10759932 and rs2770148 SNPs were in complete LD. These three SNPs were also part of a single haplotype block, which did not include the SNPs in the 3' region of *TLR4*, identified in this region. Four haplotypes were identified in the haplotype block (Tables 77 and 78). Haplotype 2 (Tables 77 and 78) was nominally associated with TB in the case-control (p = 0.03, OR = 1.31, 95% CI [1.03-1.66]) and TDT (p = 0.01) studies. This association was still significant in the TDT group after permutation testing, but became less significant in the case-control analysis. According to the calculations of the Haploview program, 3% of permutations exceeded the best p value, meaning that there is a 3% chance that the nominally significant p value was a random statistical fluctuation. However, in the case-control haplotype analysis (Table 77), there was a 9% chance of this result not being significant.

TABLE 77. *TLR4* case-control haplotype analysis by Haploview [313].

Block 1 ^a	Freq _{Cases}	Freq _{Controls}	χ^2	p value	p value _{Permutation} ^{b, c}
1 C-C-T	0.12	0.15	2.96	0.09	0.27
2 C-T-C	0.20	0.16	4.97	0.03	0.09
3 C-T-T	0.36	0.33	2.25	0.13	0.39
4 T-T-T	0.32	0.37	4.33	0.04	0.13

^a The order of the SNPs in each block corresponds to Figure 40.

^b Permutation test p values were calculated from 10 000 permutations in Haploview.

^c p values less than 0.03 were observed on 913 out of 10 000 occasions (9%).

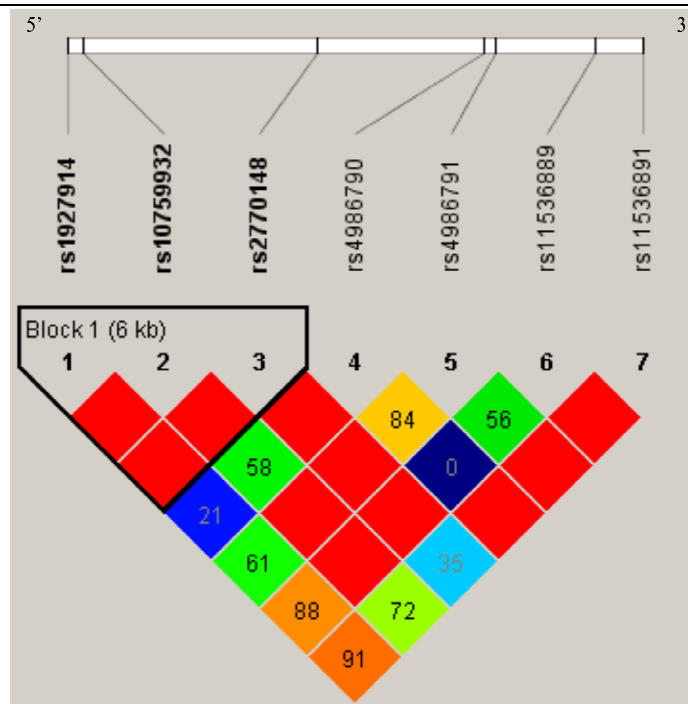


FIGURE 40. Plot of LD between *TLR4* markers analysed in control individuals of South African Coloured descent generated by Haploview v3.3. Squares are coded in the same manner as in Figure 21.

TABLE 78. *TLR4* TDT haplotype analysis by Haploview [313].

Block 1 ^a	Transmitted:untransmitted	χ^2	p value	p value _{Permutation} ^b
1 C-C-T	16 : 23	1.37	0.24	0.60
2 C-T-C	25 : 10	6.44	0.01	0.03
3 C-T-T	36 : 31	0.42	0.52	0.89
4 T-T-T	22 : 34	2.60	0.11	0.29

^a The order of the SNPs in each block corresponds to Figure 40.

^b Permutation test p values were calculated from 100 000 permutations in Haploview.

^c p values less than 0.01 were observed on 2685 out of 100 000 occasions (3%).

4.8.1.4. Comparison of allele frequencies between populations

We compared previously published allele frequencies of *TLR2* (Table 79) and *TLR4* (Table 80) from different populations with that from the South African Coloured controls of this study.

TABLE 79. Comparison of allele frequencies of rs4696480 and rs5743708 between populations.

Population	rs4696480		rs5743708	
	A	T	A	G
South African Coloured	0.36	0.64	0.01	0.99
African American [58]	0.40	0.60	Absent	
Austrians and Germans [540]	0.51	0.49	-	-
European, Utah [59,62]	-	-	0.05	0.95
French [536]	-	-	0.03	0.97
Spanish [549]	-	-	0.01	0.99
Turkish [229]	-	-	0.05	0.95

TABLE 80. Comparison of allele frequencies of rs4986790 and rs4986791 between populations.

Population	rs4986790		rs4986791	
	G	A	T	C
South African Coloured	0.04	0.96	0.03	0.97
European, Utah [59,62]	0.03	0.97	0.03	0.97
Europeans [540]	0.08	0.92	-	-
Gambians [264]	0.11	0.89	Absent	
Han Chinese, Beijing [59,62]	Absent		Absent	
Japanese, Tokyo [59,62]	Absent		Absent	
Spanish [549]	0.07	0.93	0.08	0.92
Yoruba, Ibadan, Nigeria [59,62]	0.03	0.97	Absent	

The rs4696480 and rs5743708 SNPs from *TLR2* have been genotyped in several European populations and also in African Americans (Table 79). The A allele of the rs5743708 SNP was present at low frequencies in all populations genotyped and was absent in African Americans. The T allele of rs4696480 was the major allele in the South African Coloured and African American populations, while the A allele of this SNP was the major allele in Europeans.

The allele frequencies of the rs4986790 and rs4986791 SNPs (Table 80) from *TLR4* of the South African Coloured population were compared to Europeans, Gambians and the four HapMap populations [59,62] (Section 4.9.1.) The A allele of the rs4986790

polymorphism was frequently present in the Gambian population, but absent in Asians. The South African Coloured population and the HapMap individuals from European descent had the same allele frequencies for rs4986791 (T = 0.03 and C = 0.97). This polymorphism was more polymorphic in the Spanish population (minor T allele frequency = 0.08), but not present in the three remaining populations.

4.8.2. Discussion

Several TLRs contribute to an efficient innate immune response against mycobacteria. TLR2 and TLR4 are nonredundant recognition systems for *M.tuberculosis* [522], which may initiate an antimycobacterial immune response [551]. Since these pattern-recognition receptors are important in recognising invading micro-organisms, we chose two of these genes, namely *TLR2* and *TLR4*, as candidate genes. *TLR2* encodes the TLR2 protein which plays an essential role in the responsiveness to several bacterial products, including that of *M.tuberculosis*. *TLR4* encodes TLR4 which responds to heat-labile soluble mycobacterial components as well as viable *M.tuberculosis* and will activate innate immunity when exposed to these [544,545]. We investigated polymorphisms from the two genes in the South African Coloured population to determine if they are susceptibility factors for TB.

We detected a significant association between the synonymous rs3804100 SNP of *TLR2* and TB ($p = 0.013$, OR = 1.67 95%CI [1.1-2.5]) in the case-control analysis, but not in the TDT analysis. The T allele occurred more frequently in cases than in controls. This SNP was not highly polymorphic, with a minor allele frequency of 0.07 in the controls. Since this SNP does not affect the protein structure of TLR2, we cannot speculate on the function of it, if any, in TB. A haplotype (haplotype 2, Tables 75 and 76) containing the T allele of this SNP was also associated with TB susceptibility in both the case-control and TDT studies. However, the TDT showed an opposite association, with haplotype 2 being undertransmitted to TB offspring. This contradictory result could therefore suggest that this association is an artefact, but could also be explained by the smaller sample size of the TDT-group when compared to the case-control group. We did not find a statistically significant association between the rs5743709 *TLR2* SNP and TB. This SNP was previously associated with TB in a Turkish population [229]. That study showed a slight deviation from HWE in the control population and the authors speculated that this could be due to the high incidence of consanguinity in the Turkish population [229]. However, this deviation from HWE could also contribute to a false positive association. Since this polymorphism was in HWE in both the cases and controls from the South African Coloured population, our finding of no association is probably correct. We investigated more samples, and therefore have more power to detect smaller effect sizes than the Turkish study.

The TLR2_GT microsatellite was previously associated with TB in a study by Yim et al [291]. Shorter GT repeats were more often present among TB patients from Korea and were associated with weaker promoter activity and reduced expression of TLR2 on CD14+ peripheral blood monocytes [291]. However, we did not find any association between alleles of this microsatellite and TB (Tables 73 and 74).

We did not find any association between the seven *TLR4* SNPs genotyped and TB susceptibility in the South African Coloured population. These seven SNPs included

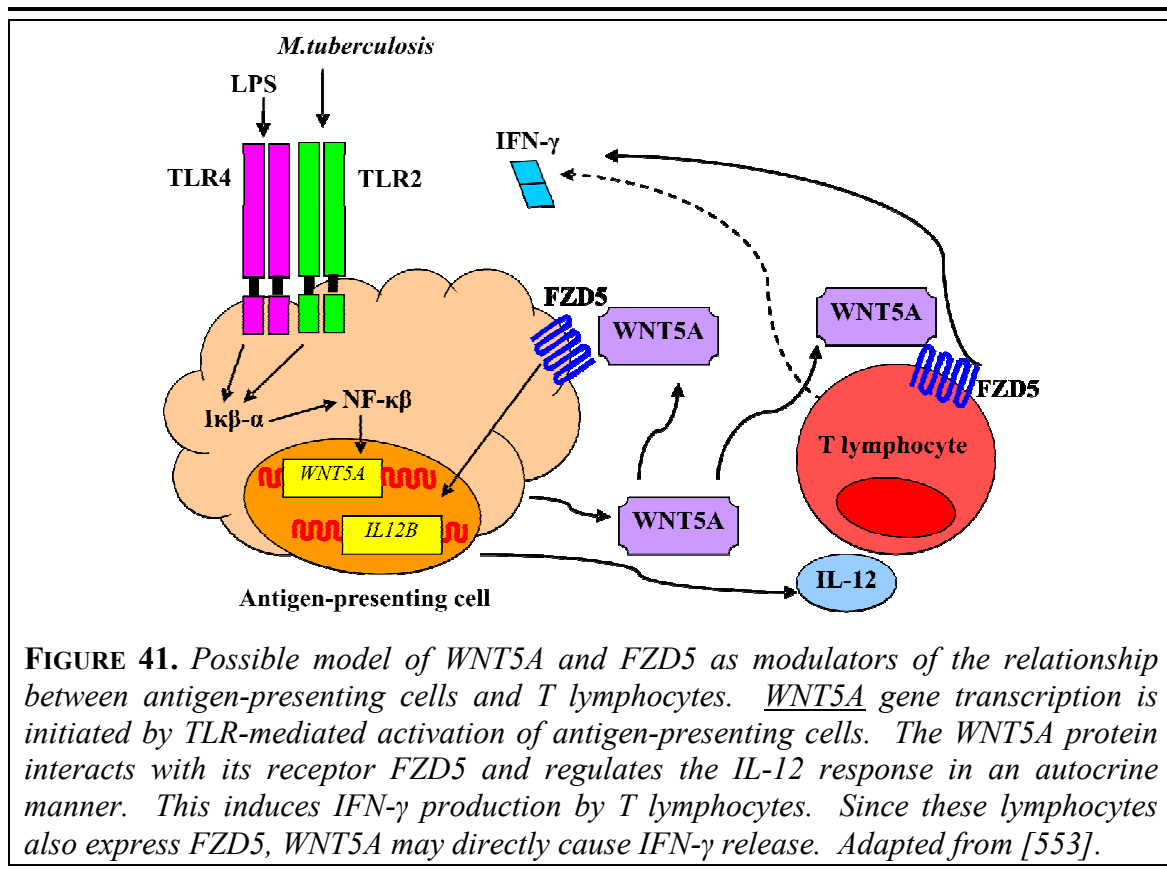
the non-synonymous rs4986790 and rs4986791 polymorphisms, which affect the extracellular structure of the TLR4 protein. The rs4986790 SNP was not a susceptibility factor in a Gambian TB association study either [264].

We also considered a haplotype block consisting of three SNPs (rs1927914, rs10759932 and rs2770148) and found that a haplotype (haplotype 2, Tables 77 and 78) was associated with TB susceptibility. The haplotype occurred more often in cases than in controls ($p = 0.03$, OR = 1.31, 95% CI [1.03-1.66]). It was also transmitted to affected offspring more often than expected ($p = 0.01$). During permutation testing for the TDT and case-control haplotype analyses, which was done to correct for multiple testing, these significant p values occurred randomly in 3% and 9% of the 10 000 runs, respectively. This suggests that this is a true significant association. These 3 SNPs were not individually associated with TB, but were part of a haplotype block associated with disease. Therefore it is possible that the true, unidentified susceptibility variant is in LD with this block, or further upstream of the promoter region of *TLR4*.

In summary, we found an association between a 5' haplotype of *TLR4* and TB susceptibility in both the case-control and TDT analyses. We also found an association between a synonymous SNP, of unknown function, in the *TLR2* case-control study, which was not detected in the TDT study.

4.9. THE WINGLESS-TYPE MMTV INTEGRATION SITE FAMILY, MEMBER 5A AND FRIZZLED HOMOLOG 5 (DROSOPHILA)

Both the innate and acquired immune systems are necessary to eradicate mycobacteria from the host [552]. Recently, a microarray-based gene-expression screening of mycobacteria-infected macrophages was done to search for novel regulatory pathways in innate responses to infection [553]. This study suggested that the Wingless /Frizzled (WNT/FZD) signalling system connects the innate and adaptive immunity during infections (Figure 41) and implicated the wingless-type MMTV integration site family, member 5A (WNT5A) protein in human defence against infection with *M.tuberculosis* [553].



Blumenthal et al [553] demonstrated that WNT5A is expressed by antigen-presenting cells when they are stimulated by mycobacteria or other bacterial structures. In addition, both WNT5A and its receptor, frizzled homolog 5 (FZD5), regulate IL-12 and IFN- γ production in antigen-presenting cells when exposed to mycobacteria. The expression of WNT5A depended on TLR signalling and the activation of the NF- κ B pathway. WNT5A and FZD5 were also present on human mononuclear cells and in granulomatous lesions of the lungs of TB patients [553]. The WNT gene family consists of several structurally related genes which encode secreted signalling proteins. Their receptors are members of the FZD gene family which encodes 7-transmembrane domain proteins [554]. The WNT proteins, which include WNT5A, have been implicated in cell proliferation [555], differentiation [556], developmental processes [557] and oncogenesis [558].

The HapMap project (*Section 1.2.2.*) attempts to determine the common arrangement of DNA variation in the human genome [59]. The first phase of the study genotyped one common SNP (with an allele frequency greater than 5%) every 5 kilobases in 269 DNA samples from 4 populations, namely 90 Yoruba individuals from Ibadan in Nigeria, 90 individuals of European descent from Utah, 45 Han Chinese individuals from Beijing and 45 Japanese from Tokyo [62]. The project proved the redundancy among SNPs located closely together and this enabled researchers to select haplotype tagging SNPs. These SNPs represent the common variation in a genomic region and selecting them for genotyping in association studies may optimise these studies. The HapMap also made it possible to extract genomic variation information without resequencing.

We selected haplotype tagging SNPs (Table 81) from all four populations of the HapMap to test for a possible gene-based association between *WNT5A* and *FZD5* and TB susceptibility. All four HapMap populations were considered, because similar populations might have contributed to the genetic composition of the admixed South African Coloured population. TagSNPs were selected with the Tagger program in Haploview, using the default criteria of $r^2 = 0.8$ and $\text{LOD} = 3$.

TABLE 81. *Tagging SNPs genotyped in WNT5A and FZD5 in this study.*

Polymorphism	Location	Genotyping method
<i>WNT5A</i>		
rs3796232	5' upstream	SNPlex
rs1795651	5' upstream	SNPlex
rs7624718	5' upstream	SNPlex
rs557077	Promoter	TaqMan
rs566926	Intron 1	SNPlex
rs648872	Intron 1	SNPlex
rs815541	Intron 1	SNPlex
rs9311564	Intron 2	SNPlex
rs472631	Intron 3	SNPlex
rs556874	Intron 4	SNPlex
rs11918967	Intron 4	SNPlex
rs7622120	Intron 4	SNPlex
rs590386	3' UTR	SNPlex
<i>FZD5</i>		
rs10188753	5' upstream	SNPlex
rs718290	5' upstream	SNPlex
rs7582078	5' upstream	SNPlex
rs2010400	Promoter	SNPlex
rs6708488	Promoter	SNPlex
rs3731568	3' UTR	SNPlex

4.9.1. Results

4.9.1.1. Single-point statistical analysis

All polymorphisms were in HWE in the control samples. Bonferroni corrections were applied using the methodology of Nicodemus et al [76].

a) *WNT5A*

We tested 13 haplotype-tagging SNPs in the *WNT5A* gene region from the HapMap to genotype all samples from the South African Coloured population. The threshold for significance was set at 0.008, since 3 independent LD blocks and 3 independent SNPs were determined for this region (Figure 42). Under this criterion, none of the SNPs reached statistical significance in either the case-control (Table 82) or TDT analyses (Table 83). However, some nominal significant p values were observed in the case-control (rs566926, rs7622120) and TDT (rs590386) analyses. None of these weak associations were found in both the case-control and TDT and are therefore probably false positives.

TABLE 82. Single-point statistical analysis in the *WNT5A* case-control study.

Polymorphism	Allele		Genotype frequencies						Cases vs Controls p value ^c
			Cases (n ^a =432)			Controls (n=482)			
	1 ^b	2	11	12	22	11	12	22	
rs3796232	G	C	0.27	0.50	0.23	0.22	0.48	0.30	0.06
rs1795651	A	G	0.10	0.41	0.49	0.11	0.43	0.46	0.76
rs7624718	G	A	0.02	0.15	0.83	0	0.18	0.82	0.07
rs557077	C	T	0.22	0.46	0.32	0.21	0.48	0.31	0.85
rs566926	A	C	0.04	0.21	0.75	0.04	0.29	0.67	0.03
rs648872	T	C	0.02	0.27	0.71	0.03	0.27	0.70	0.85
rs815541	G	C	0	0.12	0.88	0.01	0.15	0.84	0.28
rs9311564	G	A	0.01	0.08	0.91	0	0.07	0.93	0.70
rs472631	C	T	0.14	0.42	0.44	0.16	0.46	0.38	0.22
rs556874	A	G	0.09	0.39	0.52	0.10	0.46	0.44	0.06
rs11918967	G	C	0.19	0.46	0.35	0.23	0.48	0.29	0.09
rs7622120	A	G	0.10	0.38	0.52	0.09	0.47	0.44	0.02
rs590386	A	G	0.03	0.22	0.75	0.02	0.21	0.77	0.45

^a number of investigated individuals.

^b Allele 1 is the minor allele.

^c p value from a genotype-based χ^2 test

TABLE 83. Single-point statistical analysis in the *WNT5A* TDT.

SNP	Overtransmitted	Transmitted:Untransmitted	χ^2	p value
rs3796232	G	57 : 54	0.08	0.78
rs1795651	G	48 : 39	0.93	0.33
rs7624718	A	18 : 15	0.27	0.60
rs557077	T	55 : 40	2.37	0.12
rs566926	C	30 : 22	1.23	0.27
rs648872	C	32 : 27	0.42	0.52
rs815541	C	19 : 13	1.13	0.29
rs9311564	A	9 : 5	1.14	0.29
rs472631	-	47 : 47	0	1
rs556874	A	41 : 35	0.47	0.49
rs11918967	C	60 : 49	1.11	0.29
rs7622120	A	46 : 40	0.42	0.52
rs590386	A	27 : 14	4.12	0.04

b) FZD5

We tested 6 haplotype-tagging SNPs in the *FZD5* gene region from the HapMap to genotype in the case-control and family-based samples from the South African Coloured population. The threshold for significance was set at 0.017, since three SNPs (rs6708488, rs7582078 and rs10188753) were not independent of other SNPs (Figure 43). Under this criterion, none of the SNPs reached statistical significance in either the case-control (Table 84) or TDT analyses (Table 85). No nominally significant p values were observed either.

The rs2010400 and rs718290 SNPs were not highly polymorphic in the South African Coloured samples. In fact, no homozygous T (the minor allele) individuals were detected for the rs2010400 SNP, while only 1% of all the individuals genotyped for rs718290 were homozygous for the minor C allele.

TABLE 84. Single-point statistical analysis in the *FZD5* case-control study.

Polymorphism	Allele		Genotype frequencies						Cases vs Controls p value ^c
			Cases (n ^a =432)			Controls (n=482)			
	1 ^b	2	11	12	22	11	12	22	
rs10188753	G	A	0.10	0.40	0.50	0.09	0.41	0.50	0.91
rs718290	C	T	0.01	0.18	0.81	0.01	0.15	0.84	0.41
rs7582078	A	C	0.19	0.46	0.35	0.17	0.47	0.36	0.54
rs2010400	T	G	0	0.01	0.99	0	0.01	0.99	0.85
rs6708488	G	A	0.06	0.28	0.66	0.03	0.32	0.65	0.11
rs3731568	C	A	0.05	0.33	0.62	0.05	0.34	0.61	0.97

^a number of investigated individuals.

^b Allele 1 is the minor allele.

^c p value from a genotype-based χ^2 test

TABLE 85. Single-point statistical analysis in the *FZD5* TDT.

SNP	Overtransmitted	Transmitted:Untransmitted	χ^2	p value
rs10188753	G	49 : 33	3.12	0.08
rs718290	T	17 : 11	1.29	0.26
rs7582078	A	42 : 29	2.38	0.12
rs2010400	T	4 : 3	0.14	0.71
rs6708488	A	37 : 27	1.56	0.21
rs3731568	A	37 : 27	1.56	0.21

4.9.1.2. Haplotype analysis

Haplotypes and LD in the respective genes were analysed with Haploview using the 95% confidence interval definition from Gabriel et al [89]. This definition excludes all SNPs with allele frequencies lower than 0.05 from the analysis. Permutation testing was done for all haplotypes to help ensure that no false positives would be accepted as true indicators of association.

a) *WNT5A*

Three haplotype blocks were identified for the *WNT5A* gene and four SNPs with frequencies lower than 0.05 (rs590386, rs9311564, rs566926 and rs557077) were

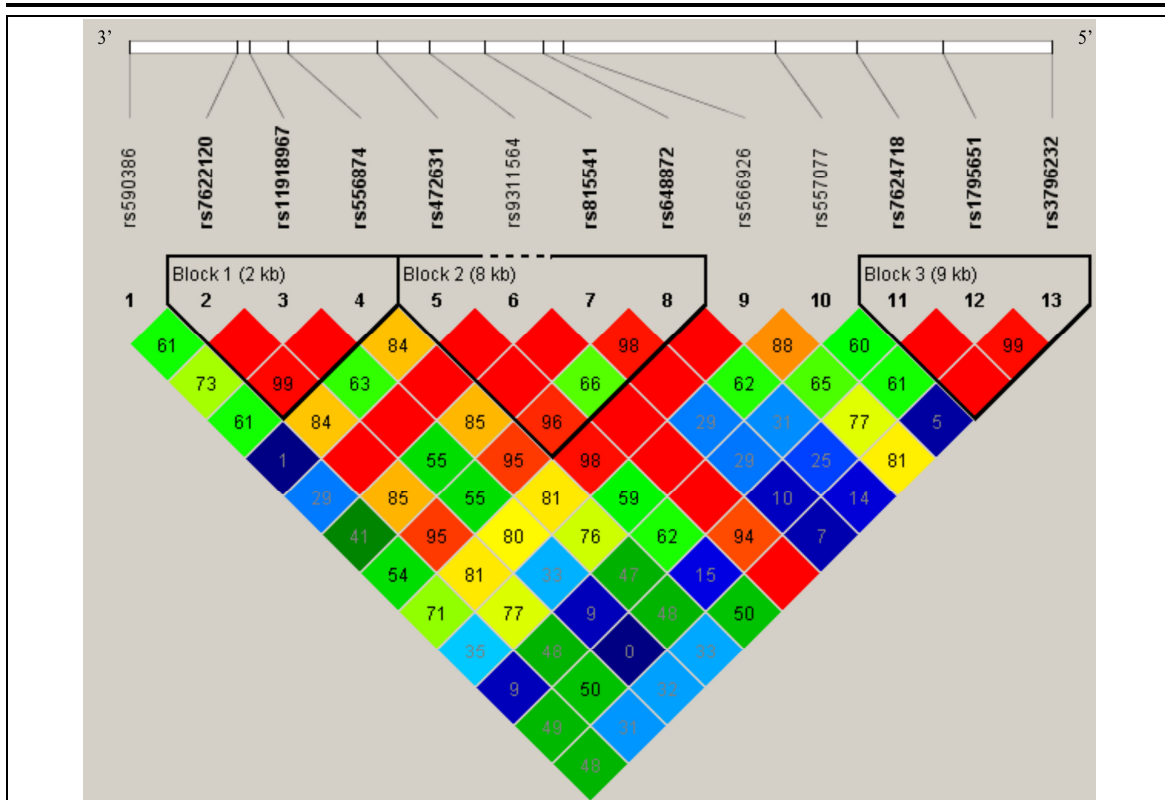


FIGURE 42. Plot of LD between *WNT5A* markers analysed in control individuals of South African Coloured descent generated by Haploview v3.3. Squares are coded in the same manner as in Figure 21.

excluded from this analysis by Haploview (Figure 42). The strength of LD across the region was varied. Eleven haplotypes were identified in three blocks in both the case-control (Table 86) and TDT (Table 87) analyses. Haplotype 1 had the highest frequency. Several of the haplotypes in the respective blocks had nominally significant p values in the case-control samples. Haplotypes 1, 2, 8 and 10 (Table 86) in the case-control analysis had p values less than 0.05, but this was not the case in the TDT analysis (Table 87).

TABLE 86. *WNT5A* case-control haplotype analysis by Haploview [313].

Block ^a	Haplotype	Freq _{Cases}	Freq _{Controls}	χ^2	p value	p value _{Permutation} ^b
1	1 G-C-G	0.58	0.53	4.85	0.03	0.21
	2 A-G-A	0.28	0.33	5.11	0.02	0.19
	3 G-G-G	0.13	0.14	0.82	0.36	0.97
2	4 T-C-C	0.49	0.45	2.49	0.11	0.67
	5 C-C-C	0.35	0.38	2.58	0.11	0.64
	6 T-C-T	0.10	0.08	1.61	0.20	0.87
	7 T-G-T	0.06	0.08	2.08	0.15	0.80
3	8 A-G-G	0.42	0.37	5.38	0.02^c	0.16
	9 A-A-C	0.31	0.32	0.55	0.46	0.99
	10 A-G-C	0.18	0.22	4.41	0.04	0.27
	11 G-G-G	0.10	0.09	0.03	0.86	1.00

^a The order of the SNPs in each block corresponds to Figure 42.

^b Permutation test p values were calculated from 10 000 permutations in Haploview.

^c p values less than 0.02 were observed on 1627 out of 10 000 occasions (16%).

TABLE 87. *WNT5A* TDT haplotype analysis by Haploview [313].

Block ^a	Haplotype	Transmitted : untransmitted	χ^2	p	p _{Permutation} ^{b, c}
1	1 G-C-G	42 : 32	1.35	0.25	0.88
	2 A-G-A	28 : 29	0.02	0.90	1
	3 G-G-G	19 : 27	1.39	0.24	0.59
2	4 T-C-C	34 : 35	0.02	0.90	1
	5 C-C-C	35 : 31	0.25	0.61	1
	6 T-C-T	14 : 12	0.07	0.79	1
	7 T-G-T	10 : 13	0.49	0.48	1
3	8 A-G-G	30 : 36	0.70	0.40	1
	9 A-A-C	36 : 32	0.25	0.61	1
	10 A-G-C	27 : 24	0.09	0.76	1
	11 G-G-G	15 : 12	0.23	0.63	1

^a The order of the SNPs in each block corresponds to Figure 42.

^b Permutation test p values were calculated from 10 000 permutations in Haploview.

^c p values less than 0.24 were observed on 5877 out of 10 000 occasions (59%)

b) *FZD5*

No haplotype blocks were identified for the *FZD5* gene region (Figure 43) and LD across the section was mostly weak. Since no haplotypes were observed using the definition from Gabriel et al, we evaluated all the SNPs together in a single haplotype (Tables 88 and 89).

Fourteen haplotypes were identified in the case-control analysis (Table 88) and 16 in the TDT (Table 89) analysis. Haplotype 1 (Table 88) had the highest frequency in the case-controls, while haplotype 2 (Table 89) had the most transmissions to affected offspring. Haplotype 9 had a nominally significant p value ($p = 0.03$) in the case-control analysis (Table 88), while haplotypes 11 and 14 in the TDT had nominally significant p values (Table 89). None of the haplotypes were associated with TB after permutation testing in either the case-control or TDT analyses.

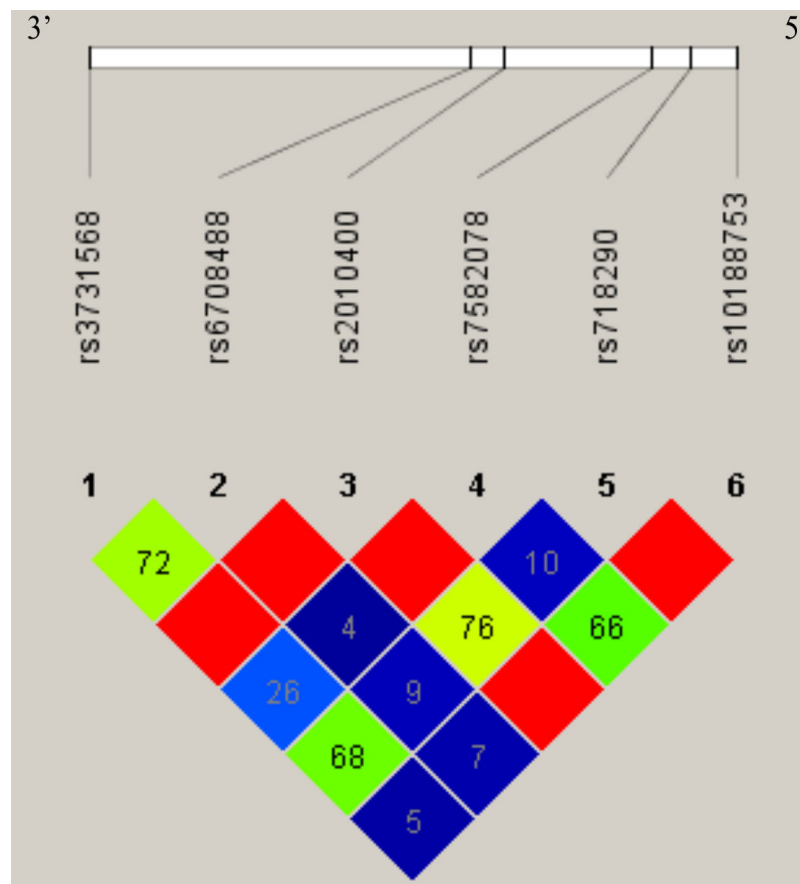


FIGURE 43. Plot of LD between *FZD5* markers analysed in control individuals of South African Coloured descent generated by Haploview v3.3. Squares are coded in the same manner as in Figure 21. No haplotype block could be calculated.

TABLE 88. *FZD5* case-control haplotype analysis by Haploview [313].

Block 1 ^a		Freq _{Cases}	Freq _{Controls}	χ^2	p value	p value _{Permutation} ^b
1	A-A-G-C-T-A	0.26	0.28	0.69	0.41	0.99
2	A-A-G-A-T-G	0.14	0.15	0.36	0.55	1
3	C-A-G-C-T-A	0.13	0.13	0.11	0.74	1
4	A-A-G-A-T-A	0.10	0.10	0	0.98	1
5	A-C-G-C-T-A	0.07	0.08	0.17	0.68	1
6	C-A-G-A-T-G	0.05	0.05	0.04	0.84	1
7	A-C-G-A-T-G	0.04	0.04	0.09	0.77	1
8	A-A-G-C-C-A	0.04	0.04	0.001	0.98	1
9	A-A-G-A-C-A	0.04	0.03	4.49	0.03 ^c	0.11
10	A-C-G-A-T-A	0.04	0.03	1.02	0.31	0.95
11	A-A-G-C-T-G	0.03	0.02	0.46	0.50	1
12	A-C-G-C-T-G	0.02	0.02	0.02	0.88	1
13	C-A-G-C-T-G	0.01	0.01	0.002	0.97	1
14	C-A-G-A-T-A	0.01	0.01	0.02	0.90	1

^a The order of the SNPs in each block corresponds to Figure 43.

^b Permutation test p values were calculated from 10 000 permutations in Haploview.

^c p values less than 0.03 were observed on 1102 out of 10 000 occasions (11%).

TABLE 89. *FZD5* TDT haplotype analysis by Haploview [313].

Block 1 ^a		Transmitted:untransmitted	χ^2	p	p _{Permutation} ^{b, c}
1	A-A-G-C-T-A	19 : 20	0.05	0.83	1
2	A-A-G-A-T-G	29 : 18	2.68	0.10	0.63
3	C-A-G-C-T-A	14 : 17	0.44	0.51	1
4	A-A-G-A-T-A	11 : 12	0.02	0.88	1
5	A-C-G-C-T-A	9 : 11	0.28	0.60	1
6	A-C-G-A-T-G	9 : 3	2.45	0.12	0.70
7	C-A-G-A-T-G	9 : 7	0.24	0.62	1
8	A-A-G-A-C-A	4 : 5	0.07	0.79	1
9	A-A-G-C-T-G	5 : 3	0.49	0.48	1
10	A-C-G-A-T-A	4 : 4	0.11	0.74	1
11	A-A-G-C-C-A	0 : 4	4.31	0.04	0.15
12	C-A-G-C-T-G	3 : 3	0.08	0.77	1
13	A-C-G-C-T-G	2 : 2	0.03	0.87	1
14	A-C-G-C-C-A	0 : 4	4.08	0.04	0.18
15	C-A-G-A-T-A	2 : 3	0.03	0.86	1
16	C-C-G-A-T-G	0 : 2	1.78	0.18	0.82

^a The order of the SNPs in each block corresponds to Figure 43.

^b Permutation test p values were calculated from 10 000 permutations in Haploview.

^c p values less than 0.04 were observed on 1455 out of 10 000 occasions (15%).

4.9.1.3. Comparison of *WNT5A* and *FZD5* with HapMap populations

We compared allele frequencies of *WNT5A* (Table 90) and *FZD5* (Table 91) polymorphisms between the South African Coloured control population and the four HapMap populations, namely 90 Yoruba individuals from Ibadan in Nigeria, 90 individuals of European descent from Utah, 45 Han Chinese individuals from Beijing and 45 Japanese from Tokyo [59,62].

The thirteen polymorphisms genotyped in *WNT5A* were compared between all populations available (Table 90). The rs9311564 SNP was dimorphic in the South

TABLE 90. Comparison of *WNT5A* allele frequencies between the South African Coloured population and the four HapMap populations [59,62].

SNP	Allele	South African Coloured	Yoruba, Ibadan, Nigeria	European, Utah	Han Chinese, Beijing	Japanese, Tokyo
rs3796232	G	0.46	0.35	0.48	0.23	0.17
	C	0.54	0.65	0.53	0.77	0.83
rs1795651	A	0.32	0.41	0.33	0.13	0.03
	G	0.68	0.59	0.67	0.87	0.97
rs7624718	G	0.09	0.09	0.14	0.06	0.13
	A	0.91	0.91	0.86	0.94	0.87
rs557077	C	0.45	0.21	0.85	0.87	0.83
	T	0.55	0.79	0.15	0.13	0.17
rs566926	A	0.18	0.02	0.21	0.60	0.75
	C	0.82	0.98	0.79	0.40	0.25
rs648872	T	0.16	0.33	0.16	0.07	0.17
	C	0.84	0.68	0.84	0.93	0.83
rs815541	G	0.08	0.01	0.16	0.07	0.14
	C	0.92	0.99	0.84	0.93	0.86
rs9311564	G	0.04	0.11	0	0	0
	A	0.96	0.89	1	1	1
rs472631	C	0.39	0.38	0.48	0.61	0.75
	T	0.61	0.62	0.53	0.39	0.25
rs556874	A	0.33	0.22	0.44	0.61	0.75
	G	0.67	0.78	0.56	0.39	0.25
rs11918967	G	0.47	- ^a	0.44	0.68	0.89
	C	0.53	-	0.56	0.32	0.11
rs7622120	A	0.33	- ^a	0.49	0.59	0.77
	G	0.67	-	0.51	0.41	0.23
rs590386	A	0.12	0.23	0.05	0.08	0.15
	G	0.88	0.78	0.95	0.92	0.85

^a These SNPs were not genotyped in this population.

African Coloured and Yoruba Nigerian populations and monomorphic in the remaining populations. This polymorphism is therefore probably of African origin. The major allele of six SNPs (rs557077, rs566926, rs472631, rs556874, rs11918967 and rs7622120) differed between the populations. For five of these six SNPs, the alternate allele was more frequent in the two Asian populations than in the other populations. However, in the South African Coloureds and Yoruba Nigerians, the rs557077 SNP major allele was different to that from the populations of European and Asian descent.

Six polymorphisms from *FZD5* (Table 91) were available for comparison between the five different populations. Two of these (rs718290 and rs2010400) were present in the South African Coloured and Yoruba Nigerian populations, but monomorphic in the other populations. These polymorphisms are therefore probably unique to populations of African descent. For the remaining polymorphisms of *FZD5*, all populations shared the same major allele for all SNPs.

TABLE 91. Comparison of *FZD5* allele frequencies between the South African Coloured population and the four HapMap populations [59,62].

SNP	Allele	South African Coloured	Yoruba, Ibadan, Nigeria	European, Utah	Han Chinese, Beijing	Japanese, Tokyo
rs10188753	G	0.29	0.33	0.47	0.27	0.32
	A	0.71	0.67	0.53	0.73	0.68
rs718290	C	0.08	0.23	0	0	0
	T	0.92	0.77	1	1	1
rs7582078	A	0.40	0.31	0.48	0.26	0.31
	C	0.60	0.69	0.52	0.74	0.69
rs2010400	T	0.01	0.05	0	0	0
	G	0.99	0.95	1	1	1
rs6708488	G	0.19	0.33	0.12	0.41	0.29
	A	0.81	0.67	0.88	0.59	0.71
rs3731568	C	0.22	0.11	0.42	0.19	0.27
	A	0.78	0.89	0.58	0.81	0.73

4.9.2. Discussion

The WNT5A protein was recently implicated in human host defence against infection with *M.tuberculosis* [553]. Both WNT5A and its receptor FZD5 may regulate IL-12 and IFN- γ production in antigen-presenting cells when exposed to mycobacteria. We used a gene-based approach and typed haplotype-tagging SNPs, selected from the HapMap, in the case-control and family-based sample collections of the South African Coloured population to test for possible associations of these novel candidate genes with TB. None of the *WNT5A* or *FZD5* SNPs genotyped was associated with disease in the case-control or TDT analyses. In addition, no haplotypes from these genes were associated with TB either.

We also compared the allele frequencies of the genotyped SNPs with those of the HapMap populations. The HapMap project genotyped one common SNP (with an allele frequency greater than 5%) every 5 kilobases in 269 DNA samples from 4 populations, namely Yoruba from Africa, European in Utah, and Chinese/Japanese [62]. We observed that certain SNPs were specific to the populations with an African ancestry. None of the other SNPs were specific to a certain population. A different major allele was noted for certain SNPs, especially in the Asian populations.

In conclusion, this is the first study to consider *WNT5A* and *FZD5* polymorphisms as susceptibility factors for TB. Since we did not detect any association with haplotype-tagging SNPs in these samples, it is possible that these genes are not involved in TB susceptibility, at least in the South African Coloured population.

CHAPTER 5

CANDIDATE GENES FROM OTHER GRANULOMATOUS DISEASES

THESE RESULTS HAVE BEEN PUBLISHED PREVIOUSLY

1. **Möller M**, Kwiatkowski R, Nebel A, van Helden P, Hoal E, Schreiber S: Allelic variation in *BTNL2* and susceptibility to tuberculosis in a South African population *Microbes and infection* 2007. 9: 522-528
2. **Möller M**, Nebel A, Kwiatkowski R, van Helden P, Hoal E, Schreiber S: Host susceptibility to tuberculosis: *CARD15* polymorphisms in a South African population *Molecular and Cellular probes* 2007. 21: 148-151

5.1. INTRODUCTION

Granulomas are present in a number of diseases. A granuloma typically consists of a centre of epithelioid macrophages surrounded by lymphocytes, and formation of this structure is thought to be a protective response against injury, inflammation or infection. The development of a granuloma is an intricate process involving the recruitment and organization of various cell types [559]. Granulomas can develop due to infectious diseases, such as tuberculosis (TB), or because of idiopathic disorders such as sarcoidosis and Crohn's disease (CD) (Table 92). Given this similarity, susceptibility genes in other granulomatous diseases provide worthwhile candidate genes for testing in TB.

TABLE 92. Comparison of TB, sarcoidosis and Crohn's disease.

	TB	Sarcoidosis	Crohn's disease
Incidence	<ul style="list-style-type: none"> • Higher in developing countries [4]. 	<ul style="list-style-type: none"> • Similar in developing and developed countries [560]. 	<ul style="list-style-type: none"> • Higher in developed countries [561].
Granuloma	<ul style="list-style-type: none"> • Caseating [19]. • Mostly present in lungs [19]. 	<ul style="list-style-type: none"> • Non-caseating [560]. • In two or more organs [560]. 	<ul style="list-style-type: none"> • Non-caseating [562]. • Mostly in the gastro-intestinal tract [562].
Aetiology	<ul style="list-style-type: none"> • Infectious pathogen. 	<ul style="list-style-type: none"> • Unknown. 	<ul style="list-style-type: none"> • Unknown.
Susceptibility factors	<ul style="list-style-type: none"> • Genetic. • Environmental. • Bacterial. 	<ul style="list-style-type: none"> • Genetic. • Environmental. 	<ul style="list-style-type: none"> • Genetic. • Environmental.

Sarcoidosis is a multiorgan disorder that affects mostly the lungs and lymphatic system. The disease is of unknown aetiology, but some evidence supports the hypothesis that it is caused by an infectious agent in genetically susceptible hosts [560]. Sarcoidosis has been considered to be a different form of TB [563], because of some similarities between the diseases, such as the presence of granulomas and coughing [564]. This is still a popular hypothesis and recently the *Mycobacterium tuberculosis* (*M.tuberculosis*) heat shock protein 70 was suggested to have an influence in the pathogenesis of sarcoidosis [565]. Susceptibility genes in sarcoidosis can therefore also be considered as candidate genes for TB and vice versa. Associations have been detected in both diseases with the same candidate genes, such as natural resistance associated macrophage protein 1 (*NRAMP1*) [210,260] and human leukocyte antigen (*HLA*) [277,566].

While TB granulomas occur mostly in the lungs, CD granulomas are found in the intestine but can also appear in the lungs. Since TB and CD share this pathophysiologic characteristic, susceptibility genes for CD are plausible candidate genes for investigation in TB. Moreover, although the cause of CD is unknown, evidence is mounting that barrier function is compromised by genetic variants that can hinder interaction with bacterial antigens [567]. While it appears that the normal flora plays a pivotal role in the pathophysiology of the disease [568], a controversial theory

[569,570] suggests that *Mycobacterium avium subspecies paratuberculosis* is one of the causative agents in CD [571]. If true, this would increase the disease similarities and the chance of overlapping genetic susceptibilities.

5.2. BUTYROPHILIN-LIKE 2 GENE

The butyrophilin-like 2 gene (*BTNL2*, *BTL-II*) was first identified by comparing the human and mouse genomic sequences at the HLA class II and class III regions [572]. It is one of the B7 receptor family genes and was hypothesised to have a possible function as a T cell co-stimulatory molecule [289], because of amino acid homology to B7-1 (CD80) and B7-2 (CD86). These co-stimulatory molecules were found to be important for an effective T cell response against mycobacteria in a mouse model of TB [573]. A recent study determined that *BTNL2* inhibits T cell activation in mice [574]. Recently, a truncating splice site variation in *BTNL2* was associated with sarcoidosis in a German population [289]. This association was subsequently confirmed by others in an additional population, namely white Americans [575]. Interestingly, this single nucleotide polymorphism (SNP) alone (rs2076530) was associated with sarcoidosis only in white populations but not in African American patients. However a three-locus haplotype including this SNP was associated with the disease in African Americans [575]. The rs2076530 SNP alters the splicing pattern of the gene and a premature stop codon is formed. The protein product of this splice form is truncated and could limit the T cell downregulatory function of *BTNL2* [289].

It has been shown that *M.tuberculosis* and lipopolysaccharides can induce *BTNL2* in monocyte-derived macrophages [289]. Variations in *BTNL2* could therefore have a possible function in genetic susceptibility or resistance to TB. In this study, we sequenced *BTNL2* and investigated the influence of the gene in TB, applying population-based and family-based association study designs in South African Coloured population.

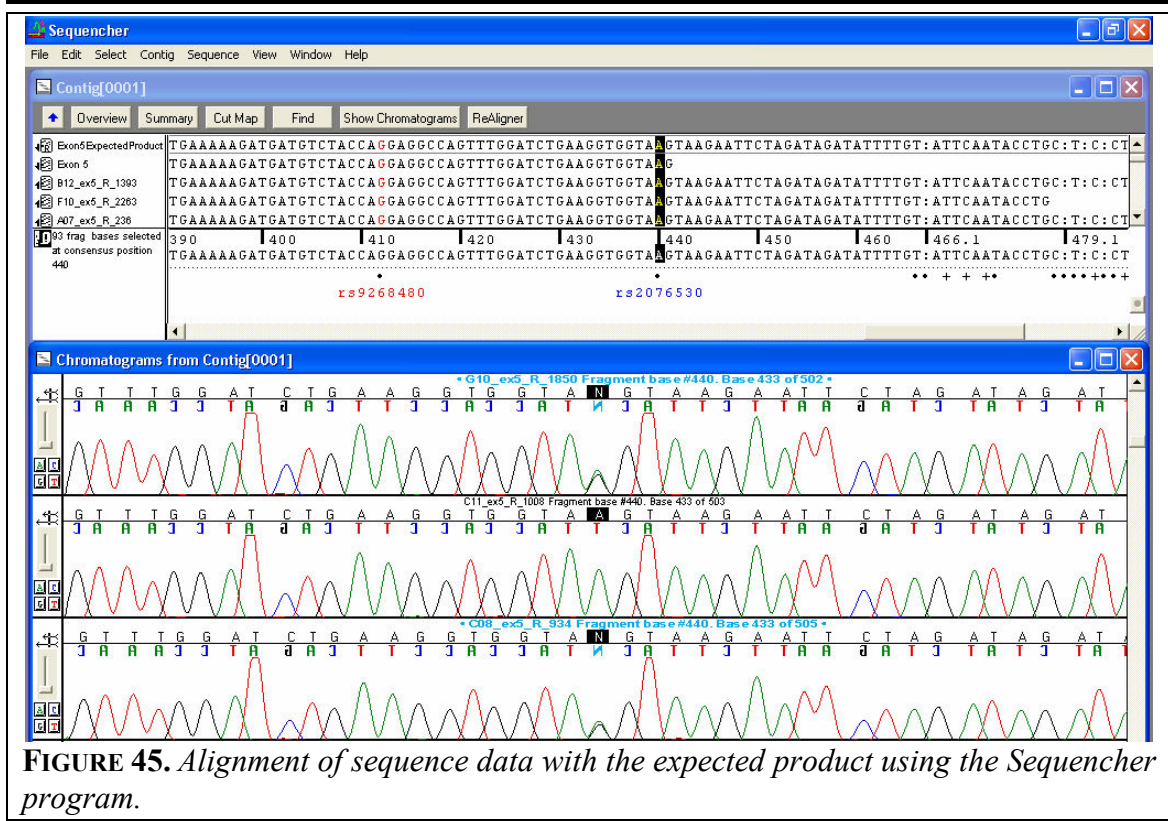
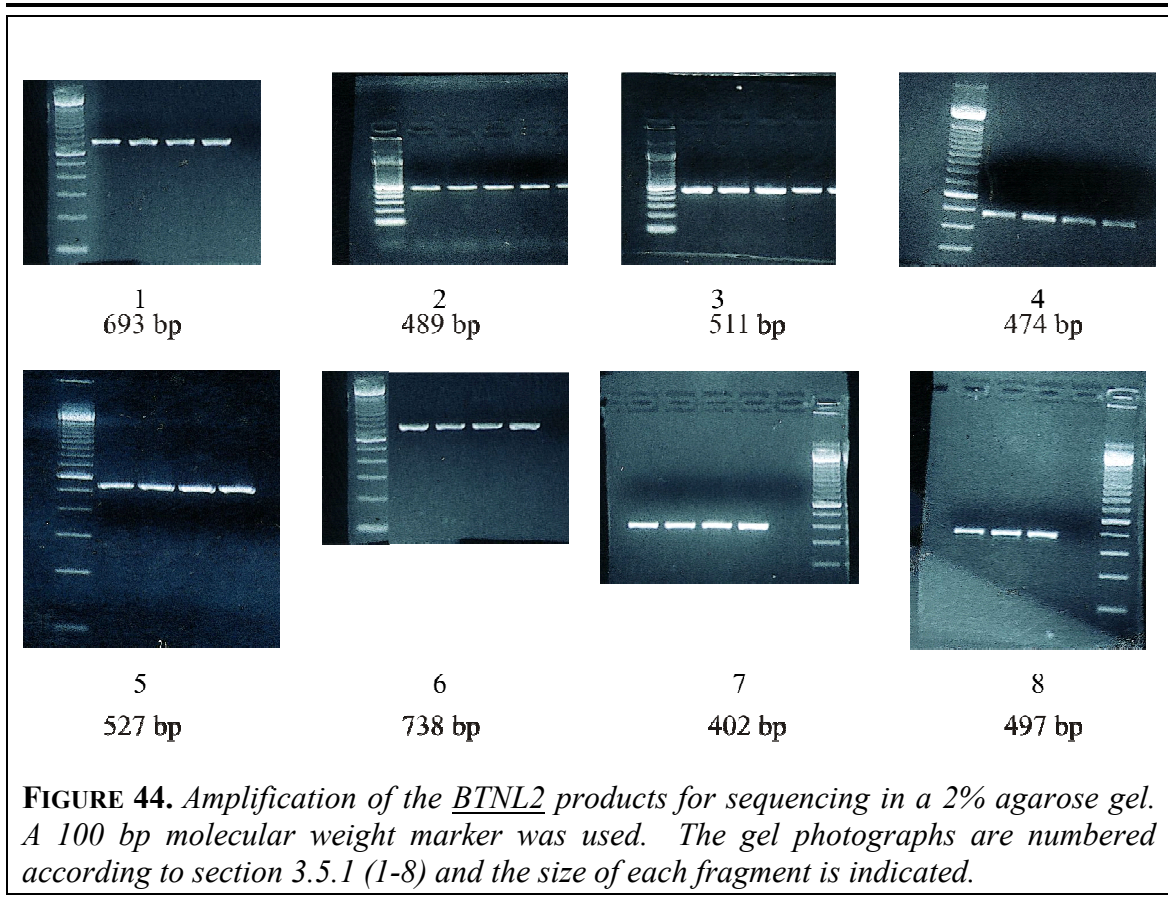
5.2.1. Results

5.2.1.1. Mutation detection

The putative promoter region, 8 exons and flanking regions of the gene were amplified (*Section 3.5.1.*) and the products were visualised in 2% agarose gels (Figure 44). Chromatograms were produced during sequencing of the PCR products and were used to determine the quality of the sequencing. Clear peaks with low background signal were an indication of good quality.

The alignment of sequence data with the expected product was performed with Sequencher version 4.2 (Figure 45).

SGCaller [288] was used to automatically genotype the observed variations in all the sequenced samples (Figure 46). The calls from the program were verified by concurrently viewing the chromatogram. Sequencing of *BTNL2* in the South African Coloured population led to the identification of 76 SNPs: 5 upstream of the gene, 29 exonic, 31 intronic and 11 downstream of the gene. Of the SNPs detected, 15 were novel, but were either in introns or had a very low allele frequency. The allele frequencies, chromosome position and rs-numbers are listed in Addendum A6.



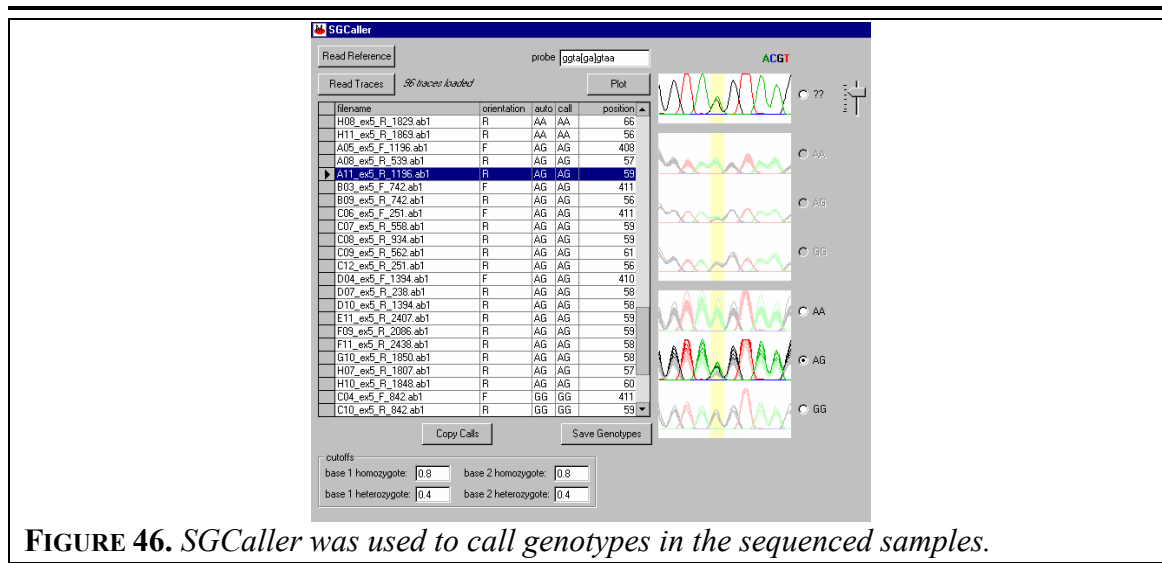


FIGURE 46. SGCcaller was used to call genotypes in the sequenced samples.

TABLE 93. Single-point statistical analysis in the *BTNL2* case-control study design.

SNP	Allele		Genotype frequencies						Cases vs Controls p value ^c
			Cases (n ^a =432)			Controls (n=482)			
	1 ^b	2	11	12	22	11	12	22	
rs3817974	G	A	0.01	0.25	0.74	0.02	0.22	0.76	0.32
rs17208699	C	A	0.01	0.19	0.80	0.01	0.17	0.82	0.44
rs3817973	A	G	0.11	0.38	0.51	0.08	0.36	0.56	0.21
rs3817972	A	G	0.01	0.20	0.79	0.02	0.17	0.81	0.28
ss65713168	T	C	0.01	0.23	0.76	0.01	0.21	0.78	0.39
ss65713167	C	A	0.01	0.23	0.76	0.02	0.21	0.77	0.36
ss65713164	A	G	0.01	0.23	0.76	0.02	0.21	0.77	0.36
ss65713163	T	C	0.01	0.23	0.76	0.02	0.21	0.77	0.37
rs2076530	G	A	0.11	0.40	0.49	0.07	0.38	0.55	0.065
rs9268480	A	G	0.05	0.25	0.70	0.02	0.22	0.76	0.023
rs2076529	G	A	0.11	0.39	0.50	0.07	0.38	0.55	0.083
ss65713152	A	T	0.01	0.25	0.74	0.02	0.24	0.74	0.74
ss65713150	T	C	0.01	0.25	0.74	0.02	0.24	0.74	0.73
rs3763317	G	A	0.18	0.48	0.34	0.18	0.47	0.35	0.94
rs9268501	A	C	0.15	0.45	0.40	0.15	0.45	0.40	0.99
rs9268502	G	A	0.15	0.45	0.40	0.15	0.45	0.40	0.98
rs9268503	G	A	0.14	0.46	0.40	0.15	0.45	0.40	0.96
rs9268504	C	G	0.15	0.45	0.40	0.15	0.45	0.40	0.98

^a number of investigated individuals.

^b Allele 1 is the minor allele.

^c p value from a genotype-based χ^2 test

5.2.1.2. Single-point SNP analyses in *BTNL2*

According to the set inclusion criteria (*Section 3.5.*), 18 SNPs were genotyped in the entire DNA collection (Table 93). The influence of these SNPs in TB was investigated with a case-control study design and a transmission disequilibrium test (TDT). All genotyped SNPs were in Hardy-Weinberg equilibrium (HWE) in the control group. The level of significance was 0.017 after Bonferroni correction. Under this criterion, none of the analysed SNPs reached statistical significance. The SNP rs2076530 was not significantly associated with TB in our case-control study ($p = 0.065$, Table 93) or TDT ($p = 0.44$, Table 94). Comparisons of the genotype frequencies of the remaining 17 SNPs analysed showed a significant p value of 0.023 for the rs9268480 SNP (Table 93) in the case-control group. Given the adjusted significance level in our study, this p value is not considered to reflect an association. None of the other SNPs examined were found to be associated with the disease. Altogether, single marker analysis detected no association between TB and any of the SNPs investigated in the *BTNL2* gene.

TABLE 94. Single-point statistical analysis in the *BTNL2* TDT.

SNP	Overtransmitted	Transmitted:Untransmitted	χ^2	p value
rs3817974	G	30-26	0.286	0.59
rs17208699	G	22-20	0.095	0.76
rs3817973	C	42-35	0.636	0.43
rs3817972	G	22-21	0.023	0.88
ss65713168	T	30-24	0.667	0.41
ss65713167	C	29-23	0.692	0.41
ss65713164	A	28-21	1	0.32
ss65713163	T	29-25	0.296	0.59
rs2076530	A	44-37	0.605	0.44
rs9268480	C	30-18	3	0.083
rs2076529	A	46-38	0.762	0.38
ss65713152	T	22-15	1.324	0.25
ss65713150	C	22-15	1.324	0.25
rs3763317	C	60-44	2.462	0.12
rs9268501	T	63-43	3.774	0.052
rs9268502	C	63-43	3.774	0.052
rs9268503	C	63-43	3.774	0.052
rs9268504	C	63-43	3.774	0.052

5.2.1.3. Haplotype analyses

The Haploview program [313] identified three haplotype blocks (Figure 47) for the markers genotyped in *BTNL2* in the case-control samples studied. Strong linkage disequilibrium (LD) exists across the *BTNL2* gene region in the South African Coloured population (Figure 47). Haplotype 6 (Table 95) was tagged by the minor A allele of rs9268480 and although it was nominally significant, evaluation of global significance between cases and controls with the haplotyping program Cocophase [101] yielded a

non-significant global p-value ($p > 0.05$) for each of the three haplotype blocks after 10 000 permutations (Table 95).

TABLE 95. *BTNL2* case-control haplotype analysis by Cocaphase.

Block ^a	Global test	Individual tests	p value
	p value ^b	Haplotype (case/control frequency)	
1	0.17	1 A-A-G-G-C-A-G-C-A-G-A (0.694/0.733)	0.067
		2 A-A-G-G-C-A-G-C-G-G-A (0.00490/0.00633)	0.39
		3 A-A-G-G-C-A-G-T-A-G-A (0/0.00106)	0.27
		4 A-A-G-G-T-C-G-C-A-G-A (0.00123/0)	0.21
		5 A-A-A-G-C-A-G-C-A-G-A (0.00123/0.00527)	0.13
		6 A-A-A-G-C-A-G-C-G-A-G (0.173/0.132)	0.017
		7 G-C-A-A-T-C-A-T-G-G-G (0.0993/0.0949)	0.76
		8 G-A-A-A-C-A-G-C-G-G-G (0.00735/0.00422)	0.38
		9 G-A-A-A-T-C-A-T-G-G-G (0.00123/0.00317)	0.38
		10 G-A-A-G-T-C-A-T-G-G-G (0.0184/0.0200)	0.80
2	0.91	11 A-T (0.139/0.141)	0.91
		12 T-C (0.861/0.859)	0.91
3	0.83	13 G-C-A-A-G (0.0454/0.0367)	0.36
		14 G-A-G-G-C (0.375/0.375)	0.98
		15 G-A-G-A-C (0.00119/0.00105)	0.93
		16 A-C-A-A-G (0.579/0.798)	0.72

^a The order of the SNPs in each block corresponds to Figure 47.

^b Permutation test p values were calculated from 10 000 permutations in Cocaphase [101], to correct for multiple testing while taking account of correlation between markers and haplotypes.

Haploview also identified the same three haplotype blocks for the markers genotyped in *BTNL2* in the samples from the TB families. None of the estimated haplotypes were nominally significant and therefore evaluation of global significance for the TDT with the TDTphase program [101] also gave a non-significant global p-value ($p > 0.05$) for each of the three haplotype blocks after 10 000 permutations (Table 96).

The TDTphase and Cocaphase programs infer haplotype phase. Only 12 haplotypes were estimated in the families in comparison with 16 estimated in the case-control haplotype analysis. Four haplotypes (haplotypes 3, 4, 9 and 15 in Table 95) were not present in the family samples (Table 96) and had extremely low frequencies in the case-control group. The absence of these haplotypes in the families can therefore be due to their low frequencies in the general population and do not represent a real difference between the sample groups.

TABLE 96. *BTNL2* TDT haplotype analysis by TDTphase.

Block ^a	Global test	Individual tests		p value
	p value ^b	Haplotype (transmitted/untransmitted)		
1	0.41	1	A-A-G-G-C-A-G-C-A-G-A (159/151)	0.32
		2	A-A-G-G-C-A-G-C-G-G-A (0/1)	0.24
		3	A-A-A-G-C-A-G-C-A-G-A (1/0)	0.24
		4	A-A-A-G-C-A-G-C-G-A-G (17/24)	0.25
		5	G-C-A-A-T-C-A-T-G-G-G (13/16)	0.56
		6	G-A-A-A-C-A-G-C-G-G-G (1/2)	0.56
		7	G-A-A-G-T-C-A-T-G-G-G (5/2)	0.25
2	0.62	8	A-T (19/22)	0.62
		9	T-C (206/203)	0.62
3	0.23	10	G-C-A-A-G (3/6)	0.31
		11	G-A-G-G-C (40/25)	0.12
		12	A-C-A-A-G (26/38)	0.20

^a The order of the SNPs in each block corresponds to Figure 47. ^b Permutation test p values were calculated from 10 000 permutations in TDTphase [101], to correct for multiple testing while taking account of correlation between markers and haplotypes.

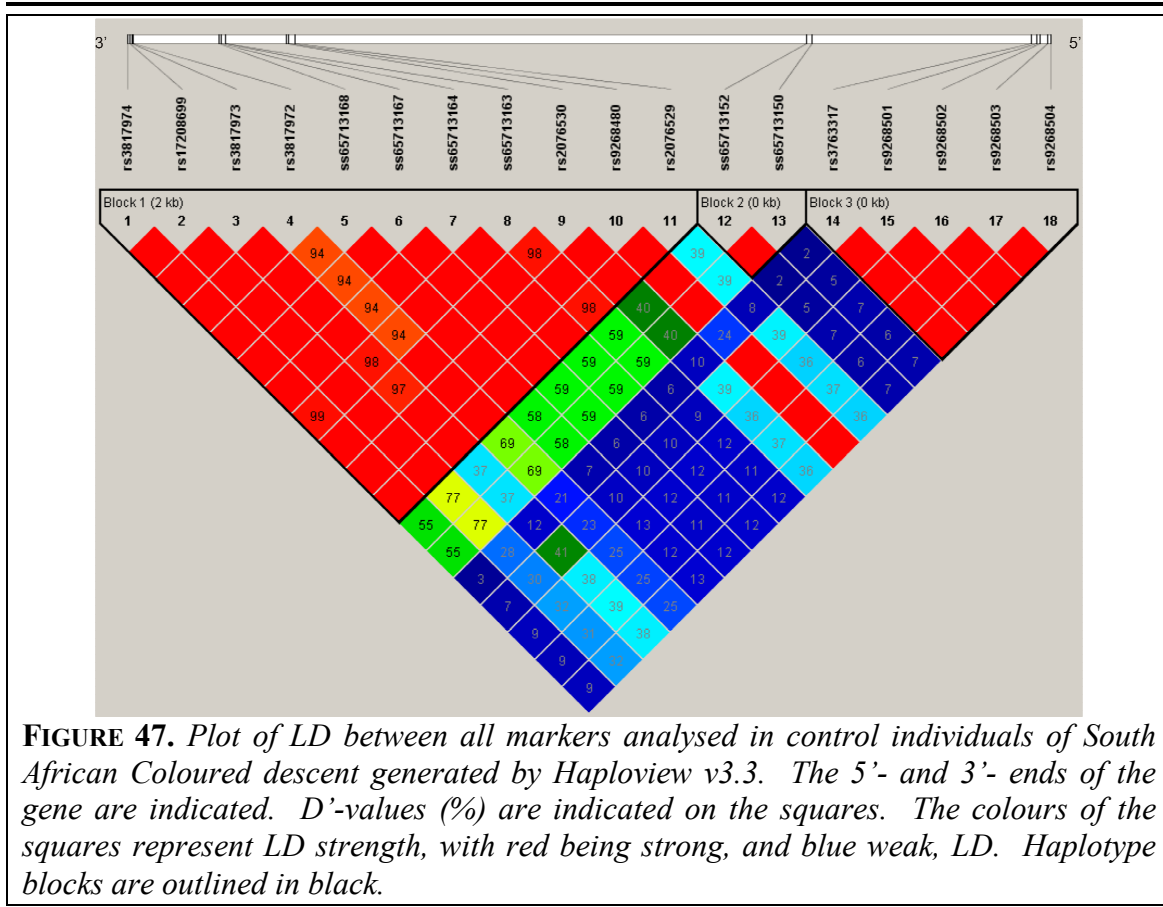


FIGURE 47. Plot of LD between all markers analysed in control individuals of South African Coloured descent generated by Haploview v3.3. The 5'- and 3'- ends of the gene are indicated. D'-values (%) are indicated on the squares. The colours of the squares represent LD strength, with red being strong, and blue weak, LD. Haplotype blocks are outlined in black.

5.2.1.4. Comparison of *BTNL2* data between South Africans and Americans

The South African control data was compared to previously published genotyping data from African-American and white American individuals [575]. Rybicki et al. [575] detected 10 SNPs in the exon 5 / intron 5 region of *BTNL2*. For the analysis, data on three overlapping SNPs, namely rs2076529, rs9268480 and rs2076530, were available from both studies. Comparison of the 3 SNPs in common revealed that the South African allele frequencies were in general more similar to those of the African-Americans than to the white Americans. In addition, 2 of the 5 other SNPs (ss38346932 and ss38346939) found by Rybicki et al. [575] were identified by sequencing 24 control individuals from the South African samples (Table 97). The frequencies of these 2 SNPs were also more similar to the African-American population than the white Americans.

5.2.1.5. Comparison of *BTNL2* data between South African Coloureds and Germans

Comparative analyses were done of 482 South African Coloured controls from this study and 533 German controls from another study [289] using 6 SNPs that were genotyped in both populations (Table 97). The allele frequencies were very similar for 4 of the SNPs, but in 2 instances differed significantly (rs2076530 $p = 0.03$ and rs2076529 $p = 0.04$). The allele frequencies for these SNPs in Germans were similar to those found in white Americans [575]. With the block definition used in Haploview, two haplotype blocks were identified in the South African population and three in the German population. The reason for this difference is the slightly stronger LD observed in South African Coloureds across this region (Figure 48).

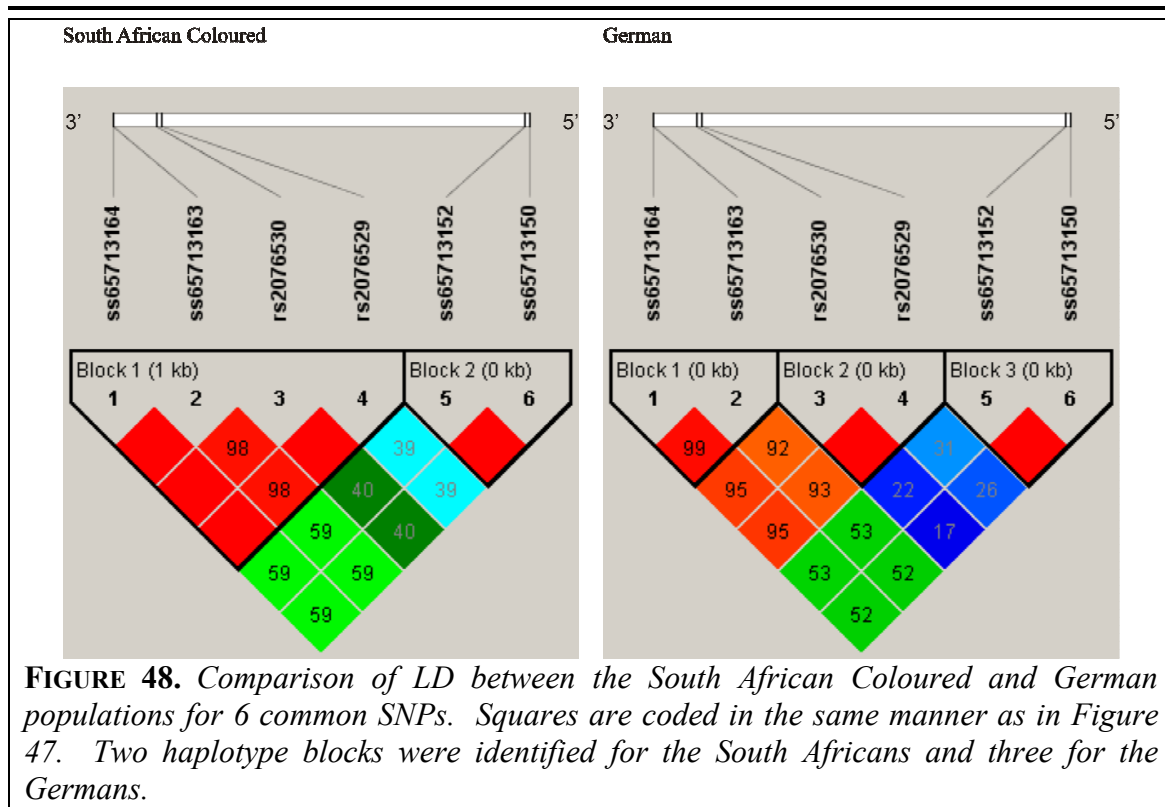


TABLE 97. Comparisons of the allele frequencies in the African-American, white American, South African Coloured and German control populations.

SNP ^a	African Americans [575]			White Americans [575]			South African Coloured controls			Germans [289]		
	Allele	Frequency	Het freq	Allele	Frequency	Het freq	Allele	Frequency	Het freq	Allele	Frequency	Het freq
ss65713164	Not genotyped	Not genotyped	Not genotyped	G	0.88	0.21	G	0.85	0.26	A	0.15	0.26
ss65713163	Not genotyped	Not genotyped	Not genotyped	C	0.87	0.21	C	0.85	0.26	T	0.15	0.26
rs34423804	A 0.98 T 0.02	0.033	0.001	A T 0.001	0.999 0.001	0.001	A T	0.98 0.02	0.042	Not genotyped	Not genotyped	Not genotyped
rs36110770	C 0.98 T 0.02	0.032	0.001	C T 0.001	0.999 0.001	0.001	Not detected by sequencing	Not detected by sequencing	Not genotyped	Not genotyped	Not genotyped	Not genotyped
rs2076529	A 0.70 G 0.30	0.42	0.52	A G 0.39	0.61 0.39	0.52	A G	0.74 0.26	0.38	A G	0.59 0.41	0.50
rs28362679	C 0.98 T 0.02	0.03	0.04	C T 0.02	0.98 0.02	0.04	Not detected by sequencing	Not detected by sequencing	Not genotyped	Not genotyped	Not genotyped	Not genotyped
rs9268480	G 0.85 A 0.15	0.26	0.42	G A 0.29	0.71 0.29	0.42	G A	0.87 0.13	0.21	Not genotyped	Not genotyped	Not genotyped
rs35037492	G 0.98 A 0.02	0.045	0.001	G A 0.001	0.999 0.001	0.001	Not detected by sequencing	Not detected by sequencing	Not genotyped	Not genotyped	Not genotyped	Not genotyped
rs2076530	A 0.70 G 0.30	0.42	0.52	A G 0.40	0.60 0.40	0.52	A G	0.74 0.26	0.38	A G	0.58 0.42	0.51
rs17202456	C 0.87 T 0.13	0.23	0.16	C T 0.09	0.91 0.09	0.16	C T	0.85 0.15	0.21	Not genotyped	Not genotyped	Not genotyped
ss65713152	Not genotyped	Not genotyped	Not genotyped	T A 0.14	0.86 0.14	0.24	T A	0.86 0.14	0.24	T A	0.92 0.08	0.15
ss65713150	Not genotyped	Not genotyped	Not genotyped	C T 0.14	0.86 0.14	0.24	C T	0.86 0.14	0.24	C T	0.92 0.08	0.15

^a SNPs marked in bold were genotyped in all the samples of the South African population, while genotyping data for the SNPs in normal text were derived from sequencing 48 individuals.

^b Het freq, Heterozygote frequency.

5.2.2. Discussion

TB and sarcoidosis are both granulomatous diseases, occur mainly in the lungs and can cause coughing and weight loss [564]. Because of these similarities in presentation of the diseases, we investigated the involvement of *BTNL2*, a susceptibility gene for sarcoidosis, in tuberculosis, using samples of South African Coloured descent. This gene was first identified as a sarcoidosis susceptibility factor by Valentonyte et al. [289], and Rybicki et al. confirmed the association between sarcoidosis and rs2076530 in white Americans, but not in African-Americans [575]. We did not observe a significant association of the truncating splice site SNP (rs2076530) with TB at the genotype level. Interestingly, the truncating allele A was more frequent in controls. Since *BTNL2* has a T cell downregulatory effect in mice [574], individuals with the truncated product could have a stronger T cell response, depending on which T cell subset was affected most. This could be an advantage in TB, since resistance to TB is dependant on the T helper 1 cell response of the infected individual. The rs9268480 SNP adjacent to rs2076530 showed a nominally significant association, but this disappeared on stringent correction. *BTNL2* is on chromosome 6, which also carries the *HLA* genes. Rybicki et al. [575] reported a high level of LD between rs9268480 and *HLA-DRB1* for both white Americans and African-Americans. Two recent studies reported that the association observed between *BTNL2*, multiple sclerosis [576] and Graves' disease [577] was secondary to *HLA-DRB1*. It is therefore possible that in the South African samples rs9268480 is also in LD with this region, which has been implicated as a susceptibility factor in TB [152,165,277]. None of the 16 other SNPs analysed in *BTNL2* were associated with disease (Tables 93 and 94) and no association was found between any of the estimated haplotypes and TB (Tables 95 and 96).

Comparisons between the South African Coloured and German populations revealed slightly stronger LD for the South African population in the *BTNL2* region. For this reason, two haplotype blocks were identified in the South African Coloured population and three in the Germans. Stronger LD is expected in admixed populations [578] such as the South African Coloured population.

The South African Coloured population has several different parental populations, including European whites and sub-Saharan African populations. The German population is outbred and genetically relatively homogeneous and could (together with the white Americans) be considered as an approximation for the white parental population, while the African-Americans could be evaluated as another admixed population [578]. The comparison of the allele frequencies for polymorphisms in exon 5 / intron 5 (Table 97) between the populations revealed that, for this restricted region of the genome, the South African Coloured population is more similar to African-Americans than to whites. Therefore, the lack of association with the functional *BTNL2* SNP and TB in the South African Coloured population could reflect the African admixture and would not preclude a role of rs2076530 in a white TB population, in a similar manner as seen with sarcoidosis [575]. Future work should evaluate *BTNL2* in a TB sample set with no African genetic contribution, to determine if the association of this gene with disease could be population-specific. Due to the high incidence of latent TB infection in the control community (unpublished results), we have tested the role of *BTNL2* in disease progression only and it is still possible that this gene may play a role in primary infection.

5.3. CASPASE RECRUITMENT DOMAIN-CONTAINING PROTEIN 15 GENE

The caspase recruitment domain-containing protein 15 gene (*CARD15*) encodes the nucleotide-binding oligomerization domain 2 protein (NOD2) [569] and was identified as a susceptibility gene for CD [308,579,580]. The R702W (rs2066844), G908R (rs2066845) and 1007fs (rs2066847) polymorphisms in *CARD15* were independently associated with this form of inflammatory bowel disease [579] in European, white American and French Canadian populations. The R702W variant, caused by a C to T transition, results in an arginine to tryptophan change in NOD2, while G908R is formed by a G to C transversion and changes a glycine to arginine. The 1007fs polymorphism is formed by a cytosine insertion at the second nucleotide in codon 1007 [580]. This results in a frameshift and substitutes leucine for proline, which leads to the truncation of NOD2 in the leucine-rich repeat region of the protein [579]. The truncated NOD2 protein formed as a consequence of the 1007fs insertion in *CARD15* leads to a decrease in the responsiveness of NF- κ B to the bacterial ligand muramyl dipeptide [580]. Muropeptides are breakdown products of peptidoglycans from Gram-negative and Gram-positive bacteria, which are also present in the cell walls of *M. tuberculosis* [581]. Mycobacterial infection induces activation of the transcription factor NF- κ B [582]. R702W and G908R, like 1007fs, also occur in the vicinity of the leucine-rich repeat region of the protein and compromise the recognition of muramyl dipeptides [583].

It was recently demonstrated that NOD2 is an essential recognition system of *M. tuberculosis* [522] and *CARD15* could consequently be considered as a candidate gene in TB. In a case-control study with Gambian TB patients [240] no associations were found with promoter polymorphisms of *CARD15*, and the three CD variants R702W, G908R and 1007fs were absent in the Gambian population. We therefore aimed to genotype these CD-associated polymorphisms in TB patients and controls from the admixed South African Coloured population and conducted a case-control association study to determine if the variants predispose to TB.

5.3.1. Results

The three CD polymorphisms in *CARD15* were genotyped in TB and control samples from the study population. Genotyping for the three variants was done on an automated platform with the TaqMan[®] system (Applied Biosystems, Foster City, USA) and assays were designed as published previously [308,309]. All variants were in HWE.

The genotype (Table 98) and allele (Table 99) frequencies for the polymorphisms were very low and the minor alleles were found only in the heterozygous state. Comparison of the genotype frequencies for the three polymorphisms revealed no association between any of the variants and TB (Table 98). Since all three *CARD15* variants are relatively rare and have been shown to occur on separate haplotypes [579], we compared the frequencies of carriers (i.e. heterozygote and double heterozygote carriers of any of the three SNPs) between the cases and the controls. The carrier frequencies were not significantly different in the two groups (0.03 vs. 0.05; $p = 0.22$).

The TDT showed similar results (Table 100) to the case-control study, with extremely low frequencies of the minor alleles. The R702W and G908R polymorphisms were

found in 5 and 4 families respectively, while the 1007fs variant was not present in any of the 185 families genotyped.

TABLE 98. Genotype frequencies of the *CARD15* variants investigated in the South African Coloured population using a case-control study design.

	R702W		G908R		1007fs	
TB	CC	0.986 (421) ^a	GG	0.995 (426)	wt/wt	0.991 (423)
	CT	0.014 (6)	GC	0.005 (2)	wt/insC ^b	0.009 (4)
	TT	0	CC	0	insC/insC	0
Controls	CC	0.977 (468)	GG	0.992 (476)	wt/wt	0.983 (423)
	CT	0.023 (11)	GC	0.008 (4)	wt/insC	0.017 (8)
	TT	0	CC	0	insC/insC	0
p-value	0.46		0.69		0.39	

^a The number of individuals with each genotype is indicated in brackets.

^b insC, cytosine insertion

TABLE 99. Allele frequencies of the *CARD15* variants investigated in the South African Coloured population in a case-control study design.

	R702W		G908R		1007fs	
TB	C	0.993 (848) ^a	G	0.998 (854)	C	0.995 (850)
	T	0.007 (6)	C	0.002 (2)	insC ^b	0.005 (4)
Controls	C	0.989 (947)	G	0.996 (956)	C	0.992 (950)
	T	0.011 (11)	C	0.004 (4)	insC	0.008 (8)
p-value	0.47		0.69		0.40	

^a The number of alleles is indicated in brackets.

^b insC, cytosine insertion

TABLE 100. TDT results for *CARD15*.

SNP	Minor allele	Families^a	Frequency^b	Transmitted:Untransmitted
R702W	T	5	0.01	4 : 1
G908R	C	4	0.006	2 : 2
1007fs	insC	0	0	0

^a The number of families where at least one individual has the minor allele is indicated.

^b Frequency of the minor allele.

5.3.2. Discussion

The pathophysiology of both TB and CD is characterised by the presence of granulomas. We investigated the role of the CD susceptibility gene *CARD15* in TB in the South African Coloured population, because NOD2, the protein encoded by *CARD15*, was found to be an essential recognition mechanism for *M. tuberculosis* [522]. To this end, three variants that affect the leucine-rich-repeat recognition domains

of NOD2 and that represent the main genetic risk associated with this gene in CD were genotyped.

This is the first report of allele frequencies for the R702W, G908R and 1007fs polymorphisms in TB samples from the South African Coloured population. The variant alleles of these polymorphisms are found at frequencies of about 2%, 3% and 3% respectively in the European populations [561], but were present at frequencies below 1.1% in this study collection in both the family and case-control groups (Tables 98 and 100). The frequencies of two (R702W, G908R) of the three variants genotyped in this control group were not significantly different ($p > 0.05$) from those observed in controls in a CD study of another South African Coloured sample set, which detected minor allele frequencies of 0.03 for R702W, 0.005 for G908R and 0.04 for 1007fs [584]. A genetic input from European ancestors is the most likely explanation for the presence of the CD variants in this admixed South African population, since these polymorphisms were not found at all in Gambians [240] or various Asian populations [561,585]. However, at least for the *CARD15* locus, it seems that the genetic contribution from Europeans was relatively small.

TB is present in epidemic proportions in the South African Coloured population from the Western Cape province, while the occurrence of CD in this population is relatively rare [586]. In other African populations with a low incidence of CD, the polymorphisms associated with CD were absent [240]. The rarity of the CD-associated *CARD15* polymorphisms in the South African Coloured population could in part explain the lower incidence of CD observed in this group [586], but other factors – such as the availability of health care and the higher frequency of infectious diarrhoea – could hinder the diagnosis of CD in this population [587]. Low frequencies for these mutations have also been observed in African-American CD patients [587]. The variants were previously investigated in CD [584] in an independent sample set of South African Coloureds ($n_{\text{cases}} = 76$, $n_{\text{controls}} = 100$), but no association was detected with the disease.

We did not find any statistically significant association between the polymorphisms studied and TB due to their low allele frequencies in both the South African Coloured case and control samples and family samples. This study therefore did not have enough power to detect small effect sizes. The common disease-common variant hypothesis states that the risk of contracting a common disease is determined by the presence of common variants in the population [588]. It can also be assumed that the detection of disease genes is facilitated in a population in which high exposure and high prevalence brings most of the genetic susceptibility alleles to manifestation. Since TB is a common disease in the South African Coloured population, the low allele frequencies of the *CARD15* mutations in this population indicate that the gene with its rare variants does not play a major role in predisposing to the disease. These variants could however still be predisposing factors in other populations where they are more frequently present. An investigation of TB cases and controls from another population with a higher frequency of *CARD15* polymorphisms is therefore desirable.

In summary, we determined that *CARD15* is not a major susceptibility factor in South African Coloured TB patients. The low frequency of these variants in the population studied makes it unlikely that they play any role in disease aetiology.

CHAPTER 6

**CANDIDATE GENES WHICH ARE
DIFFERENTIALLY EXPRESSED
BETWEEN TUBERCULOSIS CASES AND
CONTROLS**

6.1. INTRODUCTION

At any given time, four groups of individuals who respond differently to infection with *Mycobacterium tuberculosis* (*M.tuberculosis*) can be identified in a community affected by tuberculosis (TB) [271]. These are 1) patients with active TB, named the “active” group, 2) individuals who have had several episodes of TB, even though they responded to and were cured after six months of TB treatment, designated as the “recurrent” group, 3) individuals who had TB only once and who were cured after chemotherapy, classified as the “cured” group, and 4) *M.tuberculosis*-infected individuals with a positive Mantoux test, who have never presented with the active disease, designated the “latent” group. A recent paper [271], in which DNA-array analysis experiments were done in the South African Coloured population, identified 337 genes which were differentially expressed in the blood of the four groups. Of these differentially expressed biomarkers, nine (Table 101) could be used to discriminate between the outcomes after infection and disease as described above.

TABLE 101. *Genes differentially expressed between the four TB groups. From [271].*

Symbol	Gene name	Function
<i>RIN3</i>	Ras and Rab interactor 3	Intracellular trafficking, endosome
<i>LY6G6D</i>	Lymphocyte antigen 6 complex, locus G6D	Immunological, receptor
<i>TEX264</i>	Testis expressed 264	Unknown
<i>C14orf2</i>	Chromosome 14 open reading frame 2	Unknown
<i>SOCS3</i>	Suppressor of cytokine signalling 3	Immunological, apoptosis
<i>KIAA2013</i>	Hypothetical protein KIAA2013	Unknown, in silico predicted motif/domain
<i>ASNA1</i>	ArsA arsenite transporter, ATP-binding, homolog 1 (bacterial)	Solute transporter, anion transporting ATPase
<i>ATP5G1</i>	ATP synthase, H ⁺ transporting, mitochondrial F0 complex, subunit c (subunit 9)	Solute transporter, energy
<i>NOLA3</i>	Nucleolar protein family A, member 3	Nuclear, RNA processing

The genes encoding these biomarkers are therefore novel candidate genes for susceptibility to TB. In addition to these 9 biomarkers, we considered the list of 337 differentially expressed genes and sorted them according to their ability to distinguish between the latent and recurrent groups. These two groups were selected since we considered them to represent the extremes of the four phenotypes of TB i.e. latently infected individuals were considered to be more resistant, while patients with recurrent episodes were considered to be more susceptible to TB. We excluded all genes with $p \geq 0.05$. This resulted in a second list of 24 genes (Table 102) which were investigated using the published literature to determine their possible function in TB.

TABLE 102. Genes differentially expressed between the latent and recurrent groups. From [271].

Differentially expressed	Accession number	Differentially expressed	Accession number
<i>FLJ22875</i>	NM_032231	<i>TCF19</i>	NM_007109
<i>FANCE</i>	NM_021922	<i>OSBPL5</i>	NM_020896
<i>PTMA</i>	NM_002823	<i>SCARB2</i>	NM_005506
<i>TEX264</i>	NM_015926	<i>TBXA2R</i>	NM_001060
<i>GCN5L1</i>	NM_001487	<i>OSBPL10</i>	NM_017784
<i>FLJ21709</i>	NM_032206	<i>IGSF6</i>	NM_005849
<i>B2M</i>	NM_004048	<i>ITGAE</i>	NM_002208
<i>KIAA0750</i>	NM_014632	<i>AV696234</i>	AV696234
<i>CCDC55</i>	NM_032141	<i>RGC32</i>	NM_014059
<i>SFRS4</i>	NM_005626	<i>CD6</i>	NM_006725
<i>CCL3</i>	NM_002983	<i>CHODL</i>	NM_024944
<i>TIMP1</i>	NM_003254	<i>LOC51231</i>	NM_016440

From the genes listed in Tables 101 and 102, we selected two novel candidate genes, namely testis expressed 264 (*TEX264*) and suppressor of cytokine signalling 3 (*SOCS3*), to investigate their role in TB susceptibility in the South African Coloured population, based on their ability to distinguish between the phenotypes of TB.

6.1.1 Testis expressed 264

Testis expressed 264 (*TEX264*), located on chromosome 3p21.31, was selected as a candidate gene because of differential gene expression patterns in blood from patients with active, cured, latent or recurrent TB [271] and is a candidate gene by experiment. In the DNA-array expression analysis (Figure 49), this gene could distinguish between the latent and recurrent groups. This observation was confirmed with quantitative reverse-transcriptase PCR (qRT-PCR) [271].

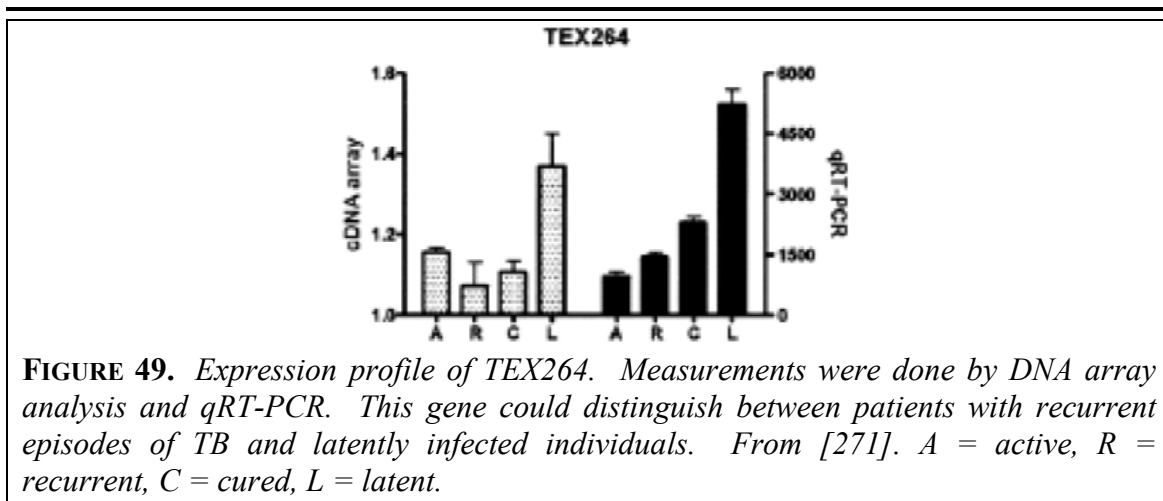


FIGURE 49. Expression profile of *TEX264*. Measurements were done by DNA array analysis and qRT-PCR. This gene could distinguish between patients with recurrent episodes of TB and latently infected individuals. From [271]. A = active, R = recurrent, C = cured, L = latent.

Thus far, *TEX264* is a gene of unknown function, but the TEX264 protein is most abundantly expressed in the larynx and lymph (EST Profile Viewer, Unigene, NCBI), indicating some possible role in the immune system. TEX264 is found mainly extracellularly [589,590] and was first identified as a novel secreted protein during a bioinformatics assessment [591]. We sequenced the coding regions of the gene to identify known and novel single nucleotide polymorphisms (SNPs) in the gene and genotyped selected polymorphisms in the samples from the South African Coloured population to investigate their possible association with TB susceptibility.

6.1.2. Suppressor of cytokine signalling 3

Members of the suppressor of cytokine signalling (SOCS) protein family act as negative feedback inhibitors [592] of cytokines which signal through the Janus kinase/signal transducer and activators of transcription (JAK/STAT) pathway. Eight SOCS family members have been identified to date [593] and transcription of their respective encoding genes is induced by cytokines, hormones and growth factors [594]. SOCS-1, SOCS-3 and the cytokine-inducible SH2-containing protein (CIS) are the best understood members of the protein family [594] and can all be induced in cells from the innate immune system, such as macrophages and dendritic cells [595]. SOCS-1 and SOCS-3 are part of a related pair, with their amino acid sequences being more similar to each other than to other proteins of the SOCS family [596]. These two related regulators were both induced in murine macrophages infected with *M.bovis* bacillus Calmette-Guérin (BCG) and it was suggested that these molecules are part of a mechanism responsible for a poor response to interferon gamma (IFN- γ) during mycobacterial infection [597]. SOCS-1 and SOCS-3 were also directly induced by lipopolysaccharides (LPS) [598] and CpG-DNA [599]. This induction was independent of protein synthesis and the JAK/STAT pathway, proving that SOCS expression can be activated by toll-like receptor (TLR) signalling pathways.

SOCS-3 is an intracellular protein, the expression of which is rapidly upregulated by IFN- γ [600]. A STAT-binding site, present in the proximal region of the *SOCS3* promoter, is needed for this response to IFN- γ [600]. In mice deficient for *SOCS3*, immune responses to infection were impaired. This was a result of the production of an excess of transforming growth factor β and interleukin-10 (IL-10) [601]. Three intracellular pathogens, namely *Leishmania donovani* [602], *Listeria monocytogenes* [603] and *Salmonella enterica* [604], can induce the expression of SOCS-3 in macrophages and thereby inhibit cytokine signalling. This response could be part of the normal immune system, where continued macrophage activation may be detrimental to the host [603], but pathogens could also exploit this inhibitory mechanism to suppress macrophage activation [602]. Since SOCS-3 is highly expressed [271] in patients with active TB (Figure 50), it is possible that *M.tuberculosis* could be upregulating transcription of *SOCS3* to avoid destruction within macrophages. The high expression levels of SOCS-3 could also be an indicator of an ongoing immune response [271], since patients with recurrent episodes of disease expressed extremely low levels (Figure 50) of this protein, indicating an impaired immune response [271]. However, low levels of SOCS-3 [271] were also observed in cured and latently infected individuals (Figure 50), supporting the hypothesis that *M.tuberculosis* infection actively stimulates the expression of SOCS-3. Since this has been observed in macrophages infected with

M.bovis BCG [597], the involvement of SOCS-3 in the *M.tuberculosis*-induced inhibition of macrophage activity is quite possible.

SOCS3 was selected as a candidate gene for TB based on results from the gene expression study of Mistry et al [271] and is a candidate gene by experiment. We selected three *SOCS3* polymorphisms (Table 103) from dbSNP [58] and the literature [605,606] to genotype in the samples from the South African Coloured population.

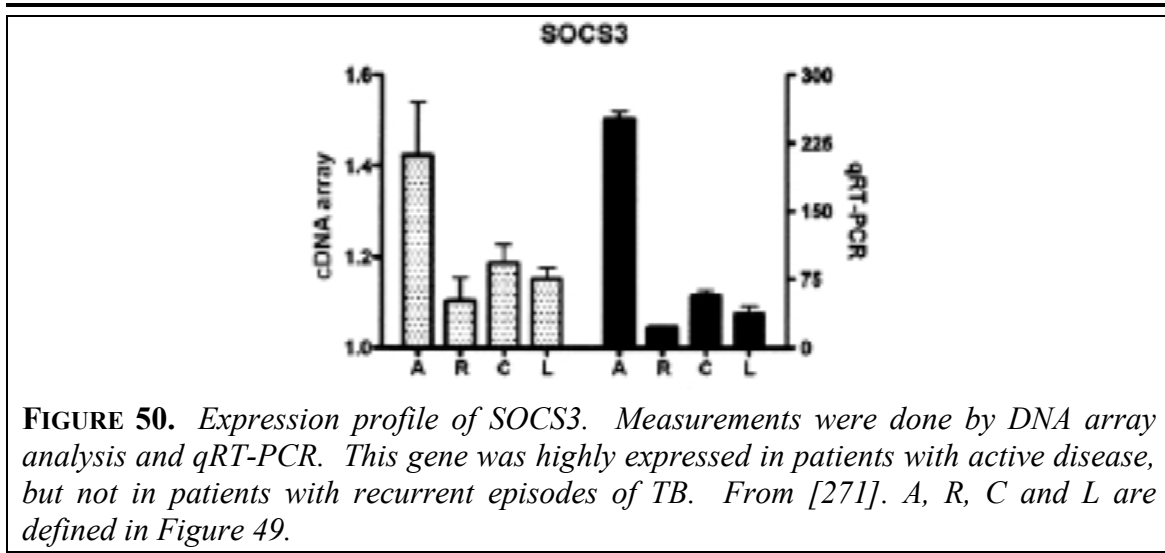


FIGURE 50. Expression profile of *SOCS3*. Measurements were done by DNA array analysis and qRT-PCR. This gene was highly expressed in patients with active disease, but not in patients with recurrent episodes of TB. From [271]. A, R, C and L are defined in Figure 49.

TABLE 103. Polymorphisms genotyped in *SOCS3*.

Polymorphism	Class	Location	Genotyping method
rs4969168	SNP	3'UTR ^a	SNPlex
rs4969169	SNP	3'UTR	SNPlex
rs7207782	SNP	Promoter	SNPlex

^a untranslated region (UTR)

6.2. RESULTS

6.2.1. *TEX264*

The coding regions of the *TEX264* were amplified (Section 3.5.5.) and the products were visualised on 2% agarose gels (Figure 51). Sequencing of the coding regions of *TEX264* identified two novel polymorphisms (Figure 52) in the South African Coloured population. These novel polymorphisms were submitted to the dbSNP database [58] and received unique rs-numbers. Base 28551 A to G (Figure 52A, rs28730659, in intron 3) is a non-coding polymorphism and has an allele frequency of 0.1. The genotyping data from this SNP did not pass quality checking as described in section 3.11.1 and was therefore excluded from further analysis. Base 33205 (Figure 52B, rs28994878, in exon 5) is a synonymous polymorphism and has an allele frequency of 0.07. This novel SNP, together with 2 polymorphisms in the 5' region of the gene (Table 104), was genotyped in all samples.

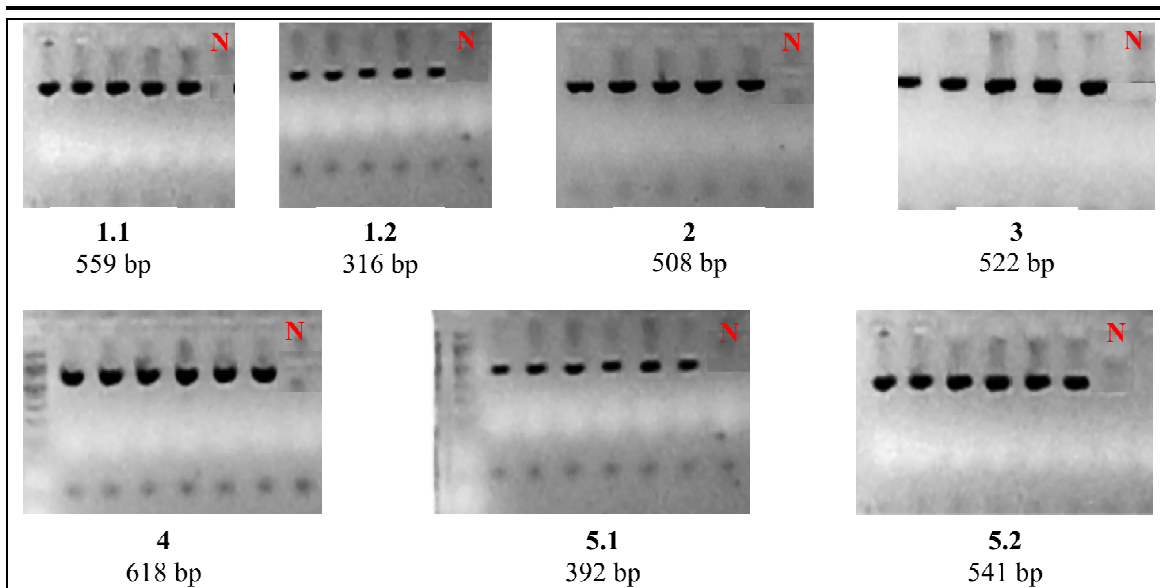


FIGURE 51. Amplification of *TEX264* products for sequencing. The five coding exons of this gene were amplified in seven separate PCR reactions. The gel photographs are numbered according to section 3.5.5 (1.1 – 5.2) and the size of each fragment is indicated. N = negative control.

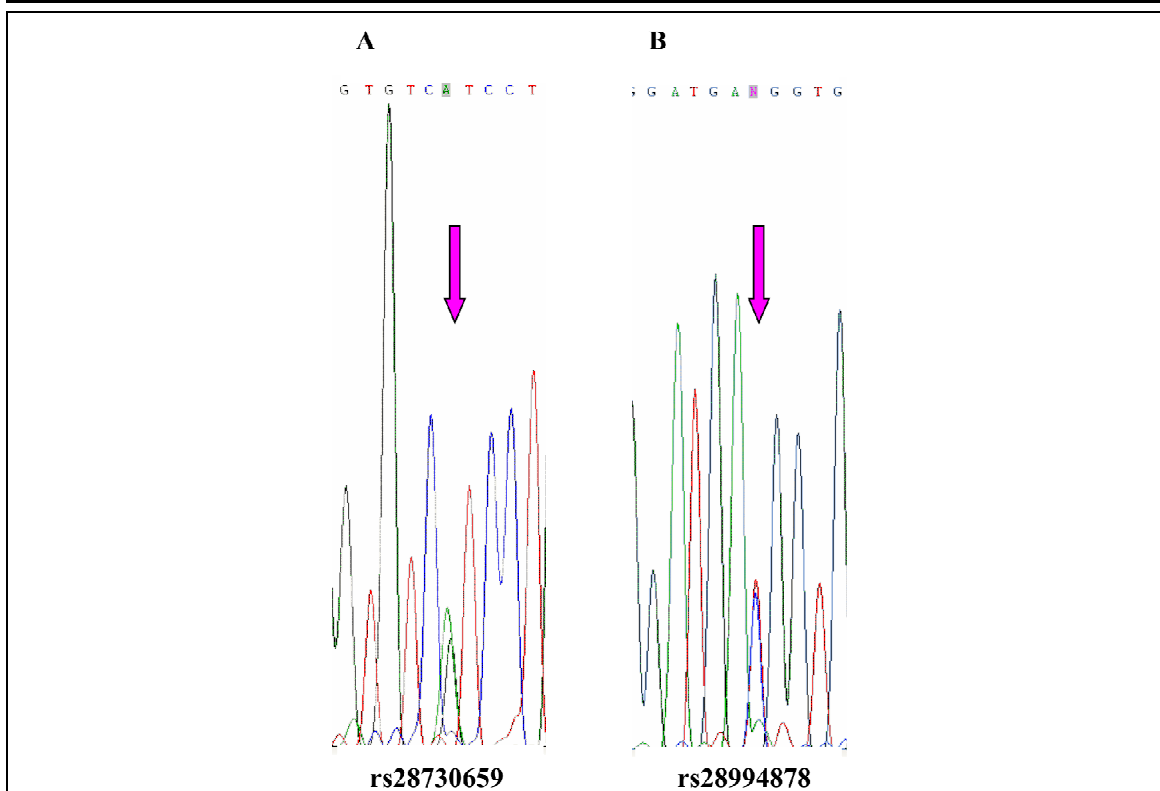


FIGURE 52. Sequencing of the coding regions of *TEX264* and the detection of novel polymorphisms, highlighted in grey. A) The rs28730659 A/G SNP is present in intron 3. B) The synonymous rs28994878 C/T SNP is in exon 5.

TABLE 104. Polymorphisms genotyped in the *TEX264* gene.

Polymorphism	Class	Location	Genotyping method
rs6775666	SNP	Promoter	SNPlex
rs4355273	SNP	5'UTR	SNPlex
rs28994878	SNP	Exon 5, synonymous	SNPlex

6.2.1.1. Single-point statistical analysis

The three *TEX264* polymorphisms were successfully genotyped in samples from the case-control and family-based studies. All SNPs were in Hardy-Weinberg equilibrium (HWE) for the control group. The threshold for significance was set at $p = 0.025$ after Bonferroni correction, since rs6775666 and rs4355273 were part of a single haplotype block and therefore not independent of each other.

The novel synonymous rs28994878 SNP was not associated with TB susceptibility in the South African Coloured population in our case-control study ($p = 0.90$, Table 105) or transmission disequilibrium test (TDT, $p = 0.40$, Table 106). The two polymorphisms rs6775666 and rs4355273 in the 5'UTR region of the gene, which could potentially modify expression of *TEX264*, were also not associated with disease susceptibility (Tables 105 and 106).

TABLE 105. Single-point statistical analysis in the *TEX264* case-control study.

Polymorphism	Allele		Genotype frequencies						Cases vs Controls p value ^c
			Cases (n ^a =432)			Controls (n=482)			
	1 ^b	2	11	12	22	11	12	22	
rs6775666	A	G	0.002	0.088	0.91	0.002	0.108	0.89	0.62
rs4355273	C	G	0.01	0.08	0.91	0.002	0.114	0.884	0.15
rs28994878	T	C	0	0.07	0.93	0	0.06	0.94	0.90

^a number of investigated individuals.

^b Allele 1 is the minor allele.

^c p value from a genotype-based χ^2 test

TABLE 106. Single-point statistical analysis in the *TEX264* TDT.

SNP	Overtransmitted	Transmitted:Untransmitted	χ^2	p value
rs6775666	G	13 : 11	0.17	0.68
rs4355273	G	13 : 8	1.19	0.28
rs28994878	T	4 : 2	0.67	0.41

6.2.1.2. Haplotype analysis

The Haploview program [313] identified a single haplotype block (Figure 53) consisting of the two polymorphisms in the 5' region of *TEX264*. The rs28994878 SNP was

excluded from this haplotype since the block definition from Gabriel et al [89] labelled its linkage with the other variants as uncertain.

Two haplotypes were determined for the case-control (Table 107) and TDT (Table 108) samples in the haplotype block. Haplotype 1 was the most frequent haplotype in both groups. None of the haplotypes were associated with TB susceptibility in either the case-control or TDT haplotype analyses (Tables 107 and 108).

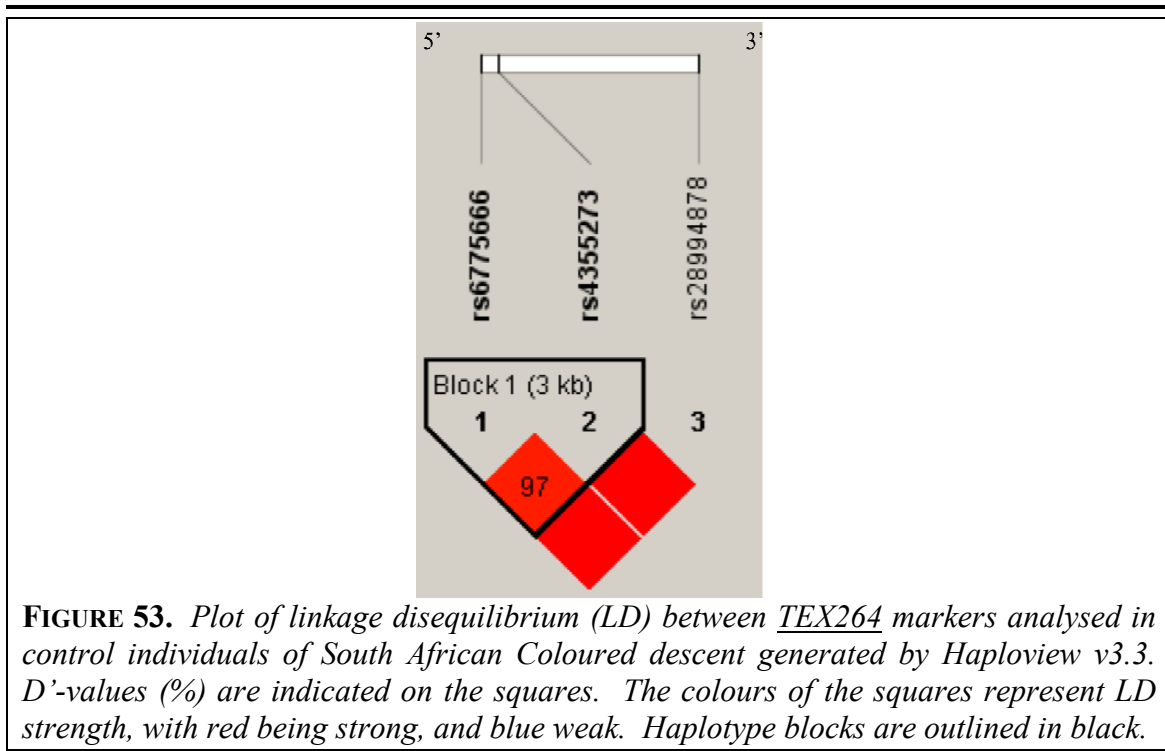


FIGURE 53. Plot of linkage disequilibrium (LD) between *TEX264* markers analysed in control individuals of South African Coloured descent generated by Haploview v3.3. *D'*-values (%) are indicated on the squares. The colours of the squares represent LD strength, with red being strong, and blue weak. Haplotype blocks are outlined in black.

TABLE 107. *TEX264* case-control haplotype analysis by Haploview [313].

Block 1 ^a		Freq _{Cases}	Freq _{Controls}	χ^2	p value	p value _{Permutation} ^b
1	G-G	0.95	0.94	1.10	0.29	0.64
2	A-C	0.05	0.06	0.87	0.35	0.67

^a The order of the SNPs in each block corresponds to Figure 53.

^b Permutation test p values were calculated from 10 000 permutations in Haploview.

^c p values less than 0.29 were observed on 6420 out of 10 000 occasions (64%).

TABLE 108. *TEX264* TDT haplotype analysis by Haploview [313].

Block 1 ^a		Transmitted:untransmitted	χ^2	p value	p value _{Permutation} ^b
1	G-G	11 : 9	0.20	0.65	1
2	A-C	9 : 10	0.05	0.82	1

^a The order of the SNPs in each block corresponds to Figure 53.

^b Permutation test p values were calculated from 10 000 permutations in Haploview.

^c p values less than 0.65 were observed on 10 000 out of 10 000 occasions (100%).

6.2.2. SOCS3

6.2.2.1. Single-point statistical analysis

The three SNPs in *SOCS3* were genotyped in the case-control and TDT samples sets. All SNPs were in HWE in the control population. After Bonferroni correction using the methodology of Nicodemus et al [76], the level of significance required was set at $p = 0.017$. According to this criterion none of the polymorphisms analysed were associated with TB susceptibility in either the case-control study (Table 109) or TDT (Table 110).

TABLE 109. Single-point statistical analysis in the *SOCS3* case-control study.

Polymorphism	Allele		Genotype frequencies						Cases vs Controls p value ^c
			Cases (n ^a =432)			Controls (n=482)			
	1 ^b	2	11	12	22	11	12	22	
rs4969168	G	A	0.17	0.49	0.34	0.22	0.50	0.28	0.07
rs4969169	T	C	0	0.11	0.89	0	0.13	0.87	0.66
rs7207782	A	G	0.04	0.34	0.62	0.06	0.35	0.59	0.34

^a number of investigated individuals.

^b Allele 1 is the minor allele.

^c p value from a genotype-based χ^2 test

TABLE 110. Single-point statistical analysis in the *SOCS3* TDT.

SNP	Overtransmitted	Transmitted:Untransmitted	χ^2	p value
rs4969168	A	58 : 41	2.92	0.09
rs4969169	C	15 : 7	2.91	0.09
rs7207782	G	30 : 27	0.16	0.69

6.2.2.2. Haplotype analysis

Haploview [313] did not calculate any haplotype blocks (Figure 54) for the three SNPs genotyped in *SOCS3* when considering the haplotype block definition of Gabriel et al [89]. However, when a definition of a single spine of LD was used, the three variants were part of a haplotype block. Six haplotypes were calculated in the haplotype block in both the case-control (Table 111) and TDT (Table 112) haplotype analyses.

Haplotype 2 was the most frequently observed haplotype in the case-control and family-based samples. This haplotype showed a nominally significant association with TB susceptibility in the case-control analysis ($p = 0.03$, Table 111).

To correct for multiple testing, we did 10 000 permutations using Haploview [313]. According to this program, 835 out of the 10 000 permutations exceeded the best observed p value of haplotype 2. This means that there is an 8% chance that the significant nominal p value of haplotype 2 was observed due to chance alone. This haplotype was not associated with disease in the TDT analysis (Table 112). However, haplotype 6 was nominally associated with TB susceptibility ($p = 0.05$, Table 112).

There was a 16% chance that this association was due to chance alone and not surprisingly, it was not detected in the case-control analysis.

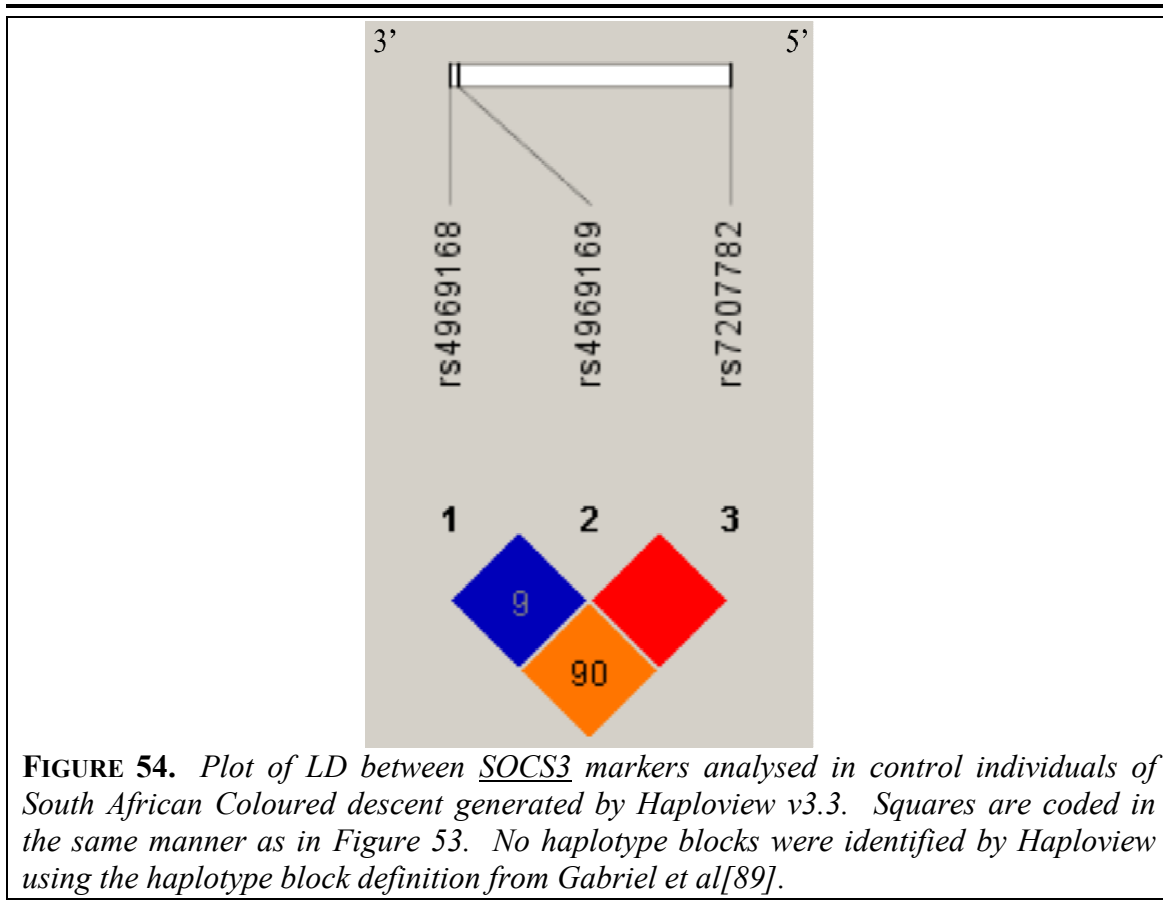


TABLE 111. *SOCS3* case-control haplotype analysis by Haploview [313].

Block 1 ^a		Freq _{Cases}	Freq _{Controls}	χ^2	p value	p value _{Permutation} ^b
1	A-C-A	0.01	0.01	0.01	0.93	1
2	A-C-G	0.54	0.48	4.88	0.03	0.08
3	A-T-G	0.03	0.04	0.16	0.69	1
4	G-C-A	0.20	0.22	1.78	0.18	0.56
5	G-C-G	0.20	0.22	0.83	0.36	0.86
6	G-T-G	0.03	0.03	0.59	0.44	0.93

^a The order of the SNPs in each block corresponds to Figure 54.

^b Permutation test p values were calculated from 10 000 permutations in Haploview.

^c p values less than 0.03 were observed on 835 out of 10 000 occasions (8%).

TABLE 112. *SOCS3* TDT haplotype analysis by Haploview [313].

Block 1 ^a	Transmitted:untransmitted	χ^2	p value	p value _{Permutation} ^b	
1	A-C-A	2 : 1	0.31	0.58	0.99
2	A-C-G	44 : 34	1.33	0.25	0.71
3	A-T-G	4 : 3	0.07	0.79	1
4	G-C-A	24 : 20	0.32	0.57	0.98
5	G-C-G	25 : 35	1.48	0.22	0.62
6	G-T-G	2 : 8	3.82	0.05	0.16

^a The order of the SNPs in each block corresponds to Figure 54.

^b Permutation test p values were calculated from 10 000 permutations in Haploview.

^c p values less than 0.05 were observed on 1630 out of 10 000 occasions (16%).

6.3. DISCUSSION

Gene expression studies can suggest novel candidate genes to investigate in diseases and such a study has been done in TB using individuals from the South African Coloured population [271]. The *TEX264* and *SOCS3* genes were investigated as factors contributing to TB susceptibility in the South African Coloured population. Expression of these genes, together with that of 7 other genes, made it possible to discriminate between the four phenotypes of TB, as defined before, namely active, recurrent, latent and cured [271]. *TEX264* could discriminate between individuals in the latent group, where it was expressed at the highest level, and those with recurrent disease, while *SOCS3* was expressed at a higher level in patients with active TB than in the other groups. We hypothesised that polymorphisms in these genes could play a role in TB susceptibility.

The SNPs investigated in *TEX264* included a novel polymorphism identified in the South African Coloured population (rs28994878). However, none of the SNPs or haplotypes were associated with TB. This was also the case for the polymorphisms and haplotypes investigated in *SOCS3*. The *SOCS3* gene was previously associated with Type 1 diabetes [607] and atopic dermatitis [605].

Although we found no association between polymorphisms in these novel candidate genes and TB susceptibility, it is still possible that they contribute to the disease by a mechanism not determined by genetic variation. This includes regulation of gene expression by *M.tuberculosis*, which has previously been shown to modulate the immune response of the host [608]. Different strains of *M.tuberculosis* can elicit different immune responses in different individuals [609-611]. Other candidates identified in the gene expression study from Mistry et al [271] could be investigated to determine if they play a role in genetic susceptibility to TB.

CHAPTER 7

A CANDIDATE GENE FROM A GENOME- WIDE LINKAGE SCREEN

7.1. INTRODUCTION

Linkage studies are used to outline chromosomal regions containing putative susceptibility genes, either by a genome-wide scan, which ensures that all major genomic regions involved in disease susceptibility are identified, or by concentrating on a candidate region. This methodology creates the possibility of identifying novel genes and pathways which were not formerly considered to be disease-contributing factors. The approach has been successfully used in the analysis of monogenic diseases, but varying results have been found in research considering complex diseases, such as tuberculosis (TB), where more than one gene can influence disease susceptibility, and individual genes are generally of small effect. An example of a linkage study which focused on a candidate region, namely chromosome 17q11-q21 region, is that of Jamieson et al [65]. A region on mouse chromosome 11 was previously shown to contain susceptibility genes to another intracellular pathogen [64]. Evaluation of the syntenic chromosome 17q11-q21 region in humans suggested that four genes contributed separately to TB susceptibility [65].

The use of genome-wide scans in TB genetics has suggested several susceptibility loci (*Section 1.2.2.1*). These include regions on chromosome 15q and Xq, implicated in TB susceptibility of the South African population [66], as well as chromosome 2q35 in aboriginal Canadians [68]. A study in Brazilians, which examined TB and leprosy families, suggested three regions (10q26.13, 11q12.3 and 20p12.1) [69]. A recent genome scan in Moroccan families suggested that chromosome 8q12-q13 may contain a major dominant-acting TB susceptibility gene [70]. Little overlap between susceptibility regions has been found when comparing the different TB linkage studies. This is understandable, because the strongest linkage observed may be exclusive to a specific family or population and thus impossible to identify in other studies [71]. The phenotyping used to identify affected individuals is also crucial. A new genome-wide scan, as yet unpublished (Cooke et al), which also investigated the South African Coloured population, has led to the identification of a novel susceptibility gene in the chromosome 20q13.31-33 region. In that study, 40 single nucleotide polymorphisms (SNPs) in the chromosomal region identified were screened. In the initial analysis, polymorphisms from two genes showed evidence of disease association in Gambian TB samples. Additional SNPs from these genes were then genotyped in populations from Guinea-Bissau and The Republic of Conakry. The most convincing candidate gene in this region was cathepsin Z (*CTSZ*), while the neighbouring TH1-like (*Drosophila*) (*TH1*) and tubulin, beta 1 (*TUBB1*) genes were also potential candidates. A significant association ($p = 0.005$) was seen with a 3' untranslated region (UTR) SNP (rs10369) of *CTSZ*, but not with the only non-synonymous polymorphism of this gene (rs9760). *TUBB1* also showed a weakly significant association ($p = 0.04$), but the association was not consistent across the study populations.

Genome-wide linkage studies in TB genetics have not conclusively established major loci for the disease. However, Cooke et al suggested that this study approach may serve as a strategy to investigate genomic regions more likely to contain susceptibility genes, an advantage to researchers in countries where SNP genotyping is still expensive. Therefore, we selected a candidate gene in the chromosomal region identified by Cooke et al to investigate its possible role in TB susceptibility in the South African Coloured population.

7.2. CATHEPSIN Z

Cathepsin Z (CTSZ [612], also known as cathepsin P [613], X [614] or Y [615]) is one of the 11 cysteine proteases of the papain family [616]. Proteins of this family are mostly localised in lysosomes, where they form a vital component of the lysosomal proteolytic system. They may also be found in the cytosol or extracellularly, usually as precursors of their active forms [613]. They are essential in biological processes such as bone resorption, apoptosis [615] and extracellular matrix remodelling (ECM). In the immune system, cathepsins are involved in antigen processing and maturation of the major histocompatibility complex (MHC) class II molecules [616]. CTSZ is mostly expressed in immune cells, such as macrophages and monocytes [617,618], and a role for the protein in the immune response has been hypothesised [619]. However, no specific immune function has been identified for this protein as far as could be ascertained. The protein has a considerable similarity to cathepsins B, C and L [620]. The enzyme has both carboxymonopeptidase and carboxydipeptidase activities, but shows little endopeptidase activity [615]. Recently it has been suggested that CTSZ may be used as a diagnostic marker in inflammatory diseases [621].

Several members of the cathepsins have been implicated in TB. It has been found that cathepsin L maturation and activity can be impaired by *Mycobacterium tuberculosis* (*M.tuberculosis*) and *M.avium* [622]. This diminished the ability of macrophages to present antigens. However, infection of macrophages with the mycobacteria had no effect on the maturation of cathepsins B or D. Cathepsins are differentially expressed during *M.tuberculosis* infection and this expression is specifically associated with macrophages present in the granuloma [623]. Another cysteine cathepsin, namely cathepsin W, was identified as a risk factor for the extrapulmonary dissemination of human TB [624]. The involvement of cathepsins in a caspase-independent pathway which modulates apoptosis by macrophages infected with *M.tuberculosis* was recently suggested [625] and confirmed in a second study [626]. As yet, there is no direct evidence of a function for CTSZ in TB. However, since this enzyme is expressed mostly in macrophages, and because of its homology to other cathepsins involved in TB, the gene encoding this protein is a candidate gene for TB susceptibility. Furthermore, *CTSZ* is a candidate gene due to its location in the chromosome 20q13.31-33 region (*Section 7.1*) and its possible role in the immune response. We selected 7 polymorphisms (Table 113) from this gene to genotype in the case-control and family-based samples from the South African Coloured population.

TABLE 113. *Polymorphisms genotyped in CTSZ in this study.*

Polymorphism	Class	Location	Genotyping method
rs448943	SNP	3'UTR ^a	SNPlex
rs10369	SNP	3'UTR	SNPlex
rs13720	SNP	3'UTR	SNPlex
rs6064734	SNP	Exon 6, synonymous	SNPlex
rs3787492	SNP	Exon 6/intron 5 boundary	SNPlex
rs163785	SNP	Exon 4, synonymous	SNPlex
rs11540881	SNP	Exon 4, synonymous	SNPlex

^a *untranslated region*

7.2.1. Results

7.2.1.1. Single-point statistical analysis

Seven SNPs were successfully genotyped in all the TB cases, controls and family-based samples from the South African Coloured population (Tables 114 and 115). Of the seven polymorphisms genotyped, three (rs6064734, rs163785 and rs11540881) were found to be monomorphic in all samples and were excluded from further analyses. The remaining four SNPs were analysed in the case-control and family-based samples and were all in Hardy-Weinberg equilibrium (HWE) for the control population.

A p value of 0.025 would be necessary to achieve statistical significance, as rs3787492 was in linkage equilibrium with the three other SNPs and considered to be an independent test. The three 3'UTR polymorphisms rs448943, rs10369 and rs13720 were in linkage disequilibrium (LD).

TABLE 114. Single-point statistical analysis in the *CTSZ* case-control study design.

Polymorphism	Allele		Genotype frequencies						Cases vs Controls p value ^c
	1 ^b	2	Cases (n ^a =432)			Controls (n=482)			
			11	12	22	11	12	22	
rs448943	G	C	0.26	0.46	0.28	0.22	0.49	0.29	0.36
rs10369	T	C	0.06	0.35	0.60	0.04	0.34	0.62	0.54
rs13720	G	A	0.04	0.27	0.69	0.04	0.30	0.66	0.58
rs6064734	Monomorphic								
rs3787492	T	C	0.05	0.38	0.57	0.06	0.41	0.53	0.33
rs163785	Monomorphic								
rs11540881	Monomorphic								

^a number of investigated individuals.

^b Allele 1 is the minor allele.

^c p value from a genotype-based χ^2 test

The 3'UTR polymorphism rs10369 was not significantly associated with TB in our case-control study ($p = 0.54$, Table 114) or TDT ($p = 0.08$, Table 115). This SNP was previously associated (Cooke et al, unpublished data) with TB in samples from The Gambia, Guinea-Bissau and The Republic of Conakry. None of the other 3'UTR SNPs were associated with disease (Tables 114 and 115).

We also investigated the rs3787492 SNP located in an exon/intron boundary of *CTSZ*, because it could be involved in alternative splicing of the gene. Two transcript variants of this gene have been found previously (*CTSZ* Entrez Gene Summary, NCBI). This SNP was not associated with TB in our study. Altogether, single marker analysis detected no association between TB and any of the SNPs investigated in the *CTSZ* gene.

TABLE 115. Single-point statistical analysis in the *CTSZ* TDT.

SNP	Overtransmitted	Transmitted:Untransmitted	χ^2	p value
rs448943	G	50 :47	0.09	0.76
rs10369	T	39 : 25	3.06	0.08
rs13720	-	28 : 28	0	1
rs6064734		Monomorphic		
rs3787492	C	32 : 30	0.07	0.80
rs163785		Monomorphic		
rs11540881		Monomorphic		

7.2.1.2. Haplotype analysis

The Haploview program [313] identified a single haplotype block (Figure 53) for three of the markers genotyped in *CTSZ* in the control samples. The LD between the rs448943, rs10369 and rs13720 3'UTR markers in this block was strong. The exon/intron boundary SNP (rs3787492) was not part of this block. Since no additional information on the haplotype block containing this SNP was available, it was not considered in the statistical analysis of *CTSZ* haplotypes with TB in this study.

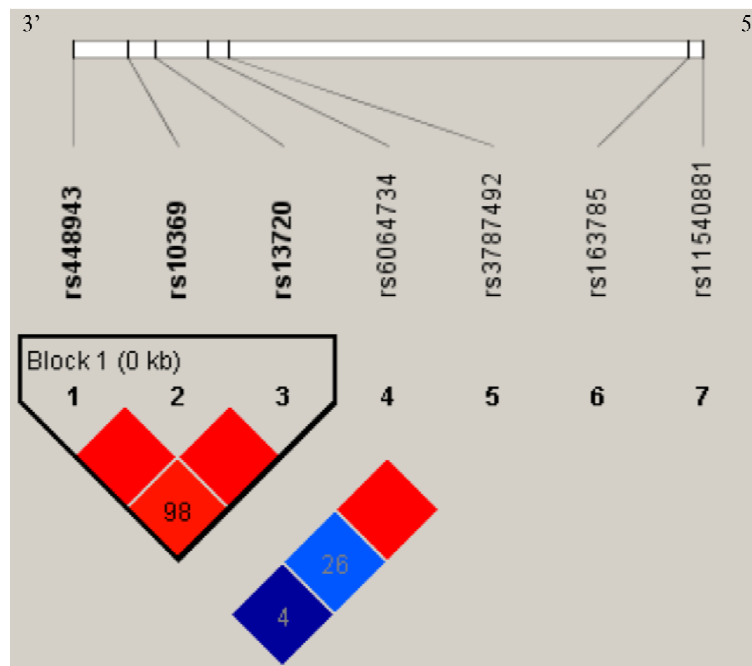


FIGURE 55. Plot of LD between *CTSZ* markers analysed in control individuals of South African Coloured descent generated by Haploview v3.3. D' -values (%) are indicated on the squares. The colours of the squares represent LD strength, with red being strong, and blue weak. Haplotype blocks are outlined in black. The rs6064734, rs163785 and rs11540881 SNPs were monomorphic in this study.

Four haplotypes were determined in both the case-control (Table 116) and TDT (Table 117) haplotype analyses. Haplotype 1 was the most frequently observed haplotype in

both analyses. Haplotypes 2 ($p = 0.01$) and 3 ($p = 0.02$) were transmitted to TB offspring more frequently than expected in the TDT analysis (Table 117). The most significant association was observed on 645 out of 10 000 occasions (6%) during permutation testing. Since the level of significance was set at 5%, these results are not reflections of true associations. None of the remaining haplotypes in the case-control or TDT analyses were associated with TB.

TABLE 116. *CTSZ* case-control haplotype analysis by Haploview [313].

Block 1 ^a		Freq _{Cases}	Freq _{Controls}	χ^2	p value	p value _{Permutation} ^b
1	C-C-A	0.33	0.34	0.18	0.67	0.97
2	G-C-A	0.26	0.26	0.06	0.81	1
3	G-T-A	0.23	0.21	0.97	0.33 ^c	0.72
4	C-C-G	0.18	0.19	0.65	0.42	0.83

^a The order of the SNPs in each block corresponds to Figure 53.

^b Permutation test p values were calculated from 10 000 permutations in Haploview.

^c p values less than 0.33 were observed on 7243 out of 10 000 occasions (72%).

TABLE 117. *CTSZ* TDT haplotype analysis by Haploview [313].

Block 1 ^a		Transmitted:untransmitted	χ^2	p value	p value _{Permutation} ^b
1	C-C-A	39 : 36	0.09	0.76	0.99
2	G-C-A	24 : 44	5.98	0.01^c	0.06
3	G-T-A	29 : 14	5.12	0.02	0.10
4	C-C-G	23 : 20	0.34	0.56	0.94

^a The order of the SNPs in each block corresponds to Figure 53.

^b Permutation test p values were calculated from 10 000 permutations in Haploview.

^c p values less than 0.01 were observed on 645 out of 10 000 occasions (6%).

7.2.2. Discussion

In this chapter, polymorphisms in *CTSZ* were tested as susceptibility factors in TB. This gene was previously associated with the disease in a two-stage genome-wide linkage study (Cooke et al, unpublished data). We attempted to replicate the association observed in their study in the case-control and family-based samples from the South African Coloured population.

We did not find any association between the 3'UTR SNP rs10369 and TB susceptibility, although the C allele of this SNP was associated with the disease in the unpublished study from Cooke et al. We also found no association between the other 3'UTR or synonymous SNPs genotyped in our study. These supposedly synonymous SNPs (rs6064734, rs163785 and rs11540881) were reported to be polymorphic in the dbSNP database and were therefore included in genotyping. However, they were monomorphic in all the samples from the South African Coloured population genotyped in this study.

The genome-wide screen from Cooke et al was done in two steps, which strengthens the case for an involvement of *CTSZ* in TB. Even so, it is possible that other genes in the

chromosome 20q13.31-33 region may be in LD with *CTSZ*. Another explanation is that *CTSZ* is not a major contributor to TB and that our study did not have enough power to detect such a small effect size. In fact, the study from Cooke et al did suggest a minor genetic effect (OR 0.49) of *CTSZ* in TB.

In conclusion, polymorphisms and haplotypes in *CTSZ* are probably not associated with TB in the South African Coloured population. Future studies should consider other TB candidate genes in the chromosome 20q13.31-33 region.

CHAPTER 8

FUNCTIONAL POLYMORPHISMS THAT INFLUENCE SUSCEPTIBILITY TO IMMUNE-MEDIATED DISEASES OTHER THAN TUBERCULOSIS

8.1. INTRODUCTION

Immune-mediated diseases may be affected by several similar genes, which can disrupt immune homeostasis and contribute to disease pathology [627]. Inflammation, a common manifestation of the innate immune system after injury or infection and also observed during tuberculosis (TB) infection [628], can disrupt immune homeostasis.

Acute inflammation, normally seen after injuries, involves the infiltration of tissues by leukocytes and plasma, while chronic inflammation (Figure 56), such as observed during *Mycobacterium tuberculosis* (*M.tuberculosis*) infection, is characterised by mononuclear immune cell infiltration of tissues. It can serve as a protective mechanism to remove or destroy the causative agent, limit the spread of infection and prepare the damaged region for healing [629].

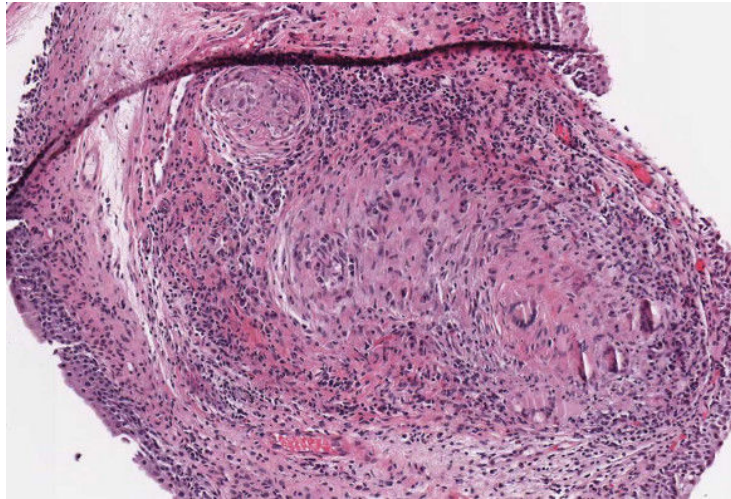


FIGURE 56. Bladder biopsy showing chronic granulomatous inflammation. From [630].

However, inflammation is a two-edged sword and could, when misdirected by the immune system, also attack the host organism [629]. Diseases where this is the case include psoriasis, rheumatoid arthritis and asthma, and several of these diseases share common susceptibility genes. In some cases, functional polymorphisms have been identified in these genes and were associated with various immune-mediated diseases as well as TB. However, opposite associations between these functional variants, previously associated with immune-mediated diseases, and TB have frequently been observed.

An example is the A allele of the -308 single nucleotide polymorphism (SNP) in the tumour necrosis factor (TNF superfamily, member 2) (*TNF*) gene promoter which leads to a higher rate of *TNF* transcription in expression studies [631,632]. This allele was associated with immune-mediated diseases such as systemic lupus erythematosus [633], rheumatoid arthritis [634] and inflammatory bowel disease (IBD, which includes the subphenotypes Crohn's disease and ulcerative colitis) [635]. High levels of TNF- α in these diseases are detrimental to the host and treatment may involve the use the TNF- α

inhibitors such as infliximab. In patients latently infected with *M.tuberculosis*, infliximab treatment led to reactivation of the disease [636], suggesting that TNF- α is necessary to control TB infection. The alternative low TNF- α producing G allele of the -308 *TNF* SNP was subsequently associated with TB susceptibility [171,172], lending support to this suggestion.

Correa et al [172] proposed that heterozygotes of the *TNF* promoter polymorphism have a survival advantage, since individuals with homozygous GG genotypes would be resistant to TB, but susceptible to autoimmune diseases. This supported the hypothesis that autoimmune diseases are a result of natural selection for an increased resistance to TB. In addition, their results suggested that some susceptibility genes are not specific to a certain disease, but that related immunogenetic systems contribute to common diseases.

To test this hypothesis, we selected from the literature several functional polymorphisms in candidate genes, previously associated with other immune-mediated disorders, to genotype in the samples from the South African Coloured population. This was an exploratory study to determine the possible association of these SNPs with TB.

8.2. CANDIDATE GENES FROM IMMUNE-MEDIATED DISEASES OTHER THAN TB

We selected 27 SNPs from 22 susceptibility genes, previously associated with various immune-mediated diseases, to genotype with the SNPlex genotyping method (*Section 3.9*) in samples from the South African Coloured population. These genes, the SNPs genotyped in them, the immune-mediated diseases they were previously associated with and their functional effects are listed in Table 118.

TABLE 118. *Functional polymorphisms previously associated with other immune-mediated diseases which were genotyped as described in this chapter.*

SNP	Selected previous disease associations and functional effect	Ref
ATP-binding cassette, subfamily B (MDR/TAP), member 1 (<i>ABCB1</i>)		
rs3789243	Intestinal bowel disease (IBD) Associated with altered protein or gene expression activity.	[637,638]
rs1045642		
rs2032582		
ATG16 autophagy related 16-like 1 (<i>S. cerevisiae</i>) (<i>ATG16L1</i>)		
rs2241880	Crohn's disease Nonsynonymous SNP, involved in autophagocytosis pathway [92,639] implicated in resistance to intracellular pathogens.	
Cytotoxic T-lymphocyte-associated protein 4 (<i>CTLA4</i>)		
rs3087243	Several autoimmune diseases, such as systemic lupus erythematosus and coeliac disease Alters the ratio of splice variants.	[640,641]
Class II, major histocompatibility complex, transactivator (<i>CIITA</i>)		
rs3087456	Systemic lupus erythematosus, rheumatoid arthritis Associated with regulation of class II MHC expression.	[642-644]

SNP	Selected previous disease associations and functional effect	Ref
Fc receptor-like 3 (<i>FCRL3</i>)		
rs7528684	<u>Rheumatoid arthritis, systemic lupus erythematosus</u> C allele was associated with higher transcriptional activity.	[645,646]
Interleukin 1 receptor antagonist (<i>IL1RN</i>)		
rs2419598	<u>Rheumatoid arthritis, ankylosing spondylitis</u> Associated with reduced protein production	[647,648]
Interleukin 23 receptor (<i>IL23R</i>)		
rs11209026	<u>Psoriasis, Crohn's disease</u> Nonsynonymous coding polymorphism affects amino acid structure of the protein.	[97,470]
Interleukin 6 signal transducer (gp130, oncostatin M reporter) (<i>IL6ST</i>)		
rs3730293	<u>Periodontitis</u>	
rs3729961	Nonsynonymous polymorphisms which change amino acid sequence of the protein.	[649]
Insulin induced gene 2 (<i>INSIG2</i>)		
rs7566605	<u>Obesity</u> May result in elevated triglyceride levels.	[650]
Membrane-spanning 4 domains, subfamily F, member 2 (<i>MS4A2</i>)		
rs569108	<u>Asthma</u> Increased expression on basophils.	[651,652]
Nel-like 1 (<i>NELL1</i>)		
rs1607616	<u>IBD</u> Novel disease gene	(Franke et al, accepted)
Peptidyl arginine deiminase, type IV (<i>PADI4</i>)		
rs2240340	<u>Rheumatoid arthritis</u> Affects stability of transcripts as part of a haplotype.	[653]
Peroxisome proliferator-activated receptor gamma (<i>PPARG</i>)		
rs2067819	<u>Crohn's disease</u>	
rs3892175	Regulate expression of an underlying genetic susceptibility in mice	[654]
Prostaglandin E receptor 4 (subtype EP4) (<i>PTGER4</i>)		
rs1992660	<u>Crohn's disease</u> Modulates expression of <i>PTGER4</i> .	[655]
Protein tyrosine phosphatase, non-receptor type 22 (<i>PTPN22</i>)		
rs2476601	<u>Diabetes, rheumatoid arthritis, TB</u> Alters P1 binding site in protein.	[656-660]
Runt-related transcription factor 1 (<i>RUNX1</i>)		
rs2268277	<u>Rheumatoid arthritis</u> Alters binding site in this transcription factor.	[661]
Solute carrier family 22,(organic cation transporter),member 4 (<i>SLC22A4</i>)		
rs1050152	<u>Crohn's disease</u> Alters transcription and transporter function of protein.	[662,663]

SNP	Selected previous disease associations and functional effect	Ref
Solute carrier family 22,(organic cation transporter),member 5(<i>SLC22A5</i>)		
rs11739135	<u>Crohn's disease</u> Alters transcription and transporter function of protein.	[662,663]
Tumour necrosis factor (TNF superfamily, member 2) (<i>TNF</i>)		
rs1799964	<u>Psoriatic arthritis, Crohn's disease</u>	[664-
rs1800629	Upregulates transcription of gene.	666]
Tumour necrosis factor superfamily, member 1A (<i>TNFRSF1A</i>)		
rs767455	<u>Crohn's disease</u> Polymorphism influences disease phenotype.	[667]
Tumour necrosis factor superfamily, member 1B (<i>TNFRSF1B</i>)		
rs3397	<u>Crohn's disease</u> Polymorphism influences disease phenotype.	[667]
Tumour necrosis factor (ligand) superfamily, member 15 (<i>TNFSF15</i>)		
rs6478108	<u>Crohn's disease</u>	[668]
rs7848647	May influence expression levels of the protein.	

8.2.1. Results

8.2.1.1. Single-point statistical analysis

All SNPs were in Hardy-Weinberg equilibrium (HWE) in the control samples. Since this was an exploratory study, no corrections for multiple testing were done and the threshold for significance was set at 0.05.

Four SNPs, from *RUNX1*, *IL1RN*, *CIITA* and *TNFRSF1B*, were associated with TB in the case-control study (Table 119). One significant association ($p = 0.02$), with a SNP in *INSIG2*, was detected in the transmission disequilibrium test (TDT, Table 120), but this polymorphism was not associated with disease in the case-control study. Since the case-control study has greater power than the TDT to detect associations, it is likely that this is not a reflection of a true genetic association.

TABLE 119. Single-point statistical analysis in the case-control study.

Polymorphism	Allele		Genotype frequencies						Cases vs Controls p value ^c
			Cases (n ^a =432)			Controls (n=482)			
	1 ^b	2	11	12	22	11	12	22	
rs1045642	T	C	0.07	0.35	0.58	0.08	0.40	0.52	0.15
rs1050152	T	C	0.01	0.16	0.83	0.01	0.19	0.80	0.35
rs11209026	A	G	0	0.06	0.94	0	0.05	0.95	0.61
rs11739135	A	G	0	0.16	0.84	0.01	0.18	0.80	0.37
rs1607616	A	T	0.15	0.45	0.40	0.13	0.48	0.39	0.61
rs1799964	C	T	0.04	0.38	0.58	0.04	0.35	0.62	0.50
rs1800629	A	G	0.01	0.24	0.75	0.02	0.27	0.71	0.55
rs1992660	G	A	0.22	0.48	0.31	0.20	0.52	0.27	0.34
rs2032582	T	G	0.06	0.31	0.63	0.06	0.35	0.59	0.44
rs2067819	A	G	0.07	0.37	0.56	0.07	0.38	0.55	0.91
rs2240340	G	A	0.16	0.49	0.35	0.21	0.45	0.35	0.17
rs2241880	C	T	0.12	0.39	0.49	0.12	0.45	0.43	0.27
rs2268277	G	C	0.17	0.47	0.36	0.23	0.51	0.26	0.006
rs2419598	A	T	0.13	0.44	0.42	0.18	0.50	0.33	0.01
rs2476601	A	G	0	0.03	0.97	0	0.03	0.97	0.74
rs3087243	A	G	0.08	0.33	0.59	0.08	0.34	0.58	0.95
rs3087456	A	G	0.05	0.37	0.58	0.10	0.44	0.46	0.0004
rs3397	T	C	0.10	0.40	0.50	0.14	0.43	0.43	0.04
rs3729961	G	C	0.10	0.44	0.46	0.09	0.39	0.52	0.16
rs3730293	C	T	0.03	0.33	0.64	0.03	0.30	0.67	0.66
rs3789243	T	C	0.14	0.45	0.41	0.15	0.46	0.39	0.70
rs3892175	A	G	0.01	0.19	0.80	0.01	0.16	0.83	0.58
rs569108	C	T	0.01	0.24	0.75	0.02	0.27	0.71	0.39
rs6478108	C	T	0.04	0.34	0.63	0.05	0.36	0.59	0.41
rs7528684	A	G	0.10	0.41	0.49	0.10	0.48	0.42	0.08
rs7566605	C	G	0.04	0.31	0.65	0.05	0.31	0.64	0.82
rs767455	C	T	0.08	0.37	0.55	0.05	0.40	0.55	0.16
rs7848647	T	C	0.04	0.34	0.62	0.07	0.33	0.60	0.26
rs8055955	A	G	0.04	0.36	0.60	0.04	0.31	0.65	0.30

^a number of individuals investigated.

^b Allele 1 is the minor allele.

^c p value from a genotype-based χ^2 test

TABLE 120. Single-point statistical analysis in the TDT.

SNP	Overtransmitted	Transmitted:Untransmitted	χ^2	p value
rs1045642	T	48 : 45	0.10	0.76
rs1050152	T	15 : 13	0.14	0.71
rs11209026	A	11 : 9	0.20	0.65
rs11739135	G	14 : 13	0.04	0.85
rs1607616	A	56 : 46	0.98	0.32
rs1799964	C	39 : 33	0.50	0.48
rs1800629	G	26 : 24	0.08	0.78
rs1992660	G	51 : 49	0.04	0.84
rs2032582	T	37 : 33	0.23	0.63
rs2067819	G	33 : 22	2.20	0.14
rs2240340	A	58 : 49	0.76	0.38
rs2241880	C	40 : 34	0.49	0.49
rs2268277	C	48 : 34	2.39	0.12
rs2419598	T	53 : 48	0.25	0.62
rs2476601	-	-	0	1
rs3087243	A	28 : 27	0.02	0.89
rs3087456	A	43 : 32	1.61	0.20
rs3397	C	46 : 36	1.22	0.27
rs3729961	G	45 : 32	2.20	0.14
rs3730293	T	32 : 31	0.02	0.90
rs3789243	C	46 : 38	0.76	0.38
rs3892175	G	14 : 9	1.09	0.30
rs569108	C	26 : 23	0.18	0.67
rs6478108	T	33 : 28	0.41	0.52
rs7528684	A	53 : 49	0.16	0.69
rs7566605	G	44 : 24	5.88	0.02
rs767455	C	34 : 31	0.14	0.71
rs7848647	C	27 : 18	1.80	0.18
rs8055955	A	24 : 22	0.09	0.77

8.2.2. Discussion

In this chapter we have tested the involvement of functional SNPs, previously associated with other immune-mediated diseases, in host genetic susceptibility to TB. These variants are TB susceptibility candidates, since it is possible that similar immunogenetic mechanisms may contribute to susceptibility to common diseases [172]. This is evident from the treatment of some immune-mediated diseases in patients latently infected with *M.tuberculosis* who then develop active TB [636,669]. Polymorphisms from four genes were significantly associated with TB in the case-control study and one SNP was associated with TB in the TDT.

The *CIITA* rs3087456 SNP was associated with TB ($p = 0.0004$, OR = 1.53 95% CI [1.24-1.88]) in the South African Coloured population. This promoter polymorphism

was previously associated with rheumatoid arthritis and systemic lupus erythematosus [642,644] and regulates the expression of the *CIITA* gene [643]. The protein encoded is essential in the expression of MHC class II molecules. Mice deficient for the *CIITA* gene are extremely susceptible to *M.tuberculosis* infection [670]. The study also showed that *CIITA* expression in mice with a functional *CIITA* was elevated after infection with the bacterium and that this increased expression decreased the growth rate of the mycobacteria three weeks after infection [670]. It is therefore likely that the A allele of the rs3087456 SNP may contribute to resistance to TB in humans. This association is the only one that would remain significant after a stringent Bonferroni correction ($p = 0.01$).

The rs2268277 SNP ($p = 0.006$, OR = 1.36 95% CI [1.12-1.65]) is present in the *RUNX1* gene and was previously associated with rheumatoid arthritis [661]. This SNP alters a binding site in the *RUNX1* protein, which is a transcription factor. The protein regulates the expression of many genes involved in the immune system, such as the macrophage colony-stimulating factor receptor and T cell receptor β genes [671]. It is therefore possible that it may also regulate the expression of genes involved in the host defense against *M.tuberculosis*.

The rs2419598 SNP in *IL1RN* was associated with TB in the South African Coloured population ($p = 0.01$, OR = 1.33 95% CI [1.10-1.61]). This SNP was previously associated with rheumatoid arthritis and ankylosing spondylitis [647,648] and reduced production of the *IL1RN* protein. Wilkinson et al determined that a haplotype containing a repeat polymorphism of *IL1RN* was associated with TB [253]. They also found that *M. tuberculosis*-induced *IL1RN* protein production was increased in healthy individuals with this haplotype compared to individuals without it. Awomoyi et al studied the same polymorphism as Wilkinson et al, but found no association with TB susceptibility in Gambians [193].

The *TNFRSF1B* rs3397 SNP was weakly associated with TB in our study ($p = 0.04$, OR = 1.22 95% CI [1.00-1.49]). The TNF- α receptor encoded by this gene, TNF receptor 2 (TNFR2), is part of the TNF superfamily and may influence the biological activity of TNF- α by acting as both an agonist and an antagonist [672]. Mice without a functional copy of the *TNFRSF1B* gene die when infected with *M.bovis* BCG, but are more resistant to infection than *TNFRSF1A* knockout mice [673]. Certain mycobacterial strains, such as H37RV, induce the release of TNFR2 by infected alveolar macrophages. This leads to the inhibition of TNF- α and decreases apoptosis of the infected macrophages [674]. In a recent study Stein et al [426] found an association between disease and the *TNFRSF1A* gene, also part of the TNF superfamily, but we did not observe any association in our study.

The *INSIG2* rs7566605 SNP was associated with TB only in the TDT ($p = 0.02$), but due to the reduced power of the TDT in comparison to the case-control study it is probably not a reflection of a true association. The rs2476601 *PTPN22* SNP was previously associated with TB in Colombians [659], but was not associated with disease in our study of the South African Coloured population. This gene encodes the Lyp protein which binds to C-terminal Src tyrosine kinase to suppress T cell activation. The rs2476601 SNP alters a P1 binding motif in the protein which prevents C-terminal Src

tyrosine kinase from binding it. This inhibits formation of the complex and suppresses T cell activation [660].

The small values of the odds ratios from the associated SNPs indicate that these polymorphisms are not major contributors to TB susceptibility. Since this was an exploratory study, no corrections for multiple testing were done. For this reason, a replication study in a different population group is necessary. Such a project is currently being done in TB samples from Ghana. If associations are found with the same markers as in this study, it will verify the involvement of these functional SNPs in genetic susceptibility to TB in the global population. However, these associations could also be specific to the South African Coloured population and might therefore not be reproducible in other populations.

CHAPTER 9

CONCLUDING REMARKS

This study contributes significantly to the TB host genetics field, since several candidate genes, never investigated before, were tested as susceptibility factors. In addition, we considered susceptibility genes previously identified in other populations. This is important, because association results should ideally be replicated in more than one population. In total, 136 polymorphisms in 39 genes (Table 121), selected because of their potential involvement in host genetic susceptibility to tuberculosis (TB), were genotyped in the South African Coloured population during this study. All polymorphisms were analysed using both case-control studies and transmission disequilibrium tests (TDTs); and linkage disequilibrium (LD) as well as haplotypes in the candidate genes were investigated. We also determined allele frequencies for polymorphisms, which allowed comparisons to other populations.

TABLE 121. *Polymorphisms genotyped and their single-point association results.*

Polymorphism	Case-control study	TDT	Section
C-C chemokine ligand-2 gene			4.2.1
rs2857654	✗	✗	
rs1024611	✗	✗	
rs1024610	✗	✗	
rs3760399	✗	✗	
rs3760396	✗	✗	
rs4586	✗	✗	
rs2530797	✗	✗	
Interleukin 4 gene			4.3.1
rs2243248	✗	✗	
rs2243250	✗	✗	
rs2070874	✗	✗	
rs2243251	✗	✗	
rs2243291	✗	✗	
Interleukin 10 gene			4.4.1
rs1800890	✗	✗	
rs1800893	✗	✗	
rs1800896	✗	✗	
rs1800871	✗	✗	
rs1800872	✗	✗	
rs3024490	✗	✗	
rs3790622	✗	✗	
rs3024496	✗	✗	
rs3024498	✗	✗	
Interleukin 12, beta gene			4.5.1.a)
rs730691	✗	✗	
rs3212217	✗	✗	
rs3212220	✗	✗	

Polymorphism	Case-control study	TDT	Section
D5S291	✗	✗	
rs3213096	✗	✗	
rs2288831	✗	✗	
rs2853696	✗	✗	
rs3212227	✗	✗	
Interleukin 12 receptor, beta 1			4.5.1.b)
rs393548	✗	✓	
rs2305743	✗	✗	
rs11086087	✗	✗	
rs429774	✗	✗	
rs375947	✗	✗	
Interleukin 12 receptor, beta 2 gene			4.5.1.c)
rs11576006	✗	✗	
rs3762317	✗	✗	
Interleukin 18 gene			4.6.1
rs1946519	✗	✗	
rs1946518	✗	✗	
rs187238	✗	✗	
rs5744229	✗	✗	
rs189667	✗	✗	
rs549908	✗	✗	
SH2 domain protein 1A gene (<i>SH2D1A</i>)			4.7.1
rs990545	✓	Not done ^a	
Toll-like receptor 2 gene (<i>TLR2</i>)			4.8.1
rs4696480	✗	✗	
TLR2_GT	✗	✗	
rs3804099	✗	✗	
rs3804100	✓	✗	
rs5743704	✗	✗	
rs5743708	✗	✗	
Toll-like receptor 4 gene			4.8.1
rs1927914	✗	✗	
rs10759932	✗	✗	
rs2770148	✗	✗	
rs4986790	✗	✗	
rs4986791	✗	✗	
rs11536889	✗	✗	
rs11536891	✗	✗	

Polymorphism	Case-control study	TDT	Section
Wingless-type MMTV integration site family, member 5A gene			4.9.1
rs3796232	x	x	
rs1795651	x	x	
rs7624718	x	x	
rs557077	x	x	
rs566926	x	x	
rs648872	x	x	
rs815541	x	x	
rs9311564	x	x	
rs472631	x	x	
rs556874	x	x	
rs11918967	x	x	
rs7622120	x	x	
rs590386	x	x	
Frizzled homolog 5 gene			4.9.1
rs10188753	x	x	
rs718290	x	x	
rs7582078	x	x	
rs2010400	x	x	
rs6708488	x	x	
rs3731568	x	x	
Butyrophilin-like 2 gene (<i>BTNL2</i>)			5.2.1
rs3817974	x	x	
rs17208699	x	x	
rs3817973	x	x	
rs3817972	x	x	
ss65713168	x	x	
ss65713167	x	x	
ss65713164	x	x	
ss65713163	x	x	
rs2076530	x	x	
rs9268480	x	x	
rs2076529	x	x	
ss65713152	x	x	
ss65713150	x	x	
rs3763317	x	x	
rs9268501	x	x	
rs9268502	x	x	
rs9268503	x	x	
rs9268504	x	x	

Polymorphism	Case-control study	TDT	Section
Caspase recruitment domain-containing protein 15 gene (<i>CARD15</i>)			5.3.1
rs2066844	✗	✗	
rs2066845	✗	✗	
rs2066847	✗	✗	
Testis expressed 264 gene (<i>TEX264</i>)			6.2.1
rs6775666	✗	✗	
rs4355273	✗	✗	
rs28994878	✗	✗	
Suppressor of cytokine signalling 3 gene (<i>SOCS3</i>)			6.2.2
rs4969168	✗	✗	
rs4969169	✗	✗	
rs7207782	✗	✗	
Cathepsin Z (<i>CTSZ</i>)			7.2.1
rs448943	✗	✗	
rs10369	✗	✗	
rs13720	✗	✗	
rs6064734	✗	✗	
rs3787492	✗	✗	
rs163785	✗	✗	
rs11540881	✗	✗	
ATP-binding cassette, subfamily B (MDR/TAP), member 1 gene			8.2.1
rs3789243	✗	✗	
rs1045642	✗	✗	
rs2032582	✗	✗	
ATG16 autophagy related 16-like 1 (<i>S. cerevisiae</i>) gene			8.2.1
rs2241880	✗	✗	
Cytotoxic T-lymphocyte-associated protein 4 gene			8.2.1
rs3087243	✗	✗	
Class II, major histocompatibility complex, transactivator gene (<i>CIITA</i>)			8.2.1
rs3087456	✓	✗	
Fc receptor-like 3 gene			8.2.1
rs7528684	✗	✗	
Interleukin 1 receptor antagonist gene (<i>IL1RN</i>)			8.2.1
rs2419598	✓	✗	
Interleukin 23 receptor gene			8.2.1
rs11209026	✗	✗	
Interleukin 6 signal transducer (gp130, oncostatin M reporter) gene			8.2.1
rs3730293	✗	✗	
rs3729961	✗	✗	

Polymorphism	Case-control study	TDT	Section
Insulin induced gene 2 (<i>INSIG2</i>)			8.2.1
rs7566605	✘	✓	
Membrane-spanning 4 domains, subfamily F, member 2 gene			8.2.1
rs569108	✘	✘	
Nel-like 1			8.2.1
rs1607616	✘	✘	
Peptidyl arginine deiminase, type IV gene			8.2.1
rs2240340	✘	✘	
Peroxisome proliferator-activated receptor gamma gene			8.2.1
rs2067819	✘	✘	
rs3892175	✘	✘	
Prostaglandin E receptor 4 (subtype EP4) gene			8.2.1
rs1992660	✘	✘	
Protein tyrosine phosphatase, non-receptor type 22 gene			8.2.1
rs2476601	✘	✘	
Runt-related transcription factor 1 gene (<i>RUNX1</i>)			8.2.1
rs2268277	✓	✘	
Solute carrier family 22,(organic cation transporter),member 4 gene			8.2.1
rs1050152	✘	✘	
Solute carrier family 22,(organic cation transporter),member 5 gene			8.2.1
rs11739135	✘	✘	
Tumour necrosis factor (TNF superfamily, member 2) gene			8.2.1
rs1799964	✘	✘	
rs1800629	✘	✘	
Tumour necrosis factor superfamily, member 1A gene			8.2.1
rs767455	✘	✘	
Tumour necrosis factor superfamily, member 1B gene (<i>TNFRSF1B</i>)			8.2.1
rs3397	✓	✘	
Tumour necrosis factor (ligand) superfamily, member 15 gene			8.2.1
rs6478108	✘	✘	
rs7848647	✘	✘	

✘ *Not associated with TB susceptibility.*

✓ *Associated with TB susceptibility.*

^a *No TDT was done for this X-linked polymorphism, since one sample from each family was used to increase the sample numbers of the case-control study.*

Twelve candidate genes which could regulate interferon gamma (IFN- γ) levels during TB were selected from literature searches. Two polymorphisms, in *SH2D1A* and *TLR2*, had significant associations with TB in the case-control studies. In addition, a novel allele of a microsatellite was identified in the study population. Since most of these IFN- γ -regulating genes on their own were not identified as genetic susceptibility factors

for TB in this study, it is possible that their combined effect on the production of this cytokine may contribute to disease in the South African Coloured population. This is a future study direction.

Susceptibility genes from granulomatous diseases other than TB, namely sarcoidosis and Crohn's disease (CD), were investigated. We did not detect any association of the polymorphisms tested in these genes (*BTNL2*, *CARD15*) with TB susceptibility in the South African Coloured population. In addition, 15 novel single nucleotide polymorphisms (SNPs) in *BTNL2* were identified from resequencing samples from the South African Coloured population.

We selected two genes (*TEX264*, *SOCS3*), derived from an expression-array analysis, to investigate as novel candidate genes for TB susceptibility. During this study, 2 novel polymorphisms were identified in the samples from the South African Coloured population. No statistical associations were found between selected polymorphisms in the two genes and TB. A candidate gene (*CTSZ*) located in a region that was previously identified as containing TB susceptibility genes was investigated as a susceptibility gene, but we detected no association.

In an exploratory study, we tested the involvement of various known functional polymorphisms involved in other immune-mediated diseases, in TB susceptibility. Four SNPs in *CIITA*, *RUNXI*, *IL1RN* and *TNFRSF1B* were associated with TB susceptibility in the case-control analyses, while one SNP in *INSIG2* was associated in the TDT.

Even though this work did not replicate several previously described associations, it is still of value. Association results that are not statistically significant or are negative are often not published, which makes it difficult to interpret the real importance of a reported association. Some of the previously published associations considered during our study had extremely low sample numbers and were therefore underpowered. The lack of reproducibility of certain associations could also be a result of ethnic-specific associations. Alternatively, polymorphisms could have smaller effect sizes in the South African Coloured population than we were able to detect with our sample.

The human immune system consists of a complex network of interacting pathways which protect against infection. Some of these pathways are redundant and when a defect is present they may compensate for each other. Within pathways, components may also be redundant. For this reason it is extremely difficult to detect the effect of a single gene polymorphism, which could be one explanation for why we did not find statistical associations with certain SNPs. A proposed solution to this is to consider all the genes in a pathway known to be involved in TB, which may mean that single SNP and candidate gene analysis might become of less importance individually, and require consolidation.

The South African Coloured population, a unique admixed population, is likely to have more variability than white populations in various genes, due to African admixture. In addition, stronger LD is expected in admixed populations than in more ethnically homogenous populations [578,675]. This was observed in the LD analyses in our study when comparative analyses between the South African Coloured and other populations

were done. LD, which refers to the linkage of alleles at different sites along the chromosome, is an important tool in disease susceptibility studies, since it can be used to identify the functional markers that contribute to the disease without genotyping every known polymorphism in the gene region. This feature of admixed populations is another advantage of using the South African Coloured population in association studies.

In spite of its strengths as discussed above, this research had certain limitations:

1) The polymorphisms selected for genotyping in this study were all located in or around protein-coding genes, based on the dogma that most DNA that determines phenotype and form is expressed as proteins. However, DNA sequences in introns or intergenic regions that do not encode proteins or cis-acting regulatory elements are also transcribed in humans and have biological functions [676]. In addition, the existence of numerous intercalated transcripts spanning the majority of the genome has been established [677]. Therefore the view of the genome as having a limited set of loci transcribed independently is oversimplified. Since these mechanisms are not currently understood, we did not include polymorphisms located in intergenic DNA in our research at this stage.

2) A hypothesis that genetic polymorphisms are not the only inherited features of DNA which may influence disease susceptibility has been considered for several years [678]. Epigenetic mechanisms, such as DNA methylation, RNA interference and histone acetylation, regulate the transcription rate and/or tissue-specific expression of certain genes without altering the primary nucleotide sequence of the DNA. These DNA features have been shown to be transmitted to subsequent generations, but they may also be influenced by the environment [678]. However, this inheritance is not as stable as DNA-based inheritance and it is not completely elucidated. It has been shown that *M.tuberculosis* may interfere with epigenetic regulation [679]. Therefore, once the inheritance is understood, it is possible that these epigenetic mechanisms may be found to play a role in host susceptibility to TB. However, this was not studied here.

3) Due to the high incidence of latent TB infection in the control community where sample collection was done [680], the possible associations of the SNPs with TB were tested in disease progression only. It is not clear whether a different set of genes may be involved in infection versus progression to disease, as has been supposed for solute carrier family 11, member 1 gene (*SLC11A1*) [213]. It is therefore still possible that the polymorphisms studied here are associated with primary infection in the South African Coloured population.

4) Certain polymorphisms could not be genotyped during this research, due to limitations of the genotyping methods, such as the SNPlex™ Genotyping System.

5) The number of families for the TDTs was not large enough to have the same power as the case-control association studies. This complicated the interpretation of the results, since TDTs were used as a secondary means to judge the validity of the significant association results found for the case-control studies. It is unlikely that the discrepant results in the TDTs were due to population stratification, as this was excluded in previous work on this population [176]. However, this does not mean that

the family-based study design is not a good validation study. If the number of families is equal to the number of cases available, TDTs will have enough power to replicate associations found in case-controls studies [280] as shown previously by Valentonyte et al [289] and others.

6) TB is a complex disease and several genes may make small contributions to the disease outcome. It is possible that our study did not have enough power to detect these small effect sizes, especially for SNPs that had allele frequencies less than 5% in the South African Coloured population. However, very few low frequency SNPs ($n = 15$) were studied.

7) During this research we did not make use of all available single-SNP tests. Since it is not known what amount of disease-susceptibility polymorphisms are dominant, recessive or additive, the design of optimal analyses is not possible and it is up to the researcher to decide which test is more suited for analysis of the available data [74]. Our research tested the null hypothesis of no association between rows and columns of the 2 x 3 table among cases and controls. Other single-SNP tests not considered in this study include the Cochran-Armitage test, logistic regression and linear regression. However, any statistical analysis can only give an indication of a genetic association. The functional validation of reported associations is therefore necessary.

8) Statistical genetics should include correction for multiple testing and this was done in this study, either by Bonferroni correction as suggested by Nicodemus et al [76] or by permutation testing. However, the possibility exists that these tests may be too stringent and that correcting for multiple tests may lead to the exclusion of a true susceptibility variant.

Future work would include the replication of the associations found in this study. This should preferably be done in a population that is ethnically distinct from the South African Coloured population. In addition, the combined effect of TB susceptibility genes should be studied and functionality of the associated polymorphisms should be tested. A genome-wide association (GWA) study in this unique TB population would provide a wealth of data if markers that represent the most common variation in the South African Coloured population and which cover the whole genome are determined. Ideally, such an experiment should be done using genome-wide SNP arrays [681]. This microarray-based method allows access to large amounts of genetic information. A SNP array that allows genotyping of more than 900 000 SNPs [682] was recently released. Alternatively, mapping by admixture linkage disequilibrium (MALD), could be considered [683]. This method employs the strong LD in admixed populations, such as the South African Coloured population, to localise disease gene polymorphisms which have different allele frequencies in the parental populations of the admixed population. The parental populations should differ in the incidence of the disease studied [578]. Such a study in the South African Coloured population will therefore be feasible, since it is known that at least two of the parental populations, namely Europeans and African blacks, have different incidences of TB.

The idea that TB is not only influenced by the bacterium, but also by both genetic and environmental factors, was restated by Rene Dubos 55 years ago [684]. Dubos was also

of the opinion that medical solutions alone would not prevent or cure TB. It is obvious that this has been the case so far. However, Dubos did not consider the possible effect of research in the field of human genetic susceptibility to TB. Even though this research is years from being clinically applied, it is likely that in the coming era of personalised medicine, which may arrive faster than we imagine, it will be possible to test individuals for their susceptibility to a number of diseases with a genetic component, including TB. This would allow identification of more vulnerable individuals in the population who may need different vaccination strategies or prolonged treatment when exposed to *M.tuberculosis*.

REFERENCES

1. Ducati RG, Ruffino-Netto A, Basso LA, and Santos DS. The resumption of consumption-- a review on tuberculosis. *Mem Inst Oswaldo Cruz* 2006; 101:697-714.
2. World Health Organization. WHO Report 2007: Global tuberculosis control - surveillance, planning, financing. WHO/HTM/TB/2006.362. 2007. Geneva.
3. World Health Organization. WHO report 2005: Global tuberculosis control - surveillance, planning, financing. WHO/HTM/TB/2005.349. 2005. Geneva.
4. World Health Organization. WHO Report 2006: Global tuberculosis control - surveillance, planning, financing. WHO/HTM/TB/2006.362. 2006. Geneva.
5. Mathema B, Kurepina NE, Bifani PJ, and Kreiswirth BN. Molecular epidemiology of tuberculosis: current insights. *Clin Microbiol Rev* 2006; 19:658-85.
6. Coovadia HM and Benatar SR. A century of tuberculosis - South African perspectives. Cape Town: Oxford University Press, 1991.
7. Daniel TM. The history of tuberculosis. *Respir Med* 2006; 100:1862-70.
8. Doherty TM and Andersen P. Vaccines for tuberculosis: novel concepts and recent progress. *Clin Microbiol Rev* 2005; 18:687-702.
9. Kaufmann SH, Baumann S, and Nasser EA. Exploiting immunology and molecular genetics for rational vaccine design against tuberculosis. *Int J Tuberc Lung Dis* 2006; 10:1068-79.
10. Canetti G, Rist N, and Grosset J. Primary drug resistance in tuberculosis. *Am Rev Respir Dis* 1964; 90:792-9.
11. Rom WN and Garay SM. Tuberculosis. Philadelphia: Lippincott Williams & Wilkins, 2004.
12. Cole ST, Brosch R, Parkhill J et al. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* 1998; 393:537-44.
13. Camus JC, Pryor MJ, Medigue C, and Cole ST. Re-annotation of the genome sequence of *Mycobacterium tuberculosis* H37Rv. *Microbiology* 2002; 148:2967-73.
14. Murray PR, Barron EJ, Pfaller MA, Tenover FC, and Tenover FC. *Manual of clinical Microbiology*. Washington: American Society of Microbiology, 1999.
15. Greenwood D, Slack RCB, and Peutherer JF. *Medical Microbiology - A guide to microbial infections: pathogenesis, immunity, laboratory diagnosis and control*. London: Churchill Livingstone, 1997.
16. Hingley-Wilson SM, Sambandamurthy VK, and Jacobs WR, Jr. Survival perspectives from the world's most successful pathogen, *Mycobacterium tuberculosis*. *Nat Immunol* 2003; 4:949-55.
17. Houben EN, Nguyen L, and Pieters J. Interaction of pathogenic mycobacteria with the host immune system. *Curr Opin Microbiol* 2006; 9:76-85.
18. Mueller P and Pieters J. Modulation of macrophage antimicrobial mechanisms by pathogenic mycobacteria. *Immunobiology* 2006; 211:549-56.
19. Quast TM and Browning RF. Pathogenesis and clinical manifestations of pulmonary tuberculosis. *Dis Mon* 2006; 52:413-9.
20. van Helden PD, Möller M, Babb C et al. TB epidemiology and human genetics. *Novartis Found Symp* 2006; 279:17-31.
21. Pieters J. Entry and survival of pathogenic mycobacteria in macrophages. *Microbes Infect* 2001; 3:249-55.

22. Frieden TR, Sterling TR, Munsiff SS, Watt CJ, and Dye C. Tuberculosis. *Lancet* 2003; 362:887-99.
23. Kaufmann SH and McMichael AJ. Annulling a dangerous liaison: vaccination strategies against AIDS and tuberculosis. *Nat Med* 2005; 11:S33-S44.
24. Smith I. Mycobacterium tuberculosis pathogenesis and molecular determinants of virulence. *Clin Microbiol Rev* 2003; 16:463-96.
25. Gleeson TD and Decker CF. Treatment of tuberculosis. *Dis Mon* 2006; 52:428-34.
26. Johnson R, Streicher EM, Louw GE, Warren RM, van Helden PD, and Victor TC. Drug resistance in Mycobacterium tuberculosis. *Curr Issues Mol Biol* 2006; 8:97-111.
27. Nations JA, Lazarus AA, and Walsh TE. Drug-resistant tuberculosis. *Dis Mon* 2006; 52:435-40.
28. Frieden TR, Sherman LF, Maw KL et al. A multi-institutional outbreak of highly drug-resistant tuberculosis: epidemiology and clinical outcomes. *JAMA* 1996; 276:1229-35.
29. Bloom BR and Murray CJ. Tuberculosis: commentary on a reemergent killer. *Science* 1992; 257:1055-64.
30. World Health Organization. Frequently asked questions - XDR-TB. <http://www.who.int/tb/xdr/faqs/en/index.html> . 2007.
31. Johnson MD and Decker CF. Tuberculosis and HIV Infection. *Dis Mon* 2006; 52:420-7.
32. Whalen C, Horsburgh CR, Hom D, Lahart C, Simberkoff M, and Ellner J. Accelerated course of human immunodeficiency virus infection after tuberculosis. *Am J Respir Crit Care Med* 1995; 151:129-35.
33. Bellamy R. Genetics and pulmonary medicine. 3. Genetic susceptibility to tuberculosis in human populations. *Thorax* 1998; 53:588-93.
34. Sorensen TI, Nielsen GG, Andersen PK, and Teasdale TW. Genetic and environmental influences on premature death in adult adoptees. *N Engl J Med* 1988; 318:727-32.
35. McShane H. Susceptibility to tuberculosis - the importance of the pathogen as well as the host. *Clinical and Experimental Immunology* 2003; 133:20-1.
36. Motulsky AG. Metabolic polymorphisms and the role of infectious diseases in human evolution. *Hum Biol* 1960; 32:28-62.
37. Dubos R and Dubos J. *The white plague: tuberculosis, man and society*. 1952.
38. Stead WW, Senner JW, Reddick WT, and Lofgren JP. Racial differences in susceptibility to infection by Mycobacterium tuberculosis. *N Engl J Med* 1990; 322:422-7.
39. Alcais A and Abel L. Application of genetic epidemiology to dissecting host susceptibility/resistance to infection illustrated with the study of common mycobacterial infections. In: Bellamy R ed. Cambridge: Cambridge University Press, 2004:7-33.
40. Comstock GW. Tuberculosis in twins: a re-analysis of the Proffit survey. *Am Rev Respir Dis* 1978; 117:621-4.
41. Kallmann FJ and Reisner D. Twin studies on the significance of genetic factors in tuberculosis. *Am Rev Tuberc* 1942; 47:549-7.
42. Simmonds B. *Tuberculosis in twins*. Pitman Medical Publishing Company, 2004.

43. Cooke GS and Hill AV. Genetics of susceptibility to human infectious disease. *Nat Rev Genet* 2001; 2:967-77.
44. Caragol I, Raspall M, Fieschi C et al. Clinical tuberculosis in 2 of 3 siblings with interleukin-12 receptor beta1 deficiency. *Clin Infect Dis* 2003; 37:302-6.
45. de Jong R, Altare F, Haagen IA et al. Severe mycobacterial and Salmonella infections in interleukin-12 receptor-deficient patients. *Science* 1998; 280:1435-8.
46. Ozbek N, Fieschi C, Yilmaz BT et al. Interleukin-12 receptor beta 1 chain deficiency in a child with disseminated tuberculosis. *Clin Infect Dis* 2005; 40:e55-e58.
47. Sakai T, Matsuoka M, Aoki M, Nosaka K, and Mitsuya H. Missense mutation of the interleukin-12 receptor beta1 chain-encoding gene is associated with impaired immunity against Mycobacterium avium complex infection. *Blood* 2001; 97:2688-94.
48. Fieschi C, Dupuis S, Catherinot E et al. Low penetrance, broad resistance, and favorable outcome of interleukin 12 receptor beta1 deficiency: medical and immunological implications. *J Exp Med* 2003; 197:527-35.
49. Sanal O, Turul T, De Boer T et al. Presentation of interleukin-12/-23 receptor beta1 deficiency with various clinical symptoms of Salmonella infections. *J Clin Immunol* 2006; 26:1-6.
50. Mansouri D, Adimi P, Mirsaeidi M et al. Inherited disorders of the IL-12-IFN-gamma axis in patients with disseminated BCG infection. *Eur J Pediatr* 2005; 164:753-7.
51. Flynn JL. Lessons from experimental Mycobacterium tuberculosis infections. *Microbes Infect* 2006; 8:1179-88.
52. Gupta UD and Katoch VM. Animal models of tuberculosis. *Tuberculosis (Edinb)* 2005; 85:277-93.
53. Allison MJ, Zappasodi P, and Lurie MB. Host-parasite relationships in natively resistant and susceptible rabbits on quantitative inhalation of tubercle bacilli. Their significance for the nature of genetic resistance. *Am Rev Respir Dis* 1962; 85:553-69.
54. Chackerian AA and Behar SM. Susceptibility to Mycobacterium tuberculosis: lessons from inbred strains of mice. *Tuberculosis (Edinb)* 2003; 83:279-85.
55. Hirschhorn JN and Daly MJ. Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet* 2005; 6:95-108.
56. Clementi M and Di Gianantonio E. Genetic susceptibility to infectious diseases. *Reprod Toxicol* 2006; 21:345-9.
57. Burgner D, Jamieson SE, and Blackwell JM. Genetic susceptibility to infectious diseases: big is beautiful, but will bigger be even better? *Lancet Infect Dis* 2006; 6:653-63.
58. Sherry ST, Ward MH, Kholodov M et al. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* 2001; 29:308-11.
59. The International HapMap Consortium. The International HapMap Project. *Nature* 2003; 426:789-96.
60. Venter JC, Adams MD, Myers EW et al. The sequence of the human genome. *Science* 2001; 291:1304-51.
61. Lander ES, Linton LM, Birren B et al. Initial sequencing and analysis of the human genome. *Nature* 2001; 409:860-921.

62. The International HapMap Consortium. A haplotype map of the human genome. *Nature* 2005; 437:1299-320.
63. Abel L and Dessein AJ. The impact of host genetics on susceptibility to human infectious diseases. *Curr Opin Immunol* 1997; 9:509-16.
64. Roberts M, Mock BA, and Blackwell JM. Mapping of genes controlling *Leishmania major* infection in CXS recombinant inbred mice. *Eur J Immunogenet* 1993; 20:349-62.
65. Jamieson SE, Miller EN, Black GF et al. Evidence for a cluster of genes on chromosome 17q11-q21 controlling susceptibility to tuberculosis and leprosy in Brazilians. *Genes Immun* 2004; 5:46-57.
66. Bellamy R, Beyers N, McAdam KP et al. Genetic susceptibility to tuberculosis in Africans: a genome-wide scan. *Proc Natl Acad Sci U S A* 2000; 97:8005-9.
67. Cervino AC, Lakiss S, Sow O et al. Fine mapping of a putative tuberculosis-susceptibility locus on chromosome 15q11-13 in African families. *Hum Mol Genet* 2002; 11:1599-603.
68. Greenwood CM, Fujiwara TM, Boothroyd LJ et al. Linkage of tuberculosis to chromosome 2q35 loci, including *NRAMP1*, in a large aboriginal Canadian family. *Am J Hum Genet* 2000; 67:405-16.
69. Miller EN, Jamieson SE, Joberty C et al. Genome-wide scans for leprosy and tuberculosis susceptibility genes in Brazilians. *Genes Immun* 2004; 5:63-7.
70. Baghdadi JE, Orlova M, Alter A et al. An autosomal dominant major gene confers predisposition to pulmonary tuberculosis in adults. *J Exp Med* 2006; 203:1679-84.
71. Daly AK and Day CP. Candidate gene case-control association studies: advantages and potential pitfalls. *Br J Clin Pharmacol* 2001; 52:489-99.
72. Risch N and Merikangas K. The future of genetic studies of complex human diseases. *Science* 1996; 273:1516-7.
73. Gambaro G, Anglani F, and D'Angelo A. Association studies of genetic polymorphisms and complex disease. *Lancet* 2000; 355:308-11.
74. Balding DJ. A tutorial on statistical methods for population association studies. *Nat Rev Genet* 2006; 7:781-91.
75. Healy DG. Case-control studies in the genomic era: a clinician's guide. *Lancet Neurol* 2006; 5:701-7.
76. Nicodemus KK, Liu W, Chase GA, Tsai YY, and Fallin MD. Comparison of type I error for multiple test corrections in large single-nucleotide polymorphism studies using principal components versus haplotype blocking algorithms. *BMC Genet* 2005; 6 Suppl 1:S78.
77. Rosenberg PS, Che A, and Chen BE. Multiple hypothesis testing strategies for genetic case-control association studies. *Stat Med* 2006; 25:3134-49.
78. Becker T and Knapp M. A powerful strategy to account for multiple testing in the context of haplotype analysis. *Am J Hum Genet* 2004; 75:561-70.
79. Nyholt DR. A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. *Am J Hum Genet* 2004; 74:765-9.
80. Gordi T and Khamis H. Simple solution to a common statistical problem: interpreting multiple tests. *Clin Ther* 2004; 26:780-6.

81. Epstein MP, Allen AS, and Satten GA. A simple and improved correction for population stratification in case-control studies. *Am J Hum Genet* 2007; 80:921-30.
82. Spielman RS, McGinnis RE, and Ewens WJ. Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am J Hum Genet* 1993; 52:506-16.
83. Schaid DJ. Transmission disequilibrium, family controls, and great expectations. *Am J Hum Genet* 1998; 63:935-41.
84. Spielman RS and Ewens WJ. The TDT and other family-based tests for linkage disequilibrium and association. *Am J Hum Genet* 1996; 59:983-9.
85. Spielman RS and Ewens WJ. A sibship test for linkage in the presence of association: the sib transmission/disequilibrium test. *Am J Hum Genet* 1998; 62:450-8.
86. Sun F, Flanders WD, Yang Q, and Khoury MJ. Transmission disequilibrium test (TDT) when only one parent is available: the 1-TDT. *Am J Epidemiol* 1999; 150:97-104.
87. Sebastiani P, Abad MM, Alpargu G, and Ramoni MF. Robust transmission/disequilibrium test for incomplete family genotypes. *Genetics* 2004; 168:2329-37.
88. Evangelou E, Trikalinos TA, Salanti G, and Ioannidis JP. Family-based versus unrelated case-control designs for genetic associations. *PLoS Genet* 2006; 2:e123.
89. Gabriel SB, Schaffner SF, Nguyen H et al. The structure of haplotype blocks in the human genome. *Science* 2002; 296:2225-9.
90. Daly MJ, Rioux JD, Schaffner SF, Hudson TJ, and Lander ES. High-resolution haplotype structure in the human genome. *Nat Genet* 2001; 29:229-32.
91. Patil N, Berno AJ, Hinds DA et al. Blocks of limited haplotype diversity revealed by high-resolution scanning of human chromosome 21. *Science* 2001; 294:1719-23.
92. Hampe J, Franke A, Rosenstiel P et al. A genome-wide association scan of nonsynonymous SNPs identifies a susceptibility variant for Crohn disease in ATG16L1. *Nat Genet* 2007; 39:207-11.
93. Ozaki K, Ohnishi Y, Iida A et al. Functional SNPs in the lymphotoxin-alpha gene that are associated with susceptibility to myocardial infarction. *Nat Genet* 2002; 32:650-4.
94. Ozaki K, Inoue K, Sato H et al. Functional variation in LGALS2 confers risk of myocardial infarction and regulates lymphotoxin-alpha secretion in vitro. *Nature* 2004; 429:72-5.
95. Ozaki K and Tanaka T. Genome-wide association study to identify single-nucleotide polymorphisms conferring risk of myocardial infarction. *Methods Mol Med* 2006; 128:173-80.
96. Klein RJ, Zeiss C, Chew EY et al. Complement factor H polymorphism in age-related macular degeneration. *Science* 2005; 308:385-9.
97. Duerr RH, Taylor KD, Brant SR et al. A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. *Science* 2006; 314:1461-3.
98. Stephens M, Smith NJ, and Donnelly P. A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* 2001; 68:978-89.

99. Stephens M and Donnelly P. A comparison of bayesian methods for haplotype reconstruction from population genotype data. *Am J Hum Genet* 2003; 73:1162-9.
100. Scheet P and Stephens M. A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. *Am J Hum Genet* 2006; 78:629-44.
101. Dudbridge F. Pedigree disequilibrium tests for multilocus haplotypes. *Genet Epidemiol* 2003; 25:115-21.
102. Levenstien MA, Ott J, and Gordon D. Are molecular haplotypes worth the time and expense? A cost-effective method for applying molecular haplotypes. *PLoS Genet* 2006; 2:e127.
103. Jeffreys AJ, Kauppi L, and Neumann R. Intensely punctate meiotic recombination in the class II region of the major histocompatibility complex. *Nat Genet* 2001; 29:217-22.
104. Gu S, Pakstis AJ, Li H, Speed WC, Kidd JR, and Kidd KK. Significant variation in haplotype block structure but conservation in tagSNP patterns among global populations. *European Journal of Human Genetics* 2007.
105. Conrad DF, Jakobsson M, Coop G et al. A worldwide survey of haplotype variation and linkage disequilibrium in the human genome. *Nat Genet* 2006; 38:1251-60.
106. Goddard KA, Hopkins PJ, Hall JM, and Witte JS. Linkage disequilibrium and allele-frequency distributions for 114 single-nucleotide polymorphisms in five populations. *Am J Hum Genet* 2000; 66:216-34.
107. Carlson CS, Eberle MA, Rieder MJ, Smith JD, Kruglyak L, and Nickerson DA. Additional SNPs and linkage-disequilibrium analyses are necessary for whole-genome association studies in humans. *Nat Genet* 2003; 33:518-21.
108. Wang N, Akey JM, Zhang K, Chakraborty R, and Jin L. Distribution of recombination crossovers and the origin of haplotype blocks: the interplay of population history, recombination, and mutation. *Am J Hum Genet* 2002; 71:1227-34.
109. Phillips MS, Lawrence R, Sachidanandam R et al. Chromosome-wide distribution of haplotype blocks and the role of recombination hot spots. *Nat Genet* 2003; 33:382-7.
110. Abecasis GR, Ghosh D, and Nichols TE. Linkage disequilibrium: ancient history drives the new genetics. *Hum Hered* 2005; 59:118-24.
111. Weiss KM and Clark AG. Linkage disequilibrium and the mapping of complex human traits. *Trends Genet* 2002; 18:19-24.
112. Abecasis GR and Cookson WOC. GOLD - Graphical Overview of Linkage Disequilibrium. *Bioinformatics* 2000; 16:182-3.
113. Ardlie KG, Kruglyak L, and Seielstad M. Patterns of linkage disequilibrium in the human genome. *Nat Rev Genet* 2002; 3:299-309.
114. Mueller JC. Linkage disequilibrium for different scales and applications. *Brief Bioinform* 2004; 5:355-64.
115. Wall JD and Pritchard JK. Haplotype blocks and linkage disequilibrium in the human genome. *Nat Rev Genet* 2003; 4:587-97.
116. Murphy PM. Molecular mimicry and the generation of host defense protein diversity. *Cell* 1993; 72:823-6.

117. Hoal EG. Human genetic susceptibility to tuberculosis and other mycobacterial diseases. *IUBMB Life* 2002; 53:225-9.
118. Janeway C, Travers P, Walport M, and Shlomchik M. *Immunobiology: the immune system in health and disease*. New York: Garland Science, 2001.
119. Vidal SM, Malo D, Vogan K, Skamene E, and Gros P. Natural resistance to infection with intracellular parasites: isolation of a candidate for Bcg. *Cell* 1993; 73:469-85.
120. Malo D, Vogan K, Vidal S et al. Haplotype mapping and sequence analysis of the mouse *Nramp* gene predict susceptibility to infection with intracellular parasites. *Genomics* 1994; 23:51-61.
121. Vidal S, Tremblay ML, Govoni G et al. The *Ity/Lsh/Bcg* locus: natural resistance to infection with intracellular parasites is abrogated by disruption of the *Nramp1* gene. *J Exp Med* 1995; 182:655-66.
122. North RJ, LaCourse R, Ryan L, and Gros P. Consequence of *Nramp1* deletion to *Mycobacterium tuberculosis* infection in mice. *Infect Immun* 1999; 67:5811-4.
123. Cellier M, Govoni G, Vidal S et al. Human natural resistance-associated macrophage protein: cDNA cloning, chromosomal mapping, genomic organization, and tissue-specific expression. *J Exp Med* 1994; 180:1741-52.
124. Gruenheid S, Pinner E, Desjardins M, and Gros P. Natural resistance to infection with intracellular pathogens: the *Nramp1* protein is recruited to the membrane of the phagosome. *J Exp Med* 1997; 185:717-30.
125. Denis M. Killing of *Mycobacterium tuberculosis* within human monocytes: activation by cytokines and calcitriol. *Clin Exp Immunol* 1991; 84:200-6.
126. Rockett KA, Brookes R, Udalova I, Vidal V, Hill AV, and Kwiatkowski D. 1,25-Dihydroxyvitamin D₃ induces nitric oxide synthase and suppresses growth of *Mycobacterium tuberculosis* in a human macrophage-like cell line. *Infect Immun* 1998; 66:5314-21.
127. Rook GA, Steele J, Fraher L et al. Vitamin D₃, gamma interferon, and control of proliferation of *Mycobacterium tuberculosis* by human monocytes. *Immunology* 1986; 57:159-63.
128. Bellamy R, Ruwende C, Corrah T et al. Tuberculosis and chronic hepatitis B virus infection in Africans and variation in the vitamin D receptor gene. *J Infect Dis* 1999; 179:721-4.
129. Selvaraj P, Narayanan PR, and Reetha AM. Association of vitamin D receptor genotypes with the susceptibility to pulmonary tuberculosis in female patients & resistance in female contacts. *Indian J Med Res* 2000; 111:172-9.
130. Wilkinson RJ, Llewelyn M, Toossi Z et al. Influence of vitamin D deficiency and vitamin D receptor polymorphisms on tuberculosis among Gujarati Asians in west London: a case-control study. *Lancet* 2000; 355:618-21.
131. Liu W, Cao WC, Zhang CY et al. *VDR* and *NRAMP1* gene polymorphisms in susceptibility to pulmonary tuberculosis among the Chinese Han population: a case-control study. *Int J Tuberc Lung Dis* 2004; 8:428-34.
132. Bornman L, Campbell SJ, Fielding K et al. Vitamin D receptor polymorphisms and susceptibility to tuberculosis in West Africa: a case-control and family study. *J Infect Dis* 2004; 190:1631-41.

133. Fitness J, Floyd S, Warndorff DK et al. Large-scale candidate gene study of tuberculosis susceptibility in the Karonga district of northern Malawi. *Am J Trop Med Hyg* 2004; 71:341-9.
134. Soborg C, Andersen AB, Range N et al. Influence of candidate susceptibility genes on tuberculosis in a high endemic region. *Mol Immunol* 2007; 44:2213-20.
135. Lewis SJ, Baker I, and Davey SG. Meta-analysis of vitamin D receptor polymorphisms and pulmonary tuberculosis risk. *Int J Tuberc Lung Dis* 2005; 9:1174-7.
136. Babb C, van der Merwe L, Beyers N et al. Vitamin D receptor gene polymorphisms and sputum conversion time in pulmonary tuberculosis patients. *Tuberculosis* 2007; 87:295-302.
137. Roth DE, Soto G, Arenas F et al. Association between Vitamin D Receptor Gene Polymorphisms and Response to Treatment of Pulmonary Tuberculosis. *J Infect Dis* 2004; 190:920-7.
138. Pan H, Yan BS, Rojas M et al. *Ipr1* gene mediates innate immunity to tuberculosis. *Nature* 2005; 434:767-72.
139. Tosh K, Campbell SJ, Fielding K et al. Variants in the *SP110* gene are associated with genetic susceptibility to tuberculosis in West Africa. *Proc Natl Acad Sci U S A* 2006; 103:10364-8.
140. Thye T, Browne EN, Chinbuah MA et al. No associations of human pulmonary tuberculosis with Sp110 variants. *J Med Genet* 2006; 43:e32.
141. Szeszko JS, Healy B, Stevens H et al. Resequencing and association analysis of the *SP110* gene in adult pulmonary tuberculosis. *Hum Genet* 2006.
142. Babb C, Keet EH, van Helden PD, and Hoal EG. *SP110* polymorphisms are not associated with pulmonary tuberculosis in a South African population. *Hum Genet* 2007; 121:521-2.
143. Balamurugan A, Sharma SK, and Mehra NK. Human leukocyte antigen class I supertypes influence susceptibility and severity of tuberculosis. *J Infect Dis* 2004; 189:805-11.
144. Brahmajothi V, Pitchappan RM, Kakkanaiah VN et al. Association of pulmonary tuberculosis and HLA in south India. *Tubercle* 1991; 72:123-32.
145. Hafez M, el Salab S, el Shennawy F, and Bassiony MR. HLA-antigens and tuberculosis in the Egyptian population. *Tubercle* 1985; 66:35-40.
146. Mahmoudzadeh-Niknam H, Khalili G, and Fadavi P. Allelic distribution of human leukocyte antigen in Iranian patients with pulmonary tuberculosis. *Hum Immunol* 2003; 64:124-9.
147. Chandanayingyong D, Maranetra N, and Bovornkitti S. HLA antigen profiles in Thai tuberculosis patients. *Asian Pac J Allergy Immunol* 1988; 6:77-80.
148. Ruggiero G, Cosentini E, Zanzi D et al. Allelic distribution of human leucocyte antigen in historical and recently diagnosed tuberculosis patients in Southern Italy. *Immunology* 2004; 111:318-22.
149. Al Arif LI, Goldstein RA, Affronti LF, and Janicki BW. HLA-Bw15 and tuberculosis in a North American black population. *Am Rev Respir Dis* 1979; 120:1275-8.
150. Delgado JC, Baena A, Thim S, and Goldfeld AE. Aspartic acid homozygosity at codon 57 of *HLA-DQ* beta is associated with susceptibility to pulmonary tuberculosis in Cambodia. *J Immunol* 2006; 176:1090-7.

151. Teran-Escandon D, Teran-Ortiz L, Camarena-Olvera A et al. Human leukocyte antigen-associated susceptibility to pulmonary tuberculosis: molecular analysis of class II alleles by DNA amplification and oligonucleotide hybridization in Mexican patients. *Chest* 1999; 115:428-33.
152. Amirzargar AA, Yalda A, Hajabolbaghi M et al. The association of *HLA-DRB*, *DQA1*, *DQB1* alleles and haplotype frequency in Iranian patients with pulmonary tuberculosis. *Int J Tuberc Lung Dis* 2004; 8:1017-21.
153. Vejbaesya S, Chierakul N, Luangtrakool K, Srinak D, and Stephens HA. Associations of HLA class II alleles with pulmonary tuberculosis in Thais. *Eur J Immunogenet* 2002; 29:431-4.
154. Dubaniewicz A, Moszkowska G, Szczerkowska Z, and Hoppe A. Analysis of *DQB1* allele frequencies in pulmonary tuberculosis: preliminary report. *Thorax* 2003; 58:890-1.
155. Lombard Z, Dalton DL, Venter PA, Williams RC, and Bornman L. Association of *HLA-DR*, *-DQ*, and vitamin D receptor alleles and haplotypes with tuberculosis in the Venda of South Africa. *Hum Immunol* 2006; 67:643-54.
156. Goldfeld AE, Delgado JC, Thim S et al. Association of an *HLA-DQ* allele with clinical tuberculosis. *JAMA* 1998; 279:226-8.
157. Singh SP, Mehra NK, Dingley HB, Pande JN, and Vaidya MC. Human leukocyte antigen (HLA)-linked control of susceptibility to pulmonary tuberculosis and association with HLA-DR types. *J Infect Dis* 1983; 148:676-81.
158. Khomenko AG, Litvinov VI, Chukanova VP, and Pospelov LE. Tuberculosis in patients with various HLA phenotypes. *Tubercle* 1990; 71:187-92.
159. Rajalingam R, Singal DP, and Mehra NK. Transporter associated with antigen-processing (TAP) genes and susceptibility to tuberculoid leprosy and pulmonary tuberculosis. *Tissue Antigens* 1997; 49:168-72.
160. Selvaraj P, Uma H, Reetha AM et al. HLA antigen profile in pulmonary tuberculosis patients & their spouses. *Indian J Med Res* 1998; 107:155-8.
161. Bothamley GH, Beck JS, Schreuder GM et al. Association of tuberculosis and M. tuberculosis-specific antibody levels with HLA. *J Infect Dis* 1989; 159:549-55.
162. Pospelov LE, Matrakshin AG, Chernousova LN et al. Association of various genetic markers with tuberculosis and other lung diseases in Tuvian children. *Tuber Lung Dis* 1996; 77:77-80.
163. Cox RA, Downs M, Neimes RE, Ognibene AJ, Yamashita TS, and Ellner JJ. Immunogenetic analysis of human tuberculosis. *J Infect Dis* 1988; 158:1302-8.
164. Pospelova LE, Matrashkin AG, Larionova EE, Ereemeev VV, and Mes'ko EM. [The association of tuberculosis with the specificities of the HLA gene *DRB1* in different regions of Tuva]. *Probl Tuberk Bolezn Legk* 2005;23-5.
165. Kim HS, Park MH, Song EY et al. Association of *HLA-DR* and *HLA-DQ* Genes With Susceptibility to Pulmonary Tuberculosis in Koreans: Preliminary Evidence of Associations With Drug Resistance, Disease Severity, and Disease Recurrence. *Hum Immunol* 2005; 66:1074-81.
166. Wang J, Song C, and Wang S. [Association of *HLA-DRB1* genes with pulmonary tuberculosis]. *Zhonghua Jie He He Hu Xi Za Zhi* 2001; 24:302-5.
167. Dubaniewicz A, Lewko B, Moszkowska G, Zamorska B, and Stepinski J. Molecular subtypes of the HLA-DR antigens in pulmonary tuberculosis. *Int J Infect Dis* 2000; 4:129-33.

168. Sriram U, Selvaraj P, Kurian SM, Reetha AM, and Narayanan PR. HLA-DR2 subtypes & immune responses in pulmonary tuberculosis. *Indian J Med Res* 2001; 113:117-24.
169. Ravikumar M, Dheenadhayalan V, Rajaram K et al. Associations of *HLA-DRB1*, *DQB1* and *DPB1* alleles with pulmonary tuberculosis in south India. *Tuber Lung Dis* 1999; 79:309-17.
170. Rajalingam R, Mehra NK, Jain RC, Myneedu VP, and Pande JN. Polymerase chain reaction--based sequence-specific oligonucleotide hybridization analysis of HLA class II antigens in pulmonary tuberculosis: relevance to chemotherapy and disease severity. *J Infect Dis* 1996; 173:669-76.
171. Scola L, Crivello A, Marino V et al. *IL-10* and *TNF-alpha* polymorphisms in a sample of Sicilian patients affected by tuberculosis: implication for ageing and life span expectancy. *Mech Ageing Dev* 2003; 124:569-72.
172. Correa PA, Gomez LM, Cadena J, and Anaya JM. Autoimmunity and tuberculosis. Opposite association with *TNF* polymorphism. *J Rheumatol* 2005; 32:219-24.
173. Hill AVS, Ruwende C, McGuire W et al. Association of the *TNF* -238 promoter polymorphism with susceptibility to tuberculosis and malaria in Africa. *Human Immunology* 1996; 47:118.
174. Gadzhiev BO, Sergeev AS, and Movsum-zade KM. [Association of mutant alleles of serum cholinesterase with various multi-factorial and infectious diseases]. *Genetika* 1987; 23:559-64.
175. Flores-Villanueva PO, Ruiz-Morales JA, Song CH et al. A functional promoter polymorphism in monocyte chemoattractant protein-1 is associated with increased susceptibility to pulmonary tuberculosis. *J Exp Med* 2005; 202:1649-58.
176. Barreiro LB, Neyrolles O, Babb CL et al. Promoter Variation in the DC-SIGN Encoding Gene *CD209* Is Associated with Tuberculosis. *PLoS Medicine* 2006; 3:e20.
177. Matrakshin AG, Tsoi KN, Pospelov LE, Kapina MA, Kholod ON, and Pushkina EI. [A genotypic study of children ill with tuberculosis and of healthy BCG-revaccinated ones of Tuvinian nationality]. *Probl Tuberk* 1993;25-7.
178. Kharakter Z, Mazhak KD, and Pavlenko AV. [The role of genetically determined haptoglobin phenotypes in patients with destructive pulmonary tuberculosis]. *Probl Tuberk* 1990;50-2.
179. Lopez-Maderuelo D, Arnalich F, Serantes R et al. Interferon-gamma and interleukin-10 gene polymorphisms in pulmonary tuberculosis. *Am J Respir Crit Care Med* 2003; 167:970-5.
180. Rossouw M, Nel HJ, Cooke GS, van Helden PD, and Hoal EG. Association between tuberculosis and a polymorphic NFkappaB binding site in the interferon gamma gene. *Lancet* 2003; 361:1871-2.
181. Lio D, Marino V, Serauto A et al. Genotype frequencies of the +874T-->A single nucleotide polymorphism in the first intron of the interferon-gamma gene in a sample of Sicilian patients affected by tuberculosis. *Eur J Immunogenet* 2002; 29:371-4.
182. Sallakci, Nilgun, Coskun, Mesut, Berber, Zafer, Gurkan, Fuat, Kocamaz, Halil, Uysal, Gulnar, Bhujju, Sabin, Yavuzer, Ugur, Singh, Mahavir, and Yegin, Olcay.

- Interferon- γ gene+874T-A polymorphism is associated with tuberculosis and gamma interferon response. *Tuberculosis* . 2007.
183. Tso HW, Ip WK, Chong WP, Tam CM, Chiang AK, and Lau YL. Association of interferon gamma and interleukin 10 genes with tuberculosis in Hong Kong Chinese. *Genes Immun* 2005; 6:358-63.
 184. Henao MI, Montes C, Paris SC, and Garcia LF. Cytokine gene polymorphisms in Colombian patients with different clinical presentations of tuberculosis. *Tuberculosis (Edinb)* 2005.
 185. Etokebe GE, Bulat-Kardum L, Johansen MS et al. Interferon-gamma gene (T874A and G2109A) polymorphisms are associated with microscopy-positive tuberculosis. *Scand J Immunol* 2006; 63:136-41.
 186. Fraser DA, Bulat-Kardum L, Knezevic J et al. Interferon-gamma receptor-1 gene polymorphism in tuberculosis patients from Croatia. *Scandinavian Journal of Immunology* 2003; 57:480-4.
 187. Rosenzweig SD, Schaffer AA, Ding L et al. Interferon-gamma receptor 1 promoter polymorphisms: population distribution and functional implications. *Clin Immunol* 2004; 112:113-9.
 188. Bulat-Kardum L, Etokebe GE, Knezevic J et al. Interferon-gamma receptor-1 gene promoter polymorphisms (G-611A; T-56C) and susceptibility to tuberculosis. *Scand J Immunol* 2006; 63:142-50.
 189. Oral HB, Budak F, Uzaslan EK et al. Interleukin-10 (*IL-10*) gene polymorphism as a potential host susceptibility factor in tuberculosis. *Cytokine* 2006; 35:143-7.
 190. Delgado JC, Baena A, Thim S, and Goldfeld AE. Ethnic-specific genetic associations with pulmonary tuberculosis. *J Infect Dis* 2002; 186:1463-8.
 191. Kusuhara K, Yamamoto K, Okada K, Mizuno Y, and Hara T. Association of *IL12RB1* polymorphisms with susceptibility to and severity of tuberculosis in Japanese: a gene-based association analysis of 21 candidate genes. *Int J Immunogenet* 2007; 34:35-44.
 192. Gomez LM, Camargo JF, Castiblanco J, Ruiz-Narvaez EA, Cadena J, and Anaya JM. Analysis of *IL1B*, *TAP1*, *TAP2* and *IKBL* polymorphisms on susceptibility to tuberculosis. *Tissue Antigens* 2006; 67:290-6.
 193. Awomoyi AA, Charurat M, Marchant A et al. Polymorphism in *IL1B*: *IL1B*-511 association with tuberculosis and decreased lipopolysaccharide-induced IL-1beta in IFN-gamma primed ex-vivo whole blood assay. *J Endotoxin Res* 2005; 11:281-6.
 194. Bellamy R, Ruwende C, Corrah T, McAdam KP, Whittle HC, and Hill AV. Assessment of the interleukin 1 gene cluster and other candidate gene polymorphisms in host susceptibility to tuberculosis. *Tuber Lung Dis* 1998; 79:83-9.
 195. Vidyarani M, Selvaraj P, Prabhu AS, Jawahar MS, Adhilakshmi AR, and Narayanan PR. Interferon gamma (*IFNgamma*) & interleukin-4 (*IL-4*) gene variants & cytokine levels in pulmonary tuberculosis. *Indian J Med Res* 2006; 124:403-10.
 196. Ma X, Reich RA, Wright JA et al. Association between interleukin-8 gene alleles and human susceptibility to tuberculosis disease. *Journal of Infectious Diseases* 2003; 188:349-55.

197. Hoal-van Helden EG, Epstein J, Victor TC et al. Mannose-binding protein B allele confers protection against tuberculous meningitis. *Pediatr Res* 1999; 45:459-64.
198. El Sahly HM, Reich RA, Dou SJ, Musser JM, and Graviss EA. The effect of mannose binding lectin gene polymorphisms on susceptibility to tuberculosis in different ethnic groups. *Scand J Infect Dis* 2004; 36:106-8.
199. Mombo LE, Lu CY, Ossari S et al. Mannose-binding lectin alleles in sub-Saharan Africans and relation with susceptibility to infections. *Genes Immun* 2003; 4:362-7.
200. Soborg C, Madsen HO, Andersen AB, Lillebaek T, Kok-Jensen A, and Garred P. Mannose-binding lectin polymorphisms in clinical tuberculosis. *J Infect Dis* 2003; 188:777-82.
201. Selvaraj P, Narayanan PR, and Reetha AM. Association of functional mutant homozygotes of the mannose binding protein gene with susceptibility to pulmonary tuberculosis in India. *Tuber Lung Dis* 1999; 79:221-7.
202. Bellamy R, Ruwende C, McAdam KP et al. Mannose binding protein deficiency is not associated with malaria, hepatitis B carriage nor tuberculosis in Africans. *QJM* 1998; 91:13-8.
203. Li CM, Campbell SJ, Kumararatne DS et al. Association of a polymorphism in the *P2X7* gene with tuberculosis in a Gambian population. *J Infect Dis* 2002; 186:1458-62.
204. Fernando SL, Saunders BM, Sluyter R et al. A polymorphism in the *P2X7* gene increases susceptibility to extrapulmonary tuberculosis. *Am J Respir Crit Care Med* 2007; 175:360-6.
205. Papiha SS, Agarwal SS, and White I. Association between phosphoglucomutase (PGM1) and group-specific component (Gc) subtypes and tuberculosis. *J Med Genet* 1983; 20:220-2.
206. Papiha SS, Singh BN, Lanchbury JS et al. Association of *HLA* and other genetic markers in South Indian patients with pulmonary tuberculosis. *Tubercle* 1987; 68:159-67.
207. Gao PS, Fujishima S, Mao XQ et al. Genetic variants of *NRAMP1* and active tuberculosis in Japanese populations. International Tuberculosis Genetics Team. *Clin Genet* 2000; 58:74-6.
208. Duan HF, Zhou XH, Ma Y et al. [A study on the association of 3'UTR polymorphisms of *NRAMP1* gene with susceptibility to tuberculosis in Hans]. *Zhonghua Jie He He Hu Xi Za Zhi* 2003; 26:286-9.
209. Liu W, Zhang CY, Tian L et al. [A case-control study on natural-resistance-associated macrophage protein 1 gene polymorphisms and susceptibility to pulmonary tuberculosis]. *Zhonghua Yu Fang Yi Xue Za Zhi* 2003; 37:408-11.
210. Hoal EG, Lewis L-A, Jamieson SE et al. *SLC11A1* (*NRAMP1*) but not *SLC11A2* (*NRAMP2*) polymorphisms are associated with susceptibility to tuberculosis in a high-incidence community in South Africa. *Int J Tuberc Lung Dis* 2004; 8:1464-71.
211. Koh WJ, Kwon OJ, Kim EJ, Lee KS, Ki CS, and Kim JW. *NRAMP1* gene polymorphism and susceptibility to nontuberculous mycobacterial lung diseases. *Chest* 2005; 128:94-101.

212. Zhang W, Shao L, Weng X et al. Variants of the natural resistance-associated macrophage protein 1 gene (*NRAMP1*) are associated with severe forms of pulmonary tuberculosis. *Clin Infect Dis* 2005; 40:1232-6.
213. Malik S, Abel L, Tooker H et al. Alleles of the *NRAMP1* gene are risk factors for pediatric tuberculosis disease. *Proc Natl Acad Sci U S A* 2005; 102:12183-8.
214. Leung KH, Yip SP, Wong WS et al. Sex- and age-dependent association of *SLC11A1* polymorphisms with tuberculosis in Chinese: a case control study. *BMC Infect Dis* 2007; 7:19.
215. Taype CA, Castro JC, Accinelli RA, Herrera-Velit P, Shaw MA, and Espinoza JR. Association between *SLC11A1* polymorphisms and susceptibility to different clinical forms of tuberculosis in the Peruvian population. *Infect Genet Evol* 2006; 6:361-7.
216. Ryu S, Park YK, Bai GH, Kim SJ, Park SN, and Kang S. 3'UTR polymorphisms in the *NRAMP1* gene are associated with susceptibility to tuberculosis in Koreans. *Int J Tuberc Lung Dis* 2000; 4:577-80.
217. Abe T, Iinuma Y, Ando M et al. *NRAMP1* polymorphisms, susceptibility and clinical features of tuberculosis. *J Infect* 2003; 46:215-20.
218. Awomoyi AA, Marchant A, Howson JM, McAdam KP, Blackwell JM, and Newport MJ. Interleukin-10, polymorphism in *SLC11A1* (formerly *NRAMP1*), and susceptibility to tuberculosis. *J Infect Dis* 2002; 186:1808-14.
219. Ma X, Dou S, Wright JA et al. 5' dinucleotide repeat polymorphism of *NRAMP1* and susceptibility to tuberculosis among Caucasian patients in Houston, Texas. *Int J Tuberc Lung Dis* 2002; 6:818-23.
220. Bellamy R, Ruwende C, Corrah T, McAdam KP, Whittle HC, and Hill AV. Variations in the *NRAMP1* gene and susceptibility to tuberculosis in West Africans. *N Engl J Med* 1998; 338:640-4.
221. Cervino AC, Lakiss S, Sow O, and Hill AV. Allelic association between the *NRAMP1* gene and susceptibility to tuberculosis in Guinea-Conakry. *Ann Hum Genet* 2000; 64:507-12.
222. Soborg C, Andersen AB, Madsen HO, Kok-Jensen A, Skinhoj P, and Garred P. Natural resistance-associated macrophage protein 1 polymorphisms are associated with microscopy-positive tuberculosis. *J Infect Dis* 2002; 186:517-21.
223. Qu YB, Tang YX, Zhang ZB et al. [Relationship between single nucleotide polymorphisms of *NRAMP1* gene and susceptibility to pulmonary tuberculosis in workers exposed to silica dusts]. *Zhonghua Lao Dong Wei Sheng Zhi Ye Bing Za Zhi* 2006; 24:531-3.
224. Shaw MA, Collins A, Peacock CS et al. Evidence that genetic susceptibility to *Mycobacterium tuberculosis* in a Brazilian population is under oligogenic control: linkage study of the candidate genes *NRAMP1* and *TNFA*. *Tuber Lung Dis* 1997; 78:35-45.
225. Floros J, Lin HM, Garcia A et al. Surfactant protein genetic marker alleles identify a subgroup of tuberculosis in a Mexican population. *J Infect Dis* 2000; 182:1473-8.
226. Malik S, Greenwood CM, Eguale T et al. Variants of the *SFTPA1* and *SFTPA2* genes and susceptibility to tuberculosis in Ethiopia. *Hum Genet* 2006; 118:752-9.
227. Madan T, Saxena S, Murthy KJ, Muralidhar K, and Sarma PU. Association of polymorphisms in the collagen region of human *SP-A1* and *SP-A2* genes with

- pulmonary tuberculosis in Indian population. Clin Chem Lab Med 2002; 40:1002-8.
228. Hawn TR, Dunstan SJ, Thwaites GE et al. A polymorphism in Toll-interleukin 1 receptor domain containing adaptor protein is associated with susceptibility to meningeal tuberculosis. J Infect Dis 2006; 194:1127-34.
 229. Ogun AC, Yoldas B, Ozdemir T et al. The Arg753Gln polymorphism of the human toll-like receptor 2 gene in tuberculosis disease. Eur Respir J 2004; 23:219-23.
 230. Selvaraj P, Kurian SM, Chandra G, Reetha AM, Charles N, and Narayanan PR. Vitamin D receptor gene variants of *Bsm1*, *Apal*, *TaqI*, and *FokI* polymorphisms in spinal tuberculosis. Clin Genet 2004; 65:73-6.
 231. Liu W, Zhang CY, Wu XM et al. [A case-control study on the vitamin D receptor gene polymorphisms and susceptibility to pulmonary tuberculosis]. Zhonghua Liu Xing Bing Xue Za Zhi 2003; 24:389-92.
 232. Cox RA, Arnold DR, Cook D, and Lundberg DI. HLA phenotypes in Mexican Americans with tuberculosis. Am Rev Respir Dis 1982; 126:653-5.
 233. Sanjeevi CB, Narayanan PR, Prabakar R et al. No association or linkage with *HLA-DR* or *-DQ* genes in south Indians with pulmonary tuberculosis. Tuber Lung Dis 1992; 73:280-4.
 234. Singh SP, Mehra NK, Dingley HB, Pande JN, and Vaidya MC. HLA haplotype segregation study in multiple case families of pulmonary tuberculosis. Tissue Antigens 1984; 23:84-6.
 235. Blackwell JM. Genetics and genomics in infectious disease susceptibility. Trends Mol Med 2001; 7:521-6.
 236. Hawkins BR, Higgins DA, Chan SL, Lowrie DB, Mitchison DA, and Girling DJ. HLA typing in the Hong Kong Chest Service/British Medical Research Council study of factors associated with the breakdown to active tuberculosis of inactive pulmonary lesions. Am Rev Respir Dis 1988; 138:1616-21.
 237. Selvaraj P, Sriram U, Mathan KS, Reetha AM, and Narayanan PR. Tumour necrosis factor alpha (-238 and -308) and beta gene polymorphisms in pulmonary tuberculosis: haplotype analysis with *HLA-A*, *B* and *DR* genes. Tuberculosis (Edinb) 2001; 81:335-41.
 238. Roh EY, Park MH, Yoon JA et al. Association of *TAP1* and *TAP2* gene polymorphisms with tuberculosis in Koreans. Human Immunology 2004; 65:S35.
 239. Möller M, Kwiatkowski R, Nebel A, van Helden PD, Hoal EG, and Schreiber S. Allelic variation in *BTNL2* and susceptibility to tuberculosis in a South African population. Microbes Infect 2007; 9:522-8.
 240. Stockton JC, Howson JM, Awomoyi AA, McAdam KP, Blackwell JM, and Newport MJ. Polymorphism in *NOD2*, Crohn's disease, and susceptibility to pulmonary tuberculosis. FEMS Immunol Med Microbiol 2004; 41:157-60.
 241. Möller M, Nebel A, Kwiatkowski R, van Helden PD, Hoal EG, and Schreiber S. Host susceptibility to tuberculosis: *CARD15* polymorphisms in a South African population. Mol Cell Probes 2007; 21:148-51.
 242. Druszczynska M, Strapagiel D, Kwiatkowska S et al. Tuberculosis bacilli still posing a threat. Polymorphism of genes regulating anti-mycobacterial properties of macrophages. Pol J Microbiol 2006; 55:7-12.

243. Pacheco E, Fonseca C, Montes C, Zabaleta J, Garcia LF, and Arias MA. *CD14* gene promoter polymorphism in different clinical forms of tuberculosis. *FEMS Immunol Med Microbiol* 2004; 40:207-13.
244. Gomez LM, Anaya JM, Sierra-Filardi E, Cadena J, Corbi A, and Martin J. Analysis of DC-SIGN (*CD209*) functional variants in patients with tuberculosis. *Hum Immunol* 2006; 67:808-11.
245. Barreiro LB, Neyrolles O, Babb CL et al. Length variation of DC-SIGN and L-SIGN neck-region has no impact on tuberculosis susceptibility. *Hum Immunol* 2007; 68:106-12.
246. Kasvosve I, Gomo ZA, Mvundura E et al. Haptoglobin polymorphism and mortality in patients with tuberculosis. *Int J Tuberc Lung Dis* 2000; 4:771-5.
247. Akahoshi M, Ishihara M, Remus N et al. Association between *IFNA* genotype and the risk of sarcoidosis. *Hum Genet* 2004; 114:503-9.
248. Moran A, Ma X, Reich RA, and Graviss EA. No association between the +874T/A single nucleotide polymorphism in the IFN-gamma gene and susceptibility to TB. *Int J Tuberc Lung Dis* 2007; 11:113-5.
249. Awomoyi AA, Nejentsev S, Richardson A et al. No association between interferon-gamma receptor-1 gene polymorphism and pulmonary tuberculosis in a Gambian population sample. *Thorax* 2004; 59:291-4.
250. Mirsaeidi SM, Houshmand M, Tabarsi P et al. Lack of association between interferon-gamma receptor-1 polymorphism and pulmonary TB in Iranian population sample. *J Infect* 2006; 52:374-7.
251. Puzyrev VP, Freidin MB, Rudko AA, Strelis AK, and Kolokolova OV. [Polymorphisms of the candidate genes for genetic susceptibility to tuberculosis in the Slavic population of Siberia: a pilot study]. *Mol Biol (Mosk)* 2002; 36:788-91.
252. Lee HW, Lee HS, Kim DK et al. Lack of an association between interleukin-12 receptor beta1 polymorphisms and tuberculosis in Koreans. *Respiration* 2005; 72:365-8.
253. Wilkinson RJ, Patel P, Llewelyn M et al. Influence of polymorphism in the genes for the interleukin (IL)-1 receptor antagonist and IL-1beta on tuberculosis. *J Exp Med* 1999; 189:1863-74.
254. Cooke GS, Campbell SJ, Fielding K et al. Interleukin-8 polymorphism is not associated with pulmonary tuberculosis in the gambia. *J Infect Dis* 2004; 189:1545-6.
255. Ninomiya S, Niimi T, Shimizu S et al. Matrix metalloproteinase-1 polymorphism of promoter region in sarcoidosis and tuberculosis patients. *Sarcoidosis Vasc Diffuse Lung Dis* 2004; 21:19-24.
256. Ma X, Wright J, Dou S et al. Ethnic divergence and linkage disequilibrium of novel SNPs in the human NLI-IF gene: evidence of human origin and lack of association with tuberculosis susceptibility. *J Hum Genet* 2002; 47:140-5.
257. Vejbaesya S, Chierakul N, Luangtrakool P, and Sermduangprateep C. *NRAMP1* and *TNF-alpha* polymorphisms and susceptibility to tuberculosis in Thais. *Respirology* 2007; 12:202-6.
258. Liaw YS, Tsai-Wu JJ, Wu CH et al. Variations in the *NRAMP1* gene and susceptibility of tuberculosis in Taiwanese. *Int J Tuberc Lung Dis* 2002; 6:454-60.

259. El Baghdadi J, Remus N, Benslimane A et al. Variants of the human *NRAMP1* gene and susceptibility to tuberculosis in Morocco. *Int J Tuberc Lung Dis* 2003; 7:599-602.
260. Dubaniewicz A, Jamieson SE, Dubaniewicz-Wybieralska M, Fakiola M, Nancy ME, and Blackwell JM. Association between *SLC11A1* (formerly *NRAMP1*) and the risk of sarcoidosis in Poland. *Eur J Hum Genet* 2005; 13:829-34.
261. Awomoyi A, Sirugo G, Newport MJ, and Tishkoff S. Global distribution of a novel trinucleotide microsatellite polymorphism (ATA)_n in intron 8 of the *SLC11A1* gene and susceptibility to pulmonary tuberculosis. *Int J Immunogenet* 2006; 33:11-5.
262. Niimi T, Sato S, Sugiura Y et al. Transforming growth factor-beta gene polymorphism in sarcoidosis and tuberculosis patients. *Int J Tuberc Lung Dis* 2002; 6:510-5.
263. Ryu YJ, Kim EJ, Koh WJ, Kim H, Kwon OJ, and Chang JH. Toll-like receptor 2 polymorphisms and nontuberculous mycobacterial lung diseases. *Clin Vaccine Immunol* 2006; 13:818-9.
264. Newport MJ, Allen A, Awomoyi AA et al. The toll-like receptor 4 Asp299Gly variant: no influence on LPS responsiveness or susceptibility to pulmonary tuberculosis in The Gambia. *Tuberculosis (Edinb)* 2004; 84:347-52.
265. Campbell SJ, Sabeti P, Fielding K et al. Variants of the *CD40* ligand gene are not associated with increased susceptibility to tuberculosis in West Africa. *Immunogenetics* 2003; 55:502-7.
266. Huang S, Hendriks W, Althage A et al. Immune response in mice that lack the interferon-gamma receptor. *Science* 1993; 259:1742-5.
267. Kamijo R, Le J, Shapiro D et al. Mice that lack the interferon-gamma receptor have profoundly altered responses to infection with *Bacillus Calmette-Guerin* and subsequent challenge with lipopolysaccharide. *J Exp Med* 1993; 178:1435-40.
268. Newport MJ, Huxley CM, Huston S et al. A mutation in the interferon-gamma-receptor gene and susceptibility to mycobacterial infection. *N Engl J Med* 1996; 335:1941-9.
269. Ottenhoff TH, Verreck FA, Hoeve MA, and Van D, V. Control of human host immunity to mycobacteria. *Tuberculosis (Edinb)* 2005; 85:53-64.
270. Bonecini-Almeida MG, Chitale S, Boutsikakis I et al. Induction of in vitro human macrophage anti-*Mycobacterium tuberculosis* activity: requirement for IFN-gamma and primed lymphocytes. *J Immunol* 1998; 160:4490-9.
271. Mistry R, Cliff JM, Clayton CL et al. Gene-expression patterns in whole blood identify subjects at risk for recurrent tuberculosis. *J Infect Dis* 2007; 195:357-65.
272. Hampe J, Wollstein A, Lu T et al. An integrated system for high throughput TaqMan based SNP genotyping. *Bioinformatics* 2001; 17:654-5.
273. Teuber M, Koch WA, Manaster C, Wachter S, Hampe J, and Schreiber S. Improving quality control and workflow management in high-throughput single-nucleotide polymorphism genotyping environments. *Journal of the Association for Laboratory Automation* 2005; 10:43-7.
274. City of Cape Town / Metropole Region. TB Control Programme. Progress Report. 1997-2002. 2002.

275. Munch Z, Van Lill SWP, Booysen CN, Zietsman HL, Enarson DA, and Beyers N. Tuberculosis transmission patterns in a high-incidence area: a spatial analysis. *International Journal of Tuberculosis and Lung Disease* 2003; 7:271-7.
276. Beyers N, Gie RP, Zietsman HL et al. The use of a geographical information system (GIS) to evaluate the distribution of tuberculosis in a high-incidence community. *S Afr Med J* 1996; 86:40-1, 44.
277. Lombard Z, Brune AE, Hoal EG et al. HLA class II disease associations in southern Africa. *Tissue Antigens* 2006; 67:97-110.
278. Barreiro LB, Neyrolles O, Babb CL et al. Length Variation of DC-SIGN and L-SIGN Neck-Region has no Impact on Tuberculosis Susceptibility. *Human Immunology* 2007; 68:106-12.
279. den Boon S, Verver S, Marais BJ et al. Association between passive smoking and infection with *Mycobacterium tuberculosis* in children. *Pediatrics* 2007; 119:734-9.
280. McGinnis R, Shifman S, and Darvasi A. Power and efficiency of the TDT and case-control design for association scans. *Behav Genet* 2002; 32:135-44.
281. Bergen AW, Haque KA, Qi Y et al. Comparison of yield and genotyping performance of multiple displacement amplification and OmniPlex whole genome amplified DNA generated from multiple DNA sources. *Hum Mutat* 2005; 26:262-70.
282. Lasken RS and Egholm M. Whole genome amplification: abundant supplies of DNA from precious samples or clinical specimens. *Trends Biotechnol* 2003; 21:531-5.
283. Hughes S, Arneson N, Done S, and Squire J. The use of whole genome amplification in the study of human disease. *Prog Biophys Mol Biol* 2005; 88:173-89.
284. Bergen AW, Qi Y, Haque KA, Welch RA, and Chanock SJ. Effects of DNA mass on multiple displacement whole genome amplification and genotyping performance. *BMC Biotechnol* 2005; 5:24.
285. Barker DL, Hansen MS, Faruqi AF et al. Two methods of whole-genome amplification enable accurate genotyping across a 2320-SNP linkage panel. *Genome Res* 2004; 14:901-7.
286. Dean FB, Nelson JR, Giesler TL, and Lasken RS. Rapid amplification of plasmid and phage DNA using Phi 29 DNA polymerase and multiply-primed rolling circle amplification. *Genome Res* 2001; 11:1095-9.
287. GenomiPhi DNA amplification kit. 2005. Amersham Biosciences.
288. Manaster C, Valentonyte R, Teuber M, Zheng W, Schreiber S, and Hampe J. SGCcaller: a program to call and review genotypes measured by sequencing. *Biotechniques* 2005; 38:544, 546.
289. Valentonyte R, Hampe J, Huse K et al. Sarcoidosis is associated with a truncating splice site mutation in *BTNL2*. *Nat Genet* 2005; 37:357-64.
290. Bergholdt R, Ghandil P, Johannesen J et al. Genetic and functional evaluation of an interleukin-12 polymorphism (IDDM18) in families with type 1 diabetes. *J Med Genet* 2004; 41:e39.
291. Yim JJ, Lee HW, Lee HS et al. The association between microsatellite polymorphisms in intron II of the human Toll-like receptor 2 gene and tuberculosis among Koreans. *Genes Immun* 2006; 7:150-5.

292. Newton CR, Graham A, Heptinstall LE et al. Analysis of any point mutation in DNA. The amplification refractory mutation system (ARMS). *Nucleic Acids Res* 1989; 17:2503-16.
293. Strachan T and Read AP. *Human Molecular Genetics*. BIOS Scientific Publishers, 1999:1-557.
294. Ye S, Dhillon S, Ke X, Collins AR, and Day IN. An efficient procedure for genotyping single nucleotide polymorphisms. *Nucleic Acids Res* 2001; 29:E88.
295. Okayama N, Fujimura K, Nakamura J, Suehiro Y, Hamanaka Y, and Hinoda Y. Evaluation of a new efficient procedure for single-nucleotide polymorphism genotyping: tetra-primer amplification refractory mutation system-polymerase chain reaction. *Clin Chem Lab Med* 2004; 42:13-6.
296. van Rietschoten JG, Westland R, van den Bogaard R et al. A novel polymorphic GATA site in the human *IL-12Rbeta2* promoter region affects transcriptional activity. *Tissue Antigens* 2004; 63:538-46.
297. Schoske R, Vallone PM, Ruitberg CM, and Butler JM. Multiplex PCR design strategy used for the simultaneous amplification of 10 Y chromosome short tandem repeat (STR) loci. *Anal Bioanal Chem* 2003; 375:333-43.
298. Jasinska A and Krzyzosiak WJ. Repetitive sequences that shape the human transcriptome. *FEBS Lett* 2004; 567:136-41.
299. Dearlove AM. High throughput genotyping technologies. *Brief Funct Genomic Proteomic* 2002; 1:139-50.
300. Pravica V, Asderakis A, Perrey C, Hajeer A, Sinnott PJ, and Hutchinson IV. In vitro production of IFN-gamma correlates with CA repeat polymorphism in the human IFN-gamma gene. *Eur J Immunogenet* 1999; 26:1-3.
301. Pravica V, Perrey C, Stevens A, Lee JH, and Hutchinson IV. A single nucleotide polymorphism in the first intron of the human IFN-gamma gene: absolute correlation with a polymorphic CA microsatellite marker of high IFN-gamma production. *Hum Immunol* 2000; 61:863-6.
302. Yim JJ, Ding L, Schaffer AA, Park GY, Shim YS, and Holland SM. A microsatellite polymorphism in intron 2 of human Toll-like receptor 2 gene: functional implications and racial differences. *FEMS Immunol Med Microbiol* 2004; 40:163-9.
303. De la Vega FM, Lazaruk KD, Rhodes MD, and Wenz MH. Assessment of two flexible and compatible SNP genotyping platforms: TaqMan SNP Genotyping Assays and the SNplex Genotyping System. *Mutat Res* 2005; 573:111-35.
304. Holland PM, Abramson RD, Watson R, and Gelfand DH. Detection of specific polymerase chain reaction product by utilizing the 5'----3' exonuclease activity of *Thermus aquaticus* DNA polymerase. *Proc Natl Acad Sci U S A* 1991; 88:7276-80.
305. Livak KJ. Allelic discrimination using fluorogenic probes and the 5' nuclease assay. *Genet Anal* 1999; 14:143-9.
306. Livak KJ, Flood SJ, Marmaro J, Giusti W, and Deetz K. Oligonucleotides with fluorescent dyes at opposite ends provide a quenched probe system useful for detecting PCR product and nucleic acid hybridization. *PCR Methods Appl* 1995; 4:357-62.
307. De la Vega FM, Dailey D, Ziegler J, Williams J, Madden D, and Gilbert DA. New generation pharmacogenomic tools: a SNP linkage disequilibrium Map, validated

- SNP assay resource, and high-throughput instrumentation system for large-scale genetic studies. *Biotechniques* 2002; Suppl:48-50, 52, 54.
308. Hampe J, Cuthbert A, Croucher PJ et al. Association between insertion mutation in *NOD2* gene and Crohn's disease in German and British populations. *Lancet* 2001; 357:1925-8.
 309. Hampe J, Grebe J, Nikolaus S et al. Association of *NOD2* (*CARD15*) genotype with clinical course of Crohn's disease: a cohort study. *Lancet* 2002; 359:1661-5.
 310. SNPLEX Genotyping System 48-plex User Guide. 2004. Applied Biosystems.
 311. Nickerson DA, Kaiser R, Lappin S, Stewart J, Hood L, and Landegren U. Automated DNA diagnostics using an ELISA-based oligonucleotide ligation assay. *Proc Natl Acad Sci U S A* 1990; 87:8923-7.
 312. SNPLEX Genotyping System 48-plex Assay Design and Ordering Guide. 2005. Applied Biosystems.
 313. Barrett JC, Fry B, Maller J, and Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005; 21:263-5.
 314. Cartharius K, Frech K, Grote K et al. MatInspector and beyond: promoter analysis based on transcription factor binding sites. *Bioinformatics* 2005; 21:2933-42.
 315. Hardy GH. Mendelian proportions in a mixed population. 1908. *Yale J Biol Med* 2003; 76:79-80.
 316. Weinberg W. On the demonstration of heredity in man. *Naturkunde in Wurttemberg*, Stuttgart 1908; 64:368-82.
 317. Wigginton JE, Cutler DJ, and Abecasis GR. A note on exact tests of Hardy-Weinberg equilibrium. *Am J Hum Genet* 2005; 76:887-93.
 318. Xu J, Turner A, Little J, Bleecker ER, and Meyers DA. Positive results in association studies are associated with departure from Hardy-Weinberg equilibrium: hint for genotyping error? *Hum Genet* 2002; 111:573-4.
 319. Griffiths AJ, Miller JH, Suzuki DT, Lewontin RC, and Gelbart WM. Patterns of inheritance. W.H. Freeman & Co., 1999.
 320. Beals, M. The chi-square test. www.tiem.utk.edu/~gross/bioed/bealsmodules/chi-square.html. 2004. 7-9-2004.
 321. Motulsky H. *Intuitive Biostatistics*. Oxford University Press, Inc, 1995.
 322. Purcell S, Cherny SS, and Sham PC. Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. *Bioinformatics* 2003; 19:149-50.
 323. Lee EY, Yim JJ, Lee HS, Lee YJ, Lee EB, and Song YW. Dinucleotide repeat polymorphism in intron II of human Toll-like receptor 2 gene and susceptibility to rheumatoid arthritis. *Int J Immunogenet* 2006; 33:211-5.
 324. Yamada N, Yamaya M, Okinaga S et al. Microsatellite polymorphism in the heme oxygenase-1 gene promoter is associated with susceptibility to emphysema. *Am J Hum Genet* 2000; 66:187-95.
 325. McGinnis RE and Spielman RS. Insulin gene 5' flanking polymorphism. Length of class 1 alleles in number of repeat units. *Diabetes* 1995; 44:1296-302.
 326. Boehm U, Klamp T, Groot M, and Howard JC. Cellular responses to interferon-gamma. *Annu Rev Immunol* 1997; 15:749-95.
 327. Isaacs A and Lindenmann J. Virus interference. I. The interferon. *Proc R Soc Lond B Biol Sci* 1957; 147:258-67.

328. Schroder K, Hertzog PJ, Ravasi T, and Hume DA. Interferon-gamma: an overview of signals, mechanisms and functions. *J Leukoc Biol* 2004; 75:163-89.
329. Perussia B. Lymphokine-activated killer cells, natural killer cells and cytokines. *Curr Opin Immunol* 1991; 3:49-55.
330. Mosmann TR and Coffman RL. TH1 and TH2 cells: different patterns of lymphokine secretion lead to different functional properties. *Annu Rev Immunol* 1989; 7:145-73.
331. Sad S, Marcotte R, and Mosmann TR. Cytokine-induced differentiation of precursor mouse CD8⁺ T cells into cytotoxic CD8⁺ T cells secreting Th1 or Th2 cytokines. *Immunity* 1995; 2:271-9.
332. Skeen MJ and Ziegler HK. Activation of gamma delta T cells for production of IFN-gamma is mediated by bacteria via macrophage-derived cytokines IL-1 and IL-12. *J Immunol* 1995; 154:5832-41.
333. Carnaud C, Lee D, Donnars O et al. Cutting edge: Cross-talk between cells of the innate immune system: NKT cells rapidly activate NK cells. *J Immunol* 1999; 163:4647-50.
334. Frucht DM, Fukao T, Bogdan C, Schindler H, O'Shea JJ, and Koyasu S. IFN-gamma production by antigen-presenting cells: mechanisms emerge. *Trends Immunol* 2001; 22:556-60.
335. Sharma M, Sharma S, Roy S, Varma S, and Bose M. Pulmonary epithelial cells are a source of interferon-gamma in response to *Mycobacterium tuberculosis* infection. *Immunol Cell Biol* 2007.
336. Billiau A. Interferon-gamma: biology and role in pathogenesis. *Adv Immunol* 1996; 62:61-130.
337. Leal IS, Smedegard B, Andersen P, and Appelberg R. Failure to induce enhanced protection against tuberculosis by increasing T-cell-dependent interferon-gamma generation. *Immunology* 2001; 104:157-61.
338. Collins HL and Kaufmann SHE. The many faces of host responses to tuberculosis. *Immunology* 2001; 103:1-9.
339. Cooper AM, Dalton DK, Stewart TA, Griffin JP, Russell DG, and Orme IM. Disseminated tuberculosis in interferon gamma gene-disrupted mice. *J Exp Med* 1993; 178:2243-7.
340. Flynn JL, Chan J, Triebold KJ, Dalton DK, Stewart TA, and Bloom BR. An essential role for interferon gamma in resistance to *Mycobacterium tuberculosis* infection. *J Exp Med* 1993; 178:2249-54.
341. Moreira AL, Tsenova L, Murray PJ et al. Aerosol infection of mice with recombinant BCG secreting murine IFN-gamma partially reconstitutes local protective immunity. *Microb Pathog* 2000; 29:175-85.
342. Altare F, Lammas D, Revy P et al. Inherited interleukin 12 deficiency in a child with bacille Calmette- Guerin and *Salmonella enteritidis* disseminated infection. *J Clin Invest* 1998; 102:2035-40.
343. Picard C, Fieschi C, Altare F et al. Inherited interleukin-12 deficiency: IL12B genotype and clinical phenotype of 13 patients from six kindreds. *Am J Hum Genet* 2002; 70:336-48.
344. Altare F, Durandy A, Lammas D et al. Impairment of mycobacterial immunity in human interleukin-12 receptor deficiency. *Science* 1998; 280:1432-5.

345. Altare F, Ensser A, Breiman A et al. Interleukin-12 receptor beta1 deficiency in a patient with abdominal tuberculosis. *J Infect Dis* 2001; 184:231-6.
346. Fieschi C, Dupuis S, Picard C, Smith CI, Holland SM, and Casanova JL. High levels of interferon gamma in the plasma of children with complete interferon gamma receptor deficiency. *Pediatrics* 2001; 107:E48.
347. Jouanguy E, Altare F, Lamhamedi S et al. Interferon-gamma-receptor deficiency in an infant with fatal bacille Calmette-Guerin infection. *N Engl J Med* 1996; 335:1956-61.
348. Jouanguy E, Lamhamedi-Cherradi S, Altare F et al. Partial interferon-gamma receptor 1 deficiency in a child with tuberculoid bacillus Calmette-Guerin infection and a sibling with clinical tuberculosis. *J Clin Invest* 1997; 100:2658-64.
349. Jouanguy E, Lamhamedi-Cherradi S, Lammas D et al. A human *IFNGR1* small deletion hotspot associated with dominant susceptibility to mycobacterial infection. *Nat Genet* 1999; 21:370-8.
350. Vogt G, Chagnier A, Yang K et al. Gains of glycosylation comprise an unexpectedly large group of pathogenic mutations. *Nat Genet* 2005; 37:692-700.
351. Dupuis S, Dargemont C, Fieschi C et al. Impairment of mycobacterial but not viral immunity by a germline human *STAT1* mutation. *Science* 2001; 293:300-3.
352. Filipe-Santos O, Bustamante J, Haverkamp MH et al. X-linked susceptibility to mycobacteria is caused by mutations in *NEMO* impairing CD40-dependent IL-12 production. *J Exp Med* 2006; 203:1745-59.
353. Newport M. The genetics of nontuberculous mycobacterial infection. *Expert Rev Mol Med* 2003; 2003:1-13.
354. Henao MI, Montes C, Paris SC, and Garcia LF. Cytokine gene polymorphisms in Colombian patients with different clinical presentations of tuberculosis. *Tuberculosis (Edinb)* 2006; 86:11-9.
355. Hirsch CS, Toossi Z, Othieno C et al. Depressed T-cell interferon-gamma responses in pulmonary tuberculosis: analysis of underlying mechanisms and modulation with therapy. *J Infect Dis* 1999; 180:2069-73.
356. Jo EK, Park JK, and Dockrell HM. Dynamics of cytokine generation in patients with active pulmonary tuberculosis. *Curr Opin Infect Dis* 2003; 16:205-10.
357. Ribeiro-Rodrigues R, Resende CT, Johnson JL et al. Sputum cytokine levels in patients with pulmonary tuberculosis as early markers of mycobacterial clearance. *Clin Diagn Lab Immunol* 2002; 9:818-23.
358. Tsao TC, Huang CC, Chiou WK, Yang PY, Hsieh MJ, and Tsao KC. Levels of interferon-gamma and interleukin-2 receptor-alpha for bronchoalveolar lavage fluid and serum were correlated with clinical grade and treatment of pulmonary tuberculosis. *Int J Tuberc Lung Dis* 2002; 6:720-7.
359. Ellner JJ, Hirsch CS, and Whalen CC. Correlates of protective immunity to *Mycobacterium tuberculosis* in humans. *Clin Infect Dis* 2000; 30 Suppl 3:S279-S282.
360. Park SK, Cho S, Lee IH et al. Subcutaneously administered interferon-gamma for the treatment of multidrug-resistant pulmonary tuberculosis. *Int J Infect Dis* 2007.
361. Ting LM, Kim AC, Cattamanchi A, and Ernst JD. *Mycobacterium tuberculosis* inhibits IFN-gamma transcriptional responses without inhibiting activation of STAT1. *J Immunol* 1999; 163:3898-906.

362. Sibley LD, Adams LB, and Krahenbuhl JL. Inhibition of interferon-gamma-mediated activation in mouse macrophages treated with lipoarabinomannan. *Clin Exp Immunol* 1990; 80:141-8.
363. Singhal A, Jaiswal A, Arora VK, and Prasad HK. Modulation of Interferon Gamma Receptor 1 by *Mycobacterium tuberculosis*: A potential immune evasive mechanism. *Infect Immun* 2007.
364. Holland SM. Immunotherapy of mycobacterial infections. *Semin Respir Infect* 2001; 16:47-59.
365. Schraufnagel DE. Tuberculosis treatment for the beginning of the next century. *Int J Tuberc Lung Dis* 1999; 3:651-62.
366. Toossi Z. Adjunctive immunotherapy of tuberculosis. *Cytokines Cell Mol Ther* 1998; 4:105-12.
367. Murray HW. Current and future clinical applications of interferon-gamma in host antimicrobial defense. *Intensive Care Med* 1996; 22 Suppl 4:S456-S461.
368. Condos R, Rom WN, and Schluger NW. Treatment of multidrug-resistant pulmonary tuberculosis with interferon-gamma via aerosol. *Lancet* 1997; 349:1513-5.
369. Koh WJ, Kwon OJ, Suh GY et al. Six-month therapy with aerosolized interferon-gamma for refractory multidrug-resistant pulmonary tuberculosis. *J Korean Med Sci* 2004; 19:167-71.
370. Hasan Z, Zaidi I, Jamil B, Khan MA, Kanji A, and Hussain R. Elevated ex vivo monocyte chemotactic protein-1 (CCL2) in pulmonary as compared with extra-pulmonary tuberculosis. *BMC Immunol* 2005; 6:14.
371. Allavena P, Bianchi G, Zhou D et al. Induction of natural killer cell migration by monocyte chemotactic protein-1, -2 and -3. *Eur J Immunol* 1994; 24:3233-6.
372. Luther SA and Cyster JG. Chemokines as regulators of T cell differentiation. *Nat Immunol* 2001; 2:102-7.
373. Kipnis A, Basaraba RJ, Orme IM, and Cooper AM. Role of chemokine ligand 2 in the protective response to early murine pulmonary tuberculosis. *Immunology* 2003; 109:547-51.
374. Mantovani A. The chemokine system: redundancy for robust outputs. *Immunol Today* 1999; 20:254-7.
375. Rovin BH, Lu L, and Saxena R. A novel polymorphism in the *MCP-1* gene regulatory region that influences MCP-1 expression. *Biochem Biophys Res Commun* 1999; 259:344-8.
376. Ueda A, Okuda K, Ohno S et al. NF-kappa B and Sp1 regulate transcription of the human monocyte chemoattractant protein-1 gene. *J Immunol* 1994; 153:2052-63.
377. Martin T, Cardarelli PM, Parry GC, Felts KA, and Cobb RR. Cytokine induction of monocyte chemoattractant protein-1 gene expression in human endothelial cells depends on the cooperative action of NF-kappa B and AP-1. *Eur J Immunol* 1997; 27:1091-7.
378. Valente AJ, Xie JF, Abramova MA, Wenzel UO, Abboud HE, and Graves DT. A complex element regulates IFN-gamma-stimulated monocyte chemoattractant protein-1 gene transcription. *J Immunol* 1998; 161:3719-28.
379. Ueda A, Ishigatsubo Y, Okubo T, and Yoshimura T. Transcriptional regulation of the human monocyte chemoattractant protein-1 gene. Cooperation of two NF-

- kappaB sites and NF-kappaB/Rel subunit specificity. *J Biol Chem* 1997; 272:31092-9.
380. Ping D, Jones PL, and Boss JM. TNF regulates the in vivo occupancy of both distal and proximal regulatory regions of the MCP-1/JE gene. *Immunity* 1996; 4:455-69.
 381. Tabara Y, Kohara K, Yamamoto Y et al. Polymorphism of the monocyte chemoattractant protein (*MCP-1*) gene is associated with the plasma level of MCP-1 but not with carotid intima-media thickness. *Hypertens Res* 2003; 26:677-83.
 382. Cho ML, Kim JY, Ko HJ et al. The *MCP-1* promoter -2518 polymorphism in Behcet's disease: correlation between allele types, MCP-1 production and clinical symptoms among Korean patients. *Autoimmunity* 2004; 37:77-80.
 383. Szalai C, Duba J, Prohaszka Z et al. Involvement of polymorphisms in the chemokine system in the susceptibility for coronary artery disease (CAD). Coincidence of elevated Lp(a) and *MCP-1* -2518 G/G genotype in CAD patients. *Atherosclerosis* 2001; 158:233-9.
 384. McDermott DH, Yang Q, Kathiresan S et al. *CCL2* polymorphisms are associated with serum monocyte chemoattractant protein-1 levels and myocardial infarction in the Framingham Heart Study. *Circulation* 2005; 112:1113-20.
 385. Gonzalez E, Rovin BH, Sen L et al. HIV-1 infection and AIDS dementia are influenced by a mutant MCP-1 allele linked to increased monocyte infiltration of tissues and MCP-1 levels. *Proc Natl Acad Sci U S A* 2002; 99:13795-800.
 386. Dawber TR and Kannel WB. The Framingham study. An epidemiological approach to coronary heart disease. *Circulation* 1966; 34:553-5.
 387. Kannel WB, Feinleib M, McNamara PM, Garrison RJ, and Castelli WP. An investigation of coronary heart disease in families. The Framingham offspring study. *Am J Epidemiol* 1979; 110:281-90.
 388. Rook GA, Hernandez-Pando R, Dheda K, and Teng SG. IL-4 in tuberculosis: implications for vaccine design. *Trends Immunol* 2004; 25:483-8.
 389. Kahnert A, Seiler P, Stein M et al. Alternative activation deprives macrophages of a coordinated defense program to *Mycobacterium tuberculosis*. *Eur J Immunol* 2006; 36:631-47.
 390. Buccheri S, Reljic R, Caccamo N et al. IL-4 depletion enhances host resistance and passive IgA protection against tuberculosis infection in BALB/c mice. *Eur J Immunol* 2007; 37:729-37.
 391. Hernandez-Pando R, Aguilar D, Hernandez ML, Orozco H, and Rook G. Pulmonary tuberculosis in BALB/c mice with non-functional *IL-4* genes: changes in the inflammatory effects of TNF-alpha and in the regulation of fibrosis. *Eur J Immunol* 2004; 34:174-83.
 392. Sugawara I, Yamada H, Mizuno S, and Iwakura Y. IL-4 is required for defense against mycobacterial infection. *Microbiol Immunol* 2000; 44:971-9.
 393. North RJ. Mice incapable of making IL-4 or IL-10 display normal resistance to infection with *Mycobacterium tuberculosis*. *Clin Exp Immunol* 1998; 113:55-8.
 394. Jung YJ, LaCourse R, Ryan L, and North RJ. Evidence inconsistent with a negative influence of T helper 2 cells on protection afforded by a dominant T helper 1 response against *Mycobacterium tuberculosis* lung infection in mice. *Infect Immun* 2002; 70:6436-43.

395. Erb KJ, Kirman J, Delahunt B, Chen W, and Le Gros G. IL-4, IL-5 and IL-10 are not required for the control of *M. bovis*-BCG infection in mice. *Immunol Cell Biol* 1998; 76:41-6.
396. Atamas SP, Choi J, Yurovsky VV, and White B. An alternative splice variant of human IL-4, IL-4 delta 2, inhibits IL-4-stimulated T cell proliferation. *J Immunol* 1996; 156:435-41.
397. Vasiliev AM, Vasilenko RN, Kulikova NL et al. Structural and functional properties of IL-4delta2, an alternative splice variant of human IL-4. *J Proteome Res* 2003; 2:273-81.
398. Demissie A, Abebe M, Aseffa A et al. Healthy individuals that control a latent infection with *Mycobacterium tuberculosis* express high levels of Th1 cytokines and the IL-4 antagonist IL-4delta2. *J Immunol* 2004; 172:6938-43.
399. Dheda K, Chang JS, Breen RA et al. In vivo and in vitro studies of a novel cytokine, interleukin 4delta2, in pulmonary tuberculosis. *Am J Respir Crit Care Med* 2005; 172:501-8.
400. Dheda K, Chang JS, Huggett JF et al. The stability of mRNA encoding IL-4 is increased in pulmonary tuberculosis, while stability of mRNA encoding the antagonistic splice variant, IL-4delta2, is not. *Tuberculosis (Edinb)* 2007; 87:237-41.
401. Luoni G, Verra F, Arca B et al. Antimalarial antibody levels and *IL4* polymorphism in the Fulani of West Africa. *Genes Immun* 2001; 2:411-4.
402. Rosenwasser LJ, Klemm DJ, Dresback JK et al. Promoter polymorphisms in the chromosome 5 gene cluster in asthma and atopy. *Clin Exp Allergy* 1995; 25 Suppl 2:74-8.
403. Bellamy, R. Genetic susceptibility and resistance to tuberculosis. PhD thesis. 1998. University of Oxford.
404. Basehore MJ, Howard TD, Lange LA et al. A comprehensive evaluation of *IL4* variants in ethnically diverse populations: association of total serum IgE levels and asthma in white subjects. *J Allergy Clin Immunol* 2004; 114:80-7.
405. Blackwell JM, Black GF, Peacock CS et al. Immunogenetics of leishmanial and mycobacterial infections: the Belem Family Study. *Philos Trans R Soc Lond B Biol Sci* 1997; 352:1331-45.
406. Blackwell JM. Genetics of host resistance and susceptibility to intramacrophage pathogens: a study of multicase families of tuberculosis, leprosy and leishmaniasis in north-eastern Brazil. *Int J Parasitol* 1998; 28:21-8.
407. Turner J, Gonzalez-Juarrero M, Ellis DL et al. In vivo IL-10 production reactivates chronic pulmonary tuberculosis in C57BL/6 mice. *J Immunol* 2002; 169:6343-51.
408. de la Barrera S, Aleman M, Musella R et al. IL-10 down-regulates costimulatory molecules on *Mycobacterium tuberculosis*-pulsed macrophages and impairs the lytic activity of CD4 and CD8 CTL in tuberculosis patients. *Clin Exp Immunol* 2004; 138:128-38.
409. Gesser B, Leffers H, Jinquan T et al. Identification of functional domains on human interleukin 10. *Proc Natl Acad Sci U S A* 1997; 94:14620-5.
410. Moore KW, de Waal MR, Coffman RL, and O'Garra A. Interleukin-10 and the interleukin-10 receptor. *Annu Rev Immunol* 2001; 19:683-765.

411. Sieling PA, Wang XH, Gately MK et al. IL-12 regulates T helper type 1 cytokine responses in human infectious disease. *J Immunol* 1994; 153:3639-47.
412. D'Andrea A, Aste-Amezaga M, Valiante NM, Ma X, Kubin M, and Trinchieri G. Interleukin 10 (IL-10) inhibits human lymphocyte interferon gamma-production by suppressing natural killer cell stimulatory factor/IL-12 synthesis in accessory cells. *J Exp Med* 1993; 178:1041-8.
413. Tripp CS, Wolf SF, and Unanue ER. Interleukin 12 and tumor necrosis factor alpha are costimulators of interferon gamma production by natural killer cells in severe combined immunodeficiency mice with listeriosis, and interleukin 10 is a physiologic antagonist. *Proc Natl Acad Sci U S A* 1993; 90:3725-9.
414. Sanchez FO, Rodriguez JI, Agudelo G, and Garcia LF. Immune responsiveness and lymphokine production in patients with tuberculosis and healthy controls. *Infect Immun* 1994; 62:5673-8.
415. Redpath S, Ghazal P, and Gascoigne NR. Hijacking and exploitation of IL-10 by intracellular pathogens. *Trends Microbiol* 2001; 9:86-92.
416. Bravo MJ, de Dios CJ, Alonso A, and Caballero A. Polymorphisms of the interferon gamma and interleukin 10 genes in human brucellosis. *Eur J Immunogenet* 2003; 30:433-5.
417. Kingo K, Koks S, Silm H, and Vasar E. *IL-10* promoter polymorphisms influence disease severity and course in psoriasis. *Genes Immun* 2003; 4:455-7.
418. Shin HD, Park BL, Kim YH, Cheong HS, Lee IH, and Park SK. Common interleukin 10 polymorphism associated with decreased risk of tuberculosis. *Exp Mol Med* 2005; 37:128-32.
419. Turner DM, Williams DM, Sankaran D, Lazarus M, Sinnott PJ, and Hutchinson IV. An investigation of polymorphism in the interleukin-10 gene promoter. *Eur J Immunogenet* 1997; 24:1-8.
420. Rosenwasser LJ and Borish L. Genetics of atopy and asthma: the rationale behind promoter-based candidate gene studies (*IL-4* and *IL-10*). *Am J Respir Crit Care Med* 1997; 156:S152-S155.
421. Shin HD, Winkler C, Stephens JC et al. Genetic restriction of HIV-1 pathogenesis to AIDS by promoter alleles of *IL10*. *Proc Natl Acad Sci U S A* 2000; 97:14467-72.
422. Edwards-Smith CJ, Jonsson JR, Purdie DM, Bansal A, Shorthouse C, and Powell EE. Interleukin-10 promoter polymorphism predicts initial response of chronic hepatitis C to interferon alfa. *Hepatology* 1999; 30:526-30.
423. Gibson AW, Edberg JC, Wu J, Westendorp RG, Huizinga TW, and Kimberly RP. Novel single nucleotide polymorphisms in the distal *IL-10* promoter affect IL-10 production and enhance the risk of systemic lupus erythematosus. *J Immunol* 2001; 166:3915-22.
424. Moraes MO, Pacheco AG, Schonkeren JJ et al. Interleukin-10 promoter single-nucleotide polymorphisms as markers for disease susceptibility and disease severity in leprosy. *Genes Immun* 2004; 5:592-5.
425. Rees LE, Wood NA, Gillespie KM, Lai KN, Gaston K, and Mathieson PW. The interleukin-10-1082 G/A polymorphism: allele frequency in different populations and functional significance. *Cell Mol Life Sci* 2002; 59:560-9.

426. Stein CM, Zalwango S, Chiunda AB et al. Linkage and association analysis of candidate genes for TB and TNF α cytokine expression: evidence for association with *IFNGR1*, *IL-10*, and TNF receptor 1 genes. *Hum Genet* 2007.
427. Meenagh A, Williams F, Ross OA et al. Frequency of cytokine polymorphisms in populations from western Europe, Africa, Asia, the Middle East and South America. *Hum Immunol* 2002; 63:1055-61.
428. Crawley E, Kay R, Sillibourne J, Patel P, Hutchinson I, and Woo P. Polymorphic haplotypes of the interleukin-10 5' flanking region determine variable interleukin-10 transcription and are associated with particular phenotypes of juvenile rheumatoid arthritis. *Arthritis Rheum* 1999; 42:1101-8.
429. Brombacher F, Kastelein RA, and Alber G. Novel IL-12 family members shed light on the orchestration of Th1 responses. *Trends Immunol* 2003; 24:207-12.
430. Wolf SF, Temple PA, Kobayashi M et al. Cloning of cDNA for natural killer cell stimulatory factor, a heterodimeric cytokine with multiple biologic effects on T and natural killer cells. *J Immunol* 1991; 146:3074-81.
431. Kobayashi M, Fitz L, Ryan M et al. Identification and purification of natural killer cell stimulatory factor (NKSF), a cytokine with multiple biologic effects on human lymphocytes. *J Exp Med* 1989; 170:827-45.
432. Oppmann B, Lesley R, Blom B et al. Novel p19 protein engages IL-12p40 to form a cytokine, IL-23, with biological activities similar as well as distinct from IL-12. *Immunity* 2000; 13:715-25.
433. Gillessen S, Carvajal D, Ling P et al. Mouse interleukin-12 (IL-12) p40 homodimer: a potent IL-12 antagonist. *Eur J Immunol* 1995; 25:200-6.
434. Ling P, Gately MK, Gubler U et al. Human IL-12 p40 homodimer binds to the IL-12 receptor but does not mediate biologic activity. *J Immunol* 1995; 154:116-27.
435. Presky DH, Yang H, Minetti LJ et al. A functional interleukin 12 receptor complex is composed of two beta-type cytokine receptor subunits. *Proc Natl Acad Sci U S A* 1996; 93:14002-7.
436. Chua AO, Chizzonite R, Desai BB et al. Expression cloning of a human IL-12 receptor component. A new member of the cytokine receptor superfamily with strong homology to gp130. *J Immunol* 1994; 153:128-36.
437. Chua AO, Wilkinson VL, Presky DH, and Gubler U. Cloning and characterization of a mouse IL-12 receptor-beta component. *J Immunol* 1995; 155:4286-94.
438. van Rietschoten JG, Smits HH, Westland R, Verweij CL, den Hartog MT, and Wierenga EA. Genomic organization of the human interleukin-12 receptor beta2-chain gene. *Immunogenetics* 2000; 51:30-6.
439. Parham C, Chirica M, Timans J et al. A receptor for the heterodimeric cytokine IL-23 is composed of IL-12Rbeta1 and a novel cytokine receptor subunit, IL-23R. *J Immunol* 2002; 168:5699-708.
440. Wang X, Wilkinson VL, Podlaski FJ et al. Characterization of mouse interleukin-12 p40 homodimer binding to the interleukin-12 receptor subunits. *Eur J Immunol* 1999; 29:2007-13.
441. Chackerian AA, Chen SJ, Brodie SJ et al. Neutralization or absence of the interleukin-23 pathway does not compromise immunity to mycobacterial infection. *Infect Immun* 2006; 74:6092-9.
442. Schluger NW and Rom WN. The host immune response to tuberculosis. *Am J Respir Crit Care Med* 1998; 157:679-91.

443. Gately MK, Renzetti LM, Magram J et al. The interleukin-12/interleukin-12-receptor system: role in normal and pathologic immune responses. *Annu Rev Immunol* 1998; 16:495-521.
444. Ma X, Reich RA, Gonzalez O et al. No evidence for association between the polymorphism in the 3' untranslated region of interleukin-12B and human susceptibility to tuberculosis. *Journal of Infectious Diseases* 2003; 188:1116-8.
445. Cooper AM, Magram J, Ferrante J, and Orme IM. Interleukin 12 (IL-12) is crucial to the development of protective immunity in mice intravenously infected with *Mycobacterium tuberculosis*. *J Exp Med* 1997; 186:39-45.
446. Holscher C, Atkinson RA, Arendse B et al. A protective and agonistic function of IL-12p40 in mycobacterial infection. *J Immunol* 2001; 167:6957-66.
447. Cooper AM, Kipnis A, Turner J, Magram J, Ferrante J, and Orme IM. Mice lacking bioactive IL-12 can generate protective, antigen-specific cellular responses to mycobacterial infection only if the IL-12 p40 subunit is present. *J Immunol* 2002; 168:1322-7.
448. Russell TD, Yan Q, Fan G et al. IL-12 p40 homodimer-dependent macrophage chemotaxis and respiratory viral inflammation are mediated through IL-12 receptor beta 1. *J Immunol* 2003; 171:6866-74.
449. Jana M, Dasgupta S, Saha RN, Liu X, and Pahan K. Induction of tumor necrosis factor-alpha (TNF-alpha) by interleukin-12 p40 monomer and homodimer in microglia and macrophages. *J Neurochem* 2003; 86:519-28.
450. Pahan K, Sheikh FG, Liu X, Hilger S, McKinney M, and Petro TM. Induction of nitric-oxide synthase and activation of NF-kappaB by interleukin-12 p40 in microglial cells. *J Biol Chem* 2001; 276:7899-905.
451. Gately MK, Carvajal DM, Connaughton SE et al. Interleukin-12 antagonist activity of mouse interleukin-12 p40 homodimer in vitro and in vivo. *Ann N Y Acad Sci* 1996; 795:1-12.
452. Khader SA, Partida-Sanchez S, Bell G et al. Interleukin 12p40 is required for dendritic cell migration and T cell priming after *Mycobacterium tuberculosis* infection. *J Exp Med* 2006; 203:1805-15.
453. Khader SA, Pearl JE, Sakamoto K et al. IL-23 compensates for the absence of IL-12p70 and is essential for the IL-17 response during tuberculosis but is dispensable for protection and antigen-specific IFN-gamma responses if IL-12p70 is available. *J Immunol* 2005; 175:788-95.
454. Remus N, El Baghdadi J, Fieschi C et al. Association of *IL12RB1* polymorphisms with pulmonary tuberculosis in adults in Morocco. *J Infect Dis* 2004; 190:580-7.
455. Akahoshi M, Nakashima H, Miyake K et al. Influence of interleukin-12 receptor beta1 polymorphisms on tuberculosis. *Hum Genet* 2003; 112:237-43.
456. Tso HW, Lau YL, Tam CM, Wong HS, and Chiang AK. Associations between *IL12B* Polymorphisms and Tuberculosis in the Hong Kong Chinese Population. *J Infect Dis* 2004; 190:913-9.
457. Sahiratmadja E, Baak-Pablo R, de Visser AW et al. Association of polymorphisms in IL-12/IFN-gamma pathway genes with susceptibility to pulmonary tuberculosis in Indonesia. *Tuberculosis (Edinb)* 2007.
458. Ohyama H, Ogata K, Takeuchi K et al. Polymorphism of the 5' flanking region of the IL-12 receptor beta2 gene partially determines the clinical types of leprosy through impaired transcriptional activity. *J Clin Pathol* 2005; 58:740-3.

459. Hall MA, McGlenn E, Coakley G et al. Genetic polymorphism of *IL-12p40* gene in immune-mediated disease. *Genes Immun* 2000; 1:219-24.
460. Huang D, Cancilla MR, and Morahan G. Complete primary structure, chromosomal localisation, and definition of polymorphisms of the gene encoding the human interleukin-12 p40 subunit. *Genes Immun* 2000; 1:515-20.
461. Morahan G, Kaur G, Singh M et al. Association of variants in the *IL12B* gene with leprosy and tuberculosis. *Tissue Antigens* 2007; 69 Suppl 1:234-6.
462. Morahan G, Huang D, Ymer SI et al. Linkage disequilibrium of a type 1 diabetes susceptibility locus with a regulatory *IL12B* allele. *Nat Genet* 2001; 27:218-21.
463. Mueller T, Mas-Marques A, Sarrazin C et al. Influence of interleukin 12B (*IL12B*) polymorphisms on spontaneous and treatment-induced recovery from hepatitis C virus infection. *J Hepatol* 2004; 41:652-8.
464. Noguchi E, Yokouchi Y, Shibasaki M et al. Identification of missense mutation in the *IL12B* gene: lack of association between *IL12B* polymorphisms and asthma and allergic rhinitis in the Japanese population. *Genes Immun* 2001; 2:401-3.
465. Orozco G, Sanchez E, Lopez-Nevot MA et al. Inducible nitric oxide synthase promoter polymorphism in human brucellosis. *Microbes Infect* 2003; 5:1165-9.
466. Feng CG, Jankovic D, Kullberg M et al. Maintenance of Pulmonary Th1 Effector Function in Chronic Tuberculosis Requires Persistent IL-12 Production. *J Immunol* 2005; 174:4185-92.
467. Morahan G, Huang D, Wu M et al. Association of *IL12B* promoter polymorphism with severity of atopic and non-atopic asthma in children. *Lancet* 2002; 360:455-9.
468. Davoodi-Semiromi A, Yang JJ, and She JX. IL-12p40 is associated with type 1 diabetes in Caucasian-American families. *Diabetes* 2002; 51:2334-6.
469. Morahan G, Boutlis CS, Huang D et al. A promoter polymorphism in the gene encoding interleukin-12 p40 (*IL12B*) is associated with mortality from cerebral malaria and with reduced nitric oxide production. *Genes Immun* 2002; 3:414-8.
470. Cargill M, Schrodi SJ, Chang M et al. A large-scale genetic association study confirms *IL12B* and leads to the identification of *IL23R* as psoriasis-risk genes. *Am J Hum Genet* 2007; 80:273-90.
471. Yin LM, Zhu WF, Wei L et al. Association of interleukin-12 p40 gene 3'-untranslated region polymorphism and outcome of HCV infection. *World J Gastroenterol* 2004; 10:2330-3.
472. Szabo SJ, Dighe AS, Gubler U, and Murphy KM. Regulation of the interleukin (IL)-12R beta 2 subunit expression in developing T helper 1 (Th1) and Th2 cells. *J Exp Med* 1997; 185:817-24.
473. Lee JS, Song CH, Kim CH et al. Profiles of IFN-gamma and its regulatory cytokines (IL-12, IL-18 and IL-10) in peripheral blood mononuclear cells from patients with multidrug-resistant tuberculosis. *Clin Exp Immunol* 2002; 128:516-24.
474. Dinarello CA, Novick D, Puren AJ et al. Overview of interleukin-18: more than an interferon-gamma inducing factor. *J Leukoc Biol* 1998; 63:658-64.
475. Vankayalapati R, Wizel B, Lakey DL et al. T cells enhance production of IL-18 by monocytes in response to an intracellular pathogen. *J Immunol* 2001; 166:6749-53.

476. Sugawara I, Yamada H, Kaneko H, Mizuno S, Takeda K, and Akira S. Role of interleukin-18 (IL-18) in mycobacterial infection in IL-18-gene-disrupted mice. *Infect Immun* 1999; 67:2585-9.
477. Verreck FA, De Boer T, Hoeve M et al. Human host defense and cytokines in mycobacterial infectious diseases: interleukin-18 cannot compensate for genetic defects in the interleukin-12 system. *Clin Infect Dis* 2002; 35:210-2.
478. Kawakami K, Kinjo Y, Uezu K et al. Interferon-gamma production and host protective response against *Mycobacterium tuberculosis* in mice lacking both IL-12p40 and IL-18. *Microbes Infect* 2004; 6:339-49.
479. Pechkovsky DV, Goldmann T, Vollmer E, Muller-Quernheim J, and Zissel G. Interleukin-18 expression by alveolar epithelial cells type II in tuberculosis and sarcoidosis. *FEMS Immunol Med Microbiol* 2006; 46:30-8.
480. Song CH, Lee JS, Nam HH et al. IL-18 production in human pulmonary and pleural tuberculosis. *Scand J Immunol* 2002; 56:611-8.
481. Vankayalapati R, Wizel B, Weis SE, Samten B, Girard WM, and Barnes PF. Production of interleukin-18 in human tuberculosis. *J Infect Dis* 2000; 182:234-9.
482. Okamura H, Kashiwamura S, Tsutsui H, Yoshimoto T, and Nakanishi K. Regulation of interferon-gamma production by IL-12 and IL-18. *Curr Opin Immunol* 1998; 10:259-64.
483. Netea MG, Azam T, Lewis EC et al. *Mycobacterium tuberculosis* induces interleukin-32 production through a caspase-1/IL-18/interferon-gamma-dependent mechanism. *PLoS Med* 2006; 3:1310-9.
484. Thompson SR and Humphries SE. Interleukin-18 genetics and inflammatory disease susceptibility. *Genes Immun* 2007; 8:91-9.
485. Arimitsu J, Hirano T, Higa S et al. *IL-18* gene polymorphisms affect IL-18 production capability by monocytes. *Biochem Biophys Res Commun* 2006; 342:1413-6.
486. Takada T, Suzuki E, Morohashi K, and Gejyo F. Association of single nucleotide polymorphisms in the *IL-18* gene with sarcoidosis in a Japanese population. *Tissue Antigens* 2002; 60:36-42.
487. Sebelova S, Izakovicova-Holla L, Stejskalova A, Schuller M, Znojil V, and Vasku A. Interleukin-18 and its three gene polymorphisms relating to allergic rhinitis. *J Hum Genet* 2007; 52:152-8.
488. Lee CC, Lin WY, Wan L et al. Interleukin-18 gene polymorphism, but not interleukin-2 gene polymorphism, is associated with rheumatoid arthritis. *Immunogenetics* 2007; 59:433-9.
489. Segat L, Bevilacqua D, Boniotto M et al. *IL-18* gene promoter polymorphism is involved in HIV-1 infection in a Brazilian pediatric population. *Immunogenetics* 2006; 58:471-3.
490. Janssen R, Grutters JC, Ruven HJ et al. No association between interleukin-18 gene polymorphisms and haplotypes in Dutch sarcoidosis patients. *Tissue Antigens* 2004; 63:578-83.
491. Zhou Y, Yamaguchi E, Hizawa N, and Nishimura M. Roles of functional polymorphisms in the interleukin-18 gene promoter in sarcoidosis. *Sarcoidosis Vasc Diffuse Lung Dis* 2005; 22:105-13.

492. Giedraitis V, He B, Huang WX, and Hillert J. Cloning and mutation analysis of the human *IL-18* promoter: a possible role of polymorphisms in expression regulation. *J Neuroimmunol* 2001; 112:146-52.
493. Huang HY, Thuita L, Strickland P, Hoffman SC, Comstock GW, and Helzlsouer KJ. Frequencies of single nucleotide polymorphisms in genes regulating inflammatory responses in a community-based population. *BMC Genet* 2007; 8:7.
494. Higa S, Hirano T, Mayumi M et al. Association between interleukin-18 gene polymorphism 105A/C and asthma. *Clin Exp Allergy* 2003; 33:1097-102.
495. Shin HD, Kim LH, Park BL et al. Association of interleukin 18 (IL18) polymorphisms with specific IgE levels to mite allergens among asthmatic patients. *Allergy* 2005; 60:900-6.
496. Imboden M, Nicod L, Nieters A et al. The common G-allele of interleukin-18 single-nucleotide polymorphism is a genetic risk factor for atopic asthma. The SAPALDIA Cohort Study. *Clin Exp Allergy* 2006; 36:211-8.
497. Heinzmann A, Gerhold K, Ganter K et al. Association study of polymorphisms within interleukin-18 in juvenile idiopathic arthritis and bronchial asthma. *Allergy* 2004; 59:845-9.
498. Tiret L, Godefroy T, Lubos E et al. Genetic analysis of the interleukin-18 system highlights the role of the interleukin-18 gene in cardiovascular disease. *Circulation* 2005; 112:643-50.
499. Kretowski A, Mironczuk K, Karpinska A et al. Interleukin-18 promoter polymorphisms in type 1 diabetes. *Diabetes* 2002; 51:3347-9.
500. Frucht DM and Holland SM. Defective monocyte costimulation for IFN-gamma production in familial disseminated *Mycobacterium avium* complex infection: abnormal IL-12 regulation. *J Immunol* 1996; 157:411-6.
501. Bustamante J, Picard C, Fieschi C et al. A novel X-linked recessive form of Mendelian susceptibility to mycobacterial disease. *J Med Genet* 2007; 44:e65.
502. Cooke GS, Campbell SJ, Sillah J et al. Polymorphism within the interferon-gamma/receptor complex is associated with pulmonary tuberculosis. *Am J Respir Crit Care Med* 2006; 174:339-43.
503. Pasquinelli V, Quiroga MF, Martinez GJ et al. Expression of signaling lymphocytic activation molecule-associated protein interrupts IFN-gamma production in human tuberculosis. *J Immunol* 2004; 172:1177-85.
504. Hron JD, Caplan L, Gerth AJ, Schwartzberg PL, and Peng SL. SH2D1A regulates T-dependent humoral autoimmunity. *J Exp Med* 2004; 200:261-6.
505. Okamoto S, Ji H, Howie D et al. Expression of the *SH2D1A* gene is regulated by a combination of transcriptional and post-transcriptional mechanisms. *Eur J Immunol* 2004; 34:3176-86.
506. Parolini O, Weinhausel A, Kagerbauer B et al. Differential methylation pattern of the X-linked lymphoproliferative (XLP) disease gene *SH2D1A* correlates with the cell lineage-specific transcription. *Immunogenetics* 2003; 55:116-21.
507. Suviolahti E, Oksanen LJ, Ohman M et al. The *SLC6A14* gene shows evidence of association with obesity. *J Clin Invest* 2003; 112:1762-72.
508. Benjafield AV, Wang WY, and Morris BJ. No association of angiotensin-converting enzyme 2 gene (*ACE2*) polymorphisms with essential hypertension. *Am J Hypertens* 2004; 17:624-8.

509. Roy SK, Wachira SJ, Weihua X, Hu J, and Kalvakolanu DV. CCAAT/enhancer-binding protein-beta regulates interferon-induced transcription through a novel element. *J Biol Chem* 2000; 275:12626-32.
510. Woltje M, Tschoke B, von B, V et al. CCAAT enhancer binding protein beta and hepatocyte nuclear factor 3beta are necessary and sufficient to mediate dexamethasone-induced up-regulation of alpha2HS-glycoprotein/fetuin-A gene expression. *J Mol Endocrinol* 2006; 36:261-77.
511. Zhang Y, Broser M, and Rom WN. Activation of the interleukin 6 gene by *Mycobacterium tuberculosis* or lipopolysaccharide is mediated by nuclear factors NF-IL6 and NF-kappa B. *Proc Natl Acad Sci U S A* 1994; 91:2225-9.
512. Bertilsson G, Heidrich J, Svensson K et al. Identification of a human nuclear receptor defines a new signaling pathway for CYP3A induction. *Proc Natl Acad Sci U S A* 1998; 95:12208-13.
513. Chen J and Raymond K. Roles of rifampicin in drug-drug interactions: underlying molecular mechanisms involving the nuclear pregnane X receptor. *Ann Clin Microbiol Antimicrob* 2006; 5:3.
514. Nishio Y, Isshiki H, Kishimoto T, and Akira S. A nuclear factor for interleukin-6 expression (NF-IL6) and the glucocorticoid receptor synergistically activate transcription of the rat alpha 1-acid glycoprotein gene via direct protein-protein interaction. *Mol Cell Biol* 1993; 13:1854-62.
515. Trinchieri G and Sher A. Cooperation of Toll-like receptor signals in innate immune defence. *Nat Rev Immunol* 2007; 7:179-90.
516. Akira S, Uematsu S, and Takeuchi O. Pathogen recognition and innate immunity. *Cell* 2006; 124:783-801.
517. Hashimoto C, Hudson KL, and Anderson KV. The Toll gene of *Drosophila*, required for dorsal-ventral embryonic polarity, appears to encode a transmembrane protein. *Cell* 1988; 52:269-79.
518. Medzhitov R, Preston-Hurlburt P, and Janeway CA, Jr. A human homologue of the *Drosophila* Toll protein signals activation of adaptive immunity. *Nature* 1997; 388:394-7.
519. Rock FL, Hardiman G, Timans JC, Kastelein RA, and Bazan JF. A family of human receptors structurally related to *Drosophila* Toll. *Proc Natl Acad Sci U S A* 1998; 95:588-93.
520. Quesniaux V, Fremond C, Jacobs M et al. Toll-like receptor pathways in the immune responses to mycobacteria. *Microbes Infect* 2004; 6:946-59.
521. Akira S and Takeda K. Toll-like receptor signalling. *Nat Rev Immunol* 2004; 4:499-511.
522. Ferwerda G, Girardin SE, Kullberg BJ et al. NOD2 and toll-like receptors are nonredundant recognition systems of *Mycobacterium tuberculosis*. *PLoS Pathog* 2005; 1:279-85.
523. Fremond CM, Yermeev V, Nicolle DM, Jacobs M, Quesniaux VF, and Ryffel B. Fatal *Mycobacterium tuberculosis* infection despite adaptive immune response in the absence of MyD88. *J Clin Invest* 2004; 114:1790-9.
524. Imanishi T, Hara H, Suzuki S, Suzuki N, Akira S, and Saito T. Cutting Edge: TLR2 Directly Triggers Th1 Effector Functions. *J Immunol* 2007; 178:6715-9.

525. Takeuchi O, Hoshino K, Kawai T et al. Differential roles of TLR2 and TLR4 in recognition of gram-negative and gram-positive bacterial cell wall components. *Immunity* 1999; 11:443-51.
526. Yoshimura A, Lien E, Ingalls RR, Tuomanen E, Dziarski R, and Golenbock D. Cutting edge: recognition of Gram-positive bacterial cell wall components by the innate immune system occurs via Toll-like receptor 2. *J Immunol* 1999; 163:1-5.
527. Brightbill HD, Libraty DH, Krutzik SR et al. Host defense mechanisms triggered by microbial lipoproteins through toll-like receptors. *Science* 1999; 285:732-6.
528. Werts C, Tapping RI, Mathison JC et al. Leptospiral lipopolysaccharide activates cells through a TLR2-dependent mechanism. *Nat Immunol* 2001; 2:346-52.
529. Schwandner R, Dziarski R, Wesche H, Rothe M, and Kirschning CJ. Peptidoglycan- and lipoteichoic acid-induced cell activation is mediated by toll-like receptor 2. *J Biol Chem* 1999; 274:17406-9.
530. Banaiee N, Kincaid EZ, Buchwald U, Jacobs WR, Jr., and Ernst JD. Potent inhibition of macrophage responses to IFN-gamma by live virulent *Mycobacterium tuberculosis* is independent of mature mycobacterial lipoproteins but dependent on TLR2. *J Immunol* 2006; 176:3019-27.
531. Lafuse WP, Alvarez GR, Curry HM, and Zwillig BS. *Mycobacterium tuberculosis* and *Mycobacterium avium* inhibit IFN-gamma-induced gene expression by TLR2-dependent and independent pathways. *J Interferon Cytokine Res* 2006; 26:548-61.
532. Pompei L, Jang S, Zamlynny B et al. Disparity in IL-12 release in dendritic cells and macrophages in response to *Mycobacterium tuberculosis* is due to use of distinct TLRs. *J Immunol* 2007; 178:5192-9.
533. Reiling N, Holscher C, Fehrenbach A et al. Cutting edge: Toll-like receptor (TLR)2- and TLR4-mediated pathogen recognition in resistance to airborne infection with *Mycobacterium tuberculosis*. *J Immunol* 2002; 169:3480-4.
534. Drennan MB, Nicolle D, Quesniaux VJ et al. Toll-like receptor 2-deficient mice succumb to *Mycobacterium tuberculosis* infection. *Am J Pathol* 2004; 164:49-57.
535. Heldwein KA, Liang MD, Andresen TK et al. TLR2 and TLR4 serve distinct roles in the host immune response against *Mycobacterium bovis* BCG. *J Leukoc Biol* 2003; 74:277-86.
536. Lorenz E, Mira JP, Cornish KL, Arbour NC, and Schwartz DA. A novel polymorphism in the toll-like receptor 2 gene and its potential association with staphylococcal infection. *Infect Immun* 2000; 68:6398-401.
537. Bochud PY, Hawn TR, and Aderem A. Cutting edge: a Toll-like receptor 2 polymorphism that is associated with lepromatous leprosy is unable to mediate mycobacterial signaling. *J Immunol* 2003; 170:3451-4.
538. Malhotra D, Relhan V, Reddy BS, and Bamezai R. *TLR2* Arg677Trp polymorphism in leprosy: revisited. *Hum Genet* 2005; 116:413-5.
539. Boraska JT, Barisic M, Drmic H, I et al. Microsatellite GT polymorphism in the toll-like receptor 2 is associated with colorectal cancer. *Clin Genet* 2006; 70:156-60.
540. Eder W, Klimecki W, Yu L et al. Toll-like receptor 2 as a major gene for asthma in children of European farmers. *J Allergy Clin Immunol* 2004; 113:482-8.
541. Poltorak A, He X, Smirnova I et al. Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in *Tlr4* gene. *Science* 1998; 282:2085-8.

542. Beutler B, Du X, and Poltorak A. Identification of Toll-like receptor 4 (Tlr4) as the sole conduit for LPS signal transduction: genetic and evolutionary studies. *J Endotoxin Res* 2001; 7:277-80.
543. Hoshino K, Takeuchi O, Kawai T et al. Cutting edge: Toll-like receptor 4 (TLR4)-deficient mice are hyporesponsive to lipopolysaccharide: evidence for TLR4 as the Lps gene product. *J Immunol* 1999; 162:3749-52.
544. Means TK, Wang S, Lien E, Yoshimura A, Golenbock DT, and Fenton MJ. Human toll-like receptors mediate cellular activation by *Mycobacterium tuberculosis*. *J Immunol* 1999; 163:3920-7.
545. Tsuji S, Matsumoto M, Takeuchi O et al. Maturation of human dendritic cells by cell wall skeleton of *Mycobacterium bovis* bacillus Calmette-Guerin: involvement of toll-like receptors. *Infect Immun* 2000; 68:6883-90.
546. Abel B, Thieblemont N, Quesniaux VJ et al. Toll-like receptor 4 expression is required to control chronic *Mycobacterium tuberculosis* infection in mice. *J Immunol* 2002; 169:3155-62.
547. Arbour NC, Lorenz E, Schutte BC et al. TLR4 mutations are associated with endotoxin hyporesponsiveness in humans. *Nat Genet* 2000; 25:187-91.
548. Lorenz E, Mira JP, Frees KL, and Schwartz DA. Relevance of mutations in the TLR4 receptor in patients with gram-negative septic shock. *Arch Intern Med* 2002; 162:1028-32.
549. Montes AH, Asensi V, Alvarez V et al. The Toll-like receptor 4 (Asp299Gly) polymorphism is a risk factor for Gram-negative and haematogenous osteomyelitis. *Clin Exp Immunol* 2006; 143:404-13.
550. Allen A, Obaro S, Bojang K et al. Variation in Toll-like receptor 4 and susceptibility to group A meningococcal meningitis in Gambian children. *Pediatr Infect Dis J* 2003; 22:1018-9.
551. Ryffel B, Jacobs M, Parida S, Botha T, Togbe D, and Quesniaux V. Toll-like receptors and control of mycobacterial infection in mice. *Novartis Found Symp* 2006; 279:127-39.
552. Flynn JL and Chan J. Immunology of tuberculosis. *Annu Rev Immunol* 2001; 19:93-129.
553. Blumenthal A, Ehlers S, Lauber J et al. The Wingless homolog WNT5A and its receptor Frizzled-5 regulate inflammatory responses of human mononuclear cells induced by microbial stimulation. *Blood* 2006; 108:965-73.
554. Vinson CR, Conover S, and Adler PN. A Drosophila tissue polarity locus encodes a protein containing seven potential transmembrane domains. *Nature* 1989; 338:263-4.
555. Masckauchan TN, Agalliu D, Vorontchikhina M et al. Wnt5a signaling induces proliferation and survival of endothelial cells in vitro and expression of MMP-1 and Tie-2. *Mol Biol Cell* 2006; 17:5163-72.
556. Lange C, Mix E, Rateitschak K, and Rolfs A. Wnt signal pathways and neural stem cell differentiation. *Neurodegener Dis* 2006; 3:76-86.
557. Li C, Hu L, Xiao J et al. Wnt5a regulates Shh and Fgf10 signaling during lung development. *Dev Biol* 2005; 287:86-97.
558. Fernandez-Cobo M, Zammarchi F, Mandeli J, Holland JF, and Pogo BG. Expression of Wnt5A and Wnt10B in non-immortalized breast cancer cells. *Oncol Rep* 2007; 17:903-7.

559. Co DO, Hogan LH, Il-Kim S, and Sandor M. T cell contributions to the different phases of granuloma formation. *Immunol Lett* 2004; 92:135-42.
560. American Thoracic Society. Statement on sarcoidosis. Joint Statement of the American Thoracic Society (ATS), the European Respiratory Society (ERS) and the World Association of Sarcoidosis and Other Granulomatous Disorders (WASOG) adopted by the ATS Board of Directors and by the ERS Executive Committee, February 1999. *Am J Respir Crit Care Med* 1999; 160:736-55.
561. Marsh S and McLeod HL. Crohn's disease: ethnic variation in *CARD15* genotypes. *Gut* 2003; 52:770.
562. Matsumoto T, Nakamura S, Jin-No Y et al. Role of granuloma in the immunopathogenesis of Crohn's disease. *Digestion* 2001; 63 Suppl 1:43-7.
563. Lyons DJ and Fielding JF. What's in a relationship? Sarcoidosis and tuberculosis. *Ir Med J* 1990; 83:76-9.
564. Litinsky I, Elkayam O, Flusser G, Segal R, Yaron M, and Caspi D. Sarcoidosis: TB or not TB? *Ann Rheum Dis* 2002; 61:385-6.
565. Dubaniewicz A, Kampfer S, and Singh M. Serum anti-mycobacterial heat shock proteins antibodies in sarcoidosis and tuberculosis. *Tuberculosis (Edinb)* 2006; 86:60-7.
566. Rossman MD, Thompson B, Frederick M et al. *HLA-DRB1*1101*: a significant risk factor for sarcoidosis in blacks and whites. *Am J Hum Genet* 2003; 73:720-35.
567. Schreiber S, Rosenstiel P, Albrecht M, Hampe J, and Krawczak M. Genetics of Crohn disease, an archetypal inflammatory barrier disease. *Nat Rev Genet* 2005; 6:376-88.
568. Ott SJ, Musfeldt M, Wenderoth DF et al. Reduction in diversity of the colonic mucosa associated bacterial microflora in patients with active inflammatory bowel disease. *Gut* 2004; 53:685-93.
569. Cohen RD. Mycobacterium in Crohn's: something to ruminate about? *Gastroenterology* 2005; 128:2167-8.
570. Naser SA. Mycobacterium in Crohn's disease is hard to digest. *Gastroenterology* 2005; 129:1359-60.
571. Chamberlin WM and Naser SA. Integrating theories of the etiology of Crohn's disease. On the etiology of Crohn's disease: questioning the hypotheses. *Med Sci Monit* 2006; 12:RA27-RA33.
572. Stammers M, Rowen L, Rhodes D, Trowsdale J, and Beck S. *BTL-II*: a polymorphic locus with homology to the butyrophilin gene family, located at the border of the major histocompatibility complex class II and class III regions in human and mouse. *Immunogenetics* 2000; 51:373-82.
573. Yadav D and Khuller G. Evaluation of the T cells and costimulatory molecules in the protective efficacy of 30 kDa secretory protein against experimental tuberculosis. *Immunol Cell Biol* 2001; 79:207-12.
574. Nguyen T, Liu XK, Zhang Y, and Dong C. *BTNL2*, a butyrophilin-like molecule that functions to inhibit T cell activation. *J Immunol* 2006; 176:7354-60.
575. Rybicki BA, Walewski JL, Maliarik MJ, Kian H, and Iannuzzi MC. The *BTNL2* gene and sarcoidosis susceptibility in African Americans and Whites. *Am J Hum Genet* 2005; 77:491-9.

576. Traherne JA, Barcellos LF, Sawcer SJ et al. Association of the truncating splice site mutation in *BTNL2* with multiple sclerosis is secondary to *HLA-DRB1*15*. *Hum Mol Genet* 2006; 15:155-61.
577. Simmonds MJ, Heward JM, Barrett JC, Franklyn JA, and Gough SC. Association of the *BTNL2* rs2076530 single nucleotide polymorphism with Graves' disease appears to be secondary to DRB1 exon 2 position beta74. *Clin Endocrinol (Oxf)* 2006; 65:429-32.
578. Smith MW, Patterson N, Lautenberger JA et al. A high-density admixture map for disease gene discovery in African Americans. *Am J Hum Genet* 2004; 74:1001-13.
579. Hugot JP, Chamaillard M, Zouali H et al. Association of *NOD2* leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature* 2001; 411:599-603.
580. Ogura Y, Bonen DK, Inohara N et al. A frameshift mutation in *NOD2* associated with susceptibility to Crohn's disease. *Nature* 2001; 411:603-6.
581. Girardin SE, Boneca IG, Viala J et al. Nod2 is a general sensor of peptidoglycan through muramyl dipeptide (MDP) detection. *J Biol Chem* 2003; 278:8869-72.
582. Yamada H, Mizuno S, Reza-Gholizadeh M, and Sugawara I. Relative importance of NF-kappaB p50 in mycobacterial infection. *Infect Immun* 2001; 69:7100-5.
583. Bonen DK, Ogura Y, Nicolae DL et al. Crohn's disease-associated *NOD2* variants share a signaling defect in response to lipopolysaccharide and peptidoglycan. *Gastroenterology* 2003; 124:140-6.
584. Zaahl MG, Winter T, Warnich L, and Kotze MJ. Analysis of the three common mutations in the *CARD15* gene (R702W, G908R and 1007fs) in South African colored patients with inflammatory bowel disease. *Mol Cell Probes* 2005; 19:278-81.
585. Inoue N, Tamura K, Kinouchi Y et al. Lack of common *NOD2* variants in Japanese patients with Crohn's disease. *Gastroenterology* 2002; 123:86-91.
586. Wright JP, Froggatt J, O'Keefe EA et al. The epidemiology of inflammatory bowel disease in Cape Town 1980-1984. *S Afr Med J* 1986; 70:10-5.
587. Reddy SI and Burakoff R. Inflammatory bowel disease in African Americans. *Inflamm Bowel Dis* 2003; 9:380-5.
588. Reich DE and Lander ES. On the allelic spectrum of human disease. *Trends Genet* 2001; 17:502-10.
589. Peri S, Navarro JD, Amanchy R et al. Development of human protein reference database as an initial platform for approaching systems biology in humans. *Genome Res* 2003; 13:2363-71.
590. Mishra GR, Suresh M, Kumaran K et al. Human protein reference database--2006 update. *Nucleic Acids Res* 2006; 34:D411-D414.
591. Clark HF, Gurney AL, Abaya E et al. The secreted protein discovery initiative (SPDI), a large-scale effort to identify novel human secreted and transmembrane proteins: a bioinformatics assessment. *Genome Res* 2003; 13:2265-70.
592. Alexander WS. Suppressors of cytokine signalling (SOCS) in the immune system. *Nat Rev Immunol* 2002; 2:410-6.
593. Alexander WS and Hilton DJ. The role of suppressors of cytokine signaling (SOCS) proteins in regulation of the immune response. *Annu Rev Immunol* 2004; 22:503-29.

594. Kile BT and Alexander WS. The suppressors of cytokine signalling (SOCS). *Cell Mol Life Sci* 2001; 58:1627-35.
595. Baetz A, Frey M, Heeg K, and Dalpke AH. Suppressor of cytokine signaling (SOCS) proteins indirectly regulate toll-like receptor signaling in innate immune cells. *J Biol Chem* 2004; 279:54708-15.
596. Krebs DL and Hilton DJ. SOCS proteins: negative regulators of cytokine signaling. *Stem Cells* 2001; 19:378-87.
597. Imai K, Kurita-Ochiai T, and Ochiai K. *Mycobacterium bovis* bacillus Calmette-Guerin infection promotes SOCS induction and inhibits IFN-gamma-stimulated JAK/STAT signaling in J774 macrophages. *FEMS Immunol Med Microbiol* 2003; 39:173-80.
598. Stoiber D, Kovarik P, Cohny S, Johnston JA, Steinlein P, and Decker T. Lipopolysaccharide induces in macrophages the synthesis of the suppressor of cytokine signaling 3 and suppresses signal transduction in response to the activating factor IFN-gamma. *J Immunol* 1999; 163:2640-7.
599. Dalpke AH, Oppen S, Zimmermann S, and Heeg K. Suppressors of cytokine signaling (SOCS)-1 and SOCS-3 are induced by CpG-DNA and modulate cytokine responses in APCs. *J Immunol* 2001; 166:7082-9.
600. Gatto L, Berlatto C, Poli V et al. Analysis of *SOCS-3* promoter responses to interferon gamma. *J Biol Chem* 2004; 279:13746-54.
601. Kinjyo I, Inoue H, Hamano S et al. Loss of SOCS3 in T helper cells resulted in reduced immune responses and hyperproduction of interleukin 10 and transforming growth factor-beta 1. *J Exp Med* 2006; 203:1021-31.
602. Bertholet S, Dickensheets HL, Sheikh F, Gam AA, Donnelly RP, and Kenney RT. *Leishmania donovani*-induced expression of suppressor of cytokine signaling 3 in human macrophages: a novel mechanism for intracellular parasite suppression of activation. *Infect Immun* 2003; 71:2095-101.
603. Stoiber D, Stockinger S, Steinlein P, Kovarik J, and Decker T. *Listeria monocytogenes* modulates macrophage cytokine responses through STAT serine phosphorylation and the induction of suppressor of cytokine signaling 3. *J Immunol* 2001; 166:466-72.
604. Uchiya K and Nikai T. *Salmonella* pathogenicity island 2-dependent expression of suppressor of cytokine signaling 3 in macrophages. *Infect Immun* 2005; 73:5587-94.
605. Ekelund E, Saaf A, Tengvall-Linder M et al. Elevated expression and genetic association links the *SOCS3* gene to atopic dermatitis. *Am J Hum Genet* 2006; 78:1060-5.
606. Jamshidi Y, Snieder H, Wang X, Spector TD, Carter ND, and O'Dell SD. Common polymorphisms in *SOCS3* are not associated with body weight, insulin sensitivity or lipid profile in normal female twins. *Diabetologia* 2006; 49:306-10.
607. Gylvin T, Nolsoe R, Hansen T et al. Mutation analysis of suppressor of cytokine signalling 3, a candidate gene in Type 1 diabetes and insulin sensitivity. *Diabetologia* 2004; 47:1273-7.
608. Nagabhushanam V, Solache A, Ting LM, Escaron CJ, Zhang JY, and Ernst JD. Innate inhibition of adaptive immunity: *Mycobacterium tuberculosis*-induced IL-6 inhibits macrophage responses to IFN-gamma. *J Immunol* 2003; 171:4750-7.

609. Blumenthal A, Lauber J, Hoffmann R et al. Common and unique gene expression signatures of human macrophages in response to four strains of *Mycobacterium avium* that differ in their growth and persistence characteristics. *Infect Immun* 2005; 73:3330-41.
610. Greenwell-Wild T, Vazquez N, Sim D et al. *Mycobacterium avium* infection and modulation of human macrophage gene expression. *J Immunol* 2002; 169:6286-97.
611. Manca C, Tsenova L, Freeman S et al. 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* 2005; 25:694-701.
612. Santamaria I, Velasco G, Pendas AM, Fueyo A, and Lopez-Otin C. Cathepsin Z, a novel human cysteine proteinase with a short propeptide domain and a unique chromosomal location. *J Biol Chem* 1998; 273:16816-23.
613. Pungercar J and Ivanovski G. Identification and molecular cloning of cathepsin P, a novel human putative cysteine protease of the papain family. *Pflugers Arch* 2000; 439:R116-R118.
614. Nagler DK and Menard R. Human cathepsin X: a novel cysteine protease of the papain family with a very short proregion and unique insertions. *FEBS Lett* 1998; 434:135-9.
615. Klemencic I, Carmona AK, Cezari MH et al. Biochemical characterization of human cathepsin X revealed that the enzyme is an exopeptidase, acting as carboxymonopeptidase or carboxydipeptidase. *Eur J Biochem* 2000; 267:5404-12.
616. Zavasnik-Bergant T and Turk B. Cysteine cathepsins in the immune response. *Tissue Antigens* 2006; 67:349-55.
617. Journet A, Chapel A, Kieffer S, Louwagie M, Luche S, and Garin J. Towards a human repertoire of monocytic lysosomal proteins. *Electrophoresis* 2000; 21:3411-9.
618. Garin J, Diez R, Kieffer S et al. The phagosome proteome: insight into phagosome functions. *J Cell Biol* 2001; 152:165-80.
619. Kos J, Sekirnik A, Premzl A et al. Carboxypeptidases cathepsins X and B display distinct protein profile in human cells and tissues. *Exp Cell Res* 2005; 306:103-13.
620. Falcone FH, Tetteh KK, Hunt P, Blaxter ML, Loukas A, and Maizels RM. The new subfamily of cathepsin-Z-like protease genes includes Tc-cpz-1, a cysteine protease gene expressed in *Toxocara canis* adults and infective stage larvae. *Exp Parasitol* 2000; 94:201-7.
621. Nagler DK, Lechner AM, Oettl A et al. An enzyme-linked immunosorbent assay for human cathepsin X, a potential new inflammatory marker. *J Immunol Methods* 2006; 308:241-50.
622. Nepal RM, Mampe S, Shaffer B, Erickson AH, and Bryant P. Cathepsin L maturation and activity is impaired in macrophages harboring *M. avium* and *M. tuberculosis*. *Int Immunol* 2006; 18:931-9.
623. Stewart JN, Rivera HN, Karls R, Quinn FD, Roman J, and Rivera-Marrero CA. Increased pathology in lungs of mice after infection with an alpha-crystallin mutant of *Mycobacterium tuberculosis*: changes in cathepsin proteases and certain cytokines. *Microbiology* 2006; 152:233-44.

624. Kim DK, Park GM, Hwang YI et al. Microarray analysis of gene expression associated with extrapulmonary dissemination of tuberculosis. *Respirology* 2006; 11:557-65.
625. Lee J, Remold HG, Jeong MH, and Kornfeld H. Macrophage apoptosis in response to high intracellular burden of *Mycobacterium tuberculosis* is mediated by a novel caspase-independent pathway. *J Immunol* 2006; 176:4267-74.
626. O'Sullivan MP, O'Leary S, Kelly DM, and Keane J. A caspase-independent pathway mediates macrophage cell death in response to *Mycobacterium tuberculosis* infection. *Infect Immun* 2007; 75:1984-93.
627. Smyth DJ, Howson JM, Payne F et al. Analysis of polymorphisms in 16 genes in type 1 diabetes that have been associated with other immune-mediated diseases. *BMC Med Genet* 2006; 7:20.
628. Saunders BM and Britton WJ. Life and death in the granuloma: immunopathology of tuberculosis. *Immunol Cell Biol* 2007; 85:103-11.
629. Oxford University Press. "inflammation". *The Oxford Companion to the Body*. Answers.com.<http://www.answers.com/topic/inflammation> . 2003.
630. Gilbertson JR, Ho J, Anthony L, Jukic DM, Yagi Y, and Parwani AV. Primary histologic diagnosis using automated whole slide imaging: a validation study. *BMC Clin Pathol* 2006; 6:4.
631. Kroeger KM, Carville KS, and Abraham LJ. The -308 tumor necrosis factor-alpha promoter polymorphism effects transcription. *Mol Immunol* 1997; 34:391-9.
632. Wilson AG, Symons JA, McDowell TL, McDevitt HO, and Duff GW. Effects of a polymorphism in the human tumor necrosis factor alpha promoter on transcriptional activation. *Proc Natl Acad Sci U S A* 1997; 94:3195-9.
633. Lee YH, Harley JB, and Nath SK. Meta-analysis of *TNF-alpha* promoter -308 A/G polymorphism and SLE susceptibility. *Eur J Hum Genet* 2006; 14:364-71.
634. Danis VA, Millington M, Hyland V, Lawford R, Huang Q, and Grennan D. Increased frequency of the uncommon allele of a tumour necrosis factor alpha gene polymorphism in rheumatoid arthritis and systemic lupus erythematosus. *Dis Markers* 1995; 12:127-33.
635. Gonzalez S, Rodrigo L, Martinez-Borra J et al. *TNF-alpha* -308A promoter polymorphism is associated with enhanced TNF-alpha production and inflammatory activity in Crohn's patients with fistulizing disease. *Am J Gastroenterol* 2003; 98:1101-6.
636. Keane J, Gershon S, Wise RP et al. Tuberculosis associated with infliximab, a tumor necrosis factor alpha-neutralizing agent. *N Engl J Med* 2001; 345:1098-104.
637. Schwab M, Schaeffeler E, Marx C et al. Association between the C3435T *MDR1* gene polymorphism and susceptibility for ulcerative colitis. *Gastroenterology* 2003; 124:26-33.
638. Brant SR, Panhuysen CI, Nicolae D et al. *MDR1* Ala893 polymorphism is associated with inflammatory bowel disease. *Am J Hum Genet* 2003; 73:1282-92.
639. Prescott NJ, Fisher SA, Franke A et al. A nonsynonymous SNP in *ATG16L1* predisposes to ileal Crohn's disease and is independent of *CARD15* and *IBD5*. *Gastroenterology* 2007; 132:1665-71.
640. King AL, Moodie SJ, Fraser JS et al. Coeliac disease: investigation of proposed causal variants in the *CTLA4* gene region. *Eur J Immunogenet* 2003; 30:427-32.

641. Barreto M, Santos E, Ferreira R et al. Evidence for *CTLA4* as a susceptibility gene for systemic lupus erythematosus. *Eur J Hum Genet* 2004; 12:620-6.
642. Koizumi K, Okamoto H, Iikuni N et al. Single nucleotide polymorphisms in the gene encoding the major histocompatibility complex class II transactivator (CIITA) in systemic lupus erythematosus. *Ann Rheum Dis* 2005; 64:947-50.
643. Martinez A, Sanchez-Lopez M, Varade J et al. Role of the *MHC2TA* gene in autoimmune diseases. *Ann Rheum Dis* 2007; 66:325-9.
644. Swanberg M, Lidman O, Padyukov L et al. *MHC2TA* is associated with differential MHC molecule expression and susceptibility to rheumatoid arthritis, multiple sclerosis and myocardial infarction. *Nat Genet* 2005; 37:486-94.
645. Kochi Y, Yamada R, Suzuki A et al. A functional variant in *FCRL3*, encoding Fc receptor-like 3, is associated with rheumatoid arthritis and several autoimmunities. *Nat Genet* 2005; 37:478-85.
646. Thabet MM, Wesoly J, Slagboom PE, Toes RE, and Huizinga TW. *FCRL3* promoter 169 CC homozygosity is associated with susceptibility to rheumatoid arthritis in Dutch Caucasians. *Ann Rheum Dis* 2007; 66:803-6.
647. Kaijzel EL, van Dongen H, Bakker AM, Breedveld FC, Huizinga TW, and Verweij CL. Relationship of polymorphisms of the Interleukin-1 gene cluster to occurrence and severity of rheumatoid arthritis. *Tissue Antigens* 2002; 59:122-6.
648. Timms AE, Crane AM, Sims AM et al. The interleukin 1 gene cluster contains a major susceptibility locus for ankylosing spondylitis. *Am J Hum Genet* 2004; 75:587-95.
649. Suzuki A, Ji G, Numabe Y et al. Single nucleotide polymorphisms associated with aggressive periodontitis and severe chronic periodontitis in Japanese. *Biochem Biophys Res Commun* 2004; 317:887-92.
650. Lyon HN, Emilsson V, Hinney A et al. The association of a SNP upstream of *INSIG2* with body mass index is reproduced in several but not all cohorts. *PLoS Genet* 2007; 3:e61.
651. Shirakawa T, Mao XQ, Sasaki S et al. Association between atopic asthma and a coding variant of Fc epsilon RI beta in a Japanese population. *Hum Mol Genet* 1996; 5:1129-30.
652. Nishiyama C, Akizawa Y, Nishiyama M et al. Polymorphisms in the Fc epsilon RI beta promoter region affecting transcription activity: a possible promoter-dependent mechanism for association between Fc epsilon RI beta and atopy. *J Immunol* 2004; 173:6458-64.
653. Suzuki A, Yamada R, Chang X et al. Functional haplotypes of *PADI4*, encoding citrullinating enzyme peptidylarginine deiminase 4, are associated with rheumatoid arthritis. *Nat Genet* 2003; 34:395-402.
654. Sugawara K, Olson TS, Moskaluk CA et al. Linkage to peroxisome proliferator-activated receptor-gamma in SAMP1/YitFc mice and in human Crohn's disease. *Gastroenterology* 2005; 128:351-60.
655. Libioulle C, Louis E, Hansoul S et al. Novel Crohn disease locus identified by genome-wide association maps to a gene desert on 5p13.1 and modulates expression of *PTGER4*. *PLoS Genet* 2007; 3:e58.
656. Wesoly J, Hu X, Thabet MM et al. The 620W allele is the *PTPN22* genetic variant conferring susceptibility to RA in a Dutch population. *Rheumatology (Oxford)* 2007; 46:617-21.

657. Chelala C, Duchatelet S, Joffret ML et al. *PTPN22* R620W functional variant in type 1 diabetes and autoimmunity related traits. *Diabetes* 2007; 56:522-6.
658. Vang T, Congia M, Macis MD et al. Autoimmune-associated lymphoid tyrosine phosphatase is a gain-of-function variant. *Nat Genet* 2005; 37:1317-9.
659. Gomez LM, Anaya JM, and Martin J. Genetic influence of *PTPN22* R620W polymorphism in tuberculosis. *Hum Immunol* 2005; 66:1242-7.
660. Bottini N, Musumeci L, Alonso A et al. A functional variant of lymphoid tyrosine phosphatase is associated with type I diabetes. *Nat Genet* 2004; 36:337-8.
661. Tokuhira S, Yamada R, Chang X et al. An intronic SNP in a *RUNX1* binding site of *SLC22A4*, encoding an organic cation transporter, is associated with rheumatoid arthritis. *Nat Genet* 2003; 35:341-8.
662. Fisher SA, Hampe J, Onnie CM et al. Direct or indirect association in a complex disease: the role of *SLC22A4* and *SLC22A5* functional variants in Crohn disease. *Hum Mutat* 2006; 27:778-85.
663. Peltekova VD, Wintle RF, Rubin LA et al. Functional variants of *OCTN* cation transporter genes are associated with Crohn disease. *Nat Genet* 2004; 36:471-5.
664. Reich K, Huffmeier U, Konig IR et al. *TNF* polymorphisms in psoriasis: Association of psoriatic arthritis with the promoter polymorphism *TNF*-857* independent of the *PSORS1* risk allele. *Arthritis Rheum* 2007; 56:2056-64.
665. Lv K, Chen R, Cai Q, Fang M, and Sun S. Effects of a single nucleotide polymorphism on the expression of human tumor necrosis factor-alpha. *Scand J Immunol* 2006; 64:164-9.
666. Yang SK, Lee SG, Cho YK, Lim J, Lee I, and Song K. Association of *TNF-alpha/LTA* polymorphisms with Crohn's disease in Koreans. *Cytokine* 2006; 35:13-20.
667. Waschke KA, Villani AC, Vermeire S et al. Tumor necrosis factor receptor gene polymorphisms in Crohn's disease: association with clinical phenotypes. *Am J Gastroenterol* 2005; 100:1126-33.
668. Yamazaki K, McGovern D, Ragoussis J et al. Single nucleotide polymorphisms in *TNFSF15* confer susceptibility to Crohn's disease. *Hum Mol Genet* 2005; 14:3499-506.
669. Brassard P, Kezouh A, and Suissa S. Antirheumatic drugs and the risk of tuberculosis. *Clin Infect Dis* 2006; 43:717-22.
670. Repique CJ, Li A, Brickey WJ, Ting JP, Collins FM, and Morris SL. Susceptibility of mice deficient in the MHC class II transactivator to infection with *Mycobacterium tuberculosis*. *Scand J Immunol* 2003; 58:15-22.
671. Roumier C, Fenaux P, Lafage M, Imbert M, Eclache V, and Preudhomme C. New mechanisms of *AML1* gene alteration in hematological malignancies. *Leukemia* 2003; 17:9-16.
672. Peschon JJ, Torrance DS, Stocking KL et al. TNF receptor-deficient mice reveal divergent roles for p55 and p75 in several models of inflammation. *J Immunol* 1998; 160:943-52.
673. Jacobs M, Brown N, Allie N, Chetty K, and Ryffel B. Tumor necrosis factor receptor 2 plays a minor role for mycobacterial immunity. *Pathobiology* 2000; 68:68-75.

674. Balcewicz-Sablinska MK, Keane J, Kornfeld H, and Remold HG. Pathogenic *Mycobacterium tuberculosis* evades apoptosis of host macrophages by release of TNF-R2, resulting in inactivation of TNF-alpha. *J Immunol* 1998; 161:2636-41.
675. Wilson JF and Goldstein DB. Consistent long-range linkage disequilibrium generated by admixture in a Bantu-Semitic hybrid population. *Am J Hum Genet* 2000; 67:926-35.
676. Mattick JS and Makunin IV. Non-coding RNA. *Hum Mol Genet* 2006; 15 Spec No 1:R17-R29.
677. Birney E, Stamatoyannopoulos JA, Dutta A et al. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* 2007; 447:799-816.
678. Richards EJ. Inherited epigenetic variation--revisiting soft inheritance. *Nat Rev Genet* 2006; 7:395-401.
679. Pennini ME, Pai RK, Schultz DC, Boom WH, and Harding CV. *Mycobacterium tuberculosis* 19-kDa lipoprotein inhibits IFN-gamma-induced chromatin remodeling of MHC2TA by TLR2 and MAPK signaling. *J Immunol* 2006; 176:4323-30.
680. den Boon S, Van Lill SWP, Borgdorff MW et al. Association between smoking and tuberculosis infection: a population survey in a high tuberculosis incidence area. *Thorax* 2005; 60:555-7.
681. Kennedy GC, Matsuzaki H, Dong S et al. Large-scale genotyping of complex DNA. *Nat Biotechnol* 2003; 21:1233-7.
682. Affymetrix Genome-Wide Human SNP Array 6.0. 2007. Affymetrix.
683. Smith MW and O'Brien SJ. Mapping by admixture linkage disequilibrium: advances, limitations and guidelines. *Nat Rev Genet* 2005; 6:623-32.
684. Dubos R and Dubos J. *The White Plague: Tuberculosis, Man and Society*. Boston: Little, Brown & Co., 1952.

ADDENDUM

A.1. BUFFERS AND SOLUTIONS

20X Sodium boride buffer (SB buffer):	19.1 g di-sodium tetraborate decahydrate 500 ml dH ₂ O
EDTA:	93.06 g EDTA.2H ₂ O Approximately 10 g NaOH pellets Add 400 ml distilled water Adjust the pH to 8.0 using the NaOH pellets. Make up to 500 ml Autoclave and store at room temperature
Luria broth (LB):	10 g Bacto-tryptone 10 g NaCl 5 g Bacto yeast Add 1 liter of water Autoclave
LB/ampicillin/X-Gal plates:	1.9 g Bacto-agar added to 1 l of autoclaved LB 400 µl 20mg/ml X-Gal 250 µl 100mM ampicillin Plates were poured and stored at 4 degrees
Super Optimal Broth (SOB):	5 g tryptone 1.25 g yeast extract 0.125 g NaCl Make up to 250 ml Autoclave 2.5 ml 0.25M KCl 1.25 ml 2M MgSO ₄ 25 µl 5M NaOH
SOB with added glucose (SOC):	12.5 ml SOB 200 µl glucose 100 µl 1mM MgCl ₂
TE:	1.21 g Tris-HCl 0.372 g EDTA Add approximately 800ml distilled water and mix. Adjust the pH to 8.0 with concentrated HCl Make up to a final volume of 1 liter Autoclave

A.2. REAGENTS

Reagent	Supplier	Supplier
100 bp DNA ladder	Invitrogen Germany
100 bp DNA ladder	Whitehead Scientific South Africa
3130 Performance Optimized Polymer 7	Applied Biosystems South Africa
3170 Performance Optimized Polymer 7	Applied Biosystems Germany
Agarose	Eurogentec Germany
Agarose D-1 LE	Whitehead Scientific South Africa
AmpliTaq Gold DNA polymerase	Applied Biosystems Germany
Bacilol	Bode Chemie Germany
BigDye Terminator v1.1 Cycle Sequencing Kit	Applied Biosystems Germany
BigDye Terminator v3.1 Cycle Sequencing Kit	Applied Biosystems South Africa
Bioline dNTPs (4x 25µmol solutions)	Celtic Diagnostics South Africa
Bromophenol blue	Sigma Germany
dNTPs (100mM solutions)	Amersham Biosciences Germany
Ethanol	Merck Germany
Ethidium bromide (10 mg/ml)	Invitrogen Germany
Ethidium bromide (10 mg/ml)	Sigma-Aldrich South Africa
Exonuclease I	Amersham Biosciences Germany
ExoSAPit	Amersham Biosciences South Africa
Genescan - 500 LIZ size standard	Applied Biosystems South Africa
GenomiPhi DNA amplification kit	Amersham Biosciences Germany
HiDi Formamide	Applied Biosystems South Africa
Hyperladder I	Bioline Germany
KpnI	Roche South Africa
MgCl ₂	Merck Germany
Nucleon BACC3 Kit for blood and cell cultures	Amersham Biosciences UK

Reagent		Supplier	
Orange loading dye solution	Inqaba Biotechnology	South Africa
PCR primers	UCT	South Africa
Promega 10X ligase buffer	Whitehead Scientific	South Africa
Promega pGL4.10[<i>luc2</i>]	Whitehead Scientific	South Africa
Promega PureYield Plasmid Midiprep System	Whitehead Scientific	South Africa
Promega Rvprimer 3	Whitehead Scientific	South Africa
Promega T4 DNA ligase	Whitehead Scientific	South Africa
Promega Wizard Plus SV Miniprep Kit	Whitehead Scientific	South Africa
Promega Wizard SV Gel and PCR Clean-up System	Whitehead Scientific	South Africa
Sephadex powder (G50 superfine)	Amersham Biosciences	Germany
Shrimp alkaline phosphatase	Amersham Biosciences	Germany
SNPlex ligation probes	Applied Biosystems	Germany
SNPlex system amplification kit	Applied Biosystems	Germany
SNPlex system assay standards kit	Applied Biosystems	Germany
SNPlex system hybridization kit	Applied Biosystems	Germany
SNPlex system ligation kit	Applied Biosystems	Germany
SNPlex system phosphorylation kit	Applied Biosystems	Germany
SNPlex system purification kit	Applied Biosystems	Germany
Super-therm GOLD Hotstart TAQ DNA polymerase	Southern Cross	South Africa
TaqMan SNP Genotyping Assay Mix	Applied Biosystems	Germany
TaqMan SNP Genotyping assays	Applied Biosystems	Germany
TaqMan Universal PCR Master Mix	Applied Biosystems	Germany
TBE buffer 10X ready pack	Amresco	USA
XhoI	Roche	South Africa
Xylene Cyanol FF	Sigma	Germany

A.3. EQUIPMENT

Equipment	Supplier	
384 deepwell storage plates	Abgene	UK
3130x/ DNA analyzer	Applied Biosystems	South Africa
3730x/ DNA analyzer	Applied Biosystems	Germany
Adhesive PCR film	Abgene	UK
ALPS-300 Platesealer	Abgene	UK
Axygen PCR tubes, strips and lids	Lasec	South Africa
BioDoc Analyze Video Documentation System	Whatman Biometra	Germany
Biometra T1 Thermal Cycler	Whatman Biometra	Germany
Bio-Rad Gene pulser	Bio-Rad	South Africa
Centri-sep 96 well plates	Princeton Separations	USA
Consort E844 Electrophoresis Power Supply	Sigma-Aldrich	South Africa
Eppendorf 0.5, 1.5 and 2 ml tubes	Merck	South Africa
Eppendorf Centrifuge 5415	Merck	South Africa
Eppendorf Centrifuge 5810 R	Merck	South Africa
Eppendorf Mastercycler	Merck	South Africa
GeneAmp PCR system 9700	Applied Biosystems	Germany
Gibco BRL Electrophoresis Power Supply 250 EX	Bio-Rad	Germany
Gibco BRL Horizontal Gel Electrophoresis Apparatus	Bio-Rad	Germany
High Performance Ultraviolet Transilluminator	VWR	Germany
Hydra 384 Robbins Scientific	Dunn Labortechnik GmbH	Germany
Hydra 96 Robbins Scientific	Dunn Labortechnik GmbH	Germany
Microtiter 384 well plates	Sarstedt	Germany
Microtiter 96 well plates	Sarstedt	Germany
Nanodrop ND-1000 Spectrophotometer	Inqaba Biotechnology	South Africa
Orbital shaker incubator LM-530	Yih Der	Japan

Equipment	Supplier	
Pipette tips	Whitehead Scientific	South Africa
Pipette tips with filter	Sarstedt	Germany
Sephadex spin column plates MAHVN4550	Amersham Biosciences	Germany
Single and multi-channel pipettes	Eppendorf	Germany
Syngene G box	Vacutec	South Africa
Tecan Genesis RSP 150	Tecan Deutschland GmbH	Germany
Tecan Genesis Workstation 150	Tecan Deutschland GmbH	Germany
Tecan Genesis Workstation 200	Tecan Deutschland GmbH	Germany
Tecan Powerwasher 384	Tecan Deutschland GmbH	Germany
Tecan Te-Mo multichannel pipetting option	Tecan Deutschland GmbH	Germany
TH15 incubator hood	Edmund Bühler Labortechnik	Germany
TiMix control shaker	Edmund Bühler Labortechnik	Germany
Vortex-Genie 2 G-560E	Scientific Industries Inc.	USA

A.4. SOFTWARE

Cocaphase	MRC Human Genome Mapping Project	UK
GeneMapper v3.5.1	Applied Biosystems	Germany
GeneMapper v3.7	Applied Biosystems	South Africa
Haploview v3.3	Daly lab, Broad Institute	USA
LIMS	IKMB	Germany
Nanodrop v3.0.1	Inqaba Biotechnology	South Africa
Sequencher v4.2	Gene Codes Corporation	USA
SDS 2.0	Applied Biosystems	Germany
SGCaller	IKMB	Germany

A.5. ONLINE RESOURCES

BLAST	www.ncbi.nlm.nih.gov/BLAST
CHIP Bioinformatics	snpper.chip.org/bio/snpper-enter
dbSNP	www.ncbi.nlm.nih.gov/SNP
Genetic Power Calculator	pngu.mgh.harvard.edu/~purcell/gpc/
HapMap	www.hapmap.org
Human BLAT search	genome.ucsc.edu/cgi-bin/hgBlat
myScience Environment	myscience.appliedbiosystems.com
SNPInspector	www.genomatix.de/products/SNPInspector
PubMed	www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=PubMed

A.6. POLYMORPHISMS DETECTED DURING *BTNL2* SEQUENCING

TABLE A1. SNPs detected during *BTNL2* sequencing.

Exon amplified	SNP location	Chromosome 6 position	rs-number	Allele frequencies		
				Allele	Cases	Controls
1	Promoter	32,485,094	rs9268504	G	0.587	0.417
				C	0.413	0.583
	Promoter	32,485,039	rs9268503	G	0.542	0.354
				A	0.458	0.646
	Promoter	32,484,910	rs9268502	G	0.541	0.583
				A	0.484	0.417
	Promoter	32,484,860	rs9268501	C	0.458	0.646
				A	0.542	0.354
	Promoter	32,484,766	rs3763317	G	0.604	0.5
				A	0.396	0.5
5' UTR	32,484,724	rs3763316	G	0.938	0.896	
			A	0.063	0.104	
Intron	32,484,632	rs3763315	C	0.979	0.979	
			A	0.021	0.021	
Intron	32,484,596	rs3763314	G	0.458	0.271	
			A	0.542	0.729	
2	Exon	32,482,843	none	G	0.042	0
				A	0.958	1
	Intron	32,482,618	rs3763308	C	0.833	0.87
				T	0.167	0.13
	Intron	32,482,600	rs3763307	T	0.938	0.891
				A	0.063	0.109
Intron	32,482,573	rs2395159	C	0.104	0.065	
			T	0.896	0.935	
3	Exon boundary	32,481,042	none	G	0.979	1
				A	0.021	0
	Exon	32,481,037	none	C	0.979	1
				T	0.021	0
	Exon	32,480,941	BTNL2_ex3_101	C	0.826	0.833
				T	0.174	0.167
	Exon	32,480,940	none	G	0.977	1
				A	0.023	0
Exon	32,480,841	BTNL2_ex3_201	A	0.167	0.167	
			T	0.833	0.833	
4	Intron	32,477,681	rs3834855	A	0.833	0.813
				-	0.167	0.188
	Intron	32,477,621	rs3763306	C	0.167	0.188
				T	0.833	0.813
	Intron	32,477,579	none	G	1	0.979
				A	0	0.021

Exon amplified	SNP location	Chromosome 6 position	rs-number	Allele frequencies		
				Allele	Cases	Controls
5	Intron	32,472,192	rs2076527	G	0.022	0.022
				A	0.978	0.978
	Intron	32,472,172	rs2076528	T	0.977	0.978
				G	0.023	0.022
	Intron	32,472,168	none	C	0.021	0.042
				T	0.979	0.958
	Exon	32,472,035	PM15830	G	0.979	0.979
				A	0.021	0.021
	Exon	32,472,030	PM15835	G	0.979	0.979
				A	0.021	0.021
	Exon	32,472,024	PM15841	A	0.979	0.979
				T	0.021	0.021
	Exon	32,471,989	PM15876	G	0.021	0.021
				A	0.979	0.979
	Exon	32,471,951	PM15914	G	0.979	0.979
				A	0.021	0.021
	Exon	32,471,933	rs2076529	G	0.25	0.271
				A	0.75	0.729
	Exon	32,471,866	PM15999	G	0.979	0.979
				A	0.021	0.021
Exon	32,471,822	rs9268480	G	0.938	0.896	
			A	0.063	0.104	
Exon boundary	32,471,794	rs2076530	G	0.25	0.271	
			A	0.75	0.729	
Intron	32,471,768	PM16097	C	0.021	0.021	
			T	0.979	0.979	
Intron	32,471,764	PM16101	C	0.958	0.979	
			T	0.042	0.021	
Intron	32,471,752	ss38346939	C	0.818	0.854	
			T	0.182	0.146	
6	Intron	32,470,908	rs12528939	G	0.833	0.854
				A	0.167	0.146
	Intron	32,470,810	rs17202435	C	0.833	0.854
				T	0.167	0.146
	Exon	32,470,723	BTNL2_ex6_58	C	0.886	0.909
				T	0.114	0.091
	Exon	32,470,719	BTNL2_ex6_62	G	0.864	0.864
				A	0.136	0.136
	Exon	32,470,681	BTNL2_ex6_100	C	0.833	0.854
				A	0.167	0.146
	Exon	32,470,680	BTNL2_ex6_101	G	0.13	0.109
				A	0.87	0.891
	Exon	32,470,647	BTNL2_ex6_134	C	0.167	0.146

Exon amplified	SNP location	Chromosome 6 position	rs-number	Allele frequencies			
				Allele	Cases	Controls	
	Exon	32,470,617	BTNL2_ex6_164	A	0.833	0.854	
				T	0.167	0.146	
	Intron	32,470,431	PM17434	C	0.833	0.854	
				T	0.188	0.146	
	Intron	32,470,425	none	G	0.813	0.854	
				A	0.958	0.938	
	Intron	32,470,411	PM17454	G	0.042	0.063	
				T	0.833	0.909	
	Intron	32,470,394	rs2076534	C	0.167	0.091	
				A	0.771	0.813	
	Intron	32,470,310	PM17479	C	0.229	0.188	
				T	0.826	0.848	
	7	Intron	32,469,872	none	T	0.174	0.152
					C	0.063	0.042
Intron		32,469,834	none	G	0.938	0.958	
				A	0.958	0.958	
Intron		32,469,820	none	C	0.042	0.042	
				T	0.958	0.958	
Intron		32,469,819	none	G	0.042	0.042	
				A	0.042	0.042	
Exon		32,469,799	rs3129953	G	0.958	0.958	
				A	0.896	0.938	
Exon		32,469,787	none	A	0.104	0.063	
				G	0.958	0.958	
Exon		32,469,740	none	C	0.042	0.042	
				T	0.958	0.958	
Exon	32,469,730	none	G	0.042	0.042		
			A	0.958	0.958		
Exon	32,469,727	none	C	0.042	0.042		
			G	0.958	0.958		
Intron	32,469,672	rs17202428	C	0.958	0.958		
			T	1	0.958		
Intron	32,469,655	rs17202414	C	0	0.042		
			T	0.958	0.958		
Intron	32,469,591	rs17202407	C	0.042	0.042		
			T	0.958	0.958		
Intron	32,469,590	rs17208769	A	0.042	0.042		
			G	0.958	0.958		
Intron	32,469,588	rs17208762	G	0.042	0.042		
			A	0.958	0.958		

Exon amplified	SNP location	Chromosome 6 position	rs-number	Allele frequencies		
				Allele	Cases	Controls
8	Exon	32,469,108	rs3817972	A	0.188	0.146
				G	0.813	0.854
	Exon	32,469,089	rs3817973	G	0.729	0.729
				A	0.271	0.271
	Exon	32,469,058	rs17208699	A	0.833	0.854
				C	0.167	0.146
	Intergenic	32,469,013	rs3817974	A	0.826	0.833
				G	0.174	0.167
	Intergenic	32,469,008	rs3817975	T	0.864	0.891
				A	0.136	0.109
	Intergenic	32,468,981	rs3817976	A	0.167	0.146
				G	0.833	0.854
	Intergenic	32,468,965	rs3817977	A	0.864	0.891
				C	0.136	0.109
	Intergenic	32,468,964	rs3817978	A	0.143	0.13
				G	0.857	0.87
	Intergenic	32,468,961	rs3817979	C	0.979	0.979
				T	0.021	0.021
	Intergenic	32,468,948	rs3817980	A	0.167	0.146
				G	0.833	0.854
Intergenic	32,468,933	rs17202316	C	0.833	0.854	
			T	0.167	0.146	
Intergenic	32,468,923	rs3817981	G	0.188	0.146	
			A	0.813	0.854	
Intergenic	32,468,893	rs13198563	C	0.042	0.104	
			T	0.958	0.896	
Intergenic	32,468,827	rs17208671	C	0.792	0.813	
			A	0.208	0.188	

A.7. PUBLICATIONS



Available online at www.sciencedirect.com



Molecular and Cellular Probes 21 (2007) 148–151

Molecular
and Cellular
Probes

www.elsevier.com/locate/ymcpr

Note

Host susceptibility to tuberculosis: *CARD15* polymorphisms in a South African population

Marlo Möller^a, Almut Nebel^b, Ruta Kwiatkowski^b, Paul D. van Helden^a,
Eileen G. Hoal^a, Stefan Schreiber^{b,*}

^aMolecular Biology and Human Genetics, MRC Centre for Molecular and Cellular Biology and the DST/NRF Centre of Excellence for Biomedical TB Research, Faculty of Health Sciences, PO Box 19063, Stellenbosch University, Tygerberg 7505, South Africa

^bInstitute for Clinical Molecular Biology, Christian-Albrechts-University Kiel, Schittenhelmstrasse 12, 24105 Kiel, Germany

Received 3 July 2006; accepted 11 October 2006
Available online 18 October 2006

Abstract

Tuberculosis (TB) is one of the leading causes of death worldwide. The nucleotide-binding oligomerisation domain 2 protein (NOD2) has recently been recognised as a non-redundant recognition mechanism of *Mycobacterium tuberculosis*. The caspase recruitment domain-containing protein 15 gene (*CARD15*), which encodes the NOD2 protein, is a susceptibility gene for Crohn's disease (CD), a granulomatous, chronic inflammatory disorder. *CARD15* was therefore investigated as a candidate gene in TB. We genotyped the R702W, G908R and 1007fs variants, previously associated with CD, in TB cases and controls from the admixed South African Coloured population. No statistically significant differences between cases and controls were observed for these variants. We determined that the CD-associated mutations occur at very low frequencies in this population. Our results indicate that *CARD15* is not a major susceptibility gene for TB in the South African Coloureds.

© 2006 Elsevier Ltd. All rights reserved.

Keywords: Tuberculosis; *CARD15*; NOD2; Case-control; Association; South African Coloured

Tuberculosis (TB) is found at a very high incidence in the Western Cape province of South Africa [1]. This infectious disease is caused by *Mycobacterium tuberculosis* (*M. tuberculosis*) and although both bacterial and environmental factors contribute to the development of the disease, it is increasingly evident that differences in host genes could also determine the eventual outcome of infection [2–4].

Granulomas can develop due to infectious diseases, such as TB, or because of idiopathic disorders such as Crohn's disease (CD). While TB granulomas occur mostly in the lungs, CD granulomas are found in the intestine but can also appear in the lungs. Since TB and CD share this pathophysiologic characteristic, susceptibility genes for CD are plausible candidate genes for investigation in TB. Moreover, although the cause of CD is unknown,

evidence is mounting that barrier function is compromised by genetic variants that can hinder interaction with bacterial antigens [5]. While it appears that the normal flora plays a pivotal role in the pathophysiology [6], a controversial theory [7,8] suggests that *Mycobacterium avium subspecies paratuberculosis* is one of the causative agents in CD [9].

The caspase recruitment domain-containing protein 15 gene (*CARD15*) encodes the nucleotide-binding oligomerisation domain 2 protein (NOD2) [8] and was identified as a susceptibility gene for CD [10–12]. The R702W (rs2066844), G908R (rs2066845) and 1007fs (rs2066847) polymorphisms in *CARD15* were independently associated with this form of inflammatory bowel disease [10]. The R702W variant, caused by a C to T transition, results in an arginine to tryptophan change in NOD2, while G908R is formed by a G to C transversion and changes a glycine to arginine. The 1007fs polymorphism is formed by a cytosine insertion at the second nucleotide in codon 1007 [11]. This

*Corresponding author. Tel.: +49 431 5972350; fax: +49 431 5971434.
E-mail address: s.schreiber@mucosa.de (S. Schreiber).

Table 1
Allele and genotype frequencies of the *CARD15* variants investigated in the South African Coloured population

Variant	R702W				G908R				1007fs			
	Genotype		Allele		Genotype		Allele		Genotype		Allele	
TB ^a	CC	0.986(421)	C	0.993	GG	0.995(426)	G	0.998	wt/wt	0.991(423)	C	0.995
	CT	0.014(6)			GC	0.005(2)			wt/insC ^b	0.009(4)		
	TT	0	T	0.007	CC	0	C	0.002	insC/insC	0	insC	0.005
Controls	CC	0.977(468)	C	0.989	GG	0.992(476)	G	0.996	wt/wt	0.983(423)	C	0.992
	CT	0.023(11)			GC	0.008(4)			wt/insC	0.017(8)		
	TT	0	T	0.011	CC	0	C	0.004	insC/insC	0	insC	0.008
<i>p</i> -Value	0.46				0.69				0.39			

The number of individuals with each genotype is indicated in brackets.

^aTB, tuberculosis.

^binsC, cytosine insertion.

results in a frameshift which leads to the truncation of NOD2 in the leucine-rich repeat region of the protein [10]. The truncated NOD2 protein formed as a consequence of the 1007fs insertion in *CARD15* leads to a decrease in the responsiveness of NF- κ B to the bacterial ligand muramyl dipeptide [11]. Muropeptides are breakdown products of peptidoglycans from Gram-negative and Gram-positive bacteria, which are also present in the cell walls of *M. tuberculosis* [13]. Mycobacterial infection induces activation of the transcription factor NF- κ B [14]. R702W and G908R, like 1007fs, also occur in the vicinity of the leucine-rich repeat region of the protein and compromise the recognition of muramyl dipeptides [15].

It was recently demonstrated that NOD2 is an essential recognition system of *M. tuberculosis* [16] and *CARD15* could consequently be considered as a candidate gene in TB. In a case-control study with Gambian TB patients [17] no associations were found with promoter polymorphisms of *CARD15*, and the three CD variants R702W, G908R and 1007fs were absent in the population. We therefore aimed to genotype these CD-associated polymorphisms in TB patients and controls from the admixed South African Coloured population and did a case-control association study to determine if the variants predispose to TB.

Subjects for this study were all from the population known as South African Coloured, unrelated and HIV-negative. This population is admixed, with genetic input from San, Khoi, Malaysians, African blacks and Europeans [18]. No significant stratification has been found in a previous study [19] using the same population as in this investigation. A total of 432 confirmed TB cases (age in years = 34 ± 14.8 , males = 53%) and 482 healthy controls (age in years = 27 ± 12.3 , males = 23%) were collected from the same suburbs in the Cape Town metropolitan area which have a high incidence of TB [20]. Blood samples were collected with informed consent and with permission from the Ethics Committee of the Faculty of Health Sciences (Stellenbosch University, South Africa). Genotyping for the three variants was done on an automated

platform with the TaqMan[®] system (Applied Biosystems, Foster City, USA) and assays were designed as published previously [12,21]. Allele and genotype frequencies were estimated by direct counting and statistical differences between the genotypes of patients and controls were tested with the Fisher's exact test (Prism version 4.00, GraphPad Software, San Diego, California, USA). Hardy-Weinberg equilibrium calculations were done.

The three CD polymorphisms in *CARD15* were genotyped in TB and control samples from the study population. All variants were in Hardy-Weinberg equilibrium.

The genotype and allele frequencies for the polymorphisms were very low and the minor alleles were found only in the heterozygous state (Table 1). Comparison of the genotype frequencies for the three polymorphisms revealed no association between any of the variants and TB (Table 1). Since all three *CARD15* variants are relatively rare and have been shown to occur on separate haplotypes [10], we compared the frequencies of carriers (i.e. heterozygote and double heterozygote carriers of any of the three SNPs) between the case sample and the controls. The carrier frequencies were not significantly different in the two groups (0.03 versus 0.05; $p = 0.22$).

The pathophysiology of both TB and CD is characterised by the presence of granulomas. We investigated the role of the CD susceptibility gene *CARD15* in TB in the South African population, because NOD2, the protein encoded by *CARD15*, was found to be an essential recognition mechanism for *M. tuberculosis* [16]. To this end, three variants that affect the leucine-rich-repeat recognition domains of NOD2 and that represent the main genetic risk associated with this gene in CD were genotyped.

This is the first report of allele frequencies for the R702W, G908R and 1007fs polymorphisms in TB samples from the South African Coloured population. The variant alleles of these polymorphisms are found at frequencies of about 2, 3 and 3% respectively in the European background populations [22], but were only present at frequencies below 1.1% in this study collection in both

the TB and control groups (Table 1). The frequencies of two (R702W, G908R) of the three variants genotyped in this control group were not significantly different ($p > 0.05$) from those observed in controls in a CD study of another South African Coloured sample set, which detected minor allele frequencies of 0.03 for R702W, 0.005 for G908R and 0.04 for 1007fs [23]. A genetic input from European ancestors is the most likely explanation for the presence of the CD variants in this admixed South African population, since these polymorphisms were not found at all in Gambians [17] or various Asian populations [22,24]. However, at least for the *CARD15* locus it seems that the genetic contribution from Europeans was relatively small.

TB is present in epidemic proportions in the South African Coloured population from the Western Cape province, while the occurrence of CD in this population is relatively rare [25]. In other African populations with a low incidence of CD, the polymorphisms associated with CD were absent [17]. The rarity of the CD polymorphisms in the South African Coloured population could in part explain the lower incidence of CD observed in this group [25], but other factors—such as the availability of health care and the higher frequency of infectious diarrhoea—could hinder the diagnosis of CD in this population [26]. Low frequencies for these mutations have also been observed in African-American CD patients [26]. The variants were previously investigated in CD [23] in a smaller sample set of South African Coloureds, but no association was detected with the disease.

We did not find any statistically significant association between the polymorphisms studied and TB due to their low allele frequencies in both the South African Coloured case and control samples. The common disease-common variant hypothesis states that the risk of contracting a common disease is determined by the presence of common variants in the population [27]. It can also be assumed that the detection of disease genes is facilitated in a population in which high exposure and high prevalence brings most of the genetic susceptibility alleles to manifestation. Since TB is a common disease in the South African Coloured population, the low allele frequencies of the *CARD15* mutations in this population indicate that the gene with its rare variants does not play a major role in predisposing to the disease. These variants could however still be predisposing factors in other populations where they are more frequently present. An investigation in TB cases from another population with a higher frequency of *CARD15* polymorphisms is therefore desirable.

In summary, we determined that *CARD15* is not a major susceptibility factor in South African Coloured TB patients. The low frequency of these variants in the population studied makes it unlikely that they play any role in disease aetiology.

Sample collection was funded by a Grant (053844/Z/98/Z) from the Wellcome Trust to PDvH and EGH. This study was funded by the German National Genome

Research Network (NGFN-2) and supported by the DFG research group FOR423.

References

- [1] City of Cape Town/Metropole Region TB Control Programme. Progress report. 1997–2002.
- [2] Bellamy R, Hill AV. Genetic susceptibility to mycobacteria and other infectious pathogens in humans. *Curr Opin Immunol* 1998;10:483–7.
- [3] Cooke GS, Hill AV. Genetics of susceptibility to human infectious disease. *Nat Rev Genet* 2001;2:967–77.
- [4] Hill AV. The genomics and genetics of human infectious disease susceptibility. *Annu Rev Genom Hum Genet* 2001;2:373–400.
- [5] Schreiber S, Rosenstiel P, Albrecht M, et al. Genetics of Crohn's disease, an archetypal inflammatory barrier disease. *Nat Rev Gen* 2005;6:376–88.
- [6] Ott SJ, Musfeldt M, Wenderoth DF, et al. Reduction in diversity of the colonic mucosa associated bacterial microflora in patients with active inflammatory bowel disease. *Gut* 2004;53:685–93.
- [7] Naser SA. Mycobacterium in Crohn's disease is hard to digest. *Gastroenterology* 2005;129:1359–60.
- [8] Cohen RD. Mycobacterium in Crohn's: something to ruminate about? *Gastroenterology* 2005;128:2167–8.
- [9] Chamberlin WM, Naser SA. Integrating theories of the etiology of Crohn's disease. On the etiology of Crohn's disease: questioning the hypotheses. *Med Sci Monit* 2006;12:RA27–33.
- [10] Hugot JP, Chamaillard M, Zouali H, et al. Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature* 2001;411:599–603.
- [11] Ogura Y, Bonen DK, Inohara N, et al. A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. *Nature* 2001;411:603–6.
- [12] Hampe J, Cuthbert A, Croucher PJ, et al. Association between insertion mutation in NOD2 gene and Crohn's disease in German and British populations. *Lancet* 2001;357:1925–8.
- [13] Girardin SE, Boneca IG, Viala J, et al. NOD2 is a general sensor of peptidoglycan through muramyl dipeptide (MDP) detection. *J Biol Chem* 2003;278:8869–72.
- [14] Yamada H, Mizuno S, Reza-Gholizadeh M, Sugawara I. Relative importance of NF-kappaB p50 in mycobacterial infection. *Infect Immun* 2001;69:7100–5.
- [15] Bonen DK, Ogura Y, Nicolae DL, et al. Crohn's disease-associated NOD2 variants share a signaling defect in response to lipopolysaccharide and peptidoglycan. *Gastroenterology* 2003;124:140–6.
- [16] Ferwerda G, Girardin SE, Kullberg BJ, et al. NOD2 and toll-like receptors are nonredundant recognition systems of *Mycobacterium tuberculosis*. *PLoS Pathog* 2005;1:279–85.
- [17] Stockton JC, Howson JM, Awomoyi AA, McAdam KP, Blackwell JM, Newport MJ. Polymorphism in NOD2, Crohn's disease, and susceptibility to pulmonary tuberculosis. *FEMS Immunol Med Microbiol* 2004;41:157–60.
- [18] Lombard Z, Brune AE, Hoal EG, et al. HLA class II disease associations in southern Africa. *Tissue Antigens* 2006;67:97–110.
- [19] Barreiro LB, Neyrolles O, Babb CL, et al. Promoter variation in the DC-SIGN encoding gene CD209 is associated with tuberculosis. *PLoS Med* 2006;3:e20.
- [20] Munch Z, Van Lill SWP, Booysen CN, Zietsman HL, Enarson DA, Beyers N. Tuberculosis transmission patterns in a high-incidence area: a spatial analysis. *Int J Tuberc Lung Dis* 2003;7:271–7.
- [21] Hampe J, Grebe J, Nikolaus S, et al. Association of NOD2 (*CARD15*) genotype with clinical course of Crohn's disease: a cohort study. *Lancet* 2002;359:1661–5.
- [22] Marsh S, McLeod HL. Crohn's disease: ethnic variation in *CARD15* genotypes. *Gut* 2003;52:770.

- [23] Zaahl MG, Winter T, Warnich L, Kotze MJ. Analysis of the three common mutations in the *CARD15* gene (R702W, G908R and 1007fs) in South African colored patients with inflammatory bowel disease. *Mol Cell Probes* 2005;19:278–81.
- [24] Inoue N, Tamura K, Kinouchi Y, et al. Lack of common NOD2 variants in Japanese patients with Crohn's disease. *Gastroenterology* 2002;123:86–91.
- [25] Wright JP, Froggatt J, O'Keefe EA, et al. The epidemiology of inflammatory bowel disease in Cape Town 1980–1984. *S Afr Med J* 1986;70:10–5.
- [26] Reddy SI, Burakoff R. Inflammatory bowel disease in African Americans. *Inflamm Bowel Dis* 2003;9:380–5.
- [27] Reich DE, Lander ES. On the allelic spectrum of human disease. *Trends Genet* 2001;17:502–10.

Author's personal copy

Original article

Allelic variation in *BTNL2* and susceptibility to tuberculosis in a South African population

Marlo Möller^a, Ruta Kwiatkowski^b, Almut Nebel^b, Paul D. van Helden^a,
Eileen G. Hoal^{a,1}, Stefan Schreiber^{b,*}

^a Molecular Biology and Human Genetics, MRC Centre for Molecular and Cellular Biology and the DST/NRF Centre of Excellence for Biomedical TB Research, Faculty of Health Sciences, P.O. Box 19063, Stellenbosch University, Tygerberg 7505, South Africa

^b Institute for Clinical Molecular Biology, Christian-Albrechts-University Kiel, Schüttenhelmstrasse 12, 24105 Kiel, Germany

Received 24 October 2006; accepted 16 January 2007

Available online 27 January 2007

Abstract

Tuberculosis and sarcoidosis show phenotypic features of granulomatous disease. The bacterium *Mycobacterium tuberculosis* can induce the expression of the sarcoidosis susceptibility gene *BTNL2* in monocyte-derived macrophages. *BTNL2* was therefore investigated as a candidate gene for tuberculosis in a case–control association study in the South African Coloured population. We sequenced the coding regions of *BTNL2* to detect known and novel polymorphisms and genotyped 18 SNPs in 432 pulmonary tuberculosis cases and 482 controls. We did not find a significant association between the truncating rs2076530 SNP, previously associated with sarcoidosis, and tuberculosis. No association was found between any of the other SNPs studied and disease and none of the estimated haplotypes showed any association with TB. Comparative analyses with the South African data from this study and published data on German and American populations revealed that, for a segment of *BTNL2*, the admixed, but not stratified, South African population resembles the African-Americans more than white populations. © 2007 Elsevier Masson SAS. All rights reserved.

Keywords: *BTNL2*; Tuberculosis; Association; Case–control; South African Coloured

1. Introduction

Tuberculosis (TB) has re-emerged as a threat to world health. The causative bacterium, *Mycobacterium tuberculosis* (*M. tuberculosis*), is thought to infect one new case every second [1] and one-third of the world's population is latently infected. Only 10% of those infected will ever develop the active disease. It is estimated that in 2003, 1.75 million people died of TB and approximately 8.8 million new cases developed the active disease [1].

Susceptibility or resistance to TB has a large genetic component as proven by twin studies [2], but environmental and

socio-economic factors also contribute to the outcome [3]. Since TB is a complex disease, genetic studies in TB can be challenging, although several susceptibility genes have been identified [4], such as natural resistance associated macrophage protein 1 (*NRAMP1*) [5], mannose binding lectin [6], dendritic cell-specific ICAM-3 grabbing non-integrin [7], vitamin D (1,25-dihydroxyvitamin D3) receptor [8] and the human leukocyte antigen (*HLA*) genes [9].

Sarcoidosis is a multiorgan disorder that affects mostly the lungs and lymphatic system. The disease is of unknown aetiology, but some evidence supports the hypothesis that it is caused by an infectious agent in genetically susceptible hosts [10]. Sarcoidosis has been considered to be a different form of TB [11], because of some similarities between the diseases, such as the presence of granulomas and coughing [12]. This is still a popular hypothesis and recently the *M. tuberculosis* heat shock protein 70 was suggested to have an influence in

* Corresponding author. Tel.: +49 431 597 2350; fax: +49 431 597 1434.
E-mail address: s.schreiber@mucosa.de (S. Schreiber).

¹ Both authors contributed equally to this paper.

the pathogenesis of sarcoidosis [13]. Susceptibility genes in sarcoidosis can therefore also be considered as candidate genes for TB and vice versa. Associations have been detected in both diseases with the same candidate genes, such as *NRAMP1* [5,14] and *HLA* [9,15].

The butyrophilin-like 2 gene (*BTNL2*, *BTL-II*) was first identified by comparing the human and mouse genomic sequences at the major histocompatibility complex (MHC) class II and class III regions [16]. It is one of the B7 receptor family genes and was hypothesised to have a possible function as a T cell co-stimulatory molecule [17], because of amino acid homology to B7-1 (CD80) and B7-2 (CD86). These co-stimulatory molecules were found to be important for an effective T cell response against mycobacteria in a mouse model of TB [18]. A recent study determined that *BTNL2* inhibits T cell activation in mice [19].

Recently, we have identified a truncating splice site variation in *BTNL2* that is associated with sarcoidosis in a German population [17]. This association was subsequently confirmed by others in an additional population, namely white Americans [20]. Interestingly, this SNP (rs2076530) was associated with sarcoidosis only in white populations but not in African-American patients, where a three-locus haplotype including this SNP was associated with the disease [20]. The rs2076530 SNP alters the splicing pattern of the gene and a premature stop codon is formed. The protein product of this splice form is truncated and could limit the T cell downregulatory function of *BTNL2* [17]. It has been shown that *M. tuberculosis* and lipopolysaccharides can induce *BTNL2* in monocyte-derived macrophages [17]. Variations in *BTNL2* could therefore have a possible function in genetic susceptibility or resistance to TB.

In this study, we sequenced *BTNL2* and investigated the influence of the gene in TB, applying a population-based association study design in the admixed, but not stratified [7] South African Coloured population.

2. Materials and methods

2.1. Study population

Samples for this study were collected in the Cape Town metropolitan area, Western Cape Province of South Africa, where the incidence of TB is greater than 500 per 100,000 population per year [21]. In some of the suburbs, where the majority of samples were collected, the incidence of TB was 1340 per 100,000 population in 1996 [22], with a low prevalence of HIV.

The study population, known as South African Coloured, has a mixed ancestry, dating back many generations, that includes San, Khoi, Malaysians, African blacks and Europeans [9]. No significant stratification has been observed in a previous investigation which genotyped 25 unlinked SNP markers using the population of the present study [7]. Blood samples were collected with permission from the Ethics Committee of the Faculty of Health Sciences (Stellenbosch University,

South Africa) and the DNA was purified by standard methods.

The 432 cases (age in years = 34 ± 14.8 , males = 53%) and 482 controls (age in years = 27 ± 12.3 , males = 23%) were unrelated. All the cases had bacteriologically confirmed TB. The controls were healthy people with no history of TB who lived in the same high incidence community as the TB patients and were therefore highly likely to have been exposed to TB. All subjects were HIV negative.

2.2. Mutation detection in *BTNL2*

In order to determine the frequency of known SNPs and the presence of novel SNPs in *BTNL2*, the eight coding exons and their flanking regions together with the putative promoter were amplified and sequenced in 48 individuals (24 controls, 24 cases, in total 96 chromosomes) using previously described primers [17]. Sequencing reactions were performed with Big-Dye v1.1 (Applied Biosystems, Foster City, USA). Sequence data were aligned with Sequencher 4.2 (Gene Codes Corporation, Michigan, USA).

2.3. Genotyping

Genotyping of the SNPs was performed with the SNPlex Genotyping System™ [23] (Applied Biosystems, Foster City, USA), according to the instructions of the manufacturer, on an automated platform and a laboratory information management system (LIMS) described previously [24,25]. Criteria for the inclusion of a SNP into the study were as follows: a minor allele frequency of at least 5% and presence in a functional region of the gene i.e. in an exon, splice site or putative promoter.

2.4. Statistical analysis

Hardy–Weinberg equilibrium (HWE) was calculated. A case–control study was done to evaluate the association of *BTNL2* with TB in the South African Coloured population. The χ^2 -test was used with 2×3 contingency tables to determine if there was any significant difference between the genotypes of cases and controls for a single marker. Power calculations were performed online with OpenEpi (<http://www.openepi.com>). With the sample size available, we had 80% power to detect an odds ratio of 2.15 with 95% confidence given an expected allele frequency of at least 5% for the SNPs studied. Bonferroni corrections for multiple testing were done by considering the number of independent LD blocks [26]. We determined three independent LD blocks (Fig. 1) and consequently a p -value of 0.017 was adopted as a threshold for significance.

The Haploview v3.2 [27] program was used for the analysis of haplotypes in cases and controls and for determining linkage disequilibrium (LD) in controls. Haplotype blocks were automatically selected by Haploview using a common block definition from Gabriel et al. [28]. In addition, we tested global significance for these haplotypes using the Cocophase

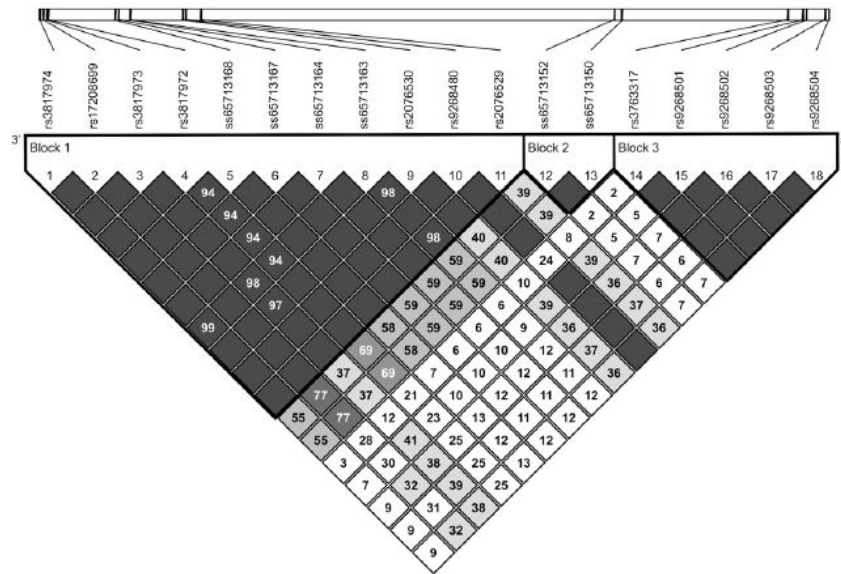


Fig. 1. Plot of LD between all markers analysed in control individuals of South African Coloured descent generated by Haploview v3.2 [27]. The 5'- and 3'- ends of the gene are indicated. D' -values (%) are indicated on the squares and squares without any number indicate $D' = 100\%$. Haplotype blocks are outlined in black.

program in the Unphased suite [29]. For each block 10,000 permutation replicates were done.

3. Results

3.1. Mutation detection and single-point SNP analyses in *BTNL2*

Sequencing of *BTNL2* in the South African Coloured population led to the identification of 76 SNPs: five upstream of the gene, 29 exonic, 31 intronic and 11 downstream of the gene. Of the SNPs detected, 15 were novel, but were either in introns or had a very low allele frequency. According to the set inclusion criteria (see Section 2), 18 SNPs were genotyped in the entire DNA collection (Table 1). The influence of these SNPs in TB was investigated with a case–control study design. All genotyped SNPs were in agreement with HWE in the control population.

The SNP rs2076530, which leads to the truncation of the *BTNL2* protein, was previously shown to be associated in sarcoidosis [17]. However, in our study, there was not a significant association ($p = 0.065$) between this SNP and TB (Table 1). Comparisons of the genotype frequencies of the remaining 17 SNPs analysed revealed a p -value of 0.023 for the rs9268480 SNP (Table 1). Given the adjusted significance level in our study, this p -value is not considered to reflect an association. None of the other SNPs examined were found to be associated with the disease. Altogether, single marker analysis revealed no association between TB and any of the identified SNPs in the *BTNL2* gene.

3.2. Haplotype analyses

The Haploview program [27] identified three haplotype blocks (Fig. 1) for the markers genotyped in *BTNL2* in the samples studied. Strong LD exists across the *BTNL2* gene region in the South African Coloured population (Fig. 1). Haplotype 6 (Table 2) was tagged by the minor A allele of rs9268480 and although it was nominally significant, the evaluation of global significance between cases and controls with the Cocophase program [29] yielded a non-significant global p -value ($p > 0.05$) for each of the three haplotype blocks after 10,000 permutations (Table 2).

3.3. Comparison of *BTNL2* data between South Africans and Americans

The South African control data were compared to previously published genotyping data from African-American and white American individuals [20]. Rybicki et al. [20] detected 10 SNPs in the exon 5/intron 5 region of *BTNL2*. For the analysis, data on three overlapping SNPs, namely rs2076529, rs9268480 and rs2076530, were available from both studies. Comparison of the three SNPs in common revealed that the South African allele frequencies were in general more similar to those of the African-Americans than to the white Americans. In addition, two of the five other SNPs (ss38346932 and ss38346939) found by Rybicki et al. [20] were identified by sequencing 24 control individuals from the South African samples (Table 3). The frequencies of these two SNPs were also more similar to the African-American population.

Table 1
Single-point statistical analysis in the case–control study design

SNP	Allele		Genotype frequencies						Cases vs controls <i>p</i> -Value ^c
			Cases (<i>n</i> ^a = 432)			Controls (<i>n</i> = 482)			
	1 ^b	2	11	12	22	11	12	22	
rs3817974	G	A	0.01	0.25	0.74	0.02	0.22	0.76	0.32
rs17208699	C	A	0.01	0.19	0.80	0.01	0.17	0.82	0.44
rs3817973	A	G	0.11	0.38	0.51	0.08	0.36	0.56	0.21
rs3817972	A	G	0.01	0.20	0.79	0.02	0.17	0.81	0.28
ss65713168	T	C	0.01	0.23	0.76	0.01	0.21	0.78	0.39
ss65713167	C	A	0.01	0.23	0.76	0.02	0.21	0.77	0.36
ss65713164	A	G	0.01	0.23	0.76	0.02	0.21	0.77	0.36
ss65713163	T	C	0.01	0.23	0.76	0.02	0.21	0.77	0.37
rs2076530	G	A	0.11	0.40	0.49	0.07	0.38	0.55	0.065
rs9268480	A	G	0.05	0.25	0.70	0.02	0.22	0.76	0.023
rs2076529	G	A	0.11	0.39	0.50	0.07	0.38	0.55	0.083
ss65713152	A	T	0.01	0.25	0.74	0.02	0.24	0.74	0.74
ss65713150	T	C	0.01	0.25	0.74	0.02	0.24	0.74	0.73
rs3763317	G	A	0.18	0.48	0.34	0.18	0.47	0.35	0.94
rs9268501	A	C	0.15	0.45	0.40	0.15	0.45	0.40	0.99
rs9268502	G	A	0.15	0.45	0.40	0.15	0.45	0.40	0.98
rs9268503	G	A	0.14	0.46	0.40	0.15	0.45	0.40	0.96
rs9268504	C	G	0.15	0.45	0.40	0.15	0.45	0.40	0.98

^a Number of investigated individuals.

^b Allele 1 is the minor allele.

^c *p*-Value from a genotype-based χ^2 -test; the level of significance was 0.017 after Bonferroni correction. Under this criterion, none of the analysed SNPs reached statistical significance.

3.4. Comparison of *BTNL2* data between South Africans and Germans

Comparative analyses were done for 482 South African controls from this study and 533 German controls from another study [17] using 6 SNPs that were genotyped in both populations (Table 3). The allele frequencies were very similar for four of the SNPs, but in two instances differed significantly

(rs2076530, *p* = 0.03 and rs2076529, *p* = 0.04). The allele frequencies for these SNPs in Germans were similar to those found in the white Americans [20].

With the block definition used in Haploview, two haplotype blocks were identified in the South African population and three in the German population. The reason for this difference is the slightly stronger LD observed in South Africans for this region (Fig. 2).

Table 2
Haplotype analysis

Block ^a	Global test	Individual tests		<i>p</i> -Value	
	<i>p</i> -Value ^b	Haplotype (case–control frequency)			
1	0.17	1	A-A-G-G-C-A-G-C-A-G-A (0.694/0.733)		0.067
		2	A-A-G-G-C-A-G-C-G-G-A (0.00490/0.00633)		0.39
		3	A-A-G-G-C-A-G-T-A-G-A (0/0.00106)		0.27
		4	A-A-G-G-T-C-G-C-A-G-A (0.00123/0)		0.21
		5	A-A-A-G-C-A-G-C-A-G-A (0.00123/0.00527)		0.13
		6	A-A-A-G-C-A-G-C-G-A-G (0.173/0.132)		0.017
		7	G-C-A-A-T-C-A-T-G-G-G (0.0993/0.0949)		0.76
		8	G-A-A-A-C-A-G-C-G-G (0.00735/0.00422)		0.38
		9	G-A-A-A-T-C-A-T-G-G-G (0.00123/0.00317)		0.38
		10	G-A-A-G-T-C-A-T-G-G-G (0.0184/0.0200)		0.8
2	0.91	11	A-T (0.139/0.141)		0.91
		12	T-C (0.861/0.859)		0.91
3	0.83	13	G-C-A-A-G (0.0454/0.0367)		0.36
		14	G-A-G-G-C (0.375/0.375)		0.98
		15	G-A-G-A-C (0.00119/0.00105)		0.93
		16	A-C-A-A-G (0.579/0.798)		0.72

^a The order of the SNPs in each block corresponds to Fig. 1.

^b Permutation test *p*-values were calculated from 10,000 permutations in Cocaphase [29], to correct for multiple testing while taking account of correlation between markers and haplotypes.

Table 3
Comparisons of the allele frequencies in the African-American, white American, South African Coloured and German control populations

SNP ^a	Allele frequencies in African-Americans [20]			Allele frequencies in white Americans [20]			Allele frequencies in South African Coloured controls			Allele frequencies in Germans				
	Allele frequency	Heterozygote frequency	Allele	Allele frequency	Heterozygote frequency	Allele	Allele frequency	Heterozygote frequency	Allele	Allele frequency	Heterozygote frequency	Allele	Allele frequency	Heterozygote frequency
rs65713164	Not genotyped		Not genotyped			G	0.88	0.21	G	0.85	0.26	G	0.85	0.26
rs65713163	Not genotyped		Not genotyped			A	0.12		A	0.15		A	0.15	
rs34423804	A 0.98	0.033	A	0.999	0.001	C	0.87	0.21	C	0.85	0.26	C	0.85	0.26
rs36110770	T 0.02	0.032	T	0.001	0.001	T	0.13	0.042	T	0.15		T	0.15	
rs2076529	C 0.98		C	0.999	0.001	A	0.98		A	Not genotyped		A	Not genotyped	
rs2076529	T 0.02		T	0.001		T	0.02		T	Not genotyped		T	Not genotyped	
rs2076529	A 0.70	0.42	A	0.61	0.52	Not detected by sequencing			Not detected by sequencing			Not detected by sequencing		
rs28362679	G 0.30	0.03	G	0.39	0.04	A	0.74	0.38	A	0.59	0.50	A	0.59	0.50
rs28362679	C 0.98		C	0.98		G	0.26		G	0.41		G	0.41	
rs28362679	T 0.02		T	0.02		Not detected by sequencing			Not detected by sequencing			Not detected by sequencing		
rs9268480	G 0.85	0.26	G	0.71	0.42	G	0.87	0.21	G	Not genotyped		G	Not genotyped	
rs9268480	A 0.15	0.045	A	0.29	0.001	A	0.13		A	Not genotyped		A	Not genotyped	
rs35037492	G 0.98		G	0.999		Not detected by sequencing			Not detected by sequencing			Not detected by sequencing		
rs35037492	A 0.02		A	0.001		A	0.74	0.38	A	0.58	0.51	A	0.58	0.51
rs2076530	A 0.70	0.42	A	0.60	0.52	G	0.26		G	0.42		G	0.42	
rs2076530	G 0.30	0.23	G	0.40	0.16	C	0.85	0.21	C	Not genotyped		C	Not genotyped	
rs17202456	C 0.87		C	0.91		T	0.15		T	Not genotyped		T	Not genotyped	
rs17202456	T 0.13		T	0.09		T	0.86	0.24	T	0.92	0.15	T	0.92	0.15
rs65713152	Not genotyped		Not genotyped			A	0.14		A	0.08		A	0.08	
rs65713150	Not genotyped		Not genotyped			C	0.86	0.24	C	0.92	0.15	C	0.92	0.15
rs65713150	Not genotyped		Not genotyped			T	0.14		T	0.08		T	0.08	

^a SNPs marked in bold were genotyped in all the samples of the South African population, while genotyping data for the SNPs in normal text were derived from sequencing.

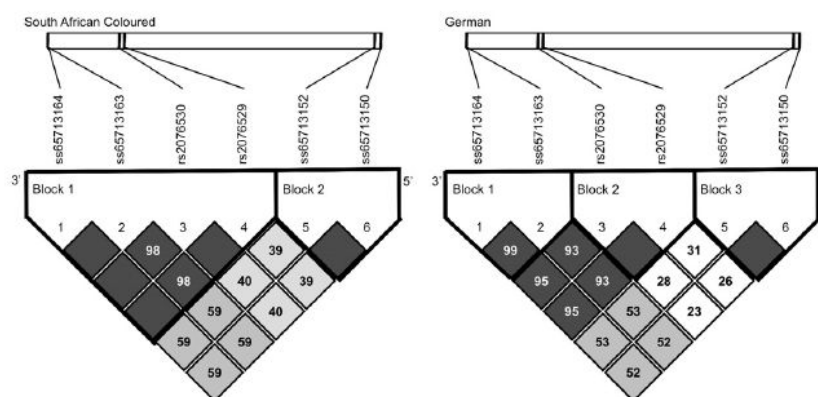


Fig. 2. Comparison of LD between the South African and German populations for 6 common SNPs. Squares are coded in the same manner as in Fig. 1. Two haplotype blocks were identified for the South Africans and three for the Germans.

4. Discussion

TB and sarcoidosis are both granulomatous diseases, occur mainly in the lungs and can cause coughing and weight loss [12]. Because of these similarities in presentation of the diseases, we investigated the involvement of *BTNL2*, a susceptibility gene for sarcoidosis in a tuberculosis sample of South African Coloured descent. We first identified this gene as a sarcoidosis susceptibility factor [17], and Rybicki et al. confirmed the association with rs2076530 in white Americans, but not in African-Americans [20]. Due to the high incidence of latent infection in the control community (own unpublished results), we have tested the role of *BTNL2* in disease progression only, though it is still possible that this gene may play a role in primary infection.

We did not observe a significant association of the truncating splice site SNP (rs2076530) with TB at the genotype level. Interestingly, the truncating allele A was more frequent in controls. Since *BTNL2* has a T cell downregulatory effect in mice [19], individuals with the truncated product could have a stronger T cell response, depending on which T cell subset was affected most. This could be an advantage in TB, since resistance to TB is dependant on the T helper 1 cell response of the infected individual. None of the 17 other SNPs analysed in *BTNL2* were associated with disease (Table 1) and no association was found between any of the estimated haplotypes and TB (Table 2).

Comparisons between the South African and German populations revealed slightly stronger LD for the South African population in the *BTNL2* region. For this reason, two haplotype blocks were identified in the South African population and three in the Germans. Stronger LD is expected in admixed populations [30] such as the South African Coloured population.

The South African Coloured population has several different parental populations, including European-descent and sub-Saharan African populations. The German population is outbred and genetically relatively homogeneous and could (together with the white Americans) be considered as an

approximation for the white parental population, while the African-Americans could be evaluated as another admixed population [30]. The comparison of the allele frequencies for polymorphisms in exon 5/intron 5 (Table 3) between the populations revealed that, for this restricted region of the genome, the South African Coloured population is more similar to African-Americans than to whites. Therefore, the lack of association with the functional *BTNL2* SNP and TB in the South African Coloured population could reflect the African admixture and would not preclude a role of rs2076530 in a white TB population, in a similar manner as seen with sarcoidosis [20]. Future work will evaluate *BTNL2* in a TB sample set with no African genetic contribution, to determine if the association of this gene with disease could be population-specific.

Acknowledgements

We thank all study participants for their cooperation; T. Wesse, L. Bossen, M. Davids, A. Dietsch and M. Friskovec for technical help and R. Vogler for database support. Sample collection was funded by a grant from the Wellcome Trust (053844/Z/98/Z) to Paul van Helden and Eileen Hoal. This study was funded by the German National Genome Research Network (NGFN-2) and supported by the DFG research group FOR423.

References

- [1] World Health Organization, Global tuberculosis control – surveillance, planning, financing, WHO Report 2005, WHO/HTM/TB/2005.349. Geneva; 2005.
- [2] G.W. Comstock, Tuberculosis in twins: a re-analysis of the Prophit survey, *Am. Rev. Respir. Dis.* 117 (1978) 621–624.
- [3] R. Bellamy, Genetics and pulmonary medicine. 3. Genetic susceptibility to tuberculosis in human populations, *Thorax* 53 (1998) 588–593.
- [4] G.S. Cooke, A.V. Hill, Genetics of susceptibility to human infectious disease, *Nat. Rev. Genet.* 2 (2001) 967–977.
- [5] E.G. Hoal, L.-A. Lewis, S.E. Jamiesson, F. Tanzer, M. Rossouw, T. Victor, R. Hillerman, N. Beyers, J.M. Blackwell, P.D. van Helden, *SLC11A1*

- (*NRAMP1*) but not *SLC11A2* (*NRAMP2*) polymorphisms are associated with susceptibility to tuberculosis in a high-incidence community in South Africa, *Int. J. Tuberc. Lung Dis.* 8 (2004) 1464–1471.
- [6] E.G. Hoal-van Helden, J. Epstein, T.C. Victor, D. Hon, L.A. Lewis, N. Beyers, D. Zurakowski, A.B. Ezekowitz, P.D. van Helden, Mannose-binding protein B allele confers protection against tuberculous meningitis, *Pediatr. Res.* 45 (1999) 459–464.
- [7] L.B. Barreiro, O. Neyrolles, C.L. Babb, L. Tailleux, H. Quach, K. McElreavey, P.D. Helden, E.G. Hoal, B. Gicquel, L. Quintana-Murci, Promoter variation in the DC-SIGN encoding gene *CD209* is associated with tuberculosis, *PLoS Med.* 3 (2006) e20.
- [8] L. Bornman, S.J. Campbell, K. Fielding, B. Bah, J. Sillah, P. Gustafson, K. Manneh, I. Lisse, A. Allen, G. Sirugo, A. Sylla, P. Aaby, K.P. McAdam, O. Bah-Sow, S. Bennett, C. Lienhardt, A.V. Hill, Vitamin D receptor polymorphisms and susceptibility to tuberculosis in West Africa: a case-control and family study, *J. Infect. Dis.* 190 (2004) 1631–1641.
- [9] Z. Lombard, A.E. Brune, E.G. Hoal, C. Babb, P.D. van Helden, J.T. Eppelen, L. Bornman, HLA class II disease associations in southern Africa, *Tissue Antigens* 67 (2006) 97–110.
- [10] American Thoracic Society, Statement on sarcoidosis, Joint Statement of the American Thoracic Society (ATS), the European Respiratory Society (ERS) and the World Association of Sarcoidosis and Other Granulomatous Disorders (WASOG) adopted by the ATS Board of Directors and by the ERS Executive Committee, February 1999, *Am. J. Respir. Crit. Care Med.* 160 (1999) 736–755.
- [11] D.J. Lyons, J.F. Fielding, What's in a relationship? Sarcoidosis and tuberculosis, *Ir. Med. J.* 83 (1990) 76–79.
- [12] I. Litinsky, O. Elkayam, G. Flusser, R. Segal, M. Yaron, D. Caspi, Sarcoidosis: TB or not TB?, *Ann. Rheum. Dis.* 61 (2002) 385–386.
- [13] A. Dubaniewicz, S. Kampfer, M. Singh, Serum anti-mycobacterial heat shock proteins antibodies in sarcoidosis and tuberculosis, *Tuberculosis (Edinb.)* 86 (2006) 60–67.
- [14] M.J. Maliarik, K.M. Chen, R.G. Sheffer, B.A. Rybicki, M.L. Major, J. Popovich, M.C. Iannuzzi, The natural resistance-associated macrophage protein gene in African Americans with sarcoidosis, *Am. J. Respir. Cell Mol. Biol.* 22 (2000) 672–675.
- [15] M.D. Rossman, B. Thompson, M. Frederick, M. Maliarik, M.C. Iannuzzi, B.A. Rybicki, J.P. Pandey, L.S. Newman, E. Magira, B. Beznik-Cizman, D. Monos, ACCESS Group, *HLA-DRB1*1101*: a significant risk factor for sarcoidosis in blacks and whites, *Am. J. Hum. Genet.* 73 (2003) 720–735.
- [16] M. Stammers, L. Rowen, D. Rhodes, J. Trowsdale, S. Beck, *BTL-II*: a polymorphic locus with homology to the butyrophilin gene family, located at the border of the major histocompatibility complex class II and class III regions in human and mouse, *Immunogenetics* 51 (2000) 373–382.
- [17] R. Valentonyte, J. Hampe, K. Huse, P. Rosenstiel, M. Albrecht, A. Stenzel, M. Nagy, K.I. Gaede, A. Franke, R. Haesler, A. Koch, T. Lengauer, D. Seegert, N. Reiling, S. Ehlers, E. Schwinger, M. Platzer, M. Krawczak, J. Müller-Quernheim, M. Schurmann, S. Schreiber, Sarcoidosis is associated with a truncating splice site mutation in *BTNL2*, *Nat. Genet.* 37 (2005) 357–364.
- [18] D. Yadav, G. Khuller, Evaluation of the T cells and costimulatory molecules in the protective efficacy of 30 kDa secretory protein against experimental tuberculosis, *Immunol. Cell Biol.* 79 (2001) 207–212.
- [19] T. Nguyen, X.K. Liu, Y. Zhang, C. Dong, *BTNL2*, a butyrophilin-like molecule that functions to inhibit T cell activation, *J. Immunol.* 176 (2006) 7354–7360.
- [20] B.A. Rybicki, J.L. Walewski, M.J. Maliarik, H. Kian, M.C. Iannuzzi, The *BTNL2* gene and sarcoidosis susceptibility in African Americans and Whites, *Am. J. Hum. Genet.* 77 (2005) 491–499.
- [21] City of Cape Town/Metropole Region TB Control Programme, Progress report, 1997–2002, 2002.
- [22] Z. Munch, S.W.P. Van Lill, C.N. Booysen, H.L. Zietsman, D.A. Enarson, N. Beyers, Tuberculosis transmission patterns in a high-incidence area: a spatial analysis, *Int. J. Tuberc. Lung Dis.* 7 (2003) 271–277.
- [23] F.M. De la Vega, K.D. Lazaruk, M.D. Rhodes, M.H. Wenz, Assessment of two flexible and compatible SNP genotyping platforms: TaqMan SNP genotyping assays and the SNPLex genotyping system, *Mutat. Res.* 573 (2005) 111–135.
- [24] J. Hampe, A. Wollstein, T. Lu, H.J. Frevel, M. Will, C. Manaster, S. Schreiber, An integrated system for high throughput TaqMan based SNP genotyping, *Bioinformatics* 17 (2001) 654–655.
- [25] M. Teuber, W.A. Koch, C. Manaster, S. Wachter, J. Hampe, S. Schreiber, Improving quality control and workflow management in high-throughput single-nucleotide polymorphism genotyping environments, *J. Assoc. Lab. Autom.* 10 (2005) 43–47.
- [26] K.K. Nicodemus, W. Liu, G.A. Chase, Y.Y. Tsai, M.D. Fallin, Comparison of type I error for multiple test corrections in large single-nucleotide polymorphism studies using principal components versus haplotype blocking algorithms, *BMC. Genet.* 6 (Suppl. 1) (2005) S78.
- [27] J.C. Barrett, B. Fry, J. Maller, M.J. Daly, Haploview: analysis and visualization of LD and haplotype maps, *Bioinformatics* 21 (2005) 263–265.
- [28] S.B. Gabriel, S.F. Schaffner, H. Nguyen, J.M. Moore, J. Roy, B. Blumenstiel, J. Higgins, M. DeFelice, A. Lochner, M. Faggart, S.N. Liu-Cordero, C. Rotimi, A. Adeyemo, R. Cooper, R. Ward, E.S. Lander, M.J. Daly, D. Altshuler, The structure of haplotype blocks in the human genome, *Science* 296 (2002) 2225–2229.
- [29] F. Dudbridge, Pedigree disequilibrium tests for multilocus haplotypes, *Genet. Epidemiol.* 25 (2003) 115–121.
- [30] M.W. Smith, N. Patterson, J.A. Lautenberger, A.L. Truelove, G.J. McDonald, A. Waliszewska, B.D. Kessing, M.J. Malasky, C. Scafe, E. Le, P.L. De Jager, A.A. Mignault, Z. Yi, G. De The, M. Essex, J.L. Sankale, J.H. Moore, K. Poku, J.P. Phair, J.J. Goedert, D. Vlahov, S.M. Williams, S.A. Tishkoff, C.A. Winkler, F.M. De la Vega, T. Woodage, J.J. Sninsky, D.A. Hafler, D. Altshuler, D.A. Gilbert, S.J. O'Brien, D. Reich, A high-density admixture map for disease gene discovery in African Americans, *Am. J. Hum. Genet.* 74 (2004) 1001–1013.

 WILEY

INNATE IMMUNITY TO PULMONARY INFECTION



Novartis Foundation

Copyright © Novartis Foundation 2006
Published in 2006 by John Wiley & Sons Ltd,
The Atrium, Southern Gate,
Chichester PO19 8SQ, UK

National 01243 779777
International (+44) 1243 779777
e-mail (for orders and customer service enquiries): cs-books@wiley.co.uk
Visit our Home Page on <http://www.wileyeurope.com>
or <http://www.wiley.com>

All Rights Reserved. No part of this book may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except under the terms of the Copyright, Designs and Patents Act 1988 or under the terms of a licence issued by the Copyright Licensing Agency Ltd, 90 Tottenham Court Road, London W1T 4LP, UK, without the permission in writing of the Publisher. Requests to the Publisher should be addressed to the Permissions Department, John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England, or emailed to permreq@wiley.co.uk, or faxed to (+44) 1243 770620.

This publication is designed to provide accurate and authoritative information in regard to the subject matter covered. It is sold on the understanding that the Publisher is not engaged in rendering professional services. If professional advice or other expert assistance is required, the services of a competent professional should be sought.

Other Wiley Editorial Offices

John Wiley & Sons Inc., 111 River Street, Hoboken, NJ 07030, USA

Jossey-Bass, 989 Market Street, San Francisco, CA 94103-1741, USA

Wiley-VCH Verlag GmbH, Boschstr. 12, D-69469 Weinheim, Germany

John Wiley & Sons Australia Ltd, 35 Park Road, Milton, Queensland 4064, Australia

John Wiley & Sons (Asia) Pte Ltd, 2 Clementi Loop #02-01, Jin Xing Distripark, Singapore 129809

John Wiley & Sons Canada Ltd, 6045 Freemont Blvd, Mississauga, Ontario, Canada L5R 4J3

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Novartis Foundation Symposium 279
x + 238 pages, 23 figures, 5 tables

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

ISBN-13 978-0-470-02656-4 (Hb)

Typeset in 10^{1/2} on 12^{1/2} pt Garamond by SNP Best-set Typesetter Ltd., Hong Kong
Printed and bound in Great Britain by T. J. International Ltd, Padstow, Cornwall.

This book is printed on acid-free paper responsibly manufactured from sustainable forestry, in which at least two trees are planted for each one used for paper production.

Contents

*Symposium on Innate immunity to pulmonary infection, held at the Wolfson Pavillion,
University of Cape Town Medical School, South Africa, 28–30 November 2005*

Editors: Derek J. Chadwick (Organizer) and Jamie Goode

This symposium is based on a proposal made by Siamon Gordon and Gordon Brown

Siamon Gordon Chair's introduction 1

Eric D. Bateman and **Anamika Jithoo** Lung diseases in South Africa: an
overview 4
Discussion 11

**Paul D. van Helden, Marlo Möller, Chantal Babb, Robin Warren, Gerhard
Walzl, Pieter Uys** and **Eileen Hoal** TB epidemiology and human
genetics 17
Discussion 31

David P. Speert Bacterial infections of the lung in normal and
immunodeficient patients 42
Discussion 51

Malik Peiris Pathogenesis of avian flu H5N1 and SARS 56
Discussion 60

**Claudia Montagnoli, Silvia Bozza, Roberta Gaziano, Teresa Zelante,
Pierluigi Bonifazi, Silvia Moretti, Silvia Bellocchio, Lucia Pitzurra** and
Luigina Romani Immunity and tolerance to *Aspergillus fumigatus* 66
Discussion 77

**Cecilia Garlanda, Barbara Bottazzi, Giovanni Salvatori, Rita De Santis,
Alessia Cotena, Livija Deban, Viriginia Maina, Federica Moalli, Andrea**

TB epidemiology and human genetics

Paul D van Helden, Marlo Möller, Chantal Babb, Robin Warren, Gerhard Walzl,
Pieter Uys and Eileen Hoal

Department of Medical Biochemistry/(US/MRC) Centre for Molecular and Cellular Biology and DST/NRF Centre of Excellence for Biomedical TB Research, Faculty of Health Sciences, University of Stellenbosch, PO Box 19063, Tygerberg 7505, South Africa

Abstract. The impact of tuberculosis (TB) is considerably lower than one may expect, since in the absence of immunosuppression, fewer than 10% of infected individuals will develop active disease. The relatively low proportion of individuals who progress to active disease after infection can probably be ascribed to innate resistance in most infected individuals, since vaccination using BCG or a previous episode of TB does not work reliably or effectively to confer protection in high burden parts of the world. Innate factors affecting resistance or susceptibility can be modulated by the environment and such external influences cannot be ignored. Specifically, we will address bacterial variability as well as environmental factors such as diet, smoking, helminths and hormones. We will also discuss host genes that may be involved in susceptibility or resistance at various stages of infection or disease. The discovery of as yet unknown genes impacting on TB susceptibility or disease course may lead to new insights into mechanisms of disease and novel therapies. With adaptive immunity being of little value and good TB control programmes being rare, innate resistance is still our best defence against this disease.

2006 Innate immunity to pulmonary infection. Wiley, Chichester (Novartis Foundation Symposium 279) p 17–41

It is estimated that approximately one third of the global population is infected by tuberculosis (TB). Fortunately, far fewer than 10% of those infected eventually become ill (in the absence of immunosuppression). Studies on families, twins and adoptees have suggested that there is clearly a genetic component involved in susceptibility or resistance to TB. On the other hand, recent research has shown that there are many different strains of *Mycobacterium tuberculosis* (*Mtb*) and that they can influence the course of disease. Clearly we have to deal with on the one hand a successful pathogen, but on the other hand a host that has learnt to live with this pathogen. It has been proposed that the bacterium is perhaps 10 000 years old and evolved as humans domesticated cattle and settled in villages. Recently, however, it has been suggested that the organism may be as old as 3 million years, which implies a long period of co-evolution and thus adaptation on both sides (Gutierrez

et al 2005). At least four successive epidemics of tuberculosis are thought to have occurred in the last 4000 years, viz. in the Nile Valley, Greece, the Americas (approximately 1000 years ago) and Europe. Each wave probably spans centuries, but incidence peaks probably last a few decades only.

In general, when *Mtb* is introduced into a naïve population living under harsh conditions, it may spread rapidly. Thus, in the UK with the massive migration into cities that occurred during the industrial revolution and the poor living conditions at that time, an ideal opportunity for an epidemic was created. In London, it was estimated that 20% of all deaths in 1667 were due to TB. TB peaked (possibly) in the UK around 1780 (early industrial revolution) at about 1120/100 000 p.a., or, it is estimated that 1.25% of the entire population died each year from TB. Thereafter, in England, TB started to decline years before other infectious diseases and long before the introduction of control programmes or antibiotics.

A recent mirror of this epidemic was seen in the Inuit, where, after introduction of TB to a naïve settlement, most individuals died and the epidemic rapidly waned. The waning of any epidemic may occur once the living conditions change or the population becomes more resistant, as susceptible individuals disappear. We argue that this is not due to adaptive immunity, but innate immunity which is likely to increase with exposure to the organism and subsequent removal (by death) of susceptible individuals.

Infection by *Mtb* is a complex and multistage process proceeding from the initial encounter with the pathogen. For this reason we need to imagine a multistep process (Fig. 1).

At each stage in this process, innate factors may play an important role. While there is a body of evidence that suggests there may be some immunity acquired from prior exposure to Mycobacterial species (e.g. BCG vaccination), there is also much evidence to suggest that prior infection does not necessarily confer any protection against further infection or progression to disease (Rook et al 2005, Cosma et al 2004).

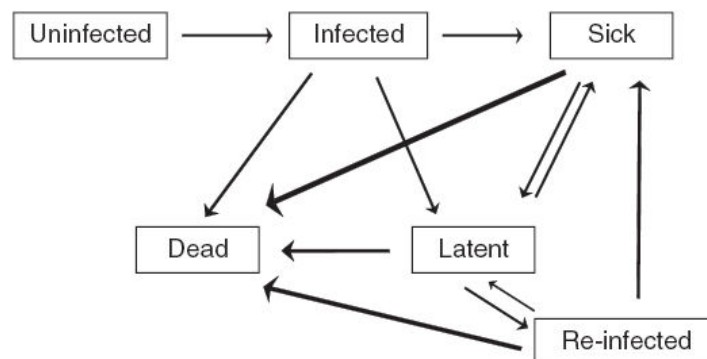


FIG. 1. Transitions in tuberculosis.

Apart from living conditions and an increased (genetic) innate resistance due to death of susceptibles, other factors can play a role: in England and much of the developed world, in the 1800s for the first time food production exceeded population growth and real wages allowed the purchase of adequate food. This is very likely one of the factors that contributed to the decline in TB after 1830, since it is known that a person who is 10% underweight has a threefold increased risk for developing TB after infection.

Excluding sociological determinants, the biological factors likely to be involved in innate resistance to *Mtb* are: (1) the bacterium, (2) the environment (nutrition and other infections), (3) the host. These will be considered below.

The bacterium

Infectious disease studies are complicated by the fact that two genomes, one prokaryotic and one eukaryotic, are interacting in an age-old contest. Genotyping studies have shown that there are thousands of different *Mtb* strains in circulation and comparative genomics has shown that the genome of *Mtb* has evolved through single nucleotide polymorphisms (SNPs), insertions and deletions. This has prompted researchers to investigate the relationship between genome variation and phenotype. A study by Tsolaki et al (2004) suggested a correlation between deletions and the severity of disease while Manca et al (2001) showed that a deletion in the pks 1–15 gene encoding production of phenolic glycolipid was responsible for an altered immune response.

Recent studies have concluded that the 'Beijing' strain is more pathogenic, causes a febrile response on infection and has a higher propensity to develop drug resistance. Furthermore, this strain induces a Th2 immune response on infection allowing for progression towards disease. This differs from the principle genetic group 2 strain CDC1551 which induces a strong Th1 response and less progression towards disease (Manca et al 2001). When the above strains were tested in an *in vivo* mouse model it was shown that the Beijing strain was more pathogenic and could outgrow the CDC1551 strain. There is ample evidence that shows that *Mtb* strains have different growth rates and prompt variable host responses, e.g. cytokine and T cell responses (Manca et al 2001, Janulionis et al 2005, Hoal-van Helden et al 2001a and 2001b). This evidence suggests that these effects are also host dependent. Despite these advances in defining different levels of pathogenicity, many mechanisms underlying these differences in the bacterium remain to be elucidated. It is hypothesised that two evolutionary scenarios may explain these observations: (1) distant evolutionary events which induce an inherited trait that is characteristic of the evolutionary lineage, and (2) recent evolutionary events which induce an inherited trait that is characteristic of a sub-population within a defined lineage.

Thus, signature polymorphisms in *Mtb* isolates may be associated with specific innate (and adaptive) reactions. Some of these also provide a growth advantage and explain the abundance of certain strain types regionally or globally.

The environment

Environmental factors that could be involved in innate resistance to TB include the infection pressure from *Mtb* in the immediate environment, nutrition, other infectious organisms, hormones (e.g. steroid hormones) and substance abuse, such as tobacco smoke (den Boon et al 2005) which suppresses macrophage activity. It is well known that TB can be associated with poverty, which in turn is associated with malnutrition, not only calorie deficit, but more importantly in the case of infectious diseases, with micronutrient imbalances.

Recent work has provided evidence that the different behaviour of omega-3 versus -6 lipids seen at the level of cells and organisms can also be detected *in vitro* in the membrane of the phagosome enclosing mycobacteria. Using phagosomal membrane actin assembly as a functional, *in vitro* readout, these studies showed that the omega-6 lipid, arachidonic acid, as well as six other pro-inflammatory lipids, could stimulate phagosome actin assembly, fusion with lysosomes and a significant increase in pathogen killing. In contrast, the addition of the omega-3 lipids, especially eicosapentanoic acid, suppresses phagosomal actin assembly and induces a significant increase in the growth of pathogenic mycobacteria (*Mtb* and *M. avium*) in macrophages (Anes et al 2003). The ability of these lipids to increase pathogen growth has also been shown to operate at the level of mice and guinea pigs, in the case of both *Mtb* and Salmonella (Paul et al 1997, Chang et al 1992). Our prediction is that dietary manipulation of omega-6 and other pro-inflammatory lipids should help to restrict the growth of pathogens within macrophage phagosomes. Even short term dietary intervention can have dramatic effects on the above-mentioned processes in animal models (Kris-Etherton et al 2002).

The gender bias in tuberculosis has never been satisfactorily explained, but may yield clues to innate resistance or susceptibility factors. Population or gender-based dietary consumption habits would influence disease prevalence. Worldwide, the same gender bias in TB disease is seen. In childhood, no significant differences are noted, but during adolescence girls experience an initially higher rate of TB, whereas in adulthood, males experience a considerably higher disease incidence. Some of this difference in adulthood may be ascribed to behavioural and cultural differences, but other factors are likely to be important. An example of this is the food consumption preferences seen between males and females (see Table 1 below). However, the gender bias may also be linked to steroid hormones, such as DHEA (dehydroepiandrosterone), which have been shown to influence the course of TB

TABLE 1 Gender-based nutritional bias of TB patients

	<i>Mean daily amount</i>	
	<i>Males (n = 23)</i>	<i>Females (n = 7)</i>
Fe (mg)	15 (6)	8.7 (8.1)
Mg (mg)	341 (346)	233 (258)
Zn (mg)	14.2 (9.4)	10.1 (6.8)
Se (µg)	71 (45)	30 (45)
Vitamin A (µg)	688 (625)	893 (500)
β-carotene (µg)	2389 (3000–6000)	3522 (3000–6000)
Vitamin B12 (µg)	12.6 (2)	4 (2)
Vitamin C (mg)	67 (75)	108 (60)
Folate (µg)	232 (320)	193 (320)

These figures based on actual food consumed while in hospital (recommended amounts are shown in brackets); see Roberts et al (2005).

disease, specifically at higher concentrations exacerbating pathology (Rook et al 1997).

The net effect of increased iron is to increase risk for active TB, as is the lower levels of some key antioxidant vitamins. The overall effect of these micronutrients and the proteins (such as NRAMP1) involved in their homeostasis (contributing to the 'ionome', Eide et al 2005) is clearly critical. For example, it is known that *Mtb* has an absolute requirement for iron, and that iron supplements should be avoided during TB disease. Furthermore, it has been shown that the total antioxidant profile (could be regarded as a general measure of iron and vitamin status) is significantly lower in TB patients than controls, however, a causative relationship has not yet been established (Wiid et al 2004).

Finally, the effect of multiple infections needs to be considered. Mathematical modelling suggests that in an area of high TB incidence and ARI (annual risk of infection), multiple infection (or super-infection) would be common (Fig. 2).

The simple probability that a particular individual will experience exactly k infection events during a stay of n years (n may be fractional) in a community where the ARI is given by p is:

$${}_n P_k = n^k p^k e^{-np} / k! \quad (n = 0, 1, 2, \dots, k \leq n)$$

Recently, superinfection or reinfection has been proven to occur frequently as predicted (van Rie et al 1999, Warren et al 2004). In an elegant experiment done in a zebrafish model with *M. marinum*, Cosma et al (2004) showed that newly infecting mycobacteria track directly to an existing granuloma harbouring bacteria from a prior infection. The net effect of this is not known, but superinfection may drive

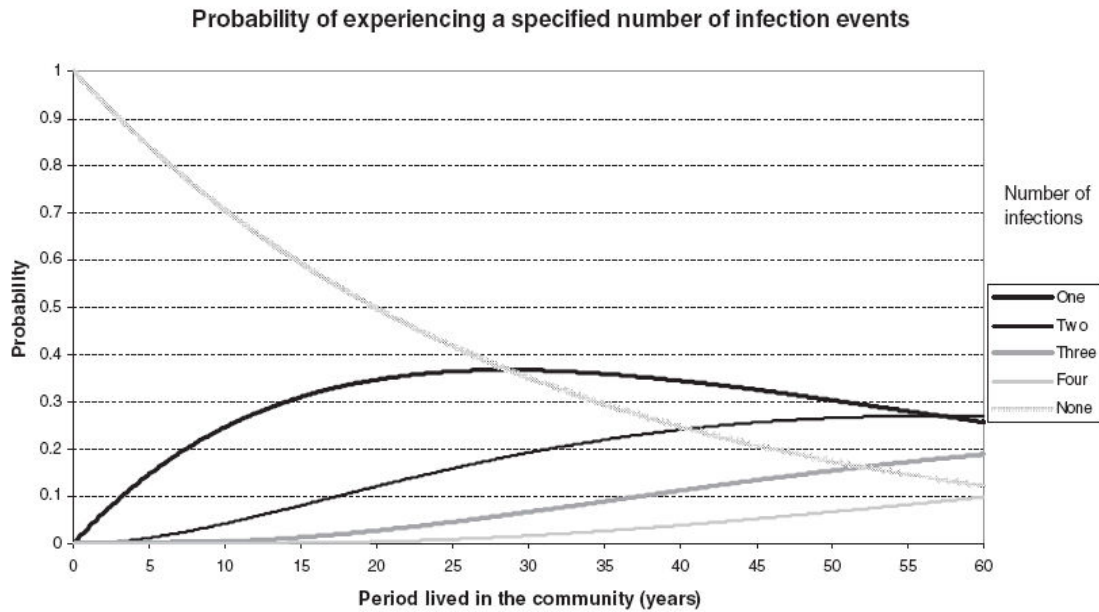


FIG. 2. Multiplicity of Infection. Graphs of ${}_n P_k$ for various k ($0, 1 \dots 4$), with n along horizontal axis (units of years) for the case of $P = 3.5\%$.

the progression from infection to active disease by either activating latent bacteria (e.g. by means of resuscitation factors, rpf) or by simply overwhelming the innate and adaptive immunity of the host. This effect may be linked to the observation that infection by high doses of mycobacteria will induce a Th2-type response (stimulation of IL4 secretion) rather than Th1, such as may occur on low dose exposure (Rook et al 2005).

Finally, the reaction to infection by *Mtb* is complicated by the intimate connection between the innate and adaptive immune systems and that in reality, most newborn humans are vaccinated with *M. bovis* BCG. Such vaccination may provide for up to 80% protection in developed countries, but far less or none in developing countries, particularly those in the tropics (Rook et al 2005). This may be linked to the mixed Th1/Th2 response in countries of the tropics, which may well be a consequence of exposure to helminths. Thus, the apparent innate response to *Mtb* infection is de facto a 'primed' response and differs in individuals according to exposure to other infecting organisms, such as environmental (myco) bacteria and helminths (Rook et al 2005).

The host

The host defences against intracellular bacteria are mainly cell-mediated but also humoral and therefore any genetic deficiencies in components that play a role in

these systems can lead to susceptibility. While there is a definite role for environmental factors, studies have indicated that genetic factors may be even more important than the environment in determining the outcome of infection. The macrophage is usually the first important cell encountering the invading pathogen, and many of the genes governing macrophage function can be expected to influence this essential first step in the innate defence system.

Investigations of TB, a multifactorial disease, have to take into account that there is likely to be an interaction between environmental factors and common polymorphisms in a number of genes. A large body of evidence points to the major role of genetic factors in the human response to a number of infectious pathogens, and these genes could also impact on treatment and vaccine efficacy. The approaches that have been used to identify the genetic component include segregation analysis, animal models and linkage analysis. Understanding the immune responses of individuals with more resistant genotypes, particularly where this can be replicated in a number of different populations, could suggest novel therapies to combat this highly successful pathogen. Complex disease, unlike monogenic conditions, can be influenced by several genes, with each gene making a small contribution to the overall susceptibility to the disease. Tuberculosis is perhaps more complex than most in that the different phenotypes or forms of the disease such as cavitary TB, pleural effusion, TB meningitis, etc. may be influenced by different genes.

Identification of common TB susceptibility genes

Complex traits such as TB can be investigated via two general designs. Firstly, family-based linkage analysis via genome-wide scanning, and secondly, population-based association studies of candidate genes.

Genome scans

The major advantage of the model-free genome scan is that novel genes may be identified. Although the phenotype is usually TB, it is possible that using intermediate phenotypes in other immune pathways could indicate as yet unsuspected genes. The first genome scan in TB was conducted on two samples of affected sibling pairs from The Gambia and South Africa, and identified two regions, on chromosome 15q and Xq (Bellamy et al 2000). The gene UBE3A in the 15q11–q13 region which encodes a ubiquitin ligase in macrophages, was subsequently associated with TB (Cervino et al 2002). A recent genome-wide scan for tuberculosis and leprosy *per se*, conducted in Brazil, found a cluster of susceptibility genes across chromosome 17q11.2 (Jamieson et al 2004) and indicated that four separate candidate genes, NOS2A, CCL18, CCL4 and STAT5B may contribute to this region of linkage.

Candidate gene association studies

Association studies can suffer from lack of reproducibility of results, and it is important that studies be done with large numbers and repeated, preferably in ethnically diverse populations. However, many of these studies have indicated genes and pathways that are important in the pathogenesis of TB (Fig. 3).

A candidate gene approach: current genes

Human leukocyte antigen

HLA-DR2 is most consistently associated with TB in many populations, including Indian, Polish, Thai, Indonesian and Russian (Lombard et al 2006). Nonetheless, inter-population variations in HLA/TB associations have been reported. HLA-DQB1*0503 was found to influence TB progression in the Cambodian population (Goldfeld et al 1998), but not in the people of the Western Cape (Goldfeld & Hoal, unpublished results). DQB1*0601 was associated with TB susceptibility in the Thai and South Indian population and the HLA haplotype DRB1*08032-DQB1*0601 was associated with genetic susceptibility to multidrug-resistant TB in Korean patients. A study of the Venda population showed an association of DRB1*1302 with TB susceptibility (Lombard et al 2006), whereas Boshoff et al (unpublished data) have shown a marginal association of DRB1*03 with TB in the South African Coloured population.

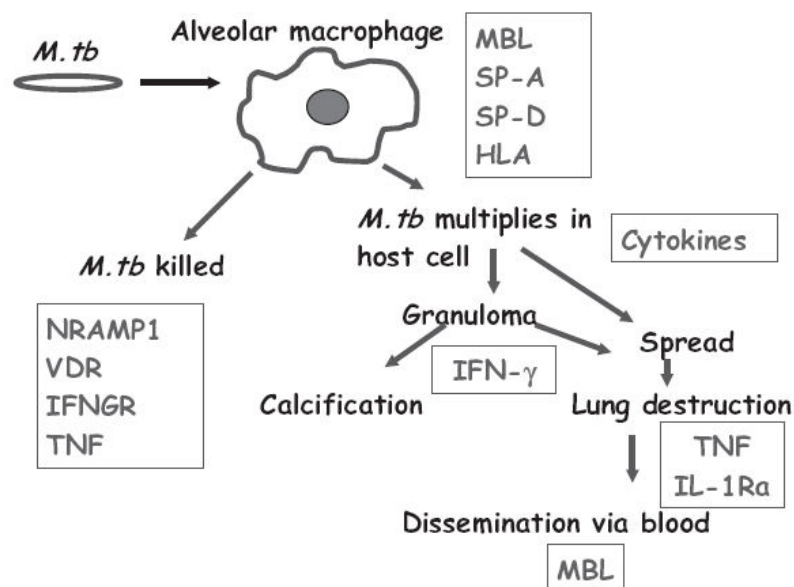


FIG. 3. A simplified representation of the TB disease process and some of the genes that may be involved at different stages.

In this context, it may be worthwhile to explore the relationship between mycobacterial strain type and genetics, e.g. HLA type. For example, in Cambodia, strain diversity is likely to be low and dominated (estimated 60–80%) by the Beijing/W strain type. In the Western Cape, Beijing type strains occur at possibly approximately 20% of total isolates (unpublished data). Therefore, it may be that HLA type is closely linked to *Mtb* strain type in a given locality, and that after extensive exposure, a skewing of HLA type may occur with concomitant resistance to certain strain types. Introduction of new strain types with new epitopes to that locality or ethnic group would then be expected to generate a new epidemic.

Natural resistance-associated macrophage protein

The Natural resistance-associated macrophage protein 1 (*NRAMP1* or *SLC11A1*) gene is a major determinant of natural resistance to intracellular infections, and was originally identified in the mouse model. It is an integral membrane protein expressed only in the lysosome of macrophages and monocytes. After phagocytosis of bacteria, NRAMP1 is targeted to the membrane of the phagosome containing the bacterium, where it may modify the environment to affect the replication of the bacterium, acting as a divalent cation pump which could remove iron or other divalent cations from the phagosome (Blackwell et al 2000).

Associations of *NRAMP1* with TB have been found in Japan, Canada, Korea, Guinea-Conakry, Vietnam, the Gambia and South Africa (Hoal et al 2004) and in most instances the allele over-represented in controls is thought to drive the highest rate of transcription of the protein. Stepwise logistic regression analysis of the South African results indicated that the 5' and 3' polymorphisms contribute separate main effects (Hoal et al 2004). More recently, it has been suggested that NRAMP1 may influence only the speed of progression from infection to disease (Malik et al 2005).

Vitamin D receptor

Vitamin D receptor (VDR) is synthesized in monocytes and activated T and B lymphocytes. Its ligand, the active metabolite of vitamin D, calcitriol, is produced in the kidney and by activated monocytes and macrophages, particularly in granulomas. Through its interaction with vitamin D, the retinoid X receptor (RXR) and the vitamin D response element (VDRE), VDR exerts several immunomodulatory effects (Selvaraj et al 2004). These include the activation of monocytes and cell mediated immunity, modulation of the Th1–Th2 host immune response, suppression of lymphocyte proliferation and restriction of *Mtb* survival in macrophages. Vitamin D deficiency is linked to TB by epidemiological evidence. It was found

that the prevalence of both vitamin D deficiency and TB was high in Asians because of their vegetarian diet and thus VDR polymorphisms should not necessarily be viewed in isolation. A particular allele was associated with female patients in an Indian population, and was found to increase susceptibility to pulmonary TB in the Gujarati population, but only in patients with a lack of serum vitamin D (Wilkinson et al 2000). The *FokI* polymorphism was associated with TB in the Chinese Han population. A large study in West Africa found no association in a case control analysis but an association was found with a particular haplotype in the transmission disequilibrium test family data (Bornman et al 2004). This dependence on the haplotype could explain many of the divergent findings on this and other genes.

Evidence of the subtlety of the effect of VDR polymorphisms in the immune response was found by Roth et al (2004) in Peru, who detected an association with time to sputum conversion in TB patients after diagnosis, but did not find a significant association with susceptibility to TB disease.

Collectins

Mannose-binding lectin. Mannose-binding lectin (MBL) is a serum lectin which acts as an opsonin to promote phagocytosis. Intracellular microorganisms may increase their infectivity by using this system, as it promotes the uptake of bacteria into macrophages where they survive. Low functional MBL-serum levels can occur because of the presence of three variant alleles which lead to an unstable protein. Low MBL levels can protect against infection with *Mtb*. This was found in case-control studies where heterozygosity for the *MBL* variant alleles was associated with protection against the disease and the B allele has also been associated with protection against TB and particularly tuberculous meningitis in South Africa (Hoal-van Helden et al 1999). Conversely, an increased susceptibility to pulmonary TB was found in homozygous carriers of the variant alleles in India and a study in Texas gave equivocal results.

Surfactant proteins (SP)-A and SP-D. Uptake of *Mtb* appears to be facilitated by SP-A and inhibited by SP-D. A Mexican population was typed for polymorphisms in both SP-A and SP-D (Floros et al 2000) and TB cases were compared with two control groups. Using multiple logistic regression analysis, an allele of SP-D was found to be associated with susceptibility to TB only when compared with the skin-test positive control group and an allele each of SP-A1 and SP-A2 was associated with TB susceptibility only when compared with the general control group. This illustrates the extreme sensitivity of association studies to definition of phenotype.

Interferon γ /IL12 pathway

Interleukin (IL)12 stimulates interferon (IFN) γ production by lymphocytes, induces type 1 helper T cell responses and is essential for resistance against infection with intracellular bacteria. It is produced by macrophages particularly when infection with intracellular microorganisms occurs. IL12 is a cytokine composed of a heavy chain (IL12B) and a light chain (IL12A). The functional response of lymphocytes to IL12 is dependent on the expression of the IL12 receptor. Any deficiency in these genes will cause a decrease in IFN γ production.

This pathway has been implicated in TB susceptibility by a wide variety of methods. In the mouse model, gene knockout experiments have indicated the importance of IFN γ , IFN γ receptor 1, and IL12 in susceptibility to mycobacterial infection. In this way, we have a gain of knowledge by loss of function. In the human equivalent, specific gene defects have been found to cause rare familial susceptibility to normally non-pathogenic mycobacteria (Casanova & Abel 2002). The first defect identified in a gene in the IFN γ pathway was the autosomal recessive IFN γ receptor ligand binding (IFN γ R1) deficiency. This resulted in an overexpression of a dominant form of the IFN γ R1, which binds IFN γ , but lacks the intracellular signalling domain. Detailed investigation of patients has led to the identification of mutations in several autosomal genes in the IFN γ pathway that can increase susceptibility to these atypical mycobacterial infections, including complete IFN γ receptor signal transduction chain (IFN γ R2) deficiency, autosomal-dominant partial deficiency of the signal transducer and activator of transcription, and autosomal-dominant partial deficiency of the IFN γ R1. An autosomal recessive IL12 deficiency was associated with BCG and *S. enteritidis* infection, as were IL12 receptor deficiencies.

At the population level, association with TB was found with a haplotype of *IL12RB1* in Japan, with an intron 2 allele and a specific haplotype in a large study in Hong Kong, and with 2 promoter polymorphisms in a family-based study in Morocco (Remus et al 2004). In a Croatian population an allele of a polymorphic microsatellite of IFN γ R1 was associated with protection against pulmonary TB, but not in a Gambian population (Newport et al 2003).

A promoter polymorphism (+874 A \rightarrow T) in the IFN γ gene itself, which appears to result in lower NF- κ B binding and lower transcription levels of IFN γ , was demonstrated in case-control studies to be associated with susceptibility to TB in Sicily, Spain and South Africa (Rossouw et al 2003). The finding in the South African population was replicated in an independent TDT study, confirming the importance of this gene in tuberculosis at the population level.

DC-SIGN

The transmembrane C-type lectin DC-SIGN (Dendritic Cell Specific Intercellular adhesion molecule [ICAM]-Grabbing Nonintegrin), or CD209, is known to be the major *Mtb* receptor on human dendritic cells. DC-SIGN was found to interact with HIV in 1992, and is now also known to be a pathogen receptor for Cytomegalovirus, Ebola, *Helicobacter pylori*, *Leishmania* and *Schistosoma mansoni*. Unifying features of all these pathogens is that they cause chronic infections that can last a lifetime, and their persistence depends on the manipulation of the Th1/Th2 balance. DC-SIGN binds strongly to mycobacteria such as *Mtb* and *M. bovis* BCG via the mannose capped cell wall component (ManLAM) of the pathogen, but does not bind to LAM that lacks the mannose cap (AraLAM). This is intriguing as ManLAM is abundant in slow growing virulent mycobacteria, such as *Mtb* and *M. leprae*, whereas AraLAM is abundant in fast growing atypical, avirulent mycobacteria, such as *M. smegmatis* and *M. chelonae*. It has been suggested that *Mtb* targets DC-SIGN both to infect dendritic cells and to down-regulate the dendritic cell mediated immune response.

We tested whether polymorphisms in DC-SIGN are associated with susceptibility to tuberculosis, and found an association between DC-SIGN promoter variation and risk of developing tuberculosis in our South African cohort. The -871G and -336A allelic combination is significantly overrepresented among healthy controls ($P = 1.6 \times 10^{-3}$) and population stratification was excluded (Barreiro et al 2006). The above allelic combination is usually confined to Eurasian populations, and it is possible that these two variants may have increased in frequency in non-African populations as a result of host genetic adaptation to a longer history of exposure to tuberculosis.

Conclusions

Genetic studies in infectious disease are usually complicated because of the presence of two different genomes and the influence their interaction can have on the disease. Although several genes have been identified as susceptibility genes for a number of intracellular bacteria, it is necessary to bear in mind that other genes and the environment can have an influence on the development of the disease, which is the reason that no single major susceptibility gene has been identified in any infectious human disease. The results from strategies used to identify candidate genes or to associate the candidate genes with infectious disease are not the final word on the subject of susceptibility, but provide important evidence on the pathways involved. A greater understanding of the immune response to TB could provide insights into novel treatments that target genetically based susceptibility, such as aerosolised IFN γ , TNF modulation, or even simple supplementation of

vitamin D. These therapies could specifically target the more vulnerable individuals in a population and lead to improved health in the entire community. It is likely that each individual has a spectrum of risk factors, which will include genes and environmental factors that will confer a risk profile on that individual. Evidence for this is work that shows that individuals who have had a prior episode have a fourfold higher risk for developing another episode of active TB than those who have never had active TB (Verver et al 2005).

Thus, it may be that innate immunity is the most important process protecting individuals against tuberculosis and by understanding this process; we may develop new ways to combat this ancient scourge.

References

- Anes E, Kühnel MP, Bos E, Moniz-Pereira J, Habermann A, Griffiths G 2003 Selected lipids activate phagosome actin assembly and maturation resulting in killing of pathogenic mycobacteria. *Nat Cell Biol* 5:793–802
- Barreiro LB, Neyrolles O, Babb CL et al 2006 Promoter variation in the DC-SIGN-encoding gene CD209 is associated with tuberculosis. *PLoS Medicine* 3:e20
- Bellamy R, Beyers N, McAdam KP et al 2000 Genetic susceptibility to tuberculosis in Africans: a genome-wide scan. *Proc Natl Acad Sci USA* 97:8005–8009
- Blackwell JM, Searle S, Goswami T, Miller EN 2000 Understanding the multiple functions of Nramp1. *Microbes Infect* 2:317–321
- Bornman L, Campbell SJ, Fielding K et al 2004 Vitamin D receptor polymorphisms and susceptibility to tuberculosis in West Africa: 1 case-control and family study. *J Infect Dis* 190:1631–1641
- Casanova JL, Abel L 2002 Genetic dissection of immunity to mycobacteria: The human model. *Annu Rev Immunol* 20:581–620
- Cervino ACL, Lakiss S, Sow O et al 2002 Fine mapping of a putative tuberculosis susceptibility locus on chromosome 15q11-13 in African families. *Hum Mol Genet* 11:1599–1603
- Chang HR, Dulloo AG, Vladoianu IR et al 1992 Fish oil decreases natural resistance of mice to infection with *Salmonella typhimurium*. *Metabolism* 41:1–2
- Cosma CC, Humbert O, Ramakrishnan L 2004 Superinfecting mycobacteria home to established tuberculous granulomas. *Nat Immunol* 5:828–835
- den Boon S, van Lill SWP, Borgdorff MW et al 2005 Association between smoking and tuberculosis infection: a population survey in a high tuberculosis incidence area. *Thorax* 60:555–557
- Eide DJ, Clark S, Nair TM et al 2005 Characterization of the yeast ionome: a genome-wide analysis of nutrient mineral and trace element homeostasis in *Saccharomyces cerevisiae*. *Genome Biology* 6:R77
- Floros J, Lin HM, Garcia A et al 2000 Surfactant protein genetic marker alleles identify a subgroup of tuberculosis in a Mexican population. *J Infect Dis* 182:1473–1478
- Goldfeld AE, Delgado JC, Thim S et al 1998 Association of an HLA-DQ allele with clinical tuberculosis. *JAMA* 279:226–228
- Gutierrez MC, Brisse S, Brosch R et al 2005 Ancient origin and gene mosaicism of the progenitor of *Mycobacterium tuberculosis*. *PLoS Pathog* 1:e5
- Hoal-van Helden EG, Epstein J, Victor TC et al 1999 Mannose-binding protein B allele confers protection against tuberculous meningitis. *Pediatr Res* 45:459–464

- Hoal-van Helden EG, Hon D, Lewis L-A, Beyers N, van Helden PD 2001a Mycobacterial growth in human macrophages: Variation according to donor, inoculum and bacterial strain. *Cell Biol Int* 25:77–81
- Hoal-van Helden EG, Stanton L-A, van Helden PD 2001b Diversity of *in vitro* cytokine responses by human macrophages to infection by *Mycobacterium tuberculosis* strains. *Cell Biol Int* 25:83–90
- Hoal-van Helden EG, Lewis LA, Jamieson S et al 2004 SLC11A1 (NRAMP1) but not SLC11A2 (NRAMP2) polymorphisms are associated with susceptibility to tuberculosis in a high incidence community in South Africa. *Int J Tuberc Lung Dis* 8:1464–1471
- Jamieson SE, Miller EN, Black GF et al 2004 Evidence for a cluster of genes on chromosome 17q11–q21 controlling susceptibility to tuberculosis and leprosy in Brazilians. *Genes Immun* 5:46–57
- Janulionis E, Sofer C, Schwander SK et al 2005 Survival and replication of clinical mycobacterium tuberculosis isolates in the context of human innate immunity. *Infect Immun* 2595–2601
- Kris-Etherton PM, Harris WS, Appel LJ 2002 Fish consumption, fish oil, omega-3 fatty acids and cardiovascular disease. *Circulation* 106:2747–2757
- Lombard Z, Brune AE, Hoal EG et al 2006 HLA class II disease associations in southern Africa. *Tissue Antigens* 67:97–110
- Malik S, Abel L, Tooker H et al 2005 Alleles of the NRAMP1 gene are risk factors for pediatric tuberculosis disease. *PNAS* 34:12183–12188
- Manca C, Tsenova L, Bergtold A et al 2001 Virulence of a *Mycobacterium tuberculosis* clinical isolate in mice is determined by failure to induce Th1 type immunity and is associated with induction of IFN-alpha/beta. *Proc Natl Acad Sci USA* 98:5752–5757
- Newport MJ, Awomoyi AA, Blackwell JM 2003 Polymorphism in the interferon-gamma receptor-1 gene and susceptibility to pulmonary tuberculosis in The Gambia. *Scand J Immunol* 58:383–385
- Paul KP, Leichsenring M, Pfisterer M et al 1997 Influence of n-6 and n-3 polyunsaturated fatty acids on the resistance to experimental tuberculosis. *Metabolism* 46:619–624
- Remus N, El Baghdadi J, Fieschi C et al 2004 Association of IL12RB1 polymorphisms with pulmonary tuberculosis in adults in Morocco. *J Infect Dis* 190:580–587
- Roberts T, Herselman M, Marais D, Labadarios D 2005 Served versus actual nutrient intake of hospitalised patients with tuberculosis as compared with energy and nutrient requirements. *South African Journal of Clinical Nutrition* 18:78–93
- Rook GA, Hernandez-Pando R 1997 Pathogenetic role, in human and murine tuberculosis, of changes in the peripheral metabolism of glucocorticoids and antiglucocorticoids. *Psychoneuroendocrinology* 22 Suppl 1:S109–113
- Rook GAW, Dheda K, Zumla A 2005 Immune responses to tuberculosis in developing countries: implications for new vaccines. *Nat Rev Immunol* 5:661–667
- Rossouw M, Nel HJ, Cooke GS, van Helden PD, Hoal EG 2003 Association between tuberculosis and a polymorphic NF- κ B binding site in the interferon γ gene. *Lancet* 361:1871–1872
- Roth DE, Soto G, Arenas F et al 2004 Association between vitamin D receptor gene polymorphisms and response to treatment of pulmonary tuberculosis. *J Infect Dis* 190:920–927
- Selvaraj P, Chandra G, Jawahar MS, Vadya Rani M, Nisha Rajeshwari D, Narayanan PR 2004 Regulatory role of vitamin D receptor gene variants of BsmI, ApaI, TaqI, and FokI polymorphisms on macrophage phagocytosis and lymphoproliferative response to mycobacterium tuberculosis antigen in pulmonary tuberculosis. *J Clin Immunol* 24:523–532
- Tsolaki AG, Hirsch AE, DeRiemer K et al 2004 Functional and evolutionary genomics of *Mycobacterium tuberculosis*: Insights from genomic deletions in 100 strains. *Proc Natl Acad Sci USA* 101:4865–4870

- Van Rie A, Warren RM, Richardson M et al 1999 Exogenous reinfection is a common cause of tuberculosis recurrence after cure. *New Eng J Medicine* 341:1174–1179
- Verver S, Warren RM, Beyers N et al 2005 Rate of reinfection tuberculosis after successful treatment is higher than rate of new tuberculosis. *Am J Respir Crit Care Med* 171: 1430–1435
- Warren RM, Victor TC, Streicher EM et al 2004 Patients with active tuberculosis often have different strains in the same sputum specimen. *Am J Respir Crit Care Med* 169:610–614
- Wiid I, Seaman T, Hoal EG, Benade AJS, Paul D van Helden 2004 Total antioxidant levels are low during active TB and rise with anti-tuberculosis therapy. *IUBMB Life* 56:101–106
- Wilkinson RJ, Llewelyn M, Toossi Z et al 2000 Influence of vitamin D deficiency and vitamin D receptor polymorphisms on tuberculosis among Gujarati Asians in west London: a case-control study. *Lancet* 355:618–621

DISCUSSION

Segal: What is known about the cellular molecular mechanisms by which the TB bacillus is killed?

van Helden: I don't know where to start in answering this question. For example, the whole notion of nitric oxide (NO) is controversial. Acidification of the phagosome is important, and *Mtb* stops that. Your question to me addresses something I have often said: we shouldn't be looking at TB patients but rather those people who are infected and don't become ill.

Segal: This is the key for understanding the effects of nutrition and IFN γ and so on. We have been doing a study on Crohn's disease, which is similar in that it is often a granulomatous lesion. We did a range of studies in humans rather than in models. We did two things relevant to TB. First, we made skin windows. Gordon Brown had one of these on his arm. You sandpaper a bit of skin off and see what cells come out. You can also apply things onto the skin and see how that can modulate what happens. In Crohn's we found there is a major defect of the accumulation of neutrophils at skin windows. We believe this to be the primary lesion. Second, we injected killed bacteria to see what effect this had on the acute inflammatory response at the site of infection. We used *Escherichia coli*, but you could use TB that had been irradiated. You might find big differences. Many of the immunological effects people are looking at in these diseases are secondary effects to the failure to clear antigen, for example. Then it is assumed that they have some causal relationship to the disease when actually they don't.

van Helden: In TB there are so many different kinds of responses. The first category of infected individuals don't even develop a granuloma and the bacilli are killed. In the next category the granuloma will form, but the bacilli will remain latent for the rest of the person's life. In the third category, the primary focus will form, and there will then be secondary foci and active disease.

Segal: In the first group, I guess one would say that the organism has been killed, digested and removed. In the second, it has been killed but not digested. In

the third it hasn't been killed at all. It would be nice to translate that into actual measurements, either in the body or test-tube.

Quesniaux: When patients are treated with anti-TNF antibodies, one of the adverse effects is the occurrence of infections (Mohan et al 2003) and TB is one of the most prominent of these (Keane et al 2001, Mohan et al 2004, Keane 2005). Here we are really acting on the second category: reactivation of previous infections.

van Helden: Perhaps the best evidence for reactivation disease was work done by Troels Lillebaek (Lillebaek et al 2002). The son developed TB 33 years after the father had had a case. It transpired that they both had exactly the same isolate and no one else in that country had the same isolate. There is no way he could have got it from anyone else, indicating that the TB is held inactivated but not dead.

Quesniaux: TB is kept under active control from the immune response.

Brown: I'd like to go back to the point Tony Segal raised about the molecular mechanism of killing. One of the big holes in TB research concerns the receptors that are involved in uptake, and how this influences the resultant response. Many of the proposed receptors don't seem to fit. For example, CR3 is not expressed in alveolar macrophages. And a recent paper shows that in a normal, uninfected individual DC-SIGN is not expressed in macrophages—it is only induced after infection (Tailleux 2005).

Quesniaux: Are there differences in DC-SIGN expression after infection by different strains of TB?

van Helden: We've never looked.

Speert: I agree completely about CR3 not being expressed in alveolar macrophages. My understanding is that cigarette smoking does up-regulate CR3. Could this be a way of tying together some of the things that we have heard? I have a comment about babies infected with BCG. The ones that get seriously ill and die are the ones with profound immunodeficiency. The other group where BCG immunization is contraindicated is chronic granulomatous disease. Reactive oxygen radicals appear to play some role in protection against tuberculosis.

van Helden: We have now found plenty of HIV-positive children with BCGosis.

Mantovani: I was intrigued by the dietary data that you showed. I remember in my medical education that the suggestion was to overfeed patients with TB. The numbers presented here were low. Were these data borne out by larger numbers?

van Helden: The hospital is very small, and there are no other data. This is a major hole in our research. The other problem with TB patients is that they don't feel well, so they don't eat well. If you measure anything in a newly diagnosed TB patient, it may not be trustworthy, because they haven't been eating well because they haven't been feeling well. But in this country, in general, they are also poor,

so they may be malnourished from poverty as well. The work done in mice involves much larger numbers and this is well substantiated. In humans nutritional work is poor.

Mantovani: One of the genetic associations you mentioned is with the IL1RA. Is this with infection or disease manifestations? In other infectious disorders, such as *H. pylori* gastritis, the manifestation is not really related to the IL1RA genotype but to the haplotype, and the balance between the pro and anti-inflammatory stimulant molecules in the IL1 system. Are there any data along these lines?

Hoal: There have been a couple of studies on IL1RA (Bellamy et al 1998, Wilkinson et al 1999). As with a number of association studies, they show slightly different things. Some show the haplotype to be associated but some show there is an effect just with the polymorphism.

Wilkinson: We showed that an extended haplotype was associated with higher delayed type hypersensitivity (DTH). The low producing IL1RA allele linked to the high-producing IL1 β associated with DTH (Wilkinson et al 1999).

Bekker: We have been doing some work in a small impoverished community with high HIV and TB rates in Cape Town, looking at Mtb strains in patients with and without HIV. We are seeing W Beijing strain much more in young people who are HIV infected. Has anyone looked at the immune competency of people infected with the W Beijing strain?

van Helden: I would predict that the Beijing in your younger age group is a reflection of an emerging strain. It has been relatively recently introduced into our society. Because it is a more able strain it spreads faster. At a Novartis Foundation Symposium in 1997 I proposed a square box scheme, where I hypothesized that you regard your strains as being represented in different parts of this box. They spread well but don't cause much disease, or they cause lots of disease but don't necessarily spread well (van Helden 1998). Beijing is going to be a position in the box. I think it is a very successful emerging strain that is spreading quickly.

Bekker: The counter to that is a set of recent data again from the community we have investigated in the south of Cape Town. In a cross sectional prevalence study in the community where we have sampled randomly, 10% of the community, it seems that people who have HIV and TB co-infection are presenting and are in the TB clinic and on treatment. People who are sitting quietly with their disease as yet undiagnosed are the older population who are not HIV infected. Immune competent people live with their disease for a long time in this population, before symptoms push them to seek healthcare, and I guess are able to spread their disease. It may be that W Beijing is the predominant strain in this scenario. The immune incompetent (HIV infected) people pick this particular strain up easily; they are the 'canaries' in the population.

van Helden: The question that we have asked is whether Beijing can go into a latency phase. I don't know. This would be interesting to address.

Wilkinson: We have looked at the proportion of strains that are Beijing in isolates from Red Cross Childrens' Hospital in Cape Town. Even within the last few years the proportion of Beijing has tended to increase. It is difficult to know what the denominator is because this is a hospital rather than a community. The second thing is that some of these strains won't grow very well in the laboratory: they appear to have a resuscitation defect. We speculate that if these strains are unable to resuscitate *in vivo*, teleologically speaking they must continue reinfecting people and therefore do best in environments where there is intense transmission.

Mizrahi: What we see depends crucially on what we can culture. Our tools are blunt, with culturability of *Mtb* strains being one of the limiting factors. The standard genotyping tool, which is based on an insertion element, is also relatively crude. However, with the introduction of higher-resolution tools such as DNA microarrays (Tsolaki et al 2004), we are now in a position to start looking at the association between the ability to reactivate and the genotype of the strain. These are studies that need to be done.

Schoub: I want to comment on the interaction between viral infections and TB. We have tried to mine data from the South African National Health Laboratory Service databank to look at seasonality, and correlate it with, for example, reactivation disease. One can speculate on the mechanism but do you have any comment on the actual observation?

van Helden: We don't have any data.

McGreal: What do we actually know at a structural and functional level when we talk about strain difference in *Mtb*? And also at a functional level what impact does strain difference have in terms of the immune response and possible immune evasion strategies. You indicated some associations with different strains, but how much do we know about that?

van Helden: Very little. We are only now developing microarray technology to answer those questions. Part of the problem is choosing the genome to base the microarray on. If you take your standard H37Rv, it has lots of deletions compared to many of our clinical strains. We need to take multiple strains and fully sequence them, and there are now five *Mtb* genomes available. The Broad Institute in Cambridge, MA, may sequence another 10. We should develop a microarray based on everything we can find.

McGreal: When you said that certain strains are prevalent in certain areas and populations, is that a result of local susceptibility? If you were to take a different genetic scheme would you see the same susceptibility?

van Helden: That's something we want to look at.

Mizrahi: It is important to mention that certain studies have revealed that the associations between strain families of *Mtb* and their human host populations are stable (Hirsh et al 2004, Baker et al 2004). This finding speaks to the issue of TB being an ancestral pathogen of humankind. The associations deduced from studies

of tuberculosis in immigrant populations imply that particular lineages of *Mtb* may be highly adapted to certain human populations, and less well adapted to others. This notion is supported by the results of a new and larger study published recently by Peter Small's group (Gagneux et al 2006).

van Helden: On the other hand it is to some extent because the immigrant populations stick together. It is a big confound.

Steinman: Can we have a little more discussion about DC-SIGN? First of all, we find it difficult to detect on most human blood leukocytes. Does anyone have experience otherwise and does it change in patients? We find that DC-SIGN is expressed on the small subset of myeloid dendritic cells (DCs) in blood following culture in IL4 (Graneli-Piperno et al 2006). It would be striking if you suddenly saw DC-SIGN increase in TB, since we are all influenced by the papers (Geijtenbeek et al 2003) saying that ligation of DC-SIGN by lipoarabinans from mycobacteria can block the maturation of DCs. The second thing is, DC-SIGN is seen on the monocyte-derived DCs. This is where it was described, and this seems to be due to induction by IL4. If you add IL4 to a monocyte, in a day you will have lots of DC-SIGN. It is not yet clear what monocyte-derived DCs correspond to *in vivo*. The third thing is what happens in the lymphoid organ. We have found that DC-SIGN is abundant on the macrophages in the lymph node medulla. It is not detectable among most DCs in the T cell area in apparently normal lymph nodes. Because of what has been shown *in vitro* with DC-SIGN we should look at it much more assiduously in disease states.

Lambrecht: We have looked at lung DCs from human lavage samples and DC-SIGN is not found on these either.

Gordon: There are other mannose recognition lectins on macrophages.

Brown: The ability of DC-SIGN to modulate the DC function is very likely. We have shown that signalling through another lectin, Dectin-1 can recruit Syk resulting in the induction of IL2 and IL10.

Steinman: That was the message of Geijtenbeek et al (2003). However as mentioned, the sites of DC-SIGN expression *in vivo* are not clear. We are all assuming that it is present on all DCs and exclusively DCs, but neither seems to be the case.

Brown: A lot of these DC receptors are not DC specific. This is another growing theme.

van Helden: How did you look for the expression?

Steinman: By monoclonal antibodies.

Gordon: What about the other genetic factors? You went over them quite quickly as though none of them really mattered, or all of them mattered a little.

van Helden: I don't think anyone has found a major genetic factor.

Mayosi: Do we know the population-attributable risk for any of the genetic factors?

Hoal: It is always small. This is the issue with genetic studies of infectious diseases or other complex diseases. We will find a number of genes that are important but only have a small contribution. I don't know whether you could have the situation where someone's susceptibility rises significantly if they have a combination of a number of these susceptibility alleles. We just don't know that yet. The field is still in the phase of finding the genes, replicating them in a second population, and working out what sort of significance they have.

Gordon: Do you think it is fair to talk about the Bronte family and genetics? How do you separate out the environmental factors and the genetic ones?

van Helden: Apparently, in the Bronte family the father had chronic TB. He died in his 80s, but I think he infected the rest and they all died young.

Gordon: Why does this indicate genetic susceptibility?

van Helden: Other families have a case but not all succumb to disease.

Lambrecht: One clinical problem we run into occasionally is elderly women with atypical mycobacteria in the middle lobe. People always claimed that there were anatomical problems with the middle lobe, but isn't there also a genetic predisposition to develop atypical mycobacterial disease in the lungs?

van Helden: I have no idea. The problem with that sort of thing is power. We looked at our first 2000 isolates, and out of these we found 60 that were attributable to MOTTs. Only now have we put some effort into trying to find this out. But if we only have 60 cases out of 2000, we don't have enough genetic power to really examine this.

Lambrecht: It would be easy to do a candidate gene approach with 60 patients.

van Helden: The populational attributable risk is quite small per gene. It won't be a monogenic effect like the IFN γ receptor defects in the Maltese kindred, for example. You could be right, but I doubt it.

Wilkinson: There was a study done by the British Thoracic Society that looked at susceptibility to *Mycobacterium malmoense* and this showed a weak association with VDR promoter polymorphisms (Gelder et al 2000). However, the predominant factors that predispose people to atypical mycobacterioses are damaged lungs and immunosuppression.

Finn: I know that in this meeting we are focused on the innate immune system. You mentioned several times that this genetic difference may signify that someone's innate immunity is stronger than other people's. One of the important roles of the innate immune system is to jump-start the adaptive immune system and generate good memory responses that will protect you for the rest of your life. The question is, how good is the memory response to the bug in those people who do not get reinfected and those who are protected under high risk conditions? And what types of effector mechanisms are involved that would instruct us in terms of generating that type of immunity through vaccination to protect the rest of the population?

van Helden: This is an important issue, but we haven't done any work on this. There is old work which suggests that the partial protective effect of infection varies between 16 and 18%. Unfortunately, I am not qualified to comment directly on how good that work is. It is based on mathematical modelling and epidemiology. I am sceptical about the interpretation. The tool that is used is generally the skin test, which is extremely blunt. Half of the T cell researchers would say that there is partial immunity. Then there are others who say that it is impossible, because reinfection is occurring and if you are reinfected you have a four times higher likelihood to progress to disease, so where is your partial immunity. To reconcile these two schools of thought I would like to suggest that you increase your categories of individuals. You have categories of individuals that will develop partial immunity and another category that won't.

Finn: Among those that develop partial immunity there will be some that will still be susceptible.

Brown: Is it going to be possible to make vaccine for TB, given the occurrence of reinfection?

van Helden: Many think so, since it could be cell-mediated immunity.

Finn: With regard to the comment you made about a low dose of BCG maintaining better protection, to an immunologist this immediately says that it generates higher affinity and avidity T cells and so on.

van Helden: The New Zealanders say that the dose and timing of repeat BCG is critical. The Irish say that a high-dose BCG vaccine given once is just as good.

Walz: I can't think of any pathogen where a vaccine provides better protection than the natural infection. But there is always a first time.

Steinman: Cowpox isn't natural for humans, but it worked as a smallpox vaccine. Immunologists just haven't got onto the vaccine scene, but when we do, I think that we can do better than complex microbial vaccines!

Finn: As you mentioned, some pathogens have evolved a symbiosis with the host, and a vaccine might do a better job eliciting immunity than the bugs.

Ryffel: Going back to Valerie's comment on the anti-TNF antibody induced reactivation of chronic/latent infection, we are able to model reactivation or tuberculosis infection in mice (Botha & Ryffel 2003). By administration of neutralizing TNF antibodies or soluble TNF-R we are able to reactivate chronic infection (unpublished). Further we demonstrated that membrane TNF provides a partial protection to infection (Fremond et al 2005) suggesting that neutralising exclusively soluble but not membrane TNF may reduce the risk of reactivation of TB infection.

Quesniaux: An individual who has had a first infection with TB and has had this under control for 20 years will have mounted a very efficient T cell memory response. If you now come along with anti-TNF, this could all go wrong with the TB being reactivated.

Finn: I disagree with your statement that this person has mounted a good memory response. This is an effector memory-type response, something that is continuously protective. It is not that good deep central memory that results from eliminating the pathogen. In the complete absence of antigen you deeply bury your central memory response that can then be reactivated on subsequent infection. You cannot say that reactivation of the bug by TNF means that it has now defeated a well established memory response.

Quesniaux: How do we explain the effect of anti-TNF?

Steinman: TNF has many effects on DCs. In the literature on NOD mice, people have tried to manipulate the onset of that spontaneous autoimmune disease by manipulating TNF levels. There is now evidence that the DC is responsible. The approach is to block TNF early in life, which then reduces the severity of the disease. If you then target antigens to the DCs in the TNF blocked mice, you can see changes in antigen presentation. Therefore, the DCs may be changing as a result of TNF manipulation.

Finn: There is constant cross-talk between the innate and the adaptive immune system. It doesn't just go from the innate to the adaptive and then stop, it goes back as well. You have a wonderfully stimulated memory CD4 response specific for the pathogen. You can arm your macrophage that is infected with a pathogen to make more IFN γ by the interaction with the CD4⁺ T cell. If you have a good adaptive immune memory, it can then stimulate much higher activity of your innate system when the new infection comes. If you have a very good CD4⁺ T cell memory response, that macrophage will be much more effective in destroying the pathogen.

van Helden: In thinking about vaccination and protective immunity, we have to remember that there is evidence that BCG does work. My feeling is that if there is no immunosuppression, 90–95% of people infected with TB will not become ill. BCG is not going to do anything for them anyway. It is a small part of the population in which vaccination can have an effect.

Finn: I would like to argue against this point. The 95% of the people will not develop disease unless they are immunosuppressed in some way, unless they live a long life and their immune systems become old. But if you generate a strong immune memory early in life that immune memory is a bit more protective from what happens to that person later in life, who therefore, if healthy and well-nourished will not be susceptible to the disease. Generating a strong and effective immune memory early in life will protect many more people than the 5% that you think will be protected by the vaccine.

Steinman: The point is how you assay memory. This is an evolving field. We used to say that HIV-infected people have good memory, because you could easily detect CD8⁺ T cells in them that make IFN γ in response to HIV antigens. By current criteria, though, these CD8 cells are dysfunctional. They don't grow in response

to antigen. From the same individuals, CMV, EBV and influenza specific cells grow fine (Arrode et al 2005). There is a dysfunction here of the HIV specific CD8⁺ T cells, and they are termed 'helpless' memory cells. How we assay memory is very important.

Finn: The bottom line is that we know a lot more than last year and 10 years ago. What we know speaks much more in favour of the vaccine being able to do this better than the natural disease.

Steinman: Then there are suppressor cells. What you described in terms of reinfection would fit what we know about suppressor cells. These cells are triggered by antigen, but then they can suppress other immune responses to other antigens, particularly when they are presented by the same presenting cell. If you respond to mycobacteria A and make a suppressor cell specific for peptides from protein A, it will block the immune response to other mycobacteria proteins, B. That is, as long as the antigen presenting cells are presenting both A and B, the suppressor cells for A will block presentation of B. However, it is still not straightforward to measure suppressor cells currently in human. This is a big gap.

Hoal: I want to return to the vaccination and susceptibility story. This is mouse work that I don't know very well, but I heard that some mouse strains that are more susceptible to mycobacteria are also less likely to be protected by vaccination. This is quite worrying if the same were to occur in the human population. We have to be careful when we do vaccination trials in humans to look at this susceptible portion of the population.

Gordon: Bernard Ryffel, do the mouse models allow you to get at some of these questions?

Ryffel: It is well established that the susceptibility to TB infection differs among mouse strains, B6 mice are typically resistant, while Balb/c mice are more susceptible; the genomic analysis will certainly provide polygenic resistance loci, which are currently unknown. What is really amazing is that mice with complete ablation of Toll-like receptor/MyD88 signalling are able to have preserved T cell response to TB antigens. But this T cell response is not protective, as the innate immune response is profoundly defective (Fremont et al 2004). However, BCG vaccination of MyD88 deficient mice provides a short-term, but not long-term protection.

Walzl: There are mouse models concerning regulatory cells. Strangely enough, people examined induction of regulatory cells by *Mtb* in conjunction with the hygiene hypothesis and the allergic response. There are several studies showing *Mtb* in circulatory cells that suppress allergic airway responses. It is strange that people look at it that way; the logical conclusion would be that those responses are there because *Mtb* wants to promote itself.

Lambrecht: This could explain why the low dose BCG works better than the high dose. If you give high doses of freeze-inactivated BCG it is a strong inducer of T_{regs}. There are moieties within the BCG which have the ability to induce

T_{regs}. If they are to the advantage of the bug, this would explain why high levels of BCG would lead to greater susceptibility. It is a balance between protecting from an over-zealous response and on the other side inducing immunity.

E Sim: Am I right in thinking that you can distinguish between reactivation and reinfection by the same strain?

van Helden: No, we can't.

Ryffel: Do you have more evidence that there is coinfection? If this is the case, with two strains, what is the prevalence? What is the contribution of each?

van Helden: We have no quantitative data.

Wilkinson: There is an interesting recent paper showing three cases of TB, all of which are pulmonary, and have another extrapulmonary site (Garcia de Viedma et al 2005). For each of the three extrapulmonary sites there is a different strain. The authors suggest that the extrapulmonary strain has greater virulence than the pulmonary strains when they infect macrophages together.

References

- Arrode G, Finke JS, Zebroski H, Siegal FP, Steinman RM 2005 CD8+ T cells from most HIV-1 infected patients, even when challenged with mature dendritic cells, lack functional recall memory to HIV gag but not other viruses. *Eur J Immunol* 35:159–170
- Baker L, Brown T, Maiden MC, Drobniowski F 2004 Silent nucleotide polymorphisms and a phylogeny for *Mycobacterium tuberculosis*. *Emerg Infect Dis* 10:1568–1577
- Bellamy R, Ruwende C, Corrah T, McAdam KP, Whittle HC, Hill AV 1998 Assessment of the interleukin 1 gene cluster and other candidate gene polymorphisms in host susceptibility to tuberculosis. *Tuber Lung Dis* 79:83–89
- Botha T, Ryffel B 2003 Reactivation of latent tuberculosis infection in TNF-deficient mice. *J Immunol* 171:3110–3118
- Fremont CM, Yeremeev V, Nicolle DM, Jacobs M, Quesniaux VF, Ryffel B 2004 Fatal *Mycobacterium tuberculosis* infection despite adaptive immune response in the absence of MyD88. *J Clin Invest* 114:1790–1799
- Fremont C, Allie N, Dambuza I et al 2005 Membrane TNF confers protection to acute mycobacterial infection. *Respir Res* 6:136
- Gagneux S, DeRiemer K, Van T, Kato-Maeda M et al 2006 Variable host-pathogen compatibility in *Mycobacterium tuberculosis*. *Proc Natl Acad Sci USA* 103:2869–2873
- Garcia de Viedma D, Lorenzo G, Cardona PJ et al 2005 Association between the infectivity of *Mycobacterium tuberculosis* strains and their efficiency for extrarespiratory infection. *J Infect Dis* 192:2059–2065
- Geijtenbeek TB, Van Vliet SJ, Koppel EA et al 2003 *Mycobacteria* target DC-SIGN to suppress dendritic cell function. *J Exp Med* 197:7–17
- Gelder CM, Hart KW, Williams OM et al 2000 Vitamin D receptor gene polymorphisms and susceptibility to *Mycobacterium mageritense* pulmonary disease. *J Infect Dis* 181:2099–2102
- Granelli-Piperno A, Shimeliovich I, Pack M, Trumpfheller C, Steinman RM 2006 HIV-1 selectively infects a subset of nonmaturing BDCA1-positive dendritic cells in human blood. *J Immunol* 176:991–998

- Hirsh AE, Tsolaki AG, DeRiemer K, Feldman MW, Small PM 2004 Stable association between strains of *Mycobacterium tuberculosis* and their human host populations. *Proc Natl Acad Sci USA* 101:4871–4876
- Keane J 2005 TNF-blocking agents and tuberculosis: new drugs illuminate an old topic. *Rheumatology (Oxford)* 44:714–720
- Keane J, Gershon S, Wise RP et al 2001 Tuberculosis associated with infliximab, a tumor necrosis factor alpha-neutralizing agent. *N Engl J Med* 345:1098–1104
- Lillebaek T, Dirksen A, Baess I, Strunge B, Thomsen VO, Andersen AB 2002 Molecular evidence of endogenous reactivation of *Mycobacterium tuberculosis* after 33 years of latent infection. *J Infect Dis* 185:401–404
- Mohan AK, Cote TR, Siegel JN, Braun MM 2003 Infectious complications of biologic treatments of rheumatoid arthritis. *Curr Opin Rheumatol* 15:179–184
- Mohan AK, Cote TR, Block JA, Manadan AM, Siegel JN, Braun MM 2004 Tuberculosis following the use of etanercept, a tumor necrosis factor inhibitor. *Clin Infect Dis* 39:295–299
- Tailleux L, Pham-Thi N, Bergeron-Lafaurie A et al 2005 DC-SIGN induction in alveolar macrophages defines privileged target host cells for mycobacteria in patients with tuberculosis. *PLoS Med* 2:e381
- Tsolaki AG, Hirsh AE, DeRiemer K et al 2004 Functional and evolutionary genomics of *Mycobacterium tuberculosis*: insights from genomic deletions in 100 strains. *Proc Natl Acad Sci USA* 101:4721–4722
- van Helden PD 1998 Bacterial genetics and strain variation. In: *Genetics and tuberculosis*. Wiley, Chichester (Novartis Found Symp 217) p 178–194
- Wilkinson RJ, Patel P, Llewelyn M et al 1999 Influence of polymorphism in the genes for the Interleukin 1 Receptor Antagonist and Interleukin IL-1 β on tuberculosis. *J Exp Med* 189:1863–1874