INVESTIGATING THE ROLES OF *CTSZ*, *MC3R* AND *MC4R* IN HOST SUSCEPTIBILITY TO TUBERCULOSIS

by

Lindsey Adams

Supervisor: Prof. Eileen Hoal

Faculty of Health Sciences

Department of Biomedical Sciences

Division of Molecular Biology and Human Genetics

December 2010

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own original work, and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

December 2010

SUMMARY

Tuberculosis (TB) is an infectious disease which has plagued society for thousands of years. Despite public health programs, anti-TB drugs and a vaccine, the absolute numbers of people infected with TB each year continue to rise as populations grow. The high TB-burden areas are also plagued by other debilitating factors including HIV/AIDS infection, poverty and malnutrition. Nutrition has been implicated in TB susceptibility in a number of studies. While most are observational reports made during times of war, famine or natural disaster, multiple studies provide convincing evidence for poor nutritional status increasing the morbidity and mortality of TB.

Numerous approaches are currently utilized in TB research, and there has been convincing evidence to support the role of host genetics in TB susceptibility. Based on previous linkage studies and a search of current literature, three genes were selected for this case-control study. Subsequently, variations located in cathepsin Z (*CTSZ*), melanocortin 3 receptor (*MC3R*) and melanocortin 4 receptor (*MC4R*) were genotyped in the South African Coloured (SAC) population to determine the existence of an association with TB disease.

CTSZ is a lysosomal cysteine protease expressed in cells of the immune system. Interaction between this 303 amino acid protein and β_2 integrin receptors lymphocyte function-associated antigen-1 (LFA-1) and macrophage antigen-1 (MAC-1) leads to altered lymphocyte proliferation. As a result, a single exonic variant in *CTSZ*, rs34069356, the same identified in a previous linkage study, showed strong evidence for association with TB susceptibility in cases (n = 410) and controls (n = 301) in the SAC population (p< 0.0001).

MC3R and *MC4R* are two of 5 melanocortin receptors. *MC3R* has been found to be a key regulator in energy expenditure and host metabolism while activation of *MC4R* leads to a decrease in food intake. Activation of these two receptors is regulated by leptin, a hormone released by adipose tissue. A variant located upstream of the *MC3R* gene, rs6127698, also showed evidence of disease association with the less frequent allele, T, being under-represented in cases (n = 540) compared to controls (n = 541) (genotypic frequency, p = 0.0039), suggesting a possible resistance phenotype. Functional analysis of this variant revealed an increase in *MC3R* expression when stimulated with BCG, with individuals homozygous for the T allele exhibiting an even larger upregulation of *MC3R* expression than individuals homozygous for the G allele, though this difference was not statistically significant. A single haplotype in *MC3R* was found to be associated with TB susceptibility (p = 0.0008) and this association remained after permutation testing to correct for multiple testing (p = 0.0061)

Three variants were selected for genotyping in *MC4R* and while none of these showed a statistically significant difference between cases (n = 510) and controls (n = 487), this gene should not be ruled out as both *MC3R* and *MC4R* have been found to work closely though not redundantly and double knockout experiments result in exacerbated obesity, suggesting that these proteins have a synergistic effect.

The results of this study support both a role of host genetics and nutritional status in TB and strongly motivate further research in both of these fields.

OPSOMMING

Tuberkulose (TB) is 'n aansteeklike siekte wat reeds vir eeue die gesondheid van die publiek bedreig. Ten spyte van publieke gesondheidsprogramme en verskeie anti-TB medikasie middele, groei die aantal van mense wat hiermee ge-infekteer word steeds jaarliks. Dit is veral in areas waar TB steeds groei, waar ook ander neerdrukkende faktore soos HIV/Vigs, armoede en wanvoeding hoogty vier. Na aanleiding van verskeie verslae tydens oorloë, hongersnood en ander natuulike rampe is dit veral duidelik dat swak nutriënt inname morbiditeit en sterftes wat met TB gepaard gaan verhoog.

Talle benaderings word tans gebruik in TB-navorsing, Bewyse is oortuigend om die rol van genetika van die gaheer met vatbaarheid vir TB te verbind. Op grond van vorige studies en die huidige literatuur, het ons drie gene gekies vir hierdie pasiënt-kontrole studie. Variante geleë in cathepsin Z (*CTSZ*), melanocortin 3 receptor (*MC3R*) en melanocortin 4 receptor (*MC4R*) is ge-genotipeer in die Suid-Afrikaanse Kleurling bevolking (SAK) (540 gevalle en 540 kontrole) om sodoende die assosiasie met TB te bepaal.

CTSZ is 'n lisosomale sisteïen protease wat uitgedruk word in immuunselle. Interaksie tussen hierdie 303 aminosuur protein en β_2 integrin reseptore nl. LFA-1 en MAK-1 bring veranderde limfosiet proliferasie mee. 'n Enkele eksoniese variant in *CTSZ*, rs34069356, dieselfde soos ge-identifiseer in 'n vorige studie, verskaf sterk bewys vir assosiasie met TB vatbaarheid in gevalle (n = 410) en kontrole (n = 301) in die SAK bevolking.

MC3R en *MC4R* is twee van 5 melanokortien reseptore. Daar is gevind dat *MC3R* 'n sleutelrol speel in die energie regulering van gasheer metabolisme, terwyl die aktivering van *MC4R* eindelik lei tot 'n afname in voedsel inname. Aktivering van hierdie twee reseptore word gereguleer deur Leptien, 'n hormoon wat vrygestel word deur adipose weefsel, 'n Variant, stroomop geleë vanaf *MC3R,* rs6127698, is ook bewys om met TB ge-assosieer te wees, met die T-alleel meer seldsaam in gevalle (n = 540) as in kontroles (n = 541) wat dui op 'n moontlike weerstandsfenotipe. Funksionele analise van hierdie variant onthul 'n toename in *MC3R* uitdrukking wanneer gestimuleer met BCG, met individue homosigoties vir die T-alleel wat selfs groter opregulation veroorsaak wanneer vergelyk word met individue homosigoties vir die G allele. Hierdie resultaat was egter nie statisties beduidend nie. 'n Enkele haplotiepe in *MC3R* is ge-assosieer met TB vatbaarheid en die assosiasie is onveranderd nadat 'n permutasie korreksie aangebring is (p = .0061).

Voorts is drie variante gekies vir genotipering in *MC4R* en ten spyte daarvan dat nie een daarvan 'n statisties beduidende verskil getoon het tussen pasiënte (n = 510) en kontroles (n = 487) nie, behoort hierdie geen nie uitgesluit word nie, Die rede hiervoor is dat beide *MC3R* en *MC4R* verskeie kere gevind is om in samewerking 'n rol te speel om vetsug te voorkom of te vererger.

Die resultate van hierdie studie beaam beide 'n rol van gasheer genetika en voedingstatus in TB en motiveer veral verdere navorsing in beide van hierdie vakgebiede.

This thesis is dedicated to my parents, Ross and Jennifer Adams

ACKNOWLEDGEMENTS

Imagination is more important than knowledge. Knowledge is limited. Imagination encircles the world.
Albert Einstein.

There are many people I would like to thank for their support and contribution to this work:

To my promoter prof Eileen Hoal-van Helden and co-promoter prof Paul van Helden. Thank you for taking a chance on me. It was overwhelming to begin a new chapter in life in a foreign country and you both made the transition much more manageable. Your support and encouragement throughout the duration of my studies will never be forgotten.

To Monster, without whom this thesis might have been much more chaotic and fluffy. Thank you for taking on the enormous task of reading (and re-reading) my work and helping me to shape it into what is presented here. Your comments and criticisms pushed me to become a better writer and a better scientist.

To Don Hawyard, Ruben van der Merwe and Zhou Fang for their constant assistance with all the functional work even though most never made it into this thesis. Thank you for putting up with your constantly disappearing lab equipment. Now that I`ve departed from South Africa, hopefully your stuff will be where you left it.

To my dear host genetics group: Marlo, Marna, Nikki, Muneeb, Chandré, Lance and Khutso. Thank you for putting up with me and my antics. You all made my time in South Africa hilariously memorable. I truly hope the hostgen elf eases up on his troublesome laboratory pranks and allows you to finish up your work with minimal complications.

To the rest of my friends in the Division of Molecular Biology, especially Erika de Wit, Andre Loxton, Kim Stanley and Carrie Kirsten. Thank you for all of your advice but even more so thank you for the distractions you provided when I needed them the most. Without you all I might have had to take things slightly more seriously.

Thank you to all the individuals who gave consent to take part in this study. Special thanks to all my honour student volunteers for agreeing to take part in the functional study.

To Jessica Chvatal, thank you for all of your encouragement and helping me stay focused on my studies. All our adventures gave me just the motivation I needed to endure those late nights and long weekends spent in the lab.

A special thank you to my parents, Ross and Jennifer Adams, without whom none of this would have been possible. Without your love and your support (and your bank account), I would never have had the opportunity to travel so far from home. No child could ask for better parents. You have given me the courage to face the world and allowed me to be just delusional enough to think that I can accomplish anything. Hopefully this thesis will provide some evidence that I did learn a few things while I was away. Although at times I found it all a little overwhelming, knowing that I had you both worrying about me the entire time I was away made me know I was never alone. Thank you.

List of Abbreviation

AIDS	acquired immune deficiency syndrome
ARMS	amplification refractory mutation system
BCG	bacillus Calmette-Guérin
BMI	body mass index
bp	base pairs
C	cytosine
CAF	Central Analytical Facility
cAMP	cyclic adenosine monophosphate
CI	confidence interval
cM	centimorgan
CTSP	cathepsin P gene
CTSX	cathepsin X gene
CTSZ	cathepsin Z gene
CTSZ	cathepsin Z protein
ddH$_2$O	double distilled water
DM	diabetes mellitus
DNA	deoxyribonucleic acid
dNTP	deoxynucleoside triphosphate
DOTS	directly observed treatment and short-course drug therapy
EDTA	ethylenediaminetetraacetic acid
g	gram
G	guanine
GPCR	g-protein coupled receptor
HIV	human immunodeficiency virus
HWE	Hardy-Weinberg equilibrium
kb	kilobases
LD	linkage disequilibrium
LFA-1	lymphocyte function-associated antigen-1
LTBI	latent tuberculosis infection
M. tuberculosis	*Mycobacterium tuberculosis*
MAC-1	macrophage antigen-1
MC3R	melanocortin 3 receptor gene
MC3R	melanocortin 3 receptor protein
MC4R	melanocortin 4 receptor gene
MC4R	melanocortin 4 receptor protein
MCR	melanocortin receptor
MDM	monocyte-derived macrophages
MDR	multidrug-resistant
MgCl$_2$	magnesium chloride
MSH	melanocyte stimulating hormone
ml	milliliter
mm	millimeter
mM	millimolar
µg	microgram
µl	microliter
µm	micrometer
NAAT	nucleic acid amplification test

nM	nanomolar
OADC	oleic-albumin-dextrose-catalase
OD	optical density
OR	odds ratio
p	probability value
PBMC	peripheral blood mononuclear cells
PCR	polymerase chain reaction
PTM	post translational modification
POMC	pro-opiomelanocortin
POW	prisoner of war
PPD	purified protein derivative
qPCR	quantitative polymerase chain reaction
RT	room temperature
RNA	ribonucleic acid
s	second
SAC	South African Coloured
SB	sodium borate
T	thymine
TB	tuberculosis
TST	tuberculin skin test
USA	United States of America
UTR	untranslated region
UV	ultra-violet
V	volts
WGA	whole genome amplification
WHO	World Health Organization
XDR	extreme drug-resistance
°C	degrees Celsius

List of Figures

List of Tables

List of Publications

Lindsey A Adams, Marlo Möller, Almut Nebel, Stefan Schreiber, Lize van der Merwe, Paul D van Helden, Eileen G Hoal: **Polymorphisms in *MC3R* promoter and *CTSZ* 3`UTR are associated with tuberculosis susceptibility** (Currently under review)

CHAPTER 1: Introduction

A. Setting the Stage

1.1 Tuberculosis

Once thought to be near elimination, *Mycobacterium tuberculosis (M. tuberculosis)* has re-emerged as a foremost concern in developing countries. In a report published by the World Health Organization (WHO) in 2009, it was estimated that tuberculosis (TB) was responsible for 1.3 million deaths in 2008 alone [1]. Despite the existence of a vaccine, numerous antimycobacterial drugs and treatment supervision programs, it is predicted that there will be 9.8 million new cases of TB in 2010, the highest in history [2]. Approximately one third of the world's population is currently infected, with the majority of infections occurring in developing countries. Progression of TB is influenced by numerous factors including genetics, stress, malnutrition, poor sanitation and crowded living conditions [3]. While several genes have been implicated as having a role in TB infection, researchers have not yet determined how *M. tuberculosis* is able to evade the host immune system and persist for years as a latent infection.

Only 10% of those infected with *M. tuberculosis* will go on to develop active disease but factors affecting how this occurs have yet to be established [4]. In order to gain control of this high burden disease and develop more effective prevention and treatment strategies, researchers must gain insight into the pathogenic mechanisms of *M. tuberculosis*.

1.2 History

TB is a disease that has plagued society for thousands of years. While effective public health programs have dramatically reduced TB morbidity in the Western world, TB remains one of the biggest killers in developing countries [5]. Epidemiological data suggests it was during the establishment of social networks consisting of about 200 to 440 people, approximately 10 000 years ago, that *M. tuberculosis* found its niche, quickly becoming endemic in human populations [6]. The migration of populations in Europe, the expansion of cities and the rise in population density during the 1600s provided optimal growing conditions for *M. tuberculosis* [6]. As living conditions worsened, TB spread and during the 18th and 19th centuries TB became the leading cause of death in Western Europe [6]. It is estimated that one quarter of Europe's population died of TB in the 19th century and TB

mortality did not decrease until sanitation and living conditions improved [7]. When Europeans emigrated to the New World in the 1800s, large urban areas like Boston and New York experienced TB mortality rates of almost 700 out of 100 000 [7].

The discovery of antibiotics effective against TB and the use of the bacillus Calmette-Guérin (BCG) vaccine resulted in a brief decline in TB morbidity and mortality in the 20th century. This decline caused health professionals to consider TB to be a disease of the past and research funding was decreased until 1972 at which point it was eliminated entirely from the budget of the Center for Disease Control in the United States of America (USA) [6]. It took 9 years for this funding to be reinstated. The rise in TB incidence in the USA during the mid-1980's was due to increases in homelessness and poverty and the emergence of the Human Immunodeficiency Virus (HIV)/Acquired Immune Deficiency Syndrome (AIDS) [8]. The USA and England were able to decrease TB morbidity and mortality only after a dramatic increase in funding which was largely allocated to the direct monitoring of antibiotic treatment [8].

The developing world continues to suffer from TB with TB incidence between 100-300 in 100 000 people in Asia and Western Russia (compared to 10 in 100 000 in North America). Even these numbers are not as high as those of Southern and Central Africa where TB incidence is over 500 per 100 000 people [8] (Figure 1.1).

1.3 Pathogen & Pathogenesis

M. tuberculosis is an aerobic pathogenic bacterium with a circular chromosome of 4 411 523 base pairs (bp) (65.6% G+C content) measuring 0.3-0.6 and 1-4 μm in width and length, respectively [9, 10]. This rod-shaped bacterium is weakly Gram-positive and classified as an acid-fast bacteria as it cannot be decolourized by acid alcohol [11]. *M. tuberculosis* has a complex cellular envelope composed of long-chain fatty acids, glycolipids and a number of other components which makes it exceedingly difficult to treat [10]. It is the slow growing nature of this bacterium that allows for chronic infection, requires long term treatment and complicates the diagnosis [12].

When *M. tuberculosis* is inhaled as an aerosol into the lungs, it is phagocytosed by alveolar macrophages which envelope it in a phagosome [10, 13]. Once in the phagosome, engulfed agents are typically processed for antigen presentation or elimination from the host. In the case of *M. tuberculosis*, elimination does not always occur. When infected macrophages are unable to remove the tuberculous material, a

granuloma is formed as infected macrophages are surrounded in concentric layers by a variety of immune cells (Figure 1.2) [14]. The composition of a granuloma depends on the infectious agent it is controlling. Tuberculous granulomas are composed of macrophages, T and B lymphocytes, dendritic cells, neutrophils, fibroblasts and epithelioid cells [10, 15].

Estimated prevalence of tuberculosis (per 100 000 population), 2007
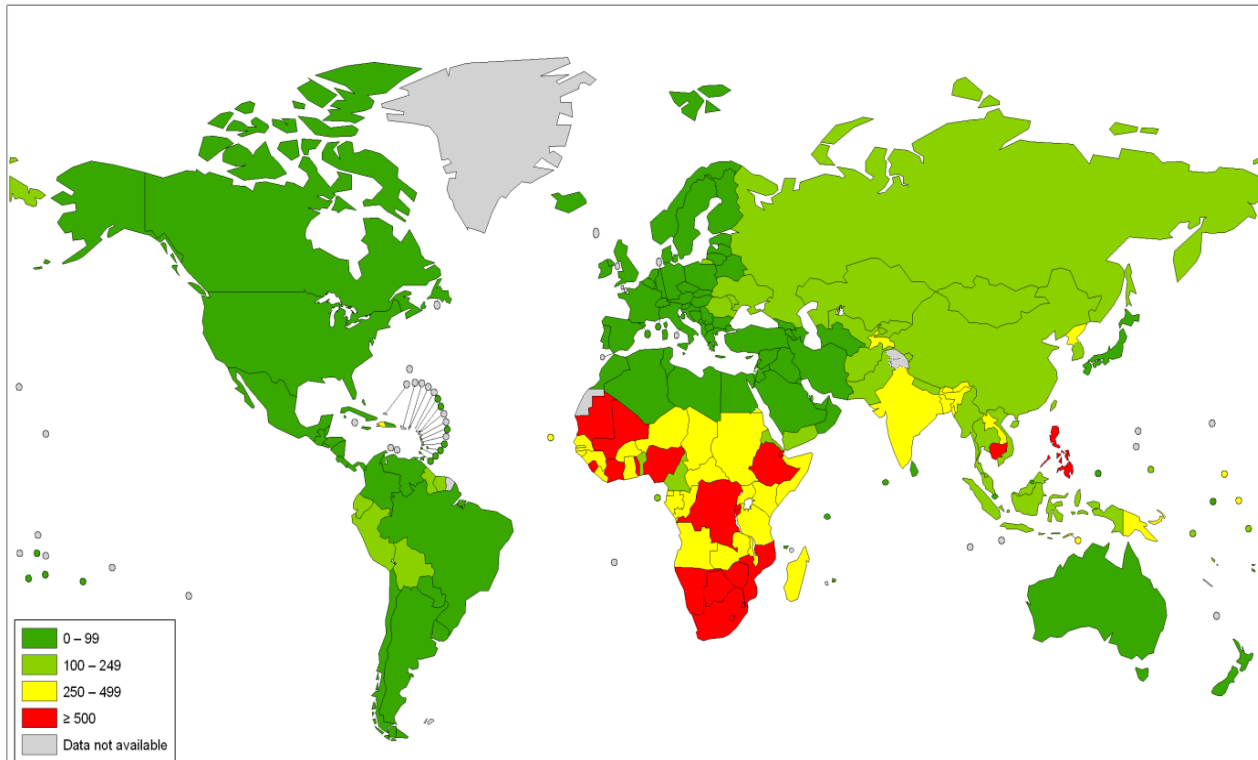
Figure 1.1 Prevalence of TB in 2007 as estimated by WHO: TB prevalence is highest in Central and sub-Saharan Africa with over 500 per 100 000 being infected with TB (Figure taken from the WHO website).

Through concurrent evolution with its host, *M. tuberculosis* has acquired the ability to both evade the host immune response and use certain components of the host immune system to its own advantage. From within the granuloma, the bacteria has access to host lipids which provide an abundant source of nutrients allowing it to persist for years [14]. Necrotic cavities can develop at the center of these granulomas which result in significant damage to the lung tissue upon granuloma rupture [4, 16]. The granuloma is beneficial to both the host and the bacilli as it protects the host from further spread of infection and shields the bacteria from elimination by the host [16]. Determining how *M. tuberculosis* is able to evade the host immune system and persist within these granulomas will play a major role in developing more effective treatments against TB infection. In times of stress (HIV/AIDS

3

infection, malnourishment, poverty, etc.) the strength of the host immune response can deteriorate. During this period of immunological stress, the bacteria can resume multiplication leading to the erosion of the granuloma wall and the release of *M. tuberculosis* into the airway [8, 16].
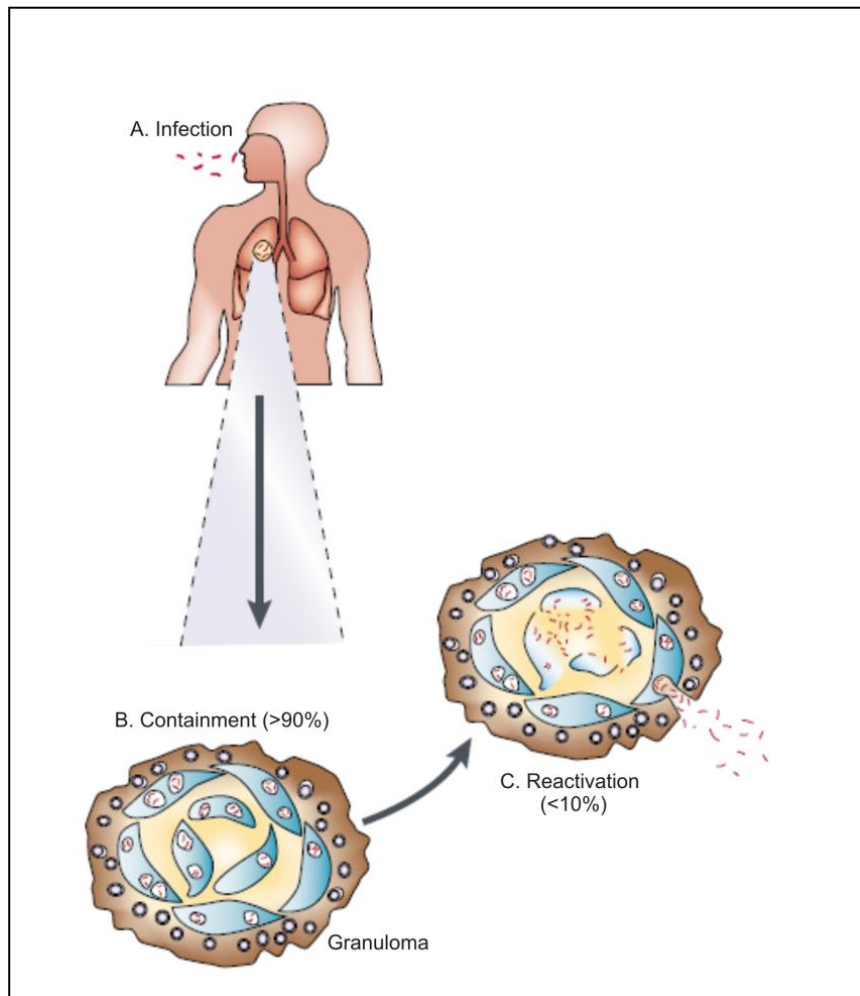
Figure 1.2 Depiction of TB Progression: A. Aerosols containing minute numbers of the TB bacilli are inhaled and become engulfed by macrophages in the lung. **B.** Ninety percent of individuals are able to effectively contain the bacteria through the formation of granulomas which can last for years, never allowing the bacteria to spread. **C.** In ten percent of individuals, the bacteria is able to persist in the granuloma and multiply until the point of granuloma rupture, releasing numerous bacteria into the surrounding lung space. Adapted from [10].

1.4 Host Metabolism

Prior to the discovery of effective antimycobacterial drugs, a diet rich in calories, protein, fats, minerals and vitamins was generally considered to be an important, if not essential factor, in treatment of TB [17]. There exists a multitude of evidence which supports an important role for nutrition in developing active TB and determining the severity of disease.

Most of the studies investigating the link between nutrition and TB are merely observational reports during times of war, famine or natural disaster [17]. The majority of these studies are flawed in a number of ways with respect to sampling methods, sample sizes, definition of malnutrition and assessment of infection [17]. Despite these complicating factors a number of studies exist which present convincing evidence for such an association. During the First World War, Denmark experienced a TB epidemic when locals suffered from a diet lacking in protein, vitamins and minerals as the country was exporting the majority of its meat, fish, poultry and dairy products [17]. It wasn't until the German blockade of Denmark, when the Danes experienced a surplus of these foods, that there was a dramatic decrease in TB rates whilst infection rates of TB in surrounding countries continued to climb [17].

The impact of proper nutrition was also seen in Norway. In the early 20th century, naval recruits in Norway experienced a higher than normal rate of TB infection believed to be caused by crowded and unhygienic living conditions [17]. Despite improvements in housing and hygiene, there was no major change in infection rates. Once recruits received a diet supplemented with margarine, cod liver oil, whole wheat bread, fresh fruit and milk, TB morbidity decreased significantly [17].

Perhaps one of the most revealing studies was that of Leyton who made observations in a prisoner of war (POW) camp in Germany in 1946 [18]. In this camp, British and Russian POWs were housed under similar living conditions and subjected to similar types of manual labour. The British POWs received supplements from the Red Cross in the form of meat, cheese, fish, milk and vitaminized margarine. This added up to approximately a thousand extra calories a day compared to their Russian counterparts [18]. Leyton investigated a number of diseases prevalent in the prison (both acute and chronic) including malaria, dysentery, nephritis and diabetes but the only disease which showed significant difference was the incidence of TB [18]. Infection rates of British POWs were found to be 1.2% while those of Russian prisoners were as high as 15% [18].

Numerous studies have shown that TB patients exhibit a reduced nutritional status when compared to healthy controls [19]. Measurements of body mass index (BMI), skin fold thickness and arm muscle circumference recorded in these studies provide convincing evidence which supports the occurrence of muscle wasting upon progression from TB infection to TB disease. One study performed in the USA reported that weight loss was a common symptom in 45% of TB patients with 26% showing signs of persistent anorexia [19]. TB-associated wasting is likely caused by a decrease in patient appetite resulting in a

decreased energy intake [20]. A study performed by Paton *et al.* [20] found that patients with TB-associated wasting weighed an average of 12 kg less than controls. TB susceptibility is also affected by an individual's BMI prior to development of active disease. A study performed in 2010 in HIV-infected individuals from South Africa found an increased BMI to reduce the risk of both morbidity and mortality of TB even after correcting for CD4 counts [21].

The cause of TB-associated wasting continues to evade researchers. One possible candidate is leptin, a key mediator of energy metabolism implicated in the development of anorexia in patients with chronic inflammatory states [22, 23]. This protein is the product of the *ob*-gene and the binding of leptin to its receptors located in the hypothalamus results in the suppression of appetite [24]. Leptin is produced by adipocytes and studies have shown that levels of leptin are decreased during times of starvation as fat stores become depleted [24-26]. In 2002, two studies published conflicting results on the association of leptin to TB-induced wasting. A study conducted in Indonesia which found that plasma leptin levels were significantly suppressed in TB patients was contradicted by a study performed in London which found no such association [23, 27]. This discrepancy may to some extent be explained by leptin being regulated by antagonistic mechanisms. Decreased plasma leptin levels have been previously associated with fasting but have also been found to increase in response to mediators of inflammation [22, 24-26]. Further studies investigating the role of leptin in both inflammation and wasting could provide valuable insights into TB-associated wasting.

B. Exacerbating the Epidemic

1.5 Bacterial Factors: Latency & Drug-Resistance

Latent TB infection (LTBI) is clinically defined as previous infection with *M. tuberculosis* with no clinical symptoms of TB disease [28]. This form of TB infection still exhibits an immune response to TB antigens as utilized by the tuberculin skin test and is the most common form of TB infection (currently infecting one third of the world's population) [28]. The ability of *M. tuberculosis* to persist within the host has proven to be one of its biggest strengths. Within the first weeks of infection, the bacilli multiplies rapidly but it slows its rate as increasing numbers of T-lymphocyte, monocytes and macrophages are recruited to the infection site [29]. As discussed in Section 1.3, during effective control of TB infection, a

granuloma is developed around infected macrophages which protect the host from the further spread of infection. These granulomas, however, also provide a safe haven for the bacteria, protecting them from complete elimination by the host. During times of immunological distress, the bacterium is able to resume multiplication and progress to active disease. How the bacterium is able to initially avoid elimination and persist within the host for years is still under investigation but the study of latent infection is complicated by the lack of an appropriate animal model. While a variety of animal models have been utilized for latency studies, none are an accurate representation of latent infection in humans [30]. LTBI is treated in some areas but the usefulness of LTBI treatment is controversial. A number of clinical trials have shown that LTBI treatment can reduce but never eliminate the risk of TB disease and poor adherence to the treatment (due to treatment length and side effects) is thought to lead to the emergence of drug-resistance [30]. In addition, current drugs used for LTBI treatment are the same as those used to treat active disease, including isoniazid and rifampin/pyrazinamide. Since most anti-TB drugs target cell wall components they are most effective on actively replicating bacteria [30]. It is imperative that drugs are developed which target the bacilli in the resting state in order to effectively eliminate the bacteria from the host.

Drug-resistance has existed almost since the first use of antibiotics. Dubos reported that *M. tuberculosis* bacilli became resistant to streptomycin a few weeks to a few months following treatment [7]. Incomplete treatments increase the likelihood of an individual developing a drug-resistant strain of TB. Drug-resistant strains fall into one of two categories: multidrug-resistant (MDR) and extensively drug-resistant (XDR). MDR strains of *M. tuberculosis* are, by definition, resistant to rifampicin and isoniazid and may or may not have other drug-resistance [31]. The treatment for this category of *M. tuberculosis* lasts even longer than regular treatment and is a combination of more costly and less effective antibiotics which are even more poorly tolerated by patients [31]. The highest incidence of MDR-TB is in Eastern Europe where 10% of new cases and 40% of previously treated cases are MDR. In the remainder of the world, it is estimated that 5% of TB patients have MDR strains of which 40% are patients who have been previously treated for TB [32]. WHO estimated that each year there are 490 000 new cases of MDR-TB resulting in over 130 000 deaths [31, 32]. XDR-TB is characterized by resistance to rifampicin and isoniazid as well as any quinolone and at least one injectable second-line agent (capreomycin, amikacin, kanamycin) [31, 32]. The present increasing prevalence of drug-resistant TB makes TB a concern once again on an international scale. This is illustrated by a study of

TB prevalence in Berlin, Germany [33]. Germany is currently classified as a low-incidence country with respect to TB prevalence with a TB incidence rate of 6.1 per 100 000 and a mortality rate of 0.2 per 100 000 in 2007 [33]. It was found that 43% of TB patients identified in Germany were born outside the country and 80% (53 of 66 patients) of those found to be infected with MDR-TB were also foreign born [33]. Another study which investigated XDR-TB in California, USA also found that 83% of XDR-TB patients were foreign born and half of those had immigrated to the USA less than 6 months prior to diagnosis [34]. Even in a developed country like the USA, patients diagnosed with XDR-TB who were HIV-negative had a mortality rate 41% (results based on known final outcomes of 12 of the 18 patients diagnosed with XDR-TB between 1993 and 2006). In developing countries concurrently battling HIV/AIDS, poverty and malnutrition, mortality rates are expected to be much higher. These studies clearly indicate that attention must be paid on an international scale to the prevalence of all forms of TB and an increase in global morbidity will exacerbate the spread of drug-resistance.

1.6 Host Environmental Factors

While it is clear that TB is most prevalent in areas suffering from poverty, the reasons for this might not be as apparent. A study conducted at The Bushullo Major Health Centre located in the Southern Nations, Nationalities and People's Region of Ethiopia in 2005 illustrates some of the issues with treating TB in such regions. Cambanis *et al.* [35] reported that based on interviews conducted with 243 TB-suspected patients, 51% had to borrow money and an additional 22% had to sell personal property (including cattle and grain) to afford the trip to the health center. This study also found that 85% of surveyed patients waited more than 10 days from the onset of symptoms before seeking medical attention and major reasons for this included cost of transport, prolonged transport times (including overnight trips) and the need to sell personal belongings [35]. A report published in 1991 also found that poverty affected treatment compliance [36]. Farmer *et al.* [36] reported that in extremely poor countries there is massive drop-out from treatment programs. However, even where TB programs are established, there exists a lack of follow-up and a need for accurate record keeping [36]. A more recent study investigating possible solutions to the ongoing battle of treatment compliance was that by Belo *et al.* [37] which reported that the poorest patients had longer transit times which resulted in longer wait times between symptom onset and visiting a health facility. A possible solution to this issue is the introduction of incentives to encourage patients to seek medical

attention sooner and complete the course of treatment. The authors found that the top five incentives as ranked in a survey conducted in the Rio de Janeiro state of Brazil were access to laboratory tests, home drug delivery, food allocation, ticket transportation and scheduled appointments [37]. With a curable disease like TB, one of the main obstacles to elimination of this disease is treatment compliance. The introduction of incentives may be a plausible method of increasing treatment compliance and subsequently decreasing TB incidence.

Based on the current global distribution of TB, it is apparent that those areas most afflicted remain those concurrently suffering from over-crowding, poverty and malnutrition. Not surprisingly, nearly half the population of sub-Saharan Africa is currently living below the international poverty line, and chronic hunger afflicts 1 in 5 people in the developing world [38-40]. Chronic hunger increases susceptibility to disease, decreases learning ability and reduces the individual's ability to work resulting in a further increase in poverty. Much as there appears to be a bi-directional effect between poverty and chronic hunger, there also appears to be a bi-directional interaction between TB and malnutrition (Figure 1.3). For this reason, studying the effects of nutritional status on the development of TB and the impact of active TB on nutritional status has proven to be a difficult task. This problem is aggravated by a number of other environmental factors including, but not limited to, poor public healthcare systems and unhygienic living conditions.

In areas plagued by poverty, malnutrition and TB, each entity seems to intensify the prevalence of the other two. It would seem that addressing any one of these issues would have a positive impact on all three. Further examples of the impact of nutritional status on TB are discussed in Section 1.4 which deals with host metabolism.

Unlike TB, a disease which has plagued society for centuries, HIV/AIDS was only recognized in 1981 and the collision of these two epidemics has had disastrous consequences. HIV/AIDS is an infectious disease transferred by the exchange of bodily fluids and by exposure to contaminated blood or blood products [41]. Infection with HIV results in a dramatic decrease in circulating levels of CD4 T-lymphocytes, effectively weakening the host immune system. Once CD4 counts drop below 200 cells/mm^3 (from 1500 cells/mm^3 in healthy individuals), patients exhibit increased susceptibility to a variety of opportunistic infections [42]. Similar to TB, the burden of HIV/AIDS is felt most in developing regions of the world including South and South-East Asia and sub-Saharan Africa (Figure 1.4). A report by UNAIDS (a joint United Nations programme on HIV/AIDS) in 2009 estimates that as of December 2008 there were 33.4 million people currently living

with HIV/AIDS, 22.4 million of which were residing in sub-Saharan Africa [41]. In various areas of the world, many people, largely women, avoid HIV/AIDS testing and subsequently receiving treatment, for fear of shaming or being disowned by their families [43].

The stigma associated with HIV/AIDS has resulted in many individuals never receiving treatment which has allowed certain opportunistic infections to flourish. TB is currently the leading cause of death in individuals infected with HIV/AIDS [44]. Not surprisingly, sub-Saharan Africa currently has the highest TB-HIV/AIDS co-infection rates worldwide with South Africa ranking number one (Figure 1.5). South Africa's history has had a major influence on its current TB-HIV/AIDS co-infection status. Two major factors contributing to this co-infection status are apartheid, which segregated different populations, and the development of the mining industry with each creating environments ideal for the spread of infectious diseases. Poor sanitation, malnutrition and crowded living conditions allowed TB to thrive. Overcrowding and poorly-ventilated hostels where mine-workers resided which were regularly serviced by commercial sex-workers created the ideal situation for the spread of both HIV/AIDS and TB [44].
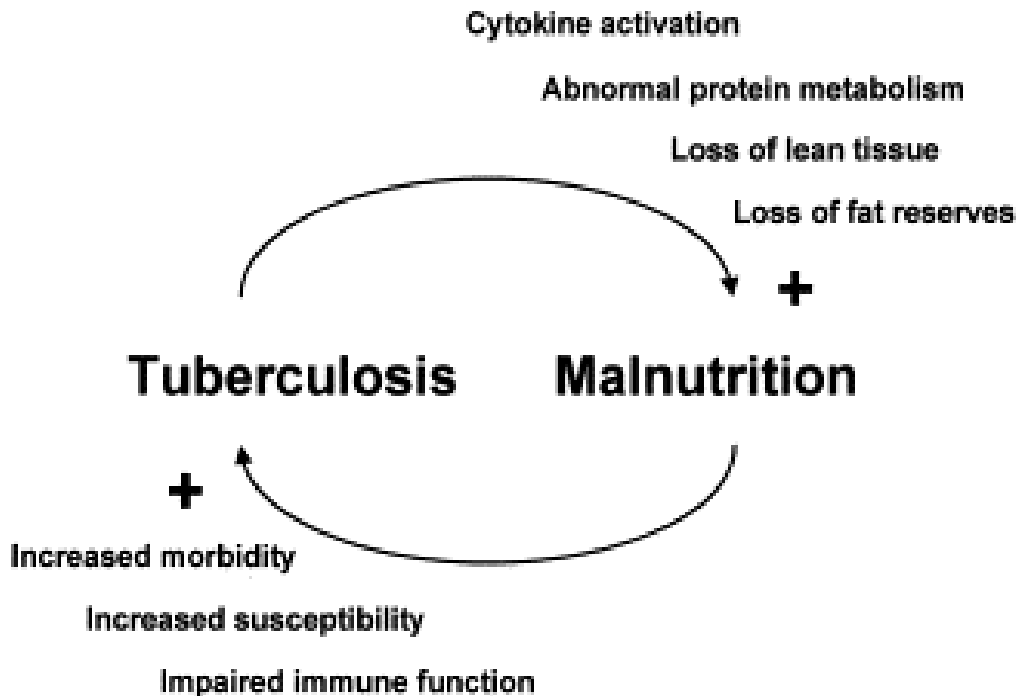
Figure 1.3 Bi-directional interaction of malnutrition and TB: Malnutrition leads to increased susceptibility to infectious diseases as a result of a compromised immune system. Increases in susceptibility increase disease prevalence. In the case of TB, infection leads to wasting in the form of lean tissue and adipocyte loss resulting in increased levels of malnutrition. Figure from Malnutrition and Tuberculosis by Macallan [40].
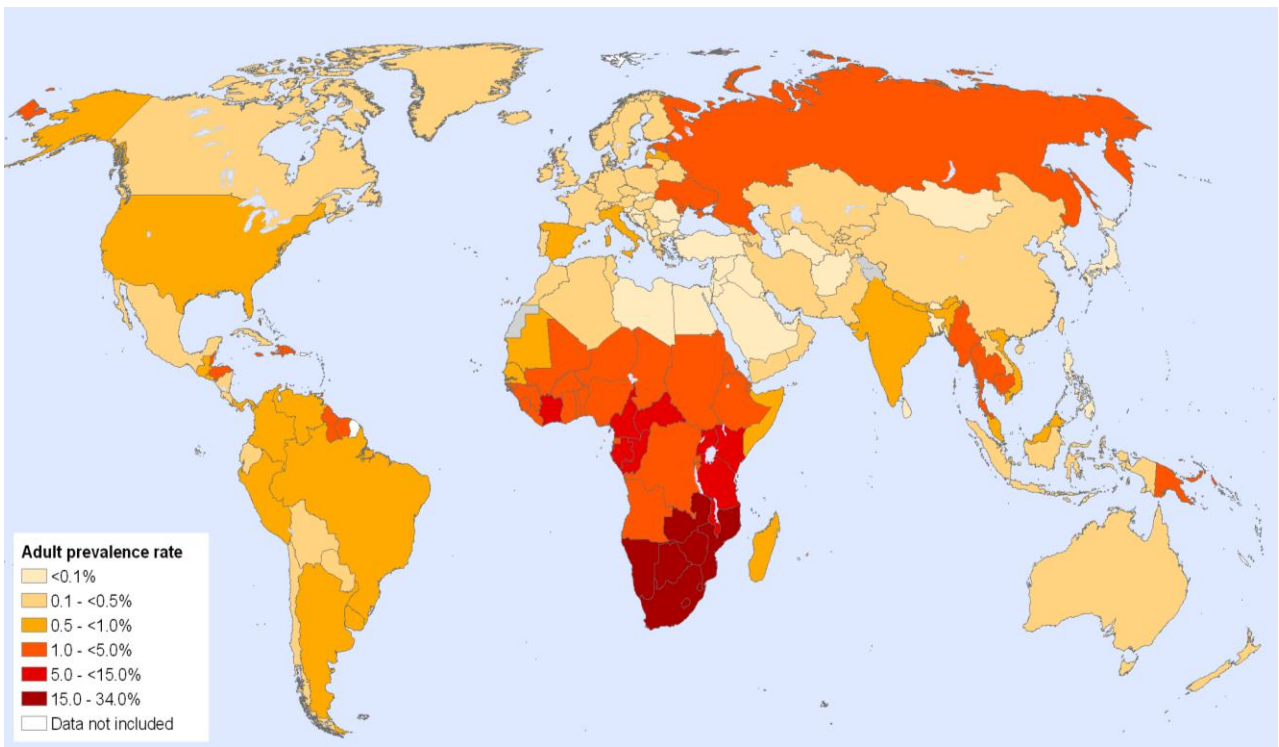
Figure 1.4 Global HIV/AIDS revalence: A WHO report published in 2007 estimated that 39.5 million people were infected with HIV/AIDS worldwide in 2006 (Figure taken from www.unaids.org)
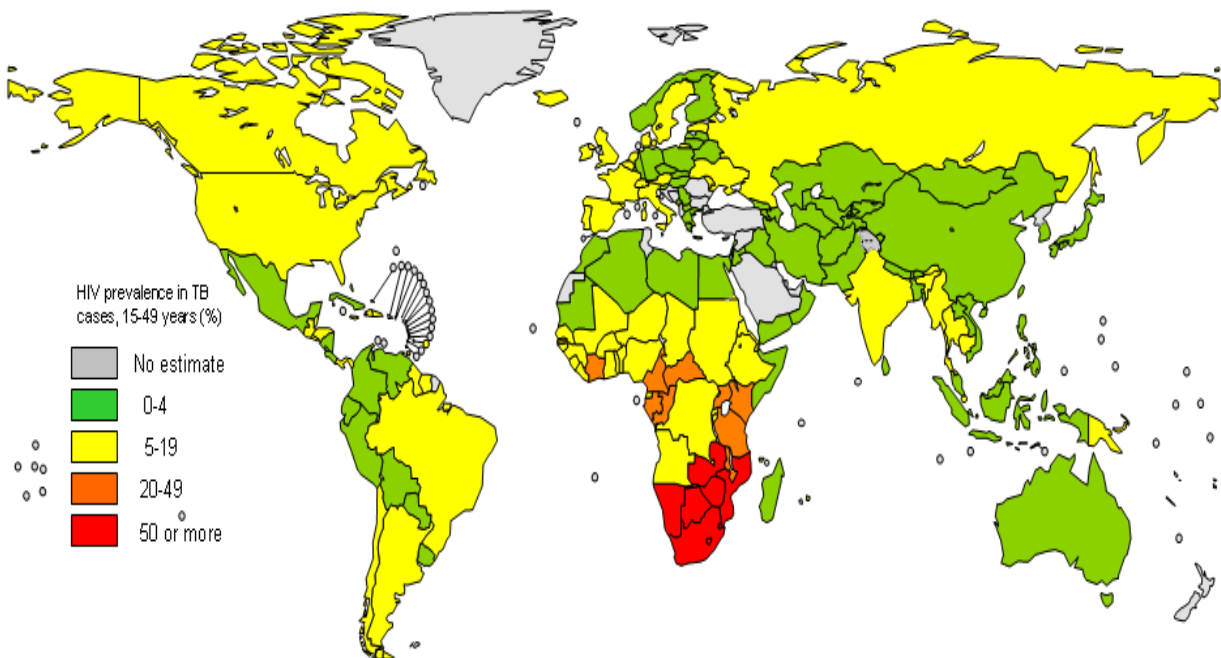
Figure 1.5 Estimated HIV/AIDS prevalence in new TB cases: In a report published in 2006, WHO estimated that HIV prevalence in TB incidence cases was highest in sub-Saharan Africa in 2005 (Figure taken from the WHO website).

1.7 The Synergistic Effect

The result of having all the issues discussed above being co-localized has been disastrous, the results of which may soon be felt on a global scale if not dealt with promptly and appropriately. Each issue appears to fuel the others resulting in some of the world's most resource-rich countries being incapacitated by rampant disease, poverty and malnutrition. Effects of these issues are also felt in the industrialized world. Studies found that in the late 1980s and the early 1990s, HIV infection was associated with outbreaks of MDR-TB, frequently with patients succumbing to disease even before MDR-TB could be bacteriologically confirmed [45]. During an XDR-TB outbreak in KwaZulu Natal, South Africa 52 of 53 patients died an average of 16 days after diagnosis [46]. These exceptionally high mortality rates were a result of co-infection with HIV/AIDS; (of the 44 XDR-TB patients tested, all tested positive for the virus) [46]. Genotyping of the TB strains showed that all XDR-TB patients were infected by genetically similar strains [46]. Having so many malnourished individuals living with HIV/AIDS and residing in poor, over-crowded living conditions dramatically increases the risk of exposure/re-exposure to TB and spread of increasingly virulent and untreatable strains of TB.

C. Approaching the Solution

1.8 Diagnostics

Currently, only one standard method exists to detect latent TB infection. The tuberculin skin test (TST) which is also referred to as the Mantoux or purified protein derivative (PPD) test [47] is widely used but it is limited in a number of ways. The PPD used in this test is composed of antigens found in *M. tuberculosis* but also *Mycobacterium bovis* and other mycobacteria which can result in false positives in areas where the BCG vaccination is used [48]. The reverse problem, false negatives, is also an issue in areas plagued with HIV/AIDS as a compromised immune system can lead to anergy [49]. Aside from erroneous results, this method also results in poor follow up due the requirement of a return visit 72 hours after administration of the test as well as variability in dose of PPD, method of application and criteria of interpretation [49].

At present there exists a variety of methods which are used to detect active TB disease in patients but a definitive diagnosis requires culturing of *M. tuberculosis* bacteria from a patient specimen [50]. Due to the slow growing nature of *M. tuberculosis*, limitations of the

current detection methods and lack of proper disease seeking programs, it is estimated that each year more than 30% of new TB cases go undetected [1]. The most commonly used method of TB diagnosis in low-income countries is the sputum smear which was introduced over 125 years ago by Robert Koch [12]. This method detects the presence of acid fast bacilli but has a limited use as it cannot distinguish species or provide any information on drug-resistance. Chest radiography is also used to detect active TB but normal results or mild radiographic evidence each play a role in missed diagnoses [50]. Another major diagnostics tool is solid culture, but this method requires approximately 2 weeks for a diagnosis [51]. Recent advances in culture techniques include the development of liquid culture diagnostics which have only a 2-day turn around period, but while they have a high specificity (83 - 100%) they lack sensitivity (21 - 83%) [52]. A variety of other techniques are currently in development but the practicality of their application may be limited in resource-poor settings. Nucleic acid amplification tests (NAAT) can provide results within one day. A number of reviews on this method suggest that it has a high specificity but exhibits poor and inconsistent results with respect to sensitivity and this method of TB diagnosis is even less sensitive when patients are TST negative (ie: HIV/AIDS endemic areas) [53]. The development of an immune-based diagnostic test have been attempted for decades but due to a wide ranging variability in the immune response, none have been able to completely replace microscopy and culture [53]. Difficulty in developing a serologically-based diagnostic test arises due to the different states in which *M. tuberculosis* may be found. It is thought that *M. tuberculosis* expresses different genes and proteins during active and latent infection [54].

Due to the currently limited ability to rapidly and conclusively diagnose active TB disease, missed diagnoses result in the perpetuation of TB transmission and increased morbidity. Even if a proper diagnostic test were to be developed, its influence on the morbidity and mortality rates of TB are likely to be minimal until proper disease seeking and drug adherence programs are in place.

1.9 Drugs, Vaccines & Public Health Programs

Where treatment is unavailable, TB mortality can be up to 60% and most of these cases occur in developing countries [9]. During the centuries in which TB has plagued society, physicians have struggled to find ways to treat diseased individuals. The more ancient procedures did nothing to cure the disease but merely alleviated the symptoms, making

patients feel as though they were getting better [7]. Today, there exist multiple antibiotics with which to treat TB which include streptomycin, isoniazid, rifampicin and pyrazinamide [8]. Despite all these treatments, TB has yet to be eliminated. Currently, the most effective treatment is the combination of 3 antituberculosis drugs administered over a 6 month period [55]. With a treatment period of this length, it is difficult to keep patients motivated to complete the course of antibiotics, especially when patient health improves in just 1-2 weeks [55]. In addition to treatment length, many of the antibiotics currently used to treat TB are poorly tolerated by patients. Isoniazid can result in adverse neurological and hepatic reactions while rifampicin can result in allergic reactions including fever, rash and flu-like symptoms [56]. Adverse reactions to other first-line TB drugs include loss of visual acuity and colour vision (ethambutol), permanent deafness and nephrotoxicity (streptomycin) and hepatotoxicity and gastrointestinal intolerance (pyrazinamide) [56].

The sole existing vaccine, the BCG vaccine, currently shows inconsistent efficiency (0-80% in randomized control trials) [57]. It is thought that the original strain generated by Calmette-Guérin, shown to be an efficient vaccine for TB, has become increasingly attenuated over the years as subsequent passages of the vaccine have been produced [57]. While attempts to develop a new vaccine have been in progress over the past several years, none have been able to replace BCG. The lack of an appropriate animal model complicates the search for an alternative vaccine. Researchers are limited to using mouse and guinea pig models to test the activity of potential vaccines despite the mouse's inability to replicate lung pathology and the guinea pig's high susceptibility to infection [58].

In an increased attempt to eradicate TB, WHO implemented a new program entitled Directly Observed Treatment and Short-course drug therapy (DOTS) [32]. Under DOTS, antibiotic treatment must be observed by health professionals to ensure that all antibiotics are taken and the entire treatment course is completed [55]. The use of patient registries which follow a patient from diagnosis to cure have a well-established precedent to improve public health [59]. With a patient registry, a patient's progress can be tracked from diagnosis to cure but in developing countries, where governments and public healthcare systems are frequently unstable, this direct patient observation is extremely difficult [8]. In addition, the concept of treatment observation is insufficient since most TB patients only arrive at the clinic once extremely ill, likely exposing family and community members prior to seeking treatment. Surveillance programs which actively seek out individuals with active disease must be implemented if control over TB is ever likely to be accomplished. Ineffective public health measures, low interest in co-operation by patients, development of

drug-resistant strains and the emergence of HIV/AIDS has led to higher levels of TB infection than ever before [60].

1.10 Basic Sciences: Mycobacteriology & Immunology

Through mycobacterial research, much information on the life cycle and metabolism of *M. tuberculosis* has already been gathered. Due to the current need for original preventative, diagnostic and treatment strategies, further research into the inner workings of *M. tuberculosis* is absolutely necessary. The release of the complete genomic sequence of *M. tuberculosis* has allowed researchers to identify new genes which may provide novel targets for drug therapy and insight into the complex interaction between host and bacteria [61]. Using the genomic sequence, researchers are able to utilize microarrays to measure gene expression which can aid in identifying critical genes and investigate the function of unknown genes [62]. Mycobacterial proteomics goes beyond the static mycobacterial genomic sequence and focuses on the actively expressed portions of the genome which may aid in identifying new areas for vaccine development [63]. Identifying differences in the protein profiles between active and latent disease may allow researchers to develop methods to diagnose the two separate disease states [62]. Determining translational control and post-translational modifications (PTMs) present in *M. tuberculosis* may provide researchers with a level at which to manage the progression to active disease and may provide novel targets for drug development. Mycobacteriological research alone is insufficient as the numerous strains have been shown to vary in fitness, virulence and persistence. Studies in conjunction with the host immune response to the various strains may provide valuable insight into the complex interaction between host and bacterium.

An increased progression to active disease in immunocompromised individuals emphasizes the importance of the host immune response in effective control of the *M. tuberculosis* pathogen. While much research into the role of immunology has been performed, current knowledge of the mechanisms which play a role in protection against TB remains limited [64]. However, immunological research may play an important role in a number of areas crucial to effective TB control. As discussed in Section 1.8, much research has gone into the development of serological diagnostic tests for TB. While none as of yet have been able to replace microscopy or solid culture, further investigations may provide useful information in distinguishing disease state based on the detection of specific biological markers. By combining mycobacteriology and immunology, proteins which elicit

an appropriate immune response may be used to identify novel targets against which a vaccine can be developed. Immunological research is limited as it allows only for investigations to be performed on genes implicated in the host immune response. Because no single gene has been shown to be responsible for TB susceptibility it is important to identify new genes which currently have an unknown role in the immune response to TB infection.

D. Genetics

1.11 Evidence for Genetic Component to TB Susceptibility

A study conducted in 1990 of nursing home residents of Arkansas, USA showed that African-Americans were twice as likely to develop TB than Caucasians [65]. This suggests that some populations are more susceptible to TB than others. A possible explanation for the non-uniform distribution of TB comes from Haldane who suggests that over the past 5000 years natural selection in humans has been influenced by microorganisms [66]. In regions plagued with malaria we can see selection in certain individuals for resistance to this protist by means of specific genetic mutations [66]. Similar trends can be seen with TB disease in the Qu'Appelle Indians who were first exposed to the *M. tuberculosis* in 1890 [60]. Following the introduction of the bacteria, half the families died from TB infection. After 40 years the annual TB mortality rate dropped from initial levels of 10% down to only 0.2% [67]. In the 18th and 19th century, TB killed one quarter of the European population, possibly selecting for a population with a higher resistance to TB [67]. Because TB has been prevalent for less than a hundred years in Africa, it is likely that African populations exhibit higher susceptibility to infection [68].

However, within a population or within a family, there are still individuals who appear to be more susceptible to TB than others. Studies have found monozygotic twins have a higher concordance for TB infection than dizygotic twins [68]. It appears that some individuals are better able to control TB infection than others. Support for this theory comes from an event which occurred in Lubeck, Germany in 1923. During what was supposed to be the routine vaccination with BCG, 249 babies accidently received virulent *M. tuberculosis*. Despite all infants receiving the same strain and same dose, only 173 survived [69]. It is therefore probable that the genetic make-up of each individual plays a role in how *M. tuberculosis* is handled by the host.

To better understand the mechanisms utilized by *M. tuberculosis*, it is necessary to further investigate the host immune response upon TB infection and identify new genes involved in TB pathogenesis and susceptibility

1.12 Finding the Genes

For many infectious diseases, there can be a complex interaction between microorganism, host and environment which influences the severity and progression of disease [70]. Genetics are known to play a significant role in the development of infectious diseases with some genetic variations being more influential than others [70]. TB is a complex disease with multiple genetic loci and multiple environmental factors contributing to its development [70]. Finding the genes involved in TB infection has proven to be a challenge [71].

Certain factors can confound the search for disease-causing genes. One such factor is population stratification. Population stratification refers to the existence of a subpopulation whose genetic makeup differs significantly from the larger population due to differences in ancestry [71]. The result of this is the detection of spurious associations when performing case-control studies [72]. Linkage disequilibrium (LD) can also lead to erroneous results. LD is a measure of the proximity of two alleles at different loci on the same chromosome. The closer two alleles are, the less chance recombination will occur between them, giving them a higher LD value [72]. As a result of LD, the allele at one locus can frequently predict the allele at a second locus [73]. This can cause false positive results when an allele in LD with the disease-causing allele is investigated and an association is found. LD can be also useful as it can lead to the identification of the true disease-causing variant.

A variety of methods are currently utilized by researchers to seek out loci, and subsequently genes, which show an association to disease.

Linkage studies locate genes linked to particular phenotypes without any prior knowledge of biological function and are most useful at identifying rare variants [74]. Linkage analysis involves using naturally occurring polymorphisms as generic markers to create a genetic map from which the transmission of chromosomal regions in families can be traced [74]. No assumptions of inheritance mode are made allowing for an unbiased identification of regions of the genome that may contain potential susceptibility genes. This technique has increased the number of disorders linked to specific genes from roughly 100 in the late 1980s to over 2200 today [75]. The affected sibling-pair method is a type of linkage study

which is frequently used to identify loci which show an association to a complex genetic trait [76]. This method attempts to identify regions of the chromosome which are inherited more frequently by affected siblings than would be expected to occur through chance alone [77, 78]. This method has been successful in diseases such as type 1 diabetes, asthma and inflammatory bowel disease [70]. Linkage analysis has proven useful as it has replicated previous associations and identified novel genetic variants which may have been missed by the hypothesis-driven approach [69]. By identifying polymorphisms located in non-genomic regions, linkage studies can lead to new hypotheses of the function of these regions and their involvement in disease [79]. Following the detection of a region showing evidence of linkage, further fine mapping of that region can allow for the identification of candidate genes with little or no previous association with that disease [70].

Hypothesis-driven studies identify candidate genes based on results from animal models or mechanisms previously elucidated in humans and known to be involved in the development of disease [80]. This method allows for targeted investigations to be performed on SNPs suspected of being associated to a disease based on priori hypotheses [80]. Hypothesis-driven candidate gene association studies are cost effective as they focus research on genomic regions showing likely association to a specific disease as oppose to scanning the entire genome [70]. This approach can identify candidate regions through a number of ways. Animal models have been used to determine biochemical pathways and identify genes crucial to the pathogenesis of a particular disease [80]. Identifying these critical pathways in animal models allows for a targeted search of similar pathways in humans. Once identified, researchers can investigate all or specific genes involved in that pathway. Computer-based algorithms have also been utilized to identify candidate genes based on methods such as sequence comparisons against genes previously associated to disease development [79].

Population-based case-control studies compare the frequency of selected genetic variants between affected and unaffected individuals [73]. This method is most frequently used to identify common polymorphisms underlying complex genetic traits [73]. Population-based case-control studies are most useful when thousands of affected samples can be collected however complications can arise when investigating a disease where the affection status is subjective [74]. One important aspect of this type of method is the selection of appropriate control samples. Controls must be selected from the same population group so that both case and control subjects have similar exposure rates and are influenced by

similar environmental and social factors. Despite careful selection of a control group and a clear definition of affection status, population-based case-control studies are prone to false positives caused by population stratification or LD since this type of study does not include population analysis [79].

1.13 Genes Currently Implicated in TB Susceptibility

A number of genes have already been implicated in TB susceptibility and have been studied to varying degrees in a number of populations. An in depth summary of these studies can be found in Table 1.1 (from a recent review by Möller *et al.* [81] which has been updated to include the most recent studies on genes associated to TB susceptibility).

The genes listed in Table 1.1 have been identified through the use of the previously mentioned methods including both hypothesis-driven and linkage studies. Due to the number of factors which can lead to spurious results, attempts to find the true disease-causing variants are frequently accomplished through the use of multiple methods, each with unique advantages and disadvantages. Using a multistage strategy, Cooke *et al.* [82] mapped a novel locus which showed an association to TB susceptibility in African populations. Beginning with an affected sibling pair linkage analysis, one genomic region, chromosome 20q13.31-33, showed evidence of association to TB susceptibility in a Coloured population from South Africa and Malawians from the Karonga district [82]. Using a population from Gambia, detailed association mapping within this region identified two genes showing evidence of disease association: cathepsin Z (*CTSZ*) and melanocortin 3 receptor (*MC3R*) [82]. Polymorphisms in either gene which showed disease association were genotyped in two more West African populations, Guinea-Bissau and the Republic of Conakry [82]. A polymorphism in the 3'-untranslated region (UTR) in *CTSZ* rs34069356 showed strong disease association with genotype distributions being similar across all three populations (p=0.005) [82]. Following regression analysis, the association seen in the *MC3R* polymorphism rs3827103 did not remain significant though a trend for a protective effect of the *MC3R* AA genotype did remain between cases and controls [82].

Confirmation of these findings or further investigations as to the role of either of these genes in TB pathogenesis has yet to be performed.

1.14 Summary

Most of the developing world already suffers from the effects of the most recent TB epidemic and as drug-resistance and HIV/AIDS become more prevalent it is likely that TB incidence will increase in all areas of the world. Further research is required to identify genes responsible for TB susceptibility so that mechanisms underlying TB pathogenesis can be elucidated. Only once the pathogenesis of TB is better understood can we hope to develop more effective prevention and treatment strategies and gain control of this disease. Evidence for a genetic component to TB susceptibility has been supported by numerous studies, however, no individual gene has been found to be associated to TB susceptibility in all populations. Inconsistent results may be explained by unequal susceptibility shown across different populations. Using a multistage strategy, a recent study identified a locus on chromosome 20 which showed association to TB susceptibility across a number of populations [82]. Further investigations will allow us to gain more insight into the role host genetics plays in TB susceptibility in order to identify new genes which impact events such as effective TB control, progression from latent to active disease and host resistance.

Table 1.1 Candidate genes previously associated to TB

Gene	Associated Population	Ref.	Not Associated Population	Ref.
HLA	India	[83, 84]	India (South)	[85]
	Mexico	[86]	China (Hong Kong)	[87]
	Iran	[88]		
	Thailand	[89, 90]		
	Poland	[91]		
	South Africa (Venda)	[92]		
	Cambodia	[93, 94]		
	Indonesia	[95]		
	India	[84]		
	Russia	[96]		
	India (South)	[97, 98]		
	India (North)	[99]		
	Russia (Tuvinian)	[100]		
	Italy	[101]		
	Korea	[102]		
	China	[103]		
	Brazil	[104]		
CCL2	Ghana	[105]	China (Hong Kong)	[87]
	Mexico	[106]	South African Coloured (SAC)	[107]
	Korea	[106]	Russia	[105]
CD209	SAC	[108]	Colombia	[109]
	Gambia	[110]	Tunisia	[111]
	Sub-Saharan Africa	[110]	Guinea-Bissau	[110]
			India (South)	[112]
			Republic of Guinea	[110]
			Guinea-Bissau	[110]
			Malawi	[110]
IFNG	Sicily	[113]	West Africa	[114]
	Croatia	[115]	India (South)	[116]
	Caucasian	[117]	African American	[118]
	China	[119]	Caucasian	[118]
	SAC	[120]	USA (Houston)	[118]
	Turkey	[121]	Japan	[122]
	China (Hong Kong)	[123]	Indonesia	[124]
	Colombia	[125]		
MBL	Spain	[126]	Tanzania	[127]
	SAC	[128]	Poland	[129]
	African America	[130]	USA (Houston)	[130]
	Turkey	[131]	Caucasian	[130]
	Italy	[132]	China	[133]
	India	[134]	Gambia	[135]
	Denmark	[136]		
NOS2A	Colombia (Paisa)	[137]	Mexico	[106]
	African American	[138]		

	Brazil	[139]			
	SAC	[107]			
SLC11A1	USA (Houston)	[140]	Japan	[122]
	China	[141-143]	Thailand	[90]
	Korea	[144]	Morocco	[145]
	Tanzania	[146]	Gambia	[147]
	Cambodia	[93]	Russia (Slavic)	[148]
	Gambia	[147]	Taiwan	[149]
	China (Han Chinese)	[150, 151]	Poland	[129]
	Japan	[152]		
	Peru	[153]		
	SAC	[154]		
	Poland	[155]		
	Gambia	[147, 156]		
	USA (Caucasian)	[157]		
	Japan	[158]		
	Guinea-Conakry	[159]		
	Denmark	[146]		
SP110	Republic of Guinea	[160]	Guinea-Bissau	[160]
	Gambia	[160]	SAC	[161]
			Russia	[162]
			Ghana	[163]
TLR1	USA (African American)	[164]		
	USA	[164]		
TLR2	Korea	[165]	Guinea-Bissau	[166]
	Turkey	[167]		
TLR4			Gambia	[168]
			Guinea-Bissau	[166]
TLR8	Indonesia	[169]		
	Russia	[169]		
VDR	Guinea-Bissau	[166]	Tanzania	[127]
	India	[170]	South Africa (Venda)	[92]
	Asia (Gujarati)	[171]	Cambodia	[93]
	China (Han Chinese)	[172]	Peru	[173]
	West Africa	[174]	Japan	[122]
	Gambia	[175]		

CHAPTER 2: Study Overview

2.1 *CTSZ*

CTSZ, also known as cathepsin P (*CTSP*) or cathepsin X (*CTSX*), encodes the 303 amino acid protein cathepsin Z (CTSZ) [176, 177]. It is one of 11 lysosomal cysteine proteases (called cathepsins) that belong to the papain family and *CTSZ* is located on chromosome 20 at position q13 [177]. Like all cathepsins, CTSZ is translated as an inactive peptide, requiring the cleavage of the proregion for activation [177, 178]. While the cathepsins vary in their physiological function and expression, *CTSZ* is expressed only in cells of the immune system [177, 179, 180].

Recent studies have found that active CTSZ co-localized with β_2 integrins in monocytes, dendritic cells and lymphocytes and was found to regulate the activity of β_2 integrins lymphocyte function-associated antigen-1 (LFA-1; CD11a/CD18) and macrophage antigen-1 (MAC-1; CD11b/CD18) [177]. Interactions between CTSZ and the β_2 integrins LFA-1 and MAC-1 regulate leukocytes proliferation, cellular adhesion and migration, phagocytosis and activation of T- lymphocyte [181]. Integrins are one of six families of cell adhesion molecules and all integrins are heterodimers composed of one α and one β subunit [181]. The integrin family is subdivided into three main classes: β_1, β_2 and β_3 and all function in cell-cell and cell-matrix adhesion [181]. While β_1 and β_3 integrins are co-expressed on most cell types, β_2 integrin expression is limited to leukocytes [182]. Integrins play an important role in the extravasation of leukocytes across the epithelium into tissue (i.e. an infected lung) [183]. LFA-1 and MAC-1 are known to have a central role in leukocyte migration into sites of infections and inflammation as well as into lymphoid organs (for immunosurveillance) [184, 185]. Using vesicular staining, Obermajer *et al.* [179] found that in pro-monocytic U-937 cells, CTSZ localized in the perinuclear region but once differentiated, protein localization was most pronounced at the perimembrane region. Within a number of cell types, CTSZ was found to co-localize with β_2 integrins near the membrane and in differentiated macrophages and dendritic cells, interaction between CTSZ and β_2 integrin MAC-1 resulted in increased cellular adhesion [179, 180]. Within T-lymphocytes it has been shown that CTSZ co-localizes with LFA-1 at cell-cell junctions and is responsible for the activation of this β_2 integrin [180]. Based on in vitro experiments, it has been postulated that CTSZ activates LFA-1 through the sequential cleavage of C-terminal amino acids F^{766}, A^{767}, E^{768} and S^{769} until it reaches P^{765} which is resistant to CTSZ proteolytic cleavage (Figure 2.1) [177]. LFA-1 activation within T-lymphocytes leads to increased lymphocyte proliferation while MAC-1 activation from within macrophages and

dendritic cells results in the suppression of lymphocyte proliferation [180]. Both lymphocyte proliferation and cellular adhesion are important in effective control of microbial infection and the formation of granulomas. Yassin *et al.* [186] found that TB patients showed increased expression of CD11/CD18 on monocytes. It is plausible that interference with leukocyte migration and proliferation or in cellular adhesion may lessen the host's ability to effectively control mycobacterial infection.

2.2 *MC3R*

MC3R is one of five melanocortin receptors (MCRs) which belong to the family of G protein coupled receptors (GPCRs), each with seven membrane spanning regions and all transmit their signals through the activation of adenylate cyclase [187]. The main agonist for each MCR is α-melanocyte stimulating hormone (α-MSH). As previously mentioned, *MC3R* is located on chromosome 20q13, as is CTSZ and encodes a 361 amino acid protein [188]. The melanocortin 3 receptor protein (MC3R) is abundantly expressed in several brain regions and in a variety of peripheral tissues including the placenta, heart, gut and immune cells [184, 189, 190]. Unlike the other 4 MCRs, MC3R is the only MCR which is activated by γ- MSH to the same extent as α- and β-MSH [188]. *MC3R* has been shown to play a role in a variety of biological systems including the regulation of energy homeostasis and fat metabolism, regulation of the cardiovascular system and inflammation [191-197]. Defects in *MC3R* can result in decreased total expression, intracellular retention and defective receptor activation [192, 198-204]. Several studies have been performed investigating the association of obesity to various polymorphisms in *MC3R* [191, 192, 194, 197, 205]. Using a mouse model, Chen *et al.* [199] showed that inactivating mutations of *MC3R* led to an increase in fat mass with a corresponding decrease in body mass without any change in eating behavior or metabolic rate. One of the key regulators of MSH levels is leptin. Leptin is released from adipocytes and acts to modulate appetite and energy expenditure. Leptin levels are proportional to the number of adipocytes and remain relatively constant despite daily variations in energy consumption and expenditure [206]. Leptin increases MSH levels by increasing expression on pro-opiomelanocortin (POMC), the precursor to MSH. When leptin crosses the blood/brain barrier, it binds to the leptin receptor expressed on melanocortin neurons and increases the expression of MSH. Binding of MSH to MC3R results in a decrease in feed efficiency (Figure 2.2).
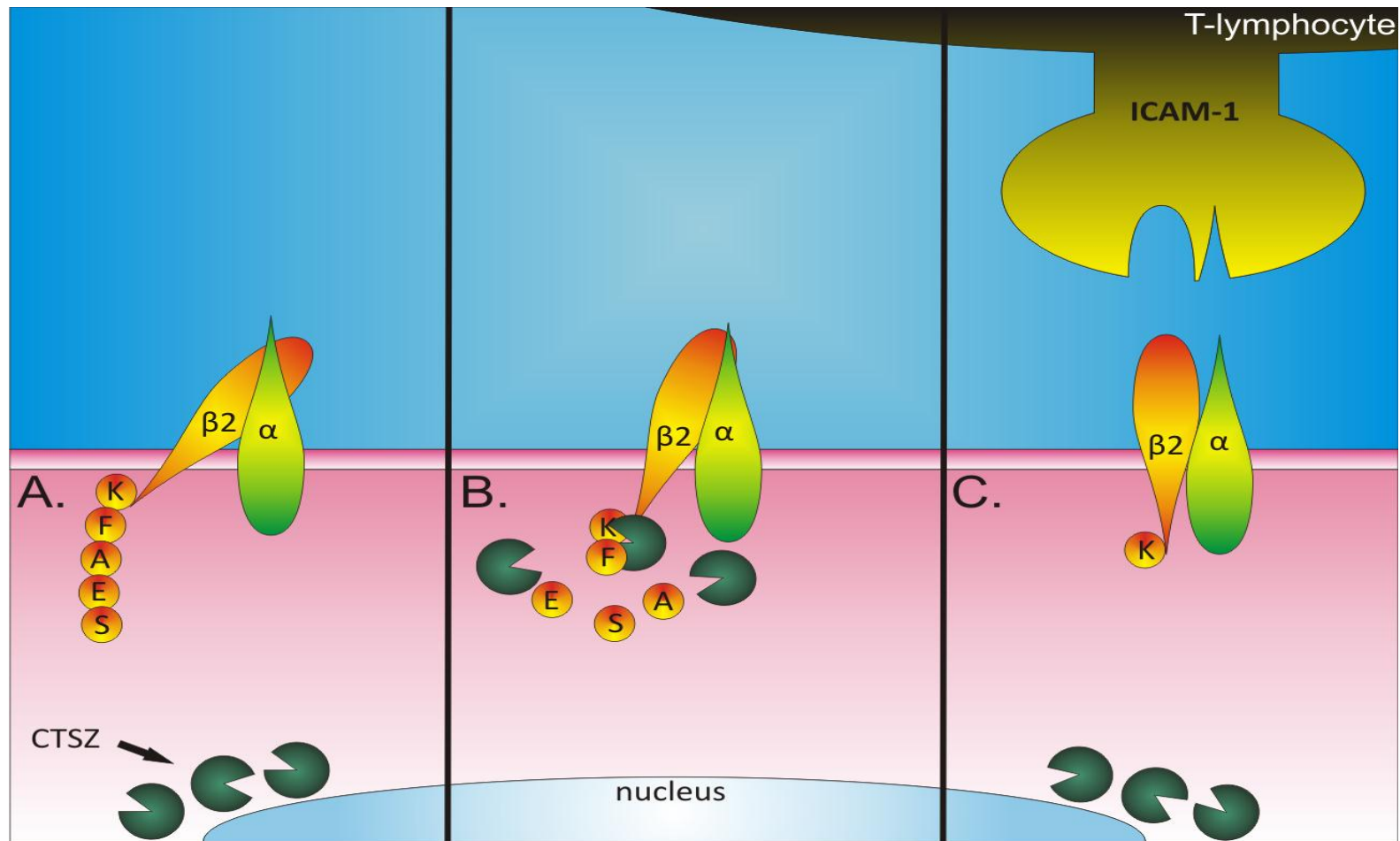
Figure 2.1 Interaction between CTSZ and β_2 integrin in a Dendritic Cell: A. In an immature, unstimulated dendritic cell, CTSZ is located in the perinuclear region and the β_2 integrin exists in its low-affinity form. **B.** Upon cellular stimulation via an antigen, CTSZ migrates to the cellular membrane and commences cleavage of the carboxyterminal amino acids from the β_2 subunit. **C**. Cleavage of the four carboxyterminal amino acids allows the β2 integrin to convert to its high-affinity form. As the dendritic cell matures, CTSZ migrates back to the perinuclear region. The fully matured dendritic celll is then able to interact with a T-lymphocyte via an interaction between the high-affinity β_2 integrin on the dendritic cell and inter-cellular adhesion molecule-1 on the T-lymphocyte. This interaction leads to the proliferation of T-lymphocytes required to initiate a strong immune response [177, 180].
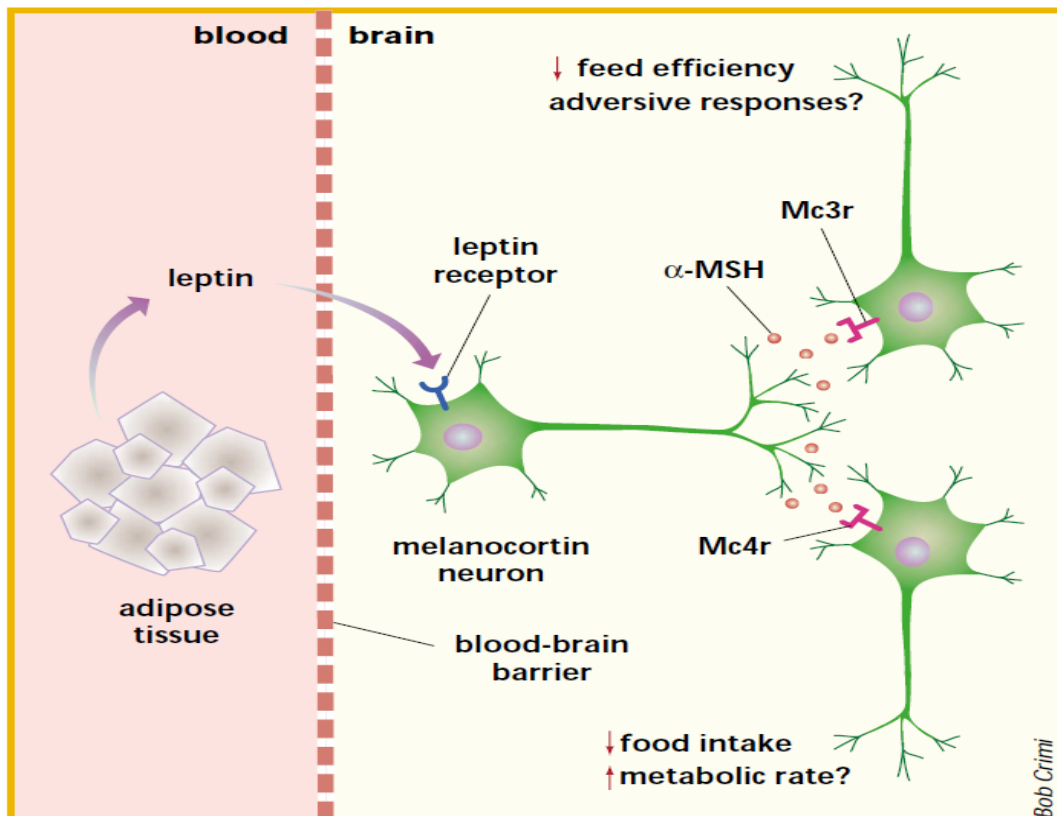
Figure 2.2 Effect of leptin on MC3R/MC4R activation: Leptin, released by adipose tissues crosses the blood/brain barrier where it binds the leptin receptor expressed on melanocortin neurons. This results in an increase in the expression of POMC and the subsequent release of MSH, an agonistic ligand of MC3R and MC4R. Activation of MC3R results in an increase in feed efficiency while activation of MC4R results in a decrease in food intake. Figure from Melanocortins and body weight: a tale of two receptors by Cummings and Schwartz [206].

A study of HIV-infected individuals in Soweto found that there was a decreasing risk of both TB morbidity and mortality with an increasing BMI, showing a dose-response effect [206]. These findings are also supported by a study of diabetic individuals conducted in 2007. Leung *et al.* [207] reported that diabetic individuals falling into either the overweight or obese BMI range had a protective effect against TB infection showing a strong linear dose-response relationship where for every unit change above a BMI of 18.5, there was a 10% reduction in risk of developing active TB. An increased BMI typically means an increase in fat stores with a corresponding increase in circulating levels of leptin. The increased leptin levels may result in increased activation of the MC3R leading to a stronger host immune response to infection.

MC3R is typically studied in combination with the melanocortin 4 receptor (*MC4R*) as these two genes have been found to work closely, but not redundantly with each other and both have been previously linked to obesity.

2.3 *MC4R*

The *MC4R* gene is found on chromosome 18q22 and encodes a 322 amino acid protein which is expressed in the brain, autonomic nervous system and the spinal cord [188, 208]. *MC4R* has shown an association to BMI which measures an individual's level of adiposity and variants in this gene show an association with early onset obesity [201, 209]. Like MC3R, the melanocortin 4 receptor protein (MC4R) transmits its signal through regulating the intracellular concentration of cyclic adenine monophosphate (cAMP) by increasing levels of adenylyl cyclase upon activation via binding of α- and β-MSH and to a lesser extent γ-MSH [188, 210]. The role of this gene in obesity has been extensively studied [198, 200, 201, 204, 211, 212]. A number of studies which, combined, screened 2 600 subjects led to the identification of 46 polymorphisms in *MC4R* found to be associated to obesity [213]. The activation of MC4R leads to a reduction in food intake (Figure 2.2) and in mice, the homozygous deletion of the gene results in an obese phenotype [213, 214]. *MC4R* is able to regulate food intake by producing a satiety signal after interaction with α-MSH (*MC4R* agonist) or an orexigenic signal after interaction with agouti-related protein [214, 215]. The exacerbated obesity seen in mice lacking both *MC3R* and *MC4R* suggest that these two genes perform non-redundant functions. The obesity caused by the increase in food intake seen in MC4R knockout mutants is exacerbated by the increase in feed efficiency caused by the lack of *MC3R*.

Not only is the prevalence of communicable diseases increasing in Africa, there is also a global trend for rising prevalence of non-communicable diseases, like diabetes [216]. It has been projected that the largest increase will occur mainly in Asia and Africa. Currently, 70% of individuals with diabetes mellitus (DM) are living in developing countries and much evidence exists which supports a link between DM and an increased risk of developing TB [216]. A detailed review of 13 studies investigating such an association found that regardless of the study design, geographical region or background TB incidence, individuals with DM were three times as likely to develop active TB compared to those without DM [217]. SNPs in *MC4R* have also been associated with an increased risk of developing diabetes and this association was more evident in women than men [209, 218].

2.4 South Africa

A major obstacle to genetic research is that frequently a gene or variant within a gene is found to be associated in one population but not in another. Inconsistencies like this may

be explained through the knowledge that with complex diseases, environmental factors can influence disease development while genetic factors can influence an individual's susceptibility. Different environmental factors may cause a genetic variant to have a larger or smaller affect in one population than in another. The complex ancestry of the Coloured population of South Africa (SAC) provides researchers a unique opportunity to study complex diseases like TB. The SAC population has a complex history of admixture receiving genetic input from Khoisan, South East Asian, Bantu-speaking and European antecedents [219]. While the possibility of population stratification can be problematic when performing case-control association studies, studies have shown that the SAC population, with its wide range of genetic input, currently represents a relatively homogenous population [108].

2.5 Summary

Due to prevalence of TB in Southern Africa and the rise of non-communicable diseases like DM in the developing world, it is imperative that further research is conducted investigating the impact of both human genetics and proper nutrition on TB susceptibility. Using samples from the SAC population, TB researchers have a unique opportunity to study TB infection and the role of host genetics in an admixed population.

2.6 Aims

The aim of this study was two-fold.

a) Initially, investigations were made to identify any significant association with TB susceptibility and selected polymorphisms in *CTSZ*, *MC3R* and *MC4R* in the SAC population.

b) Secondly, functional analysis was performed in an attempt to provide motivation for future studies to determine the role of host metabolism and nutritional status on the impact of development of active TB disease.

CHAPTER 3: Methods & Materials

3.1 Study Participants

Samples were collected from Ravensmead/Uitsig and surrounding suburbs in metropolitan Cape Town in the Western Cape province of South Africa. This area was selected due to the high incidence of TB and low prevalence of HIV infection [108]. Previous studies have found no significant stratification within this population [108]. Ethics approval for this study was obtained from the Institutional Ethics Committee (Stellenbosch University, Ethics Number 95/072 and Amendments).

For each gene, a case/control study was performed. The total number of participants in each study can be found in Table 3.1.

Table 3.1 Sample Sizes for each case-control study

Gene	Cases	Controls
CTSZ	410	301
MC3R	540	541
MC4R	510	487

The variation in sample number for each study is a result of the studies occurring separately. Samples for the *CTSZ* investigation were selected by identifying the overlap between the investigation of *MC3R* performed in this study and a previous case-control study which also investigated *CTSZ* (M Möller: unpublished results). This previous study was performed in the same SAC population but had not genotyped rs34069356. The same samples used in the *MC3R* study were selected for use in the *MC4R* study however due to sample exhaustion not all samples were available analysis. Case samples consisted of affected individuals belonging to one of two groups: Active TB patients (bacteriologically confirmed TB with clinical symptoms) or previous TB patients (previously received treatment and had no clinical symptoms at time of sampling). The control samples were collected from individuals residing in the same community, living under similar conditions including socio-economic status and availability of healthcare facilities, who were healthy at the time of sampling. Subjects used in this study were HIV-negative and unrelated. Control samples were obtained from individuals who were above the age of 16 years to ensure a high probability of exposure to *M. tuberculosis*. Age and gender matching was not performed between cases and controls.

3.2 DNA Samples

DNA was purified from whole blood using Illustra DNA Extraction Kit BACC3 (GE Healthcare, Waukesha, WI, USA) according to the manufacturer's instructions. The

concentrations of DNA samples were determined using a NanoDrop®ND-1000 Spectrophotometer and the NanoDrop® v3.0.1 Software (Thermo Fisher Scientific, Boston, MA, USA). Isolated genomic DNA was diluted in H_2O to a final concentration of 100 ng/µl and stored at -20°C.

For the *CTSZ* and *MC4R* case-control studies, whole-genome amplified DNA was used. In whole genome amplification (WGA), an identical copy of a DNA sample is generated that is indistinguishable in sequence but has a much higher concentration than the original sample [220].The use of WGA is ideal when DNA samples are limited and when multiple genes are being investigated. One reaction can generate micrograms of DNA from nanograms of starting material with an average product length of over 10 kb [220]. The DNA is replicated extremely accurately due to the use of proofreading 3' – 5' exonuclease activity of DNA polymerase [220]. An unrelated group of samples was amplified using the Illustra GenomiPhi V2 DNA Amplification Kit (GE Healthcare) according to the manufacturer's instructions. One microliter of DNA was mixed with 9 µl of Sample Buffer containing random hexamers which bind non-specifically to the single stranded DNA. The reaction mixture was denatured for 3 minutes at 95°C [220]. Following denaturation, 9 µl of master mix containing additional random hexamers, nucleotides, salts and buffers was added to the denatured DNA and samples were incubated in an isothermal reaction at 30°C for 1.5 hour to allow for DNA replication. Finally, samples were heated to 65°C for 10 minutes to terminate the reaction [220].

3.3 Primer Design

DNAman v4.1 (Lynnon Corporation, Quebec, Canada) was used to design primers with similar melting temperatures, GC content and lacking complementarity. The sequences of all primers used in this study can be found in Table 3.2. Primers were made by IDT (San Jose, USA).

3.4 Genotyping Methods

3.4.1 TaqMan

TaqMan combines polymerase chain reaction (PCR) amplification and detection into a single step by using fluorogenic probes in the 5` nuclease assay [221]. The 5' nuclease assay was originally described by Holland *et al.* [221]. The introduction of PCR meant the

introduction of a DNA amplification technique that is not only both powerful and sensitive but also has a wide range of applications [221]. The use of *Thermus aquaticus* (Taq) DNA polymerase improved both the yield and specificity of the PCR product and Holland *et al.* [221] used this enzyme to develop a method where amplification and detection could be combined into a single step. With a biallelic system, fluorogenic probes (each specific for a particular allele) are included in the PCR reaction [222]. Following PCR, fluorescence measurements are taken of each sample and computer software automatically calls each of the genotypes. Results typically show 4 distinct clusters for allele 1 homozygotes, allele 2 homozygotes, heterozygotes and samples which exhibited no amplification as can be seen in Figure 3.1 [222].
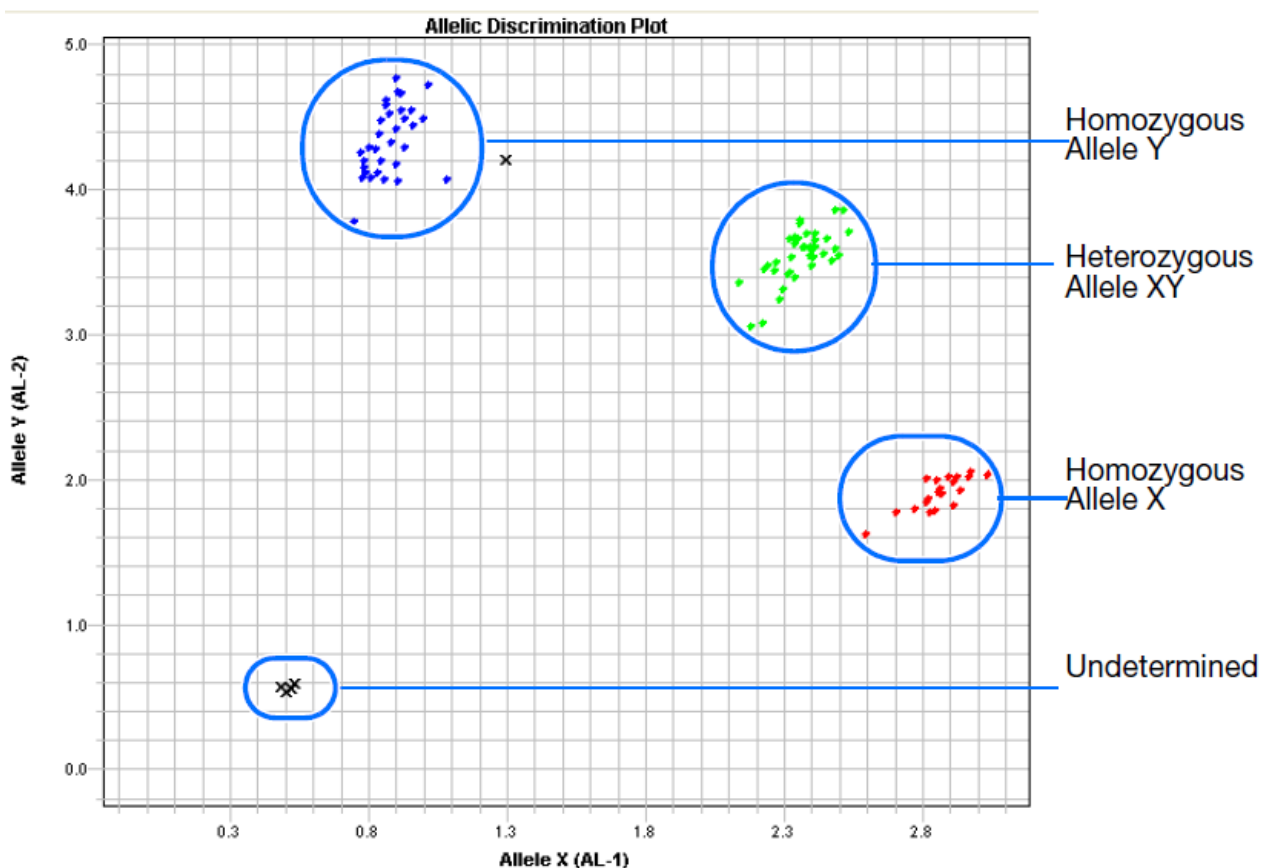
Figure 3.1: Allelic Discrimination Using TaqMan: The first of three calculations made by the software involves determining the contribution of each individual dye to the measured fluorescence spectrum. Next, these results are normalized against negative control samples run on the same reaction. From this an individual score for allele one and allele two are determined then the final step involves normalizing the extent of the reaction to a blank, also run on the same plate (From Applied Biosystems 7900HT Fast Real-Time PCR System Allelic Discrimination Getting Started Guide).

Table 3.2 Primer Sequences for all studies

Gene Study	Primer Name	Application	Sequence (5` - 3`)
CTSZ	CTSZ1 Forward	PCR/Sequencing	CTGACGGCAGCAAGGGGAGC
	CTSZ1 Reverse	PCR/Sequencing	TCCCTCCCTCCCCCACCCTT
MC3R	MC3R Upstream Forward	PCR/Sequencing/Tetra-Primer ARMS	CTTGCTCACCACTGTATTTCCT
	MC3R Upstream Reverse	PCR/Sequencing	AAGCATTCATTGCTGTCAGAAG
	MC3R Exon Forward	PCR/Sequencing	GAATGAGCATCCAAAAGACG
	MC3R Exon Reverse	PCR/Sequencing	CTGCATCCTATCCCAAGTTCAT
	SNP (-373) Forward	Tetra-Primer ARMS	TGTCTTGCCATGAAAAGAGTTT
	SNP (-373) Reverse	Tetra-Primer ARMS	CACCGGCTGCTACAGTCAC
	Tubulin Forward Primer[a]	Real-time PCR	GCCAGATGCCAAGTGAC
	Tubulin Reverse Primer[a]	Real-time PCR	CTGAAGAAGGTGTTGAAGGAAT
	MC3R Forward Primer[a]	Real-time PCR	CGTGGTGTTCATCGTCTAC
	MC3R Reverse Primer[a]	Real-time PCR	AGAACATGGTGATGAGGC
MC4R	MC4R Forward	PCR/Sequencing	ATGCGATGAGCCCCACTGTC
	MC4R Reverse	PCR	AGCTGCCTTTTCCCACCCAC
	Seq Primer 2a	Sequencing	GCAAAATATTCAGAAACTT
	Seq Primer 3	Sequencing	GACTTCAGGAGCTTCTGC
	Seq Primer 4	Sequencing	GCCTCACAACTTTCAGAC
	Seq Primer 5	Sequencing	CTGTAGCTCCTTGCTTGC
	Seq Primer 6	Sequencing	CCTCTGATTTATGCACTCCGG

[a] **Primers designed and prepared by Roche**

TaqMan was used to genotype 1 SNP in *CTSZ* and 3 SNPs in *MC4R*. The reaction was set up according to an adjusted TaqMan protocol developed and verified in house. The manufacturer's instructions (Applied Biosytems®, Foster City, USA) require 1 to 20 ng of purified genomic DNA to be used in the reaction, however, in both the *CTSZ* and *MC4R* studies, 100 ng of WGA-DNA was used instead. Using the Eppendorf EpMotion® 5070, TaqMan plates were prepared with 5 µL of 20 ng/µL WGA DNA, 2.5 µL Master Mix, 0.0625 µL SNP specific TaqMan assay and 2.4375 µL ddH$_2$O samples for a final reaction volume of 10 µL. PCR was carried out on a GeneAMP® PCR System 9700 (Applied Biosystems) and the protocol was as follows: 95°C for 10 minutes, 45 cycles of 95°C for 10 seconds then 60°C for 1 minute, followed by a hold at 10°C. Immediately after amplification, fluorescence of each reaction was measured on the 7900HT Fast Real-Time PCR System (Applied Biosystems). Sequence Detection Systems v2.3 software (Applied Biosystems) was used to automatically call genotypes.

3.4.2 Direct Sequencing

3.4.2.1 *MC3R* Initial Screening

To determine which *MC3R* polymorphisms were present in the SAC population, the single exon of *MC3R* and 1000 bp upstream of the gene were sequenced in 10 cases and 10 controls. These two regions were first amplified by PCR. Each reaction contained 100 ng of genomic DNA, 2.5 µl of 10X Reaction Buffer containing 15mM MgCl$_2$ (Southern Cross Biotechnologies, Cape Town, South Africa), 1 µl of 2.5 mM dNTPs (Bioline, London, UK), 1 µl each of forward and reverse primers (all primers used in this study were of a concentration of 10 nM) (Integrated DNA Technologies, Coralville, USA) and 0.075 µl of Super-Therm Gold DNA polymerase (Southern Cross Biotechnologies). Reactions were made up to a volume of 25 µl with sterile ddH$_2$O. MC3R Upstream Forward and MC3R Upstream Reverse primers were used to amplify the upstream region and the exon was amplified by MC3R Forward and MC3R Reverse primers (Table 3.2). An Eppendorf Mastercycler® PCR System (Eppendorf, Hamburg, Germany) was used for the following cycling program: 10 minutes of denaturation at 95°C followed by 30 cycles of 1 minute at 94°C, 1 minute at 54.1°C and 2 minutes at 72°C. The reaction was terminated at 72°C for 10 minutes. To control for contamination, a negative control sample containing only master mix was included in every run.

Successful amplification was verified via agarose gel electrophoresis. A 1% agarose sodium borate (SB) gel was prepared with 1 g of agarose (Seakem, Rockland, USA), 100 ml of 1X SB buffer and 0.5 µl of 10 mg/ml ethidium bromide (Invitrogen, San Diego USA). Five microliters of PCR product together with 5 µl of 2X loading buffer (0.25% bromophenol blue, 0.25% xylene cyanol FF, 30% glycerol in water) were loaded onto the gel. Five microliters of a 100 bp DNA ladder (Thermo Fisher Scientific) was electrophoresed with the samples as a molecular weight marker. The samples were run for 1 hour at 150 V and the gel was photographed with UV light using G:Box (Syngene, Frederick, USA).

Prior to sequencing, PCR products were purified using ExoSAP-IT clean up (Affymetrix, Santa Clara, USA) according to the manufacturer's instructions. Five microliters of amplicon was mixed with 2 µl ExoSAP-IT and the sample was incubated in an Eppendorf Mastercycler®ep for 37°C for 15 minutes to degrade remaining primers and nucleotides. This was followed by 80°C for 15 minutes to inactivate ExoSAP-IT, 2 minute incubation at 5°C and a 15°C hold. The fragment was sequenced in both forward and reverse direction by the Central Analytical Facility (CAF) (Stellenbosch, South Africa). Sequencing data was analyzed using Sequencher v4.10 (Gene Codes Corporation, Ann Arbour, USA). To determine which polymorphisms had been previously reported, the sequence containing the polymorphism was aligned to the human genome using Blat (http://genome.ucsc.edu/cgi-bin/hgBlat). If a polymorphism had been previously reported Blat listed the rs number (a unique record identifier assigned by Single Nucleotide Polymorphism Database (dbSNP)). This database was established by the National Human Genome Research Institute and The National Center for Biotechnology Information (NCBI). SNPs selected for genotyping can be found in Table 3.3.

Table 3.3 SNPs genotyped in *CTSZ*, *MC3R* and *MC4R*

GENE	rs Number	Position	Genotyping Method
CTSZ	rs34069356	Exon 6, missense	Taqman
MC3R	rs72650656	Promoter	Sequencing
	rs6127698	Promoter	Sequencing
	rs11575886	Promoter	Sequencing
	rs72650657	Promoter	Sequencing
	rs72650658	Promoter	Sequencing
	rs3827103	Exon 1, missense	Sequencing
MC4R	rs11872992	Promoter	Taqman
	rs8087522	Promoter	Taqman
	rs34114122	5` UTR[a]	Taqman

[a] Untranslated Region

The single SNP selected for the investigation of *CTSZ* in this study was the same SNP identified by Cooke *et al.* [82] which showed evidence of association to TB susceptibility. SNPs were selected for the *MC3R* study if they were identified in the SAC population during the initial screening. A similar process was used for the selection of *MC4R* SNPs, with the three SNPs showing the highest prevalence in the SAC population being selected for genotyping.

3.4.2.2 *MC3R* Genotyping

An unrelated case-control study was performed by genotyping TB case and control samples via direct sequencing as described in Section 3.4.2.1. Purification of PCR product was performed by CAF. Variations detected in the initial screen were automatically genotyped in all the samples using the software SGcaller developed by Christian-Albrechts-University (Kiel, Germany) [223].

3.4.2.3 *MC4R* Initial Screening

To determine which *MC4R* polymorphisms were present in the SAC population, the single exon of *MC4R* and 1000 bp upstream of the gene were sequenced in 10 cases and 10 controls. One 3300 bp fragment was amplified via PCR. Each reaction contained 100 ng of genomic DNA, 2.5 µl of 10X reaction buffer containing 15mM $MgCl_2$ (Southern Cross Biotechnologies), 1 µl of 2.5 mM dNTPs (Bioline), 1 µl of MC4R Forward primer, 1 µl MC4R Reverse primer (Integrated DNA Technologies) and 0.075 µl of Super-Therm Gold DNA polymerase (Southern Cross Biotechnologies). Reactions were made up to a volume of 25 µl with ddH_2O. An Eppendorf Mastercycler PCR System was used for the following cycling program: 10 minutes of denaturation at 95°C followed by 40 cycles of 1 minute at 94°C, 2 minute at 65°C and 2.5 minutes at 72°C. The reaction was terminated at 72°C for 10 minutes. To control for contamination, a negative control sample containing only master mix was included in every run. The remainder of the screen was performed as described in Section 3.4.2.1. Polymorphisms selected for genotyping can be found in Table 3.3.

3.4.3 Tetra-primer Amplification Refractory Mutation System

Genotyping of the *MC3R* SNP associated with TB susceptibility in volunteer blood samples was performed by Tetra-primer Amplification Refractory Mutation System (ARMS) as previously described by Ye *et al.* [223]. Tetra-primer ARMS is a reliable method which

genotypes a sample in a single PCR reaction. The tetra-primer ARMS concept is depicted in Figure 3.2. Previous studies have shown that this method is robust and reproducible using different thermal cyclers in different laboratories with fully concordant results [224]. Each reaction contained 100 ng of genomic DNA, 2.5 µl of 10X reaction buffer containing 1.5 mM MgCl$_2$ (Southern Cross Biotechnologies), 1 µl of 2.5 mM dNTPs (Bioline), 0.5 µl of MC3R Upstream Forward primer, 0.5 µl MC3R Upstream Reverse primer, 4 µl of SNP(-373) Forward primer, 2 µl of SNP(-373) Reverse primer, 0.05 µl of Super-Therm Gold DNA polymerase (Southern Cross Biotechnologies). Reactions were made up to a volume of 25 µl with sterile ddH$_2$O. An Eppendorf Mastercycler PCR System was used for the following cycling program: 10 minutes of denaturation at 95°C followed by 30 cycles of 1 minute at 94°C, 1 minute at 58°C and 1 minute at 72°C. The reaction was terminated at 72°C for 10 minutes. To control for contamination, a negative control sample containing only master mix was included in every run. A positive control for each genotype was run with every reaction. Gel electrophoresis was performed as described in Section 3.4.2.1.

3.5 Statistical Analysis

Epi Info^{TM} (CDC, Atlanta, USA) was used to determine the sample size required to generate results with 80% power and a 95% confidence interval (CI). Assuming an allele frequency of at least 5%, a sample size of approximately 500 cases and 500 controls was selected for all case-control studies to enable the detection of an odds ratio (OR) = 2.0. For each SNP that was genotyped, Hardy-Weinberg equilibrium (HWE) was assessed. Prism v4.02 (GraphicPad Software, La Jolla, USA) was used to calculate the OR, 95% CI and p-values for the single point analysis. It was used to determine the mean and standard deviation of results obtained from functional analysis. Genotypic distribution between cases and controls was calculated by either χ^2 test or Fisher's exact test. For SNPs where all 3 genotypes were represented in both cases and controls, a p-value was determined using the χ^2 test. The Fisher's exact test was used when the allelic frequency was low resulting in at least one of the three possible genotypes being absent in either the cases or controls. A significant p-value was defined as < 0.05. Haploview v4.1 (Broad Institute, Cambridge, USA) was used to evaluate differences in haplotypes between cases and controls and assess which SNPs were in LD. Permutation testing was performed to determine if the observed difference in allele frequency was greater than could have occurred by chance [108] .
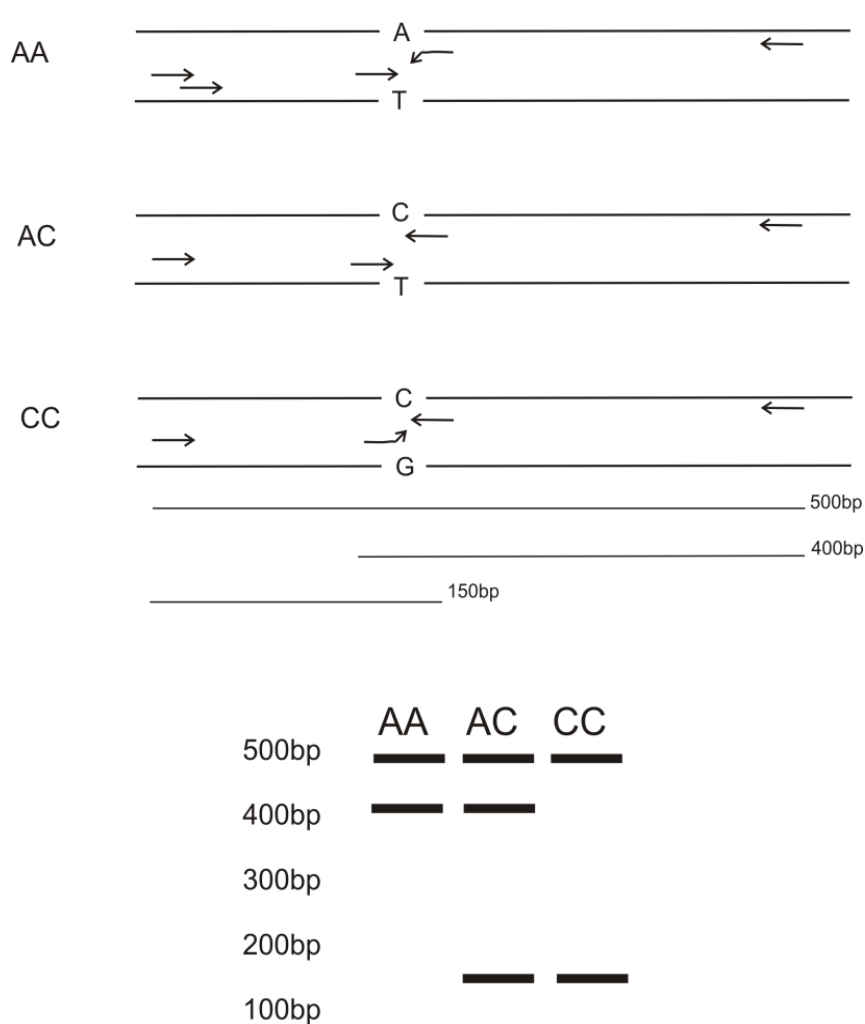
Figure 3.2 Concept of Tetra-primer ARMS: One primer set (the outer set) will be amplified in every reaction to act as a positive control to ensure success of the reaction. These outer primers are designed in such a way that they are asymmetrically located around the SNP to ensure that the allele specific amplicons differ enough in size to be easily separated by gel electrophoresis [222]. The nested primer set is allele specific and will amplify fragments only if a particular allele is present. 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.

3.6 Functional Analysis

3.6.1 BCG Cultures and Stocks

Five milliliters of starter culture of BCG obtained in the division of molecular biochemistry and human genetics was inoculated into 35 ml of Middlebrook 7H9 broth (BD Scientific, Sparks, USA) enriched with 0.2% glycerol, 0.1% Tween80 and 10% Middlebrook oleic-albumin-dextrose-catalase (OADC) (BD Scientific) and cultured until optical density (OD) reached between 0.8 and 0.9 as measured at 600 nm. Thirty milliliters of BCG culture was transferred to 50 ml Falcon tube and centrifuged at 3000 rpm for 5 min at room temperature (RT). The supernatant was discarded and the pellet was resuspended in 20 ml sterile 15% glycerol. Stocks were prepared by aliquoting 1 ml of resuspended culture into 1.5 ml centrifuge tubes which were stored at -80°C for future use. Bacterial concentrations were ascertained by determining the number of colony forming units per

milliliter by plating 100 µl aliquots of a serial dilution of 100%, 10%, 1% and 0.1% BCG stock on agar plates. Plates were prepared according to the manufacturer's instructions. Each plate contained 50 ml of Middlebrook 7H11 Agar enriched with 0.5% glycerol and 10% Middlebrook OADC (BD Scientific). Plates were incubated at $37^{o}C$ for 2.5 weeks before determining the colony forming units per milliliter stock solution.

3.6.2 Culturing Monocyte-Derived Macrophages

Monocyte-derived macrophages (MDMs) were isolated from peripheral blood mononuclear cells (PBMCs) obtained from fresh human blood. Peripheral venous blood samples were collected from healthy adult volunteers using Vacutainers containing EDTA and DNA was isolated as previously described (Section 3.2). Volunteers were genotyped using tetra-primer ARMS (Section 3.4.3) for the *MC3R* promoter SNP found to be associated with TB susceptibility. Individuals homozygous for either the T or G allele were then sequenced by the CAF to detect any additional polymorphisms either in the single exon of the gene or up to 1000 bp upstream. Volunteers with no additional polymorphisms in the sequenced region were selected for functional analysis. Two volunteers with the GG genotype and one volunteer with the TT genotype were selected. Forty milliliters of peripheral blood was collected into heparinised tubes (BD Scientific) and PBMCs were isolated by Ficoll-Histopaqe density gradient centrifugation. Twenty milliliters of heparinised fresh whole blood was made up to a final volume of 35 ml with GIBCO® GlutaMAX™ media (Invitrogen) and pipetted over 15 ml of Histopaque (Sigma-Alrich, St. Louis, USA). Samples were centrifuged at 400g for 25 minutes at RT with no brake or accelerator using an Eppendorf Centrifuge 5810R. Mononuclear cells were harvested with a pipette and washed twice with GIBCO® GlutaMAX™ media. Cells were pelleted between washes by centrifugation at 250g for 5 minutes. After the final wash the cell pellet was resuspended in 2 ml GIBCO® GlutaMAX™ media enriched with 10% Human AB Serum (Sigma-Aldrich) referred to henceforth as complete media.

PBMCs were seeded into 6-well tissue culture plates (Nunc, Rochester, USA) at 1×10^{7} cells/ml and were incubated in an autoflow CO_2 water jacketed incubator (Thermofisher Scientific) at $37^{o}C$ containing a humidified atmosphere and 5% CO_2 for 16h to allow mononuclear cells to adhere to the plate. Following the incubation period the conditioned media was removed and centrifuged at 700g for 5 minutes to remove all non-adherent cells and cellular debris. This supernatant was transferred to a fresh 15 ml centrifuge tube. Adherent cells were then stimulated with BCG at a multiplicity of infection of 2 for 4 hours.

Following stimulation, medium was removed and wells were washed 3 times with GIBCO[®] GlutaMAX™ medium to remove all non-phagocytosed bacteria. Complete medium enriched with 10% conditioned medium was added at 2 ml per well and cells were cultured for 6 days. Following the incubation period, RNA was extracted.

3.6.3 Analysis of RNA transcript levels

The use of reverse transcription of RNA into cDNA followed by quantitative PCR (qPCR) is a powerful method used to detect minute levels of RNA transcripts [225]. Using qPCR, RNA transcript levels can be determined using absolute quantification which requires a calibration curve, meticulously validated methodology, identical amplification efficiencies for target and control genes and must be run on the same LightCycler[®] System [226]. An alternative method of RNA quantification normalizes the target gene against an endogenous standard and is currently the recommended method for analysis of gene expression. In this method, the target gene expression is normalized against the expression of a reference gene which is typically a non-regulated reference gene or a housekeeping gene which is present in all nucleated cell types [226]. In this study, *MC3R* gene expression was normalized against the expression of the housekeeping gene *α-tubulin* as has been previously used in studies investigating MC3R protein expression in macrophages [190].

Following the qPCR reaction, two main mathematical models can be applied to analyze the results. Both will be discussed below.

3.6.3.1 RNA extraction

RNA was extracted using TRIzol[®] reagent (Invitrogen) according to the manufacturer's instructions. Following a 6 day culture, cells were washed with fresh media to remove any cellular debris. Cells were lysed with 1 ml of TRIzol and cells were removed using a rubber policeman and lysate was transferred to a 1.5 ml centrifuge tube. Lysates were incubated at RT for 5 minutes after which 200 µL of chloroform was added and samples were vortexed. Samples were incubated at RT for 5 minutes followed by centrifugation for 15 minutes at 12 000g at 4°C. Immediately following centrifugation, the top clear organic phase was removed and added to 500 µl ice cold isopropanol in a new 1.5 ml centrifuge tube. Samples were vortexed and incubated for 15 minutes at RT then centrifuged for 15 minutes at 12 000g at 4°C. The supernatant was discarded and the RNA pellet was washed with 1 ml RNase free 70% ethanol. Samples were vortexed and centrifuged for 5

minutes at 7500*g* at 4°C. The supernatant was removed and pellets were allowed to air dry at RT. Once all ethanol had evaporated, RNA pellets were resuspended in 20 µl RNase free water and the OD was measured using a NanoDrop®ND-1000 Spectrophotometer and the NanoDrop® v3.0.1 Software (Thermo Fisher Scientific).

3.6.3.2 RT-PCR

Purified RNA was reverse transcribed to cDNA using the Transcriptor First Strand cDNA Synthesis Kit (Roche, Basel, Switzerland) according to the manufacturer's instructions. All reagents were centrifuged before use and the entire reaction was set up on ice. One microgram of RNA was added to 2 µl of Random Hexamer Primer and the volume was made up to 13 µl with PCR-grade water. Samples were denatured by incubation for 10 minutes at 65°C in an Eppendorf Mastercycler PCR System. Samples were immediately cooled on ice. To each sample, 4 µl of Transcriptor Reverse Transcriptase Reaction buffer, 0.5 µl Protector RNase Inhibitor, 2 µl of dNTPs (10mM) and 0.5 µl of Transcriptor Reverse Transcriptase was added to each sample. All samples were incubated for 10 minutes at 25°C followed by second incubation for 30 minutes at 55°C. The reaction was terminated with a 5 minute incubation at 85°C. The concentration of cDNA was determined using a NanoDrop®ND-1000 Spectrophotometer and the NanoDrop® v3.0.1 Software (Thermo Fisher Scientific).

3.6.3.3 qPCR

Prior to the analysis of the RNA transcript levels, the *MC3R* and *α-tubulin* assays were optimized to have an assay efficiency between 80 and 120%. Assays were optimized using a standard curve containing the following dilution series: 100 ng, 75 ng, 50 ng, 25 ng, 12.5 ng, 10 ng, 6.25 ng and 0 ng. *MC3R* RNA transcripts were quantified using LightCycler® FastStart DNA Master SYBR Green I Kit (Roche) according to the manufacturers' instructions. The reaction was repeated with primers specific for *α-tubulin* to act as an internal control. To measure MC3R transcript levels, a 62 bp fragment of the transcript was amplified with 1 µl of each PCR primer (0.5 µM), 2 µl dNTPs, 1.5 µl MgCl$_2$, 2.5 µl LightCycler® FastStart DNA Master SYBR Green I (Roche) and 5 µl PCR grade H$_2$O which were mixed and transferred to a pre-cooled LightCycler® Capillary (Roche). To each capillary, 100 ng cDNA was added and each capillary was sealed with a stopper. Capillaries were centrifuged at 0.7*g* prior to being transferred to the LightCycler® Sample Carousel (Roche) and placed in the LightCycler® instrument. A LightCycler® 1.5 (Roche) was used for the following cycling program: 15 minutes of enzyme activation at 95°C

followed by 50 cycles of 15 seconds at 95°C, 15 seconds at 50°C and 8 seconds at 72°C. Melt curve analysis was performed using the following reaction conditions: an incubation of 0s at 95°C, a 15s incubation at 65°C with temperature increase up to 95°C at a ramp rate of 0.1°C/second. The reaction was terminated at 40°C for 10 minutes. To amplify a 61 bp fragment of the α-tubulin transcript, 2 µl of each PCR Primer (0.5 µM), 2 µl dNTPs, 1.5 µl MgCl$_2$, 2 µl LightCycler$^®$ FastStart DNA Master SYBR Green I (Roche) and 8.5 µl PCR grade H$_2$O were mixed and transferred to a pre-cooled LightCycler$^®$ Capillary (Roche). To each capillary, 65 ng cDNA was added and each capillary was sealed with a stopper. Capillaries were centrifuged at 0.7g prior to being transferred to the LightCycler$^®$ Sample Carousel (Roche) and placed in the LightCycler$^®$ instrument. A LightCycler$^®$ 1.5 (Roche) was used for the following cycling program: 15 minutes of enzyme activation at 95°C followed by 40 cycles of 15s at 95°C, 15s at 61°C and 8s at 72°C. Melting curve analysis was performed using the following reaction conditions: an incubation of 0s at 95°C, 15s incubation at 65°C with temperature increase up to 95°C at a ramp rate of 0.1°C/second. The reaction was terminated at 40°C for 10 minutes. To ensure success of the reactions, a positive control sample was included in every run. Contamination was controlled for by the inclusion of a negative control sample containing only master mix in every run.

3.6.3.4 Livak Method

RNA quantification based on the average threshold cycle (C$_t$) value is performed without correction for differences in amplification efficiency. This method assumes that the amplification product is less than 150 bp and the primer and Mg^{2+} concentrations have been properly optimized, yielding an assay efficiency close to 1 [227]. With this method the equation used to analyze the relative gene expression is:

$$amount_{t\arg et} = 2^{-\Delta\Delta C_P}$$

(eq. 1)

The entire derivation of this equation can be found in a paper published by Livak and Schmittgen in 2001 [227]. For this method to be valid the amplification efficiency for the target and the reference genes must be approximately equal. A final value of 1 indicates there is no change in gene expression while a value of less than one indicates a decrease in expression and a value larger than one indicates an increase in expression.

3.7.3.5 Pfaffl Method

The mathematical model developed by Pfaffl allows for differences in the amplification efficiency to be incorporated into the RNA quantification [226]. When a target gene is normalized against a single reference gene one uses the equation:

$$ratio = \frac{\left(E_{t\,arg\,et}\right)^{\Delta CP_{t\,arg\,et}\,(control-sample)}}{\left(E_{reference}\right)^{\Delta CP_{reference}(control-sample)}}$$

(eq. 2)

This ratio compares the change in expression of a target gene in a treated sample compared to a control sample versus the change in expression of a reference gene in a treated sample compared to a control sample. E_{target} and $E_{reference}$ refer to the predetermined amplification efficiency of the target and the reference genes, respectively [226]. Amplification efficiencies are calculated from the slope produced by the standard curve used during reaction optimization.

CHAPTER 4: Results

4.1 Introduction

To investigate the association of polymorphisms in *CTSZ, MC3R* and *MC4R* to TB susceptibility, 3 independent case-controls studies were performed using an SAC population previously shown to be free of population stratification [82]. A number of SNPs which were all found to be in HWE were selected for genotyping (Table 3.3). Haplotype analysis was performed on each gene for the SNPs genotyped. Although *CTSZ* and *MC3R* are within close proximity on chromosome 20, the two genes were determined to be in linkage equilibrium and were subsequently analyzed separately. Following the detection of a statistically significant association between a promoter polymorphism in *MC3R*, functional analysis was performed to investigate the influence of this SNP on gene expression when stimulated with BCG.

4.2 *CTSZ*

4.2.1 Single Point Analysis

For *CTSZ*, the same SNP reported by Cooke *et al.* [82] to be associated in West Africa (rs34069356) also showed evidence of disease association in an SAC population with the AA genotype being over represented in TB patients (genotypic OR 0.28(0.17-0.45); p <0.0001) (Table 4.1). This remained significant even after stringent Bonferroni correction for multiple testing. This SNP is located within the last (sixth) exon of the gene and is a missense mutation resulting in an amino acid change from alanine to threonine.

4.2.2 Haplotype Analysis

Four SNPs from a previous study investigating the association of *CTSZ* to TB susceptibility in the SAC population, *viz.* rs448943, rs10369, rs13720 and rs3787492, were included in the haplotype analysis of *CTSZ* (Adams, submitted) (Table 4.2). The SNPs rs34069356 and rs3787492 was excluded from the analysis since their allele frequencies were below 1% (Figure 4.1). Three of the markers previously genotyped in *CTSZ*, rs448943, rs10369 and rs13720, were identified as a single block. Following haplotype analysis there did not appear to be any association with TB. These results were repeated when analyzed with the Cocaphase program in the UNPHASED suite (LITBIO, Italy).

Table 4.1 Single Point Analysis of *CTSZ*

SNP	Group	N^a	Genotype Frequency			Genotypic		Allelic	
			1^b1	12	22	p-value [c]	OR^d (95% CI)	p-value	OR (95% CI)
CTSZ									
rs34069356 A/G	TB	410	0.00	0.23	0.77	<0.0001	0.28 (0.17 – 0.45)	<0.0001	0.31 (0.19 – 0.49)
	Controls	301	0.00	0.08	0.92				

[a] number of individuals investigated
[b] Allele 1 is the minor allele
[c] p-value based on a genotype-based Fisher exact test
[d] odd ratio calculated by Prism

Table 4.2 Haplotype analysis of CTSZ

GENE	Haplotype		Haplotype Frequency		χ^2	P value	P value$_{permutation}$ [b]	OR (95% CI)
			Cases	Controls				
	Block[a]: rs448943 – rs10369 – rs13720							
	N1	A-C-A	0.337	0.342	0.05	0.8228	0.9999	1
CTSZ	N2	G-C-A	0.258	0.255	0.02	0.8894	1.0000	0.95 (0.74 – 1.22)
	N3	G-T-A	0.242	0.202	3.29	0.0697	0.3706	1.19 (0.92 – 1.55)
	N4	A-C-G	0.163	0.199	3.12	0.0775	0.4003	0.92 (0.70 – 1.20)

[a] The order of the SNPS in detected haplotype block corresponds to Figure 4.1
[b] Permutation test p values were calculated from 10 000 permutations in Haploview.

Figure 4.1 LD Plots of *CTSZ* SNPs analyzed: LD plot generated by Haploview v4.1. r^2-values (%) are indicated on the squares. The colours of the squares represent LD strength according to D` values, with red being strong and blue being weak. Haplotype block is outlined in black. CTSZ_34069356 and CTSZ_3787492 were excluded as allele frequencies were less than 1%.
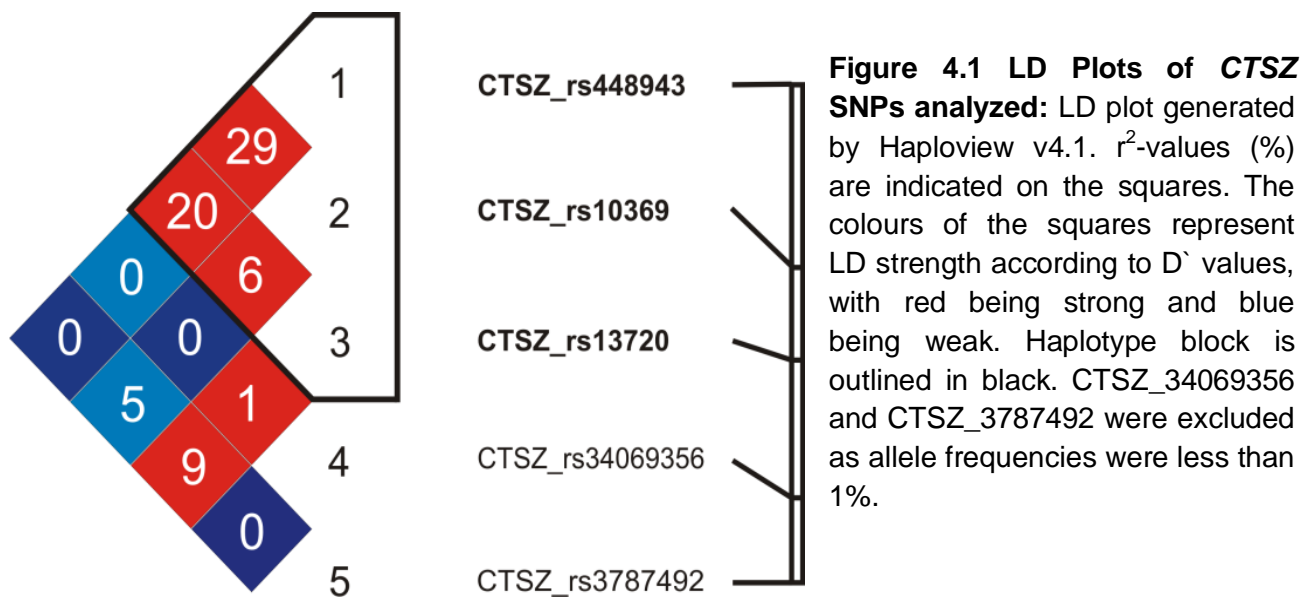
4.3 *MC3R*

4.3.1 Single Point Analysis

One of the SNPs in *MC3R*, rs6127698, is located 373bp upstream of the start codon and showed statistically significant difference in allelic frequency distribution between TB patients and controls (Table 4.3). The least common allele, T, was found more frequently in controls than in cases, indicating a possible resistance phenotype for the TT genotype. This association remained significant after adjusting for multiple testing (p = 0.0039, OR of the GG genotype for disease, 1.5 (1.18 – 1.92)).

4.3.2 Haplotype Analysis

Four of the SNPs in *MC3R, viz.* rs72650656, rs11575886, rs72650657 and rs72650658, were excluded from the analysis since their allele frequencies were below 1%. SNPs rs6127698 and rs3827103 in *MC3R* were grouped as one haplotype block (Figure 4.2). Within this block, three haplotypes were identified with haplotype N1 being the most frequent (Table 4.4). Haplotype N3 showed a significant association with case-control status (p=0.0008) which was observed only 61 out of 10 000 times during permutation testing done to correct for multiple testing (Table 4.4). A similar result was found with the

Table 4.3 Single Point Analysis of *MC3R*

SNP	Group	N[a]	Genotype Frequency 1[b]1	12	22	Genotypic p-value	OR[e] (95% CI)	Allelic p-value	OR (95% CI)
MC3R									
rs72650656 G/C	TB	534	0.00	0.01	0.99	1.0000[c]	0.88 (0.31 – 2.45)	0.9854	0.88 (0.32 – 2.44)
	Controls	529	0.00	0.01	0.99				
rs6127698 T/G	**TB**	**540**	**0.07**	**0.32**	**0.61**	**0.0039[d]**	**1.50 (1.18 – 1.92)**	**0.0012**	**1.39 (1.14 – 1.68)**
	Controls	**540**	**0.08**	**0.39**	**0.51**				
rs11575886 C/T	TB	537	0.01	0.06	0.93	0.1240[c]	1.19 (0.74 – 1.92)	0.8333	1.08 (0.68 – 1.71)
	Controls	540	0.00	0.07	0.93				
rs72650657 G/C	TB	537	0.00	0.04	0.96	0.5666[c]	1.2 (0.69 – 2.15)	0.5892	1.21 (0.69 – 2.12)
	Controls	541	0.00	0.05	0.95				
rs72650658 G/C	TB	532	0.00	0.02	0.98	0.1767[c]	0.40 (0.12 – 1.27)	0.1823	0.40 (0.12 – 1.28)
	Controls	530	0.00	0.01	0.99				
rs3827103 G/A	TB	529	0.14	0.42	0.44	0.1751[d]	0.84 (0.66 – 1.07)	0.1056	0.88 (0.72 – 1.03)
	Controls	529	0.10	0.40	0.49				

[a] number of individuals investigated
[b] Allele 1 is the minor allele
[c] p-value based on a genotype-based Fisher exact test
[d] p-value based on a genotype-based χ^2 test
[e] odd ratio calculated by Prism

Cocaphase program in the UNPHASED suite which also found a significant association with the same haplotype in *MC3R*.
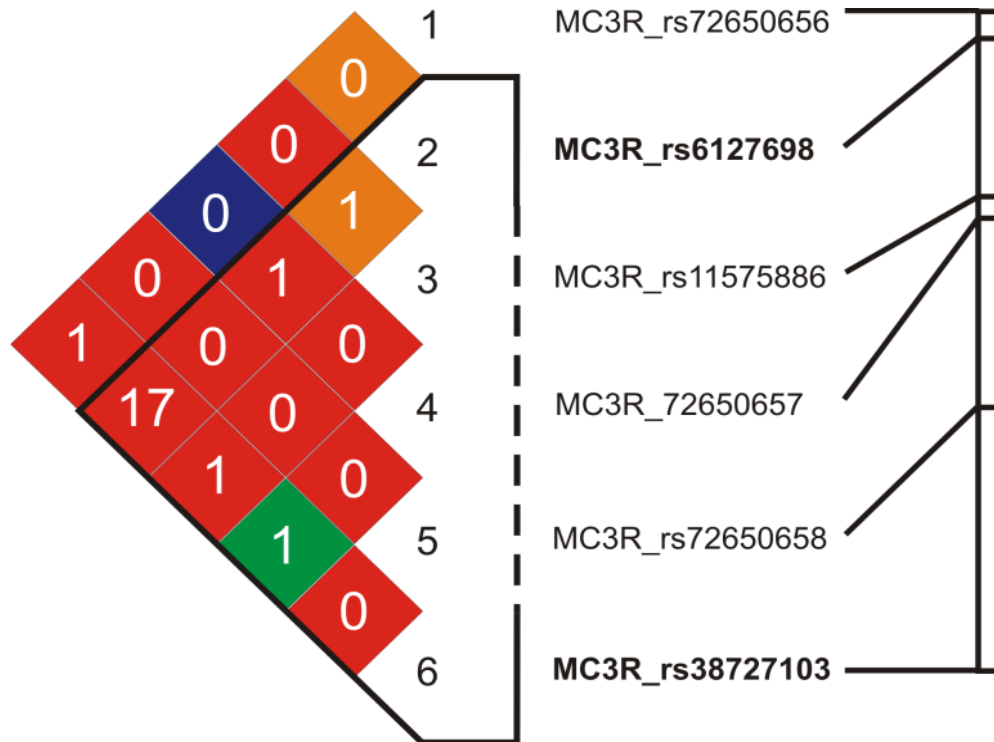

Figure 4.2 LD Plots of *MC3R* SNPs analyzed: LD plot generated by Haploview v4.1. r^2-values (%) are indicated on the squares. The colours of the squares represent LD strength according to D` values, with red being strong and blue being weak. Haplotype block is outlined in black. MC3R_rs72650656, MC3R_rs11575886, MC3R_rs72650657 and MC3R_rs72650658, were excluded as allele frequencies were less than 1%.

4.3.3 RNA Transcript Analysis

Prior to transcript analysis, the assay was optimized using a standard curve to determine the assay efficiency for both the target and the reference genes. Assay efficiencies can be found in Table 4.5. During assay optimization, varying quantities of cDNA were used to select the ideal amount for each the *MC3R* and *α-tubulin* assays. Differences in the concentrations used for each of the assays can be attributed to differences in the overall amount of each gene's transcript present in the total RNA isolated and differences in primer melting temperature. In order to optimize amplification of both assays when run with an annealing temperature of 61°C, different amounts of RNA transcript were required for each assay.

After RNA was isolated as described in Section 3.6.3.1, cDNA was prepared via a reverse transcription reaction as described in Section 3.6.3.2. cDNA was subsequently used for a qPCR reaction to determine relative gene expression of the *MC3R* gene normalized against the endogenous housekeeping gene (*α-tubulin*) (Section 3.6.3.3). Each experiment was performed in triplicate and each assay was run in duplicate. Melt curve analysis indicates a lack of primer dimer in all samples (Figure 4.3). The amplification curves show an absence of amplification in the blank suggesting there was no nonspecific amplification in the samples (Figure 4.4).

Figure 4.3 Melt Curve Analysis: Melt curve analysis was performed with each run to ensure that the reaction contained only a single amplicon. All samples showed a single peak indicating that samples lacked primer-dimer formation and additional non-specific amplification.
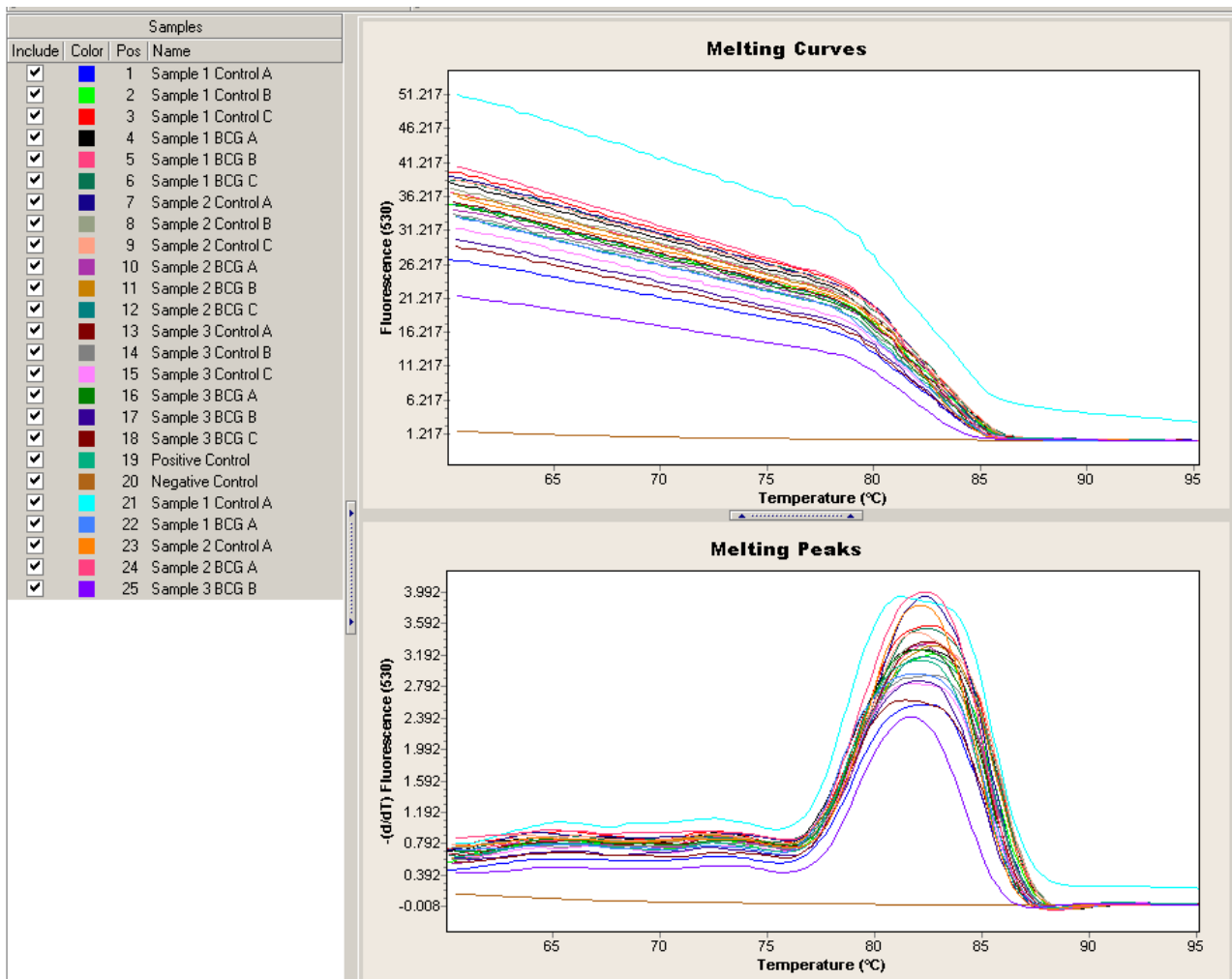
Table 4.4 Haplotype analysis of *MC3R*

GENE	Haplotype		Haplotype Frequency		χ^2	P value	P value$_{permutation}$[b]	OR (95% CI)
			Cases	Controls				
	Block[a]: rs6127698 – rs3827103							
MC3R	N1	G-G	0.429	0.403	1.46	0.2272	0.7777	0.95 (0.79 – 1.17)
	N2	G-A	0.3464	0.308	3.41	0.0647	0.3469	1
	N3	T-G	0.225	0.288	11.22	0.0008	0.0061[c]	0.71 (0.56 – 0.89)

[a] The order of the SNPS in detected haplotype block corresponds to Figure 4.2
[b] Permutation test p values were calculated from 10 000 permutations in Haploview.
[c] p values less than 0.0008 were observed on 61 out of 10 000 occasions (0.61%) during permutation testing.

Table 4.5 Assay Efficiency: Efficiencies for each assay was determined based on the slope produced from a standard curve during assay optimization.

Gene	Efficiency Value
Target (*MC3R*)	1.936
Reference (*α-tubulin*)	1.885

Figure 4.4 Amplification Curve: The amplification curve indicates that each sample reaction ran to completion and the lack of amplification in the negative control sample shows absence of non-specific amplification, the presence of which could negatively impact the calculated C_T values.
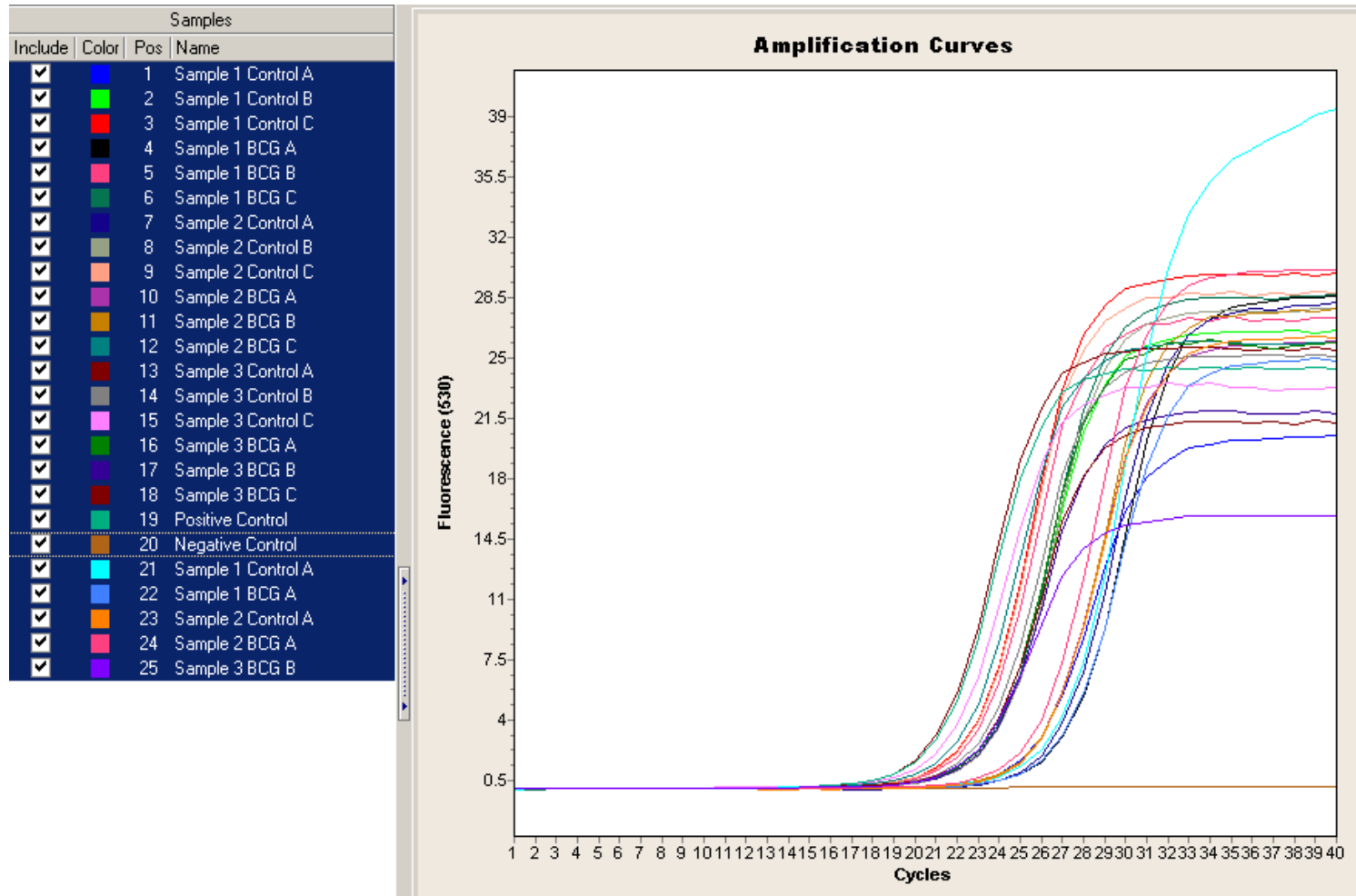
4.3.3.1 Livak Method

The average change in gene expression between macrophages stimulated with BCG and not stimulated with BCG was determined. Using the Livak method (equation 1), the average change in *MC3R* gene expression was normalized against the average change in *α-tubulin* expression. The change in gene expression was calculated for each qPCR reaction. These results were then analyzed by Prism to determine the mean and the standard deviation for each genotype. The results for the analysis using the Livak method can be seen in Figure 4.5. While results did not show a statistically significant difference in the change in *MC3R* gene expression between the GG and TT genotypes, both showed an upregulation in *MC3R* gene expression upon stimulation with BCG. The TT genotype showed a trend for an even larger increase in *MC3R* transcription than the GG genotype, suggesting that upregulation of this gene may be associated with resistance in TB.
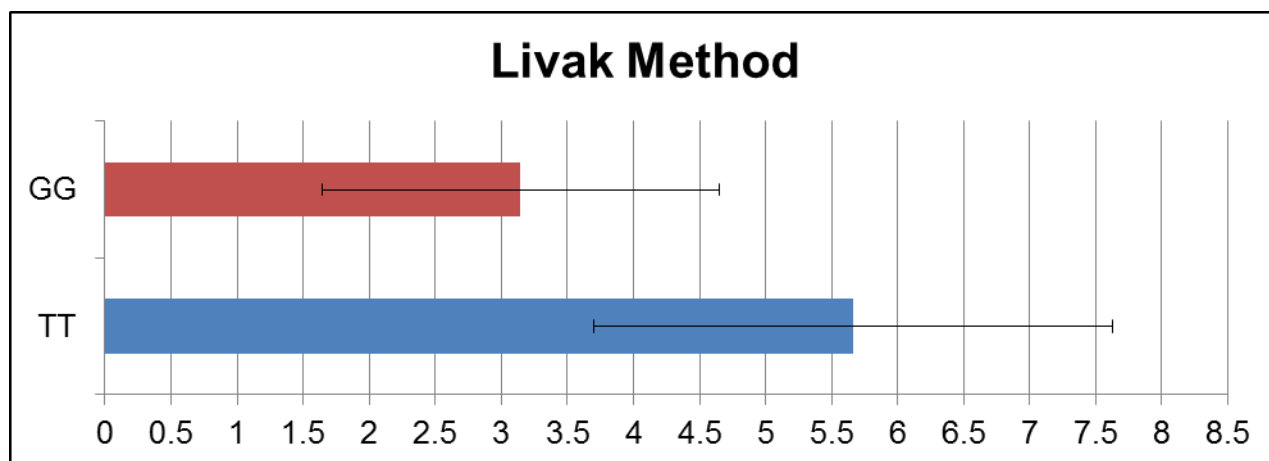

Figure 4.5 Graphic Representation of the change in *MC3R* gene expression using the Livak method: Values were determined for each qPCR run then the average value was taken for each genotype. Average values for each genotype are shown with standard deviation.

4.3.3.2 Pfaffl Method

Using the Pfaffl method (equation 2), the average change in *MC3R* gene expression was normalized to the average change in *α-tubulin* expression when stimulated with BCG. Efficiency values used in the calculation can be found in Table 4.5. The change in gene expression was calculated for each qPCR reaction. These results were then analyzed by Prism to determine the mean and the standard deviation for each genotype. Results for the analysis using the Pfaffl method can be seen in Figure 4.6. Results of the analysis of the change in *MC3R* gene expression upon stimulation with BCG as determined with the Pfaffl method were similar to those obtained by the Livak method. The smaller standard

deviation values obtained with the Pfaffl method may be a consequence of the Pfaffl method taking into account the differences in amplification efficiency between each assay.
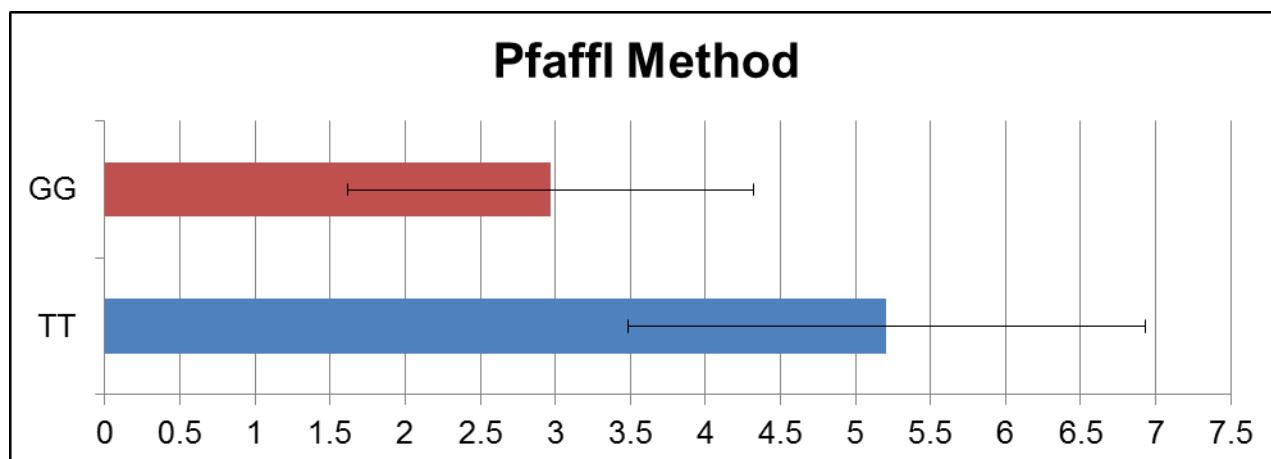

Figure 4.6 Graphic Representation of the change in *MC3R* gene expression using the Pfaffl method: Values were determined for each qPCR run then the average value was taken for each genotype. Average values for each genotype are shown with standard deviation.

4.4 *MC4R*

4.4.1 Single Point Analysis

Two of the SNPs in *MC4R*, rs11872992 and rs8087522, both located in the upstream region of the *MC4R* gene showed statistically significant differences in allele frequency distribution between TB patients and controls (Table 4.6). In both SNPs, the most common allele was found more frequently in cases than in controls but this association was not seen in the genotypic p-value.

4.4.2 Haplotype Analysis

Haplotype analysis of the 3 SNPs analyzed in *MC4R* identified two SNPs, rs8087522 and rs11872992, found to be in LD (Figure 4.7). While there appeared to be a statistically significant association between a single haplotype with TB disease, significance was lost when corrections were made for multiple testing.
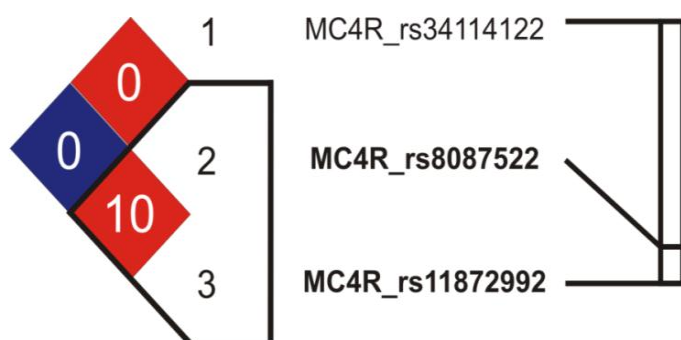

Figure 4.7 LD Plots of MC4R SNPs analyzed: LD plot generated by Haploview v4.1. r^2-values (%) are indicated on the squares. The colours of the squares represent LD strength according to D` values, with red being strong and blue being weak. Haplotype block is outlined in black.

Table 4.6 Single Point Analysis of *MC4R*

SNP	Group	N^a	1^{b}1	12	22	Genotypic p-value	OR^e (95% CI)	Allelic p-value	OR (95% CI)
MC4R									
rs11872992 G/A	TB	509	0.00	0.10	0.90	0.0689^c	1.438 (0.99 – 2.10)	0.0372	1.48 (1.04 – 2.12)
	Controls	487	0.01	0.14	0.85				
rs8087522 G/A	TB	498	0.18	0.46	0.36	0.0908^d	1.344 (1.02 – 1.76)	0.0393	1.21 (1.01 – 1.45)
	Controls	471	0.22	0.49	0.29				
rs34114122 C/A	TB	510	0.00	0.15	0.85	0.8205^c	0.9285 (0.66 – 1.31)	0.6757	0.92 (0.66 – 1.28)
	Controls	476	0.00	0.15	0.85				

[a] number of individuals investigated
[b] Allele 1 is the minor allele
[c] p-value based on a genotype-based Fisher exact test
[d] p-value based on a genotype-based χ^2 test
[e] odd ratio calculated by Prism

Table 4.7 Haplotype Analysis of *MC4R*

GENE	Haplotype		Cases	Controls	χ^2	P value	P value$_{permutation}$[b]	OR (95% CI)
	Block^a: rs8087522 – rs11872992							
MC4R	N1	C-T	0.587	0.541	4.30	0.0381	0.0850	1
	N2	C-C	0.360	0.381	0.93	0.3340	0.0526	1.03 (0.74 – 1.44)
	N3	T-C	0.053	0.078	5.17	0.0210	0.056	0.65 (0.45 – 0.94)

[a] The order of the SNPS in detected haplotype block corresponds to Figure 4.2
[b] Permutation test p values were calculated from 10 000 permutations in Haploview.

CHAPTER 5: Discussion

This study has validated the association between TB and both *CTSZ* and *MC3R* which was first identified in a genome-wide linkage study. A number of linkage analyses have been conducted in an attempt to identify novel loci involved in susceptibility to TB. In 2000, Bellamy *et al.* [228] found evidence that chromosomes 15q and Xq may have linkage to TB while Greenwood *et al.* [229] found significant linkage with chromosome region 2q35. A study of the Brazilian population implicated chromosomes 10, 11 and 20, [230] and in a study of a Moroccan population, Baghdadi *et al.* [230] showed significant linkage between TB and chromosome 8q12-q13.

A genome-wide linkage study performed in 2008 by Cooke *et al.* [82] identified a locus on chromosome 20q13.31-33 containing *CTSZ* and *MC3R,* which showed linkage with TB susceptibility. Later in 2008, Stein *et al.* [231] performed a genome linkage study in a large population from Uganda and replicated this finding (p = 0.002), identifying a 25 cM long region containing both *CTSZ* and *MC3R*.

An independent, unrelated case-control study has now been performed in an SAC population and the results of this study show that the same SNP implicated in *CTSZ* by Cooke *et al.* [82], rs34069356, was associated with TB disease. The OR of 0.28 found in this study falls at the lower limits of the CI found by Cooke *et al.* [82], with a smaller CI. This indicates the possibility of an even larger effect size than initially thought. This polymorphism results in a non-conservative amino acid change of a non-polar alanine to a polar, uncharged threonine.

CTSZ acts as a carboxypeptidase, similar to cathepsin B, but with specificity to a different profile of substrates and inhibitors. While mRNA expression of this gene is found in a large number of locations including lung, liver, kidney, pancreas and small intestines [232], it has been found that protein levels are highest in lung bronchoepithelial cells, alveolar and tingible body macrophages [178]. In all cell types investigated, *CTSZ* showed a higher distribution in lysosomal areas than in the cytoplasm [180]. Localization of this protein differs between immature and mature cells and Obermajer *et al.* [180] showed that in the presence of a *CTSZ* inhibitor, the migration of mature dendritic cells was significantly reduced. *CTSZ* also showed co-localization with β_2 integrins in differentiated U-937 cells, some Mo-T-lymphocytes and PBMCs, with the strongest signal occurring in cell-cell junctions [233]. Expression of the β_2 integrin family is characteristic of cells belonging to the immune system and it is known to play a central role in immunity due to the ability of β_2 integrins to mediate leukocyte adhesion and migration [234]. These findings seem to

suggest a role for *CTSZ* in the processes of inflammation and immune response as well as cell signaling and adhesion. Though it does not appear that the amino acid substitution investigated here occurs in an active site on *CTSZ*, the introduction of a hydroxyl side chain has many possible implications. Threonine has an uncharged, polar side chain, making the amino acid hydrophilic. Unlike alanine (a hydrophobic amino acid typically located on the interior of a protein), threonine is typically located on the exterior of a protein where the hydroxyl side chain is free to interact with surrounding water molecules. Threonine, but not alanine, is also subject to a number of PTMs including phosphorylation by threonine kinases, O-linked glycosylation and acetylation [24, 192, 235]. Phosphorylation is known to regulate the activity of proteins and since the amino acid substitution introduced by SNP rs34069356 occurs close to the N-terminal of the protein it is likely that this threonine is available for phosphorylation. The introduction of a hydroxyl group and a number of PTMs is likely to affect protein folding, intracellular localization and protein activity. While this case-control study is limited by the sample size, it does provide convincing evidence of a role for *CTSZ* in TB pathogenesis. Future studies should investigate the effect of the associated variant on PTMs of this protein and their effect on protein function. Understanding the effects of this variant could lead to a better understanding of this protein's role in the pathogenesis of TB.

A variant located 373bp upstream of the *MC3R* gene which resulted in a change from a G allele to a T allele (rs6127698) was significantly associated with TB, even after stringent Bonferroni correction for multiple testing. This SNP is predicted by Genomatix to create an alternative transcription factor binding site (http://www.genomatix.de/). The effects of the creation of an alternative transcription factor binding site are difficult to predict as a polymorphism in this region may result in either an increase or decrease in the transcription of the *MC3R* gene. The SNP in the single exon of *MC3R* identified by Cooke *et al.* [82] (rs3827103) was not significantly associated in the SAC population used in this study. SNP rs6127698 was not genotyped by Cooke *et al.* [82] and it can be noted that the r^2 value between rs3827103 and rs6127698 (Figure 4.2) is quite low, indicating that these two alleles are not predictive of each other in this SAC population. This might explain why Cooke *et al.* [82] did not find rs3827103 to be associated in their study.

Functional analysis of the associated SNP in *MC3R* revealed an increase in gene expression in individuals homozygous for either allele when MDMs were stimulated with BCG. Although no statistically significant difference was found, individuals of the TT genotype showed a trend toward increased *MC3R* transcription than individuals of the GG

genotype. The T allele was found to be overrepresented in controls compared to TB cases suggesting that this allele may contribute to a resistance phenotype. The larger increase in *MC3R* expression in individuals homozygous for the T allele suggests that upregulation of this gene may somehow aid the host immune system in more effectively containing the bacteria upon infection. Mice lacking a functional MC3R protein exhibit an increase in fat mass, a decrease in lean mass with no change in overall weight, and even a slight decrease in food intake [199]. This increase in feed efficiency is thought to result from an increase in the level of consumed energy being sent directly to fat storage. It could be reasoned that individuals with less energy available to the immune system might be more susceptible to infections. The increase in energy consumption which occurs upon infection results from the immune system initiating defense functions including increasing the activation and propagation of immune cells, DNA replication, RNA expression and protein synthesis and secretion and antigen processing and presentation [235]. It is therefore not surprising that undernourishment greatly increases human susceptibility to infectious disease, especially in low-income countries [27]. Based on the results of the functional analysis performed in this study, *MC3R* expression is increased upon stimulation with BCG. It is plausible that an increased level of *MC3R* transcript decreases the amount of energy being shunted to fat storage making more energy available for the host immune response. The limited sample size used in this study, however, makes it difficult to draw any definitive conclusions and this study should be repeated using a larger sample size. The sensitivity of the qPCR reaction can result in major variation from one run to the next. Slight differences in sample purity, sample concentration and assay efficiency can have a dramatic impact on cycle threshold. In addition to increasing the sample size, the standard deviation could be reduced by running each RNA sample in a serial dilution and by normalizing the results of the target gene against more than one reference gene. Due to the limited volume of RNA available, running the samples in a serial dilution was not an option for this pilot study. It would also be interesting to investigate the change in *MC3R* expression of heterozygous individuals to see if they exhibit an intermediate level of increased *MC3R* expression. Future studies should also investigate the functional role *MC3R* has in the pathogenesis of TB.

The case-control study investigating the association between TB and three SNPs in *MC4R* did not find any significant difference between cases and controls. Despite this, *MC4R* should not be ruled out as possibly playing a role in TB disease. *MC3R* and *MC4R* have been found to work together closely performing non-redundant functions. Chen *et al.* [199]

showed that the level of obesity observed in *MC4R* knockout mice is exacerbated when mice also lack *MC3R*. The synergistic effect of this double knockout may be a result of the increase in food intake caused by the lack of *MC4R* with even more of this increased energy consumption being sent directly to fat storage as a result of the increase in feed efficiency caused by the lack of a functional MC3R protein. The action of MCRs is coupled to adenylyl cyclase activation, and cAMP is able to down-regulate leptin mRNA expression in cultured rat adipocytes [27]. Previous studies dealing with the association between leptin and TB pathogenesis have been contradictory. In 2002, it was found that TB patients in Indonesia showed significantly suppressed plasma leptin concentrations [27] while in 2003, Schwenk *et al.* [23] could find no correlation with leptin concentrations and proinflammatory cytokine response in active TB, suggesting that this protein was not the missing link between immune defense and wasting associated with pulmonary TB [27]. These contradictory studies may be due to the fact that the production of leptin is regulated by two antagonistic mechanisms [27]. Wasting that is associated with active TB may result in a decrease in leptin production while activation of the inflammatory response may increase leptin levels [236]. Differences in diet and nutritional status between the populations used in each study may also contribute to conflicting results with respect to leptin`s involvement in both TB and TB-associated wasting.

The substantially validated role of *MC4R* in both obesity and diabetes further supports a possible role in TB. Higher BMI values have been shown to increase resistance to TB while decreased BMI values showed an increased susceptibility to both the development and the severity of TB. A study of HIV-infected individuals in Soweto found a decreasing risk of both TB morbidity and mortality with an increasing BMI, showing a dose-response effect [206]. These findings are also supported by a study of diabetic individuals conducted in 2007. Leung *et al.* [207] reported that diabetic individuals falling into either the overweight or obese BMI range had some protection against TB infection showing a strong linear dose-response relationship where for every unit change above a BMI of 18.5, there was a 10% reduction in the risk of developing active TB. A higher BMI value typically means a larger level of fat stores with a corresponding increase in circulating levels of leptin. The increased leptin may result in increased activation of the MC3R leading to a stronger host immune response to infection. These findings are contradicted by numerous studies which have shown that individuals diagnosed with DM exhibit an increased risk of TB infection compared to healthy controls. The above opposing findings are due to the complicated relationship between obesity and leptin. It remains debatable whether leptin is

a cause or effect of obesity but what is clear is that individuals suffering from obesity are chronically exposed to high levels of circulating leptin [237]. This chronic exposure results in a decrease in leptin receptor expression, reduced signaling and a decreased responsiveness to exogenous leptin [237].

The opposite of the problem of obesity is that of wasting, a common symptom of TB disease with numerous studies showing a decrease in BMI, skin fold thickness and arm muscle circumference in TB patients when compared to healthy controls [19]. One cause for this muscle wasting may be the result of a decrease in energy consumption as a direct result of a decrease in patient appetite, a function shown to be regulated by MC4R activation and suppression [20]. A study performed in the USA reported weight loss in 45% of TB patients, 26% of which exhibited signs of persistent anorexia [238]. Paton *et al.* [20] reported that patients exhibiting TB-associated wasting weighed an average of 12 kg less than healthy controls.

It would be informative to perform a case-control study of both *MC3R* and *MC4R* variants where TB patient samples are grouped into those exhibiting the wasting phenotype and those who do not. Investigating the effect of inactivating mutations in both *MC3R* and *MC4R* and their impact on TB disease may also provide valuable insights into the role of host metabolism on the development of active TB.

Observational studies concerning the influence of nutrition on the development of active TB go back as far as World War I where Denmark experienced an increase in TB morbidity and mortality as the country was exporting the majority of its meat, fish, poultry and dairy products [17]. TB rates did not decrease until the Danes experienced a surplus of these items following the German blockade of Denmark. As TB rates in Denmark decreased, those in surrounding countries continued to rise [17]. Chanarin *et al.* [236] found that in an Asiatic Indian population, life-long vegetarians who were deficient in cobalamin (vitamin B_{12} - a vitamin not found in the plant kingdom) exhibited a significantly higher level of TB susceptibility compared to omnivores. Studies investigating the influence of weight or nutrition on TB disease noted a significantly lower BMI in those infected with TB when compared to healthy controls [199].

It has been proposed that wasting is one of the determinants of disease severity and outcome and this symptom is a prominent feature of TB, particularly in resource-poor countries [27]. The mechanism by which poor nutrition leads to lowered cellular immunity remains unclear [239]. A number of studies have investigated the effects of nutritional

supplements on treatment of the disease but it is equally, if not more, important to study the effects of proper nutrition in the prevention of disease. A better understanding of the relationship between nutrition and disease susceptibility may provide valuable insights into the mechanisms involved in disease progression and suggest cost-effective interventions to counter this. The introduction of incentive programs which provide nutritious foods in exchange for treatment adherence or encouraging individuals to participate in TB screening programs may provide cost-effective ways to improve both treatment adherence and detection rates of TB incidence.

CHAPTER 6: Concluding Remarks

This study provides additional support for host genetics influencing susceptibility to the development of TB. The results obtained here have validated previous findings implicating *CTSZ* and *MC3R* in TB disease. The variant in *CTSZ* previously associated with TB susceptibility in three West African populations was also found to be associated with TB in an SAC population. These findings provide convincing evidence for *CSTZ* playing a role in the host immune response to TB infection and motivate future research into the function of this gene in the immune response to TB infection and pathogenesis. Determining the functional effect of the amino acid substitution caused by the associated variant could provide valuable information both on this gene's role in immunity and how TB is so effective at evading elimination by the host immune response. Genotypic and functional analyses of *MC3R* provided evidence that energy expenditure and host metabolism may influence an individual's susceptibility to TB. These results suggest that upregulation of this gene may reduce the risk of developing TB by influencing the amount of energy available to the host immune system. Due to the fact that *MC3R* is found to work closely while not redundantly with *MC4R* it is possible that proper nutrition may also play an important role in preventing TB. Results of this study encourage further investigations to be made into determining the impact of nutritional status both on the prevention and progression of TB. Understanding the impact of proper nutrition prior to infection may provide a novel strategy for TB prevention programs and may provide a cost-effect incentive to improve treatment adherence. It is possible that a better understanding of the impact nutritional status has on TB prevention and prevalence may allow for revisions of current public health programs leading to a more successful outcome of intervention.

Genome-wide linkage studies have not always been overly successful at identifying genes impacting on complex diseases, but this study illustrates that genes identified this way, and not previously suspected to be involved in TB disease, can prove to be convincing candidates.

Reference List

1. World Health Organization. WHO report 2009: Global Tuberculosis Control 2009 - epidemiology, strategy, financing. (2009)

2. Dye C, Williams BG. The population dynamics and control of tuberculosis. Science 2010;328:856-861.

3. Bellamy R, Beyers N, McAdam KP, Ruwende C, Gie R, Samaai P, Bester D, Meyer M, Corrah T, Collin M, Camidge DR, Wilkinson D, Hoal-Van Helden E, Whittle HC, Amos W, van Helden P, Hill AV. Genetic susceptibility to tuberculosis in Africans: a genome-wide scan. Proc Natl Acad Sci U S A 2000;97:8005-8009.

4. Pan H, Yan BS, Rojas M, Shebzukhov YV, Zhou H, Kobzik L, Higgins DE, Daly MJ, Bloom BR, Kramnik I. Ipr1 gene mediates innate immunity to tuberculosis. Nature 2005;434:767-772.

5. Russell DG, Barry CE, III, Flynn JL. Tuberculosis: what we don't know can, and does, hurt us. Science 2010;328:852-856.

6. Geiter, L. (2000). *Ending Neglect: The Elimination of Tuberculosis in the United States*. Washington, D.C.: National Academies press.

7. Dubos, R. & Dubos, J. (1952). *The White Plague: Tuberculosis, Man and Society*. Boston: Little, Brown & Co.

8. Smith I. Mycobacterium tuberculosis pathogenesis and molecular determinants of virulence. Clin Microbiol Rev 2003;16:463-496.

9. Ducati RG, Ruffino-Netto A, Basso LA, Santos DS. The resumption of consumption - a review on tuberculosis. Mem Inst Oswaldo Cruz 2006;101:697-714.

10. Kaufmann SH. How can immunology contribute to the control of tuberculosis? Nat Rev Immunol 2001;1:20-30.

11. Meena LS, Rajni. Survival mechanisms of pathogenic Mycobacterium tuberculosis H37Rv. FEBS J 2010;277:2416-2427.

12. Chaisson RE, Harrington M. How research can help control tuberculosis. Int J Tuberc Lung Dis 2009;13:558-568.

13. Pieters J. Mycobacterium tuberculosis and the macrophage: maintaining a balance. Cell Host Microbe 2008;3:399-407.

14. Russell DG, Cardona PJ, Kim MJ, Allain S, Altare F. Foamy macrophages and the progression of the human tuberculosis granuloma. Nat Immunol 2009;10:943-948.

15. Cosma CL, Sherman DR, Ramakrishnan L. The secret lives of the pathogenic mycobacteria. Annu Rev Microbiol 2003;57:641-676.

16. Kramnik I. Genetic dissection of host resistance to Mycobacterium tuberculosis: the sst1 locus and the Ipr1 gene. Curr Top Microbiol Immunol 2008;321:123-148.

17. Cegielski JP, McMurray DN. The relationship between malnutrition and tuberculosis: evidence from studies in humans and experimental animals. Int J Tuberc Lung Dis 2004;8:286-298.

18. Leyton G. Effects of Slow Starvation. (1946) pp. 73-79
Ref ID:

19. Miller LG, Asch SM, Yu EI, Knowles L, Gelberg L, Davidson P. A population-based survey of tuberculosis symptoms: how atypical are atypical presentations? Clin Infect Dis 2000;30:293-299.

20. Paton NI, Ng YM. Body composition studies in patients with wasting associated with tuberculosis. Nutrition 2006;22:245-251.

21. Hanrahan CF, Golub JE, Mohapi L, Tshabangu N, Modisenyane T, Chaisson RE, Gray GE, McIntyre JA, Martinson NA. Body mass index and risk of tuberculosis and death. AIDS 2010;24:1501-1508.

22. Sarraf P, Frederich RC, Turner EM, Ma G, Jaskowiak NT, Rivet DJ, III, Flier JS, Lowell BB, Fraker DL, Alexander HR. Multiple cytokines and acute inflammation raise mouse leptin levels: potential role in inflammatory anorexia. J Exp Med 1997;185:171-175.

23. Schwenk A, Hodgson L, Rayner CF, Griffin GE, Macallan DC. Leptin and energy metabolism in pulmonary tuberculosis. Am J Clin Nutr 2003;77:392-398.

24. Zhang Y, Proenca R, Maffei M, Barone M, Leopold L, Friedman JM. Positional cloning of the mouse obese gene and its human homologue. Nature 1994;372:425-432.

25. Ahima RS, Prabakaran D, Mantzoros C, Qu D, Lowell B, Maratos-Flier E, Flier JS. Role of leptin in the neuroendocrine response to fasting. Nature 1996;382:250-252.

26. Boden G, Chen X, Mozzoli M, Ryan I. Effect of fasting on serum leptin in normal human subjects. J Clin Endocrinol Metab 1996;81:3419-3423.

27. van CR, Karyadi E, Netea MG, Verhoef H, Nelwan RH, West CE, van der Meer JW. Decreased plasma leptin concentrations in tuberculosis patients are associated with wasting and inflammation. J Clin Endocrinol Metab 2002;87:758-763.

28. Ehlers S. Lazy, dynamic or minimally recrudescent? On the elusive nature and location of the mycobacterium responsible for latent tuberculosis. Infection 2009;37:87-95.

29. Manabe YC, Bishai WR. Latent Mycobacterium tuberculosis-persistence, patience, and winning by waiting. Nat Med 2000;6:1327-1329.

30. Frothingham R, Stout JE, Hamilton CD. Current issues in global tuberculosis control. Int J Infect Dis 2005;9:297-311.

31. Maartens G, Wilkinson RJ. Tuberculosis. Lancet 2007;370:2030-2043.

32. World Health Organization. WHO Report 2008: Global tuberculosis control - surveillance, planning, financing. (2008)

33. Kunitz F, Brodhun B, Hauer B, Haas W, Loddenkemper R. [The current status of tuberculosis in Germany and impact of the global situation]. Pneumologie 2007;61:467-477.

34. Banerjee R, Allen J, Westenhouse J, Oh P, Elms W, Desmond E, Nitta A, Royce S, Flood J. Extensively drug-resistant tuberculosis in california, 1993-2006. Clin Infect Dis 2008;47:450-457.

35. Cambanis A, Yassin MA, Ramsay A, Bertel SS, Arbide I, Cuevas LE. Rural poverty and delayed presentation to tuberculosis services in Ethiopia. Trop Med Int Health 2005;10:330-335.

36. Farmer P, Robin S, Ramilus SL, Kim JY. Tuberculosis, poverty, and "compliance": lessons from rural Haiti. Semin Respir Infect 1991;6:254-260.

37. Belo MT, Selig L, Luiz RR, Hanson C, Luna AL, Teixeira EG, Trajman A. Choosing incentives to stimulate tuberculosis treatment compliance in a poor county in Rio de Janeiro state, Brazil. Med Sci Monit 2006;12:H1-H5.

38. United Nations. The Millennium Development Goals Report. (2010)

39. Atinmo T, Mirmiran P, Oyewole OE, Belahsen R, Serra-Majem L. Breaking the poverty/malnutrition cycle in Africa and the Middle East. Nutr Rev 2009;67 Suppl 1:S40-S46.

40. Macallan DC. Malnutrition in tuberculosis. Diagn Microbiol Infect Dis 1999;34:153-157.

41. Kallings LO. The first postmodern pandemic: 25 years of HIV/ AIDS. J Intern Med 2008;263:218-243.

42. UNAIDS, WHO. AIDS epidemic update. (2009)

43. Rankin WW, Brennan S, Schell E, Laviwa J, Rankin SH. The stigma of being HIV-positive in Africa. PLoS Med 2005;2:e247.

44. Abdool Karim SS, Churchyard GJ, Abdool KQ, Lawn SD. HIV infection and tuberculosis in South Africa: an urgent need to escalate the public health response. Lancet 2009;374:921-933.

45. Wells CD, Cegielski JP, Nelson LJ, Laserson KF, Holtz TH, Finlay A, Castro KG, Weyer K. HIV infection and multidrug-resistant tuberculosis: the perfect storm. J Infect Dis 2007;196 Suppl 1:S86-107.

46. Gandhi NR, Moll A, Sturm AW, Pawinski R, Govender T, Lalloo U, Zeller K, Andrews J, Friedland G. Extensively drug-resistant tuberculosis as a cause of death in patients co-infected with tuberculosis and HIV in a rural area of South Africa. Lancet 2006;368:1575-1580.

47. Perkins MD. New diagnostic tools for tuberculosis. Int J Tuberc Lung Dis 2000;4:S182-S188.

48. Pai M, Riley LW, Colford JM, Jr. Interferon-gamma assays in the immunodiagnosis of tuberculosis: a systematic review. Lancet Infect Dis 2004;4:761-776.

49. Reid MJ, Shah NS. Approaches to tuberculosis screening and diagnosis in people with HIV in resource-limited settings. Lancet Infect Dis 2009;9:173-184.

50. Jeong YJ, Lee KS. Pulmonary tuberculosis: up-to-date imaging and management. AJR Am J Roentgenol 2008;191:834-844.

51. Cruciani M, Scarparo C, Malena M, Bosco O, Serpelloni G, Mengoli C. Meta-analysis of BACTEC MGIT 960 and BACTEC 460 TB, with or without solid media, for detection of mycobacteria. J Clin Microbiol 2004;42:2321-2325.

52. Lange C, Mori T. Advances in the diagnosis of tuberculosis. Respirology 2010;15:220-240.

53. Pai M, Kalantri S, Dheda K. New tools and emerging technologies for the diagnosis of tuberculosis: part I. Latent tuberculosis. Expert Rev Mol Diagn 2006;6:413-422.

54. Davidow A, Kanaujia GV, Shi L, Kaviar J, Guo X, Sung N, Kaplan G, Menzies D, Gennaro ML. Antibody profiles characteristic of Mycobacterium tuberculosis infection state. Infect Immun 2005;73:6846-6851.

55. Sacchettini JC, Rubin EJ, Freundlich JS. Drugs versus bugs: in pursuit of the persistent predator Mycobacterium tuberculosis. Nat Rev Microbiol 2008;6:41-52.

56. Frydenberg AR, Graham SM. Toxicity of first-line drugs for treatment of tuberculosis in children: review. Trop Med Int Health 2009;14:1329-1337.

57. Behr MA, Small PM. Has BCG attenuated to impotence? Nature 1997;389:133-134.

58. Ryan JJ, Orme SK, Wexler H, Ketcham AS. Successful long-term coumadin anticoagulation of mice with prothrombin time monitoring. J Surg Oncol 1969;1:115-121.

59. Frieden TR. Lessons from tuberculosis control for public health. Int J Tuberc Lung Dis 2009;13:421-428.

60. Stead WW, Senner JW, Reddick WT, Lofgren JP. Racial differences in susceptibility to infection by Mycobacterium tuberculosis. N Engl J Med 1990;322:422-427.

61. Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, Gordon SV, Eiglmeier K, Gas S, Barry CE, III, Tekaia F, Badcock K, Basham D, Brown D, Chillingworth T, Connor R, Davies R, Devlin K, Feltwell T, Gentles S, Hamlin N, Holroyd S, Hornsby T, Jagels K, Krogh A, McLean J, Moule S, Murphy L, Oliver K, Osborne J, Quail MA, Rajandream MA, Rogers J, Rutter S, Seeger K, Skelton J, Squares R, Squares S, Sulston JE, Taylor K, Whitehead S, Barrell BG. Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. Nature 1998;393:537-544.

62. Steyn AJ, Chan J, Mehra V. Recent developments in mycobacterial research. Curr Opin Infect Dis 1999;12:415-424.

63. Mollenkopf HJ, Grode L, Mattow J, Stein M, Mann P, Knapp B, Ulmer J, Kaufmann SH. Application of mycobacterial proteomics to vaccine design: improved protection by Mycobacterium bovis BCG prime-Rv3407 DNA boost vaccination against tuberculosis. Infect Immun 2004;72:6471-6479.

64. Barker LF, Brennan MJ, Rosenstein PK, Sadoff JC. Tuberculosis vaccine research: the impact of immunology. Curr Opin Immunol 2009;21:331-338.

65. Stead WW. Genetics and resistance to tuberculosis. Could resistance be enhanced by genetic engineering? Ann Intern Med 1992;116:937-941.

66. Motulsky AG. Metabolic polymorphisms and the role of infectious diseases in human evolution. Human Biology 1960;32:28-62.

67. Bellamy R. Susceptibility to mycobacterial infections: the importance of host genetics. Genes Immun 2003;4:4-11.

68. Secko D. A gene for susceptibility to tuberculosis. CMAJ 2005;172:1436.

69. Clementi M, Di Gianantonio E. Genetic susceptibility to infectious diseases. Reprod Toxicol 2006;21:345-349.

70. Kraft P, Cox DG. Study designs for genome-wide association studies. Adv Genet 2008;60:465-504.

71. Kathiresan S, Newton-Cheh C, Gerszten RE. On the interpretation of genetic association studies. Eur Heart J 2004;25:1378-1381.

72. Collins A. Allelic association: linkage disequilibrium structure and gene mapping. Mol Biotechnol 2009;41:83-89.

73. Zondervan KT, Cardon LR. Designing candidate gene and genome-wide case-control association studies. Nat Protoc 2007;2:2492-2501.

74. Altshuler D, Daly MJ, Lander ES. Genetic mapping in human disease. Science 2008;322:881-888.

75. Weeks DE, Lange K. The affected-pedigree-member method of linkage analysis. Am J Hum Genet 1988;42:315-326.

76. Kennedy S, Bennett S, Weeks DE. Affected sib-pair analysis in endometriosis. Hum Reprod Update 2001;7:411-418.

77. Davies JL, Kawaguchi Y, Bennett ST, Copeman JB, Cordell HJ, Pritchard LE, Reed PW, Gough SC, Jenkins SC, Palmer SM, . A genome-wide search for human type 1 diabetes susceptibility genes. Nature 1994;371:130-136.

78. Satsangi J, Parkes M, Louis E, Hashimoto L, Kato N, Welsh K, Terwilliger JD, Lathrop GM, Bell JI, Jewell DP. Two stage genome-wide search in inflammatory bowel disease provides evidence for susceptibility loci on chromosomes 3, 7 and 12. Nat Genet 1996;14:199-202.

79. Möller M, de Wit E, Hoal EG. Past, present and future directions in human genetic susceptibility to tuberculosis. FEMS Immunol Med Microbiol 2010;58:3-26.

80. Jorgensen TJ, Ruczinski I, Kessing B, Smith MW, Shugart YY, Alberg AJ. Hypothesis-driven candidate gene association studies: practical design and analytical considerations. Am J Epidemiol 2009;170:986-993.

81. Möller M, Hoal EG. Current findings, challenges and novel approaches in human genetic susceptibility to tuberculosis. Tuberculosis (Edinb) 2010.

82. Cooke GS, Campbell SJ, Bennett S, Lienhardt C, McAdam KP, Sirugo G, Sow O, Gustafson P, Mwangulu F, van HP, Fine P, Hoal EG, Hill AV. Mapping of a novel susceptibility locus suggests a role for MC3R and CTSZ in human tuberculosis. Am J Respir Crit Care Med 2008;178:203-207.

83. Selvaraj P, Uma H, Reetha AM, Kurian SM, Xavier T, Prabhakar R, Narayanan PR. HLA antigen profile in pulmonary tuberculosis patients & their spouses. Indian J Med Res 1998;107:155-158.

84. Singh SP, Mehra NK, Dingley HB, Pande JN, Vaidya MC. Human leukocyte antigen (HLA)-linked control of susceptibility to pulmonary tuberculosis and association with HLA-DR types. J Infect Dis 1983;148:676-681.

85. Sanjeevi CB, Narayanan PR, Prabakar R, Charles N, Thomas BE, Balasubramaniam R, Olerup O. No association or linkage with HLA-DR or -DQ genes in south Indians with pulmonary tuberculosis. Tuber Lung Dis 1992;73:280-284.

86. Teran-Escandon D, Teran-Ortiz L, Camarena-Olvera A, Gonzalez-Avila G, Vaca-Marin MA, Granados J, Selman M. 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;115:428-433.

87. Hawkins BR, Higgins DA, Chan SL, Lowrie DB, MITCHISON DA, 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;138:1616-1621.

88. Amirzargar AA, Yalda A, Hajabolbaghi M, Khosravi F, Jabbari H, Rezaei N, Niknam MH, Ansari B, Moradi B, Nikbin B. The association of HLA-DRB, DQA1, DQB1

alleles and haplotype frequency in Iranian patients with pulmonary tuberculosis. Int J Tuberc Lung Dis 2004;8:1017-1021.

89. Vejbaesya S, Chierakul N, Luangtrakool K, Srinak D, Stephens HA. Associations of HLA class II alleles with pulmonary tuberculosis in Thais. Eur J Immunogenet 2002;29:431-434.

90. Vejbaesya S, Chierakul N, Luangtrakool P, Sermduangprateep C. NRAMP1 and TNF-alpha polymorphisms and susceptibility to tuberculosis in Thais. Respirology 2007;12:202-206.

91. Dubaniewicz A, Moszkowska G, Szczerkowska Z, Hoppe A. Analysis of DQB1 allele frequencies in pulmonary tuberculosis: preliminary report. Thorax 2003;58:890-891.

92. Lombard Z, Dalton DL, Venter PA, Williams RC, 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;67:643-654.

93. Delgado JC, Baena A, Thim S, Goldfeld AE. Ethnic-specific genetic associations with pulmonary tuberculosis. J Infect Dis 2002;186:1463-1468.

94. Goldfeld AE, Delgado JC, Thim S, Bozon MV, Uglialoro AM, Turbay D, Cohen C, Yunis EJ. Association of an HLA-DQ allele with clinical tuberculosis. JAMA 1998;279:226-228.

95. Bothamley GH, Beck JS, Schreuder GM, D'Amaro J, de Vries RR, Kardjito T, Ivanyi J. Association of tuberculosis and M. tuberculosis-specific antibody levels with HLA. J Infect Dis 1989;159:549-555.

96. Khomenko AG, Litvinov VI, Chukanova VP, Pospelov LE. Tuberculosis in patients with various HLA phenotypes. Tubercle 1990;71:187-192.

97. Brahmajothi V, Pitchappan RM, Kakkanaiah VN, Sashidhar M, Rajaram K, Ramu S, Palanimurugan K, Paramasivan CN, Prabhakar R. Association of pulmonary tuberculosis and HLA in south India. Tubercle 1991;72:123-132.

98. Ravikumar M, Dheenadhayalan V, Rajaram K, Lakshmi SS, Kumaran PP, Paramasivan CN, Balakrishnan K, Pitchappan RM. Associations of HLA-DRB1, DQB1 and DPB1 alleles with pulmonary tuberculosis in south India. Tuber Lung Dis 1999;79:309-317.

99. Rajalingam R, Singal DP, Mehra NK. Transporter associated with antigen-processing (TAP) genes and susceptibility to tuberculoid leprosy and pulmonary tuberculosis. Tissue Antigens 1997;49:168-172.

100. Pospelov LE, Matrakshin AG, Chernousova LN, Tsoi KN, Afanasjev KI, Rubtsova GA, Yeremeyev VV. Association of various genetic markers with tuberculosis and other lung diseases in Tuvinian children. Tuber Lung Dis 1996;77:77-80.

101. Ruggiero G, Cosentini E, Zanzi D, Sanna V, Terrazzano G, Matarese G, Sanduzzi A, Perna F, Zappacosta S. Allelic distribution of human leucocyte antigen in historical and recently diagnosed tuberculosis patients in Southern Italy. Immunology 2004;111:318-322.

102. Kim HS, Park MH, Song EY, Park H, Kwon SY, Han SK, Shim YS. 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.

103. Wang J, Song C, Wang S. [Association of HLA-DRB1 genes with pulmonary tuberculosis]. Zhonghua Jie He He Hu Xi Za Zhi 2001;24:302-305.

104. Figueiredo JF, Rodrigues ML, Deghaide NH, Donadi EA. HLA profile in patients with AIDS and tuberculosis. Braz J Infect Dis 2008;12:278-280.

105. Thye T, Nejentsev S, Intemann CD, Browne EN, Chinbuah MA, Gyapong J, Osei I, Owusu-Dabo E, Zeitels LR, Herb F, Horstmann RD, Meyer CG. MCP-1 promoter variant -362C associated with protection from pulmonary tuberculosis in Ghana, West Africa. Hum Mol Genet 2009;18:381-388.

106. Flores-Villanueva PO, Ruiz-Morales JA, Song CH, Flores LM, Jo EK, Montano M, Barnes PF, Selman M, Granados J. A functional promoter polymorphism in monocyte chemoattractant protein-1 is associated with increased susceptibility to pulmonary tuberculosis. J Exp Med 2005;202:1649-1658.

107. Möller M, Nebel A, Valentonyte R, van Helden PD, Schreiber S, Hoal EG. Investigation of chromosome 17 candidate genes in susceptibility to TB in a South African population. Tuberculosis (Edinb) 2009;89:189-194.

108. Barreiro LB, Neyrolles O, Babb CL, Tailleux L, Quach H, McElreavey K, Helden PD, Hoal EG, Gicquel B, Quintana-Murci L. Promoter Variation in the DC-SIGN Encoding Gene CD209 Is Associated with Tuberculosis. PLoS Medicine 2006;3:e20.

109. Gomez LM, Anaya JM, Sierra-Filardi E, Cadena J, Corbi A, Martin J. Analysis of DC-SIGN (CD209) functional variants in patients with tuberculosis. Hum Immunol 2006;67:808-811.

110. Vannberg FO, Chapman SJ, Khor CC, Tosh K, Floyd S, Jackson-Sillah D, Crampin A, Sichali L, Bah B, Gustafson P, Aaby P, McAdam KP, Bah-Sow O, Lienhardt C, Sirugo G, Fine P, Hill AV. CD209 genetic polymorphism and tuberculosis disease. PLoS ONE 2008;3:e1388.

111. Ben-Ali M, Barreiro LB, Chabbou A, Haltiti R, Braham E, Neyrolles O, Dellagi K, Gicquel B, Quintana-Murci L, Barbouche MR. Promoter and neck region length variation of DC-SIGN is not associated with susceptibility to tuberculosis in Tunisian patients. Hum Immunol 2007;68:908-912.

112. Selvaraj P, Alagarasu K, Swaminathan S, Harishankar M, Narendran G. CD209 gene polymorphisms in South Indian HIV and HIV-TB patients. Infect Genet Evol 2009;9:256-262.

113. Lio D, Marino V, Serauto A, Gioia V, Scola L, Crivello A, Forte GI, Colonna-Romano G, Candore G, Caruso C. 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;29:371-374.

114. Cooke GS, Campbell SJ, Sillah J, Gustafson P, Bah B, Sirugo G, Bennett S, McAdam KP, Sow O, Lienhardt C, Hill AV. Polymorphism within the interferon-gamma/receptor complex is associated with pulmonary tuberculosis. Am J Respir Crit Care Med 2006;174:339-343.

115. Etokebe GE, Bulat-Kardum L, Johansen MS, Knezevic J, Balen S, Matakovic-Mileusnic N, Matanic D, Flego V, Pavelic J, Beg-Zec Z, Dembic Z. Interferon-gamma gene (T874A and G2109A) polymorphisms are associated with microscopy-positive tuberculosis. Scand J Immunol 2006;63:136-141.

116. Vidyarani M, Selvaraj P, Prabhu AS, Jawahar MS, Adhilakshmi AR, Narayanan PR. Interferon gamma (IFNgamma) & interleukin-4 (IL-4) gene variants & cytokine levels in pulmonary tuberculosis. Indian J Med Res 2006;124:403-410.

117. Lopez-Maderuelo D, Arnalich F, Serantes R, Gonzalez A, Codoceo R, Madero R, Vazquez JJ, Montiel C. Interferon-gamma and interleukin-10 gene polymorphisms in pulmonary tuberculosis. Am J Respir Crit Care Med 2003;167:970-975.

118. Moran A, Ma X, Reich RA, 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;11:113-115.

119. Ding S, Li L, Zhu X. Polymorphism of the interferon-gamma gene and risk of tuberculosis in a southeastern Chinese population. Hum Immunol 2008;69:129-133.

120. Rossouw M, Nel HJ, Cooke GS, van Helden PD, Hoal EG. Association between tuberculosis and a polymorphic NFkappaB binding site in the interferon gamma gene. Lancet 2003;361:1871-1872.

121. Sallakci N, Coskun M, Berber Z, Gurkan F, Kocamaz H, Uysal G, Bhuju S, Yavuzer U, Singh M, Yegin O. Interferon-γ gene+874T-A polymorphism is associated with tuberculosis and gamma interferon response. (2007)

122. Kusuhara K, Yamamoto K, Okada K, Mizuno Y, 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.

123. Tso HW, Ip WK, Chong WP, Tam CM, Chiang AK, Lau YL. Association of interferon gamma and interleukin 10 genes with tuberculosis in Hong Kong Chinese. Genes Immun 2005;6:358-363.

124. Sahiratmadja E, Baak-Pablo R, de Visser AW, Alisjahbana B, Adnan I, van Crevel R, Marzuki S, van Dissel JT, Ottenhoff TH, Van D, V. Association of polymorphisms in IL-12/IFN-gamma pathway genes with susceptibility to pulmonary tuberculosis in Indonesia. Tuberculosis (Edinb) 2007;87:303-311.

125. Henao MI, Montes C, Paris SC, Garcia LF. Cytokine gene polymorphisms in Colombian patients with different clinical presentations of tuberculosis. Tuberculosis (Edinburgh, Scotland) 2006;86:11-19.

126. Garcia-Laorden MI, Pena MJ, Caminero JA, Garcia-Saavedra A, Campos-Herrero MI, Caballero A, Rodriguez-Gallego C. Influence of mannose-binding lectin on HIV infection and tuberculosis in a Western-European population. Mol Immunol 2006;43:2143-2150.

127. Soborg C, Andersen AB, Range N, Malenganisho W, Friis H, Magnussen P, Temu MM, Changalucha J, Madsen HO, Garred P. Influence of candidate susceptibility genes on tuberculosis in a high endemic region. Mol Immunol 2007;44:2213-2220.

128. Hoal-van Helden EG, Epstein J, Victor TC, Hon D, Lewis LA, Beyers N, Zurakowski D, Ezekowitz AB, van Helden PD. Mannose-binding protein B allele confers protection against tuberculous meningitis. Pediatr Res 1999;45:459-464.

129. Druszczynska M, Strapagiel D, Kwiatkowska S, Kowalewicz-Kulbat M, Rozalska B, Chmiela M, Rudnicka W. Tuberculosis bacilli still posing a threat. Polymorphism of genes regulating anti-mycobacterial properties of macrophages. Pol J Microbiol 2006;55:7-12.

130. El Sahly HM, Reich RA, Dou SJ, Musser JM, Graviss EA. The effect of mannose binding lectin gene polymorphisms on susceptibility to tuberculosis in different ethnic groups. Scand J Infect Dis 2004;36:106-108.

131. Cosar H, Ozkinay F, Onay H, Bayram N, Bakiler AR, Anil M, Can D, Ozkinay C. Low levels of mannose-binding lectin confers protection against tuberculosis in Turkish children. Eur J Clin Microbiol Infect Dis 2008;27:1165-1169.

132. Capparelli R, Iannaccone M, Palumbo D, Medaglia C, Moscariello E, Russo A, Iannelli D. Role played by human mannose-binding lectin polymorphisms in pulmonary tuberculosis. The Journal of infectious diseases 2009;199:666-672.

133. Liu W, Zhang F, Xin ZT, Zhao QM, Wu XM, Zhang PH, de VS, Richardus JH, Habbema JD, Yang H, Cao WC. Sequence variations in the MBL gene and their relationship to pulmonary tuberculosis in the Chinese Han population. Int J Tuberc Lung Dis 2006;10:1098-1103.

134. Selvaraj P, Narayanan PR, Reetha AM. Association of functional mutant homozygotes of the mannose binding protein gene with susceptibility to pulmonary tuberculosis in India. Tuber Lung Dis 1999;79:221-227.

135. Bellamy R, Ruwende C, McAdam KP, Thursz M, Sumiya M, Summerfield J, Gilbert SC, Corrah T, Kwiatkowski D, Whittle HC, Hill AV. Mannose binding protein deficiency is not associated with malaria, hepatitis B carriage nor tuberculosis in Africans. QJM 1998;91:13-18.

136. Soborg C, Madsen HO, Andersen AB, Lillebaek T, Kok-Jensen A, Garred P. Mannose-binding lectin polymorphisms in clinical tuberculosis. J Infect Dis 2003;188:777-782.

137. Gomez LM, Anaya JM, Vilchez JR, Cadena J, Hinojosa R, Velez L, Lopez-Nevot MA, Martin J. A polymorphism in the inducible nitric oxide synthase gene is associated with tuberculosis. Tuberculosis (Edinb) 2007;87:288-294.

138. Velez DR, Hulme WF, Myers JL, Weinberg JB, Levesque MC, Stryjewski ME, Abbate E, Estevan R, Patillo SG, Gilbert JR, Hamilton CD, Scott WK. NOS2A, TLR4, and IFNGR1 interactions influence pulmonary tuberculosis susceptibility in African-Americans. Hum Genet 2009;126:643-653.

139. Jamieson SE, Miller EN, Black GF, Peacock CS, Cordell HJ, Howson JM, Shaw MA, Burgner D, Xu W, Lins-Lainson Z, Shaw JJ, Ramos F, Silveira F, Blackwell JM. Evidence for a cluster of genes on chromosome 17q11-q21 controlling susceptibility to tuberculosis and leprosy in Brazilians. Genes Immun 2004;5:46-57.

140. Malik S, Abel L, Tooker H, Poon A, Simkin L, Girard M, Adams GJ, Starke JR, Smith KC, Graviss EA, Musser JM, Schurr E. Alleles of the NRAMP1 gene are risk factors for pediatric tuberculosis disease. Proc Natl Acad Sci U S A 2005;102:12183-12188.

141. Zhang W, Shao L, Weng X, Hu Z, Jin A, Chen S, Pang M, Chen ZW. Variants of the natural resistance-associated macrophage protein 1 gene (NRAMP1) are associated with severe forms of pulmonary tuberculosis. Clin Infect Dis 2005;40:1232-1236.

142. Qu YB, Tang YX, Zhang ZB, Zhu R, Liu J, Gu SY, Lu GL, Xia ZL. [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;24:531-533.

143. Jin J, Sun L, Jiao W, Zhao S, Li H, Guan X, Jiao A, Jiang Z, Shen A. SLC11A1 (Formerly NRAMP1) gene polymorphisms associated with pediatric tuberculosis in China. Clin Infect Dis 2009;48:733-738.

144. Ryu S, Park YK, Bai GH, Kim SJ, Park SN, Kang S. 3'UTR polymorphisms in the NRAMP1 gene are associated with susceptibility to tuberculosis in Koreans. Int J Tuberc Lung Dis 2000;4:577-580.

145. El Baghdadi J, Remus N, Benslimane A, El Annaz H, Chentoufi M, Abel L, Schurr E. Variants of the human NRAMP1 gene and susceptibility to tuberculosis in Morocco. Int J Tuberc Lung Dis 2003;7:599-602.

146. Soborg C, Andersen AB, Madsen HO, Kok-Jensen A, Skinhoj P, Garred P. Natural resistance-associated macrophage protein 1 polymorphisms are associated with microscopy-positive tuberculosis. J Infect Dis 2002;186:517-521.

147. Bellamy R, Ruwende C, Corrah T, McAdam KP, Whittle HC, Hill AV. Variations in the NRAMP1 gene and susceptibility to tuberculosis in West Africans. N Engl J Med 1998;338:640-644.

148. Puzyrev VP, Freidin MB, Rudko AA, Strelis AK, 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;36:788-791.

149. Liaw YS, Tsai-Wu JJ, Wu CH, Hung CC, Lee CN, Yang PC, Luh KT, Kuo SH. Variations in the NRAMP1 gene and susceptibility of tuberculosis in Taiwanese. Int J Tuberc Lung Dis 2002;6:454-460.

150. Liu W, Zhang CY, Tian L, Li CZ, Wu XM, Zhao QM, Zhang PH, Yang SM, Yang H, Cao WC. [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;37:408-411.

151. Duan HF, Zhou XH, Ma Y, Li CY, Chen XY, Gao WW, Zheng SH. [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;26:286-289.

152. Abe T, Iinuma Y, Ando M, Yokoyama T, Yamamoto T, Nakashima K, Takagi N, Baba H, Hasegawa Y, Shimokata K. NRAMP1 polymorphisms, susceptibility and clinical features of tuberculosis. J Infect 2003;46:215-220.

153. Taype CA, Castro JC, Accinelli RA, Herrera-Velit P, Shaw MA, Espinoza JR. Association between SLC11A1 polymorphisms and susceptibility to different clinical forms of tuberculosis in the Peruvian population. Infect Genet Evol 2006;6:361-367.

154. Hoal EG, Lewis L-A, Jamieson SE, Tanzer F, Rossouw M, Victor T, Hillerman R, Beyers N, Blackwell JM, van Helden PD. *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;8:1464-1471.

155. Dubaniewicz A, Jamieson SE, Dubaniewicz-Wybieralska M, Fakiola M, Nancy ME, Blackwell JM. Association between SLC11A1 (formerly NRAMP1) and the risk of sarcoidosis in Poland. Eur J Hum Genet 2005;13:829-834.

156. Awomoyi AA, Marchant A, Howson JM, McAdam KP, Blackwell JM, Newport MJ. Interleukin-10, polymorphism in SLC11A1 (formerly NRAMP1), and susceptibility to tuberculosis. J Infect Dis 2002;186:1808-1814.

157. Ma X, Dou S, Wright JA, Reich RA, Teeter LD, El Sahly HM, Awe RJ, Musser JM, Graviss EA. 5' dinucleotide repeat polymorphism of NRAMP1 and susceptibility to tuberculosis among Caucasian patients in Houston, Texas. Int J Tuberc Lung Dis 2002;6:818-823.

158. Gao PS, Fujishima S, Mao XQ, Remus N, Kanda M, Enomoto T, Dake Y, Bottini N, Tabuchi M, Hasegawa N, Yamaguchi K, Tiemessen C, Hopkin JM, Shirakawa T, Kishi F. Genetic variants of NRAMP1 and active tuberculosis in Japanese populations. International Tuberculosis Genetics Team. Clin Genet 2000;58:74-76.

159. Cervino AC, Lakiss S, Sow O, Hill AV. Allelic association between the NRAMP1 gene and susceptibility to tuberculosis in Guinea-Conakry. Ann Hum Genet 2000;64:507-512.

160. Tosh K, Campbell SJ, Fielding K, Sillah J, Bah B, Gustafson P, Manneh K, Lisse I, Sirugo G, Bennett S, Aaby P, McAdam KP, Bah-Sow O, Lienhardt C, Kramnik I, Hill AV. Variants in the SP110 gene are associated with genetic susceptibility to tuberculosis in West Africa. Proc Natl Acad Sci U S A 2006;103:10364-10368.

161. Babb C, Keet EH, van Helden PD, Hoal EG. SP110 polymorphisms are not associated with pulmonary tuberculosis in a South African population. Hum Genet 2007;121:521-522.

162. Szeszko JS, Healy B, Stevens H, Balabanova Y, Drobniewski F, Todd JA, Nejentsev S. Resequencing and association analysis of the SP110 gene in adult pulmonary tuberculosis. Hum Genet 2006;121:155-160.

163. Thye T, Browne EN, Chinbuah MA, Gyapong J, Osei I, Owusu-Dabo E, Niemann S, Rusch-Gerdes S, Horstmann RD, Meyer CG. No associations of human pulmonary tuberculosis with Sp110 variants. J Med Genet 2006;43:e32.

164. Ma X, Liu Y, Gowen BB, Graviss EA, Clark AG, Musser JM. Full-exon resequencing reveals toll-like receptor variants contribute to human susceptibility to tuberculosis disease. PLoS ONE 2007;2:e1318.

165. Yim JJ, Lee HW, Lee HS, Kim YW, Han SK, Shim YS, Holland SM. The association between microsatellite polymorphisms in intron II of the human Toll-like receptor 2 gene and tuberculosis among Koreans. Genes Immun 2006;7:150-155.

166. Olesen R, Wejse C, Velez DR, Bisseye C, Sodemann M, Aaby P, Rabna P, Worwui A, Chapman H, Diatta M, Adegbola RA, Hill PC, Ostergaard L, Williams SM, Sirugo G. DC-SIGN (CD209), pentraxin 3 and vitamin D receptor gene variants associate with pulmonary tuberculosis risk in West Africans. Genes Immun 2007;8:456-467.

167. Ogus AC, Yoldas B, Ozdemir T, Uguz A, Olcen S, Keser I, Coskun M, Cilli A, Yegin O. The Arg753GLn polymorphism of the human toll-like receptor 2 gene in tuberculosis disease. Eur Respir J 2004;23:219-223.

168. Newport MJ, Allen A, Awomoyi AA, Dunstan SJ, McKinney E, Marchant A, Sirugo G. The toll-like receptor 4 Asp299Gly variant: no influence on LPS responsiveness or susceptibility to pulmonary tuberculosis in The Gambia. Tuberculosis (Edinb) 2004;84:347-352.

169. Davila S, Hibberd ML, Hari DR, Wong HE, Sahiratmadja E, Bonnard C, Alisjahbana B, Szeszko JS, Balabanova Y, Drobniewski F, van CR, Van D, V, Nejentsev S, Ottenhoff TH, Seielstad M. Genetic association and expression studies indicate a role of toll-like receptor 8 in pulmonary tuberculosis. PLoS Genet 2008;4:e1000218.

170. Selvaraj P, Kurian SM, Chandra G, Reetha AM, Charles N, Narayanan PR. Vitamin D receptor gene variants of BsmI, ApaI, TaqI, and FokI polymorphisms in spinal tuberculosis. Clin Genet 2004;65:73-76.

171. Wilkinson RJ, Llewelyn M, Toossi Z, Patel P, Pasvol G, Lalvani A, Wright D, Latif M, Davidson RN. Influence of vitamin D deficiency and vitamin D receptor polymorphisms on tuberculosis among Gujarati Asians in west London: a case-control study. Lancet 2000;355:618-621.

172. Liu W, Cao WC, Zhang CY, Tian L, Wu XM, Habbema JD, Zhao QM, Zhang PH, Xin ZT, Li CZ, Yang H. 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;8:428-434.

173. Roth DE, Soto G, Arenas F, Bautista CT, Ortiz J, Rodriguez R, Cabrera L, Gilman RH. Association between Vitamin D Receptor Gene Polymorphisms and Response to Treatment of Pulmonary Tuberculosis. J Infect Dis 2004;190:920-927.

174. Bornman L, Campbell SJ, Fielding K, Bah B, Sillah J, Gustafson P, Manneh K, Lisse I, Allen A, Sirugo G, Sylla A, Aaby P, McAdam KP, Bah-Sow O, Bennett S, Lienhardt C, Hill AV. Vitamin D receptor polymorphisms and susceptibility to tuberculosis in West Africa: a case-control and family study. J Infect Dis 2004;190:1631-1641.

175. Bellamy R, Ruwende C, Corrah T, McAdam KP, Thursz M, Whittle HC, Hill AV. Tuberculosis and chronic hepatitis B virus infection in Africans and variation in the vitamin D receptor gene. J Infect Dis 1999;179:721-724.

176. Santamaria I, Velasco G, Pendas AM, Fueyo A, Lopez-Otin C. Cathepsin Z, a novel human cysteine proteinase with a short propeptide domain and a unique chromosomal location. Journal of Biological Chemistry 1998;273:16816-16823.

177. Kos J, Jevnikar Z, Obermajer N. The role of cathepsin X in cell signaling. Cell Adh Migr 2009;3:164-166.

178. Kos J, Sekirnik A, Premzl A, Zavasnik B, V, Langerholc T, Turk B, Werle B, Golouh R, Repnik U, Jeras M, Turk V. Carboxypeptidases cathepsins X and B display distinct protein profile in human cells and tissues. Exp Cell Res 2005;306:103-113.

179. Obermajer N, Premzl A, Zavasnik BT, Turk B, Kos J. Carboxypeptidase cathepsin X mediates beta2-integrin-dependent adhesion of differentiated U-937 cells. Exp Cell Res 2006;312:2515-2527.

180. Obermajer N, Repnik U, Jevnikar Z, Turk B, Kreft M, Kos J. Cysteine protease cathepsin X modulates immune response via activation of beta2 integrins. Immunology 2008;124:76-88.

181. Freemont AJ. Demystified ... adhesion molecules. Mol Pathol 1998;51:175-184.

182. Saunders BM, Cooper AM. Restraining mycobacteria: role of granulomas in mycobacterial infections. Immunol Cell Biol 2000;78:334-341.

183. Evans R, Patzak I, Svensson L, De FK, Jones K, McDowall A, Hogg N. Integrins in immunity. J Cell Sci 2009;122:215-225.

184. Wang SX, Fan ZC, Tao YX. Functions of acidic transmembrane residues in human melanocortin-3 receptor binding and activation. Biochem Pharmacol 2008;76:520-530.

185. Lembertas AV, Perusse L, Chagnon YC, Fisler JS, Warden CH, Purcell-Huynh DA, Dionne FT, Gagnon J, Nadeau A, Lusis AJ, Bouchard C. Identification of an obesity quantitative trait locus on mouse chromosome 2 and evidence of linkage to body fat and insulin on the human homologous region 20q. J Clin Invest 1997;100:1240-1247.

186. Yassin RJ, Hamblin AS. Altered expression of CD11/CD18 on the peripheral blood phagocytes of patients with tuberculosis. Clin Exp Immunol 1994;97:120-125.

187. Lee EJ, Lee SH, Jung JW, Lee W, Kim BJ, Park KW, Lim SK, Yoon CJ, Baik JH. Differential regulation of cAMP-mediated gene transcription and ligand selectivity by MC3R and MC4R melanocortin receptors. Eur J Biochem 2001;268:582-591.

188. Cooray SN, Clark AJ. Melanocortin receptors and their accessory proteins. Mol Cell Endocrinol 2010.

189. Versteeg DH, Van BP, Adan RA, De Wildt DJ. Melanocortins and cardiovascular regulation. Eur J Pharmacol 1998;360:1-14.

190. Getting SJ, Riffo-Vasquez Y, Pitchford S, Kaneva M, Grieco P, Page CP, Perretti M, Spina D. A role for MC3R in modulating lung inflammation. Pulm Pharmacol Ther 2008;21:866-873.

191. Feng N, Young SF, Aguilera G, Puricelli E, dler-Wailes DC, Sebring NG, Yanovski JA. Co-occurrence of two partially inactivating polymorphisms of MC3R is associated with pediatric-onset obesity. Diabetes 2005;54:2663-2667.

192. Lee YS, Poh LK, Loke KY. A novel melanocortin 3 receptor gene (MC3R) mutation associated with severe obesity. J Clin Endocrinol Metab 2002;87:1423-1426.

193. Mencarelli M, Walker GE, Maestrini S, Alberti L, Verti B, Brunani A, Petroni ML, Tagliaferri M, Liuzzi A, Di Blasio AM. Sporadic mutations in melanocortin receptor 3 in morbid obese individuals. Eur J Hum Genet 2008;16:581-586.

194. Rached M, Buronfosse A, Begeot M, Penhoat A. Inactivation and intracellular retention of the human I183N mutated melanocortin 3 receptor associated with obesity. Biochim Biophys Acta 2004;1689:229-234.

195. Tao YX, Segaloff DL. Functional characterization of melanocortin-3 receptor variants identify a loss-of-function mutation involving an amino acid critical for G protein-coupled receptor activation. J Clin Endocrinol Metab 2004;89:3936-3942.

196. Tao YX. Functional characterization of novel melanocortin-3 receptor mutations identified from obese subjects. Biochim Biophys Acta 2007;1772:1167-1174.

197. Lee YS, Poh LK, Kek BL, Loke KY. The role of melanocortin 3 receptor gene in childhood obesity. Diabetes 2007;56:2622-2630.

198. Calton MA, Ersoy BA, Zhang S, Kane JP, Malloy MJ, Pullinger CR, Bromberg Y, Pennacchio LA, Dent R, McPherson R, Ahituv N, Vaisse C. Association of functionally significant Melanocortin-4 but not Melanocortin-3 receptor mutations with severe adult obesity in a large North American case-control study. Hum Mol Genet 2009;18:1140-1147.

199. Chen AS, Marsh DJ, Trumbauer ME, Frazier EG, Guan XM, Yu H, Rosenblum CI, Vongs A, Feng Y, Cao L, Metzger JM, Strack AM, Camacho RE, Mellin TN, Nunes CN, Min W, Fisher J, Gopal-Truter S, MacIntyre DE, Chen HY, Van der Ploeg LH. Inactivation of the mouse melanocortin-3 receptor results in increased fat mass and reduced lean body mass. Nat Genet 2000;26:97-102.

200. Cole SA, Butte NF, Voruganti VS, Cai G, Haack K, Kent JW, Jr., Blangero J, Comuzzie AG, McPherson JD, Gibbs RA. Evidence that multiple genetic variants of MC4R play a functional role in the regulation of energy expenditure and appetite in Hispanic children. Am J Clin Nutr 2010;91:191-199.

201. Liu G, Zhu H, Lagou V, Gutin B, Barbeau P, Treiber FA, Dong Y, Snieder H. Common variants near melanocortin 4 receptor are associated with general and visceral adiposity in European- and African-American youth. J Pediatr 2010;156:598-605.

202. Savastano DM, Tanofsky-Kraff M, Han JC, Ning C, Sorg RA, Roza CA, Wolkoff LE, Anandalingam K, Jefferson-George KS, Figueroa RE, Sanford EL, Brady S, Kozlosky M, Schoeller DA, Yanovski JA. Energy intake and energy expenditure among children with polymorphisms of the melanocortin-3 receptor. Am J Clin Nutr 2009;90:912-920.

203. Butler AA, Kesterson RA, Khong K, Cullen MJ, Pelleymounter MA, Dekoning J, Baetscher M, Cone RD. A unique metabolic syndrome causes obesity in the melanocortin-3 receptor-deficient mouse. Endocrinology 2000;141:3518-3521.

204. Yurtcu E, Yilmaz A, Ozkurt Z, Kolukisa E, Yilmaz M, Keles H, Ergun MA, Yetkin I, Menevse A. Melanocortin-4 receptor gene polymorphisms in obese patients. Biochem Genet 2009;47:295-300.

205. Li WD, Joo EJ, Furlong EB, Galvin M, Abel K, Bell CJ, Price RA. Melanocortin 3 receptor (MC3R) gene variants in extremely obese women. Int J Obes Relat Metab Disord 2000;24:206-210.

206. Cummings DE, Schwartz MW. Melanocortins and body weight: a tale of two receptors. Nat Genet 2000;26:8-9.

207. Leung CC, Lam TH, Chan WM, Yew WW, Ho KS, Leung G, Law WS, Tam CM, Chan CK, Chang KC. Lower risk of tuberculosis in obesity. Arch Intern Med 2007;167:1297-1304.

208. Gantz I, Konda Y, Tashiro T, Shimoto Y, Miwa H, Munzert G, Watson SJ, DelValle J, Yamada T. Molecular cloning of a novel melanocortin receptor. J Biol Chem 1993;268:8246-8250.

209. Loos RJ, Lindgren CM, Li S, Wheeler E, Zhao JH, Prokopenko I, Inouye M, Freathy RM, Attwood AP, Beckmann JS, Berndt SI, Jacobs KB, Chanock SJ, Hayes RB, Bergmann S, Bennett AJ, Bingham SA, Bochud M, Brown M, Cauchi S, Connell JM, Cooper C, Smith GD, Day I, Dina C, De S, Dermitzakis ET, Doney AS, Elliott KS, Elliott P, Evans DM, Sadaf F, I, Froguel P, Ghori J, Groves CJ, Gwilliam R, Hadley D, Hall AS, Hattersley AT, Hebebrand J, Heid IM, Lamina C, Gieger C, Illig T, Meitinger T, Wichmann HE, Herrera B, Hinney A, Hunt SE, Jarvelin MR, Johnson T, Jolley JD, Karpe F, Keniry A, Khaw KT, Luben RN, Mangino M, Marchini J, McArdle WL, McGinnis R, Meyre D, Munroe PB, Morris AD, Ness AR, Neville MJ, Nica AC, Ong KK, O'Rahilly S, Owen KR, Palmer CN, Papadakis K, Potter S, Pouta A, Qi L, Randall JC, Rayner NW, Ring SM, Sandhu MS, Scherag A, Sims MA, Song K, Soranzo N, Speliotes EK, Syddall HE, Teichmann SA, Timpson NJ, Tobias JH, Uda M, Vogel CI, Wallace C, Waterworth DM, Weedon MN, Willer CJ, Wraight, Yuan X, Zeggini E, Hirschhorn JN, Strachan DP, Ouwehand WH, Caulfield MJ, Samani NJ, Frayling TM, Vollenweider P, Waeber G, Mooser V, Deloukas P, McCarthy MI, Wareham NJ, Barroso I, Jacobs KB, Chanock SJ, Hayes RB, Lamina C, Gieger C, Illig T, Meitinger T, Wichmann HE, Kraft P, Hankinson SE, Hunter DJ, Hu FB, Lyon HN, Voight BF, Ridderstrale M, Groop L, Scheet P, Sanna S, Abecasis GR, Albai G,

Nagaraja R, Schlessinger D, Jackson AU, Tuomilehto J, Collins FS, Boehnke M, Mohlke KL. Common variants near MC4R are associated with fat mass, weight and risk of obesity. Nat Genet 2008;40:768-775.

210. Breit A, Buch TR, Boekhoff I, Solinski HJ, Damm E, Gudermann T. Alternative G protein-coupling and biased agonism: new insights into melanocortin-4 receptor signalling. Mol Cell Endocrinol 2010.

211. Demiralp DO, Berberoglu M, Akar N. Melanocortin-4 Receptor Polymorphisms in Turkish Pediatric Obese Patients. Clin Appl Thromb Hemost 2010.

212. Geller F, Reichwald K, Dempfle A, Illig T, Vollmert C, Herpertz S, Siffert W, Platzer M, Hess C, Gudermann T, Biebermann H, Wichmann HE, Schafer H, Hinney A, Hebebrand J. Melanocortin-4 receptor gene variant I103 is negatively associated with obesity. Am J Hum Genet 2004;74:572-581.

213. Lubrano-Berthelier C, Cavazos M, Le SC, Haas K, Shapiro A, Zhang S, Bougneres P, Vaisse C. The human MC4R promoter: characterization and role in obesity. Diabetes 2003;52:2996-3000.

214. Huszar D, Lynch CA, Fairchild-Huntress V, Dunmore JH, Fang Q, Berkemeier LR, Gu W, Kesterson RA, Boston BA, Cone RD, Smith FJ, Campfield LA, Burn P, Lee F. Targeted disruption of the melanocortin-4 receptor results in obesity in mice. Cell 1997;88:131-141.

215. Lu D, Willard D, Patel IR, Kadwell S, Overton L, Kost T, Luther M, Chen W, Woychik RP, Wilkison WO, . Agouti protein is an antagonist of the melanocyte-stimulating-hormone receptor. Nature 1994;371:799-802.

216. Harries AD, Billo N, Kapur A. Links between diabetes mellitus and tuberculosis: should we integrate screening and care? Trans R Soc Trop Med Hyg 2009;103:1-2.

217. Jeon CY, Murray MB. Diabetes mellitus increases the risk of active tuberculosis: a systematic review of 13 observational studies. PLoS Med 2008;5:e152.

218. Qi L, Kraft P, Hunter DJ, Hu FB. The common obesity variant near MC4R gene is associated with higher intakes of total energy and dietary fat, weight change and diabetes risk in women. Hum Mol Genet 2008;17:3502-3508.

219. de Wit E., Delport W, Rugamika CE, Meintjes A, Moller M, van Helden PD, Seoighe C, Hoal EG. Genome-wide analysis of the structure of the South African Coloured Population in the Western Cape. Hum Genet 2010;128:145-153.

220. GE Healthcare. illustra GenomiPhi V2 DNA Amplification Kit. (2006)

221. Holland PM, Abramson RD, Watson R, 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;88:7276-7280.

222. Livak KJ. Allelic discrimination using fluorogenic probes and the 5' nuclease assay. Genet Anal 1999;14:143-149.

223. Ye S, Dhillon S, Ke X, Collins AR, Day IN. An efficient procedure for genotyping single nucleotide polymorphisms. Nucleic Acids Res 2001;29:E88.

224. Piccioli P, Serra M, Pedemonte S, Balbi G, Loiacono F, Lastraioli S, Gargiulo L, Morabito A, Zuccaro D, Del ML, Pistillo MP, Venturini M, De AM, Notaro R. Hexaprimer amplification refractory mutation system PCR for simultaneous single-tube genotyping of 2 close polymorphisms. Clin Chem 2008;54:227-229.

225. Heid CA, Stevens J, Livak KJ, Williams PM. Real time quantitative PCR. Genome Res 1996;6:986-994.

226. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res 2001;29:e45.

227. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 2001;25:402-408.

228. Bellamy R, Hill AV. Genetic susceptibility to mycobacteria and other infectious pathogens in humans. Curr Opin Immunol 1998;10:483-487.

229. Greenwood CM, Fujiwara TM, Boothroyd LJ, Miller MA, Frappier D, Fanning EA, Schurr E, Morgan K. Linkage of tuberculosis to chromosome 2q35 loci, including NRAMP1, in a large aboriginal Canadian family. Am J Hum Genet 2000;67:405-416.

230. Baghdadi JE, Orlova M, Alter A, Ranque B, Chentoufi M, Lazrak F, Archane MI, Casanova JL, Benslimane A, Schurr E, Abel L. An autosomal dominant major gene confers predisposition to pulmonary tuberculosis in adults. J Exp Med 2006;203:1679-1684.

231. Stein CM, Zalwango S, Malone LL, Won S, Mayanja-Kizza H, Mugerwa RD, Leontiev DV, Thompson CL, Cartier KC, Elston RC, Iyengar SK, Boom WH, Whalen CC. Genome scan of M. tuberculosis infection and disease in Ugandans. PLoS ONE 2008;3:e4094.

232. Menard R, Nagler DK, Zhang R, Tam W, Sulea T, Purisima EO. Human cathepsin X. A novel cysteine protease with unique specificity. Adv Exp Med Biol 2000;477:317-322.

233. Imhof BA, urrand-Lions M. Adhesion mechanisms regulating the migration of monocytes. Nat Rev Immunol 2004;4:432-444.

234. Cooper, G. M. (2000). *The Cell: A Molecular Approach*, Second Edition edn.

235. Schaible UE, Kaufmann SH. Malnutrition and infection: complex mechanisms and global impacts. PLoS Med 2007;4:e115.

236. Chanarin I, Stephenson E. Vegetarian diet and cobalamin deficiency: their association with tuberculosis. J Clin Pathol 1988;41:759-762.

237. Scarpace PJ, Zhang Y. Elevated leptin: consequence or cause of obesity? Front Biosci 2007;12:3531-3544.

238. Greene WC. A history of AIDS: looking back to see ahead. Eur J Immunol 2007;37 Suppl 1:S94-102.

239. Dooley KE, Chaisson RE. Tuberculosis and diabetes mellitus: convergence of two epidemics. Lancet Infect Dis 2009;9:737-746.

Appendix

Publications

Poster: (presented at 2008 South African Society of Human Genetics)

Polymorphisms in *MC3R* promoter and *CTSZ* 3`UTR are associated with tuberculosis susceptibility

Lindsey A Adams [a], Marlo Möller [a], Almut Nebel [b], Stefan Schreiber [b], Lize van der Merwe [c,d], Paul D van Helden [a], Eileen G Hoal [a]

[a] DST/NRF Centre of Excellence for Biomedical Tuberculosis Research / MRC Centre of Molecular and Cellular Biology, Faculty of Health Sciences, Stellenbosch University, South Africa.

[b] Institute of Clinical Molecular Biology, Christian-Albrechts-University, Schittenhelmstraβe 12, 24105 Kiel, Germany

[c] Biostatistics Unit, Medical Research Council, Tygerberg, South Africa

[d] Department of Statistics, University of Western Cape, South Africa

Corresponding author: EG Hoal, DST/NRF Centre of Excellence for Biomedical Tuberculosis Research / MRC Centre of Molecular and Cellular Biology, Division of Molecular Biology and Human Genetics, Faculty of Health Sciences, University of Stellenbosch, Tygerberg, PO Box 19063, 7505, South Africa. Tel: +27 21 938 9412, Fax: +27 21 938 9863, egvh@sun.ac.za

Running Head: MC3R, CTSZ and TB susceptibility

Abstract: 201 words

Text: 3548 words

1

Abstract:

We have validated the association of two genes on chromosome 20q13.31-33 with tuberculosis susceptibility. A previous genome-wide linkage study performed by Cooke *et al* identified the genes melanocortin-3-receptor (*MC3R*) and cathepsin Z (*CTSZ*) as possible candidates in tuberculosis susceptibility. MC3R has been implicated in obesity studies and is known to play a role in many biological systems including the regulation of energy homeostasis and fat metabolism. CTSZ has been detected in immune cells, such as macrophages and monocytes and it is hypothesized that the protein may play a role in the immune response. In our South African population, a case-control study confirmed the previously reported association with a SNP in *CTSZ* and found an association in *MC3R* with a SNP not previously implicated in TB susceptibility. Six single nucleotide polymorphisms (SNPs) in *MC3R* and eight in *CTSZ* were genotyped and haplotypes were inferred. SNP rs6127698 in the promoter region of *MC3R* (cases = 498, controls = 506) and rs34069356 in the 3`UTR of *CTSZ* (cases = 396, controls = 298) both showed significant association with tuberculosis susceptibility (p = 0.0004 and <0.0001, respectively) indicating that pathways involving these proteins, not previously researched in this disease, could yield novel therapies for TB.

2

Introduction

Genome-wide investigations which are free of initial assumptions and are not directly hypothesis-driven can identify novel genes involved in complex diseases such as tuberculosis.[1] Most of the genes currently identified by hypothesis-driven studies belong to the innate immune system, the first line of defense against pathogens, but methods based on our current knowledge of disease may overlook genes for which a role has yet to be determined. Amongst the genes that have been implicated in tuberculosis susceptibility, many have well defined roles in the immune response including HLA class II, *NRAMP1* and *IFNG*.[2] Genome-wide linkage studies are particularly useful because they can identify genes for which a role in infection may not have been suspected. Susceptibility loci for leprosy, a disease related to tuberculosis, have been found on chromosomes 10p13 and 20p12.3 in a south Indian population while a Vietnamese study identified chromosome 6q25 as being linked to leprosy susceptibility, which led to the identification of the *PARK 2* and *PACRG* genes. [3-6] The first genome-wide linkage study for tuberculosis was conducted in 2000 by Bellamy *et al.*, and identified chromosomes 15q and Xq as showing suggestive evidence of linkage to tuberculosis susceptibility.[7]

Recently, a multistage strategy was employed by Cooke *et al* to identify a novel locus for tuberculosis susceptibility in African populations.[8] An affected sibling pair linkage analysis performed on families recruited from the South African Coloured population in metropolitan Cape Town, and Malawians from the Karonga district indicated one genomic region, 20q13.31-33, as being linked to tuberculosis susceptibility. Forty SNPs within this region were used to screen a large independent Gambian population, and two genes, melanocortin 3 receptor (*MC3R*) and cathepsin Z (*CTSZ*), showed evidence of disease association.[8] Polymorphisms in these genes were further genotyped in populations from Guinea-Bissau and the Republic of Conakry.[8] A

3

polymorphism in the 3'UTR in *CTSZ*, viz. CTSZ3P (rs34069356), showed strong disease association (P=0.005), with genotype distributions being similar across all three West African populations.[8] Following regression analysis, the initial association seen in the *MC3R* polymorphism MC3R241 (rs3827103) did not remain significant (p = 0.26), though a trend toward a protective effect of the *MC3R* genotype AA remained.[8]

CTSZ is one of the 11 cysteine proteases of the papain family.[9] In the immune system, cathepsins are involved in antigen processing and maturation of the major histocompatibility complex class II molecules.[10] CTSZ is mostly expressed in immune cells, such as macrophages and monocytes, and a role for the protein in the immune response has been hypothesized.[11-13] Cathepsins form a vital component of the lysosomal proteolytic system and are differentially expressed during *Mycobacterium tuberculosis* infection.[14] This expression is specifically associated with macrophages present in the granuloma.[14] Several members of the cathepsins have been implicated in TB.[15] For example, cathepsin L maturation and activity can be impaired by *M. tuberculosis* and *M. avium*, and cathepsin W was identified as a risk factor for the extrapulmonary dissemination of human TB.[15]

MC3R belongs to a family of 7-transmembrane G-protein coupled receptors which transmit their signals through the activation of adenylate cyclase.[16] This receptor is abundantly expressed in brain regions and in a variety of peripheral tissues and has been shown to play a role in many biological systems including the regulation of energy homeostasis and fat metabolism, as well as inflammation.[17, 18] Polymorphisms in *MC3R* have been associated with severe obesity and defects in this gene can result in decreased total expression, intracellular retention and defective receptor activation.[19-25] Chen *et al* showed that inactivating mutations of *MC3R* led to an increase

4

in fat mass with a corresponding decrease in body mass without any change in eating behavior or metabolic rate.[26]

The initial stage of the original linkage study was conducted by Cooke *et al* using Malawian and South African Coloured sibling pairs to identify the genomic region. In the second stage, fine-mapping was carried out in various West African populations. To validate these findings, we conducted case-control studies for *MC3R* and *CTSZ* in unrelated South African Coloured individuals to determine if polymorphisms in these genes show evidence of disease association.

Methods

Study Population

The population is located in the metropolitan area of Cape Town in the Western Cape Province in South Africa. This area was selected due to the high incidence of tuberculosis in the area as well as the uniform ethnicity, socio-economic status and low prevalence of HIV.[27] We did a population-based case-control association study using unrelated individuals from the South African Coloured population (Table 1). Tuberculosis patients were identified through bacteriological confirmation (smear positive and/or culture positive). Controls were selected from the same community living under the same conditions including socio-economic status and availability of health facilities. Our previous study of healthy children and young adults from the control community found that 80% of children older than 15 years had positive tuberculin skin tests (TST), an indication of latent infection with *M.tuberculosis*.[28] The majority of the control population is therefore TST positive, and with the average age of the controls in this study being 27 years (Table 1), we estimate a TST positivity of ~80% or above. These healthy individuals

5

had no previous history of tuberculosis disease or treatment and were unrelated to all others included in the study. Additional sample characteristics are given in Table 1. There was no overlap between the samples used in this study and the previous linkage study done by Cooke *et al.*[8] Approval from the Ethics Committee of the Faculty of Health Sciences, Stellenbosch University (project number 95/072) was obtained before blood samples were collected with informed consent, and known human immunodeficiency virus (HIV) positive individuals were excluded from the study. DNA was purified using standard extraction protocols.

Despite having received genetic input from Khoisan, Bantu-speaking, European and Asian antecedents, the South African Coloured currently represents a relatively homogenous population.[27] A previous study genotyped 351 cases and 360 controls for a panel of 25 independent SNPs markers which were not in linkage disequilibrium (LD), were randomly distributed along the genome and polymorphic among the major contributing ethnic groups.[27] This study showed no significant population stratification.[27] Of the 1186 samples genotyped in our study, 505 overlapped with the previous study that investigated population stratification.[27] We can however not explicitly exclude the possibility of stratification in the South African Coloured population.

Genotyping

Given the size and lack of introns of the *MC3R* gene it was possible to sequence the full gene to gain maximum information about the polymorphisms in and surrounding it. The single exon of the gene as well as 1000 base pairs upstream was sequenced in 10 controls and 10 TB patients. In total, six SNPs were detected, including the polymorphism reported by Cooke *et al*, rs3827103 (Table 2) and all were selected for further analysis. Two of the polymorphisms were

6

located within the single exon of the gene and 4 were located up to 413 base pairs upstream of the start codon. All six SNPs were within 654 base pairs of one another and this fragment was directly sequenced following polymerase chain reaction (PCR) amplification. In order to sequence the *MC3R* polymorphisms, PCR was used to amplify a 995bp fragment. PCR reactions were carried out in a total volume of 25 μl. Each reaction contained 100ng of genomic DNA, 2.5 μl of 10X Reaction Buffer containing 15mM $MgCl_2$ (JMR Holdings, UK), 1μl of 2.5mM dNTPs (Bioline), 0.5μl of 10μM forward and reverse primers (5`-AGAATCTCAGGGCCAGGTA-3` and 5-`GTCCTCGAAGGTCAGGTAGTC-3`, respectively) (Integrated DNA Technologies, UK) and 0.05μl of Super-Therm Gold DNA polymerase (JMR Holdings, UK). An Eppendorf Mastercycler PCR System was used for the following cycling program: 10 minutes of denaturation at 95°C followed by 30 cycles of 1 minute at 94°C, 1 minute at 65°C and 1 minute at 72°C. This reaction was ended by incubation at 72°C for 10 minutes then a hold at 4°C. Control for contamination was done by the inclusion of master mix blanks in every batch of samples amplified. Amplicons were sequenced in both forward and reverse direction.

CTSZ SNPs that showed an allele frequency above 5% and were located either in the exons or the 3`UTR region of the gene were selected for genotyping from previous publications [8] or online databases such as dbSNP. Eight polymorphisms were selected in total, four of which were located in the 3'UTR region (including rs34069356, the associated SNP reported by Cooke *et al*) and four in the coding region of the gene. Seven of the *CTSZ* SNPs were genotyped using the SNPlex Genotyping System[TM] (Applied Biosystems) (Table 1) on an automated platform and data was managed by laboratory information system (LIMS) as described previously.[29] The SNPs were submitted online at the myScience Environment of Applied Biosystems website (http://myscience.appliedbiosystems.com) for assay design. The assay was done according to the

7

manufacturer's instructions, CEPH controls were included, and assay output files were evaluated with the GeneMapper Analysis Software v3.5.1 (Applied Biosystems Darmstadt, Germany). Alleles were called automatically. The results were verified by inspection of the cluster plots. 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. SNP rs34069356 was genotyped using a predesigned Custom Taqman SNP Genotyping Assay kit (Applied Biosystems) and fluorescent data was read using the ABI 9700.

Statistical Analysis

With the samples successfully genotyped in *MC3R* (504 cases and 516 controls) and *CTSZ* (481 cases and 376 controls) and an expected allele frequency of at least 5% for SNPs genotyped, this study had 95% confidence and 97% and 94% power, respectively, to detect an odds ratio of 2.5. Of the 1186 samples genotyped, 290 TB cases and 246 controls were genotyped for all the *MC3R* and *CTSZ* SNPs. Hardy-Weinberg equilibrium (HWE) was assessed for all SNPs in the TB and control groups. Linkage disequilibrium (LD) patterns based on D' values were summarized with LD heatmaps.

Logistic regression was used to compare the TB and the control group to facilitate adjustment for confounders. As the age and gender both differed significantly between TB cases and controls, all analyses were adjusted for age and gender by including them in the logistic regression models as covariates. We modeled each genotype as the number of minor alleles (additive term), and a dominance term, which is only non-zero for heterozygotes.[30] This model is equivalent to labeling the genotypes. The dominance term was discarded if not significantly different from zero. We inferred haplotypes, of all possible sizes, for both genes, together with

8

the probabilities of the haplotypes being harboured by each indivdual, and analysed them using the methods of Schaid et al.[31] We summarized and discussed those models showing significant results (p-values below 0.05). Results corresponding to p-values below 0.01 are described as highly significant.

We did not use the Bonferroni correction for multiple testing, as such a correction is considered over-conservative when several genetic associations are tested in the same group of individuals [32], risking the rejection of important findings. Bonferroni correction might also be inappropriate in a situation such as this where there is a priori evidence that the genes are associated with TB [33], while Bayesian methods for correction rely on knowledge of prior probability of involvement, which is currently unknown for most genetic variants.[34]

The freely available (from www.r-project.org) programming environment, R and R packages were used for all statistics. The R package, genetics, was used to estimate genotype and allele frequencies and Hardy-Weinberg equilibrium probabilities [35]. Haplotype frequencies were inferred and analysed using the haplo.stats package [36] and LD heatmaps were drawn with LDheatmap using methods described in Shin et al [37].

Results

Single-point analysis of SNPs in MC3R

All SNPs were found to be in HWE in the control group. In the TB cases, two SNPs, (rs11575886 and rs3827103) were not in HWE, which might indicate an association with TB (Table 3). SNPs rs6127698, which is located 373bp upstream of the start codon, showed a statistically highly significant allelic association with TB susceptibility (p-value = 0.0004). The minor allele, T, was found less frequently in TB cases than in controls. Specifically, the odds of

9

TB is multiplied by 0.69/reduced by more than 30% for each additional T allele, compared to the GG homozygote (OR = 0.69; 95% CI: 0.56 – 0.85), after adjusting for age and gender.

SNPs rs11575886, showed a statistically significant association with TB susceptibility (p-value = 0.0423). The minor homozygote, CC, was found in 3 cases and not in controls, which resulted in the significant effect we detected. When we combined it with the CT heterozygote, effectively creating a dominant model, the genetic effect was no longer significant (p-value = 0.5490; OR=0.86; 95% CI: 0.52 – 1.42) for CC and CT versus TT, after adjusting for age and gender.

Single-point analysis of SNPs in CTSZ

Of the eight SNPs that were successfully genotyped, three (rs6064734, rs163785 and rs11540881) were monomorphic and excluded from further analyses. Each of the five remaining SNPs was in HWE in controls, and all except rs34069356 was in HWE in TB cases. The same SNP reported by Cooke *et al* to be associated in West Africa (rs34069356) also showed highly significant evidence of disease association in the South African Coloured population (p-value < 0.0001). No minor TT genotypes were found, and the TC heterozygote was over represented in TB patients compared to CC (OR = 3.45; 95% CI: 2.10 – 5.86.).

SNP rs13720 showed a significant additive allelic effect (p-value = 0.0487), with each G allele reducing the odds of TB by 12% (OR = 0.78; 95% CI: 0.600 – 0.999). Fitting a dominant model for G (grouping GG and AG) provided a slightly better fit (p-value = 0.0412), with AA and AG reducing the odds compared to the wild type GG (OR = 0.73; 95% CI: 0.53 – 0.99).

Haplotype analysis

10

The LD heatmaps, based on D', do not show obvious haplotype blocks inside the genes (Supplemental Figure 1). Each SNP is tightly linked with at least one other SNP in the same gene, not necessarily the closest neighbour. Two of the SNPs in *MC3R, viz.* rs72650656 and rs72650658 were not included in haplotype analysis since their minor allele frequencies were too low. We examined TB-haplotype association with a progressively larger sliding window. Supplemental Table 1 gives global p-values, adjusted for age and gender. It is clear that all the haplotypes showing significant associations in *MC3R* contain rs6127698, which had a significant effect individually. Three haplotypes were inferred for the 2-SNP *MC3R* rs6127698-rs11575886 haplotype, G-C (frequency = 0.04 in both cases and controls), T-T (frequency = 0.29 in controls and 0.22 in cases) and G-T (frequency = 0.67 in controls and 0.75 in cases); T-C was not observed at all (Supplemental Table 2). The significant odds ratio for TB was with T-T versus G-T (OR = 0.69; 95% CI: 0.56 – 0.84). Longer haplotypes in *MC3R* showed similar effects (Supplemental Table 2). Haplotype rs6127698-rs11575886-rs72650657 showed a significant OR of 0.68 (95% CI: 0.55 – 0.83) for each T-T-C haplotype (frequency = 0.29 in controls and 0.22 in TB cases) compared to the reference G-T-C haplotype (frequency = 0.64 in controls and 0.73 in cases). No other haplotype occurred at a frequency of more than 0.05. Four possible haplotypes were not observed in our study group.

In *CTSZ*, rs34063956 with and without rs3787492, which were both significant in the single SNP analyses, were included in the highly significant TB-association haplotypes. Supplemental Table 2 summarizes the model for haplotype rs13720-rs34063956-rs3787492. The A-C-G haplotype (frequency = 0.20 in controls and 0.17 in cases) appears to protect (OR = 0.73; 95% CI: 0.53 – 0.99) against TB compared to the G-T-A haplotype (frequency = 0.20 in controls and 0.23 in cases). Four possible haplotypes were not observed in our study group.

11

To assess the combined effect of the *MC3R* rs6127698 and *CTSZ* rs34069656 SNPs, we calculated odds ratios and p values for the allele combinations and adjusted for age and gender (Supplementary Table 3). The global p value was < 0.0001 and two highly significant effects, one protective (T-C combination, frequency = 0.29 in controls and 0.20 in cases, OR = 0.71, 95% CI: 0.56 – 0.89) and the other risk (G-T combination, frequency = 0.03 in controls and 0.10 in cases, OR = 3.16, 95% CI: 1.75 – 5.73), compared to the reference G-C (frequency = 0.68 in both groups) were detected.

Discussion

We have validated the association between TB and both *CTSZ* and *MC3R* which was first identified in a genome-wide linkage study. A number of linkage analyses have been conducted in an attempt to identify novel loci involved in susceptibility to tuberculosis. In 2000, Bellamy *et al* found evidence that chromosomes 15q and Xq may have linkage to tuberculosis while Greenwood *et al* found significant linkage with chromosome region 2q35.[7, 38] A study of the Brazilian population implicated chromosomes 10, 11 and 20, and in a Moroccan population, Baghdadi *et al* showed significant linkage between tuberculosis and chromosome 8q12-q13.[39, 40]

A genome-wide linkage study by Cooke *et al* identified a locus on chromosome 20q13.31-33 containing *MC3R* and *CTSZ*, which showed linkage with tuberculosis susceptibility. In 2008, Stein *et al* performed a genome linkage study in a large population from Uganda and replicated this finding (p = 0.002), identifying a 25cM long region containing both *MC3R* and *CTSZ*.[41]

We have now conducted an independent, unrelated case-control study and found that the same SNP implicated in *CTSZ* by Cooke *et al*, rs34069356, showed evidence of disease

12

association in the South African Coloured population (p < 0.0001, adjusted for age and gender). Cooke *et al* determined that TT homozygous individuals were more susceptible to TB, but we found no individuals with the TT genotype. TC heterozygotes were however significantly overrepresented in TB patients (23% in cases versus 7% in controls). This polymorphism results in a non-conservative amino acid change of a non-polar alanine to a polar, uncharged threonine. Though it does not appear that the amino acid substitution occurs in an active site on CTSZ, the introduction of a hydroxyl side chain has many possible implications. Threonine has an uncharged, polar side chain, making the amino acid hydrophilic. Unlike alanine (a hydrophobic amino acid typically located on the interior of a protein), threonine is typically located on the exterior of a protein where the hydroxyl side chain is free to interact with surrounding water molecules. Threonine, but not alanine, is also subject to a number of post-translational modifications (PTMs) including phosphorylation by threonine kinases, O-linked glycosylation and acetylation.[42] Phosphorylation is known to regulate the activity of proteins and since the amino acid substitution introduced by SNP rs34069356 occurs close to the N-terminal of the protein it is likely that this threonine is available for phosphorylation. The introduction of a hydroxyl group and a number of PTMs is likely to affect protein folding, intracellular localization and protein activity. [42]

A SNP located 373bp upstream of the *MC3R* gene (rs6127698) was significantly associated with TB (p = 0.0004, adjusted for age and gender). This SNP is predicted by Genomatix to create an alternative transcription factor binding site (http://www.genomatix.de/). The SNP in the single exon of *MC3R* associated in the study of Cooke *et al* (rs3827103) was not significantly associated with TB susceptibility in the South African Coloured population. SNP rs6127698 was not genotyped by Cooke *et al* and it should be noted that the r^2 value between

13

rs3827103 and rs6127698 is quite low (in controls: $r^2 = 0.16$, D' = 0.99, data not shown), indicating that these two alleles are not completely predictive of each other in our population. This might explain why we did not find rs3827103 to be associated with TB in our study. The effects of the creation of an alternative transcription factor binding site are difficult to predict as a polymorphism in this region may result in either an increase or decrease in the transcription of the *MC3R* gene. Further studies must be performed to obtain a better understanding of the effect of such a polymorphism.

A combined analysis of the individually significant *CTSZ* and *MC3R* SNPs revealed two highly significant effects on TB susceptibility, one protective and the other risk. In the study done by Cooke *et al* the strongest evidence for linkage was within the region 20q13.31-33 with a single point LOD score of 3.1, $p = 10^{-4}$ and a maximum likelihood score MLS of 2.8, p = 0.00008. In our study, the G-T allele combination of the associated SNPs had an odds ratio of 3.16, 95% CI: 1.75 – 5.73, p = 0.0001. The linkage LOD score of Cooke *et al* and our association OR are high and, although these are different measures, it is tempting to speculate that the associated alleles together gave rise to the original observed linkage signal that led us to focus on this region. However, it is also possible that these SNPs are in LD with SNPs located in other adjacent genes and are therefore essentially tag SNPs of a non-neighboring group of highly correlated SNPs ('bin') as was the case for the *PARK2/PACRG* gene in leprosy.[43] We did not detect any bins for the associated SNPs using the LDselect algorithm as implemented by the Genome Variation Server (http://gvs.gs.washington.edu/GVS/index.jsp) [44].

This study has validated the findings implicating *MC3R* and *CTSZ* in tuberculosis susceptibility and provides convincing evidence to motivate further investigation into the mechanisms of action of their respective pathways in TB progression.

14

Conflict of interest statement: The authors declare no conflict of interest.

Supplementary information is available at The European Journal of Human Genetics' website.

15

References

1. Clementi M and Di Gianantonio E: Genetic susceptibility to infectious diseases. *Reprod Toxicol* 2006; **21:** 345-349

2. Möller M and Hoal EG: Current findings, challenges and novel approaches in human genetic susceptibility to tuberculosis. *Tuberculosis (Edinb)* 2010; **90:** 71-83

3. Mira MT, Alcaïs A, Nguyen VT *et al*: Susceptibility to leprosy is associated with *PARK2* and *PACRG*. *Nature* 2004; **427:** 636-640

4. Siddiqui MR, Meisner S, Tosh K *et al*: A major susceptibility locus for leprosy in India maps to chromosome 10p13. *Nat Genet* 2001; **27:** 439-441

5. Tosh K, Meisner S, Siddiqui MR *et al*: A region of chromosome 20 is linked to leprosy susceptibility in a South Indian population. *J Infect Dis* 2002; **186:** 1190-1193

6. Mira MT, Alcaïs A, Van Thuc N *et al*: Chromosome 6q25 is linked to susceptibility to leprosy in a Vietnamese population. *Nat Genet* 2003; **33:** 412-415

7. 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:** 8005-8009

8. Cooke GS, Campbell SJ, Bennett S *et al*: Mapping of a novel susceptibility locus suggests a role for *MC3R* and *CTSZ* in human tuberculosis. *Am J Respir Crit Care Med* 2008; **178:** 203-207

9. 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. *Journal of Biological Chemistry* 1998; **273:** 16816-16823

10. Zavasnik-Bergant T and Turk B: Cysteine cathepsins in the immune response. *Tissue Antigens* 2006; **67:** 349-355

16

11. Journet A, Chapel A, Kieffer S, Louwagie M, Luche S, and Garin J: Towards a human repertoire of monocytic lysosomal proteins. *Electrophoresis* 2000; **21:** 3411-3419

12. Garin J, Diez R, Kieffer S *et al*: The phagosome proteome: insight into phagosome functions. *J Cell Biol* 2001; **152:** 165-180

13. 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:** 103-113

14. 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:** 233-244

15. Kim DK, Park GM, Hwang YI *et al*: Microarray analysis of gene expression associated with extrapulmonary dissemination of tuberculosis. *Respirology* 2006; **11:** 557-565

16. Lee EJ, Lee SH, Jung JW *et al*: Differential regulation of cAMP-mediated gene transcription and ligand selectivity by MC3R and MC4R melanocortin receptors. *Eur J Biochem* 2001; **268:** 582-591

17. Wang SX, Fan ZC, and Tao YX: Functions of acidic transmembrane residues in human melanocortin-3 receptor binding and activation. *Biochem Pharmacol* 2008; **76:** 520-530

18. Getting SJ, Lam CW, Leoni G, Gavins FN, Grieco P, and Perretti M: [D-Trp8]-gamma-melanocyte-stimulating hormone exhibits anti-inflammatory efficacy in mice bearing a nonfunctional MC1R (recessive yellow e/e mouse). *Mol Pharmacol* 2006; **70:** 1850-1855

19. Lee YS, Poh LK, Kek BL, and Loke KY: The role of melanocortin 3 receptor gene in childhood obesity. *Diabetes* 2007; **56:** 2622-2630

17

20. Tao YX: Functional characterization of novel melanocortin-3 receptor mutations identified from obese subjects. *Biochim Biophys Acta* 2007; **1772:** 1167-1174

21. Tao YX and Segaloff DL: Functional characterization of melanocortin-3 receptor variants identify a loss-of-function mutation involving an amino acid critical for G protein-coupled receptor activation. *J Clin Endocrinol Metab* 2004; **89:** 3936-3942

22. Mencarelli M, Walker GE, Maestrini S *et al*: Sporadic mutations in melanocortin receptor 3 in morbid obese individuals. *Eur J Hum Genet* 2008; **16:** 581-586

23. Rached M, Buronfosse A, Begeot M, and Penhoat A: Inactivation and intracellular retention of the human I183N mutated melanocortin 3 receptor associated with obesity. *Biochim Biophys Acta* 2004; **1689:** 229-234

24. Feng N, Young SF, Aguilera G *et al*: Co-occurrence of two partially inactivating polymorphisms of *MC3R* is associated with pediatric-onset obesity. *Diabetes* 2005; **54:** 2663-2667

25. Lee YS, Poh LK, and Loke KY: A novel melanocortin 3 receptor gene (*MC3R*) mutation associated with severe obesity. *J Clin Endocrinol Metab* 2002; **87:** 1423-1426

26. Chen AS, Marsh DJ, Trumbauer ME *et al*: Inactivation of the mouse melanocortin-3 receptor results in increased fat mass and reduced lean body mass. *Nat Genet* 2000; **26:** 97-102

27. 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

28. Gallant CJ, Cobat A, Simkin L *et al*: The impact of age and sex on anti-mycobacterial immunity of children and adolescents in an area of high tuberculosis incidence. *International Journal of Tuberculosis and Lung Disease* 2010; **14:** 952-958

18

29. Hampe J, Wollstein A, Lu T *et al*: An integrated system for high throughput TaqMan based SNP genotyping. *Bioinformatics (Oxford, England)* 2001; **17**: 654-655

30. Cordell HJ and Clayton DG: Genetic association studies. *Lancet* 2005; **366**: 1121-31

31. Schaid DJ, Rowland CM, Tines DE, Jacobson RM, and Poland GA: Score tests for association between traits and haplotypes when linkage phase is ambiguous. *American Journal of Human Genetics* 2002; **70**: 425-434

32. Nyholt DR: A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. *Am J Hum Genet* 2004; **74**: 765-769

33. Perneger TV: What's wrong with Bonferroni adjustments. *BMJ* 1998; **316**: 1236-1238

34. Campbell H and Rudan I: Interpretation of genetic association studies in complex disease. *Pharmacogenomics J* 2002; **2**: 349-360

35. Warnes G, Gorman D, Leisch F, and Man M: genetics: Population Genetics. http://CRAN.R-project.org/package=genetics. 2008.

36. Sinwell JP and Schaid DJ: haplo.stats: Statistical Analysis of Haplotypes with Traits and Covariates when Linkage Phase is Ambiguous. http://CRAN.R-project.org/package=haplo.stats. 2010.

37. Shin J-H, Blay S, McNeney B, and Graham J: LDheatmap: An R Function for Graphical Display of Pairwise Linkage Disequilibria Between Single Nucleotide Polymorphisms. *Journal of Statistical Software* 2006; **16**: Code Snippet 3

38. 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**: 405-416

19

39. Miller EN, Jamieson SE, Joberty C *et al*: Genome-wide scans for leprosy and tuberculosis susceptibility genes in Brazilians. *Genes Immun* 2004; **5:** 63-67

40. 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:** 1679-1684

41. Stein CM, Zalwango S, Malone LL *et al*: Genome scan of *M. tuberculosis* infection and disease in Ugandans. *PLoS ONE* 2008; **3:** e4094

42. Cooper GM (2000) The Cell: A Molecular Approach. Second Edition ed.

43. Alter A, Alcaïs A, Abel L, and Schurr E: Leprosy as a genetic model for susceptibility to common infectious diseases. *Hum Genet* 2008; **123:** 227-235

44. Carlson CS, Eberle MA, Rieder MJ, Yi Q, Kruglyak L, and Nickerson DA: Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. *Am J Hum Genet* 2004; **74:** 106-120

Table 1. Characteristics of the samples used in this study.

All samples (n = 1186)	Number	Female (%)	Age, mean±SD
TB cases	628	327 (52%)	29.2±12.2
Controls	558	411 (74%)	33.2±14.3
Samples genotyped for *MC3R* (n = 1020)			
TB cases	504	256 (51%)	33.7±14.1
Controls	516	383 (74%)	29.8±12.0
Samples genotyped for *CTSZ* (n = 857)			
TB cases	481	256 (53%)	32.5±14.7
Controls	376	288 (77%)	26.6±12.3

21

Table 2. SNPs genotyped in *MC3R* and *CTSZ*.

Gene	rs Number	Position	Genotyping Method
MC3R	rs72650656	Promoter	Sequencing
	rs6127698	Promoter	Sequencing
	rs11575886	Promoter	Sequencing
	rs72650657	Promoter	Sequencing
	rs72650658	Promoter	Sequencing
	rs3827103	Exon 1, missense	Sequencing
CTSZ	rs448943	3` UTR[a]	SNPlex
	rs10369	3` UTR	SNPlex
	rs13720	3` UTR	SNPlex
	rs34069356	3` UTR	Taqman
	rs6064734	Exon 6, synonymous	SNPlex
	rs3787492	Exon 6/intron 5 boundary	SNPlex
	rs163785	Exon 4, synonymous	SNPlex
	rs11540881	Exon 4 synonymous	SNPlex

[a] Untranslated Region

22

99

Table 3. Single-SNP statistical analysis of *MC3R* and *CTSZ*.

	Controls		HWE[c]	TB Cases		HWE	Association
	Count[a]	Freq[b]	p-value	Count	Freq	p-value	p-value[d]
MC3R rs72650656							
Typed	496		1.0000	492		1.0000	
C/C	489	0.99		484	0.98		0.7854
C/G	7	0.01		8	0.02		
C	985	0.99		976	0.99		0.7854
G	7	0.01		8	0.01		
MC3R rs6127698							
Typed	506		0.3334	498		0.1134	
G/G	258	0.51		311	0.62		0.0016
G/T	200	0.40		157	0.32		
T/T	48	0.09		30	0.06		
G	716	0.71		779	0.78		0.0004
T	296	0.29		217	0.22		
MC3R rs11575886							
Typed	506		1.0000	496		0.0216	
T/T	469	0.93		463	0.93		0.0423
T/C	37	0.07		30	0.06		
C/C	0	0.00		3	0.01		

23

	Controls		HWE[c]	TB Cases		HWE	Association
	Count[a]	Freq[b]	p-value	Count	Freq	p-value	p-value[d]
T	975	0.96		956	0.96		0.8854
C	37	0.04		36	0.04		
MC3R rs72650657							
Typed	507		1.0000	495		1.0000	
C/C	480	0.95		474	0.96		0.4066
C/G	27	0.05		21	0.04		
C	987	0.97		969	0.98		0.4066
G	27	0.03		21	0.02		
MC3R rs72650658							
Typed	497		1.0000	490		1.0000	
C/C	493	0.99		480	0.98		0.1205
C/G	4	0.01		10	0.02		
C	990	1.00		970	0.99		0.1205
G	4	0.00		10	0.01		
MC3R rs3827103							
Typed	495		0.1159	488		0.0377	
G/G	242	0.49		214	0.44		0.1939
G/A	197	0.40		202	0.41		
A/A	56	0.11		72	0.15		
G	681	0.69		630	0.65		0.0730

24

	Controls		HWE[c]	TB Cases		HWE	Association
	Count[a]	Freq[b]	p-value	Count	Freq	p-value	p-value[d]
A	309	0.31		346	0.35		

CTSZ rs448943

	Controls		HWE	TB Cases		HWE	Association
Typed	371		1.0000	467		0.0784	
A/A	107	0.29		136	0.29		0.3383
A/G	185	0.50		214	0.46		
G/G	79	0.21		117	0.25		
A	399	0.54		486	0.52		0.4399
G	343	0.46		448	0.48		

CTSZ rs10369

	Controls		HWE	TB Cases		HWE	Association
Typed	375		1.0000	478		0.7984	
C/C	241	0.64		282	0.59		0.3201
C/T	119	0.32		169	0.35		
T/T	15	0.04		27	0.06		
C	601	0.80		733	0.77		0.1673
T	149	0.20		223	0.23		

CTSZ rs13720

	Controls		HWE	TB Cases		HWE	Association
Typed	376		0.1965	477		0.0748	
A/A	245	0.65		333	0.70		0.1221
A/G	112	0.30		124	0.26		
G/G	19	0.05		20	0.04		

25

	Controls		HWE[c]	TB Cases		HWE	Association
	Count[a]	Freq[b]	p-value	Count	Freq	p-value	p-value[d]
A	602	0.80		790	0.83		0.0487
G	150	0.20		164	0.17		
CTSZ rs34069356							
Typed	298		1.0000	396		0.0031	
C/C	276	0.93		303	0.77		< 0.0001
C/T	22	0.07		93	0.23		
C	574	0.96		699	0.88		< 0.0001
T	22	0.04		93	0.12		
CTSZ rs3787492							
Typed	356		0.2206	451		0.4479	
C/C	188	0.53		253	0.56		0.6111
C/T	148	0.42		174	0.39		
T/T	20	0.06		24	0.05		
C	524	0.74		680	0.75		0.3919
T	188	0.26		222	0.25		

[a] Genotype and allelic count

[b] Genotype and allelic frequencies

[c] p-values for exact test of Hardy-Weinberg equilibrium (HWE) stratified by TB susceptibility status

26

^d p-value for association between polymorphism and TB susceptibility, adjusted for age and gender

Supplemental Material

Supplemental Table 1. Sliding window haplotype analysis of *MC3R* and *CTSZ*.

|	Number of SNPs			
First SNP	**2**	**3**	**4**	**5**
MC3R: rs6127698-rs11575886-rs72650657-rs3827103				
rs6127698	0.0016[a]	0.0052	0.0214	
rs11575886	0.7292	0.4003		
rs72650657	0.2057			
rs3827103				
CTSZ: rs448943-rs10369-rs13720-rs34063956-rs3787492				
rs448943	0.3470	0.2319	0.0001	0.0001
rs10369	0.1209	< 0.000001	< 0.000001	
rs13720	< 0.000001	< 0.0001		
rs34063956	< 0.0001			
rs3787492				

[a] Global p-values for TB association with haplotypes of increasing length, adjusted for age and gender.

28

Supplemental Table 2. Summary of highly significant haplotype-TB association models, adjusted for age and gender.

Haplotype	Frequency in controls	Frequency in cases	OR	95% CI	
MC3R: rs6127698-rs11575886					
G-C	0.04	0.04	0.90	0.56	1.44
G-T	0.67	0.75	0.69	0.56	0.84
T-T	0.29	0.22	1	Reference haplotype	
MC3R: rs6127698-rs11575886-rs72650657					
G-C-C	0.04	0.04	0.88	0.54	1.42
G-T-G	0.03	0.02	0.67	0.36	1.25
T-T-C	0.29	0.22	**0.68**	**0.55**	**0.83**
G-T-C	0.64	0.73	1	Reference haplotype	
CSTZ: rs13720-rs34063956-rs3787492					
A-C-A	0.34	0.35	0.93	0.70	1.23
A-C-G	0.20	0.17	**0.73**	**0.53**	**0.99**
G-C-A	0.26	0.25	0.86	0.64	1.17
G-T-A	0.20	0.23	1	Reference haplotype	

29

Supplemental Table 3. *MC3R* rs6127698 – *CTSZ* rs34069356 analysis.

Haplotype	Frequency in controls	Frequency in cases	OR	95% CI	
Global p value < 0.0001					
G-T	0.03	0.10	3.16	1.75	5.73
T-C	0.29	0.20	0.71	0.56	0.89
T-T	0.01	0.02	3.53	0.44	28.12
G-C	0.68	0.68	1	Reference haplotype	

30
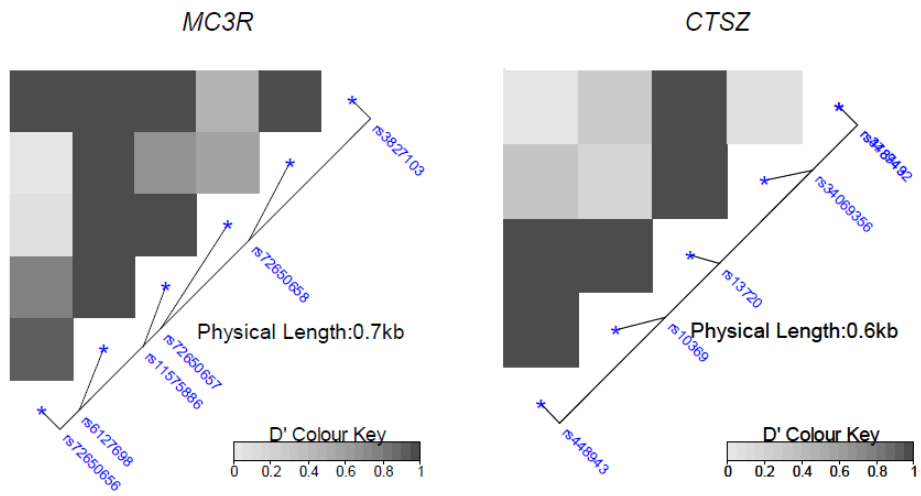
Supplemental Figure 1. LD heatmaps of *MC3R* and *CTSZ* in controls based on D'.