The Role of Single Nucleotide Polymorphisms in C-Type Lectin Receptors and the Signaling Molecules in their Pathways in Susceptibility towards developing Pulmonary Tuberculosis in an Indian Population.

Inaugural-Dissertation
to obtain the academic degree
Doctor rerum naturalium (Dr. rer. nat.)
submitted to the Department of Biology, Chemistry and Pharmacy
of Freie Universität Berlin

by

Surabhi Goyal

Berlin

2017

I performed my doctoral studies from 01.01.2014 to 31.03.2016 at the Institute of Microbiology and Hygiene from the Charité Medical School Berlin under the supervision of Prof. Dr. Hortense Slevogt.

1st Reviewer: Prof. Dr. Hortense Slevogt

2nd Reviewer: Prof. Dr. Rupert Mutzel

Date of defence: 14th February, 2018

Summary

Tuberculosis (TB) is a multifactorial disease being affected by bacterial, host as well as environmental factors. *Mycobacterium tuberculosis* (Mtb) engages a number of receptors during infection and the pathogenesis not only depends upon the receptors involved but all the adaptors down the signaling pathways and their cross-talk and involvement with other receptors/pathways. The members of C-type lectin receptors (CTLRs) have recently been described as important in TB pathogenesis. Some of these receptors can directly recognize the potent virulence factors of Mtb cell wall and induce several pro-inflammatory immune responses. Not many studies have been performed looking into the genetic variations among these receptors and their effects of TB disease. In this study, important CTLRs and their signaling molecules involved in TB infection and pathogenesis were investigated for potential TB associated single nucleotide polymorphisms (SNPs). We performed a case-control study consisting of 144 HIV-negative new pulmonary TB cases and 181 healthy controls recruited in Hyderabad, India, and the DNA was collected from the blood of each subject. A two stage sequencing approach was adopted in which candidate common variations were first screened out in a pilot explorative phase by Ampliseq based next generation sequencing consisting of 80 samples. An adjustment for population stratification was performed assuming a heterogeneous population and the candidate SNPs were genotyped and verified in all the samples in the validation phase. The results showed no association of SNPs in CTLRs with the occurrence of (pulmonary TB) PTB. However, while also focusing on signaling proteins related to CTLRs we found that SNP rs3774275 in *MASP1,* which is downstream of the MBL pathway, is significantly associated with pulmonary TB (PTB) in our population (meta-analysis p=0.034). The G allele occurs more frequently among controls and seems to provide a protective effect against TB in this study population. Furthermore, the MASP-1 and Map44 serum levels are significantly higher in TB patients when compared to healthy controls. A further *in vitro* experiment with recombinant human MASP-1 (rhMASP-1) demonstrated that addition of MASP-1 in serum increases the lectin pathway activity, suggesting a functional role of MASP-1 in TB pathogenesis. In conclusion, this study demonstrates a significant relationship between MASP-1 polymorphisms and serum levels and development of pulmonary TB, suggesting an important role of lectin pathway in TB pathogenesis. Moreover, the results propose MASP-1 as a potential genetic marker for TB resistance.

Zusammenfassung

Tuberkulose (TB) ist eine multifaktorielle Erkrankung, deren Verlauf von Bakterien-, Wirt- und Umweltfaktoren beeinflusst wird. In Falle einer Infektion wird Mtb, der Erreger der Tuberkulose von eine Anzahl verschiedener Rezeptoren des innaten Immunsystems erkannt und eine Immunreaktion initiiert. Die Mitglieder der C-Typ-Lektin-Rezeptoren (CTLRs) wurden kürzlich als wichtig in der Pathogenese der TB beschrieben. Es konnte gezeigt werden, dass Rezeptoren dieser Familie direkt bestimmte Oberflächenmoleküle der Mtb-Zellwand erkennen und verschiedene proinflammatorische Reaktionen induzieren. Ziel dieser Arbeit war es in einer klinischen Studie zu testen, ob CTLRs, die an der TB-Infektion und der Pathogenese beteiligt sind, bestimmte Nukleotidpolymorphismen aufweisen, die mit dem Risiko an einer TB zu erkranken assoziiert sind. Dafür haben wir eine Fall-Kontroll-Studie, bestehend aus 144 neuen pulmonalen TB-Fällen und 181 gesunden Kontrollen, durchgeführt, die im Bhagwan Mahavir Medical Research Center (BMMRC), Hyderabad, Indien, rekrutiert wurden. Von jedem Patienten/Probanden wurden Blutproben zur dann Extraktion gesammelt und wichtige demografische und sozioökonomische Daten erhoben. Es wurde ein zweiphasiger Sequenzierungsansatz angewendet, bei dem die Varianz innerhalb der Patientenpopulation im Vergleich zu gesunden Kontrollpersonen zuerst in einer Pilot-Explorationsphase, bestehend aus 80 Probandenproben, bestimmt wurde. Eine Anpassung an die jeweilige Abstammung wurde unter der Annahme einer heterogenen Population durchgeführt. Die identifizierten SNP-Kandidaten wurden anschließend in allen Probanden-Proben in der Validierungsphase genotypisiert. Die Ergebnisse zeigen, dass keine SNP den Genen der CTLR mit dem Risiko einer TB assoziiert war. Allerdings fanden wir, dass SNP rs3774275 in *MASP1* signifikant mit dem Auftreten einer pulmonalem TB (PTB) in unserer Population assoziiert ist (Metaanalyse $p = 0,034$). MASP-1 ist ein wichtiges Mitglied des Lektin-Weges des Komplementssystems, dessen PRR mannose-binding lektin (MBL) in der Lage ist, Mtb zu erkennen. Das G-Allel tritt häufiger bei den Kontrollen auf und scheint eine schützende Wirkung gegen TB in dieser Studienpopulation zuhaben. In weiterführenden Untersuchungen konnten wir zeigen, dass die MASP-1 Serumkonzentrationen bei TB-Patienten signifikant höher sind als bei gesunden Kontrollen. Ein weiteres *in vitro*-Experiment mit rekombinantem humanem MASP-1 (rhMASP-1) zeigte, dass die Zugabe von MASP-1 im Serum die Aktivität der

Aktivierung des Lektin-Komplementweges erhöht, was auf eine funktionelle Rolle von MASP-1 in der TB-Pathogenese hindeutet. Zusammenfassend konnten wir zeigen, dass eine signifikante Beziehung zwischen dem hier beschriebenen *MASP-1* SNP mit den Serumkonzentrationen dieses Proteins sowie mit dem Risiko assoziiert ist, an einer TB zu erkranken.Unsere Ergebnisse geben einen interessanten Hinweis darauf, dass das Komplementsystem einen wichtigen Beitrag zur TB-induzierten Immunabwehr beiträgt. Dieses sollte in weiteren Studien genauer adressiert werden.

TABLE OF CONTENTS

ABBREVIATIONS……8

CHAPTER 1: INTRODUCTION……10

1.1 TUBERCULOSIS DISEASE AND EPIDEMIOLOGY……10

1.1.1 EPIDEMIOLOGY……10

1.1.2 CLINICAL PRESENTATION……11

1.1.3 DIAGNOSIS……11

1.1.4 TREATMENT AND CARE……13

1.2 HOST- MTB INTERACTION……14

1.2.1 MTB VIRULENCE……15

1.2.2 HOST RESPONSE AGAINST MTB……16

1.2.3 EARLY RESPONSE……17

1.2.4 GRANULOMA FORMATION……18

1.3 PRRs IN MYCOBACTERIAL IMMUNITY……19

1.3.1 C-TYPE LECTIN RECEPTORS……21

1.3.1.1 GENERAL STRUCTURE AND SIGNALLING MECHANISMS……21

1.3.1.2 CTLRs INTERACTION WITH MYCOBACTERIUM TUBERCULOSIS……22

1.4 THE IMPACT OF SNPs IN HOST RESPONSE TO TB……25

1.5 AIM OF THE STUDY……27

CHAPTER 2: MATERIALS AND METHODS……28

2.1 OBSERVATIONAL STUDY……28

2.1.1 COHORT RECRUITMENT……28

2.1.2 TB CASES AND CONTROLS……28

2.1.3 SAMPLE AND DATA COLLECTION……29

2.1.4 DATA VARIABLES……29

2.1.5 STUDY SIZE……30

2.2 DETECTION OF POLYMORPHISMS……31

2.2.1 DISCOVERY (PHASE I)……31

2.2.1.1 AMPLISEQ TARGET SLELECTION……31

2.2.1.2 DNA SAMPLE SLECETION & QUALITY ASSESSMENT……36

2.2.1.3 DNA AMPLISEQ PRIMERS AND LIBRARY PRERPARATION……37

2.2.1.4 ION TORRENT NEXT GENERATION SEQUENCING……37

2.2.1.5 DATA ANALYSIS……38

2.2.2 VALIDATION (PHASE II)……39

2.2.2.1 SEQUENOM MASSARRAY SEQUENCING……39

2.2.2.2 META-ANALYSIS OF ASSOCIATION RESULTS……40

2.3 FUNCTIONAL ANALSYIS OF rs3774275……40

2.3.1 MEASUREMENT OF MASP-1, MASP-3 AND MAp44 IN SERUM……40

2.3.2 MEASUREMENT OF LECTIN PATHWAY ACTIVITY……40

CHAPTER 3: RESULTS 42
3.1 POPULATION CHARACTERISTICS 42
3.2 EXPLORATIVE PHASE RESULTS 44
3.2.1 ASSOCIATION ANALYSIS YIELDED MORE THAN 600 POTENTIAL TB-ASSOCIATED SNPs 44
3.2.2 STRATIFICATION ADJUSTMENTS 46
3.2.3 SCREENING OF SNPs FOR PHASE II 47
3.3 VALIDATION PHASE RESULT: ONLY *MASP1* SNP IS SIGNIFICANTLY ASSOCIATED WITH TB SUSCEPTIBILITY 49
3.4 MASP-1 LEVELS ARE HIGHER IN TB CASES THAN HEALTHY CONTROLS 51
3.5 GENOTYPE DEPENDENT DISTRIBUTION OF MASP-1 SERUM LEVELS 53
3.6 ASSOCIATION OF SERUM MASP-1 LEVELS WITH BLOOD GLUCOSE AND BLOOD CHOLESTEROL LEVELS 54
3.7 MASP-1 LEVELS AFFECT THE LECTIN COMPEMENT PATHWAY ACTIVITY 55
CHAPTER 4: DISCUSSION 56
REFERENCES 64
APPENDICES 73
ACKNOLEDGEMENTS 96

ABBREVIATIONS

AM	Alveolar Macrophages
BCG	Bacillus Calmette–Guérin
BCL10	B-Cell Lymphoma/leukemia 10
CRD	Carbohydrate Recognition Domain
CTLD	C-type Lectin like Domain
CTLR	C-type Lectin Receptor
CXR	Chest X-Ray
DAP	DNAX-Activating Protein
DC-SIGN	Dendritic Cell-Specific Intercellular adhesion molecule-3-Grabbing Non-integrin
DOTS	Directly Observed Treatment, Short-course
DST	Drug Sensitivity Test
EPTB	Extra-pulmonary TB
FcRγ	Fc-gamma Receptor
GATK	Genome Analysis Toolkit
GWAS	Genome-wide Association Study
HIV	Human Immunodeficiency Virus
IGRA	IFN-gamma Release Assay
ISP	Ion Sphere Particles
ITAM	Immunoreceptor Tyrosine-based Activation Motif
ITIM	Immunoreceptor Tyrosine-based Inhibition Motif
LD	Linkage Disequilibrium
LM	Lipomannan
LPA	Line Probe Assay
LTBI	Latent Tuberculosis Infection
MAF	Minor Allele Frequency
MALT1	Mucosa-associated lymphoid tissue lymphoma translocation protein 1
ManLAM	Mannose-capped Lipoarabinomannan
Map44	MBL-Associated Protein of 44 kDa
MASP	MBL-Associated Serine Protease
MBL	Mannose Binding Lectin
MCL	Macrophage C-type Lectin
MDRTB	Multidrug Resistant Tuberculosis
MINCLE	Macrophage Inducible C-type Lectin
MR	Macrophage Receptor

Mtb	*Mycobacterium tuberculosis*
MyD88	Myeloid Differentiation primary response gene 88
NAAT	Nucleic Acid Amplification Test
NCBI	National Center for Biotechnology Information
NGS	Next Generation Sequencing
NKC	Natural Killer Cell
NLR	NOD-Like Receptor
NOD	Nucleotide-binding Oligomerization Domain
PAMP	Pathogen Associated Molecular Patterns
PCA	Principal Component Analysis
PIM	Phosphatidylmyo-inositol Mannosides
PKC	Protein Kinase C delta
PRR	Pattern Recognition Receptor
PTB	Pulmonary Tuberculosis
rhMASP-1	Recombinant Human MBL-Associated Serine Protease-1
RIP2	Receptor Interacting Protein-2
RNS	Reactive Nitrogen Species
ROS	Reactive Oxygen Species
SNP	Single Nucleotide Polymorphisms
SNV	Single Nucleotide Variations
SRA	Sequence Read Archive
TB	Tuberculosis
Tc	T-cytotoxic cell
TDM	Trehalose-6,6-dimycolate
Th	T-helper
TLR	Toll-like Receptor
TNF	Tumor Necrosis Factor
TRIF	TIR-domain-containing adapter-inducing interferon-β
TST	Tubereculin Skin Test
WHO	World Health Organization

CHAPTER 1: INTRODUCTION

1.1 TUBERCULOSIS DISEASE AND EPIDEMIOLOGY

1.1.1 EPIDEMIOLOGY

Tuberculosis (TB), a communicable infectious disease, is the second leading cause of deaths from an infectious disease worldwide after human immunodeficiency virus (HIV). About 10.4 million new cases of TB and about 1.8 million TB related deaths were reported in 2015 according to Global Tuberculosis Report 2016 [1]. In 2015, the largest numbers of incident cases were reported in Asian (61%) and African (26%) countries whereas the Eastern Mediterranean, European and American regions together comprised only 13% of incidence. India alone accounts for more than a quarter of world's TB cases and deaths (Fig-1). India is also one of the countries where highest number of relapse and multi-drug resistant (MDR-TB) TB cases have been reported [1].

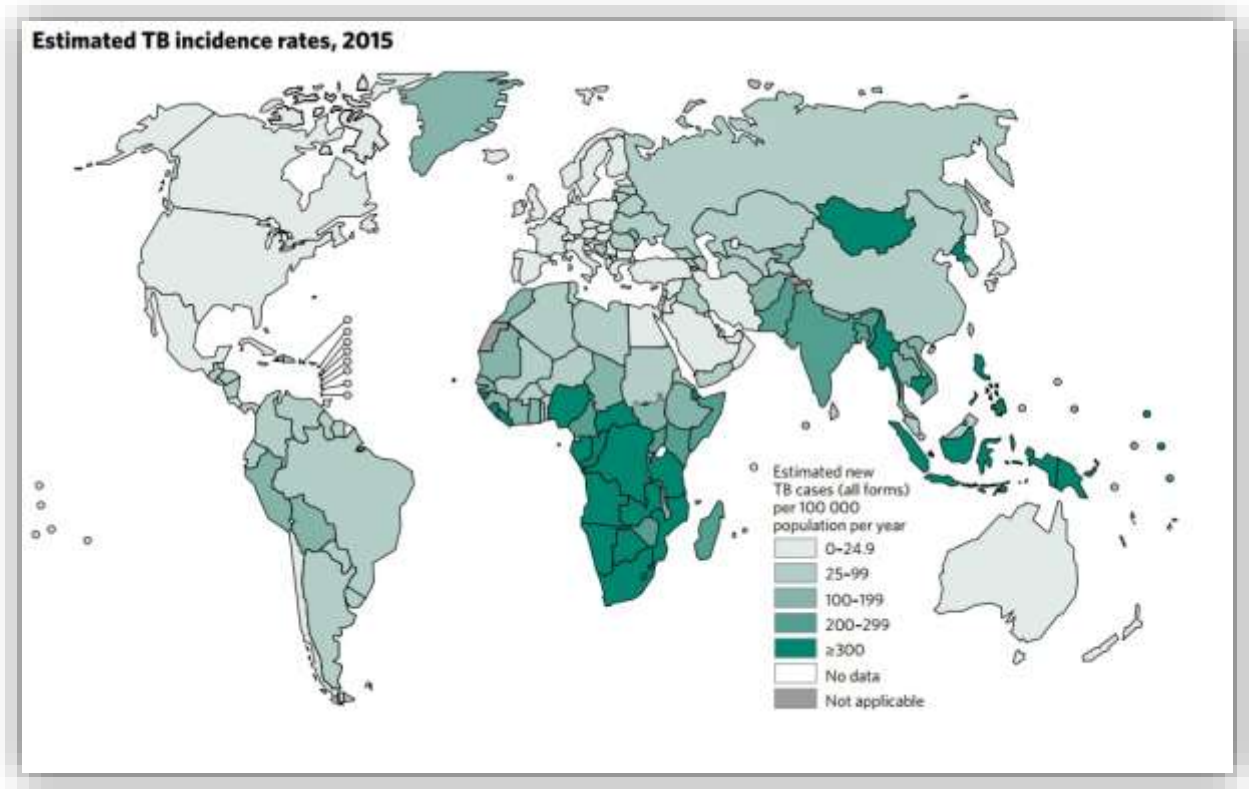


Figure 1: Worldwide estimated TB (all forms) incidence in 2015. (Source: WHO Tuberculosis Report 2016 [1]).

The incidence of TB in India in 2016 is expected to be much higher than previously reported [1]. The reason for high TB burden in most low or middle income countries could be related to the risk factors such as poverty, malnutrition, crowded living conditions, poor ventilation and hygiene, which are more pronounced in these countries [2, 3]. Other risks include alcohol misuse, smoking, HIV, immunosuppressive treatments, diabetes, advanced malignancy etc.[4]. New studies are also demonstrating that individual mutations in genes may also pose a risk towards developing TB in different populations [5].

1.1.2 CLINICAL PRESENTATION

Tuberculosis is an air borne communicable disease transmitted when the aerosols (1-4 µm in length) containing the bacteria *Mycobacterium tuberculosis* (Mtb) coughed up by an infected person are inhaled by healthy individuals [6]. The bacteria primarily affects lungs causing pulmonary TB (PTB) but may also spread to other body parts leading to extra-pulmonary TB (EPTB) depending upon factors such as age, immune status, Mtb strain, host genetics etc [7]. Once inside the body, the bacteria may cause an active (primary TB) or a latent TB infection (LTBI). About one-third of the world's population is believed to be infected with Mtb but in many cases the bacteria are latently contained without any illness or symptoms. However, 5-10% of these latently infected individuals are at a risk to develop an active infection during their lifetime (reactivated TB) depending upon factors such as age, malnutrition, HIV infection etc. that can affect the status of their immune system [7].

An active infection is mostly accompanied by productive cough, sputum production, appetite and weight loss, fever, night sweats as well as haemoptysiss [7]. However, in early stages of infection, the symptoms may be mild or completely absent leading to delay in treatment and high chances of spreading the disease [8]. Improper diagnosis and/or incomplete treatment regimens may lead to development of drug resistant strains causing MDR-TB which is difficult to treat and requires longer and use of more powerful drugs [7].

1.1.3 DIAGNOSIS

The resource-constrained countries, which share about 90% of burden of TB, mostly rely on sputum smear microscopy and chest radiography as primary methods of

active TB detection [8]. About three-fourth of the adults with active TB can be detected with smear microscopy. The grading, as per the WHO recommendations, is as follows- No bacteria in 100 immersion fields :Negative; 1-9 AFB in 100 immersion fields: Positive scanty, exact number recorded; 10-99 AFB in 100 immersion fields: 1+ ; 1 to 10 AFB per field in 50 fields: 2+ ; More than 10 AFB per field in 20 fields: 3+ [9]. The method is inexpensive and requires minimal biosafety standards and thus has a high value in TB diagnosis albeit with low sensitivity. Nevertheless, it doesn't provide any information about the drug susceptibility of the bacteria [10, 11].

Microscopy results are usually supplemented by chest X-rays which can show the presence of characteristic fluffy upper zone shadowing representing alveolar macrophages, lymphadenopathy and cavitation [12]. Radiography can prove helpful in evaluating suspicious cases of TB where the microscopy or PCR based tests have shown negative results. Chest X-rays have a high sensitivity of 98% but a lower specificity of 75% when used alone as a diagnosis method. Therefore, this method cannot be relied upon as a sole diagnosis for TB [13, 14]. Scoring systems have been devised which may be helpful in ruling out PTB but have lower specificity [15].

Automated liquid cultures are considered the gold standard for TB diagnosis and first-line drug susceptibility testing. However, their costliness, and requirement of highly skilled staff and well equipped laboratory restrict their use in low resource countries [8]. Alternative culture methods and drug sensitivity tests such as use of solid culture, microscopically observed drug susceptibility and nitrate reductase assay are more cost-effective in resource poor countries [7, 8]. Although the solid culture takes weeks before the results can be obtained, it is the only available method to test the susceptibility to second-line drugs [10].

New molecular methods of detection such as Nucleic Acid Amplification Tests (NAAT) have been developed that are faster and highly sensitive. Xpert® MTB/RIF (Cepheid,USA) and line probe assays (LPAs) have been approved by the World Health Organisation (WHO) to be used to detect especially the first-line drug sensitivity and can give results in a few hours [7, 8]. Xpert® MTB/RIF in combination with microscopy has a sensitivity of 68% and a specificity of 99% when compared with culture techniques. WHO recommends to use Xpert® MTB/RIF as the initial diagnostic test in cases that are suspected to have MDR-TB or HIV-TB [14]. However, in resource poor countries these methods are not readily available and it is estimated

that only a quarter of the MDR cases are diagnosed and only 50% of them are successfully treated [10].

Besides this, LTBI can be detected by tuberculin skin tests although it can't distinguish between Mtb infection and other mycobacterial exposures such as BCG. More recent method includes the IFNγ release assays (IGRAs) [8]. Tuberculin skin test is as sensitive as IGRAs but is less specific. However, because of low costs it is the primary method used in low income countries for LTBI detection and treatment [7]. Latent infection serves as an enormous reservoir for potential disease and it is important to screen high risk groups such as recent contacts of active TB patients, foreign individuals born in high prevalence areas and patients with HIV infection or diabetes, so that a timely and proper treatment can be provided [7]. The diagnosis of EPTB is more complicated since the smear microscopy is of little or no use in this case. The microbial and histological examination of specimens collected from suspected sites of infection is recommended. Xpert® MTB/RIF on cerebrospinal fluid has been recommended in cases such as tuberculous meningitis where a rapid diagnosis is essential [14]. A timely TB diagnosis has long been a standing problem. As the concept of personalized medicines is becoming popular in case of infectious diseases, an individual risk assessment with the help of genetic markers such as SNPs would be helpful in determining the susceptibility/resistivity status [16].

1.1.4 TREATMENT AND CARE

The standard treatment regimen for active drug susceptible TB consists of four drugs, namely, rifampicin (H), isoniazid (R), pyrazinamide (Z) and ethambutol (E) and sometimes Streptomycin (S), given during the induction phase of two months. During this period, most of the fast growing bacteria are expected to be eliminated. The induction phase is followed by a consolidation phase when isoniazid and rifampicin are continued for additional four months [17]. The prolonged regimen often leads to patient non-adherence to the treatment due to several factors like drug toxicity, social stigma and the belief that the infection is cured when the cough resolves and there is no more bacteria observed in sputum [17]. Directly Observed Therapy Short course (DOTS) strategy for TB treatment, recommended internationally by WHO, helps to ensure treatment compliance as the whole treatment is carried out under a close,

direct observation where a second person (a health worker, volunteer or a family member) directly observes the patient swallowing the medicines [14].

The patient is considered to be cured if the sputum smear gives negative result at the end of the treatment and at least at one previous occasion [18]. New patients are usually considered to be having the drug susceptible TB except in the cases where the incidence of isoniazid resistance in the population is high or when there is a suspected contact with a patient having MDR-TB. In such cases, the new patients are given ethambutol in addition to other two drugs in the continuation phase [18]. Relapse cases are more likely to be susceptible to one or more first line drugs. In these and in MDR-TB cases, drug sensitivity test (DST) is recommended to be performed for each patient. In many countries with inadequate laboratory equipment, the DST test results are not readily available. Empirical regimens are recommended to be used in these cases at least until the test results are available [17]. For MDR-TB cases, WHO recommends a minimum twenty months of chemotherapy, with a longer intensive phase of about nine months using second line drugs that are more toxic.

EPTB is treated the same way as PTB. Some experts recommend a longer treatment for TB affecting central nervous system, bone or joint [18]. LTBI cases as well as children <5 years of age who are close contacts of patients with active TB, and HIV patients who do not have active TB infection, are recommended to be given isoniazid for six months as a preventive therapy [14].

The transition of medicinal therapy management from programmed to personalized is hindered in case of TB, due to lack of appropriate genetic markers, apart from other factors. SNPs help in mapping a complex genetic trait and their association with susceptibility and drug response can lead us closer to the final goal of providing tailor made therapeutics [16, 19].

1.2 HOST- MTB INTERACTION

TB development is the result of the interactions between host, pathogen as well as environment. The pathogenesis and disease outcome depends on both the bacterium and host characteristics. Some studies also suggest that a particular combination of host and Mtb genotypes are associated with increased risk and disease severity [20].

While some individuals are able to counter the bacteria and eliminate or contain it, some proportion develop the active disease. However, the factors governing the disease development and progression are still not fully understood [8].

1.2.1 MTB VIRULENCE

Mycobacterium tuberculosis is an aerobic, acid-fast, non-motile, non-encapsulated and non-spore forming bacillus. The cell wall has a high lipid content which makes it impermeable to basic dyes as such [8]. The thick cell wall impairs the entry of nutrients leading to slow growth but also protects the bacterium from stress, many antibiotics and degradation [21]. Consequently, its slow replication rate and ability to persist in latent state presents challenges in timely diagnostics and treatment of TB [8].

The mycobacterial cell wall consists of an inner and an outer layer. The inner layer is constituted by mycolic acids, arabinogalactan and peptidoglycan. The outer layer comprises lipoproteins and lipopolysaccharides like mannose-capped lipoarabinomannan (ManLAM), lipomannan (LM), phthiocerol dimycocerosate, trehalose-1,6-dimycolate (TDM), sulfolipids and phosphatidylinositol mannosides (PIMs). Many of these molecules are soluble and serve as pathogen associated molecular patterns (PAMPs) interacting directly with the immune system [22].

The exceptionally successful intracellular survival tendency of Mtb can be attributed to its ability to reprogram the macrophages after primary infection to avoid its elimination, initiating granuloma formation and helping bacteria to stay walled off from the host immune responses while maintaining a dormant state which is extremely resistant to immune attack and drugs [23]. Apart from various proteins and enzymes required for persistence, mycobacterial surface lipids play an important role in its pathogenicity. Loss of surface lipids is associated with loss of virulence [24]. ManLAM and TDM are two very potent immunomodulators and deserve a special mention. ManLAM, characteristically present on only slow growing mycobacteria, can interact with a number of PRRs including mannose receptor (MR), dendritic-cell-specific intracellular adhesion molecule 3-grabbing nonintegrin (DC-SIGN) and CD14. ManLAM has been shown to inhibit phagosome maturation and IL-12 secretion in dendritic cells (DCs) and apoptosis and IFNγ signalling in macrophages [25]. It may

also interact with cells in a receptor independent manner by directly inserting into the cell membrane and in that way has been recently shown to inhibit T-cell signalling [26]. TDM or the cord factor, is the most abundant soluble component of the cell wall of virulent mycobacteria. TDM is implicated in the inhibition of phagosome-lysosome fusion and formation of granulomas [27]. It has been recently shown to interact with the members of a family of PRRs called C-type lectin receptors (CTLRs) such as macrophage inducible C-type lectin (Mincle) and macrophage C-type lectin (MCL) inducing Th1/Th17 immune responses [28, 29].

1.2.2 HOST RESPONSE AGAINST MTB

Mtb and humans are thought to have co-evolved during the early periods of human history, which could be a reason why Mtb has developed an efficient arsenal to tackle most of the human immune system strategies to counter it. Mtb is able to manipulate both the innate and adaptive arms of immunity to its favour. According to the current model of Mtb infection and its control, it is believed that the bacterium is taken up by the macrophages, where it proliferates and the cell either dies enabling the bacteria to infect healthy cells, or the antigen specific $CD4^+$ T cells response promotes the killing of bacterium or limits its growth [30]. The hallmark of TB infection is the development of granulomas which are a caseous mass of immune cells, a dynamic structure thought to be maintained by the adaptive immunity in an attempt to contain and restrict the bacterium from spreading [24]. Activated $CD4^+$ T cells, IL-12, IFNγ and TNFα are required for TB control. However, there are several other host factors which are involved in the decision of whether an active disease will develop or not [20].

The lack of our understanding about the protective immune response in the lung and thereby, the role of CD4+ T cells in this response are major limitations in the quest to develop a better effective vaccine for TB control [30]. Studies have shown that the currently used vaccine Bacillus Calmette-Guérin (BCG), is more effective in preventing disseminated infection than pulmonary infection in humans as well as mouse models [31-33]. The strain also lacks many virulence genes compared with Mtb.

1.2.3 EARLY RESPONSE

Following aerosol transmission, alveolar macrophages (AMs) in the lung alveoli are the first host cell types targeted by mycobacteria. Macrophages serve as the major cellular niche and reservoir for the bacteria to survive and stay transmissible. In a successful attempt of eliminating the phagocytosed bacteria, the macrophage will employ strategies such as acidification of the phagosome after phagosome-lysosome fusion that activates the protein and lipid degrading enzymes, production of reactive nitrogen and oxygen species (RNS and ROS), microbial peptides and hydrolytic enzymes. All these phenomena work together to kill and eliminate the bacteria by suppressing its metabolism as well as manipulating its DNA and cell wall. The cell may also undergo autophagy in order to destroy the bacteria from spreading further [23]. Whether the bacterium will be killed or not depends not only on the immune status of the host but is also attributed vastly to the virulence of the Mtb that has caused infection. Mtb is very effective in arresting the phagosome maturation and therefore has high chances of intracellular persistence [34].

Infection of AMs triggers their activation and production of various cytokines such as TNF-α, IL-1β, IL-6 and IL-12. Fig-2 shows an overview of the interaction between Mtb and AMs. Together, these cytokines play roles in the recruitment and activation of macrophages and neutrophils towards the site of infection, stimulating the T helper Type 1 (Th1) response, proliferation of cytotoxic T cells (Tc) and NK cells as well in the formation of granuloma. Th1 response leads to the production of IFNγ which is considered to be very important for protection against intracellular pathogens [35]. *IFNG* deficient mice have been shown to rapidly succumb to TB infection compared to control mice [36]. Also, the IFNγ levels have been found to be higher in latently infected subjects [37]. Th17 response, initiated mainly by γδ T cells in response to IL-6 and characterized by the secretion of IL-17, induces leukocyte recruitment, and has been shown to play some role in early stages of Mtb infection. But they also have been implicated in TB related tissue damage inflicted by excessive neutrophil influx [38].

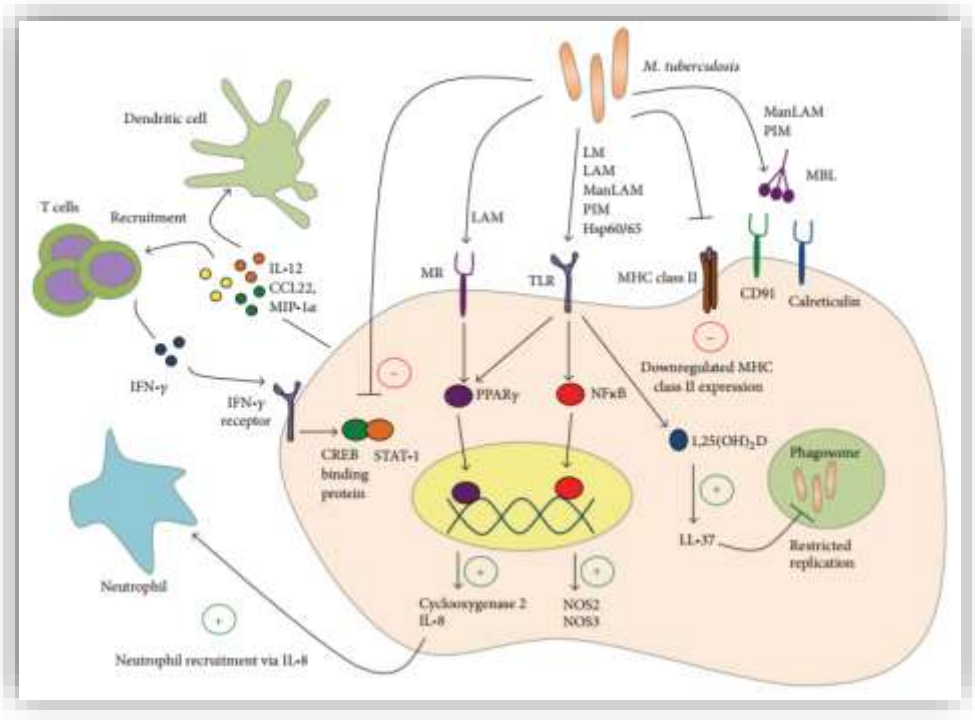


Figure 2: Innate immune response against Mtb as it is encountered by alveolar macrophages. (Source: Sia K.J. et al. 2015 [39])

1.2.4 GRANULOMA FORMATION

Apart from AMs, resident DCs also phagocytose the Mtb and carry them to the nearest lymph nodes to activate the T-cells via MHC class II mediated antigen presentation. Thus, DCs serve as a link between innate and adaptive immunity but may also contribute to dissemination of bacteria outside the primary infection site [30]. Once the T-cells are activated and the adaptive response has been initiated, the cells move towards the site of infection in response to the chemokines generated from infected cells, which is predominated by the presence of macrophages and neutrophils and is termed by some authors as an ‘innate granuloma’ [30, 34]. The recruited CD4+ T-cells surround the uninfected macrophages and neutrophils that are surrounding in turn the infected cells, and wall them off restricting their growth, movement and further spread [40]. Thus an organized dynamic structure, the so called granuloma, is formed. The granuloma has been said to be beneficial for the host in restricting bacterial transmission, resulting in a latent state of infection [41]. This is the classical view of granuloma and latent infection in humans which is, in fact, based mostly on our studies on animal models of TB. However, many authors are

now questioning the credibility of granuloma in host protection [42-44]. There is a growing belief that the mycobacteria may in fact promote the granuloma formation in order to recruit fresh cells for infection and transmission to other sites [45, 46]. Many mycobacterial ligands including TDM, PIMs and ESAT-6 have been shown to promote host granulomatous response [46-51].

1.3 PRRs IN MYCOBACTERIAL IMMUNITY

The antigen presenting cells, including macrophages, have a variety of PRRs expressed on their surfaces which help in recognizing and phagocytosing the bacteria. The receptor activation also triggers specific downstream pathways that lead to responses like cytokine and chemokine secretions. Mtb is recognized by a number of receptors including those from the families of the Toll-like receptors (TLRs) and Nucleotide Oligomerisation domain (NOD) -like receptors (NLRs). Recently several CTLRs have been newly found to recognize and mediate signalling in response to Mtb [29, 47, 52]. Ligation of different PRRs induces distinct effects at different stages of infections and thus contributes to a complex network of co-ordinated interactions between different cells and receptors that ultimately plays role in the outcome of the disease [39, 53].

Among the PRRs, TLRs have been extensively studied and well characterized for their interaction with Mtb and role in modulation of bacterial pathogenicity. They are transmembrane proteins expressed either on cell surface or on endocytic vesicles in immune cells. Out of the ten members known in humans, mainly *TLR2*, *TLR4* and *TLR9* are involved in recognition of Mtb ligands such as PIMs, lipoproteins, LAM, LM, 38-kDa and 19-kDa glycoproteins [54-58]. The signaling through TLRs is mediated by TIR domain containing adaptors such as myeloid differentiation primary response gene 88 (MyD88) and TIR-domain-containing adapter-inducing interferon-β (TRIF), which initiate a cascade of events that leads to NF-κB activation and subsequent inflammatory response [59]. *TLR2* is probably the most investigated amongst TLRs in mycobacterial infections. *TLR2*$^{-/-}$ mice are highly susceptible to H37Rv (a virulent Mtb strain) infection displaying defective granulomatous response, reduced bacterial clearance and finally succumbing to the infection [60]. A similar response was also demonstrated in *TLR4*$^{-/-}$ mice [61]. TLR-2, in association with TLR-1 and TLR-6, can induce signaling in response to specific Mtb ligands [62, 63]. Double knockout mice

for *TLR2* and *TLR9* show enhanced susceptibility to Mtb aerosol infection and display altered lung pathology [64]. Contrastingly, some knock out murine studies suggest a redundant role of TLRs and demonstrate that many Mtb dependent responses do not require specific TLR activation. TLR2 is important for post translational control of TNF-α release but it was not needed for transcriptional induction [65]. No differences in lung pathologies were observed for *TLR2*-, *TLR4*- and *TLR6*-deficient mice on aerosol infection with Mtb. There are also evidences that Mtb recognition may also mediate cellular responses in an MyD88-independent manner [56, 65-67]. *MYD88* deficient mice showed a similar NF-κB activity as wild type on aerial Mtb infection [68]. The differences in the results from knockout mice studies may be attributed to several factors including the strains of mice and infecting bacteria used, genetics etc. but they also suggest that the role of TLRs in Mtb infection is redundant.

NOD-like receptors are multi-domain PRRs located intracellularly and thus implicated in the recognition of intracellular pathogens. There are 23 members of NLR family known in humans, consisting of NLRs and LRPs [69]. NOD-2, also known as CARD15, recognizes the muramyl dipeptide of peptidoglycan layer of bacterial cell wall and its role has also been studied in immunity against Mtb infection [70]. Upon ligand binding, NOD-2 recruits receptor-interacting protein 2 (Rip2/RICK) via CARD-CARD interactions which leads to nuclear translocation of NF-κB via a series of steps [69]. *NOD2* deficient mice had reduced inflammatory response to aerosol infection with Mtb and exhibited higher bacterial burden than their wild type counterparts 6 months post infection [71]. Another member, NLRP3 can recognize the ESAT-6 antigen of Mtb and trigger the secretion of protective IL-1β and lL-18 by activating inflammasome assembly [72, 73]. TDB, the mycobacterial cord factor analogue has been recently shown to activate NLRP3 inflammasome [74]. There have been several studies and also similar contradictions as observed in the case of TLRs in murine studies [21, 75]. Apart from intensifying the complexity of mycobacterial pathogenicity, the discrepancies also open a window of thought that there may be receptors other than TLRs and NLRs which might be playing an additional important role in Mtb recognition and immune responses.

1.3.1 C-TYPE LECTIN RECEPTORS

1.3.1.1 GENERAL STRUCTURE AND SIGNALLING MECHANISMS

CTLRs are transmembrane or soluble receptors, characterized by the presence of one or more carbohydrate recognition domain (CRD) which is responsible for binding to carbohydrate molecules in a Ca^{2+} dependent manner. However, it was later found that not all proteins having a CRD could recognize a carbohydrate, many in fact could bind to lipids, proteins or even inorganic compounds. Thus, a more general term called C-type lectin like domains (CTLD) was introduced [76-78]. The characteristic double loop structure of the CTLD fold is stabilized at its base by two highly conserved disulfide bridges, as well as by polar and hydrophobic interactions. CTLRs can bind a number of carbohydrate moieties by virtue of the second long loop that is structurally and evolutionarily flexible. The CTLDs can be characterized into EPN-motif containing group (mannose sugars binding) and QPD-motif containing (galactose sugars binding) group [78].

Soluble CTLRs, also known as collectins, include receptors such as surfactant proteins and mannose binding lectin (MBL), which are found in body fluids such as serum and mucosal fluids. Collectins are involved in various processes including complement activation, agglutination and opsonization of pathogens aiding in phagocytosis, and orchestrating the adaptive immunity [79]. They usually organize into oligomers which increases their affinity for multivalent ligands present on pathogen cell surface. The monomeric unit consists of an N-terminal cysteine rich domain, a collagen domain, a coiled coil neck domain and a C-terminal domain harboring the CRD [80]. When present on the surface of myeloid cells, CTLRs can help in pathogen recognition and subsequent cellular response.

CTLRs have been divided into 17 groups based on their domain architecture [78]. The transmembrane CTLRs may be classified broadly into three types depending upon the signaling transduction mechanisms employed by these receptors. 1) Receptors signaling via immunoreceptor tyrosine based activation motif (ITAM) - Receptors such as Dectin-1 and CLEC-1B have an ITAM like motif (hemITAM) in their cytoplasmic tails. An ITAM motif is characterized by the presence of YxxI/L (Y = Tyrosine, I = Isoleucine, L = Leucine, x = any residue) repeats, which upon its tyrosine phosphorylation by Src kinases recruits the SH2 domain containing Syk kinase that

can further bind to several other adaptors and coordinate the downstream signaling [81, 82]. Some receptors, such as Dectin-2 and CLEC-5 lack a cytoplasmic tail but signal indirectly via coupling with an ITAM containing adaptor molecules such as Fc Receptor γ chain (FcRγ), DNAX-activating protein of 12KDa (DAP12) and DAP10 [83-85]. 2) Receptors signaling via Immunoreceptor tyrosine based inhibitory motif (ITIM) – An ITIM motif contains I/V/L/SxYxxI/L/V (S = Serine, V = Valine) consensus sequence. Tyrosine phosphorylation recruits the negative regulators such as inositol phosphatase SHIP and tyrosine phosphatases SHP1 and SHP2, which dephosphorylate their substrates and inhibit cellular activation [86]. Receptors such as DCIR and MICL harbor an ITIM motif in their cytoplasmic tail and have been shown to regulate the cell signaling negatively [87, 88]. The previous belief that ITAM containing receptors send activating signals and ITIM containing ones send inhibitory signals has been broken up by several studies suggesting that either could transmit activating or inhibitory signals [86, 89]. 3) The receptors lacking any ITIM or ITAM motif – These receptors include DC-SIGN, Langerin, LOX-1, MR etc. The signaling pathways of such receptors differ and many have not been elucidated yet [81].

The CTLR signaling in myeloid cells is mainly mediated by Syk recruitment. Syk activation can influence several pathways including MAPK, NFAT and, of interest, NF-κB, producing pro-inflammatory signals. Syk activates the various isoforms of protein kinase C (PKC), such as PKCδ [90] which can further activate CARD9 that forms a trimeric complex with B cell lymphoma 10 (Bcl10) and mucosa-associated lymphoid tissue lymphoma translocation protein 1 (MALT1). The CARD9-Bcl10-MALT1 complex can then induce NF-κB mediated pro-inflammatory signaling [81].

1.3.1.2 CTLRs INTERACTION WITH MYCOBACTERIUM TUBERCULOSIS

Mtb membrane is composed of a number of different carbohydrate moieties such as PIMs, TDM, LM, LAM, ManLAM and many others that make ideal ligands for recognition by carbohydrate binding receptors such as CTLRs [82]. There has been recently a great interest in Mtb recognition by CTLRs as well as the related pathways, and their subsequent role in disease development as many more CTLRs have joined the list since the last decade. The interactions of the MTB with the different CTLRs are summarized in Fig-3. Since CLTRs are either soluble or differently expressed on

various cell types these expression patterns differentially impact the particular immune function in response to MTB.

Among these receptors, surfactant proteins, MBL, mannose receptor and complement receptor-3 (CR-3) perhaps are the earliest known Mtb recognizing CTLRs [91, 92].Surfactant proteins SP-A and SP-D, are found in lung alveoli and participate in innate immune responses such as bacterial agglutination and opsonisation that leads to enhanced phagocytosis [79, 93]. SP-A has been shown to upregulate the expression of surface MR on human macrophages, which may further aid in bacterial engulfment [94]. In pulmonary TB patients, lower SP-A levels have been observed in broncho-alveolar lavage while an increase was documented in serum levels, which could be due to the destruction of air-blood barrier in affected lung segments [95, 96]. Serum MBL enhances the complement mediated phagocytosis and is known to act as a first line defence against a number of pathogens. MBL can recognize Mtb ManLAM [97]. Mtb binding to MBL activates MASPs, which catalyse the cleavage of downstream complement proteins to activate the complement system [98]. MR is expressed on the surface of macrophages, including AMs and can also recognize ManLAM and PIMs [82]. While, on one hand, MR has been shown to assist in Mtb phagocytosis, on the other there are indications that MR-ManLAM engagement may inhibit phagocyte maturation as well as induce anti-inflammatory cytokines production such as IL-10, IL-1R, and suppress IL-12 production [99-103]. DC-SIGN is another mannose binding receptor expressed on the surface of many immune cells. ManLAM binding induces the phosphorylation of NF-κB p56 subunit via Raf-1 and subsequent IL-12, IL-10 and IL-6 secretion [81]. DC-SIGN enhances Mtb uptake by DCs [104]. However, the DC-SIGN stimulation has been shown to inhibit the Dectin-1 induced IL-17 responses suggesting how Mtb may be able to modulate the responses through different receptors [105].

In the natural killer cell (NKC) complex located in human chromosome 12, two clusters of genes are found, called Dectin-1 and Dectin-2 clusters. The Dectin-1 cluster (220 kb) comprises *CLEC12A, CLEC1B, CLEC12B, CLEC9A, CLEC1A, CLEC7A* and *OLR1*, and the Dectin-2 cluster (810 kb) consists of *BDCA2, DCIR, CLEC6A, CLEC4D* (MCL) and *CLEC4E* (Mincle) (Fig-3) [94, 95].

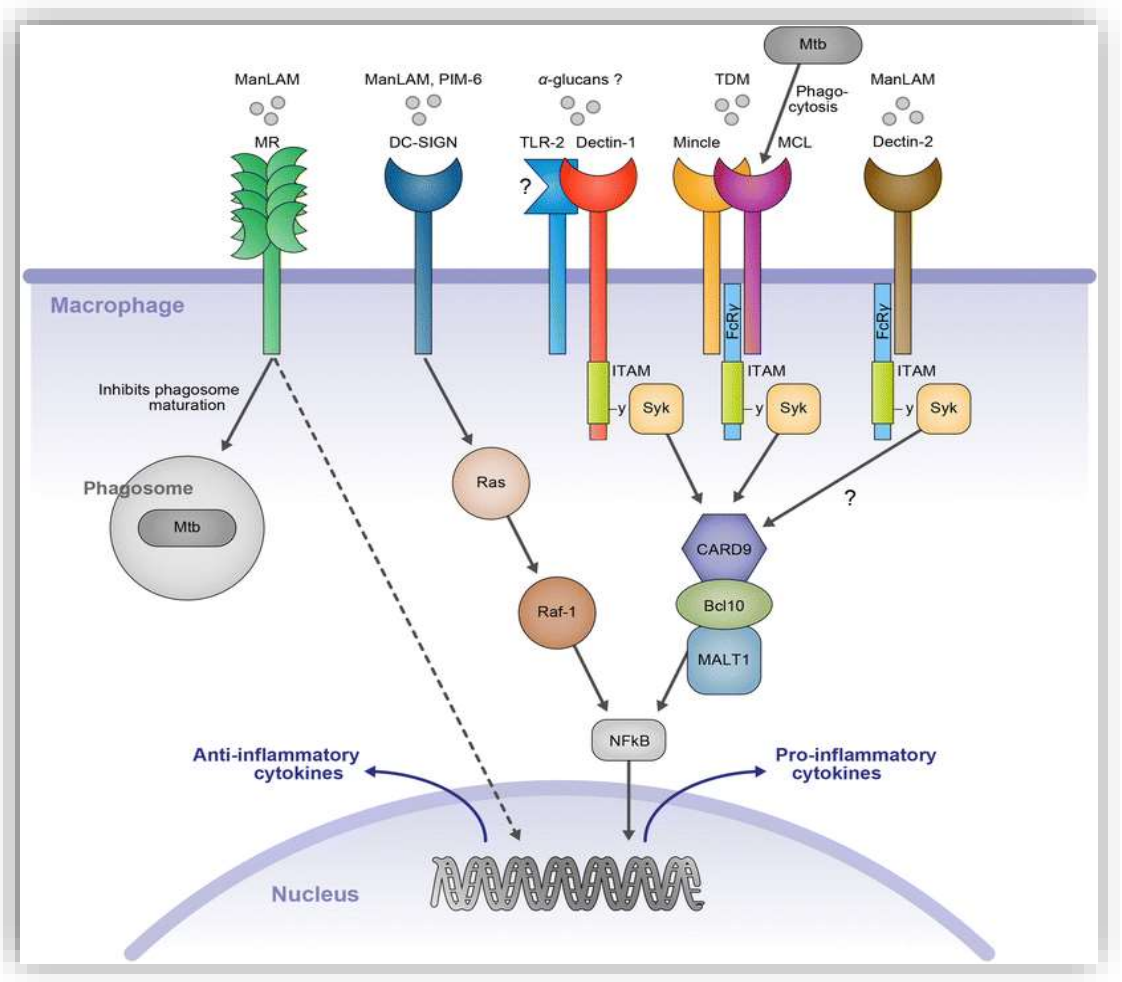


Figure 3: Membrane bound C-type lectin receptors recognizing *M. tuberculosis* in humans and their respective signalling pathways. (Source: Goyal, S. et al. [82])

Mincle and MCL can recognize TDM and mediate Th1 and Th17 responses in mice as well as human APCs inducing pro-inflammatory cytokines such as IL-6, IL-8, G-CSF via the Syk mediated pathway. [28, 106]. In mice bone marrow derived DCs, a TDM analogue was also able to induce NLRP3 dependent IL-1β secretion [74]. Mincle transduces signals via an ITAM containing adaptor called FcRγ whereas it is not clear whether MCL can associate with an ITAM containing adaptor or not. Both MCL and Mincle associated responses are mediated by CARD9-Bcl10-MALT1 pathway but there are speculations as to whether they form heteromers or coordinate each other's responses individually. Whereas some authors claim association of the two receptors, others have demonstrated completely opposite findings [29, 107-109]. *CLEC4D* (MCL) deficient mice show increased mycobacterial burden in the lungs associated with their decreased phagocytosis [110].

Dectin-2 is a recent new candidate in Mtb recognizing CLTRs. It can directly bind another virulent factor of Mtb, ManLAM and induce IL-6, IL-10, TNF and MPI-2 production in mice BMDCs. Mice lacking *CELC6A* or Dectin-2 also show abrogated IL-17 responses [52]. Besides this, studies with Dectin-1 on human cells have shown that the receptor works in collaboration with TLRs such as TLR-2 to induce anti-mycobacterial responses such as ROS production and enhanced bacterial internalization [111, 112]. Mtb induced IL-17A responses may also be in part induced by Dectin-1 and TRL-4 stimulation [113].

Our knowledge about the link between CTLRs and Mtb is still limited and these receptors serve as a desirable target to understand the pathology of one of the most complicated infectious diseases.

1.4 THE IMPACT OF SNPs IN HOST RESPONSE TO TB

One of the most pressing evidences for the role of host innate immunity in TB infection comes from the effect of individual mutations on the pathogenicity and development of the disease. Genome Wide Association studies (GWAS) have been performed in various populations and ethnic groups which demonstrate the role of ancestry and host genetics in susceptibility/resistivity towards TB [114-116]. TLRs have been widely studied for their mutations affecting the susceptibility/resistivity towards TB [117, 118]. Studies among different populations regarding the SNPs in CTLRs such as *DCSIGN, MBL, SFTPA1*, SFTPA2 and *SFTPD* which have been long known to bind Mtb, have demonstrated how ethnicity affects the individual susceptibility. *SFTPA1* and *SFTPD* mutations were studied in certain populations and different population dependent effects were observed, e.g. the A allele of *SFTPA1* SNP rs1136451 was shown to provide protection to Mexicans while it was a risk to Ethiopians [119, 120]. Furthermore, another allele 1649G of *SFTPA2* that disturbs the triple helix structure SP-A, was significantly correlated to TB in Indian, Mexican and Chinese populations [121, 122]. *DCSIGN* promotor mutation -336GG, which is associated with lower DC-SIGN expression, has been studied in several populations and was shown to be a risk factor among an Asian population in a meta-analysis of 14 studies [123].

Structural variants of MBL that influence hugely the serum levels of this protein (by virtue of mutations in promotor and Exon1 regions) also affect the pathogenicity patterns among different ethnic groups [124-126]. A meta-analysis of about 17 studies concluded no significant association between TB and common *MBL* mutations. The study also highlighted the limitations of the genetic studies such as lack of association of SNPs with protein expression or serum levels in this case as well as large heterogeneity among the studied populations [127]. Besides MBL, other components of the lectin complement pathway such as Ficolins and MBL-2 were also addressed in recent studies and showed contrasting results [128, 129]. In case of MR, not many genetic studies have been performed for search of associated SNPs. A non-synonymous *MRC1* SNP rs34039386, is the only SNP associated with TB among Chinese populations [130, 131]. Mincle polymorphisms have been only addressed in a single study, where no association was found between TB and a group of four tagging SNPs studied [132]. A comprehensive overview of the known SNPs in CTLRs associated with TB has been reviewed in Goyal *et al.* [82].

The sequencing methods have evolved enormously in last decades and now high throughput next generation sequencing (NGS) methods offer more rapid, cost-effective and reliable alternative to the classical Sanger sequencing methods or electrophoresis based genotyping/fingerprinting methods for large scale DNA-sequencing. Methods such as barcode-tagging of the amplicons with sample-specific adaptors, has allowed multiplex PCRs and sequencing of multiple samples with high sequencing depth. Moreover, genotyping by sequencing requires low amount of input DNA, provides better coverage and quality [133]. NGS serves as a good method for the screening of a targeted region for potential SNPs in large population samples [134, 135]. For the purpose of exploring the candidate SNPs in targeted sequences, we used Ion torrent NGS which uses semiconductor technology to detect nucleotide incorporation as a proton is released in the chemical process. The other method used in our study for validation of candidate SNPs is called Sequenom MassARRAY platform which uses mass spectrometry to deduce amplicon sizes and provides highly accurate results [136, 137].

To our knowledge, no study has yet addressed the whole CTLR genes for SNP analysis in association with TB. In addition most newly found CTLRs interacting with Mtb have not yet been investigated for their SNPs in TB infected patients.

Furthermore, the genetic studies addressing the SNPs in CTLRs related signalling molecules and their association with TB are also limited. In this study, we used a biphasic sequencing approach to investigate 33 genes, including entire genes of important CTLRs, as well as tagged SNPs and known SNPs in the signalling molecules of various CTLR mediated pathways, for the presence of potential polymorphisms associated with pulmonary tuberculosis in a well phenotyped cohort from the Hyderabad region of Southern India.

1.5 AIM OF THE STUDY

This study mainly aimed at investigating the presence of genetic variations among TB associated C-type lectin receptors in an Indian population. We hypothesized that the SNPs in the CTLR genes and the associated signalling molecules might affect the susceptibility/resistivity status towards developing active pulmonary TB infection. The present study also aimed at investigating the potential functional relevance of significantly associated SNP(s) in developing TB.

CHAPTER 2: MATERIALS AND METHODS

2.1 OBSERVATIONAL STUDY

2.1.1 COHORT RECRUITMENT

The cohort recruitment was performed at Bhagwan Mahavir Medical Research Center (BMMRC), Hyderabad, Telangana, India in collaboration with University of Hyderabad, Hyderabad, India, between July 2011 until November 2013.

The TB cases were recruited through the public private mix DOTS program being executed at BMMRC since 1995. In March 2011, a pilot study was performed to ascertain the feasibility of the required field work and accompanying tests.

2.1.2 TB CASES AND CONTROLS

The study was approved by the Institutional Ethics Committee for Bio-medical Research of BMMRC, Hyderabad India (Appendix II). All the recruited participants gave a written informed consent for their participation in the study (Appendix III). Newly diagnosed cases were recruited based on the following diagnostic criteria according to the Technical and Operational Guidelines for Tuberculosis Control [138] - First sputum smear positive for the presence of AFB; or Chest radiograph suggestive of TB and a second smear test positive for the presence of AFB; or a Chest radiograph suggestive of TB and culture positive for the presence of Mtb. All the relapse cases i.e. the cases with a history of active TB and subsequent treatment, and physically feeble cases were excluded. All the cases were screened for HIV and all the HIV positive cases were excluded from the study. All the cases were treatment naïve and belonged to the Hyderabad region of India. All TB patients were followed up until the end of the anti-TB treatment i.e. six months. Loss to follow-up (mainly due to unwillingness to continue participating in the study), death and relapse were considered end-points for the follow up.The healthy controls were recruited from the same geographical area of Hyderabad and were physically and clinically healthy at the time of recruitment. All controls were unrelated to the recruited TB cases. Subjects with any history of anti-TB treatment, presence of any acute or chronic diseases, heart diseases or chronic diseases of any other nature as well as pregnant subjects were not included.

2.1.3 SAMPLE AND DATA COLLECTION

All the recruited participants were interviewed personally to collect the demographic information as mentioned in the standard DOTS program questionnaire. 10 ml blood samples from the cases at the start of the treatment i.e. the day zero and from the controls on the day of recruitment were extracted by a trained medical nurse. All the blood samples from cases and controls were collected after overnight fasting. The serum was isolated by centrifuging the clotted blood and stored at -80 °C until further use. DNA was also extracted from blood sample using the commercially available kit and stored at -80 °C. In brief, the blood sample is first lysed using a lysis buffer and Proteinase K enzyme and the lysate is passed through a spin column for selective DNA adsorption on the column membrane. The bound DNA is washed twice and finally eluted out of the membrane using elution buffer. In addition, other clinical data and relevant information were collected from the hospital staff and volunteers working under the DOTS program.

2.1.4 DATA VARIABLES

Each subject was interviewed individually to collect the information mentioned in the standard questionnaire available at DOTS clinic (Appendix IV), which includes following variables: age, sex, BMI, fasting blood glucose, diastolic and systolic blood pressure, history of any chronic disease, smoking and liquor consumption, socio-economic factors that are risk for TB infection like monthly income, number of family members and earning members, sputum microscopy results, prior BCG vaccination (documented by the presence or absence of BCG scar) and tuberculin skin test (TST) results. Data information like HIV infection status, serum creatinine levels, lipid profile including the levels of triglycerides, low density & high density lipoproteins, cholesterol, complete blood picture including red and white blood cell count, haemoglobin and platelet count and chest radiographs was obtained from the hospital staff in the pathology laboratory and the records from the hospital TB clinic.

The serum samples were used to measure the levels of important cytokines like IFNγ and TNFα. In addition, the levels of acute phase reactant C-reactive protein (CRP), terminal complement complex (TCC) C5b9, MBL associated serine protease -1 (MASP1), MASP-3 and mannose-binding lectin-associated protein of 44 kDa

(MAp44) in serum were also investigated. The DNA samples were subjected to different sequencing techniques including Next Generation Sequencing, High Resolution Melting and Sequenom MassARRAY, as explained in detail later.

2.1.5 STUDY SIZE

After an initial pilot study performed at the Mahavir Hospital, a target of recruiting 200 cases and controls was set for two years. A total of 144 new pulmonary TB cases and 181 healthy controls were recruited during the whole study.

Materials and Equipment used (Observational Study)

1. Measurement of fasting blood glucose - Accu-Chek Performa, Roche Diagnostics GmbH, Germany
2. Measurement of the blood pressure - MX2 digital automatic blood pressure monitor, Omron GmbH, Germany
3. Measurement of lipid profile (total cholesterol, LDL cholesterol, triglycerides and HDL cholesterol) and serum creatinine - Konelab 20 clinical chemistry analyser, Thermo scientific, USA
4. Measurement of complete peripheral blood count - Sysmex KX-21 blood analyser, Sysmex, Japan
5. Chest radiography (CXR) analysis - Siemens Heliophos D X-rays generator, Siemens, India
6. Measurement of serum TNFα - Human TNF-(α) ELISA kit, BD biosciences, Germany
7. Measurement of serum IFNγ - Human IFN-(γ) ELISA kit, BD biosciences, Germany
8. Isolation of peripheral blood genomic DNA - QIAamp DNA Blood Mini Kit, Qiagen, Germany
9. NanoDrop 2000 – Thermo Fisher Scientific, USA
10. Measurement of DNA quality and quantity - UV-Vis Spectrophotometer, Thermo Scientific, USA

2.2 DETECTION OF POLYMORPHISMS

The next step was to screen for genetic variations in specific genes (CTLRs) which might be potentially associated with PTB-susceptibility in the given population. This part was divided into two phases- i) Discovery stage (Phase I). In this step, a subpopulation of 80 individuals (40 cases and 40 controls) was sequenced using Ion Torrent Next Generation Sequencing to identify potential SNP-candidates in targeted CTLRs and related genes, ii) Validation stage (Phase II). The rest of the population was analysed for the candidate SNPs obtained in the pilot phase. In addition, we sequenced the significantly associated SNPs in a German cohort of healthy individuals to compare the distribution of their alleles and genotypes among two different populations.

2.2.1 DYSCOVERY (PHASE I)

The explorative screening of potentially associated SNPs was performed using an AmpliSeq-based approach on an Ion Proton Next Generation Sequencing platform, provided by Prof. Hortense Slevogt's laboratory at ZIK Septomics, University Hospital, Jena, Germany.

2.2.1.1 AMPLISEQ TARGET SLELECTION

All the known C-type lectins were considered for potential SNPs that could be associated with TB. Table-1 shows all the targeted regions and SNPs addressed in this study. The genes and SNPs to be targeted in the AmpliSeq approach were chosen on the following bases-

1. Dectin-1 and Dectin-2 clusters: The two main CTLR-gene clusters in chromosome 12 were the major targets of this study. These clusters include a total of 12 genes, which code for CLEC12A, CLEC1B, CLEC12B, CLEC9A, CLEC1A, CLEC7A, OLR1, BDCA2, DCIR, CLEC6A, CLEC4D and CLEC4E. Whole genes including introns were targeted. In addition, a segment 1000 bp upstream in 5' UTR region of each gene was also included in the sequencing-design, to screen any potential SNPs in regulatory promotor regions. The sequences were extracted from the Ensembl Genome Browser

(http://www.ensembl.org/index.html). The human genome assembly GRCh37/hg19 was used as reference.

2. Tag SNPs of additional CTLR receptor/adaptor genes: The UCSC Genome Browser was used to extract the most informative SNPs from selected CTLR genes potentially associated with TB. The data from Affymetrix Genome-Wide Human SNP Array 6.0 was used as indicated in the Fig-4.

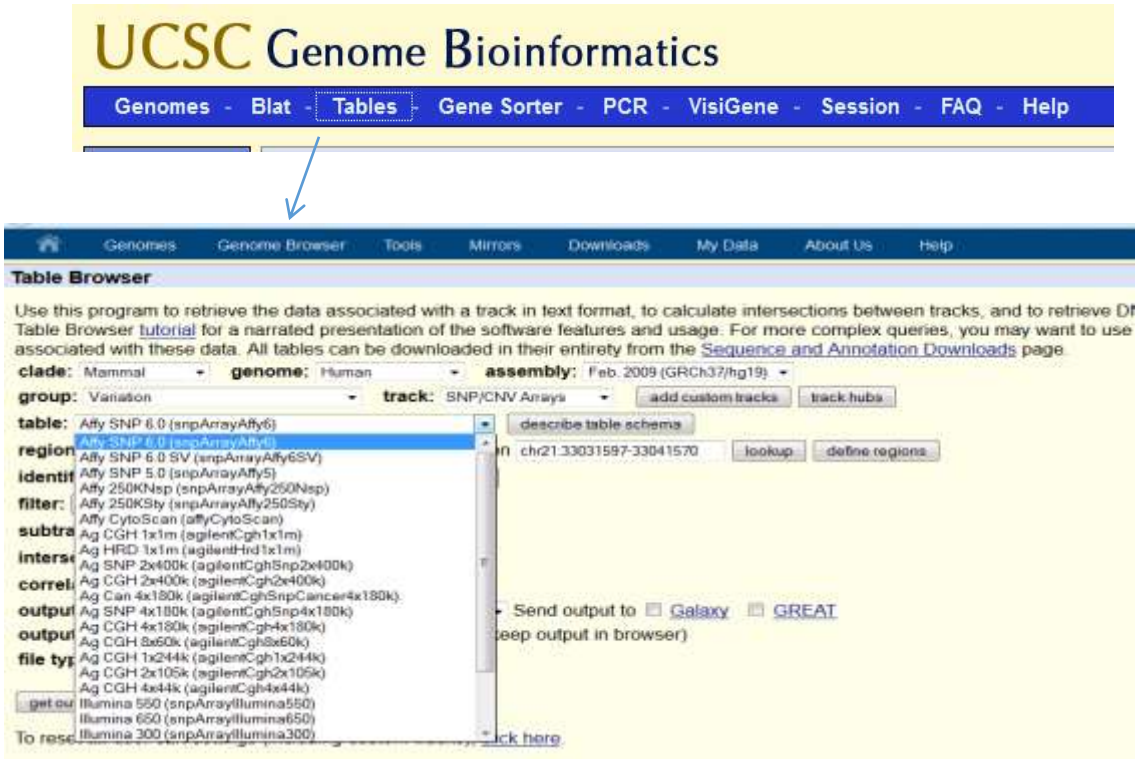


Figure 4: Selection of Tagged SNPs from UCSC Genome Browser

3. Disease associated SNPs: We also considered the known SNPs in CTLRs which have already been shown to be functionally important for the given receptor or have been already associated with development of TB and certain infectious diseases such as lung infections, granulomatous diseases, and fungal infections in different populations. For this, an extensive literature research was performed to find CTLR gene mutations associated with infectious diseases [119, 127, 130, 131, 141-148].

Chromosome #	Gene name (SNP rs#)	Start position	End position
chr1	MASP2 rs12711521	11090916	11090917
chr1	MASP2 1630A>G	11106666	11106667
chr1	Bcl10 rs3768235	85733374	85733375
chr1	Bcl10 rs1060843	85742157	85742158
chr1	Bcl10 rs2735593	85742338	85742339
chr1	Bcl10 rs2735591	85744472	85744473
chr2	Langerin rs57302492	71058230	71058231
chr2	Langerin rs13383830	71058306	71058307
chr2	Langerin rs741326	71058834	71058835
chr2	Langerin rs3815556	71059152	71059153
chr2	Langerin rs10204437	71059762	71059763
chr2	Langerin rs17662453	71061107	71061108
chr2	Langerin rs17719042	71061398	71061399
chr2	Langerin rs17006438	71062052	71062053
chr3	PKCdelta rs2230494	53220215	53220216
chr3	MASP1 rs3774275	186965268	186965269
chr3	MASP1 rs190590338	187011884	187011885
chr7	MDL-1 rs1285933	141627148	141627149
chr7	MDL-1 rs1285935	141627938	141627939
chr7	MDL-1 rs2204608	141632351	141632352
chr9	Syk rs290997	93559351	93559352
chr9	Syk rs2991216	93628027	93628028
chr9	CARD9 rs10870077	139263891	139263892
chr9	CARD9 rs121918338	139264814	139264815
chr9	CARD9 rs4077515	139266496	139266497
chr10	MRC-1 rs691005	17962854	17962855
chr10	MRC-1 rs34039386	18138630	18138631
chr10	MBL2 rs10824793	54529479	54529480
chr10	MBL2 rs1800451	54531226	54531227
chr10	MBL2 rs1800450	54531235	54531236
chr10	MBL2 rs5030737	54531242	54531243
chr10	MBL2 rs7095891	54531460	54531461
chr10	MBL2 rs7095891	54531461	54531462
chr10	MBL2 rs11003123	54531533	54531534
chr10	MBL2 rs7096206	54531684	54531685
chr10	MBL2 rs11003125	54532014	54532015
chr10	SP-A2 rs1965708	81317045	81317046

chr10	SP-A2 rs17886395	81318663	81318664
chr10	SP-A2 rs17880349	81318820	81318821
chr10	SP-A2 rs1250943	81318930	81318931
chr10	SP-A2 rs1650232	81319267	81319268
chr10	SP-A1 rs1059047	81371637	81371638
chr10	SP-A1 rs1136450	81371729	81371730
chr10	SP-A1 rs1914663	81371953	81371954
chr10	SP-A1 rs1136451	81372081	81372082
chr10	SP-A1 rs1059058	81373728	81373729
chr10	SP-A1 rs10351	81373770	81373771
chr10	SP-D rs2181204	81704511	81704512
chr10	SP-D rs6413523	81706134	81706135
chr10	SP-D rs721917	81706323	81706324
chr10	SP-D rs721917	81706324	81706325
chr10	SP-D rs726289	81706950	81706951
chr10	SP-D rs2819096	81707612	81707613
chr10	SP-D rs11200982	81732348	81732349
chr10	SP-D rs11200984	81732365	81732366
chr10	SP-D rs11200985	81732652	81732653
chr10	SP-D rs10788338	81733021	81733022
chr10	SP-D rs4255480	81735887	81735888
chr10	SP-D rs3923564	81735980	81735981
chr10	SP-D rs11201000	81736239	81736240
chr12	PTPN6 rs2301262	7055860	7055861
chr12	CLEC4C	7882011	7905201
chr12	CLEC4A	8275228	8291203
chr12	CLEC6A	8607522	8630926
chr12	CLEC4D	8661071	8674962
chr12	CLEC4E	8685901	8694559
chr12	CLEC12A	10102915	10148293
chr12	CLEC1B	10138241	10167023
chr12	CLEC12B	10162226	10171218
chr12	CLEC9A	10182276	10218565
chr12	CLEC1A	10222153	10265226
chr12	CLEC7A	10269376	10283857
chr12	OLR1	10310902	10325737
chr12	PTPN11 rs2301756	112890776	112890777
chr12	PTPN11 rs3741983	112939853	112939854
chr16	CR3 CD11b rs7193943	31271062	31271063

chr16	CR3 CD11b rs1143679	31276811	31276812
chr16	CR3 CD11b rs9929801	31283471	31283472
chr16	CR3 CD11b rs13338129	31284321	31284322
chr16	CR3 CD11b rs9937837	31298938	31298939
chr16	CR3 CD11b rs9938063	31302937	31302938
chr16	CR3 CD11b rs9888879	31310371	31310372
chr16	CR3 CD11b rs8056264	31332654	31332655
chr16	CR3 CD11b rs11150610	31334235	31334236
chr16	CR3 CD11b rs1143683	31336887	31336888
chr16	CD3 CD11b rs4077810	31340908	31340909
chr16	CD3 CD11b rs7193268	31340996	31340997
chr17	MRC2 rs2465412	60708576	60708577
chr17	MRC2 rs8078112	60720361	60720362
chr17	MRC2 rs8068977	60734815	60734816
chr17	MRC2 rs2302242	60742278	60742279
chr17	MRC2 rs7209331	60746273	60746274
chr17	MRC2 rs4968617	60753903	60753904
chr17	MRC2 rs2465429	60766482	60766483
chr17	MRC2 rs2460290	60768845	60768846
chr19	DC-SIGN rs11465413	7805950	7805951
chr19	DC-SIGN rs11465403	7806590	7806591
chr19	DC-SIGN rs1544767	7806867	7806868
chr19	DC-SIGN rs10403018	7807549	7807550
chr19	DC-SIGN rs4804802	7807609	7807610
chr19	DC-SIGN rs8105572	7809326	7809327
chr19	DC-SIGN rs17159889	7809630	7809631
chr19	DC-SIGN rs2287886	7812535	7812536
chr19	DC-SIGN rs2287886	7812536	7812537
chr19	DC-SIGN rs11465366	7812598	7812599
chr19	DC-SIGN rs4804803	7812733	7812734
chr19	DC-SIGN rs11465362	7813141	7813142
chr19	DC-SIGN rs735239	7813267	7813268
chr19	DC-SIGN rs735239	7813268	7813269
chr19	DC-SIGN rs735240	7813335	7813336
chr19	L-SIGN	7830502	7831170
chr19	L-SIGN rs2277998	7831627	7831628
chr19	L-SIGN rs560634	7831952	7831953
chr19	L-SIGN rs874492	7832000	7832001
chr19	L-SIGN rs558705	7832182	7832183

chr19	L-SIGN rs557094	7832285	7832286
chr19	L-SIGN rs3745376	7833689	7833690
chr19	L-SIGN rs1045998	7833993	7833994
chr19	L-SIGN rs15282	7834273	7834274
chr19	DAP10 rs16960862	36390123	36390124
chr21	CR3 CD18 rs684	46306160	46306161
chr21	CR3 CD18 rs2838726	46315090	46315091
chr21	CR3 CD18 rs3788145	46317125	46317126
chr21	CR3 CD18 rs2235133	46321171	46321172
chr21	CR3 CD18 rs2838732	46322944	46322945
chr21	CR3 CD18 rs760459	46328834	46328835
chr21	CR3 CD18 rs3788147	46329668	46329669
chr21	CR3 CD18 rs2838734	46329739	46329740
chr21	CR3 CD18 rs2280965	46330627	46330628
chr21	CR3 CD18 rs2838735	46335281	46335282
chr21	CR3 CD18 rs2838737	46335579	46335580
chr21	CR3 CD18 rs1474552	46337289	46337290
chr21	CR3 CD18 rs9306118	46338402	46338403
chr21	CR3 CD18 rs9976299	46338650	46338651
chr21	CR3 CD18 rs760453	46340511	46340512
chr21	CR3 CD18 rs2070947	46340842	46340843
chr21	CR3 CD18 rs2070946	46341196	46341197
chr21	CR3 CD18 rs2838738	46344425	46344426

Table 1: Individual SNPs and genes targeted for the discovery stage (Phase-I) via Next Generation Sequencing.

2.2.1.2 DNA SAMPLE SLECETION & QUALITY ASSESSMENT

40 DNA samples from each cases and controls group were chosen to proceed with an AmpliSeq approach using the Ion Torrent semiconductor sequencing technology. In the TB-subpopulation we included individuals with the most severe presentation of pulmonary TB based on the sputum test results and Chest X-ray results while the choice of forty controls was based on the age factor, the oldest assumed to be most resistant to developing TB, having never developed TB before.

Before library construction, the purity and quality of all the eighty DNA samples were determined by spectrophotometry and by PCR of a 339 bp fragment of a randomly selected gene (*CEACAM5*) followed by gel electrophoresis. In brief, 20ng genomic

DNA was used to run a 25µl PCR reaction on a S1000TM Thermal Cycler (BioRad, UK) The amplification mixture consisted of 200µM dNTPs, 0.2 µM of each forward and reverse primers for *CAECAM5* (Fw: CATTTGCAACAGCTACAGTC, Rv: AGTGCAGTGGTATCAGAAAC) and 1 U Taq Polymerase (5-Prime, UK). Thermal conditions included an initial 95 °C denaturation step for 3 min, and then 30 cycles of 30 s at 94 °C, 30 s at 60 °C and 30 s at 72 °C.

2.2.1.3 DNA AMPLISEQ PRIMERS AND LIBRARY PRERPARATION

The primer pairs for the targeted sequences were designed using the Ampliseq Designer version 3.0.1. (https://www.ampliseq.com/browse.action). The final design comprised two pools of primers consisting of 739 and 731 amplicons respectively, of sizes ranging from 125-275bp. The total estimated coverage of the targeted regions was 83%. The total Ampliseq panel size was 253.12 kb.

DNA-AmpliSeq libraries were prepared separately for control and cases samples following manufacturer instructions. In brief, 10 ng of DNA (for each pool) from 80 samples (40 cases & 40 controls) was used as input for the HiFi-amplification with the two designed primer pools. After 17 cycles, the resulting PCR-products were partially digested using FuPa reagent. The reagent does a dual job of removing the primers as well as phosphorylating the ends of the amplicons to facilitate adaptor ligation. The adaptors or barcodes were ligated to the digested amplicons so that individual samples can be recognized after pooling. After purification using Dynabeads (Thermo Fisher Scientific), the libraries were normalized to a concentration of ~100pM using the Ion Library Kit (Thermo Fisher Scientific) and pooled for sequencing.

2.2.1.4 ION TORRENT NEXT GENERATION SEQUENCING

For semiconductor sequencing, the library template pools were clonally amplified on Ion PI™ Ion Sphere™ Particles (ISPs) according to the manufacturer's instructions. The process includes preparation of template-positive ISPs, enrichment of template positive ISPs and quality control. The libraries were finally loaded onto Ion PI Chips, which were sequenced on an Ion Proton Sequencer (Thermo Fisher Scientific, USA).

40 samples from cases and 40 from controls were each multiplexed on two different chips. The raw sequence data will be stored in the Sequence Read Archive (SRA) at National Center for Biotechnology Information (NCBI) in fastq (or bam) format.

Materials and Equipment used (Library Preparation & Sequencing (Phase I))

1. Library preparation - Ion AmpliSeq™ Library Kit 2.0, Thermo Fisher Scientific, USA.
2. Barcodes – Ion Xpress™ Barcode Apators 1-16 Kit, 17-32 Kit and 33-48 Kit, Thermo Fisher Scientific, USA.
3. Library purification - Agencourt AMPure XP Kit, Beckman Coulter, Switzerland.
4. Library Purification - DYNAL™ DynaMag™ Dynabeads™ DynaMag -2 Magnet, Fisher Scientific, USA.
5. Normalization of Library - Ion Library Equalizer™ Kit, Thermo Fisher Scientific, USA.
6. Nuclease-free water, Thermo Fisher Scientific, USA.
7. 1.5 ml Eppendorf LoBind Tubes, Eppendorf, Germany.
8. Chip preparation - Ion PI Hi-Q Chef Kit, Thermo Fisher Scientific, USA.
9. Chip preparation - Ion Chef Instrument, Thermo Fisher Scientific, USA.
10. Sequencing machine - Ion Proton Sequencer, Thermo Fisher Scientific, USA.

2.2.1.5 DATA ANALYSIS

The generated raw data was groomed for further analysis by using various bioinformatics tools. The data generated were in fastq format and their quality was verified using FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). A tool called cutadapt (DOI: http://dx.doi.org/10.14806/ej.17.1.20) was used to remove the barcode/adaptor sequences and the low-quality regions from the raw sequences. The maximum amplicon size was kept up to 240 nucleotides and the Phred score was restricted to >20. The trimmed reads for each sample were aligned onto the human genome hg19 (reference sequence for targeted regions) with Bowtie2 to annotate the known SNPs [149].

Next step was to identify the single nucleotide variations (SNVs) and subsequently the SNPs in the aligned sequence. SNVs were identified using the Variant Caller plugin of the Partek Genomics Suite 6.6 (Partek Inc., St Louis, MO, USA). All the

SNPs with an odds ratio of more than 25 were exported as a text file. The data was subjected to following filters – all the genotype calls with sequencing depth <20 reads/sample were excluded; SNPs with Minor Allele Frequency (MAF) <0.02 were excluded and SNPs that failed to comply with Hardy-Weinberg Equilibrium in control samples were also excluded from further analysis. Any SNPs not falling into the targeted region were also not considered for association studies. Finally, all known biallelic SNPs which passed the above filters, were subjected to association analysis assuming an additive inheritance model, and logistic regression was used for calculating genotype association. Chi-square tests were applied for allele frequencies.

The adjustment for population stratification was done using LASER (Locating Ancestry from SEquence Reads) v. 2.01 software [150]. The software plots each sample on a Principal Component Analysis (PCA)-space defined by reference samples. The reference sequences were extracted from 1000 Genomes Project (ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/release/20130502/). All the entries corresponding to the superpopulation code SAS (South Asian), which included Gujarati (GIH), Punjabi (PJL), Bengali (BEB), Sri Lankan Tamil (STU) and Indian Telugu (ITU) populations were extracted as vcf files. Then the SNVs covered by the amplicon regions corresponding to those targeted by our ampliseq primers were filtered out using VCFtools [151] and the ones with MAF< 0.05 were excluded using the Genome Analysis Tool Kit (GATK) [135, 152]. SNV list was shortened for the most informative SNPs by removing the less informative SNPs in linkage disequilibrium (LD) using PLINK [153]. The resulting SNPs were then used by LASER to define the reference PCA space and the background coordinates for ancestry adjustment. After Genomic Control correction for multiple testing, the success of stratification was determined by inflation estimates (lambda) and quantile-quantile plots.

2.2.2 VALIDATION (PHASE II)

2.2.2.1 SEQUENOM MASSARRAY SEQUENCING

The selected candidate SNPs from the explorative phase were further sequenced for 245 additional samples (141 controls and 104 TB-patients), on Sequenom

MassARRAY iPLEX Platform at Agena Bioscience GmbH (Hamburg). Sequenom MassARRAY provides a two-fold specificity with a locus-specific PCR amplification of target DNA and then a locus-specific primer extension reaction in the presence of mass-modified dideoxynucleotide terminators which anneal directly upstream of the SNP, if complementary. The mass of the extended primer is determined using MALDI-TOF mass spectroscopy. The SNPs included all the candidates which passed the cut off value of p <0. 2 for genotype frequency distribution analysis after adjustment and additionally showed a significant difference ($p < 0.05$) for allele frequency distributions between cases and controls. Statistical analysis of all samples was performed using logistic regressions with PLINK.

2.2.2.2 META-ANALYSIS OF ASSOCIATION RESULTS

METAL was used to combine the results per SNP from association studies in the discovery and validation phases. For this, meta-analysis used the p-values across the two phases taking sample size and direction of effect into account.

2.3 FUNCTIONAL ANALSYIS OF rs3774275

2.3.1 MEASUREMENT OF MASP-1, MASP-3 AND MAp44 IN SERUM

The MASP-1 and its splice variants MASP-3 and MAp44 were measured in stored cases and controls serums using commercially available ELISA kits following the manufacturer's instructions. In brief, the diluted serum samples were incubated in antibody coated 96-well ELISA plates followed by addition of conjugated antibody. The amount of MASP-1 bound conjugated antibody was measured with a spectrophotometer in a coloured reaction at 450nm.

2.3.2 MEASUREMENT OF LECTIN PATHWAY ACTIVITY

Blood serum from healthy donors was isolated by centrifugation at 3000g at 4°C for 10 minutes and serum MASP-1 and MBL levels were measured using ELISAs. To investigate the effect of MASP-1 levels on the lectin pathway complement activity, recombinant human MASP-1 was added (+13%) to the donor serum and complement

activity measured using commercially available ELISA kit following manufacturer's instructions. Briefly, diluted serum was added in triplicates into the ELISA plate wells along with blank, positive and negative controls and incubated at 37°C for one hour. After washing, the formation of terminal complement complex, C5b-9 was detected using conjugate antibody and absorbance was measured at 405nm on a microplate reader. The experiment was performed in triplicates.

Materials and Equipment used (Measurement of serum MASPs and lectin pathway complement activity):

1. Measurement of serum MASP-1 - Human MASP-1 ELISA Kit, Cloud-Clone Corp., USA.
2. Measurement of serum MASP-3 - Human MASP-3 ELISA Kit, Hycult biotech Inc., USA.
3. Measurement of serum MAp44 - Human MAp44 ELISA Kit, Hycult biotech Inc., USA.
4. Recombinant human MASP-1 - MASP1-137H, Creative BioMart, USA.
5. Measurement of lectin pathway activity - Complement system MBL pathway WIESLAB®, Euro Diagnostica AB, Sweden.
6. Spectrophotometer - TECAN SpectraFluor Plus, MTX Lab Systems, LLC, USA.

CHAPTER 3: RESULTS

3.1 POPULATION CHARACTERISTICS

Table-2 shows the parameters that were documented for each patient and control in Hyderabad, India. With the developed cohort, the data was obtained and adjusted with respect to the confounders that may affect their values (mainly age, BMI and gender).

Parameter	Cases (n=145)	Controls (n=181)	P-value
Age (years)	27 ± 11	31 ± 10	0.001
Sex (M/F)	72/73	103/78	0.184
BMI (kg/m^2)	16 ± 2.6	24 ± 4.7	< 0.0001*
Drinking (Yes/No)	27/118	4/176	< 0.0001*
Smoking (Yes/No)	33/112	29/151	0.002*
No. of people living in the house	6 ± 3	5 ± 3	.039
Monthly income (EUR)	106 ± 79.8	173 ± 226	.001
Blood glucose (mg/dL)	97 ± 29	104 ± 33	.054
Systolic BP (mm/Hg)	104 ± 15	123 ± 17	< 0.0001‡
Diastolic BP (mm/Hg)	75 ± 11	80 ± 12	0.855‡
Cholesterol (mg/dL)	167 ± 28	183 ± 34	0.465‡
LDL (mg/dL)	104 ± 26	115 ± 26	0.752‡
HDL (mg/dL)	41 ± 8	44 ± 38	0.272
Triglycerides (mg/dL)	112 ± 39	141 ± 89	0.144‡
Creatinin (mg/dL)	0.91 ± 0.13	0.97 ± 0.17	0.298‡
IFNy (pg/mL)	7.1 ± 18	3.9 ± 12	0.247
TNFα (pg/mL)	13.4 ± 21.8	7.1 ± 22.8	0.026

Table 2: Cohort characteristics and related variables (***** Age and Gender adjusted**, ‡** BMI adjusted) (Statisitcs- Mean,SD, Student T-test ; Ratio and Chi-squared test) (BMI = Body mass index, EUR= Euro, BP = Blood Pressure, LDL, Low density lipids, HDL, High density lipids)

Tobacco smoking is a risk factor for developing active pulmonary tuberculosis due to impaired immunity and non-specific inflammatory response [154]. Furthermore, there is plentiful evidence in literature that alcohol drinking presents as a risk factor for TB incidence [155]. Consistent with this, we observed that the frequency of smokers and alcohol drinkers was significantly higher ($p = .002$ and $p < .0001$ respectively) among the cases compared to healthy controls in our study population after adjustment with confounders such as age and sex.

Besides that, abnormal blood pressure and blood glucose might indicate towards metabolic disorders, influencing inflammation and immune response. Higher BMI has been associated with high systolic blood pressure in several studies [156] while lower BMI has been strongly related to higher risk of developing TB [157]. In line with these studies, the TB patients in our cohort had significantly low mean systolic BP (104 ± 15) and BMI (16 ± 2.6) compared to controls (123 ± 17 and 24 ± 4.7) and were physically weak and frail. During the course of treatment, a steady increase in the BMI of patients was observed along with success of treatment based on sputum test results and chest X-ray (data not shown).

Low BMI may also be associated with the low socio-economic status as observed in cases, indicated by their low monthly income. In India, low socio-economic conditions are associated with malnutrition which is a root cause and risk for several diseases and infections. Furthermore, crowded households, as suggested in our data by a high number of individuals living in a single house, aid in spreading TB rapidly [158], therefore, as the number of people living in a household increase, so does the risk of developing TB.

TNF-α is one of the major cytokines involved in initial and long term control of TB [159]. The mean serum levels of TNF-α were higher in TB patients (13.4 ± 21.8) at the start of treatment compared to the controls (7.1 ± 22.8), while there were no significant difference observed in mean serum IFNγ levels among the two groups.

3.2 EXPLORATIVE PHASE RESULTS

3.2.1 ASSOCIATION ANALYSIS YIELDED MORE THAN 600 POTENTIAL TB-ASSOCIATED SNPs

Our Ampliseq sequencing was focused on dectin-1 and dectin-2 clusters genes on chromosome 12, as well as other CTLR related polymorphisms with an aim of screening out potential variants that may be associated with PTB in our study population. The design covered 83% of targeted region, which included the introns as well as exons of entire genes of dectin-1 and dectin-2 clusters. A total of 19000 SNVs were detected after the first screen using the odds ratio of 25 to get rid of most low abundant reads. After further filtering for sequencing depth (> 20 reads/sample), Minor allele frequency (MAF >0.02) and Hardy-Weinberg equilibrium, we ended up with 634 known SNPs for which association analysis was performed along with stratification adjustments.The allele frequency distribution between the case and control groups was analysed using Chi-squared test. In addition, the genotype associations were also computed assuming an additive model of inheritance.

CHR	SNP	SNP_ID	Genotype (UNADJ-p-value)	Genotype Ancestry-ADJ (p-value)	Allele freq (p-value)
10	10:81371953	rs1914663	0,02501	0,07434	0,00943
12	12:10110431	rs35333643	0,04537	0,1111	0,01697
12	12:10116631	rs3110948	0,02449	0,0733	0,01754
3	3:186965268	rs3774275	0,02122	0,06657	0,02088
12	12:10159846	rs114421141	0,08831	0,1748	0,02427
12	12:10116733	rs2961541	0,05992	0,1341	0,02809
12	12:10156769	rs79967076	0,08831	0,1748	0,03064
12	12:10170823	rs112915340	0,08831	0,1748	0,03064
12	12:10150905	rs374147676	0,03482	0,09288	0,03263
12	12:10103208	rs76427726	0,09018	0,1773	0,03279
12	12:10117091	rs190925857	0,09018	0,1773	0,03279
12	12:10123787	rs193214822	0,09018	0,1773	0,03279
12	12:10112586	rs148864420	0,09018	0,1773	0,03279
12	12:10280862	rs143386125	0,01908	0,06201	0,03348

2	2:71058835	rs741326	0,02093	0,06597	0,03697
12	12:10116629	rs3110949	0,05992	0,1341	0,04830
12	12:10116577	rs648985	0,05992	0,1341	0,04830
2	2:71058226	rs2080390	0,02805	0,0803	0,04965
12	12:10158903	rs17205441	0,08831	0,1748	0,05621
12	12:10170154	rs3736850	0,08831	0,1748	0,05621
12	12:10160808	rs75972360	0,08831	0,1748	0,05643
12	12:10156147	rs7132948	0,08831	0,1748	0,05655
12	12:10151747	rs118077922	0,1799	0,2856	0,06186
12	12:8615602	rs12099687	0,09296	0,181	0,06186
12	12:10151172	rs7980698	0,9454	0,9565	0,06289
12	12:10125497	rs147529241	0,09018	0,1773	0,06305
12	12:10121476	rs182670340	0,09018	0,1773	0,06350
12	12:10122346	rs183462894	0,09018	0,1773	0,06350
12	12:10121072	rs185612837	0,09018	0,1773	0,06359
12	12:8626362	rs4334073/rs199783308	0,06281	0,1385	0,09032
12	12:8290569	rs7302963	0,08688	0,1728	0,09179
12	12:8614238	rs4883148	0,1048	0,1965	0,10000
12	12:10114753	rs145937548	0,09018	0,1773	0,10245
19	19:7812536	rs2287886	0,1131	0,2071	0,11491
12	12:8289877	rs144658267	0,03736	0,0974	0,11596
2	2:71061108	rs17662453	0,07139	0,1511	0,12536
12	12:8617976	rs7302011	0,02766	0,07955	0,12536
12	12:10270773	rs10845048	0,07197	0,152	0,12821
10	10:54531235	rs1800450	0,1195	0,2152	0,13006
12	12:10156051	rs116924172	0,1397	0,2396	0,13086
12	12:7886732	rs77567151	0,08146	0,1654	0,13252
2	2:71058906	rs3213749	0,07546	0,157	0,14037
2	2:71058811	rs3213748	0,07546	0,157	0,14037
2	2:71059153	rs3815556	0,07546	0,157	0,14106
2	2:71058306	rs13383830	0,07546	0,157	0,14106
12	12:8626878	rs4883164	0,06281	0,1385	0,14140
10	10:81706951	rs726289	0,1157	0,2104	0,14493
12	12:8627924	rs11045619	0,0445	0,1096	0,14735
12	12:10311916	rs76039690	0,1038	0,1952	0,15095

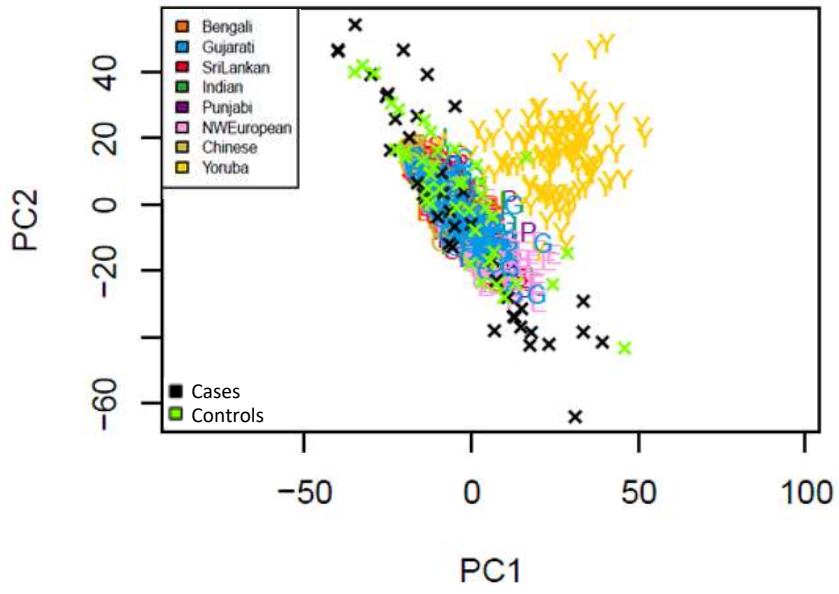
Table 3: Known SNPs and their allele and genotype frequency distributions among the two groups screened out after the exploration phase using ion torrent next generation sequencing. (UNADJ- nominal logistic regression p-values, ADJ- Ancestry adjusted p-values using Genomic control).

Table-3 shows the first 50 SNPs with highest significance with respect to allele (nominal) and genotype (nominal and ancestry adjusted, explained in the next section) frequency distributions (cut off p<0.2 for adjusted genotype frequency distributions). For a full list of the filtered known SNPs and their respective allele and genotype distribution frequencies, refer to Appendix I.

3.2.2 STRATIFICATION ADJUSTMENTS

Assuming a heterogeneous population structure of our cohort, we decided to do a control for ancestry based on the sequence data available online from 1000 Genomes Project. Initial estimates of population stratification showed a high variability in our samples. When plotted against the data from the reference samples from 8 different populations including Bengali, Sri Lankan Tamil, Gujarati, Indian Telugu, Punjabi, North West European, Chinese and Yoruba, our samples indeed showed some heterogeneity and prompted us to perform stratification adjustments (Fig-5).

Figure 5: Principal Component Analysis performed using LASER. Variability in our population samples



was observed when projected onto the PCA space defined by all SNVs (obtained from 1000 Genomes Project) from Bengali, Sri Lankan Tamil, Gujarati, Indian Telugu, Punjabi, North West European, Chinese and Yoruba populations.

The further stratification was performed only using the data from Bengali, Sri Lankan Tamil, Gujarati, Indian Telugu and Punjabi populations from 1000 Genomes Project. We observed a λ_{GC} value of 1.607019 before applying stratification correction and

genomic control tests to our test statistic, which reduced to 1.003906 after adjustments, as assessed by the Q-Q plots in Fig-6.

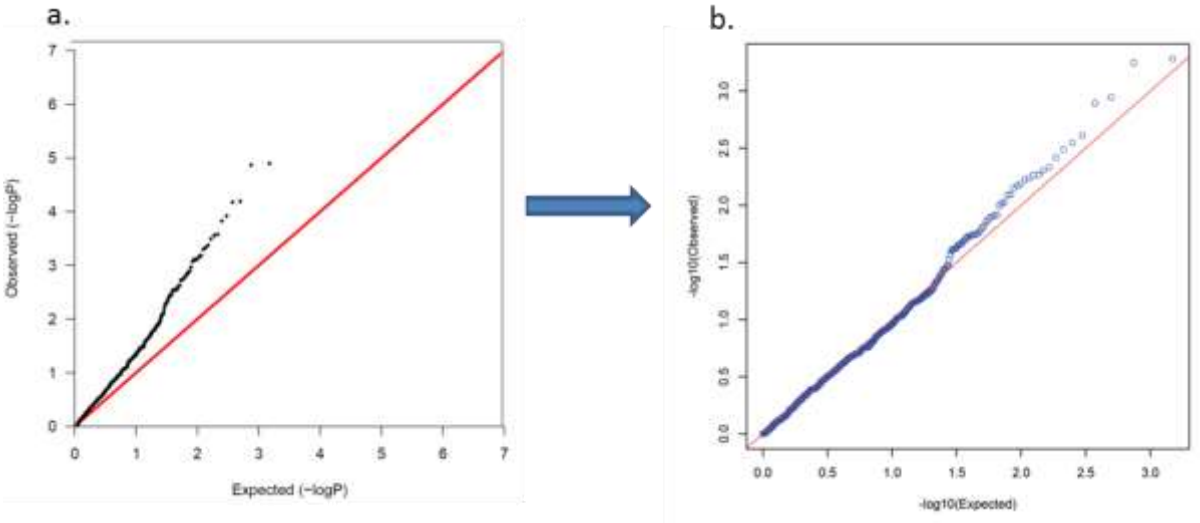


Figure 6: Assessment of ancestry adjustment using Q-Q Plots. (a) The trend of observed and expected p-values for genotype frequency distributions before adjustment and GC control. (b) The trend of observed and expected p-values for genotype frequency distributions after adjustment and GC control.

3.2.3 SCREENING OF SNPs FOR PHASE II

After applying the tests for stratification adjustment and genomic control, a cut off of $p<0.2$ was set for the ancestry adjusted genotype frequency distributions and 18 known SNPs showing lowest p-values for ancestry adjusted genotype frequency distribution were chosen. Table-4 provides details about the gene locations, positions as well as the allele and genotype frequency distributions of selected SNPs. All the selected variants showed significant nominal allele frequency distribution between TB cases and healthy controls.

The candidate SNPs included two missense mutations, each in exons of genes CD207 (Chr.2) and CLEC1B (Chr.12). Another CD207 variant was a synonymous SNP. Many SNPs from the dectin-1 and dectin-2 cluster genes Dectin-1 (CLEC7A), CLEC12A and CLEC12B met the cutoff, all coding for intronic mutations. Apart from these, two intronic variants from surfactant protein-A gene (SFTPA-1) and MASP-1, an adaptor of lectin complement pathway, respectively, also fell below the cut off p-value.

SNP_ID	Location				Type of mutation	Alleles		Association (p-values)		
	Gene	Chr	Position (GRCh38)	Gene location		Major	Minor	Allele freq (nominal)	Genotype freq (nominal)	Genotype freq (adj. + GC)
rs741326	CD207 (Langerin)	2	70831704	Exonic	Missense variant	G	A	0.03697	0.02093	0.06597
rs2080390	CD207 (Langerin)	2	70831095	Exonic	Synonymous	C	T	0.04965	0.02805	0.0803
rs3774275	MASP-1	3	187247480	Intronic	Intron variant	A	G	0.02088	0.02122	0.06657
rs1914663	SFTPA1	10	79612197	Intronic	Intron variant	C	T	0.00943	0.02501	0.07434
rs143386125	CLEC7A	12	10128263	Intronic	Intron variant	C	A	0.03348	0.01908	0.06201
rs76427726	CLEC12A	12	9950609	Intronic	Intron variant	T	C	0.03279	0.09018	0.1773
rs35333643	CLEC12A	12	9957832	Intronic	Intron variant	A	G	0.01697	0.04537	0.1111
rs148864420	CLEC12A	12	9959987	Intronic	Intron variant	C	A	0.03279	0.09018	0.1773
rs648985	CLEC12A	12	9963978	Intronic	Intron variant	G	C	0.04830	0.05992	0.1341
rs3110949	CLEC12A	12	9964030	Intronic	Intron variant	T	C	0.04830	0.05992	0.1341
rs3110948	CLEC12A	12	9964032	Intronic	Intron variant	A	C	0.01754	0.02449	0.0733
rs2961541	CLEC12A	12	9964134	Intronic	Intron variant	T	C	0.02809	0.05992	0.1341
rs190925857	CLEC12A	12	9964492	Intronic	Intron variant	T	C	0.03279	0.09018	0.1773
rs193214822	CLEC12A	12	9971188	Intronic	Intron variant	G	T	0.03279	0.09018	0.1773
rs114421141	CLEC12B	12	10007247	Intronic	Intron variant	T	C	0.02427	0.08831	0.1748
rs79967076	CLEC12B	12	10004170	Intronic	Intron variant	G	A	0.03064	0.08831	0.1748
rs112915340	CLEC12B	12	10018224	Intronic	Intron variant	T	G	0.03064	0.08831	0.1748
rs374147676	CLEC1B	12	9998306	Exonic	Missense variant	C	A/T	0.03263	0.03482	0.09288

Table 4: Detailed list of the 18 candidate SNPs included in phase II after association analysis of the AmpliSeq data. P-values are shown for the differences observed in allele frequency and genotype distribution between TB-patients and healthy controls (nominal and ancestry adjusted).

3.3 VALIDATION PHASE RESULT: ONLY *MASP1* SNP IS SIGNIFICANTLY ASSOCIATED WITH TB SUSCEPTIBILITY

Above mentioned 18 SNPs were sequenced in 245 additional samples (141 controls and 104 TB-patients) via Sequenom MASSArray technique. Four of the selected variants were excluded from the validation due to failed primer design. SNP rs374147676 being monomorphic was also excluded. Further analysis showed that out of the remaining 13 candidate SNPs, only rs3774275 in *MASP1* was significantly associated with pulmonary TB in our study population with p= 0.034 (Table-5).

SNP_ID	Gene	Strand (genotyped)	Test Allele (MA on + strand)	Other Allele	Effect	Association (p-value)
rs741326	CD207	-	A	G	Risk	0.6727
rs2080390	CD207	+	T	C	Risk	0.7578
rs3774275	**MASP1**	**+**	**G**	**A**	**Protective**	**0.0340***
rs1914663	SFTPA1	+	T	C	Protective	0.9470
rs76427726	CLEC12A	+	C	T	Risk	0.0995
rs35333643	CLEC12A	+	G	A	Risk	0.2096
rs148864420	CLEC12A	+	A	C	Protective	0.2497
rs648985	CLEC12A	-	C	G	Protective	0.7202
rs2961541	CLEC12A	+	C	T	Protective	0.9937
rs193214822	CLEC12A	+	T	G	Protective	0.2497
rs114421141	CLEC12B	+	C	T	Protective	0.8407
rs79967076	CLEC12B	+	A	G	Protective	0.6343
rs112915340	CLEC12B	+	G	T	Protective	0.9492

Table 5: List of 13 known SNPs addressed in the validation phase. Only SNP rs3774275 (*MASP1*) was significantly associated with TB susceptibility with G allele providing protective effect. (***** P-value<0.05, MA-Minor allele)

The finding was confirmed when we performed a meta-analysis of the results from explorative and validation phases, showing a very highly significant association (p=0.00281) between the SNP rs3774275 and TB (Table-6). The analysis of other candidate SNPs showed no significance for association with TB.

SNP_ID	Gene	Test Allele (MA on + strand)	Other Allele	Effect	Association (p-value)
rs741326	CD207	A	G	Risk	0.1292
rs2080390	CD207	T	C	Risk	0.1726
rs3774275	**MASP1**	**G**	**A**	**Protective**	**0.002819(**)**
rs1914663	SFTPA1	T	C	Protective	0.2902
rs76427726	CLEC12A	C	T	Risk	0.5581
rs35333643	CLEC12A	G	A	Risk	0.9281
rs148864420	CLEC12A	A	C	Protective	0.8767
rs648985	CLEC12A	C	G	Protective	0.5305
rs2961541	CLEC12A	C	T	Protective	0.3519
rs193214822	CLEC12A	T	G	Protective	0.8767
rs114421141	CLEC12B	C	T	Protective	0.4977
rs79967076	CLEC12B	A	G	Protective	0.6586
rs112915340	CLEC12B	G	T	Protective	0.4278

Table 6: Meta-analysis of Phase I and II results revealed a highly significant association of *MASP1* SNP rs3774275 with pulmonary TB (***** P-value<0.05, MA-Minor allele).

In case of significant SNP rs3774275, the G allele showed a protective effect, being more frequent among controls (39%) than in TB cases (28%). Tables-7a and 7b show the distribution of the major and minor alleles, and the GG, AG and AA genotypes frequencies, respectively among the study population. The GG genotype was twice as much frequent in the healthy group (15%) as in the TB group (7%).

a. RS3774275 allele frequencies (n=321)

	All subjects		Healthy Control		PTB Cases	
Allele	**Count**	**Proportion**	**Count**	**Proportion**	**Count**	**Proportion**
A	421	0.66	219	0.61	202	0.72
G	221	0.34	141	0.39	80	0.28

b. RS3774275 genotype frequencies (n=321)

	All subjects		Healthy Controls		PTB Cases	
Genotype	**Count**	**Proportion**	**Count**	**Proportion**	**Count**	**Proportion**
A/A	137	0.43	66	0.37	71	0.5
A/G	147	0.46	87	0.48	60	0.43
G/G	37	0.12	27	0.15	10	0.07

Table 7: Allele (a) and genotype (b) frequencies among TB cases and healthy controls for MASP-1 SNP rs3774275.

3.4 MASP-1 LEVELS ARE HIGHER IN TB CASES THAN HEALTHY CONTROLS

As the next step to investigate the functional relevance of rs3774275 in TB infection and pathogenesis, we measured the serum concentrations of MASP-1 and its splice variants in our study population. *MASP1* is located on chromosome 3q27-28 and encompasses 18 exons. The primary transcript generated is spliced into three products called MASP-1, MASP-3 and MAp44 by splicing at mutually exclusive splicing (MES) region located between exons 8 to 13 as shown in Fig-7 [145]. The SNP rs3774275 is present in intron 8 and is responsible for regulation of differential serum levels of the three splice variants [145].

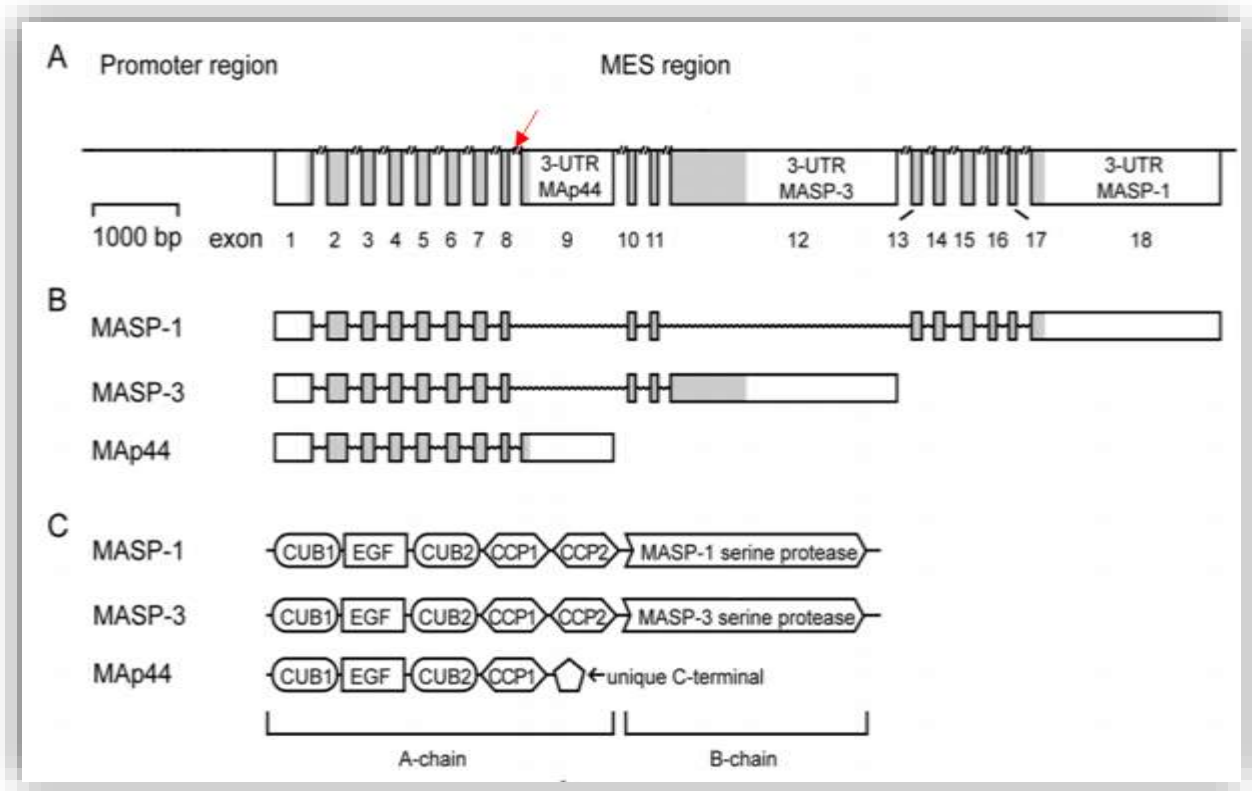


Figure 7: MASP-1 gene and transcripts. The red arrow depicts the position of SNP rs3774275. (Source: Adapted from Ammitzbøll et.al 2013 [145])

A total of 195 samples, consisting of 100 TB cases and 95 healthy controls were investigated for their serum MASPs concentrations. We found a significant difference in MASP-1 and MAp44 serum levels among the two groups. The median concentrations of both MASP1 and MAp44 were significantly higher in TB cases (8.14

µg/ml and 1.66 µg/ml, respectively) than in healthy control group (6.68 µg/ml and 1.29 µg/ml, respectively; Fig-8). However, we did not observe any significant differences in MASP-3 levels between the cases and controls.

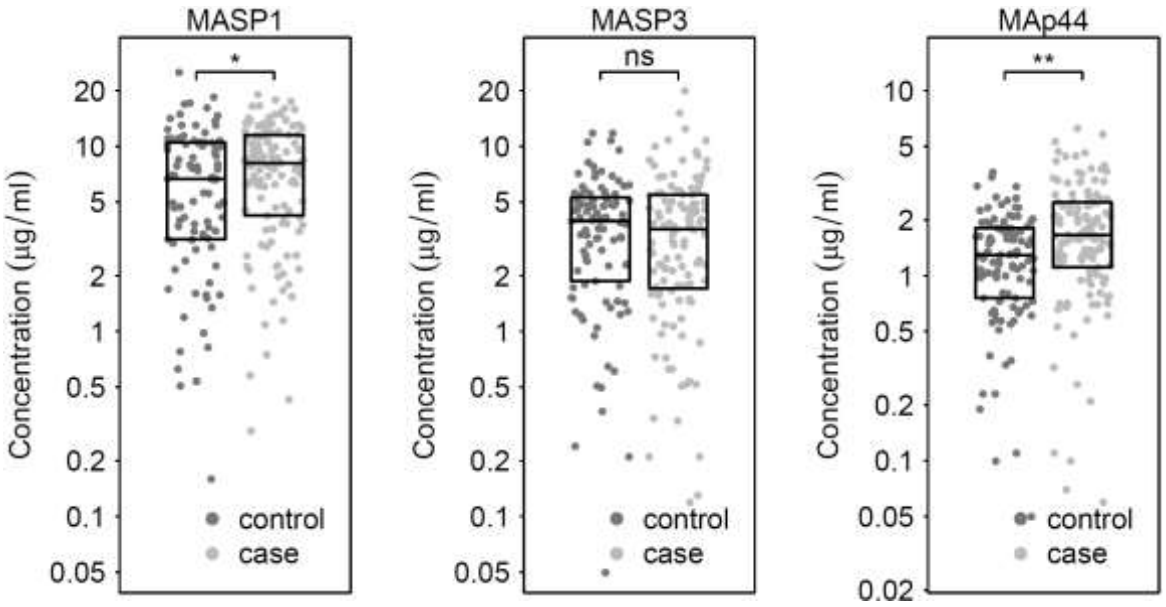


Figure 8: Median serum levels of MASP-1, MASP-3 and MAp44 in TB cases and healthy controls. MASP-1 and MAp44 serum levels were significantly higher among cases when compared to controls (*p<0.05; **p<0.01).

3.5 GENOTYPE DEPENDENT DISTRIBUTION OF MASP-1 SERUM LEVELS

We further analysed the genotype based distribution of MASP-1 concentrations among the studied serum samples. A tendency to higher MASP-1 levels for GG genotype in healthy group (7.74 µg/ml) than in the TB group (6.11 µg/ml) was observed. Similar increase in case of Mapp44 levels for GG genotype in both the groups was noted (Fig-9). A marginal allelic dose-response of G allele could be observed for MAp44 levels, with AA genotype individuals showing lowest while GG individuals having highest, and AG with intermediate protein serum concentrations in both the groups. However, the observed genotype dependent differences in serum levels could not reach a statistical significance. Low number of GG genotype individuals as well as relatively smaller sample sizes could be the reason for lack of significant values.

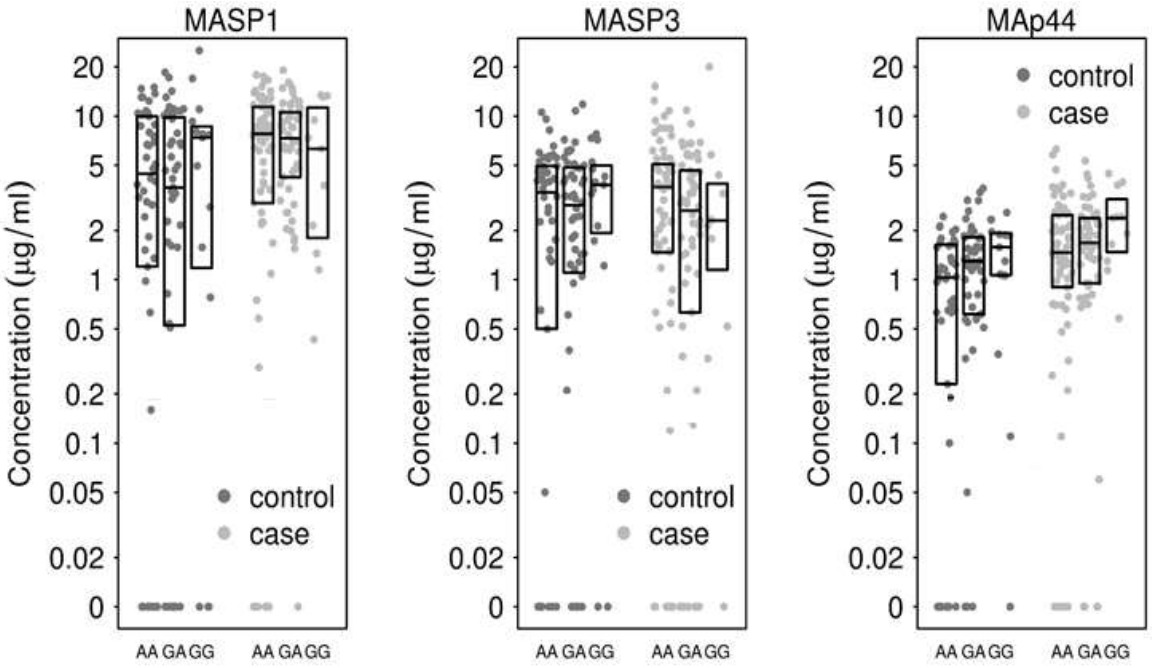


Figure 9: The distribution of serum MASP-1, MASP-3 and MAp44 levels among cases and controls based on AA, AG and GG genotype for the SNP rs3774275.

The mutation rs3774275 has been associated with MASPs and MAp44 levels in other studies as well [145, 160]. GG genotype lead to an increase of upto 13% in MASP-1 and 29% MAp44 serum levels compared to AA genotype in a Danish population [145]. In our study population we observed an increase of 21% serum MASP-1 levels in GG genotype individuals in comparison to AA and AG genotypes.

3.6 ASSOCIATION OF SERUM MASP-1 LEVELS WITH BLOOD GLUCOSE AND BLOOD CHOLESTEROL LEVELS

Next we attempted to correlate the MASP-1 serum concentrations with blood glucose and cholesterol levels as parameters for the assessment of the associated risk of diabetes mellitus or/and metabolic syndrome [161]. A correlation analysis between fasting blood glucose levels and blood cholesterol levels; and MASP-1, MASP-3 and MAp44 serum concentrations was performed (Fig-10a and 10b). The result showed no relationship between MASP-1, MASP-3 and MAp44 concentrations and blood glucose or cholesterol after adjustment with disease status.

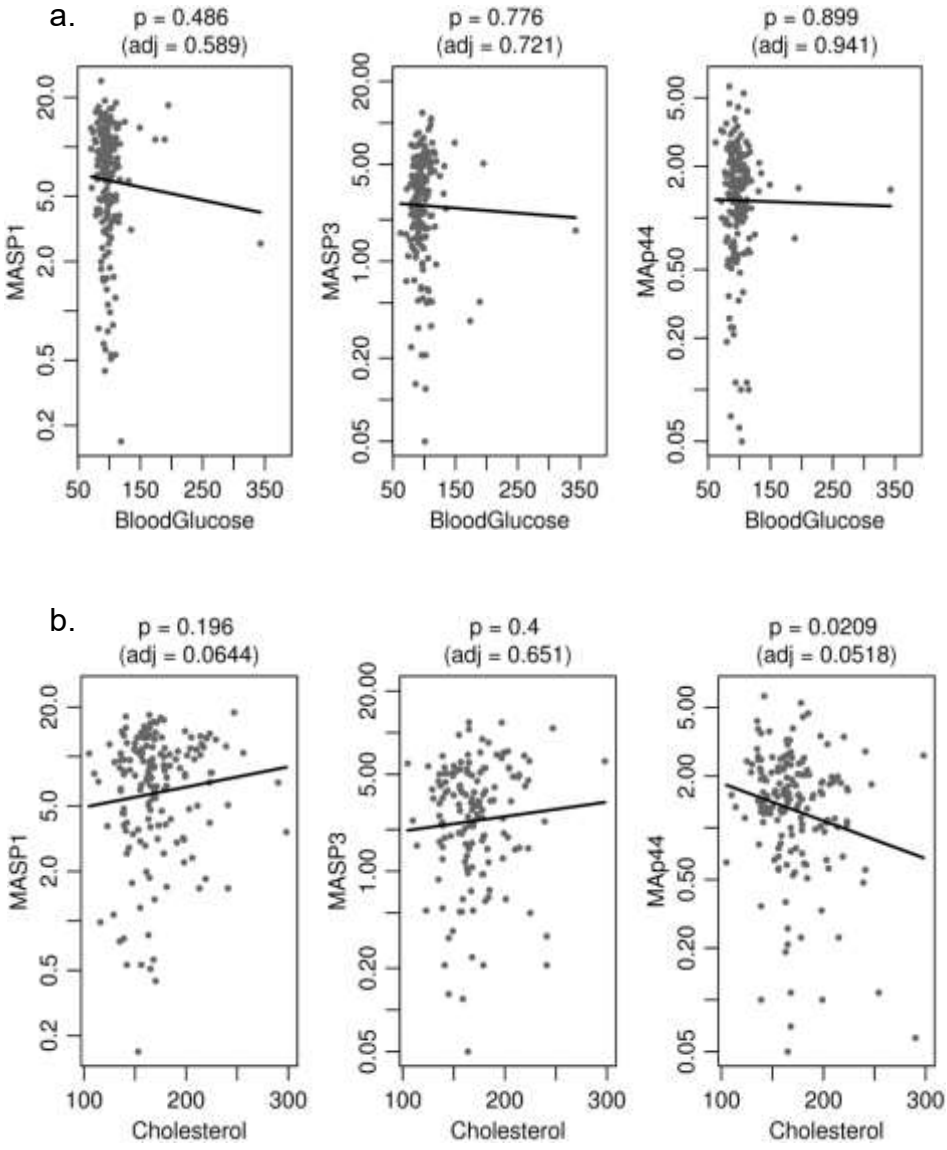


Figure 10: Correlation analysis of MASP-1, MASP-3 and MAp44 serum concentrations with (a) fasting blood glucose and (b) Blood cholesterol (adjusted for disease status). No association was observed between any of the variables involved.

3.7 MASP-1 LEVELS AFFECT THE LECTIN COMPEMENT PATHWAY ACTIVITY

We investigated whether increase in MASP-1 serum levels influences the activity of the lectin pathway of complement activation. We added recombinant human MASP-1 (rhMASP-1) to donor serums and measured the lectin complement activity by ELISA. A 13.22% increase in complement activity (measured by the formation of C5b9, membrane attack complex) on addition of 13% more rhMASP-1 was observed which was statistically significant (Fig-11). The results suggest that the increase in MASP-1 levels correlates linearly with the MBL-pathway of complement which may enhance bacterial clearance by direct killing or opsonophagocytosis.

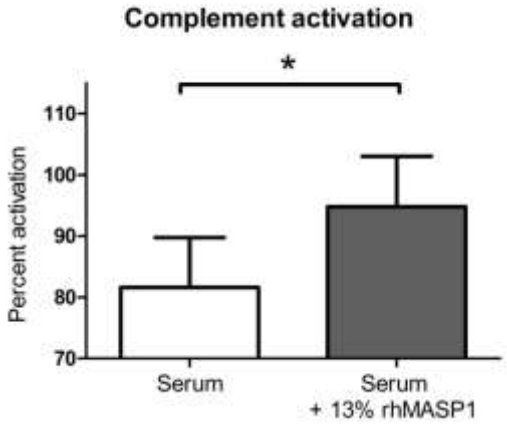


Figure 11: Lectin pathway complement activation as measured by ELISA before and after addition of recombinant human MASP-1 (n=6). A significant increase in activity was observed on addition of 13% rhMASP-1 to the donor serum (*p<0.05).

CHAPTER 4: DISCUSSION

In this study we investigated the presence of potential SNPs in the CTLR genes and their signalling molecules that could affect the susceptibility/resistivity status to develop PTB in an Indian population of Hyderabad region. The main findings of this study are- i) No SNPs in the CTLRs are found to be associated with the development of PTB in the given population, ii) From the investigation of CTLR pathway related molecules, one intronic SNP in *MASP1* ,rs3774275, is significantly associated with PTB risk, iii) The minor allele G is protective and is present twice as frequently in the healthy controls as in PTB patients, iv) MASP-1 and its splice variant MAp44 serum levels are significantly elevated in PTB patients at commencement of chemotherapy when compared to healthy controls, v) addition of recombinant MASP-1 to serum increases the lectin pathway activity *in vitro*. This work suggests a possible new role of lectin pathway and its components in TB infection and pathogenesis.

The cohort was generated in a southern state of India where TB incidence is very high. According to the TB India 2012 annual report, the state of Telangana was the third worst hit state in India with 111,915 registered TB cases in 2011, and Hyderabad city itself accounting for 6,985 cases (16%) in that year (http://www.tbcindia.nic.in/showfile.php?lid=3141). In our population we observed that the cases are more exposed to the risk factors for developing TB than controls, which includes factors such as smoking and alcohol drinking, malnutrition, hyper/hypotension, as well as socio-economic factors such as low incomes and crowded living conditions (Table-2). The observed factors have been intensively studied in low-income, high burden settings [3, 4]. Furthermore, presence of clinical factors such as aberrant blood pressure and blood glucose levels might either falsify or hinder the reproducibility of results among different populations in studies aimed at finding biomarkers for TB. Therefore, these factors must be considered and adjusted for in such studies.

To screen for relevant polymorphisms, we used an AmpliSeq-based discovery approach targeting 33 genes belonging to the C-type lectin receptor family or their related signaling adaptors. In this study, statistical adjustments were done for controlling population stratification, which is an important confounder in most genetic studies. An undetected population structure may lead to both false-negative and false positive results detection in genetic association studies. Unknown population

structure or stratification, especially in admixed populations, is a factor that often contributes to non-reproducibility of case-control studies [162]. Stratification in our Indian population is expected due to historical ethnic, religious and language barriers existing in the community, which might exert important genetic effects and should be addressed in association studies such as this one [163]. Therefore, we used LASER software for population stratification adjustments which helped improve the analysis of our dataset (Fig-5 and Fig-6). Screening for relevant SNPs in 33 CTLR genes in the first phase using NGS yielded several polymorphisms associated with TB in our population. These candidate SNPs were sequenced in 325 samples using Sequenom MassArray in the validation phase. The combined analysis of the data from the two phases revealed one polymorphism (rs37742752) in *MASP1* that could be significantly associated with the disease.

The complement system plays an important role in maintaining immune homeostasis of the body causing inflammation, opsonisation, phagocytosis, direct killing of pathogen and activation of the adaptive arm of immunity [164]. The lectin pathway is one of the three pathways of complement activation whose pattern recognition receptors include MBL and ficolins that can recognize Mtb ManLAM and PIMs [97]. MBL and ficolins are typically found in serum in association with MBL associated serine proteases or MASPs as zymogens. On binding to the ligand, associated MASPs (mainly MASP-1 and MASP-2) are activated and they cleave C2 and C4 plasma proteins which further form the C3 convertase. C3 convertase is in turn responsible for the formation of final terminal complement complex (TCC) that ultimately kills the pathogens while the by-products of the reaction, such as anaphylatoxins C3a and C5a, are recognized by cells of the innate and adaptive immunity and can initiate cellular responses such as opsonophagocytosis [164]. The genetic analysis of our population revealed that the SNP rs3774275 in *MASP1* was significantly associated with pulmonary TB. G allele seems to provide protection against developing TB, being present more frequently among healthy controls compared to TB infected patients. The SNP rs3774275 is present in intron 8 and is responsible for regulation of differential serum levels of the three splice variants [145]. MASP-1 is an important player in the lectin pathway. It is responsible for 60% of C2 cleavage that is needed for C3 convertase formation. Moreover, in normal human sera, it has been demonstrated to be an exclusive activator of MASP-2 which can further cleave C2 and C4 proteins [165, 166]. MASP-3 has been linked to the

activation of alternative pathway of complement by cleaving the pro-Factor D to its active form, which can catalyse the formation of the C3 convertase [167]. MAp44, although lacking serine protease activity, retains the MBL binding domain and therefore acts a competitive inhibitor of MASP-1 and MASP-2 to bind MBL, thus regulating the lectin pathway activity [168]. Among the components of the lectin pathway, SNPs in *MBL* have been widely associated with development of TB in several populations as summarized in Goyal *et al.* [82], however the results have been inconsistent. Nevertheless, a recent study investigated the serum levels and polymorphisms in other components of lectin pathway, including MBL, MASP2, ficolin-2 and ficolin-3, but found no association with resistivity/susceptibility to TB [128], while another study detected a plausible association of MASP-2 SNPs with TB in a Chinese population [129]. MASP-1 is increasingly gaining evidences towards its importance in lectin pathway, but to this end no studies have been performed demonstrating the role of its polymorphisms in TB development. Our findings show for the first time that MASP1 gene polymorphism rs3774275 is associated with susceptibility to TB.

We also sequenced TB related whole genes namely, *CLEC4E*, *CLEC7A*, *CLEC6A*, and *CLEC4D*, along with others from the dectin-1 and dectin-2 clusters, to screen for SNPs potentially associated with disease susceptibility. However, no SNPs in these genes reached a significant level of association after statistical analysis. A *CLEC4D* (MCL) SNP rs4304840 has been associated with TB in a recent work [110], but in our study, the SNP did not even pass the first filter. Another study addressed four tagged SNPs in *CLEC4E* (Mincle) with regards to their role in PTB, although they concluded no association. [132]. In our study, two of the four SNPs addressed by Bowker and colleagues i.e. rs10841847 and rs10841856 passed the first filter but could not reach significance to be included in the validation phase (genotype frequencies p=0.75, p=0.89 respectively, Appendix I). Consistently, we conclude that Mincle SNPs might not be associated with PTB in our population. Similarly, other common variants in *MBL*, surfactant proteins, and *CD209*, which have been associated with TB in several other populations, did not show any association with TB in our study. Besides, several SNPs chosen in validation phase belonged to CLEC12B and CLEC12A, which are both shown to exhibit inhibitory responses. Although none of them were significant after phase II analysis, they might have some implications in TB pathogenesis.

We further wanted to elucidate the functional relevance of the SNP rs3774275 in our population. For this we first measured the serum levels of MASPs and MAp44 among TB cases and healthy controls in our study cohort and observed significantly higher serum MASP-1 levels in TB patients compared to healthy controls. Among MASPs, MASP-2, which is said to be the main effector of the lectin pathway, has been associated with several infectious diseases such as Hepatitis C virus (HCV) infection, *Pseudomonas* infection, leprosy, as well as TB with respect to its polymorphisms and serum/plasma levels [98]. MASP-1 has not yet been studied in relation to many infectious diseases. Some studies on HCV infection have demonstrated a high association between MASP-1 activity and severe hepatic fibrosis [169, 170]. In another study, a synonymous mutation in *MASP1* in the MASP-3 serine protease domain was associated with early *Pseudomonas aeruginosa* colonization in cystic fibrosis patients [171]. Apart from these, MASP-1 polymorphisms are also implicated in developmental disorders associated with 3MC syndrome [172]. In addition, MASP-1 serum levels have been found higher among patients with cardio- or cerebrovascular diseases and higher MASP-1 levels have also been observed in the plasma of patients with rheumatoid arthritis [173, 174]. In a more recent work, Jenny and colleagues found that patients with type 1 diabetes mellitus have higher MASP-1 serum levels [175]. Ours is the first study to demonstrate an association between MASP-1 serum levels and pulmonary TB.

Lectin pathway is also activated during tissue oxidative stress, as has been demonstrated by Collard *et al.* The study showed that during rat myocardial reperfusion, MBL and C3 deposition increased throughout the ischemic area *in vivo* and MBL inhibition attenuated the lectin pathway activity *in vitro* [176]. Oxidative stress is one of the several defence mechanisms by which the activated macrophages attempt to kill Mtb [177]. Since MBL serum levels have been shown to be increased in TB patients [127], it could be hypothesized that MASP-1 expression is also upregulated by similar mechanisms in TB cases.

MASP-1 serum concentrations have also been associated with the SNP rs3774275 in two other studies [145, 160]. The allele G for rs3774275 has been related to higher MASP-1 levels in a dose dependent manner [145]. In our study we also observed higher MASP-1 levels in the GG-genotype of the control population, but failed to reach statistical significance, probably due to the small sample size for this minor genotype.

Nevertheless, an association between rs3774275 and MASP-1 levels could also be observed in an additional study by Krogh *et al.* [160], who investigated the relevance of MASPs in type 2 diabetes. Although they could also find a positive correlation between the G allele of rs3774275 and the MASP-1 levels, they failed to find an association between the SNP and blood glucose levels or diabetes [160].

The relation between TB and diabetes mellitus has been long known. Not only diabetes is a risk factor for developing TB but is also associated with disease severity as well as effects on disease presentation and treatment response [178]. Furthermore, a number of studies have established a strong link between complement and diabetes related complications [179]. Elevated serum MASP-1 levels as observed in our study in TB patients and in the diabetics as shown by Jenny *et al.* and Krogh *et al.* [160, 175], suggest another possible way in which diabetes may affect the pathogenesis of TB. It would of interest to investigate how the serum MASP-1 levels are affected during TB infection in diabetic patients and whether the protease could serve as a potential biomarker in such cases. Our study results, thus, may have important implications for the role of MASP-1 and related pathways in research targeting TBDM comorbidity.

Notably, the mean MASP-1 levels were lower in our healthy group (6.68 µg/ml) compared to the healthy groups in other three studies which included Danish cohorts, where the mean serum MASP-1 levels were much higher in healthy groups (9-11 µg/ml) [145, 160, 180]. In a bigger study involving 1063 healthy Japanese subjects, the mean value of MASP-1 (+MASP-3 + MAp44) serum concentration measured was 6.27 µg/ml [181]. The observed differences could be because of different methodology involved in establishing the MASP-1 assay. But the effects of genetic diversity due to differences in the ethnic backgrounds of the populations in question can’t be excluded.

We investigated whether the MASPs and MAp44 levels correlate with risk factors such as blood glucose/cholesterol levels but found no association between proteases serum concentrations and blood glucose and cholesterol. Our results are in line with a study on patients with cardiovascular diseases where no correlation between MASP plasma levels and blood glucose or cholesterol was observed [173].

It might be possible that during an active infection with Mtb MASP-1 expression is upregulated leading to an enhanced activation of lectin pathway activity. To

investigate whether an increase in MASP-1 concentration could affect the lectin pathway activity, we did an *in vitro* experiment and demonstrated that a 13% increase in MASP-1 levels lead to a significant increase in % activity of lectin pathway. A direct explanation for this result would be that as the amount of MASP-1 in the serum increases, more MASP-2 is activated which could cleave more C2 and C4 molecules, which in turn leads to the formation of more C5b9 complexes that can kill the bacterial cells. The result implicates that even a small increase in the amount of MASP-1 can significantly enhance the lectin pathway activity. However, in vivo experiments would be needed to confirm the effect of higher MASP-1 levels on lectin pathway and the course of TB infection.

Further dissection of the physiological function of MASP-1 in last couple of years has revealed a broader spectrum of its action. MASP-1 is a promiscuous receptor and is shown to bind several ligands. MASP-1 can not only activate the complement lectin pathway but also other cellular processes such as activation of signaling pathways. Megyeri and colleagues demonstrated that MASP-1 could activate the NF-κB, p38-MAPK and Ca^{2+} signaling in endothelial cells *in vitro* by cleaving surface protease activated receptor-4 (PAR-4) [182]. Moreover, the p38-MAPK activation in endothelial cells by rMASP-1 lead to IL-6 and IL-8 secretion along with other cytokines in vitro, which were able to recruit neutrophils [183]. PARs are also expressed on lung epithelium [184], and it could therefore be speculated that high levels of MASP-1 may help induce a similar response in lung tissue, activating cellular responses and recruiting phagocytes which may together enhance clearing of bacteria or contribute to tissue damage in severe cases.

MASP-1 is also known to have a thrombin like activity and its role has been suggested in coagulation [185, 186]. MASP-1 and thrombin share various similarities between there serine protease domains [187]. Both MASP-1 and thrombin bind a number of substrates, which include fibrinogen, prothrombin, Factor XIII, PARs and thrombin-activatable fibrinolysis inhibitor (TAFI) [188, 189]. Another substrate of thrombin is osteopontin [190], a chemoattractant, which has been shown to display some role in granuloma formation as well as in generating Th1 immunity which is essential for anti-TB immunity [191]. Furthermore, osteopontin levels have been shown to elevate in PTB patients and subside with successful chemotherapy [192, 193]. To our knowledge, there has been no link established between MASP-1 and osteopontin,

however, given the similarity in the substrate specificity of thrombin and MASP-1, it may be hypothesized that MASP-1 may also cleave osteopontin and have role in TB infection and pathogenesis in lung. Nevertheless, these speculations remain unanswered until further *in vitro* and *in vivo* studies are performed to dissect the importance of MASP-1 in TB infected lungs.

In the present study, we used a customized AmpliSeq approach to screen for relevant polymorphisms in 33 genes. Although this approach led to the identification of a TB-associated polymorphism in *MASP1* which was strongly supported by the findings in the validation phase, our study suffers one limitation. The sample size used in this study for phase-I may limit the detection of SNPs with low MAF or small effect on complex disease such as the one addressed in this study. Thus, a potential association of additional SNPs in other members of the C-type Lectin Receptor family cannot be excluded by this study. Fruthermore, the control group in our sample population was not screened for latent TB. Studying the effects of MASP polymorphisms and serum levels in this group warrants further investigation.

In conclusion, we investigated the susceptibility to developing TB with respect to host genetic polymorphisms in C-type lectin receptors. We have demonstrated for the first time a contribution of MASP-1 in tuberculosis pathogenesis. MASP-1 had been considered the underdog of the lectin pathway of complement until recently when a more prominent role of this protein has been dissected in lectin pathway. We found a highly significant association between PTB and MASP-1 polymorphism rs3774275 in a well-defined ancestry adjusted population case control study. Additionally, no SNPs in other studied CTLRs were associated with TB in this population. Elevated MASP-1 serum levels in PTB patients suggest the importance of lectin pathway in immunity against Mtb infection. Our results suggest a possible mechanistic relationship between MASP-1 and Mtb lung infection. The complement is activated by MTB by ligation of ManLAM and PIMs to MBL or ficolins, however, the role of lectin pathway in Mtb infection has been considered controversial and needs more attention. Further studies are needed to elucidate the precise mechanism underlining TB-related increase in MASP-1 serum concentrations and its functional effect on the immune cells *in vivo*. Moreover, investigating the MASPs and Map44 serum levels at the end of a successful chemotherapy would also help us gain insight into the role of these proteases in TB pathology. Studies investigating the role of MASP-1 in TB with

regards to its huge substrate repertoire warrants further future work. TB infection and disease progression are not only affected by the bacterium but also host genetics and the introduction of appropriate genetic markers not only help us understand the complexity of this disease but may aid in a better diagnosis as well as treatment strategies in the coming years.

REFERENCES

1. *Global Tuberculosis Report 2016*. 2016, World Health Organisation.
2. Narasimhan, P., et al., *Risk factors for tuberculosis.* Pulm Med, 2013. **2013**: p. 828939.
3. Oxlade, O. and M. Murray, *Tuberculosis and Poverty: Why Are the Poor at Greater Risk in India?* Plos One, 2012. **7**(11).
4. Marais, B.J., et al., *Tuberculosis comorbidity with communicable and non-communicable diseases: integrating health services and control efforts.* Lancet Infect Dis, 2013. **13**(5): p. 436-48.
5. Azad, A.K., W. Sadee, and L.S. Schlesinger, *Innate immune gene polymorphisms in tuberculosis.* Infect Immun, 2012. **80**(10): p. 3343-59.
6. Schluger, N.W. and W.N. Rom, *The host immune response to tuberculosis.* Am J Respir Crit Care Med, 1998. **157**(3 Pt 1): p. 679-91.
7. Zumla, A., et al., *Tuberculosis.* N Engl J Med, 2013. **368**(8): p. 745-55.
8. Lawn, S.D. and A.I. Zumla, *Tuberculosis.* Lancet, 2011. **378**(9785): p. 57-72.
9. *Standard Manual for Laboratory Technicians on Sputum Smear Microscopy*. 2011.
10. Organisation, W.H. *Global Tuberculosis Report 2015*. 2015; Available from: http://www.who.int/tb/publications/global_report/en/.
11. Singhal, R. and V.P. Myneedu, *Microscopy as a diagnostic tool in pulmonary tuberculosis.* International Journal of Mycobacteriology, 2015. **4**(1): p. 1-6.
12. Campbell, I.A. and O. Bah-Sow, *Pulmonary tuberculosis: diagnosis and treatment.* British Medical Journal, 2006. **332**(7551): p. 1194-1197.
13. Organisation, W.H., *Systematic Screening of Active tuberculosis- Principles and recommendations*. 2013.
14. Organisation, W.H., *International Standards for Tuberculosis Care*. 3rd ed. 2014.
15. Pinto, L.M., et al., *Scoring systems using chest radiographic features for the diagnosis of pulmonary tuberculosis in adults: a systematic review.* European Respiratory Journal, 2013. **42**(2): p. 480-494.
16. Mirsaeidi, M., *Personalized medicine approach in mycobacterial disease.* Int J Mycobacteriol, 2012. **1**(2): p. 59-64.
17. Horsburgh, C.R., C.E. Barry, and C. Lange, *Treatment of Tuberculosis.* New England Journal of Medicine, 2015. **373**(22): p. 2149-2160.
18. Organisation, W.H., *Treatment of Tuberculosis- guidelines.* 2009.
19. Salzer, H.J., et al., *Personalized Medicine for Chronic Respiratory Infectious Diseases: Tuberculosis, Nontuberculous Mycobacterial Pulmonary Diseases, and Chronic Pulmonary Aspergillosis.* Respiration, 2016. **92**(4): p. 199-214.
20. O'Garra, A., et al., *The immune response in tuberculosis.* Annu Rev Immunol, 2013. **31**: p. 475-527.
21. Kleinnijenhuis, J., et al., *Innate immune recognition of Mycobacterium tuberculosis.* Clin Dev Immunol, 2011. **2011**: p. 405310.
22. Hett, E.C. and E.J. Rubin, *Bacterial growth and cell division: a mycobacterial perspective.* Microbiol Mol Biol Rev, 2008. **72**(1): p. 126-56, table of contents.
23. Gengenbacher, M. and S.H.E. Kaufmann, *Mycobacterium tuberculosis: success through dormancy.* Fems Microbiology Reviews, 2012. **36**(3): p. 514-532.
24. Pieters, J., *Mycobacterium tuberculosis and the macrophage: maintaining a balance.* Cell Host Microbe, 2008. **3**(6): p. 399-407.
25. Briken, V., et al., *Mycobacterial lipoarabinomannan and related lipoglycans: from biogenesis to modulation of the immune response.* Mol Microbiol, 2004. **53**(2): p. 391-403.
26. Mahon, R.N., et al., *Mycobacterium tuberculosis ManLAM inhibits T-cell-receptor signaling by interference with ZAP-70, Lck and LAT phosphorylation.* Cell Immunol, 2012. **275**(1-2): p. 98-105.

27. Hunter, R.L., et al., *Multiple roles of cord factor in the pathogenesis of primary, secondary, and cavitary tuberculosis, including a revised description of the pathology of secondary disease.* Ann Clin Lab Sci, 2006. **36**(4): p. 371-86.
28. Schoenen, H., et al., *Cutting Edge: Mincle Is Essential for Recognition and Adjuvanticity of the Mycobacterial Cord Factor and its Synthetic Analog Trehalose-Dibehenate.* Journal of Immunology, 2010. **184**(6): p. 2756-2760.
29. Miyake, Y., et al., *C-type Lectin MCL Is an FcR gamma-Coupled Receptor that Mediates the Adjuvanticity of Mycobacterial Cord Factor.* Immunity, 2013. **38**(5): p. 1050-1062.
30. Orme, I.M., R.T. Robinson, and A.M. Cooper, *The balance between protective and pathogenic immune responses in the TB-infected lung.* Nat Immunol, 2015. **16**(1): p. 57-63.
31. Colditz, G.A., et al., *Efficacy of BCG vaccine in the prevention of tuberculosis. Meta-analysis of the published literature.* JAMA, 1994. **271**(9): p. 698-702.
32. Cooper, A.M., et al., *Expression of memory immunity in the lung following re-exposure to Mycobacterium tuberculosis.* Tuber Lung Dis, 1997. **78**(1): p. 67-73.
33. North, R.J., R. LaCourse, and L. Ryan, *Vaccinated mice remain more susceptible to Mycobacterium tuberculosis infection initiated via the respiratory route than via the intravenous route.* Infect Immun, 1999. **67**(4): p. 2010-2.
34. Shaler, C.R., et al., *Within the Enemy's Camp: contribution of the granuloma to the dissemination, persistence and transmission of Mycobacterium tuberculosis.* Front Immunol, 2013. **4**: p. 30.
35. Sakamoto, K., *The pathology of Mycobacterium tuberculosis infection.* Vet Pathol, 2012. **49**(3): p. 423-39.
36. Flynn, J.L., et al., *An essential role for interferon gamma in resistance to Mycobacterium tuberculosis infection.* J Exp Med, 1993. **178**(6): p. 2249-54.
37. Huaman, M.A., G.S. Deepe, Jr., and C.J. Fichtenbaum, *Elevated Circulating Concentrations of Interferon-Gamma in Latent Tuberculosis Infection.* Pathog Immun, 2016. **1**(2): p. 291-303.
38. Jasenosky, L.D., et al., *T cells and adaptive immunity to Mycobacterium tuberculosis in humans.* Immunol Rev, 2015. **264**(1): p. 74-87.
39. Sia, J.K., M. Georgieva, and J. Rengarajan, *Innate Immune Defenses in Human Tuberculosis: An Overview of the Interactions between Mycobacterium tuberculosis and Innate Immune Cells.* Journal of Immunology Research, 2015.
40. Ulrichs, T. and S.H.E. Kaufmann, *Mycobacterial persistence and immunity.* Frontiers in Bioscience, 2002. **7**: p. D458-D469.
41. Ulrichs, T. and S.H.E. Kaufmann, *New insights into the function of granulomas in human tuberculosis.* Journal of Pathology, 2006. **208**(2): p. 261-269.
42. Strauss, E., *MICROBIOLOGY TB Bacteria May Reign Over Cells Intended to Bridle Them.* Science, 2009. **323**(5911): p. 196-196.
43. Rubin, E.J., *The granuloma in tuberculosis--friend or foe?* N Engl J Med, 2009. **360**(23): p. 2471-3.
44. Shaler, C.R., et al., *Within the Enemy's Camp: contribution of the granuloma to the dissemination, persistence and transmission of Mycobacterium tuberculosis.* Frontiers in Immunology, 2013. **4**.
45. Davis, J.M. and L. Ramakrishnan, *The role of the granuloma in expansion and dissemination of early tuberculous infection.* Cell, 2009. **136**(1): p. 37-49.
46. Volkman, H.E., et al., *Tuberculous Granuloma Induction via Interaction of a Bacterial Secreted Protein with Host Epithelium.* Science, 2010. **327**(5964): p. 466-469.
47. Ishikawa, E., et al., *Direct recognition of the mycobacterial glycolipid, trehalose dimycolate, by C-type lectin Mincle.* J Exp Med, 2009. **206**(13): p. 2879-88.
48. Lee, W.B., et al., *Neutrophils Promote Mycobacterial Trehalose Dimycolate-Induced Lung Inflammation via the Mincle Pathway.* Plos Pathogens, 2012. **8**(4).
49. Volkman, H.E., et al., *Tuberculous granuloma induction via interaction of a bacterial secreted protein with host epithelium.* Science, 2010. **327**(5964): p. 466-9.

50. Taylor, J.L., et al., *Role for matrix metalloproteinase 9 in granuloma formation during pulmonary Mycobacterium tuberculosis infection.* Infect Immun, 2006. **74**(11): p. 6135-44.
51. Gilleron, M., et al., *Acylation state of the phosphatidylinositol mannosides from Mycobacterium bovis bacillus Calmette Guerin and ability to induce granuloma and recruit natural killer T cells.* J Biol Chem, 2001. **276**(37): p. 34896-904.
52. Yonekawa, A., et al., *Dectin-2 Is a Direct Receptor for Mannose-Capped Lipoarabinomannan of Mycobacteria.* Immunity, 2014. **41**(3): p. 402-413.
53. Lerner, T.R., et al., *The innate immune response in human tuberculosis.* Cellular Microbiology, 2015. **17**(9): p. 1277-1285.
54. Jones, B.W., et al., *Different Toll-like receptor agonists induce distinct macrophage responses.* J Leukoc Biol, 2001. **69**(6): p. 1036-44.
55. Means, T.K., et al., *Differential effects of a Toll-like receptor antagonist on Mycobacterium tuberculosis-induced macrophage responses.* J Immunol, 2001. **166**(6): p. 4074-82.
56. Reiling, N., et al., *Cutting edge: Toll-like receptor (TLR)2- and TLR4-mediated pathogen recognition in resistance to airborne infection with Mycobacterium tuberculosis.* J Immunol, 2002. **169**(7): p. 3480-4.
57. Pecora, N.D., et al., *Mycobacterium tuberculosis LprA is a lipoprotein agonist of TLR2 that regulates innate immunity and APC function.* Journal of Immunology, 2006. **177**(1): p. 422-429.
58. Tapping, R.I. and P.S. Tobias, *Mycobacterial lipoarabinomannan mediates physical interactions between TLR1 and TLR2 to induce signaling.* Journal of Endotoxin Research, 2003. **9**(4): p. 264-268.
59. Jo, E.K., et al., *Intracellular signalling cascades regulating innate immune responses to Mycobacteria: branching out from Toll-like receptors.* Cell Microbiol, 2007. **9**(5): p. 1087-98.
60. Drennan, M.B., et al., *Toll-like receptor 2-deficient mice succumb to Mycobacterium tuberculosis infection.* Am J Pathol, 2004. **164**(1): p. 49-57.
61. Abel, B., et al., *Toll-like receptor 4 expression is required to control chronic Mycobacterium tuberculosis infection in mice.* Journal of Immunology, 2002. **169**(6): p. 3155-3162.
62. Bulut, Y., et al., *Cooperation of Toll-like receptor 2 and 6 for cellular activation by soluble tuberculosis factor and Borrelia burgdorferi outer surface protein A lipoprotein: role of Toll-interacting protein and IL-1 receptor signaling molecules in Toll-like receptor 2 signaling.* J Immunol, 2001. **167**(2): p. 987-94.
63. Sanchez, D., et al., *Role of TLR2- and TLR4-mediated signaling in Mycobacterium tuberculosis-induced macrophage death.* Cell Immunol, 2010. **260**(2): p. 128-36.
64. Bafica, A., et al., *TLR9 regulates Th1 responses and cooperates with TLR2 in mediating optimal resistance to Mycobacterium tuberculosis.* J Exp Med, 2005. **202**(12): p. 1715-24.
65. Shi, S., et al., *Expression of many immunologically important genes in Mycobacterium tuberculosis-infected macrophages is independent of both TLR2 and TLR4 but dependent on IFN-alphabeta receptor and STAT1.* J Immunol, 2005. **175**(5): p. 3318-28.
66. Holscher, C., et al., *Containment of aerogenic Mycobacterium tuberculosis infection in mice does not require MyD88 adaptor function for TLR2, -4 and -9.* Eur J Immunol, 2008. **38**(3): p. 680-94.
67. Sugawara, I., et al., *Mycobacterial infection in TLR2 and TLR6 knockout mice.* Microbiol Immunol, 2003. **47**(5): p. 327-36.
68. Sugawara, I., et al., *Mycobacterial infection in MyD88-deficient mice.* Microbiol Immunol, 2003. **47**(11): p. 841-7.
69. Franchi, L., et al., *Function of Nod-like receptors in microbial recognition and host defense.* Immunol Rev, 2009. **227**(1): p. 106-28.
70. Girardin, S.E., et al., *Nod2 is a general sensor of peptidoglycan through muramyl dipeptide (MDP) detection.* Journal of Biological Chemistry, 2003. **278**(11): p. 8869-8872.

71. Divangahi, M., et al., *NOD2-Deficient Mice Have Impaired Resistance to Mycobacterium tuberculosis Infection through Defective Innate and Adaptive Immunity.* Journal of Immunology, 2008. **181**(10): p. 7157-7165.
72. Schroder, K. and J. Tschopp, *The inflammasomes.* Cell, 2010. **140**(6): p. 821-32.
73. Mishra, B.B., et al., *Mycobacterium tuberculosis protein ESAT-6 is a potent activator of the NLRP3/ASC inflammasome.* Cellular Microbiology, 2010. **12**(8): p. 1046-1063.
74. Schweneker, K., et al., *The mycobacterial cord factor adjuvant analogue trehalose-6,6 '-dibehenate (TDB) activates the Nlrp3 inflammasome.* Immunobiology, 2013. **218**(4): p. 664-673.
75. Mortaz, E., et al., *Interaction of Pattern Recognition Receptors with Mycobacterium Tuberculosis.* J Clin Immunol, 2014.
76. Geijtenbeek, T.B. and S.I. Gringhuis, *Signalling through C-type lectin receptors: shaping immune responses.* Nat Rev Immunol, 2009. **9**(7): p. 465-79.
77. Hoving, J.C., G.J. Wilson, and G.D. Brown, *Signalling C-type lectin receptors, microbial recognition and immunity.* Cell Microbiol, 2014. **16**(2): p. 185-94.
78. Zelensky, A.N. and J.E. Gready, *The C-type lectin-like domain superfamily.* FEBS J, 2005. **272**(24): p. 6179-217.
79. Gupta, G. and A. Surolia, *Collectins: sentinels of innate immunity.* Bioessays, 2007. **29**(5): p. 452-64.
80. van de Wetering, J.K., L.M. van Golde, and J.J. Batenburg, *Collectins: players of the innate immune system.* Eur J Biochem, 2004. **271**(7): p. 1229-49.
81. Sancho, D. and C. Reis e Sousa, *Signaling by myeloid C-type lectin receptors in immunity and homeostasis.* Annu Rev Immunol, 2012. **30**: p. 491-529.
82. Goyal, S., T.E. Klassert, and H. Slevogt, *C-type lectin receptors in tuberculosis: what we know.* Medical Microbiology and Immunology, 2016. **205**(6): p. 513-535.
83. Sato, K., et al., *Dectin-2 is a pattern recognition receptor for fungi that couples with the Fc receptor gamma chain to induce innate immune responses.* Journal of Biological Chemistry, 2006. **281**(50): p. 38854-38866.
84. Yamasaki, S., et al., *Mincle is an ITAM-coupled activating receptor that senses damaged cells.* Nature Immunology, 2008. **9**(10): p. 1179-1188.
85. Bakker, A.B.H., et al., *Myeloid DAP12-associating lectin (MDL)-1 is a cell surface receptor involved in the activation of myeloid cells.* Proceedings of the National Academy of Sciences of the United States of America, 1999. **96**(17): p. 9792-9796.
86. Billadeau, D.D. and P.J. Leibson, *ITAMs versus ITIMs: striking a balance during cell regulation.* Journal of Clinical Investigation, 2002. **109**(2): p. 161-168.
87. Kanazawa, N., et al., *DCIR acts as an inhibitory receptor depending on its immunoreceptor tyrosine-based inhibitory motif.* Journal of Investigative Dermatology, 2002. **118**(2): p. 261-266.
88. Marshall, A.S.J., et al., *Identification and characterization of a novel human Myeloid inhibitory c-type lectin-like receptor (MICL) that is predominantly expressed on granulocytes and monocytes.* Journal of Biological Chemistry, 2004. **279**(15): p. 14792-14802.
89. Barrow, A.D. and J. Trowsdale, *You say ITAM and I say ITIM, let's call the whole thing off: the ambiguity of immunoreceptor signalling.* Eur J Immunol, 2006. **36**(7): p. 1646-53.
90. Strasser, D., et al., *Syk kinase-coupled C-type lectin receptors engage protein kinase C-sigma to elicit Card9 adaptor-mediated innate immunity.* Immunity, 2012. **36**(1): p. 32-42.
91. Downing, J.F., et al., *Surfactant protein a promotes attachment of Mycobacterium tuberculosis to alveolar macrophages during infection with human immunodeficiency virus.* Proc Natl Acad Sci U S A, 1995. **92**(11): p. 4848-52.
92. Garred, P., et al., *Dual role of mannan-binding protein in infections: another case of heterosis?* Eur J Immunogenet, 1994. **21**(2): p. 125-31.

93. Gaynor, C.D., et al., *Pulmonary surfactant protein A mediates enhanced phagocytosis of Mycobacterium tuberculosis by a direct interaction with human macrophages.* J Immunol, 1995. **155**(11): p. 5343-51.
94. Beharka, A.A., et al., *Pulmonary surfactant protein A up-regulates activity of the mannose receptor, a pattern recognition receptor expressed on human macrophages.* J Immunol, 2002. **169**(7): p. 3565-73.
95. Gold, J.A., et al., *Surfactant protein A modulates the inflammatory response in macrophages during tuberculosis.* Infect Immun, 2004. **72**(2): p. 645-50.
96. Hu, H., et al., *Clinical value of surfactant protein-A in serum and sputum for pulmonary tuberculosis diagnosis.* Genet Mol Res, 2013. **12**(4): p. 4918-24.
97. Bartlomiejczyk, M.A., et al., *Interaction of lectin pathway of complement-activating pattern recognition molecules with mycobacteria.* Clin Exp Immunol, 2014. **178**(2): p. 310-9.
98. Beltrame, M.H., et al., *MBL-associated serine proteases (MASPs) and infectious diseases.* Mol Immunol, 2015. **67**(1): p. 85-100.
99. Schlesinger, L.S., S.R. Hull, and T.M. Kaufman, *Binding of the terminal mannosyl units of lipoarabinomannan from a virulent strain of Mycobacterium tuberculosis to human macrophages.* J Immunol, 1994. **152**(8): p. 4070-9.
100. Astarie-Dequeker, C., et al., *The mannose receptor mediates uptake of pathogenic and nonpathogenic mycobacteria and bypasses bactericidal responses in human macrophages.* Infect Immun, 1999. **67**(2): p. 469-77.
101. Nigou, J., et al., *Mannosylated lipoarabinomannans inhibit IL-12 production by human dendritic cells: evidence for a negative signal delivered through the mannose receptor.* J Immunol, 2001. **166**(12): p. 7477-85.
102. Chieppa, M., et al., *Cross-linking of the mannose receptor on monocyte-derived dendritic cells activates an anti-inflammatory immunosuppressive program.* J Immunol, 2003. **171**(9): p. 4552-60.
103. Kang, P.B., et al., *The human macrophage mannose receptor directs Mycobacterium tuberculosis lipoarabinomannan-mediated phagosome biogenesis.* J Exp Med, 2005. **202**(7): p. 987-99.
104. Tailleux, L., et al., *DC-SIGN is the major Mycobacterium tuberculosis receptor on human dendritic cells.* Journal of Experimental Medicine, 2003. **197**(1): p. 121-127.
105. Zenaro, E., M. Donini, and S. Dusi, *Induction of Th1/Th17 immune response by Mycobacterium tuberculosis: role of dectin-1, Mannose Receptor, and DC-SIGN.* J Leukoc Biol, 2009. **86**(6): p. 1393-401.
106. Ostrop, J., et al., *Contribution of MINCLE-SYK Signaling to Activation of Primary Human APCs by Mycobacterial Cord Factor and the Novel Adjuvant TDB.* J Immunol, 2015. **195**(5): p. 2417-28.
107. Zhao, X.Q., et al., *C-type lectin receptor dectin-3 mediates trehalose 6,6'-dimycolate (TDM)-induced Mincle expression through CARD9/Bcl10/MALT1-dependent nuclear factor (NF)-kappaB activation.* J Biol Chem, 2014. **289**(43): p. 30052-62.
108. Miyake, Y., O.H. Masatsugu, and S. Yamasaki, *C-Type Lectin Receptor MCL Facilitates Mincle Expression and Signaling through Complex Formation.* J Immunol, 2015. **194**(11): p. 5366-74.
109. Kerscher, B., et al., *Mycobacterial receptor, Clec4d (CLECSF8, MCL), is coregulated with Mincle and upregulated on mouse myeloid cells following microbial challenge.* European Journal of Immunology, 2016. **46**(2): p. 381-389.
110. Wilson, G.J., et al., *The C-type lectin receptor CLECSF8/CLEC4D is a key component of anti-mycobacterial immunity.* Cell Host Microbe, 2015. **17**(2): p. 252-9.
111. Lee, H.M., et al., *Dectin-1 is Inducible and Plays an Essential Role for Mycobacteria-Induced Innate Immune Responses in Airway Epithelial Cells.* Journal of Clinical Immunology, 2009. **29**(6): p. 795-805.

112. Romero, M.M., et al., *Reactive oxygen species production by human dendritic cells involves TLR2 and dectin-1 and is essential for efficient immune response against Mycobacteria.* Cellular Microbiology, 2016. **18**(6): p. 875-886.
113. van de Veerdonk, F.L., et al., *Mycobacterium tuberculosis induces IL-17A responses through TLR4 and dectin-1 and is critically dependent on endogenous IL-1.* J Leukoc Biol, 2010. **88**(2): p. 227-32.
114. Mahasirimongkol, S., et al., *Genome-wide association studies of tuberculosis in Asians identify distinct at-risk locus for young tuberculosis.* J Hum Genet, 2012. **57**(6): p. 363-7.
115. Png, E., et al., *A genome wide association study of pulmonary tuberculosis susceptibility in Indonesians.* BMC Med Genet, 2012. **13**: p. 5.
116. Thye, T., et al., *Genome-wide association analyses identifies a susceptibility locus for tuberculosis on chromosome 18q11.2.* Nat Genet, 2010. **42**(9): p. 739-41.
117. Khalilullah, S.A., et al., *Host genome polymorphisms and tuberculosis infection: What we have to say?* Egypt J Chest Dis Tuberc, 2014. **63**(1): p. 173-185.
118. Lai, Y.F., et al., *Functional polymorphisms of the TLR7 and TLR8 genes contribute to Mycobacterium tuberculosis infection.* Tuberculosis (Edinb), 2016. **98**: p. 125-31.
119. Floros, J., et al., *Surfactant protein genetic marker alleles identify a subgroup of tuberculosis in a Mexican population.* J Infect Dis, 2000. **182**(5): p. 1473-8.
120. Malik, S., et al., *Variants of the SFTPA1 and SFTPA2 genes and susceptibility to tuberculosis in Ethiopia.* Hum Genet, 2006. **118**(6): p. 752-9.
121. Madan, T., et al., *Association of polymorphisms in the collagen region of human SP-A1 and SP-A2 genes with pulmonary tuberculosis in Indian population.* Clinical Chemistry and Laboratory Medicine, 2002. **40**(10): p. 1002-1008.
122. Yang, H.Y., et al., *Correlation analysis between single nucleotide polymorphisms of pulmonary surfactant protein A gene and pulmonary tuberculosis in the Han population in China.* International Journal of Infectious Diseases, 2014. **26**: p. 31-36.
123. Chang, K., et al., *Association between CD209-336A/G and-871A/G Polymorphisms and Susceptibility of Tuberculosis: A Meta-Analysis.* Plos One, 2012. **7**(7).
124. El Sahly, H.M., et al., *The effect of mannose binding lectin gene polymorphisms on susceptibility to tuberculosis in different ethnic groups.* Scand J Infect Dis, 2004. **36**(2): p. 106-8.
125. Liu, W., et al., *Sequence variations in the MBL gene and their relationship to pulmonary tuberculosis in the Chinese Han population.* International Journal of Tuberculosis and Lung Disease, 2006. **10**(10): p. 1098-1103.
126. Wojitani, M.D.K.H., et al., *Association between mannose-binding lectin and interleukin-1 receptor antagonist gene polymorphisms and recurrent vulvovaginal candidiasis.* Archives of Gynecology and Obstetrics, 2012. **285**(1): p. 149-153.
127. Denholm, J.T., E.S. McBryde, and D.P. Eisen, *Mannose-binding lectin and susceptibility to tuberculosis: a meta-analysis.* Clin Exp Immunol, 2010. **162**(1): p. 84-90.
128. Chalmers, J.D., et al., *No Strong Relationship Between Components of the Lectin Pathway of Complement and Susceptibility to Pulmonary Tuberculosis.* Inflammation, 2015. **38**(4): p. 1731-1737.
129. Chen, M., et al., *Impact of MBL and MASP-2 gene polymorphism and its interaction on susceptibility to tuberculosis.* BMC Infect Dis, 2015. **15**: p. 151.
130. Zhang, X., et al., *Polymorphic allele of human MRC1 confer protection against tuberculosis in a Chinese population.* Int J Biol Sci, 2012. **8**(3): p. 375-82.
131. Zhang, X., et al., *The novel human MRC1 gene polymorphisms are associated with susceptibility to pulmonary tuberculosis in Chinese Uygur and Kazak populations.* Mol Biol Rep, 2013. **40**(8): p. 5073-83.
132. Bowker, N., et al., *Polymorphisms in the Pattern Recognition Receptor Mincle Gene (CLEC4E) and Association with Tuberculosis.* Lung, 2016.

133. de la Fuente, G., et al., *Pros and cons of ion-torrent next generation sequencing versus terminal restriction fragment length polymorphism T-RFLP for studying the rumen bacterial community.* PLoS One, 2014. **9**(7): p. e101435.
134. Quail, M.A., et al., *A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers.* BMC Genomics, 2012. **13**: p. 341.
135. DePristo, M.A., et al., *A framework for variation discovery and genotyping using next-generation DNA sequencing data.* Nature Genetics, 2011. **43**(5): p. 491-+.
136. Gabriel S. , Z.L., Tabbaa D., *SNP Genotyping Using the Sequenom MassARRAY iPLEX Platform*, in *Current Protocols in Human genetics*. 2009.
137. He, C., J. Holme, and J. Anthony, *SNP genotyping: the KASP assay.* Methods Mol Biol, 2014. **1145**: p. 75-86.
138. Central TB Division, D.G.o.H.S., Ministry of Health and Family Welfare, Govt. of India, *Technical and Operational Guidelines for Tuberculosis Control*, C.T.D. Directorate General of Health Services, Editor. 2005: New Delhi.
139. Dittrich, N., et al., *Toll-like receptor 1 variations influence susceptibility and immune response to Mycobacterium tuberculosis.* Tuberculosis (Edinb), 2015.
140. Oh, D.Y., et al., *A functional toll-like receptor 8 variant is associated with HIV disease restriction.* Journal of Infectious Diseases, 2008. **198**(5): p. 701-709.
141. Glocker, E.O., et al., *A homozygous CARD9 mutation in a family with susceptibility to fungal infections.* N Engl J Med, 2009. **361**(18): p. 1727-35.
142. Hattori, T., et al., *Genetic variants in the mannose receptor gene (MRC1) are associated with asthma in two independent populations.* Immunogenetics, 2009. **61**(11-12): p. 731-8.
143. Hattori, T., et al., *Genetic variants in mannose receptor gene (MRC1) confer susceptibility to increased risk of sarcoidosis.* BMC Med Genet, 2010. **11**: p. 151.
144. Silveyra, P. and J. Floros, *Genetic variant associations of human SP-A and SP-D with acute and chronic lung injury.* Front Biosci (Landmark Ed), 2012. **17**: p. 407-29.
145. Ammitzboll, C.G., et al., *Polymorphisms in the MASP1 gene are associated with serum levels of MASP-1, MASP-3, and MAp44.* PLoS One, 2013. **8**(9): p. e73317.
146. Feinberg, H., et al., *Common polymorphisms in human langerin change specificity for glycan ligands.* J Biol Chem, 2013. **288**(52): p. 36762-71.
147. Vaid, M., et al., *Association of SP-D, MNL and I-NOS genetic variants with pulmonary tuberculosis.* Indian Journal of Human Genetics, 2006. **12**(3): p. 105-110.
148. Alter, A., et al., *Genetic and functional analysis of common MRC1 exon 7 polymorphisms in leprosy susceptibility.* Hum Genet, 2010. **127**(3): p. 337-48.
149. Langmead, B. and S.L. Salzberg, *Fast gapped-read alignment with Bowtie 2.* Nat Methods, 2012. **9**(4): p. 357-9.
150. Wang, C.L., et al., *Improved Ancestry Estimation for both Genotyping and Sequencing Data using Projection Procrustes Analysis and Genotype Imputation.* American Journal of Human Genetics, 2015. **96**(6): p. 926-937.
151. Danecek, P., et al., *The variant call format and VCFtools.* Bioinformatics, 2011. **27**(15): p. 2156-2158.
152. McKenna, A., et al., *The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data.* Genome Research, 2010. **20**(9): p. 1297-1303.
153. Purcell, S., et al., *PLINK: a tool set for whole-genome association and population-based linkage analyses.* Am J Hum Genet, 2007. **81**(3): p. 559-75.
154. Bothamley, G.H., *Smoking and tuberculosis: a chance or causal association?* Thorax, 2005. **60**(7): p. 527-8.
155. Rehm, J., et al., *The association between alcohol use, alcohol use disorders and tuberculosis (TB). A systematic review.* Bmc Public Health, 2009. **9**.

156. Mhurchu, C.N., et al., *Body mass index and cardiovascular disease in the Asia-Pacific Region: an overview of 33 cohorts involving 310 000 participants.* International Journal of Epidemiology, 2004. **33**(4): p. 751-758.
157. Yen, Y.F., et al., *Association of Body Mass Index With Tuberculosis Mortality A Population-Based Follow-Up Study.* Medicine, 2016. **95**(1).
158. Schmidt, C.W., *Linking TB and the environment: an overlooked mitigation strategy.* Environ Health Perspect, 2008. **116**(11): p. A478-85.
159. Harris, J. and J. Keane, *How tumour necrosis factor blockers interfere with tuberculosis immunity.* Clinical and Experimental Immunology, 2010. **161**(1): p. 1-9.
160. Krogh, S.S., et al., *Plasma levels of MASP-1, MASP-3 and MAp44 in patients with type 2 diabetes: influence of glycaemic control, body composition and polymorphisms in the MASP1 gene.* Clin Exp Immunol, 2017.
161. Alshehri, A.M., *Metabolic syndrome and cardiovascular risk.* J Family Community Med, 2010. **17**(2): p. 73-8.
162. Marchini, J., et al., *The effects of human population structure on large genetic association studies.* Nat Genet, 2004. **36**(5): p. 512-7.
163. Bittles, A.H., *Population stratification and genetic association studies in South Asia.* J Mol Genet Med, 2005. **1**(2): p. 43-8.
164. Merle, N.S., et al., *Complement System Part I - Molecular Mechanisms of Activation and Regulation.* Front Immunol, 2015. **6**: p. 262.
165. Heja, D., et al., *Revised mechanism of complement lectin-pathway activation revealing the role of serine protease MASP-1 as the exclusive activator of MASP-2.* Proceedings of the National Academy of Sciences of the United States of America, 2012. **109**(26): p. 10498-10503.
166. Megyeri, M., et al., *Quantitative characterization of the activation steps of mannan-binding lectin (MBL)-associated serine proteases (MASPs) points to the central role of MASP-1 in the initiation of the complement lectin pathway.* J Biol Chem, 2013. **288**(13): p. 8922-34.
167. Iwaki, D., et al., *The role of mannose-binding lectin-associated serine protease-3 in activation of the alternative complement pathway.* J Immunol, 2011. **187**(7): p. 3751-8.
168. Degn, S.E., et al., *MAp44, a human protein associated with pattern recognition molecules of the complement system and regulating the lectin pathway of complement activation.* J Immunol, 2009. **183**(11): p. 7371-8.
169. Brown, K.S., et al., *Severe fibrosis in hepatitis C virus-infected patients is associated with increased activity of the mannan-binding lectin (MBL)/MBL-associated serine protease 1 (MASP-1) complex.* Clinical and Experimental Immunology, 2007. **147**(1): p. 90-98.
170. Saeed, A., et al., *Mannan binding lectin-associated serine protease 1 is induced by hepatitis C virus infection and activates human hepatic stellate cells.* Clinical and Experimental Immunology, 2013. **174**(2): p. 265-273.
171. Haerynck, F., et al., *Polymorphisms in the lectin pathway genes as a possible cause of early chronic Pseudomonas aeruginosa colonization in cystic fibrosis patients.* Human Immunology, 2012. **73**(11): p. 1175-1183.
172. Rooryck, C., et al., *Mutations in lectin complement pathway genes COLEC11 and MASP1 cause 3MC syndrome.* Nat Genet, 2011. **43**(3): p. 197-203.
173. Frauenknecht, V., et al., *Plasma levels of mannan-binding lectin (MBL)-associated serine proteases (MASPs) and MBL-associated protein in cardio- and cerebrovascular diseases.* Clin Exp Immunol, 2013. **173**(1): p. 112-20.
174. Petri, C., et al., *Investigation of Complement-activating Pattern Recognition Molecules and Associated Enzymes as Possible Inflammatory Markers in Oligoarticular and Systemic Juvenile Idiopathic Arthritis.* Journal of Rheumatology, 2015. **42**(7): p. 1252-1258.
175. Jenny, L., et al., *Plasma levels of mannan-binding lectin-associated serine proteases MASP-1 and MASP-2 are elevated in type 1 diabetes and correlate with glycaemic control.* Clinical and Experimental Immunology, 2015. **180**(2): p. 227-232.

176. Collard, C.D., et al., *Complement activation after oxidative stress: role of the lectin complement pathway.* Am J Pathol, 2000. **156**(5): p. 1549-56.
177. Ehrt, S. and D. Schnappinger, *Mycobacterial survival strategies in the phagosome: defence against host stresses.* Cellular Microbiology, 2009. **11**(8): p. 1170-1178.
178. Kumar Nathella, P. and S. Babu, *Influence of diabetes mellitus on immunity to human tuberculosis.* Immunology, 2017.
179. Ghosh, P., et al., *Role of complement and complement regulatory proteins in the complications of diabetes.* Endocr Rev, 2015. **36**(3): p. 272-88.
180. Thiel, S., et al., *Mannan-binding lectin (MBL)-associated serine protease-1 (MASP-1), a serine protease associated with humoral pattern-recognition molecules: normal and acute-phase levels in serum and stoichiometry of lectin pathway components.* Clinical and Experimental Immunology, 2012. **169**(1): p. 38-48.
181. Terai, I., et al., *Human serum mannose-binding lectin (MBL)-associated serine protease-1 (MASP-1): determination of levels in body fluids and identification of two forms in serum.* Clin Exp Immunol, 1997. **110**(2): p. 317-23.
182. Megyeri, M., et al., *Complement Protease MASP-1 Activates Human Endothelial Cells: PAR4 Activation Is a Link between Complement and Endothelial Function.* Journal of Immunology, 2009. **183**(5): p. 3409-3416.
183. Jani, P.K., et al., *MASP-1 Induces a Unique Cytokine Pattern in Endothelial Cells: A Novel Link between Complement System and Neutrophil Granulocytes.* Plos One, 2014. **9**(1).
184. Lan, R.S., G.A. Stewart, and P.J. Henry, *Role of protease-activated receptors in airway function: a target for therapeutic intervention?* Pharmacol Ther, 2002. **95**(3): p. 239-57.
185. Jenny, L., et al., *MASP-1 of the complement system promotes clotting via prothrombin activation.* Mol Immunol, 2015. **65**(2): p. 398-405.
186. Jenny, L., et al., *MASP-1 Induced Clotting--The First Model of Prothrombin Activation by MASP-1.* PLoS One, 2015. **10**(12): p. e0144633.
187. Dobo, J., et al., *MASP-1, a promiscuous complement protease: structure of its catalytic region reveals the basis of its broad specificity.* J Immunol, 2009. **183**(2): p. 1207-14.
188. Krarup, A., et al., *The action of MBL-associated serine protease 1 (MASP1) on factor XIII and fibrinogen.* Biochim Biophys Acta, 2008. **1784**(9): p. 1294-300.
189. Hess, K., et al., *Effects of MASP-1 of the complement system on activation of coagulation factors and plasma clot formation.* PLoS One, 2012. **7**(4): p. e35690.
190. Senger, D.R., et al., *Adhesive properties of osteopontin: regulation by a naturally occurring thrombin-cleavage in close proximity to the GRGDS cell-binding domain.* Mol Biol Cell, 1994. **5**(5): p. 565-74.
191. Ashkar, S., et al., *Eta-1 (osteopontin): an early component of type-1 (cell-mediated) immunity.* Science, 2000. **287**(5454): p. 860-4.
192. Hasibuan, F.M., et al., *Evaluation of matricellular proteins in systemic and local immune response to Mycobacterium tuberculosis infection.* Microbiol Immunol, 2015. **59**(10): p. 623-32.
193. Ridruechai, C., et al., *ASSOCIATION BETWEEN CIRCULATING FULL-LENGTH OSTEOPONTIN AND IFN-gamma WITH DISEASE STATUS OF TUBERCULOSIS AND RESPONSE TO SUCCESSFUL TREATMENT.* Southeast Asian Journal of Tropical Medicine and Public Health, 2011. **42**(4): p. 876-889.

APPENDICES

APPENDIX I: List of known SNPs generated after filtering and annotating the SNVs obtained by Ampliseq Sequencing (Phase I)

CHR	SNP	SNP_ID	Genotype (UNADJ-p-value)	Genotype Ancestry-ADJ (p-value)	Allele freq (p-value)
10	10:81371953	rs1914663	0,02501	0,07434	0,00943
12	12:10110431	rs35333643	0,04537	0,1111	0,01697
12	12:10116631	rs3110948	0,02449	0,0733	0,01754
3	3:186965268	rs3774275	0,02122	0,06657	0,02088
12	12:10159846	rs114421141	0,08831	0,1748	0,02427
12	12:10116733	rs2961541	0,05992	0,1341	0,02809
12	12:10156769	rs79967076	0,08831	0,1748	0,03064
12	12:10170823	rs112915340	0,08831	0,1748	0,03064
12	12:10150905	rs374147676	0,03482	0,09288	0,03263
12	12:10103208	rs76427726	0,09018	0,1773	0,03279
12	12:10117091	rs190925857	0,09018	0,1773	0,03279
12	12:10123787	rs193214822	0,09018	0,1773	0,03279
12	12:10112586	rs148864420	0,09018	0,1773	0,03279
12	12:10280862	rs143386125	0,01908	0,06201	0,03348
2	2:71058835	rs741326	0,02093	0,06597	0,03697
12	12:10116629	rs3110949	0,05992	0,1341	0,04830
12	12:10116577	rs648985	0,05992	0,1341	0,04830
2	2:71058226	rs2080390	0,02805	0,0803	0,04965
12	12:10158903	rs17205441	0,08831	0,1748	0,05621
12	12:10170154	rs3736850	0,08831	0,1748	0,05621
12	12:10160808	rs75972360	0,08831	0,1748	0,05643
12	12:10156147	rs7132948	0,08831	0,1748	0,05655
12	12:10151747	rs118077922	0,1799	0,2856	0,06186
12	12:8615602	rs12099687	0,09296	0,181	0,06186
12	12:10151172	rs7980698	0,9454	0,9565	0,06289
12	12:10125497	rs147529241	0,09018	0,1773	0,06305
12	12:10121476	rs182670340	0,09018	0,1773	0,06350
12	12:10122346	rs183462894	0,09018	0,1773	0,06350
12	12:10121072	rs185612837	0,09018	0,1773	0,06359
12	12:8626362	rs4334073/rs199783308	0,06281	0,1385	0,09032
12	12:8290569	rs7302963	0,08688	0,1728	0,09179
12	12:8614238	rs4883148	0,1048	0,1965	0,10000
12	12:10114753	rs145937548	0,09018	0,1773	0,10245
19	19:7812536	rs2287886	0,1131	0,2071	0,11491
12	12:8289877	rs144658267	0,03736	0,0974	0,11596
2	2:71061108	rs17662453	0,07139	0,1511	0,12536
12	12:8617976	rs7302011	0,02766	0,07955	0,12536
12	12:10270773	rs10845048	0,07197	0,152	0,12821
10	10:54531235	rs1800450	0,1195	0,2152	0,13006

12	12:10156051	rs116924172	0,1397	0,2396	0,13086
12	12:7886732	rs77567151	0,08146	0,1654	0,13252
2	2:71058906	rs3213749	0,07546	0,157	0,14037
2	2:71058811	rs3213748	0,07546	0,157	0,14037
2	2:71059153	rs3815556	0,07546	0,157	0,14106
2	2:71058306	rs13383830	0,07546	0,157	0,14106
12	12:8626878	rs4883164	0,06281	0,1385	0,14140
10	10:81706951	rs726289	0,1157	0,2104	0,14493
12	12:8627924	rs11045619	0,0445	0,1096	0,14735
12	12:10311916	rs76039690	0,1038	0,1952	0,15095
10	10:81733022	rs10788338	0,2348	0,3442	0,15431
12	12:8623095	rs4462413	0,1063	0,1985	0,15542
12	12:8623101	rs4301857	0,1063	0,1985	0,15542
12	12:8618971	rs7309596	0,1063	0,1985	0,15629
12	12:8620339	rs7963053	0,1063	0,1985	0,15629
12	12:8622599	rs4883157	0,1063	0,1985	0,15629
12	12:8617757	rs7315590	0,1063	0,1985	0,15629
12	12:7900184	rs10845821	0,114	0,2082	0,16017
12	12:7904111	rs11055602	0,114	0,2082	0,16017
10	10:54531685	rs7096206	0,1092	0,2022	0,16453
10	10:81707613	rs2819096	0,2763	0,386	0,16574
12	12:10166105	rs57789100	0,1939	0,3009	0,16819
19	19:7807482	rs4804801	0,2584	0,3682	0,17090
12	12:8664351	rs7135960	0,1088	0,2016	0,17489
12	12:8666118	rs11045983	0,1308	0,2289	0,18185
12	12:10108433	rs676397	0,1104	0,2037	0,18185
12	12:8662185	rs10841788	0,1367	0,2361	0,18300
2	2:71061051	rs41285965	0,1083	0,201	0,18545
19	19:7834274	rs15282	0,07257	0,1528	0,18577
17	17:60742348	rs2252814	0,1431	0,2436	0,18624
12	12:8664719	rs28485567	0,1398	0,2398	0,19024
12	12:8666378	rs11045985	0,1398	0,2398	0,19024
12	12:7892937	rs11055538	0,1263	0,2235	0,19126
12	12:8674727	rs7976134	0,1049	0,1966	0,19308
12	12:10324431	rs35688880	0,1864	0,2927	0,19397
12	12:10139024	rs112633624	0,1037	0,1951	0,19805
12	12:10313448	rs11053646	0,1288	0,2266	0,20092
12	12:10322882	rs2742113	0,1517	0,2537	0,20364
12	12:10322817	rs2742114	0,1517	0,2537	0,20364
12	12:10241261	rs117440358	0,1368	0,2361	0,20660
9	9:93628105	rs2991215	0,1824	0,2884	0,20816
2	2:71058795	rs3213747	0,09108	0,1785	0,20938
12	12:8672848	rs138110414	0,1282	0,2259	0,20997
12	12:10199326	rs373178652	0,1007	0,1912	0,20997
10	10:81704559	rs2255601	0,205	0,3129	0,22406
12	12:7902281	rs73056636	0,1424	0,2428	0,22679
12	12:8291122	rs1133104	0,1508	0,2526	0,22976

12	12:8620065	rs7978179	0,1063	0,1985	0,23061
12	12:8622964	rs116034720	0,1063	0,1985	0,23061
12	12:8622997	rs114668263	0,1063	0,1985	0,23061
12	12:8619792	rs4427631	0,1063	0,1985	0,23106
12	12:8618563	rs7134303	0,1063	0,1985	0,23106
12	12:8623261	rs4438120	0,1063	0,1985	0,23147
12	12:8623815	rs4329732	0,1063	0,1985	0,23147
12	12:8618488	rs4460883	0,1063	0,1985	0,23190
12	12:10320212	rs3912640	0,1202	0,216	0,23726
9	9:93628027	rs2991216	0,1724	0,2773	0,23782
12	12:8622488	rs4883156	0,1162	0,211	0,24056
12	12:10148203	rs3818343	0,2778	0,3875	0,24348
12	12:10270813	rs11053593	0,174	0,279	0,24348
12	12:10270821	rs11053594	0,174	0,279	0,24348
12	12:10270822	rs11053595	0,174	0,279	0,24348
12	12:8288403	rs17728942	0,1784	0,2839	0,24678
12	12:8615809	rs4242889	0,1837	0,2898	0,24768
21	21:46330674	rs11088969	0,233	0,3423	0,24794
12	12:7897184	rs6488614	0,2321	0,3414	0,24800
12	12:7904154	rs12422412	0,2245	0,3335	0,24800
12	12:8668281	rs57832572	0,1308	0,2289	0,25188
12	12:8666801	rs4534636	0,1919	0,2988	0,25740
16	16:31271063	rs7193943	0,2202	0,329	0,26089
12	12:10139082	rs574097	0,2822	0,3918	0,26137
12	12:10275072	rs59913193	0,1996	0,3071	0,26431
12	12:10315562	rs12827232	0,2085	0,3166	0,27007
12	12:10133047	rs79246913	0,2856	0,3952	0,27253
12	12:10245972	rs368830228	0,1974	0,3048	0,27337
19	19:7833982	rs1045997	0,1894	0,2961	0,27337
12	12:8284104	rs201561995	0,1437	0,2444	0,27648
12	12:8609845	rs4242888	0,2951	0,4045	0,28078
12	12:8611768	rs4264222	0,2951	0,4045	0,28078
12	12:8612481	rs12300621	0,2951	0,4045	0,28078
12	12:8612484	rs12302015	0,2951	0,4045	0,28078
12	12:8612791	rs10770737	0,2951	0,4045	0,28078
12	12:8613204	rs10770739	0,2951	0,4045	0,28078
12	12:8610352	rs4402377	0,2951	0,4045	0,28078
12	12:8610317	rs4623978	0,2951	0,4045	0,28078
12	12:8610291	rs4255605	0,2951	0,4045	0,28078
12	12:8607626	rs4882942	0,2951	0,4045	0,28078
12	12:8609136	rs7968198	0,2951	0,4045	0,28078
12	12:7882602	rs6488608	0,1471	0,2484	0,28457
3	3:53220215	rs2230494	0,3216	0,43	0,29896
9	9:139266405	rs10781499	0,2013	0,3089	0,30872
9	9:139266496	rs4077515	0,2013	0,3089	0,30872
2	2:71061399	rs17719042	0,2851	0,3947	0,31018
2	2:71058184	rs13421115	0,1244	0,2212	0,31122

12	12:8608482	rs202005669	0,2032	0,3109	0,31323
12	12:10315695	rs2634159	0,2303	0,3395	0,31479
12	12:10150961	rs2273987	0,3074	0,4164	0,31824
12	12:8617982	rs7315873	0,1063	0,1985	0,31924
12	12:10312291	rs12316150	0,228	0,3371	0,32474
12	12:10311166	rs1050289	0,228	0,3371	0,32474
12	12:8622107	rs4883155	0,1063	0,1985	0,32726
12	12:7894192	rs12823478	0,2937	0,4032	0,32787
12	12:10277113	rs11053613	0,119	0,2145	0,32787
12	12:10247561	rs76240769	0,2665	0,3763	0,32894
12	12:10146004	rs75575903	0,3254	0,4336	0,33111
12	12:7896655	rs7314869	0,2321	0,3414	0,33399
12	12:7895628	rs56318901	0,2321	0,3414	0,33399
10	10:81318663	rs17886395	0,2893	0,3988	0,33575
12	12:10318348	rs2010655	0,307	0,416	0,33852
12	12:10148703	rs659928	0,2812	0,3909	0,33974
12	12:10325128	rs2742112	0,2846	0,3942	0,34223
12	12:7897408	rs113955791	0,2135	0,3219	0,35000
12	12:7903171	rs76947756	0,2135	0,3219	0,35000
12	12:7904101	rs74903322	0,2135	0,3219	0,35000
12	12:10143498	rs1359081	0,2822	0,3918	0,35307
12	12:10320853	rs35335503	0,8161	0,8531	0,35713
12	12:10319805	rs35492478	0,4037	0,5061	0,35758
10	10:81704512	rs2181204	0,3127	0,4215	0,36247
12	12:10315915	rs35311664	0,4037	0,5061	0,36388
12	12:10273769	rs58924693	0,2642	0,374	0,36570
12	12:10275529	rs7136680	0,2642	0,374	0,36570
12	12:10274849	rs56140555	0,2642	0,374	0,36570
12	12:10274029	rs11053603	0,2642	0,374	0,36570
12	12:10104003	rs118182007	0,1988	0,3062	0,36570
12	12:7892802	rs7971286	0,3569	0,4632	0,36570
12	12:7896564	rs7304977	0,3569	0,4632	0,36570
12	12:7896605	rs7137970	0,3569	0,4632	0,36570
12	12:8693789	rs7132177	0,3226	0,4309	0,36605
12	12:10147722	rs61917158	0,2387	0,3481	0,36686
12	12:10125814	rs184227507	0,2212	0,3301	0,36717
12	12:10275769	rs7137840	0,179	0,2846	0,36776
12	12:10270938	rs11053597	0,174	0,279	0,36905
19	19:7813336	rs735240	0,2992	0,4084	0,37615
12	12:8608502	rs4628756	0,2951	0,4045	0,37773
12	12:8610223	rs4638374	0,2951	0,4045	0,38301
12	12:8611048	rs10743393	0,2951	0,4045	0,38301
12	12:8610268	rs4435079	0,2951	0,4045	0,38301
12	12:8607833	rs10734709	0,2951	0,4045	0,38361
12	12:8611129	rs10770734	0,2951	0,4045	0,38415
12	12:8610119	rs4641551	0,2951	0,4045	0,38415
12	12:8611176	rs10770735	0,2951	0,4045	0,38415

12	12:8611428	rs10743394	0,2951	0,4045	0,38415
12	12:8611429	rs61920556	0,2951	0,4045	0,38415
12	12:8613079	rs10770738	0,2951	0,4045	0,38415
12	12:8609674	rs4497482	0,2951	0,4045	0,38415
12	12:8609593	rs4883147	0,2951	0,4045	0,38415
12	12:8608226	rs10770722	0,2951	0,4045	0,38472
12	12:8611964	rs4623977	0,2951	0,4045	0,38765
12	12:8612336	rs10770736	0,2951	0,4045	0,38765
12	12:8613412	rs10770740	0,2951	0,4045	0,38765
12	12:10154249	rs10772223	0,3932	0,4966	0,38981
12	12:10122377	rs640817	0,3285	0,4366	0,39390
12	12:10248927	rs12370211	0,656	0,7228	0,39641
12	12:8609795	rs4459385	0,436	0,5351	0,39885
12	12:10132335	rs61913543	0,3319	0,4397	0,40196
12	12:10129599	rs1447875/rs201768158	0,3319	0,4397	0,40196
12	12:10131044	rs61913542	0,3319	0,4397	0,40253
21	21:46338403	rs9306118	0,375	0,48	0,40485
12	12:10160231	rs7957224	0,4207	0,5215	0,41273
12	12:10149116	rs4764177	0,4185	0,5195	0,41273
12	12:10137259	rs684134	0,4185	0,5195	0,41273
10	10:81319267	rs1650232/rs72659393	0,2204	0,3292	0,41432
12	12:10320444	rs3741860	0,2461	0,3557	0,42211
16	16:31340909	rs4077810	0,3279	0,436	0,42698
12	12:10138708	rs77140517	0,4921	0,5844	0,42731
12	12:10121927	rs563651	0,3549	0,4614	0,42872
12	12:10122204	rs560906	0,3549	0,4614	0,43142
12	12:10124533	rs1797526	0,3549	0,4614	0,43142
12	12:10147434	rs150432914	0,3611	0,4671	0,43925
12	12:7882555	rs370237272	0,2481	0,3578	0,44261
12	12:8291104	rs11043532	0,2335	0,3429	0,44261
12	12:7897898	rs77511202	0,2241	0,3331	0,44462
12	12:10226448	rs189753650	0,7332	0,7861	0,44479
12	12:10109175	rs673173	0,7988	0,8392	0,45005
12	12:10113083	rs2012504	0,3701	0,4755	0,45067
12	12:7882038	rs79787224	0,325	0,4332	0,45883
12	12:10168563	rs10845019	0,3933	0,4967	0,46422
12	12:8616149	rs4883154	0,1837	0,2898	0,46941
12	12:8694267	rs4883167	0,549	0,6333	0,47590
12	12:10150614	rs477658	0,4043	0,5066	0,47929
12	12:10150668	rs1746123	0,4043	0,5066	0,47929
12	12:10151344	rs7977955	0,4043	0,5066	0,47929
12	12:10151405	rs7980924	0,4043	0,5066	0,47929
12	12:10149657	rs582968	0,4043	0,5066	0,48209
12	12:10150069	rs584856	0,4043	0,5066	0,48209
12	12:10152679	rs935538	0,4043	0,5066	0,48320
12	12:10123861	rs525013	0,437	0,536	0,48496
12	12:10120548	rs611819	0,437	0,536	0,48496

12	12:10120556	rs611821	0,437	0,536	0,48496
12	12:10153996	rs12826100	0,3153	0,424	0,48691
12	12:10123677	rs526157	0,437	0,536	0,48774
12	12:10121976	rs562839	0,437	0,536	0,48774
12	12:10131865	rs608418	0,5683	0,6497	0,48891
12	12:10130601	rs570082	0,5683	0,6497	0,48891
12	12:10127894	rs609216	0,5683	0,6497	0,48891
12	12:10134345	rs650368	0,5565	0,6396	0,49014
12	12:7884351	rs17199006	0,4183	0,5193	0,49220
12	12:10282987	rs11833740	0,3583	0,4645	0,49225
12	12:7892112	rs117121134	0,3094	0,4183	0,49507
12	12:10152364	rs11053538	0,3074	0,4164	0,49507
12	12:10157066	rs12820108	0,3153	0,424	0,49524
12	12:10148609	rs35890903	0,3153	0,424	0,49524
12	12:8693689	rs11046135	0,6275	0,6992	0,49739
12	12:7903567	rs7300199	0,3654	0,4711	0,49913
12	12:10272593	rs7309123	0,4887	0,5815	0,50438
1	1:85733374	rs3768235	0,3336	0,4414	0,50772
12	12:10212093	rs61918595	0,9954	0,9963	0,51002
12	12:10247896	rs7953120	0,3653	0,471	0,51652
7	7:141627149	rs1285933	0,3083	0,4173	0,51764
12	12:10248147	rs7956327	0,6125	0,6867	0,51922
12	12:10222356	rs7136826	0,5111	0,6009	0,52328
12	12:10248046	rs7956208	0,6425	0,7117	0,52559
12	12:10251385	rs3816845	0,6425	0,7117	0,52559
12	12:10249513	rs10845040	0,5236	0,6116	0,52559
12	12:10320202	rs11053653	0,4097	0,5115	0,52694
12	12:10244822	rs7953886	0,5051	0,5957	0,52720
12	12:10235535	rs10845032	0,5051	0,5957	0,52720
12	12:10244716	rs7969125	0,5051	0,5957	0,52720
12	12:10244371	rs6488253	0,5051	0,5957	0,52720
12	12:10249129	rs11053577	0,7838	0,827	0,52878
12	12:10249145	rs11053578	0,7838	0,827	0,52878
12	12:10249204	rs11053579	0,7838	0,827	0,52878
12	12:10230072	rs1948185	0,6425	0,7117	0,52878
21	21:46328835	rs760459	0,5276	0,615	0,53142
12	12:10153815	rs117114115	0,3145	0,4232	0,53142
12	12:10137296	rs476844	0,5178	0,6066	0,53363
12	12:10137557	rs479499	0,5178	0,6066	0,53363
12	12:10129002	rs12298261	0,5178	0,6066	0,53363
12	12:7888555	rs10505733	0,3569	0,4632	0,53368
12	12:7890549	rs1894823	0,3569	0,4632	0,53368
12	12:10116615	rs3110950	0,4017	0,5043	0,53408
12	12:7890776	rs7311932	0,3569	0,4632	0,53408
12	12:7894188	rs73056607	0,3569	0,4632	0,53408
12	12:7896633	rs7305088	0,3569	0,4632	0,53408
12	12:7898308	rs11055567	0,3569	0,4632	0,53408

12	12:7901729	rs11055588	0,3569	0,4632	0,53408
12	12:10271701	rs56371657	0,2642	0,374	0,53408
12	12:10278489	rs11053617	0,2642	0,374	0,53408
12	12:7885652	rs7964329	0,3569	0,4632	0,53428
12	12:7895744	rs7310649	0,3569	0,4632	0,53428
12	12:7895804	rs7300836	0,3569	0,4632	0,53428
12	12:7897479	rs116928974	0,3569	0,4632	0,53428
12	12:7897735	rs6488616	0,3569	0,4632	0,53428
12	12:7898433	rs11055569	0,3569	0,4632	0,53428
12	12:10277733	rs11053615	0,2642	0,374	0,53428
12	12:7898192	rs147781690	0,3569	0,4632	0,53497
12	12:7894056	rs73056605	0,3569	0,4632	0,53698
12	12:10136672	rs2961544	0,5178	0,6066	0,53874
12	12:10130524	rs570931	0,5178	0,6066	0,54136
12	12:10159793	rs10400564	0,4207	0,5215	0,54394
12	12:10160081	rs7960084	0,4207	0,5215	0,54394
12	12:10157932	rs4764187	0,4207	0,5215	0,54394
12	12:10136199	rs1060648	0,4185	0,5195	0,54394
12	12:10247571	rs79074764	0,2988	0,4081	0,55117
7	7:141627939	rs1285935	0,4729	0,5677	0,55310
10	10:81732366	rs11200984	0,4933	0,5854	0,56193
10	10:81732349	rs11200982	0,4933	0,5854	0,56193
10	10:81706973	rs726288	0,4933	0,5854	0,56193
12	12:8666296	rs73250517	0,611	0,6855	0,56468
12	12:10313722	rs11053647	0,6903	0,751	0,56519
12	12:10129536	rs61913541	0,5974	0,6741	0,58013
12	12:10155412	rs11053543	0,4896	0,5822	0,58504
12	12:10150974	rs2273986	0,4043	0,5066	0,59127
12	12:8693807	rs7306903	0,5231	0,6112	0,59325
12	12:10112269	rs7957596	0,4687	0,564	0,59325
12	12:10149406	rs581949	0,4043	0,5066	0,59631
12	12:10152073	rs4764179	0,4043	0,5066	0,59631
12	12:10152015	rs4764178	0,4043	0,5066	0,59631
12	12:10149207	rs59400725	0,5167	0,6056	0,59821
12	12:10149204	rs35665084	0,5167	0,6056	0,59821
12	12:10132978	rs623728	0,6936	0,7538	0,59864
12	12:10126427	rs592206	0,6936	0,7538	0,59864
12	12:10131034	rs566229	0,6936	0,7538	0,60149
21	21:46341197	rs2070946	0,5028	0,5937	0,60603
12	12:10252208	rs6488258	0,5072	0,5974	0,60890
12	12:10121411	rs478829	0,6128	0,687	0,60951
12	12:10273166	rs12829123	0,5689	0,6501	0,61393
12	12:10243143	rs11053575	0,5102	0,6001	0,61677
12	12:10201646	rs4399401	0,6681	0,7328	0,61734
12	12:10242865	rs7961436	0,5102	0,6001	0,61818
12	12:10251445	rs2306894	0,5129	0,6024	0,61882
10	10:81371698	rs4253512/rs72659390	0,2242	0,3331	0,62005

12	12:10317246	rs34733039	0,4037	0,5061	0,62024
12	12:10324081	rs35553961	0,4037	0,5061	0,62024
10	10:81317064	rs17096771	0,2971	0,4064	0,62024
12	12:10144863	rs17807046	0,2317	0,3409	0,62024
12	12:10323471	rs34624528	0,4037	0,5061	0,62030
12	12:10149947	rs73050636	0,4002	0,503	0,62030
10	10:81373728	rs1059058	0,2971	0,4064	0,62030
12	12:8662809	rs7978532	0,7442	0,795	0,62044
12	12:10251498	rs148373579	0,4831	0,5766	0,62062
12	12:10231362	rs10845031	0,5102	0,6001	0,62195
12	12:10138475	rs602320	0,3735	0,4785	0,62195
12	12:8687704	rs10841846	0,5093	0,5993	0,62323
12	12:10204478	rs10505750	0,6694	0,7339	0,62378
12	12:10275336	rs4764271	0,6542	0,7214	0,62439
17	17:60742279	rs2302242	0,6163	0,6899	0,62600
12	12:10251157	rs6488256	0,6425	0,7117	0,63059
12	12:10248289	rs7970083	0,7754	0,8203	0,63065
12	12:10250366	rs3912645	0,6425	0,7117	0,63065
12	12:10250729	rs7977902	0,6425	0,7117	0,63065
12	12:10242161	rs3825300	0,9397	0,952	0,63204
12	12:10235943	rs3886143	0,6532	0,7205	0,63282
12	12:10251941	rs11053581	0,6425	0,7117	0,63282
12	12:10107191	rs4763394	0,4048	0,5071	0,63314
12	12:8608583	rs4528410	0,2166	0,3252	0,63463
12	12:10250514	rs376725596	0,6425	0,7117	0,63498
12	12:10247732	rs4764250	0,6425	0,7117	0,63498
12	12:10247358	rs4764247	0,6425	0,7117	0,63498
12	12:10251772	rs2306891	0,6425	0,7117	0,63498
12	12:10263205	rs2401601	0,5762	0,6563	0,63498
12	12:10263164	rs2401602	0,5762	0,6563	0,63498
12	12:10262275	rs2087307	0,5762	0,6563	0,63498
12	12:10263139	rs2401603	0,5762	0,6563	0,63498
12	12:10262923	rs2401606	0,5762	0,6563	0,63498
12	12:10249098	rs10845039	0,9397	0,952	0,63633
12	12:10233652	rs7980801	0,7932	0,8346	0,63633
12	12:10241829	rs2277416	0,7932	0,8346	0,63633
12	12:10263026	rs2401605	0,7264	0,7806	0,63696
12	12:10263656	rs2401600	0,7264	0,7806	0,63696
12	12:7889656	rs10845806	0,5461	0,6308	0,64048
12	12:7896785	rs117012397	0,492	0,5843	0,65333
12	12:10273625	rs73068857	0,4039	0,5063	0,67337
12	12:10184861	rs11835234	0,7172	0,773	0,67947
12	12:8624495	rs373404974	0,3975	0,5005	0,68133
12	12:10131684	rs607567	0,5178	0,6066	0,68201
12	12:10128795	rs2896048	0,5178	0,6066	0,68201
12	12:10128794	rs2401640	0,5178	0,6066	0,68201
16	16:31334236	rs11150610	0,4881	0,5809	0,68201

1	1:85744472	rs2735591	0,604	0,6796	0,68244
12	12:10157474	rs4764183	0,495	0,5869	0,68493
19	19:7805951	rs11465413	0,5965	0,6734	0,68959
12	12:10282736	rs11053623	0,4115	0,5132	0,69004
12	12:10195233	rs1844186	0,4375	0,5364	0,69454
12	12:10311962	rs10505755	0,5005	0,5917	0,70482
12	12:10324381	rs11053654	0,5299	0,617	0,70622
12	12:10316205	rs11053649	0,6903	0,751	0,70855
12	12:10312691	rs17174597	0,6903	0,751	0,70855
12	12:10313358	rs3736232	0,6903	0,751	0,70855
12	12:10313265	rs3736233	0,6903	0,751	0,70855
12	12:10312289	rs1050283	0,6903	0,751	0,70855
12	12:10312648	rs13306593	0,6903	0,751	0,70855
12	12:10311563	rs1050286	0,6903	0,751	0,70855
12	12:10312914	rs3816844	0,6903	0,751	0,70855
12	12:10140184	rs636554	0,7506	0,8002	0,71002
12	12:10279527	rs78646223	0,3147	0,4234	0,71129
12	12:10315014	rs11053648	0,6903	0,751	0,71133
12	12:10160527	rs12578560	0,6164	0,69	0,71405
21	21:46337271	rs56056043	0,6928	0,7531	0,71431
12	12:10163379	rs1359083	0,6239	0,6962	0,71578
12	12:10132283	rs7313235	0,5974	0,6741	0,71578
12	12:10163147	rs1075996	0,6239	0,6962	0,71751
12	12:10157971	rs10466840	0,6239	0,6962	0,71751
12	12:10157666	rs4764185	0,6239	0,6962	0,71751
12	12:10160068	rs7970682	0,6239	0,6962	0,71751
12	12:10126249	rs7309256	0,5974	0,6741	0,71751
19	19:7812729	rs79078188	0,6724	0,7364	0,71771
12	12:10281421	rs17807926	0,5916	0,6692	0,71771
12	12:10137886	rs1349027	0,3726	0,4778	0,71771
12	12:10163072	rs1807355	0,6239	0,6962	0,71811
12	12:10281255	rs16910631	0,3583	0,4645	0,71948
12	12:10142897	rs679982	0,7506	0,8002	0,71990
12	12:10139924	rs544783	0,6248	0,6969	0,71990
12	12:10139021	rs704230	0,7116	0,7684	0,71990
12	12:10142538	rs678208	0,6378	0,7078	0,72161
16	16:31298939	rs9937837	0,5796	0,6592	0,72161
12	12:10183166	rs61918590	0,903	0,9227	0,72359
12	12:10149206	rs60114913	0,5167	0,6056	0,72400
12	12:10124336	rs2984956	0,437	0,536	0,72471
12	12:10132878	rs623269	0,6936	0,7538	0,72872
12	12:10134961	rs519291	0,6936	0,7538	0,72872
16	16:31336888	rs1143683	0,7232	0,778	0,72942
12	12:7896427	rs10845818	0,5236	0,6116	0,73014
12	12:10275684	rs11053608	0,6842	0,7461	0,73468
12	12:8692843	rs10841856	0,8893	0,9118	0,73708
12	12:8691142	rs7139227	0,4361	0,5352	0,73801

12	12:10248964	rs7957278	0,6621	0,7279	0,73813
12	12:8688075	rs7307228	0,6429	0,712	0,73985
12	12:10196783	rs6488246	0,7962	0,837	0,74233
12	12:10196088	rs7297150	0,7962	0,837	0,74233
12	12:10199083	rs10845028	0,7962	0,837	0,74233
12	12:10202228	rs1074060	0,7962	0,837	0,74233
10	10:81319214	rs1059046/rs35576782	0,8577	0,8865	0,74241
12	12:10277083	rs3901533	0,7641	0,8112	0,74328
12	12:10276562	rs2078178/rs59180874	0,7641	0,8112	0,74328
12	12:10169041	rs12824889	0,5035	0,5943	0,74399
12	12:10122857	rs11053533	0,5877	0,666	0,74542
12	12:10132773	rs7312943	0,5877	0,666	0,74586
17	17:60720362	rs8078112	0,6225	0,6951	0,74718
12	12:10234603	rs12300167	0,7932	0,8346	0,74995
12	12:10235671	rs10772231	0,7932	0,8346	0,74995
12	12:10249358	rs11053580	0,7932	0,8346	0,74995
12	12:10230416	rs7960611	0,7932	0,8346	0,74995
12	12:10241593	rs1352477	0,7932	0,8346	0,74995
12	12:10251992	rs10845041	0,7932	0,8346	0,74995
12	12:10241553	rs5008763	0,7932	0,8346	0,74995
12	12:10237014	rs6488250	0,7932	0,8346	0,74995
12	12:10237071	rs3994131	0,7932	0,8346	0,74995
12	12:10229931	rs1948184	0,7932	0,8346	0,75148
12	12:10247461	rs4764248	0,7932	0,8346	0,75148
12	12:10236797	rs4265666	0,7932	0,8346	0,75148
12	12:10236824	rs3994132	0,7932	0,8346	0,75148
12	12:10236921	rs6416262	0,7932	0,8346	0,75148
12	12:10236942	rs6488249	0,7932	0,8346	0,75148
12	12:10240944	rs2401614	0,7932	0,8346	0,75148
12	12:10250513	rs374305068	0,7804	0,8243	0,75178
10	10:81736312	rs77043863	0,7427	0,7938	0,76399
12	12:10167277	rs637790	0,7379	0,7899	0,76548
10	10:81318848	rs370004219	0,525	0,6128	0,76548
12	12:10165469	rs476474	0,7379	0,7899	0,76560
10	10:81735981	rs3923564	0,4933	0,5854	0,76560
10	10:81706135	rs6413523	0,4933	0,5854	0,76560
10	10:81732653	rs11200985	0,4933	0,5854	0,76622
12	12:10226463	rs16910035	0,9179	0,9346	0,77653
12	12:8289226	rs4322490	0,9432	0,9547	0,78062
12	12:10170727	rs11053548	0,7513	0,8008	0,78087
12	12:8661951	rs79143728	0,5206	0,609	0,78088
12	12:8286762	rs7315526/rs77201531	0,9432	0,9547	0,78157
12	12:8288098	rs10840746	0,7064	0,7642	0,79172
12	12:8288524	rs10770233	0,7064	0,7642	0,79172
19	19:7831713	rs367752410	0,8151	0,8523	0,80414
21	21:46330628	rs2280965	0,8876	0,9104	0,80440
21	21:46322853	rs2026882	0,8814	0,9054	0,81011

12	12:10264994	rs1532087	0,6527	0,7201	0,81049
12	12:10182292	rs71450017	0,7515	0,8009	0,81168
12	12:7882695	rs17198999	0,7039	0,7622	0,81168
12	12:7882569	rs17198992	0,7039	0,7622	0,81168
12	12:7881995	rs3764005	0,7039	0,7622	0,81185
12	12:10231904	rs16910075	0,631	0,7021	0,82044
12	12:7902927	rs117331019	0,6437	0,7127	0,82522
21	21:46340843	rs2070947	0,5249	0,6127	0,83385
12	12:8276048	rs4882913	0,882	0,9059	0,83877
12	12:10157471	rs4764182	0,7484	0,7984	0,84063
12	12:8276432	rs7295783	0,882	0,9059	0,84124
19	19:7832001	rs874492	0,8092	0,8475	0,84384
12	12:10109162	rs1447877	0,8613	0,8893	0,84547
12	12:10158204	rs1868211	0,7396	0,7913	0,84603
12	12:10162511	rs185772088	0,6249	0,6971	0,85061
12	12:10312776	rs17174598	0,6903	0,751	0,85282
12	12:10313134	rs3736234	0,6903	0,751	0,85282
12	12:10313075	rs3736235	0,6903	0,751	0,85282
16	16:31302938	rs9938063	0,8462	0,8773	0,85311
12	12:10157513	rs4764184	0,757	0,8054	0,85325
19	19:7831166	rs868875	0,7002	0,7591	0,85439
19	19:7831628	rs2277998	0,7002	0,7591	0,85462
12	12:10125493	rs73048876	0,7071	0,7648	0,85588
12	12:10163375	rs1359082	0,7544	0,8033	0,85667
12	12:10121089	rs7953702	0,9663	0,9732	0,85675
12	12:8687897	rs78521210	0,53	0,617	0,85789
12	12:10183123	rs61918589	0,903	0,9227	0,85873
12	12:10154011	rs643798	0,9476	0,9583	0,85988
10	10:81317045	rs1965708	0,8012	0,8411	0,86147
12	12:10145341	rs613871	0,8827	0,9065	0,86167
12	12:10148031	rs522837	0,8893	0,9117	0,86183
12	12:10153192	rs564844	0,9476	0,9583	0,86229
12	12:10156501	rs7134176	0,9937	0,9949	0,86249
12	12:10146236	rs617956	0,7502	0,7999	0,86648
12	12:7902657	rs113639661	0,3871	0,491	0,86660
12	12:8691242	rs4562874	0,6429	0,712	0,86668
12	12:10250938	rs7313750	0,7932	0,8346	0,86673
12	12:10165593	rs620449	0,5767	0,6568	0,86845
12	12:10201778	rs1074059	0,7962	0,837	0,86864
12	12:10206925	rs7315231	0,7962	0,837	0,86864
12	12:8687812	rs10841847	0,7535	0,8025	0,86864
12	12:10207771	rs1488818	0,7962	0,837	0,86946
12	12:10215582	rs10772230	0,7962	0,837	0,86946
12	12:10204286	rs10505749	0,7962	0,837	0,86946
12	12:10131939	rs536947	0,5872	0,6656	0,86994
12	12:10170161	rs1054611	0,7171	0,7729	0,87012
12	12:10247829	rs4764251	0,6401	0,7097	0,87081

12	12:112890776	rs2301756	0,8368	0,8697	0,87087
12	12:8280908	rs2377422	0,856	0,8851	0,87107
12	12:10190537	rs10845025	0,9155	0,9326	0,87277
12	12:10186284	rs6488244	0,9155	0,9326	0,87439
12	12:10247117	rs10845034	0,937	0,9498	0,87447
12	12:10247137	rs10845035	0,937	0,9498	0,87447
12	12:10156852	rs7134681	0,9937	0,9949	1,00000
19	19:7813268	rs735239	0,9933	0,9947	1,00000
12	12:10186409	rs6488245	0,9902	0,9922	1,00000
12	12:10247544	rs3912642	0,9892	0,9914	1,00000
10	10:81372081	rs1136451	0,9839	0,9872	1,00000
10	10:54531226	rs1800451	0,9832	0,9866	1,00000
10	10:54531461	rs7095891	0,9718	0,9775	1,00000
12	12:10190953	rs10772227	0,9641	0,9714	1,00000
12	12:10189652	rs4474534	0,9641	0,9714	1,00000
21	21:46328779	rs367847454	0,9597	0,9679	1,00000
12	12:10184298	rs113211113	0,9577	0,9663	1,00000
12	12:10183263	rs11053561	0,9577	0,9663	1,00000
21	21:46344426	rs2838738	0,9575	0,9662	1,00000
10	10:81373674	rs4253526/rs148138544	0,9529	0,9625	1,00000
7	7:141627899	rs13222726	0,9524	0,9621	1,00000
12	12:7885689	rs7964345	0,9519	0,9617	1,00000
10	10:54532014	rs11003125	0,9501	0,9602	1,00000
12	12:8277556	rs11043470	0,949	0,9593	1,00000
12	12:8284656	rs11043488	0,9432	0,9547	1,00000
12	12:8287048	rs11043498	0,9432	0,9547	1,00000
12	12:10215654	rs11831360	0,943	0,9546	1,00000
12	12:10199257	rs10505747	0,943	0,9546	1,00000
12	12:10210736	rs61459404	0,943	0,9546	1,00000
12	12:10213234	rs73259794	0,943	0,9546	1,00000
12	12:10124659	rs201144260	0,9411	0,9531	1,00000
12	12:8279784	rs4883072	0,9411	0,9531	1,00000
19	19:36390123	rs16960862	0,9374	0,9502	1,00000
12	12:10248608	rs374624895	0,9351	0,9483	1,00000
12	12:10136297	rs1323461	0,9267	0,9416	1,00000
12	12:10191494	rs2054888	0,9237	0,9392	1,00000
12	12:10126141	rs770750	0,9232	0,9388	1,00000
12	12:10246314	rs6488254	0,9212	0,9372	1,00000
12	12:10246431	rs7962341	0,9212	0,9372	1,00000
12	12:10246484	rs7976945	0,9212	0,9372	1,00000
12	12:10246675	rs5012088	0,9212	0,9372	1,00000
12	12:10246711	rs952546	0,9212	0,9372	1,00000
12	12:10253808	rs60567086	0,9179	0,9346	1,00000
12	12:10224552	rs56000846	0,9179	0,9346	1,00000
12	12:10279526	rs79314785	0,9176	0,9344	1,00000
21	21:46335282	rs2838735	0,9171	0,934	1,00000
12	12:8622267	rs145656638	0,9135	0,9311	1,00000

12	12:8284102	rs77283400	0,9107	0,9289	1,00000
12	12:10280852	rs11053621	0,9088	0,9273	1,00000
21	21:46338651	rs9976299	0,9079	0,9266	1,00000
12	12:10191459	rs5011270	0,9074	0,9262	1,00000
12	12:10155706	rs1447888	0,8976	0,9184	1,00000
12	12:10280845	rs11053620	0,8975	0,9183	1,00000
12	12:10126098	rs80042268	0,8947	0,9161	1,00000
12	12:10160241	rs510630	0,8943	0,9158	1,00000
12	12:10163583	rs590429	0,8943	0,9158	1,00000
12	12:10191482	rs5011269	0,8935	0,9151	1,00000
12	12:10191483	rs5011268	0,8935	0,9151	1,00000
12	12:10202685	rs59462804	0,8755	0,9007	1,00000
12	12:10156646	rs7305054	0,8734	0,899	1,00000
12	12:10201254	rs76883345	0,8727	0,8985	1,00000
12	12:10204371	rs1488817	0,8727	0,8985	1,00000
12	12:10192809	rs78064987	0,8727	0,8985	1,00000
12	12:10203926	rs1488816	0,8727	0,8985	1,00000
12	12:10199357	rs10505748	0,8727	0,8985	1,00000
12	12:10199853	rs17807434	0,8727	0,8985	1,00000
12	12:10113258	rs999185	0,8645	0,8919	1,00000
12	12:10114476	rs1535652	0,8645	0,8919	1,00000
12	12:10117369	rs12230244	0,8645	0,8919	1,00000
12	12:10114475	rs1535651	0,8645	0,8919	1,00000
12	12:10111199	rs7314437	0,8645	0,8919	1,00000
12	12:10113662	rs6488238	0,8645	0,8919	1,00000
12	12:10110742	rs7306520	0,8645	0,8919	1,00000
12	12:10110724	rs7303131	0,8645	0,8919	1,00000
12	12:10255442	rs77549003	0,8639	0,8914	1,00000
12	12:10109449	rs1447879	0,8613	0,8893	1,00000
12	12:10109086	rs1447876	0,8613	0,8893	1,00000
12	12:10108854	rs11525545	0,8613	0,8893	1,00000
12	12:10104289	rs7957464	0,8613	0,8893	1,00000
12	12:10156825	rs7305223	0,8601	0,8884	1,00000
12	12:10271087	rs16910526	0,8509	0,8811	1,00000
12	12:10144246	rs531425	0,8475	0,8783	1,00000
16	16:31310372	rs9888879	0,8462	0,8773	1,00000
12	12:10187941	rs10845023	0,8447	0,8761	1,00000
12	12:10241554	rs61918622	0,8392	0,8717	1,00000
12	12:8283348	rs10840731	0,824	0,8594	1,00000
10	10:54531534	rs11003123	0,8239	0,8594	1,00000
12	12:10273804	rs79522375	0,8232	0,8588	1,00000
19	19:7809327	rs8105572	0,8218	0,8576	1,00000
12	12:10191746	rs10845026	0,7943	0,8355	1,00000
12	12:10246296	rs7979762	0,7928	0,8343	1,00000
12	12:8612928	rs61922366	0,7794	0,8235	1,00000
12	12:10281847	rs7311598	0,7739	0,819	1,00000
12	12:10139025	rs113815287	0,773	0,8184	1,00000

12	12:10127344	rs553104	0,7642	0,8112	1,00000
12	12:10275533	rs4764272	0,7491	0,799	1,00000
12	12:10209711	rs10505751	0,7478	0,7979	1,00000
12	12:10202639	rs11053569	0,7478	0,7979	1,00000
16	16:31276811	rs1143679	0,7434	0,7944	1,00000
12	12:10158924	rs4764188	0,7376	0,7896	1,00000
12	12:7902283	rs142503346	0,7354	0,7879	1,00000
12	12:10186622	rs4764208	0,7194	0,7749	1,00000
12	12:10191719	rs5011267	0,7172	0,773	1,00000
19	19:7831953	rs560634	0,7163	0,7723	1,00000
12	12:10109648	rs7302029	0,7159	0,772	1,00000
21	21:46322938	rs186089759	0,7126	0,7693	1,00000
12	12:10271055	rs7959451	0,7124	0,7691	1,00000
12	12:10236586	rs3887490	0,7101	0,7673	1,00000
12	12:10189624	rs190497665	0,7093	0,7666	1,00000
12	12:8283536	rs4424738	0,7064	0,7642	1,00000
12	12:8287961	rs10840744	0,7064	0,7642	1,00000
12	12:8288026	rs12099836	0,7064	0,7642	1,00000
12	12:10202239	rs111554735	0,6956	0,7554	1,00000
12	12:10124373	rs2961542	0,6424	0,7116	1,00000
12	12:10224336	rs11838264	0,6335	0,7042	1,00000
12	12:10265095	rs1001449	0,6335	0,7042	1,00000
12	12:10229281	rs11053572	0,6335	0,7042	1,00000
12	12:10231393	rs11053573	0,631	0,7021	1,00000
12	12:10158376	rs1868212	0,6239	0,6962	1,00000
21	21:46322945	rs2838732	0,6153	0,6891	1,00000
12	12:10129694	rs686148	0,5872	0,6656	1,00000
12	12:10195202	rs4237956	0,5836	0,6625	1,00000
21	21:46317126	rs3788145	0,5825	0,6616	1,00000
21	21:46337226	rs56332357	0,487	0,58	1,00000
12	12:10151179	rs7980702	0,2418	0,3514	1,00000

Appendix II: Institutional Ethical Committee Approval

Institutional Ethics Committee for Bio – Medical Research
Bhagwan Mahavir Medical Research Centre
#10-1-1, Bhagwan Marg. A.C. Guards, Hyderabad – 5000004, A.P., India
Phone: 23316057, 23497360, 23303134

Justice Bhaskar Rao
Chariman

Prof. P. P. Reddy
Member Secretary

Certificate

This is to certify that the research proposal entitled "The role of Single Nucleotide Polymorphisms (SNPs) of the Toll-Like Receptor (TLR) Signaling System and the receptors RAGE, Mincle and Dectin1 for the Susceptibility and pathogenesis in India - Functional Epidemiologic Analysis" was presented by Dr.Suman Latha, Co-investigator, Dept. of Immunology, Mahavir Hospital & Research Centre before the Institutional Ethics Committee on 11th March, 2011.

The objectives of the proposal and methodology proposed are satisfactory. The expected outcome of the project is beneficial to the study group and clinicians.

The Investigators clarified the comments made by the members of the committee.

The Investigator was advised to confine their research to the parameters mentioned in the project and not to deviate from the proposal without approval of the committee.

The committee was pleased to approve the implementation of the project at Bhagwan Mahavir Medical Research Centre subject to the review of the progress of the project from time to time.

This Certificate is issued under the seal of Institutional Ethics Committee, Bhagwan Mahavir Medical Research Centre, Hyderabad, India.

[redacted]

Prof. P. P. Reddy
Member Secretary

Appendix III: Informed consent form

BHAGWAN MAHAVIR HOSPITAL AND RESEARCH CENTRE
Mahavir Marg, Masab Tank, Hyderabad

INFORMED CONSENT

THE INDO-GERMAN GRK-1673 PROGRAM

Title of study: The role of Single Nucleotide Polymorphisms (SNPs) of the Toll-Like Receptor (TLR) Signaling System and the receptors RAGE, Mincle and Dectin1 for the Susceptibility and pathogenesis in India - Functional Epidemiologic Analysis.

Principle investigators: Ralf R. Schumann and Hortense Slevogt.
Institute of Microbiology and Hygiene, Charité-Universitätsmedizin Berlin.
Co-Investigators: Dr. Vijayalakshmi Valluri, Dr.Suman Latha, Prof. Niyaz Ahmed.
Bhagwan Mahavir Medical Research Centre, University of Hyderabad

PATIENT CONSENT FORM

Explanation of the condition: Tuberculosis is a multifactorical disease. With the availability of the complete genome sequence of M.tb hopes were raised about new therapies and interventions against the disease that still kills one person /minute in the world. Variations in the human gene are likely to regulate susceptibility or resistance to tuberculosis and disease progression. This necessitates discovery of new therapeutic interventions with through investigation of M.tb physiology in survival, persistence and replication inside host and a comprehensive understanding of the host factors involved in the disease process. The present study evaluates the importance of molecular pathogenesis to identify new targets, clinical and molecular epidemiology to understand the polymorphism in the target genes and design intervention strategies to come with new small inhibitor molecules, vaccine candidate and nutritional intervention.

Why am I being asked to take part in this study ?
You are requested to participate in this research study because either you have TB or you are free from TB.
Number of subjects expected to participate:

I will be one approximately----- subject in Mahavir PPM-DOTS who will participate in this study.
If you decide to be in this study your part would involve:
In providing 10ml blood sample. The blood will be collected using sterile needle.

Risks/Discomforts:

Since we are not giving any drug orally or by injection, there are no risks, as such, involved with your participation in the study. However, as we would be drawing blood using a sterile needle, there may be a little discomporft during pricking the needle. The

1

discomfort is only temporary. Only experts with adequate knowledge would be collecting your blood sample.

Benefits:

The patient does not directly benefit from the study, but the study now, and that of the stored blood sample, would help in the research on TB related problems and therefore may help other patients in future.

Confidentiality:

The following procedure will be followed in an effort to keep yourpersonal information confidential in this study:
Your identity would be coded, and all date will be kept in a secured, limited access location.

Your identity will not be revealed in any publication or presentation of the results of this research. As a result of being in this study, identifiable health information about you will be used, generated, and or reported for the purpose(s) outlined in the beginning of this consent form.

As such, there is certain specific information that you need to know.
Health information is any information that can be linked to you, and that relates your past, present, or future physical or mental health or condition. For the purpose of this study, your health information means your:

- Medical history
- Results of physical examination
- Laboratory (blood, urine) tests

Your health information will be accessible for use by:

- The research team for this study
- The sponsor(s) of this study-Department of Biotechnology, Govt. of India
- The institutional Ethical Committee

Access to your health information for this reason would be for an indefinite time period. All of the individuals or groups referenced above are obligated to protect the privacy of your health information.

You have the right to revoke (remove) your consent for allowing access to your health information at any time in writing. If you revoke this consent, you may no longer participate in the research activity. Revoking your consent means that all access to your identifiable up to that point may still be used.

Consent to participate:
I have read or had read to me and understood the above information before signing this consent form I have been offered ample opportunity to ask questions and have received satisfactory answers. I hereby volunteer to take part in this study. I also consent to the investigators storing a specimen of my
Blood:
Or a component of my blood suh as
Serum:
Plasma:
DNA:
For their later use as a part of this study.

Subjects Rights:

- Your participation in this study is voluntary. You do not have to be in this study if you do not want to be.
- You have the right to change your mind and leave the study at any time without giving any reasons and without any penalty.
- Any new information that may make you change your mind about being in this study will be given to you.
- You will get a copy of this consent form to keep
- You do not waive any of your legal rights by signing this consent form.

Questions about the study or your rights as a research subject

If you have any questions about the study, you may contact:

1. Medical Officer: Dr.P.S.Raju, Bhagwan Mahavir Hospital & Research Centre, Hyderabad. 500 004. Tel-040- 23497303.

2. Principal Investigator:

If you sign below it means it means that you have read(or have had read to you) the information given in this consent form and would like to be a volunteer in this study.

[redacted] April 26, 2012

Signature and name of the subject [redacted] Date& Time

Signature and name of investigator obtaining the consent Date & Time

Signature of Witness Date & Time

Appendix IV: DOTS program questionnaire

FIRST TIME VISIT

MAHAVIR PUBLIC–PRIVATE MIX DOTS

Protocol No.
1.) CUE
2.) Sputum pos
3.) Follow up.

PATIENTS HISTORY FORM

1. TB.NO .....
2. Religion ..... (Hindu / Muslim / Christian / Jain)
3. Occupation ..... (Student / Unemployed / Daily Wage / Monthly Wage)
4. Sp. Examination ..... (PPM / PVT (ORT) / Govt) Sc ..... Hc ..... Sc° .....
5. SP Culture ____ Done (+ VE / -VE / ND) S / R R INH EZ
6. Date of Registration .....
7. Category I II III
8. Disease Classification P EP
   - CODE 01=LA, 02= PLEF, 03= BONES, 04=JOINTS, 05=SPINE, 06=TB ABD, 07=GENUR, 08=PC, 09=TBM, 10=MTB, 11= OTHER

   If extra pulmonary, site (please write code).....
9. Weight WO..... W2/3..... W 4/5..... W 6/8 .....
10. Reff – by .....(PP / TBC / GOVT / OTHERS)

    a) PP code
    b) IF TBC, SOR.....
    c) Others – Family members / Old TB Patient / Neighbours

11. Personal History Part –I

a) Sex ..... (M / F)
b) Status .....(Married – pregnant / Abortion / Unmarried)
c) No of Children .....Age of last Child .....
d) No of People Living ..... No of Rooms .....
e) No of people Earning ..... Total amount / month .....
f) Habits .....(Smoking / Drinking / Tobacco / Pan)
g)Cormorbid conditions DM/HTN/HIV
h) Knowledge about TB Yes / No
i) Fuel used Fire wood / coal / kerosene / LPG Gas

Personal History Part – 2

a) Past History TB / Antibiotic ..... (Yes / No)

b) If Yes Details (Drugs/duration) _____

12. Family history

c) Relation to Pt (With Sp Code : 01- + Vc, 02- -Ve , 03 – EP):

Mother | Father | Husband | Wife | Son | Daughter | Sister | Brother | In-Laws

Code:

d) Effected person has/had ATT from _____ (PPM / Govt / Pvt)

e) Effected persons Trt Date _____ Duration (No. of Months) _____

f) Treatment outcome

13. No. of Health Providers Visited Before Mahavir _____

Details _____

a) Suspicion of TB (In Weeks Only) _____

b) First seeking medical advise. (In Weeks Only) _____

c) Direct Cost

	Provider 1	Provider 2	Provider 3
1. Fee	Rs. _____	_____	_____
2. Invest	Rs. _____	_____	_____
3. Drugs	Rs. _____	_____	_____

d) Indirect Cost

4. Transport	Rs. _____	_____	_____
5. Wages	Rs. _____	_____	_____
6. Total Cost	Rs. _____	_____	_____

e) Indebtedness (SAVING /LOAN/MORTGUAGE/SALE)

14. Investigations (ESR/X-RAY / MANTOUX/BIOSY / OTHER INVT)

a) BCG Scar - (Yes /No)

b) X-Ray Findings:

1) Right-Lung (Cavitation/Infil No Cavity / Cons./PLEEF / Normal)

2) Left-Lung (Cavitation/Infil No Cavity / Cons./PLEFF / Normal)

15. Bronchodilators (Yes / No) _____

16. 'N' hood DOTS Center Code No _____

17. INH Chemo _____

18. Remarks (Any) _____

19. Treatment outcome (CURED/TRT COMPLETED/DIED/FAILURE/DEFAULTED)

APPENDIX V: Curriculum Vitae

For reasons of data protection, the curriculum vitae is not published in the electronic version.

For reasons of data protection, the curriculum vitae is not published in the electronic version.

PUBLICATIONS

Publications

- AmpliSeq -screening of gene polymorphisms in C-type lectin receptors and their signaling adaptors reveals a common variant in MASP1 to be associated with pulmonary tuberculosis in an Indian Population. Klassert TE*, Goyal S*, Stock M, Driesch D, Hussain A, Netha R, Sumanlatha G, Valluri V, Ahmed N, Schumann RR, Carlos Flores, Slevogt H. *Frontiers in Imunology . 2018*
  * Equal contribution by the authors

- S100A12 is up-regulated in pulmonary tuberculosis and predicts the extent of alveolar infiltration on chest radiography: an observational study. Berrocal Almanza LC, Goyal S, Hussain A, KlassertTE, Driesch D,GrozdanovicZ, Sumanlatha G, Ahmed N, Valluri V,Conrad ML, Dittrich N, Schumann RR, Lala B, Slevogt H. *Scientific Reports. 2016*

- C-Type lectin receptors in Tuberculosis: What we know. Goyal S, Klassert TE,

Slevogt H. *Med. Microbiol. Immun. 2016*

- A Novel Reading Scheme for Assessing the Extent of Radiographic Abnormalities and Its Association with Disease Severity in Sputum Smear-Positive Tuberculosis: An Observational Study in Hyderabad/India. Grozdanovic Z, BerrocalAlmanza LC, Goyal S, Hussain A, Klassert TE, Driesch D, Tokaryeva V, Löschmann YY, Sumanlatha G, Ahmed N, Valluri V, Schumann RR, Lala B, Slevogt H. *PloSone 2015*

- Toll-like receptor 1 variations influence susceptibility and immune response to Mycobacterium tuberculosis. Dittrich N, Berrocal Almanza LC, Thada S, Goyal S, Slevogt H, Sumanlatha G, Hussain A, Sur S, Burkert S, Oh DY, Valluri V, Schumann RR, Conrad ML. *Tuberculosis 2015*

ACKNOWLEDGEMENTS

There are a number of people for whose support I would forever be grateful.

At work in Germany, I have to enormously thank my supervisor and doctor mother Prof. Dr. Hortense Slevogt for giving me an opportunity to work with her group in Berlin and Jena, providing me all the resources in whatever way possible and constantly pushing me to do more and better. I have learned from her to never give up and to keep on working towards the goal. I would also like to thank Dr. Tilman E. Klassert, who had been of great help for my coping up with the culture when I first arrived in Germany and without whose help I would not have been able to accomplish important parts of my thesis. Along with him, I wish to thank Dominik Driesch and Dr. Magdalena Stock for their help regarding statistical challenges that I faced during my work. My sincere thanks to Prof. Ralf Schumann who gave me the lab space in Berlin to be able to finish my work and also gave important suggestions every now and then. I also had Prof. Susanne Hartmann and Prof. Lothar H. Wieler as the members of my thesis committee and I thank them for providing me their valuable suggestions during biannual meetings. Apart from them, I also thank all my colleagues including Frauke Schreiber, Luis Carlos Berrocal Almanza, Melanie Conrad, Nickel Dittrich, Mona Binding Liermann, Annina Heinrich, Shruti Thada, Anjana Gopinath with whom I shared the lab space at one time or another and who helped me in everyday tasks in the laboratory. I must also thank the coordinators of the GRK1673 program Dr. Esther Antao, Claudia Prätor, Dr. Maliha Shah and Dr. Elisabeth Otto for their administrative assistance during the program. Also, I thank the German Research Foundation (DFG) who funded this study and provided me with financial support through scholarship.

In India, the difficult task of sampling for two and a half years would have been impossible without the great help of Dr. G. Sumanlatha and Abid Hussain. I thank my colleagues at the Mahavir hospital who were right away friendly and helpful from the first day. I express sincere gratitude towards all the tuberculosis patients and their families and other individuals who participated in this study and wish that through the contribution to the scientific community with this work, I would be able to redeem the discomfort they may have experienced during sampling procedures. I would also like to thank Dr. Niyaz Ahmed and Dr. Vijayalakshmi Valluri for letting me use the facilities in their laboratories when needed and helping me with administrative issues. I am thankful to the Council of Scientific and Industrial Research (CSIR) during my time in India who awarded me scholarship and thus, financial support during my time in India.

Lastly, I must acknowledge that I would not have been able to reach this point without the continuous support of my parents Mr. Subhash Chander Goyal and Mrs. Usha Goyal. I especially thank my brother Prince Kameshwar Goyal and Juan Sebastián Ruiz Moreno who provided me mental and emotional support whenever needed and with whom I have shared my good and bad times. I also thank all my friends who have helped me in one way or another at different times.