Mycobacterium tuberculosis Antigen Arrays for the in-depth Characterization of the Serological Response of TB Infected Patient Cohorts

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Stefanie Rimmele
aus Reutlingen

Tübingen
2012

Tag der mündlichen Prüfung: 03.07.2012

Dekan:	Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Prof. Dr. Stefan Stevanović
2. Berichterstatter:	Prof. Dr. Ulrich Rothbauer

Acknowledgements

This thesis was prepared at the Natural and Medical Sciences Institute at the University of Tübingen in the department of biochemistry headed by Dr. Thomas Joos.

Initially I would like to thank my advisor Prof. Dr. Stefan Stevanović for the supervision of this thesis. I would like to acknowledge Prof. Dr. Ulrich Rothbauer for the agreement to supervise my thesis as the second examiner.

I gratefully thank Dr. Thomas Joos and Dr. Nicole Schneiderhan-Marra for giving me the opportunity to do this thesis at the NMI and for providing me with great support throughout all work stages.

Special thanks go to Dr. Gerd Michel from FIND for providing me with valuable sample material and for the scientific support.

Many thanks also go to my friends and colleagues at the NMI for their assistance, help and for the good working atmosphere. Especially I would like to thank Hannes for writing the algorithm and supporting me in the statistical and bioinformatics analysis of my data and Jens, Michael, Nadja, Conny, Yvonne and David for a lot of fun in the lab.

Finally I thank my parents for their encouraging support throughout my life and my sister Theresa.

Table of contents

List of figures

List of tables

List of abbreviations

AcOH	acetic acid
AFB	acid fast bacilli
AUC	area under curve
BCG	Bacille Calmette-Guérin
BSA	bovine serum albumin
CBS	Carboxy Block Store
CDR	complementary determining region
CV	coefficient of variation
Da	dalton
DC	dendritic cell
E. coli	*Escherichia coli*
EDC	N-3-(dimethylaminopropyl) N'-ethylcarbodiimide
EDTA	ethylendiaminetetraacetate
e.g.	exempli gratia (for example)
EtOH	ethanol
FACS	fluorescence activated cell sorting
FCS	fetal calf serum
FP	fusion protein
gt	goat
HCL	hierarchical clustering
HIV	human immunodeficiency virus
hu	human
IGRA	interferon gamma release assay
LCB	Low Cross Buffer
LDS	lithium dodecyl sulfate
LTBI	latent tuberculosis infection
M. tuberculosis	*Mycobacterium tuberculosis*
MDR-TB	multi-drug resistant tuberculosis
MES	2-(N-morpholino) ethanesulfonic acid
MEV	MultiExperiment Viewer
MFI	median fluorescence intensity

mio	million
MOPS	3-(N-morpholino) propanesulfonic acid
ms	mouse
NAAT	nucleic acid amplification test
NaN_3	sodium azide
NHS	N-hydroxysuccinimid
Ni-NTA	Nickel-nitrilotriacetic acid
PAM	predicition analysis of microarrays
PBS	phosphate buffered saline
PE	Phycoerythrin
RD1	region of difference 1
ROC	receiver operating characteristics
rpm	rounds per minute
RT	room temperature
SAM	significance analysis of microarrays
SD	standard deviation
S/N	signal to noise ratio
SVM	support vector machine
TB	tuberculosis
Tris	Tris(hydroxymethyl)-aminomethane
TST	tuberculin skin test
V	volt
XDR-TB	extensively drug-resistant TB

Abstract

Tuberculosis (TB) is a global disease with an annual death toll of about two million people. To date, no diagnostic test is available, which enables the accurate diagnosis of the tuberculosis disease states. The serological responses of TB patients are heterogeneous, but there are efforts to identify antibody signatures to a number *of M. tuberculosis* antigens, whose abundance and presence could correlate to the disease state and may predict treatment outcome. Therefore, a whole TB proteome screen was recently performed, to identify the TB antigens which may be used in a serodiagnostic TB test.

The task of this thesis was to validate the results of the whole TB proteome screen, using a bead-based platform, and to test, whether a defined subset of these TB antigens could be used to build a serodiagnostic TB test.

Therefore, a serological assay for 112 purified TB antigens was developed and a technical validation was performed. Precision, linearity, antigen stability and coupling reproducibility were assessed.
To verify the performance of the assay and to enable a comparison between different runs reference control sera were generated by pooling samples from active tuberculosis patients.
Screening of more than 1000 human serum samples from non-TB, active TB and latent TB patients and subsequent statistical analysis were performed to identify a panel of antigens that might discriminate between different TB states.

In total, 41 differentially expressed proteins were identified, which enable active TB to be distinguished from healthy people. For the discrimination of patients with latent tuberculosis infection from healthy persons, a serological response pattern to 18 antigens could be identified. Multivariate statistical studies revealed that the heterogeneous response pattern to these 41 and 18 antigens did not correlate with clinical indices of disease status. The current results of this study do not support the hypothesis that it is easily possible to identify a serological response to defined TB antigen panels that correlate to certain disease states.

1 Introduction

1.1 Tuberculosis - background

Tuberculosis (TB) is one of the major health problems worldwide. It is estimated that about one third of the worlds' population is infected with *M. tuberculosis* and that every year nine million new cases of tuberculosis are diagnosed and up to two million deaths are attributed to the disease [1-4]. Most TB cases occur in Africa (30%), and Asia (55%, mostly in China and India, see Figure 1-1).

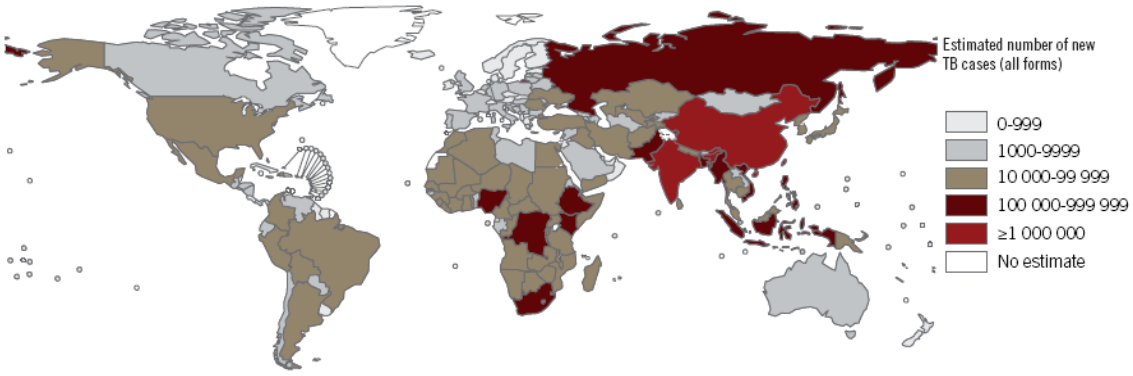


Figure 1-1: Estimated tuberculosis incidence by country [5]

The causative agent of tuberculosis is the gram-positive bacillary bacterium *Mycobacterium tuberculosis* (*M. tuberculosis*). As the growth of the bacterium is highly dependent on oxygen it is obvious that the pathogen infects mainly the respiratory system. *M. tuberculosis* was first described by Robert Koch in 1882. He received the Nobel Prize in Medicine and Physiology in 1905 "for his investigations and discoveries in relation to tuberculosis".

The genus *Mycobacterium* comprises about 100 species. The most common ones are *M. tuberculosis* and *M. leprae*.

M. tuberculosis strains are genetically diverse. This results in significant phenotypic differences between clinical isolates. Different strain lineages are associated with different geographic regions. Phenotypic studies suggest that this strain variation never had implications for the development of new diagnostics and vaccines [6].

The genome of the most common strain H37Rv was deciphered in 1998. It comprises more than four million base pairs with more than 4000 genes [7] (see

Figure 1-2). In total, 4016 open reading frames were identified in the genome, accounting for more than 90% of the potential coding capacity [8].

The huge amount of information released from this project has facilitated proteome analysis of *M. tuberculosis*. Two-dimensional polyacrylamide gel electrophoresis (2-D PAGE) was generally applied to fractions derived from *M. tuberculosis* culture filtrate, cell wall, and cytosol. For each fraction several hundred spots were found but only a small number (~20 - 40) of proteins were identified and mapped by microsequencing and immunodetection.

Culture filtrate proteins (CFPs) are well defined in terms of function as most of them contribute to the immunogenicity of *M. tuberculosis* [9]. Due to variations in the protocols used for preparation of CFPs, an absolute understanding of the protein composition of this fraction is difficult to obtain from the current literature. It varies between 12 and 38 individual proteins that could be identified with a collection of monoclonal antibodies and polyclonal sera or by molecular characterization [10-13].

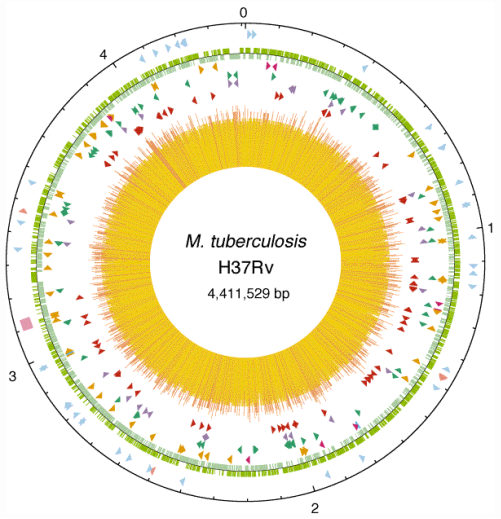


Figure 1-2: Circular map of the chromosome of M. tuberculosis *H37Rv strain [7].*

Function	No.
virulence, detoxification, adaptation	226
lipid metabolism	238
information pathways	232
cell wall and cell processes	751
insertion seqs and phages	147
PE/PPE*	168
intermediary metabolism and respiration	898
regulatory proteins	193
conserved hypotheticals	1148
unknown	15
total	4016

Table 1-1: Functional classification of M. tuberculosis *H37Rv genes [14].*

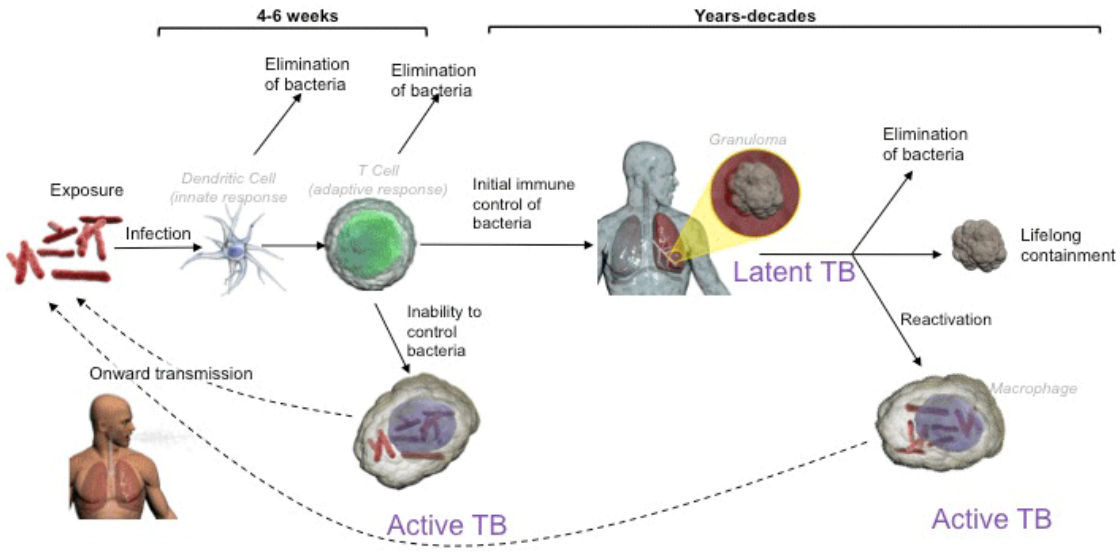
** P: Proline, E: Glutamate*

1.1.1 Pathogenesis of tuberculosis

Compared with many other diseases, the duration of the tuberculosis infection is long and there is a large variation of the response between different individuals. An infection begins with the inhalation of the tubercle bacilli. The

bacilli are often ingested by alveolar macrophages and destroyed. In addition to macrophages, dendritic cells (DC) are also readily infected with *Mycobacterium tuberculosis* [15, 16]. They are present in the airway epithelium and lung parenchyma [17] and it is likely that they play a principal role in priming anti-mycobacterial T-cell responses. At this stage, the destruction of the bacteria depends on the combination of the intrinsic microbicidal capacity of host phagocytes and virulence factor of the ingested mycobacteria. Mycobacteria that escape this initial destruction can multiply and this leads to the disruption of the macrophages. Afterwards, monocytes and other inflammatory cells are attracted to the lung and differentiate to macrophages that ingest, but do not destroy the bacteria. Within two weeks after infection, T-cell immunity with antigen-specific T-lymphocytes develops that activates macrophages to kill the intracellular mycobacteria. Several cytokines and chemokines are produced by activated macrophages and dendritic cells and are essential for the stimulation of T-lymphocytes. They produce the type I cytokines IL-12, IL-18 and IL-23 [18]. In addition, IL-1 [19] and TNFα [20] have important T-cell stimulatory properties. The production of anti-inflammatory cytokines like IL-10 and TGFβ is relevant as well [21-23]. In this stage, the disease can stay dormant or stationary. After disease progression, hematogenous dissemination takes place and the bacilli spread through the airways to other parts of the lung and the outside environment (see Figure 1-3 A and B) [24].

A



B

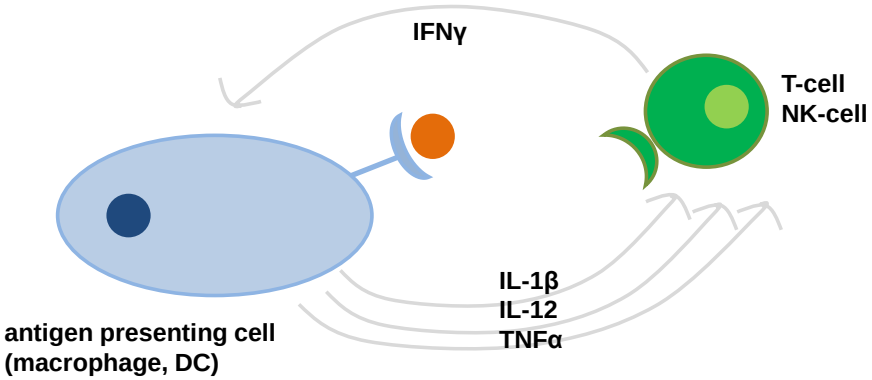


Figure 1-3: Pathogenesis of tuberculosis

A *A* Mycobacterium tuberculosis *infection starts with the inhalation of bacilli. Infection with Mtb most frequently does not directly result in active tuberculosis. Mycobacteria may be destroyed by alveolar macrophages, in which case no real infection will take place. Alternatively,* M. tuberculosis *may not be immediately killed, and so a primary complex consisting of a small infiltrate and a draining lymph node will develop. Most often, infection is stabilized at this point. In a minority of cases active disease now develops (primary tuberculosis), either in the lungs or anywhere else after dissemination of* M. tuberculosis*. Months or years afterwards, usually under conditions of failing immune surveillance, latent infection may reactivate (postprimary TB) [25].*
B *Cytokines and chemokines are involved in the type I immunity in tuberculosis.*
A major effector mechanism is the activation of M. tuberculosis*- infected macrophages by IFNγ that is produced by NK-and T-cells. Activated macrophages release several cytokines and chemokines that regulate the production of IFNγ. Adapted from [24].*

1.2 Prevention and treatment

The application of the vaccine Bacille Calmette-Guérin (BCG) has been used since 1921. Bacille Calmette-Guérin (BCG) vaccines are live attenuated strains of *Mycobacterium bovis* administered to prevent tuberculosis [26]. The vaccine provides some protection against pediatric TB, but is unreliable against pulmonary TB and its efficacy in adult TB is highly variable [6]. It is also not recommended for children infected with HIV because of the risk of disseminated BCG disease. Comparative genomic studies have documented that the evolution of BCG vaccine strains is ongoing. BCG strains differ from each other and are also different from the original BCG first used in 1921 [26]. The genetic differences affect antigenic proteins and these changes result in limited efficacy and influence the tuberculin skin test (TST) [27]. The limited efficacy of BCG is also highly variable in geographically distinct populations [6].

The low available protection affords effective treatment. TB is a not incurable disease, but effective treatment is time- and cost-intensive and even more complicated by multi-drug resistant (MDR-TB) *M. tuberculosis* strains.
While latent TB treatment usually uses a single antibiotic (mostly isoniazid) it is recommended to use a combination of at least two or three antibiotics in active TB. This reduces the risk of developing MDR-TB. The most commonly used antibiotics are rifampicin and isoniazid. Resistance of the *M. tuberculosis* bacterium against those two first-line drugs is defined as MDR-TB. In extensively drug-resistant TB (XDR-TB), the *M. tuberculosis* bacterium is resistant to three or more of the so called second-line TB drugs. Depending on the therapy regimen, therapy lasts six to twelve months [28].

1.3 Latent tuberculosis infection (LTBI)

About 90% of people who get infected with TB develop a latent TB infection, which means the infecting bacteria are alive in the body, but inactive and encapsulated. People who have latent infections do not have TB symptoms and cannot spread the infection to others, but they are at risk of developing an active infection that is both symptomatic and contagious [29]. Whereas in the general population the lifetime risk of progression from latent TB to active disease is about 10%, HIV positive persons who are infected with *M. tuberculosis* have a 30% lifetime risk of developing active TB [30]. Also children and persons who are immunosuppressed by drugs are more vulnerable developing active TB.

1.3.1 Diagnostic tests for LTBI

Tuberculin skin test (TST)

The tuberculin skin test (TST) is the standard tool for detection of latent TB. TST is based on the detection of delayed-type hypersensitivity to purified protein derivative (PPD), a mixture of antigens shared by several mycobacteria that cause a skin reaction. The test is relatively cheap and can be performed without the need for a specialist laboratory but difficulties in test administration and interpretation often leads to false results [31]. The sensitivity is substantially reduced during HIV infection [32-34]. Also, BCG vaccination and previous exposure to other Mycobacterium species can lead to false-positive results [35,

36]. The test is sometimes also used for active tuberculosis but with the same limitations.

Immune-based blood tests

In recent years new immunodiagnostic tests for (latent) tuberculosis have been developed: the commercially available Quantiferon TB Gold In Tube (Cellestis, Carnegie, Australia) and T-SPOT TB (Oxford Immunotec, Abingdon, UK). Both assays measure interferon gamma (IFNγ) release by sensitized T-cells after stimulation with peptides of *M. tuberculosis* specific antigens - Early Secretory Antigenic Target ESAT6, Culture Filtrate Protein CFP10 plus TB7.7 in the Quantiferon system [37-40]. These interferon gamma release assays (IGRA) are less influenced than the TST by factors frequently associated with childhood tuberculosis in developing countries, such as malnutrition and HIV co-infection [41-43].

1.4 Active tuberculosis

People suffering from active pulmonary disease expel infectious aerosols by cough, sneeze, speak, or spit [44]. The infectious dose of tuberculosis is very low and inhaling less than ten bacteria can cause an infection [45, 46]. The transmission is also facilitated by substance abuse, poverty, overcrowding, malnutrition, and, most importantly co-infection with human immunodeficiency virus (HIV) [44]. For people infected with HIV the risk for developing tuberculosis is 20-40% higher than for those people without an HIV infection. Estimates by the World Health Organization suggest that at least 5.6 million persons are co-infected and >1.4 mio cases of HIV-related TB will occur in 2000 [2]. The immune dysfunction caused by the HIV infection leads to a high rate of reactivation of latent tuberculosis, increased susceptibility to primary disease, and an accelerated course of disease progression [2, 47-50].

When the immune system cannot cope with the bacteria the disease becomes active. 75% of the cases are pulmonary, meaning TB in the lungs. Infected patients have symptoms like fever, dry cough, poor appetite, weight loss as well as night sweats [51, 52]. These symptoms can easily be confused with a cold or flu. In a later stage of the disease, patients suffer from chest pain, shortness of breath and coughing up of sputum.

In the other 25% of the cases the infection moves from the lungs to e.g. the brain, kidneys or the spine causing extrapulmonary TB. This occurs more commonly in immunosuppressed patients and children. In extrapulmonary TB, the symptoms depend on which part of the body is infected.

There is an arbitrary distinction between acid fast bacilli (AFB) smear positive and smear negative tuberculosis. The detection threshold for AFB smear microscopy is about 10E4 AFB/mL of sputum. The sputum bacillary load in some patients characterized as either smear positive or smear negative may fluctuate around this threshold. The infectiousness of an individual case of tuberculosis is likely to be a continuous function of the number of bacilli found in sputum such that a single patient may, over time, change from AFB smear negative to weakly smear positive and fall back to smear negative. This cycle may recur over time until, as the disease worsens, the individual becomes more firmly smear positive. During these cycles, the number of bacilli in the sputum fluctuates above and below the theoretical value of 10E4/mL needed give an AFB smear positive result [53].
During this period, the infectiousness also varies, depending upon the number of organisms in the sputum. Smear positive, culture positive cases are more infectious than smear negative culture positive cases which, in turn, are more infectious than smear negative, culture negative cases. About 15% of all cases of tuberculosis resulting from recent transmission are due to smear negative cases.

1.4.1 Diagnostic tests for active TB

Culture

Growth of the bacteria in solid media is the most sensitive (up to 98%) test that is currently available and therefore the gold standard test for active TB. However, cultivation of the bacteria lasts six to eight weeks and is more expensive than microscopy [54]. *M. tuberculosis* can be cultured from a variety of specimens including sputum, central spinal fluid (CSF), pleural effusion, bronchoalveolar lavage (BAL), gastric aspirate etc. and can thus be used to detect pulmonary as well as non-pulmonary disease. By assessing the effect of antibiotics on the cultured bacilli, this technique can also identify the antibiotic

susceptibility of the particular strain of TB infecting the patient. It is therefore the main method for identifying if a person has multi-drug resistant (MDR) TB. However, it is not always possible to obtain mycobacteria in the sample, especially in non-pulmonary TB so culture is not a sensitive test. If performed correctly it should have very high specificity and can distinguish *M. tuberculosis* from other mycobacteria.

Automated liquid culture systems are faster than traditional solid culture (six to twelve days), but a more expensive alternative compared with solid culture.

Microscopy

In low-income and middle-income countries, direct sputum smear microscopy for the detection of acid fast bacilli is the primary method for diagnosing pulmonary tuberculosis. This method is fast, inexpensive, and specific for *Mycobacterium tuberculosis* in high incidence areas. The main limitation of direct microscopy is its relatively low sensitivity (65-70%) [55], especially in individuals co-infected with HIV [56, 57] and in pediatric TB [58]. This lack of sensitivity of the sole diagnostic test in many parts of the world results in delays in diagnosis, enabling the disease to progress and increasing the potential for transmission of *M. tuberculosis* [45].

Nucleic acid amplification tests (NAAT)

NAAT is a molecular system that detects small amounts of genetic material from the micro-organism (DNA or RNA). The nucleic acid must be extracted from the clinical specimen and afterwards amplified by PCR. The PCR method is specific but not sensitive enough [31]. The main use of NAAT is not to diagnose TB per se, but to rule out infections caused by atypical mycobacteria in a sputum smear positive patient, before culture results are known.

1.5 Serological assays for TB diagnostics

Immune-based tests offer the potential to improve case detection, as some of the test formats (e.g., immunochromatographic tests) are suitable for resource-limited regions. The major advantages of immune-based tests are their speed and simplicity compared with microscopy [59]. The most common of these tests rely on detection of the serological antibody immune response to *M.*

tuberculosis as opposed to the T-cell-based cellular immune response (e.g., interferon-gamma release assays).

A number of in-house antibody detection tests have been developed but are not marketed. These tests use different antigens, distinct protocols and techniques. Currently, in developing countries, where diagnostic tests are rarely subjected to regulatory review or approval [60, 61], test manufacturers and distributors are marketing dozens of different commercial antibody detection diagnostic kits. These tests differ in a number of their features, including antigen composition, antigen source (native or recombinant), chemical composition (protein, carbohydrate, or lipid), extent and manner of purification of the antigen(s), and the class of detected immunoglobulin (e.g., IgG, IgM, or IgA). An antibody detection test can be developed into a number of formats depending on the membrane, antigen(s) coating, and incubation technique. Common designs include the enzyme-linked immunosorbent assay (ELISA) format and the immunochromatographic test format. The ELISA format has the advantages that many serum samples can be tested in parallel. The process can be completely automated, making this technique attractive in fully equipped laboratories that test a large number of samples. However, time to results still can take hours. For developing countries with limited laboratory resources and access, a simple immunochromatographic test would be the preferred method. This format requires only visual inspection of the antigen-containing lines and can be performed as a point-of-care assay without laboratory equipment. Such a test can be performed within a few minutes.

The relative importance of the individual characteristics of a diagnostic test depends upon the setting in which the test is to be performed and the intended use of the results. Technical simplicity, for example, is essential if a test is to be used in a primary health-care clinic or basic health laboratory in low-income countries. If test results are to be used to exclude a diagnosis of TB in patients with respiratory symptoms in TB-endemic countries, then tests with a high sensitivity (high negative predictive value) are required even if the test is only moderately specific. On the other hand, if a test is to be used to identify patients with respiratory symptoms in endemic countries for anti-TB treatment, a high specificity (high positive predictive value) is required.

An alternative to detection of the serological antibody immune response to *M. tuberculosis* is the direct detection of antigens in specimens other than serum (e.g., lipoarabinomannan detection in urine [62, 63] and pleural fluid [64]). The collection and processing of sputum has been a limiting factor for TB diagnosis and monitoring. Some patients, such as children, are unable to produce adequate specimens. In others, there is substantial inhomogeneity within and among repeated specimens. Urine has been viewed as a potential alternative source of diagnostic material that might be more readily available, and that might more uniformly reflect total body bacillary burden. Several studies have examined mycobacterial DNA fragments, lipoarabinomannan (LAM) [62, 65, 66] and other protein antigens in urine [67-70]. No studies have yet been performed to examine the clearance of these antigens during treatment or established a correlation to clinical outcome or to another surrogate end-point.

Until recently the tuberculin skin test (TST) has been applied for the diagnosis of latent tuberculosis. The TST uses a relatively crude mixture of TB antigens. As a result, false-positive reactions can occur because of previous Bacille Calmette-Guérin (BCG) vaccination or sensitization to non-tuberculous mycobacteria. False-negative results on tuberculin skin tests can occur because of severe illness, including active tuberculosis, or immune suppression, often due to HIV infection. With the identification of three *Mycobacterium tuberculosis* specific antigens (ESAT6, CFP10 and TB7.7) a new generation of diagnostic tests has emerged. The genes encoding these antigens are found in the region of difference 1 (RD1) of the *M. tuberculosis* genome, which is deleted from the genome of *M. bovis* BCG, and certain non-tuberculous mycobacteria, such as *M. avium*. These tests measure the interferon gamma (IFNγ) release after stimulation of whole blood with these antigens (QuantiFERON and Elispot) and are known as Interferon Gamma Release Assays (IGRA). These tests are highly specific and sensitive for the detection of TB infection [71]. However, new reports suggest a decreased sensitivity as expected in active TB patients and severe immunosuppressed persons like patients with an advanced HIV infection [72-77]. These tests are also not able to distinguish between active and latent TB infections [71]. Alternative biomarkers expressed in higher amounts could be

less susceptible to immune suppression and be used in IGRA tests. CCL2/MCP-1, CCL7/MCP-3, IL-1RA as well as the newly discovered antigens CCL8/MCP-2 and CXCL10/IP-10 seem to have diagnostic potential when compared with IFNγ. But further studies are needed to fully evaluate these markers [78, 79]. In the future it is conceivable that a combination of antigen and cytokine assays show an improved sensitivity and specificity compared with each single type of assay.

An alternative approach to whole antigen arrays is a high-content peptide microarray. These peptide microarrays represent a comprehensive set of *Mycobacterium tuberculosis* antigens in the form of linear peptide stretches and enable a detailed epitope mapping of the humoral immune response. Using these high-content planar peptide arrays, Gaseitsiwe *et al.*, 2008 [80] could segregate between TB positive and TB negative patients not by recognition of certain *Mycobacterium tuberculosis* proteins, but rather by specific peptide epitopes at different locations within the same protein. Using PAM (prediction analysis of microarrays) they identified 89 peptides differentially recognized by TB positive and TB negative individuals.
The substitution of linear epitopes of the protein antigen with short, synthetic peptides is virtually a straightforward approach to capture antigen-specific antibodies from serum samples. However, both the biologically active surface of peptides and the establishment of a well performing peptide microarray immunoassay have not yet been successfully applied on large sample cohorts.

In 2008 Antigen Discovery Inc. and the Public Health Research Institute (PHRI, USA) systematically investigated the serological response directed against the whole *Mycobacterium tuberculosis* proteome, in a rational approach towards identifying a panel of diagnostically relevant TB antigens. Using a high-throughput cloning and expression system, over 96% of the 4000 proteins in the *M. tuberculosis* proteome were cloned, expressed, and printed onto a microscopic glass slides. The TB proteome array was then screened with more than 500 sera from TB suspects worldwide and controls. These sera recognized approximately 10% of the bacterial proteome that is defined as the immunoproteome of *M. tuberculosis* which is rich in membrane-associated and

extracellular proteins. Additional analysis revealed that during active TB antibody response focused on 0.5% of the proteome enriched for extracellular proteins, relative target preference varies among patients, and response correlated with bacillary burden [81].

In order to further investigate the diagnostic potential of these immunoreactive antigens a robust and multiplexed assay system for large scale and high-output biology is needed. Such assay systems allow screening the serological reactivity of hundreds of samples against a multitude of antigens. The bead based Luminex technology belongs to such multiplexed assay systems.

The Luminex xMAP technology is a bead-based system that is based on polystyrene microspheres (6.5 µm in diameter). Each microsphere is impregnated with up to ten different amounts of three internal dyes (red, infrared and orange). The principles of flow cytometry are used as the read out.

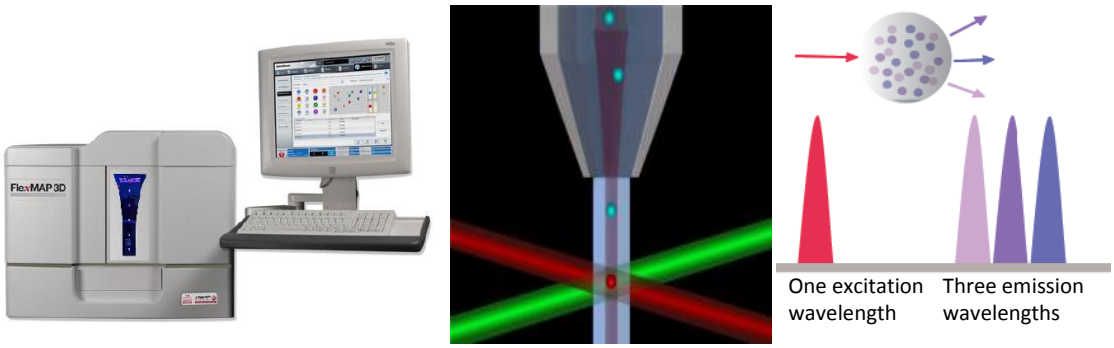


Figure 1-4: Luminex xMAP technology. Adapted from [82]

The instrument uses two different lasers. A red laser and three different photomultipliers classify the internal dyed microspheres and a green laser activates the fluorescence-labeled reporter molecules. Using the color-coding up to 500 microspheres can be individually distinguished [82]. In comparison to classical microtiter plate based ELISAs, sensitivity, accuracy and reliability of this technology are in a comparable range [83, 84].

1.6 Bioinformatic data analysis

Huge amounts of data are generated when a microarray-based screening task is performed [85]. Today, bioinformatics tools are available to effectively deal

with the analysis of high-dimensional datasets obtained from such screening methods. For example, clustering analysis to distinguish different diseases, or their respective symptoms can be useful for disease taxonomy [86]. Correct diagnosis of clusters of symptoms is also a mandatory prerequisite for supporting medical therapy [87]. In this study different data analysis tools were applied.

In machine learning programs, feature selection is used to select a subset of relevant features for learning models. Feature selection also helps people to acquire better understanding about their data by telling them which are the important features and how they are related to each other.

WEKA is a machine learning software for solving data mining problems developed by the University of Waikato, New Zealand. It contains an extensive collection of machine learning algorithms and data pre-processing methods complemented by graphical user interfaces for data exploration and the experimental comparison of different machine learning techniques on the same problem [88, 89].

Alternatively, unsupervised hierarchical clustering analysis can be performed using the MultiExperiment Viewer (MeV) [90, 91], a microarray data analysis tool incorporating sophisticated algorithms for clustering, visualization, classification, statistical analysis, and biological theme discovery developed by TIGR (The Institute for Genomic Research, Rockville, MD, USA). It is possible to generate common clustering data such as hierarchical clustering (HCL) [86]. The goal of clustering is to subdivide a set of items (in this case, serum samples from different patient groups) into a smaller number of categories. Similar items (similar serum expression patterns) fall into the same cluster, whereas dissimilar items fall in different clusters. The approach distills the complex dataset down to a more comprehensible level and can be easily visualized by the generation of heat maps. MeV also includes significance analysis of microarrays (SAM) [92], a tool that enables users to define significant parameters based on statistical analysis.

2 Purpose of the thesis

This work is part of a project driven by FIND (Foundation for Innovative New Diagnostics, Geneva, Switzerland) with the aim to develop a serodiagnostic test for tuberculosis.

Identified immunoreactive tuberculosis antigens from a whole TB proteome screen performed in 2008 [81] were provided by FIND as well as TB positive and negative human serum samples.

The aims of this thesis were:

i) Confirmation of the serological response to the TB antigens identified with the whole TB proteome screening approach.
ii) Development, optimization and validation of a robust multiplexed serological assay using a set of purified TB antigens.
iii) Screening of extensive patient and control sample cohorts.
iv) Data analysis to evaluate, whether a small number of TB antigens can be applied in a serological diagnostic test.

A graphical overview of the workflow of the whole project is given in Figure 2-1.

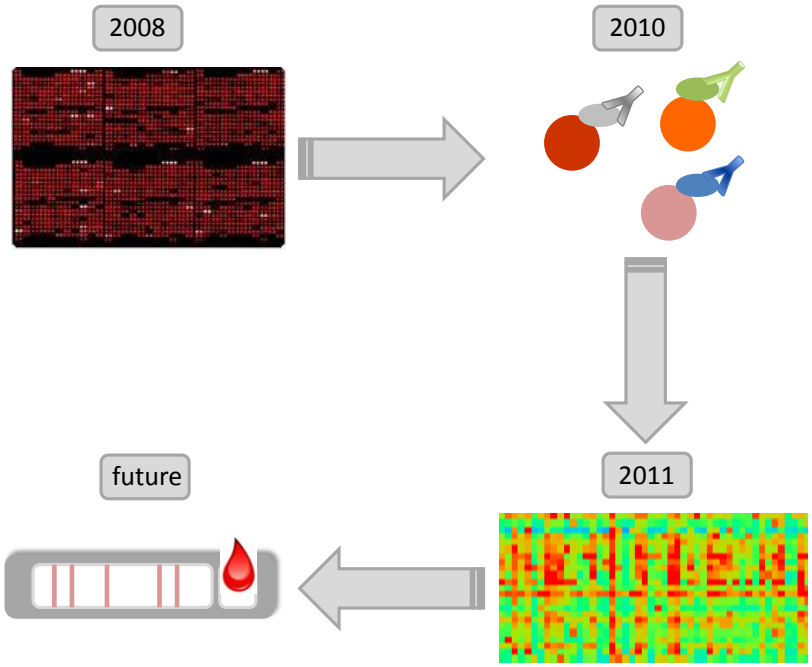


Figure 2-1: Overview of the serological assay development process.

A whole proteome screen of Mycobacterium tuberculosis *was performed in 2008 [81] and the immunoreactive proteome of* M. tuberculosis *was identified. Identified TB antigens were recombinantly expressed and purified. The task of the thesis was to use these antigens to set up a serodiagnostic test on the bead-based Luminex platform. In the future, based on the generated data, a limited number of antigens should be implemented in a low cost point-of-care test for resource limited countries with high TB incidence.*
Planar protein microarray picture adapted from [93].

3 Material and Methods

3.1 Consumables and devices

3.1.1 Antibodies

Table 3-1: Antibodies

Name	Supplier	Cat. #	Lot. #
ms-α-penta His IgG, BSA-free	Qiagen, Hilden, Germany	34660	136231967
gt-α-ms-IgG R-PE	Jackson Dianova, Hamburg, Germany	115-116-146	89764
gt-α-hu-IgG R-PE, Fc-specific	Jackson Dianova, Hamburg, Germany	109-116-098	93567, 96748 and 97773
gt-α-hu-IgG R-PE, Fab-specific	Jackson Dianova, Hamburg, Germany	109-116-097	Not available
gt-α-hu-IgG, F(ab')$_2$-specific	Jackson Dianova, Hamburg, Germany	109-005-097	98086
Human IgG	Sigma-Aldrich, St. Louis, USA	I2511	010M4840
Mouse IgG	Sigma-Aldrich, St. Louis, USA	I8765	Not available

3.1.2 Antigens

In total 112 different antigens were provided by five different suppliers.

Table 3-2: Antigens

Rv number*	symbol	description	supplier
Rv0212c	nadR	asnC-family transcriptional regulator	1 (2x**)
Rv0222		Enoyl-CoA hydratase/isomerase family protein	3
Rv0272c		hypothetical protein	1 (2x)
Rv0302		transcriptional regulator, tetR/acrR-family	1
Rv0379	secE2	protein transport protein	1
Rv0394c		secreted protein	1
Rv0440	groEL2	60 kDa chaperonin 2	1
Rv0456c	echA2	enoyl-CoA hydratase	1
Rv0577	Cfp30B	27 kDa antigen	3, 5 (3x***)
Rv0583c	lpqN	lipoprotein	1

Rv0632c	echA3	enoyl-CoA hydratase	1
Rv0798c	cfp29	29 kDa antigen	1
Rv0801		conserved hypothetical protein	1
Rv0831		Putative uncharacterized protein	2
Rv0934	pstS1	periplasmic phosphate-binding lipoprotein (PhoS1)	1, 2
Rv0944		formamidopyrimidine-DNA glycosylase	1
Rv0984	moaB2	pterin-4-alpha-carbinolamine dehydratase	1
Rv1009	rpfB	Probable resuscitation-promoting factor	2
Rv1030	kdpB	potassium-transporting ATPase B chain	1
Rv1099		Fructose-1,6-bisphosphatase class 2	2
Rv1175c	fadH	NADPH dependent 2,4-dienoyl-CoA reductase	1
Rv1196	PPE18	PPE family protein	1
Rv1242		conserved hypothetical protein	1 (2x)
Rv1284		conserved hypothetical protein	1
Rv1411c	lprG	lipoprotein	1
Rv1566c		invasion-associated protein	1
Rv1586		Probable phiRv1 integrase	3
Rv1629	polA	DNA polymerase I	1
Rv1636		Universal stress protein	3
Rv1837c	glcB	malate synthase G	1
Rv1860	apa	alanine and proline rich secreted protein (ApaC)	1, 2, 4
Rv1886c	fbpB	secreted fibronectin-binding protein antigen (Ag85B)	1, 2
Rv1926c	mpt63	immunogenic protein	1
Rv1980c	mpt64	immunogenic protein	1, 2, 3
Rv1984c	cfp21	cutinase precursor	1 (2x), 2, 3
Rv2031c	hspX	heat shock protein	1, 2
Rv2032		Putative NAD(P)H nitroreductase acg	2
Rv2094c	tatA	SEC-independent protein translocase membrane-bound protein	1
Rv2151c	ftsQ	cell division protein	1
Rv2185c		Putative uncharacterized protein TB16.3	
Rv2220		Glutamine synthetase 1	2
Rv2252		conserved hypothetical protein	1
Rv2282c		transcriptional regulator, lysR-family	1 (2x)
Rv2462		Trigger factor	3

Rv2544	lppB	lipoprotein	1
Rv2618		conserved hypothetical protein	1
Rv2870c	dxr	1-deoxy-D-xylulose 5-phosphate reductoisomerase	1 (2x)
Rv2873	mpt83	cell surface lipoprotein	1, 2
Rv2875	mpt70	major secreted immunogenic protein	1, 2
Rv2927c		conserved hypothetical protein	1
Rv2984	ppk	polyphosphate kinase	1 (2x)
Rv3050c		transcriptional regulator, asnC-family	1
Rv3243c		hypothetical protein	1 (2x)
Rv3248c	sahH	adenosylhomocysteinase	1
Rv3319	sdhB	succinate dehydrogenase iron-sulphur protein subunit	1 (2x)
Rv3326		transposase	1 (2x)
Rv3354		Putative uncharacterized protein	3
Rv3362c		ATP/GTP-binding protein	1 (2x)
Rv3376		conserved hypothetical protein	1 (2x)
Rv3495c	lprN	MCE-family lipoprotein mce4E	1
Rv3616c		conserved alanine and glycine rich protein	1
Rv3628	ppa	inorganic pyrophosphatase	1
Rv3762c		hydrolase	1
Rv3763	lpqH	19 kda lipoprotein antigen precursor	1 (2x)
Rv3775	lipE	lipase	1
Rv3804c	fbpA	secreted fibronectin-binding protein antigen (Ag85A)	1, 2
Rv3841		Ferritin BfrB	2
Rv3864		conserved hypothetical protein	1
Rv3872		Uncharacterized PE family protein PE35	3
Rv3874	esxB	10 kDa culture filtrate antigen (CFP10)	1, 2, 4
Rv3875	esxA	6 kDa early secretory antigenic target (ESAT6)	4
Rv3878		conserved alanine rich protein	1
Rv3879 aa 1-181	EspK	ESX-1 secretion-associated protein	3
Rv3881c		conserved alanine and glycine rich protein	1, 2
Rv3874-Rv3875			2, 3
Fusion		consists of 3 proteins	2

protein 1		
Fusion protein 2	consists of 3 proteins	2
Fusion protein 3	consists of 3 proteins	2
Rv2031-Rv2873		3
Antigen Mix	consists of 4 proteins	4

**The Rv number is the gene number assigned to the open reading frame. Common antigen names are in brackets.*
***Some antigens were expressed in different batches*
****Rv0577 (supplier 5) antigens of* Mycobacterium tuberculosis *and* Mycobacterium bovis *BCG strain (100% homology) were expressed in* E. coli *and* Pichia *(yeast).*

3.1.3 Patient serum samples and characteristics

Three serum sample sets derived from patients diagnosed with active, latent and non-TB were provided by FIND (Foundation for Innovative New Diagnostics, Switzerland) and PHRI (Public Health Research Institute, US).
TB suspects were defined as patients having persistent cough (more than three weeks) and at least another clinical feature indicative of active TB e.g. night sweat or fever. Active TB was defined when sputum was positive for a *M. tuberculosis* culture. Non-tuberculosis patients were TB suspects having negative AFB smears and *M. tuberculosis* cultures, chest X-ray results not suggestive of active TB, and who did not receive empiric treatment with anti-tuberculosis drugs. On mandatory two month follow-up, non-TB patients showed significant clinical and radiographic improvement, and remained AFB-smear and culture negative.

Negative control sera (assay development sample set) were collected in a low-endemicity setting from healthy persons without latent *M. tuberculosis* infection, as indicated by negative results of two tests, T-SPOT.TB and QuantiFERON, and also to tuberculin skin test, when performed.

Table 3-3: Clinical characteristics of the study population of the assay development sample set.

characteristic	active TB (n=161)	negative controls (n=17)
HIV status		
negative	161 (100%)	17 (100%)
positive	0 (0%)	0 (0%)
AFB smear		
negative	39 (24%)	17 (100%)
positive	94 (58%)	0 (0%)
n/a	28 (17%)	0 (0%)

Table 3-4: Clinical characteristics of the study population of the screening sample set.

Characteristic	non-TB (n=644)	active TB (n=592)	Latent TB (n=85)
HIV status			
negative	506 (79%)	350 (59%)	85 (100%)
positive	138 (21%)	242 (41%)	0 (0%)
AFB smear			
negative	446 (100%)	249 (42 %)	0 (0%)
positive	0 (0%)	343 (58%)	0 (0%)
n/a	0 (0%)	0 (0%)	85 (100%)
Past TB disease			
no	199 (31%)	188 (32%)	33 (39%)
yes	43 (7%)	42 (7%)	19 (22%)
unknown	402 (62%)	362 (61%)	33 (39%)

Table 3-5: Patients' parameters provided with the samples

Demographical information
Country
Sex
Age
Clinical information
Previous history of TB
BCG scar or history
HIV test
Chest X-ray

Final diagnosis
Microbiology
AFB smear
AFB smear grade
Culture
Culture grade
Liquid culture
Follow up*
Response to routine antibiotics
Repeat microbiology as above
Chest X-ray

**Diagnosis of non-TB people was confirmed in a two month follow-up.*

3.1.4 Chemicals

Table 3-6: Chemicals

Name	Supplier	Cat. #
EDC	Pierce, Rockford, USA	22981
sulfo-NHS	Pierce, Rockford, USA	24510
BSA	Carl Roth, Karlsruhe, Germany	T844.3
Tween20	Sigma-Aldrich, St. Louis, USA	P1379
MES	Sigma-Aldrich, St. Louis, USA	M2933
Triton X-100	Sigma-Aldrich, St. Louis, USA	T9284
human IgG	Sigma-Aldrich, St. Louis, USA	I2511
mouse IgG	Sigma-Aldrich, St. Louis, USA	I8765
E. coli lysate lyophilized, resolved in H_2O	ImmPORT Therapeutics, Inc., Irvine, USA	Not available
E. coli lysate	Microcoat, Bernried, Germany	601000
NaN_3	Merck KGaA, Darmstadt, Germany	822335

3.1.5 Consumables

Table 3-7: Consumables

Consumable	Supplier
Luminex MagPlex microspheres	Luminex, Austin, USA
Microcentrifuge tubes 1.5 mL	Starlab, Ahrensburg, Germany
96 half well non-binding plate	Corning, Cambridge, USA
96 well PCR plates	Thermo Fisher Scientific, Schwerte, Germany

Assay block, 1 mL	Costar, Cambridge, USA
NuPAGE 4-12 % Bis-Tris Gel	Invitrogen, Carlsbad, Germany
NuPAGE LDS Sample Buffer (4 x)	Invitrogen, Carlsbad, Germany
NuPAGE Sample Reducing Agent (10 x)	Invitrogen, Carlsbad, Germany
NuPAGE MOPS Running Buffer (20 x)	Invitrogen, Carlsbad, Germany
NuPAGE Antioxidant	Invitrogen, Carlsbad, Germany
SeeBlue Plus2 Pre-Stained Standard	Invitrogen, Carlsbad, Germany

3.1.6 Buffers

Table 3-8: Buffers

Name	**Composition**	
Coomassie staining solution	Coomassie Brilliant Blue R	0.2% (w/v)
	250	50% (v/v)
	MeOH	10% (v/v)
	AcOH	
Destaining solution 1	EtOH	50% (v/v)
	AcOH	10% (v/v)
Destaining solution 2	EtOH	10% (v/v)
	AcOH	5% (v/v)
PBS, pH 7.4	NaCl	140 mM
	KCl	2,7 mM
	$Na_2HPO_4H_2O$	10 mM
	KH_2PO_4	1,8 mM
Activation buffer, pH 6.2	Na_2HPO_4	100 mM
	Triton X-100	0.005% (v/v)
Coupling buffer, pH 5.0	MES	50 mM
	Triton X-100	0.005% (v/v)
Wash buffer for bead coupling	PBS	
	Triton X-100	0.005% (v/v)
Carboxy block store buffer (CBS)	PBS	
	BSA	1% (w/v)
Assay buffer	CBS	50% (v/v)
	LCB	50% (v/v)
Wash buffer for assays	PBS	
	Tween20	0.05% (v/v)

3.1.7 Devices

Table 3-9: Devices

Device	Supplier
Vortex Genie2	Carl Roth, Karlsruhe, Germany
Thermomixer comfort	Eppendorf, Hamburg, Germany
Ultrasonic bath	Bandelin Electronic, Berlin, Germany
Luminex FlexMAP3D	Luminex, Austin, USA
Centrifuge 5415D	Eppendorf, Hamburg, Germany
CanoScan 8800F	Canon, Krefeld, Germany

3.1.8 Software

Table 3-10: Software

Software	Supplier
Origin 7.5	OriginLab Corporation, Northampton, USA
Microsoft Office	Microsoft, Redmont, USA
xPONENT 4.0	Luminex, Austin, USA
RExcel	Statconn
WEKA	University of Waikato

3.2 Biochemical methods

3.2.1 LDS-polyacrylamide gel electrophoresis (LDS-PAGE)

Five µg of each purified antigen was mixed with LDS sample buffer (4x) and sample reducing agent (10x), denatured for 10 min at 80°C and subsequently loaded on a NuPAGE gel. A pre-stained marker was used as reference (SeeBlue Plus2 Pre-Stained Standard, Invitrogen). Running time was 50 min at 200 V (running buffer 1 x NuPAGE MOPS with 500 µL NuPAGE antioxidant).
Afterwards the gel was stained with Coomassie for 30 min and destained with destaining solution 1 and 2. A picture of the gel was taken with a scanner (CanoScan8800F, Canon).

3.2.2 Coupling of antigens on Luminex MagPlex microspheres

The automation of the coupling procedure allowed the coupling of 96 different antigens in parallel and thus improved throughput and consistency of results.
A magnetic particle handler (KingFisher 96) was used to immobilize the antigens on Luminex MagPlex microspheres. Antigens were immobilized on

magnetic Luminex beads using a bead coupling procedure based on EDC/sulfo-NHS reactions.

The bead stock was vortexed and sonicated thoroughly for at least 10 s. Three hundred µL beads from bead stock (1.25 x 10^7 beads/mL) were transferred to respective wells. Beads were washed with 250 µL activation buffer and carboxyl groups were activated with 120 µL activation buffer + 15 µL EDC (50 mg/mL) + 15 µL sulfo-NHS (50 mg/mL in water-free DMSO) for 20 min at RT with agitation. Activated beads were washed two times with 250 µL coupling buffer. Antigens were diluted to a concentration of 100 µg/mL in coupling buffer and activated beads were incubated for 2 h at RT with agitation. Coupled beads were washed with 250 µL wash buffer and resuspended in 200 µL block store buffer containing 0.05% NaN_3 and stored at 4°C until further use.

3.2.3 Coupling control

The coupling efficiency of the His-tagged antigens was controlled with a monoclonal mouse-α-5xHis antibody followed by the detection using a secondary R-PE conjugated goat-α-mouse antibody.

The mouse-α-5xHis antibody (Qiagen) was diluted in block store buffer (10 µg/mL, 1 µg/mL, 0.1 µg/mL, 0.01 µg/mL). For each antibody concentration 30 µL of the prepared bead suspension was distributed over a 96 half-well plate (1 µL of each bead sort). A plate magnet was placed under the plate, capturing all magnetic particles, and the supernatant was removed by quickly inverting the plate. Beads were resuspended in 50 µL of diluted antibody and incubated on a shaker for 45 min at RT in the dark. Beads were washed twice with 100 µL wash buffer with the plate magnet as described before. Detection was carried out using 50 µL of a PE-conjugated goat-α-mouse antibody (5 µg/mL in block store buffer) and incubating on a shaker for 30 min at RT in the dark. After washing twice with 100 µL wash buffer the beads were resuspended in 100 µL block store buffer. Read out was performed on a Luminex FlexMAP3D instrument (settings: sample size: 80 µL, time out: 60, bead count: 100 per bead type).

After a successful coupling control the antigen-coupled beads were pooled to generate a master mix.

3.3 Bead-based antigen arrays

A serological bead-based antigen array was developed for the detection of human antibodies against tuberculosis. The bead-based Luminex technology allows a miniaturized and parallelized processing of the assay.

3.3.1 General assay setup

Sample preparation

Serum samples were stored on ice. Serum dilution and incubation was performed at room temperature.

Samples were diluted in assay buffer supplemented with 10% *E. coli* lysate [81, 94] as indicated in the particular experiment. After dilution, samples were incubated for 20 min on a shaker.

Incubation protocol

The assay was performed in a semi-automated fashion using a magnetic particle handler (KingFisher Flex).

A master bead mix containing sufficient antigen-coated and control beads was prepared in assay buffer without *E.coli* lysate and distributed over a 96 well PCR plate. The beads were transferred from the bead source plate to 50 µL of the diluted human serum samples and were incubated for 2 h at room temperature. Unbound antibodies were removed by washing the beads twice with 100 µL PBS + 0.05% Tween20. Bound human antibodies were visualized with 50 µL of an R-PE labeled goat-α-human IgG antibody as reporter molecule (5 µg/mL unless otherwise indicated). Incubation time was 1 h at room temperature. After washing twice with 100 µL PBS + 0.05% Tween20, the beads were resuspended in 100 µL assay buffer without *E.coli* lysate and measured on a Luminex FlexMAP 3D instrument. A scheme of the assay is shown in Figure 3-1.

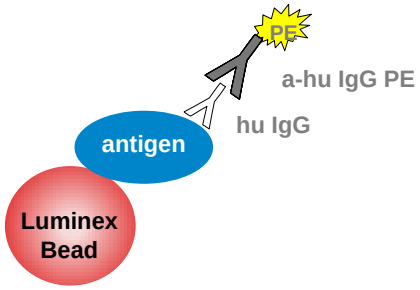


Figure 3-1: General setup for antigen assays

Beads were coupled with antigens as described in section 3.2.2. Coupled beads were incubated with serum samples and bound antigens were detected with an R-PE labeled goat-α-human antibody (see 3.3).

3.3.2 Control beads

For internal assay controls, beads were coupled with *E. coli* lysate, mouse IgG antibodies and human IgG antibodies in different dilutions. In addition an empty bead was processed in the coupling procedure.

***E. coli* coupled beads:**

For coupling of *E. coli* lysate to Luminex MagPlex beads, the lysate was diluted 1:10 in MES, pH 5.0.

Mouse IgG coupled beads:

Mouse IgG (Sigma-Aldrich, cat. no. I8765) was diluted to a final concentration of 20 µg/mL and 100 µg/mL in MES, pH 5.0.

Human IgG coupled beads:

Human IgG (Sigma-Aldrich, cat.no. I2511) was diluted to a final concentration of 20 µg/mL and 100 µg/mL in MES, pH 5.0.

Empty beads:

Empty beads were processed in the coupling procedure equal to all other beads except that they were incubated in coupling buffer (MES, pH 5.0) without any protein.

3.3.3 Generation of control samples

Negative control:

Serum of a healthy lab member served as negative control. The serum was diluted 1:200 in assay buffer + 10% *E. coli* lysate, aliquotted at 60 µL and stored at -80 C.

Quality/positive control:

The purpose of a quality/positive control is to report that all experimental steps were executed correctly in an assay experiment and to be able to compare data over a longer period of time.

[1]The screening of a set of samples $S_1, S_2, \ldots, S_n$ generated a dataset M and $m_j(S_i)$ designates the MFI for target j in sample S_i.

Prerequisite is that the assay has a predominantly linear character in the range of interest. If a pooled sample P is created from the sample S_k and S_l the MFI would be approximately additive

$$m_j(P) \approx m_j(S_k) + m_j(S_l) .$$

If a sample is diluted using a factor $\alpha \leq 1$ the MFI will exhibit a linear change:

$$\alpha m_j\ P \approx m_j(\alpha\text{-diluted } P) .$$

The threshold for a positive signal is defined by a multiple of the negative control population. The threshold for target j is designated t_j. Furthermore the decision variable x_i describes whether the sample S_i is included in the pool or not.

In this approach the number of samples is fixed to a maximum, and the number of antigens covered by the resulting QC pool is the target to be optimized. By allowing integer values for x_i, the constraint

$$\sum_{i=0}^{n} x_i = X_{max}$$

fixes the number of parts of which the pool consists, to X_{max}. E.g. a pool with $X_{max} = 5$ could consist of 3 parts S_3 and 2 parts S_7.

The set of decision variables $a_1, a_2, \ldots, a_m$, indicate whether an antigen should be covered by the pool or not. By modifying the previous coverage constraint to

$$\sum_{i=0}^{n} x_i m_j\ S_i \geq a_j t_j \mid 1 \leq j \leq m$$

it is ensured, that only if $a_j = 1$, the sum of MFIs has to exceed the threshold t_j. The term to maximize in this case is the number of covered antigens

[1] The formulas in this chapter were developed by Hannes Planatscher.

$$\max_{j=0}^{m} a_j$$

Another constraint is that the resulting pool should have the same matrix dilution as used in the normal sample preparation. If the input samples have been measured in a 1:n dilution resulting in values $m_j(S_l)$, the values need to be scaled accordingly to 1:nX_{max}. The upper bound for X_{max} is defined by the limit of dilutional linearity. E.g, if the limit is 1:2000 and the original dilution was 1:200 the maximum for X_{max} would be 10.

Algorithm[2]:

INPUT: $M, X_{maxupperbound}$
WHILE $X_{max} \leq X_{maxupperbound}$
BEGIN
$m_j\ S_l \leftarrow \frac{m_j(S_l)}{X_{max}}$;
solve ILP:
$max \quad \sum_{j=0}^{m} a_j$
subject to
$\sum_{i=0}^{n} x_i m_j\ S_i \geq a_j t_j \mid 1 \leq j \leq m$
$\sum_{i=0}^{n} x_i = X_{max}$
$a_j \in 0,1 \quad \forall j \quad 1, \dots, m$
$x_i \in \mathbb{N}_0$
$X_{max} \leftarrow X_{max} + 1$;
END

For pooling the samples indicated by the algorithm 5 µL of each of the three samples were diluted 1:200 in assay buffer + 10% *E. coli* lysate, aliquotted at 60 µL and stored at -80°C.

3.4 Data analysis

3.4.1 Scaling of data

Scaling involves the transformation of axes in such a way as to highlight features or trends in the data that might be obscured by dominant signals. To carry out analyses using MeV or for boxplots, raw MFI data were scaled prior to

[2] The algorithm was created by Hannes Planatscher.

input into the program. The raw MFI value of each antigen was log transformed to generate data values in the range of 0 to 5.

3.4.2 Box plots

Box plots are helpful in first glance qualitative interpretation and comparison of data sets. It uses the median, the approximate quartiles, and the lowest and highest data points to convey the level, spread, and symmetry of a distribution of data values. It can also be used to identify outlier data values [95].

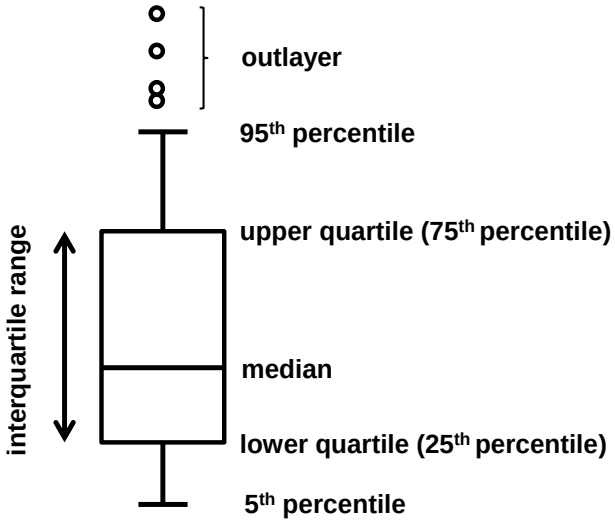


Figure 3-2: How to interpret a box plot.

3.4.3 Mann-Whitney U-test

The Mann-Whitney U-test, which is also known as the Wilcoxon rank sum test, tests for differences between two groups on a single, ordinal variable with no specific distribution [96].

The non-parametric Mann-Whitney U-test can be used for testing the null hypothesis that two samples come from the same population (i.e. have the same median). Although non-parametric, the Mann-Whitney U-test nevertheless assumes that the distributions of the datasets to be compared are similar in shape. The major difference between the Mann-Whitney U-test and Student's t-test is that the latter involves the concept of normal distribution.

3.4.4 P-values

A significance test is performed to determine the probability of obtaining the value of a test statistic (for example: t-, U-, F-, Chi-square test) given that the null hypothesis is true [97].

The p-value expresses the probability of rejecting the null-hypothesis although it is true. When the p-value is below a predetermined cut-off point, known as the significance level, it is termed significant. In clinical studies, the significance level is commonly defined as 0.05. The smaller the p-value, the more untenable is the null-hypothesis.

3.4.5 MeV and significance analysis

Statistics and clustering analyses were carried out using MultiExperiment Viewer (MeV Version 4.0). MeV is part of the TM4 Microarray Software Suite [90, 91]. TIGR MultiExperiment Viewer is one member of a suite of microarray data management and analysis applications developed at The Institute for Genomic Research (TIGR, www.tigr.org). Tools, services and detailed MeV user manuals can be accessed at www.tm4.org/mev.html.

The SAM (significance analysis of microarray) tool of the MeV software is applied for finding key genes in a set of microarray experiments. The input to SAM is in the form of expression measurements from microarray data, as well as the response variable from each experiment.

4 Results

4.1 Analysis of the quality of TB antigens

The TB antigens for the multiplexed serological assay setup were provided by five different suppliers. All antigens were expressed as His-Tag recombinant proteins in *E. coli* (one antigen, Rv0577 from supplier 5, was expressed in yeast).

The expression yield and the level of purity of all available 112 recombinant antigens was monitored by LDS-PAGE and subsequent Coomassie staining as a quality control prior bead coupling. The amount and concentration of the different antigens differed considerable as well as the degree of purity. Purity of the antigens was defined based on the Coomassie staining of the LDS gels. Good purity was defined when only minor additional protein bands were visible on the SDS gel, moderate purity when more and stronger additional protein bands were visible and a poor purity level revealed many additional and strong protein bands.

Antigens provided by supplier 4 were excellently purified (no additional protein bands visible). Antigens provided by suppliers 2, 3 and 5 revealed good purity (only minor additional bands). Antigens from supplier 1 showed different degrees of purity depending on the provided antigen batch. In total about 70% of all 112 recombinant antigens showed good to excellent purity (for an overview see Table 4-1).

Table 4-1: Overview of antigens with good/excellent purity

Supplier	n (%)
1	30/71 (42)
2	14/21 (67)
3	10/13 (77)
4	4/4 (100)
5	2/3 (67)

Protein concentrations were given by the respective antigen supplier. All images of Coomassie stained gels are shown in the supplementary material section (Figure 9-1).

4.2 Serological assay development

4.2.1 General assay development procedures

Coupling control

All 112 available recombinant antigens were successfully immobilized on Luminex MagPlex microspheres using active ester chemistry.

Coupling control was performed using a mouse-α-His antibody followed by detection with goat-α-mouse-PE. Depending on the antigen, signal intensities between 2071 and 93562 MFI (mean MFI 36910) were achieved. 80% of the antigens revealed signals between 10000 MFI and 50000 MFI (see Table 4-2). Seven antigens showed signals <10000 MFI. The reason for these low MFI signals could be due to low protein expression or poor purity of the antigens. However, for proteins Rv2874-Rv3875, Fusion protein 1 and Rv2031 from supplier 2, as well as Rv2462 and Rv3872 from supplier 3, only a strong single band was visible on the Coomassie stained gel. Many additional bands were seen on the gel for protein Rv3864 from supplier 2 but the corresponding antigen was present in a high amount. Only protein 3878 from supplier 1 revealed a low expression rate with an additional strong band at a wrong molecular weight (see SDS-PAGEs in Figure 9-1).

Proteins with a lot of additional bands visible on the Coomassie stained SDS gel had no negative influence on the coupling procedure.

Table 4-2: Overview of the results of the coupling control

	no. of antigens
<10000 MFI	7 (6%)
10000 MFI - 50000 MFI	98 (80%)
>50000	18 (15%)
mean MFI	36910
range MFI	2071 - 93462

Internal control beads to monitor the assay performance

As internal assay controls, beads were loaded with the following reagents:

i) the same *E. coli* lysate as used for sample blocking to visualize the binding of human antibodies to *E. coli proteins*.

ii) different concentrations of human IgGs for monitoring the performance of the RPE-labeled goat-α-human IgG detection antibody.

iii) mouse IgGs for monitoring the coupling control procedure with mouse-α-His and RPE-labeled goat-α-mouse antibody.

iv) empty beads as controls for unspecific binding of assay components to the polystyrol carboxy beads.

Low signals were observed with 10 µg/mL mouse-α-His and 5 µg/ml goat-α-mouse-PE antibody on the empty control beads (192 MFI) as well as the beads loaded with *E. coli* lysate (421 MFI). Cross-reactivity of the mouse-α-His detection antibody to huIgGs is observed on the beads immobilized with 20 µg/mL and 100 µg/mL human IgG (1800 MFI and 5000 MFI, respectively). Beads immobilized with mouse IgG showed the expected high signals of ~26000 and ~32000 MFI, respectively (see Figure 4-1).

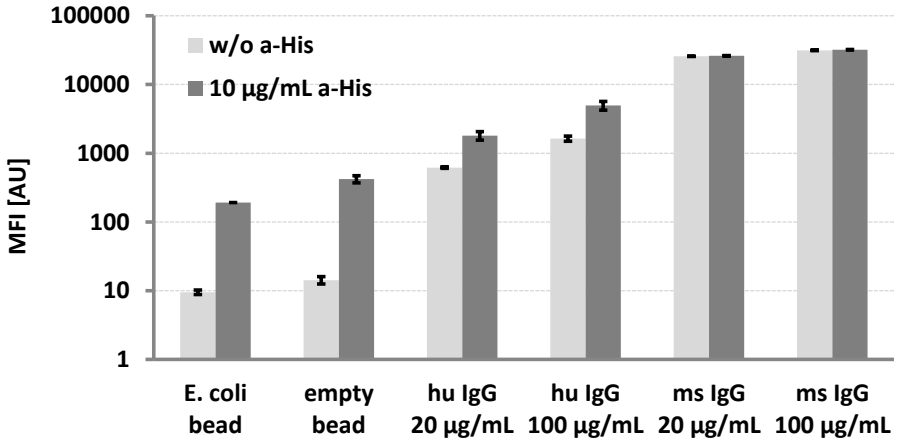


Figure 4-1: Monitoring of internal assay control beads.

Immobilized control beads were incubated with mouse-α-His (0, 0.01, 0.1 and 10 µg/mL) and 5 mg/mL RPE-labeled goat-α-mouse antibody. Shown are the results of the incubation with 0 and 10 µg/mL mouse-α-His antibody.

Assay buffer optimization

The recombinant tuberculosis antigens were expressed in *E. coli,* which means that *E. coli* host cell proteins were still present to a certain extend in the purified fusion proteins. The immune system of human beings might be confronted with naturally occurring *E. coli* bacteria and subsequently could generate antibodies against *E. coli* proteins. Incubation of such types of anti-*E. coli* antibodies present in human serum samples with on beads co-immobilized *E. coli* proteins could lead to false positive signals. Thus different assay buffer conditions were evaluated to minimize unspecific and cross reactivity signals.

For assay buffer optimization 16 randomly chosen human serum samples (TB and non-TB) were pre-incubated with assay buffer (PBS + Low Cross Buffer + 5% BSA) only and assay buffer supplemented with 10% and 20% *E. coli* lysate. Figure 4-2 depicts a comparison of sample pre-incubated with different buffer compositions. 1:200 dilutions of the serum samples and incubation in assay buffer without *E. coli* lysates can generate signals of more than 10000 MFI on the *E. coli* lysate coupled beads. In contrast, a 20 min pre-incubation of serum samples in assay buffer supplemented with *E. coli* lysate significantly reduced signals on the *E. coli* bead.

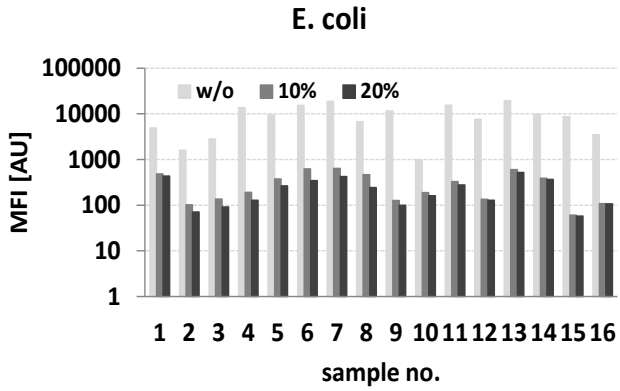


Figure 4-2: Comparison of different buffer compositions on the antibody reactivity to E. coli *lysate immobilized on beads.*

The results gained with the 20 min pre-incubation of the serum samples with different assay buffer compositions are depending on the immobilized antigen and the serum sample used.

In Figure 4-3 examples of two antigens, Rv3881 and Rv1886, are depicted. In the case of antigen Rv1886 most samples revealed considerably lower signals when pre-incubated with 20% *E. coli* lysate compared with 10% or w/o *E. coli* lysate. For antigen Rv3881 a pre-adsorption of the antibodies in the samples

with *E. coli* lysate had no influence on the reactivity of the samples with minor exceptions, e.g. sample no. 9 where the signal of the sample without pre-incubation revealed a higher signal compared with the pre-adsorbed sera. This pattern was observed for most antigens.

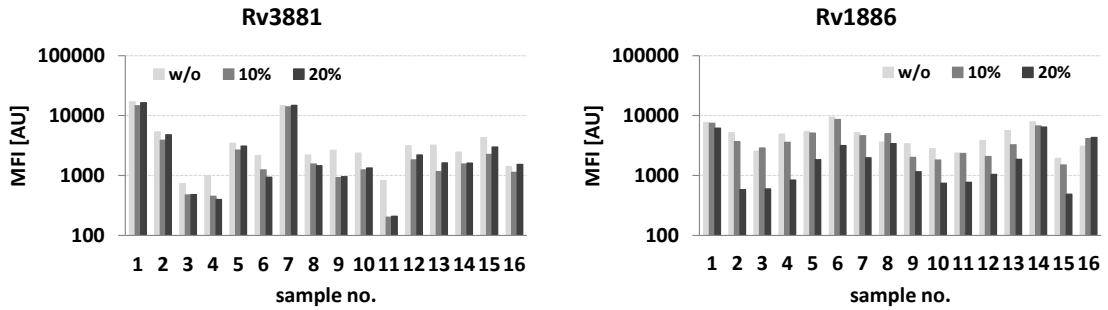


Figure 4-3: Influence of the pre-incubation of serum samples with different buffer compositions on the serum reactivity towards the immobilized antigens.

It seems that a pre-adsorption of the human serum sample with a buffer containing *E. coli* lysate had only a minor influence on the overall signal intensity and also on low signals <500 MFI. However, regarding the signals in Figure 4-2 when no pre-adsorption was performed, a buffer with 10% *E. coli* lysate was chosen for sample incubation.

Optimization of detection antibody concentration

The optimal detection antibody concentration for the detection of bound human IgGs has a major influence on the performance of the assay. Usually the detection antibody concentration needs to be increased when the overall amount of antigens in the assay is increased. However, to minimize unspecific binding of the detection antibody and to optimize the signal to noise ratio (S/N, see next section for the definition) the antibody concentration should be kept as low as possible.

Three different detection antibody concentrations were tested (2.5 µg/mL, 5 µg/mL and 10 µg/mL) on six active TB and one negative control serum at a dilution of 1:200. The optimal detection antibody concentration was determined using an 85-plex assay. The quotient of the signal of a TB positive serum sample (termed as signal S) divided by the signal of a negative reference serum

(termed as noise N) measured in an identical dilution was defined as the signal to noise ratio (S/N).

The number of maximum S/N for each antigen (n = 85) was evaluated over all samples (n = 6) and the result is depicted in Figure 4-4.

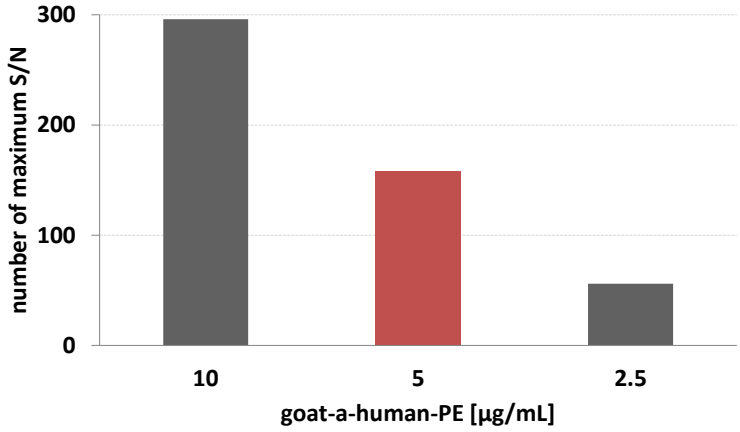


Figure 4-4: Number of maximum S/N for all 85 antigens and six samples for the tested detection antibody concentrations of 2.5 µg/mL, 5 µg/mL, and 10 µg/mL.

Best signal to noise ratios were achieved with 10 µg/mL of secondary goat-α-human-PE antibody. However, with respect to large sample screenings and thus high antibody consumption, the lower detection antibody concentration of 5 µg/mL was selected.

Determination of optimal serum dilution

It is necessary to determine the optimal serum dilution to avoid false positive or false negative results.

The optimal serum sample dilution was figured out for the 112-plex assay. Samples with different signal intensities for the antigens were chosen based on previous results. A two-fold dilution series from 1:25 to 1:3200 of seven selected active TB serum samples and the 1:200 ready-to-use positive control pool (sample no. 6) was performed. The negative control sample (sample no. 7) was prepared by initially diluting the undiluted serum in assay buffer supplemented with 10% *E. coli* lysate 1:25 and was subsequently further diluted as described.
The log2 MFI values of the dilution series for all samples and antigens is depicted as a heat map in Figure 4-5. In general, a linear dilution of the samples is seen. For a small number of antigens (Rv3874-Rv3875, FP1, FP2, and FP3) in sample no. 4, only a slight signal decrease is observed even with the highest sample dilution factor of 1:3200.

Based on high human IgG concentrations for the lower dilutions of control sample no. 7, a high rate of (false) positive signals was observed.

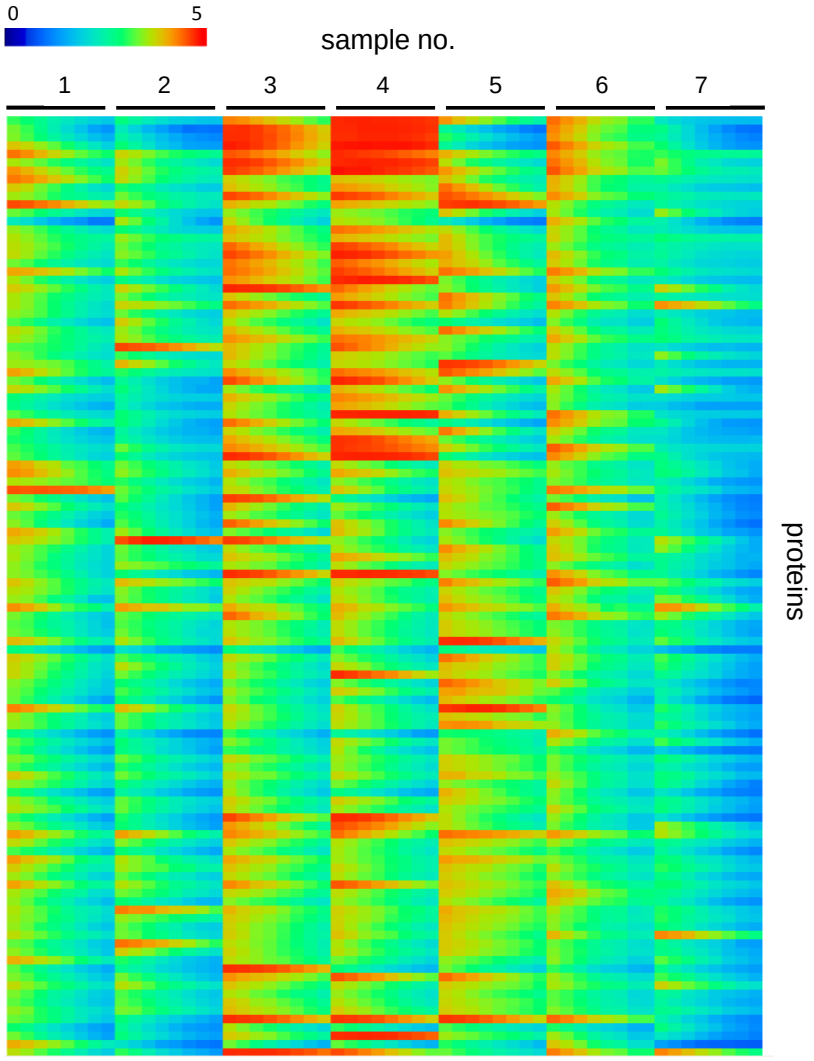


Figure 4-5: Serum sample dilution series.

Dilutions for each sample from left to right 1:25, 1:50, 1:100, 1:200, 1:400, 1:800, 1:1600 and 1:3200. Sample no. 1-5: TB positive samples, 6: positive control pool, 7: negative control. Log-transformed signal intensities are visualized as a rainbow spectrum as indicated (blue: low signal, green: mid signals, red: high signal).

Signal to noise ratios were calculated by dividing the signal of each of the TB positive serum samples by the signal of the negative control serum sample at the same dilution factor for each of the 112 individual antigens. The number of maximum S/N was counted over all six samples and 112 individual antigens and the result is depicted in Figure 4-6.

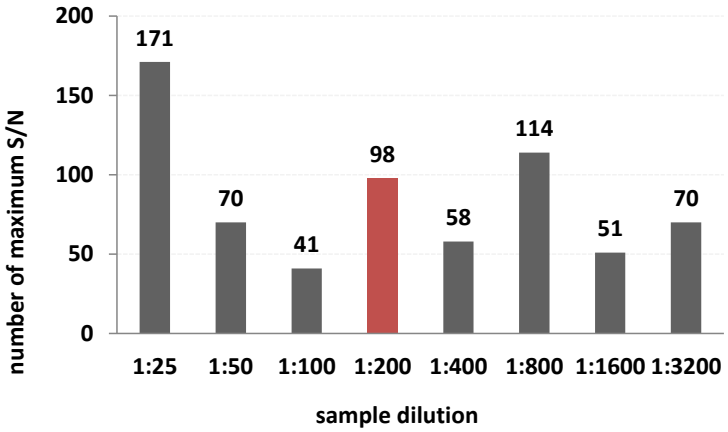


Figure 4-6: Number of maximum signal to noise ratios for all 112 antigens and six serum samples for defining the optimal sample dilution.

Most maximum signal to noise ratios were counted for the serum sample dilution of 1:25. However, with a low sample dilution and thus a high concentration of human IgGs, the risk of measuring unspecific signals is comparatively high. In addition, a low dilution of e.g. 1:25 for healthy donors leads to false positive signals (sample no. 7). Taking a strong dilution of 1:800 of the samples, there is a high risk of generating false negative results. Thus, for all further assay development and sample screenings a serum sample dilution of 1:200 was chosen.

4.2.2 Assay validation

After this first phase of assay development and optimization, assay specific quality parameters were determined. The performance of the assays was assessed based on the intra- and inter-assay CVs.

Precision

Assay precision is indicated by the coefficient of variation (%CV) which is the standard deviation divided by the mean and expressed as a percentage value. Whereas intra-assay variation measures the variability between replicate measurements in the same experiment, inter-assay variation reveals the variation between different experiments. Intra-assay coefficients of variation were calculated on the basis of eight replicates of nine different TB positive serum samples. Samples with low, medium and high signals were chosen based on the results of previous experiments. The technical replicates were generated from a single source. For the negative control one aliquot per replicate of the ready-to-use 1:200 dilution (see section 3.3.3) was used. Inter-

assay coefficients of variation were generated in triplicates in three different runs in addition to the eight replicates of the intra-assay CV experiment.

The results for intra- and inter-assay CV are summarized separately by antigen supplier in Table 4-3 and Table 4-4.

Intra-assay CVs varied strongly between <10% and 70% (mean about 15% to 20%) for all nine tested serum samples and all antigens. Inter-assay CVs are in a comparable range (mean about 30%). The wide range of intra- and inter-assay CVs between the different antigens suggests an antigen dependent effect. CVs are also dependent on the overall signal intensity. Low MFI values (<1000 MFI) usually give higher CVs than high MFI values (>5000 MFI). CVs of control beads were <10%. Intra- as well as inter-assay CVs of the antigens from different suppliers were in the same range for each of the nine tested samples.
Centrifugation of the serum samples as first step of the assay processing to spin down possible IgG cluster did not improve the assay (data not shown).
In general, CVs <20% are preferable. In this high multiplex assay with 112 antigens, this variation of CVs is in an acceptable range but should be considered when selecting the set of antigens for future screenings.

Table 4-3 Intra-assay coefficients of variation for nine active TB serum samples

	sample no.								
CV [%]	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
mean antigens									
supplier 1	18	16	6	30	4	6	9	22	34
supplier 2	14	17	3	25	2	5	10	25	38
supplier 3	15	15	5	26	2	7	8	22	13
control beads	2	1	1	2	2	1	1	2	2
range antigens									
supplier 1	3 - 30	2 - 25	1 - 12	24 - 41	1 - 9	1 - 19	2 - 22	1 - 54	4 - 68
supplier 2	3 - 24	5 - 24	2 - 7	18 - 32	1 - 4	2 - 12	2 - 26	9 - 51	8 - 24
supplier 3	4 - 28	8 - 29	2 - 11	18 - 40	1 - 5	2 - 18	4 - 14	8 - 67	3 - 26
no. of antigens < 10%									
supplier 1	13	11	67	0	71	63	49	8	4
supplier 2	6	1	21	0	21	20	13	2	0
supplier 3	5	3	12	0	13	10	10	2	0
no. of antigens 10% - 20%									
supplier 1	20	42	4	0	0	8	21	26	9
supplier 2	10	15	0	2	0	1	7	5	15
supplier 3	4	8	1	13	0	3	3	6	2
no. of antigens > 20%									
supplier 1	38	18	0	71	0	0	1	37	58
supplier 2	5	5	0	19	0	0	1	14	2
supplier 3	4	2	0	13	0	0	2	5	11

Table 4-4: Inter-assay coefficients of variation for nine active TB serum samples

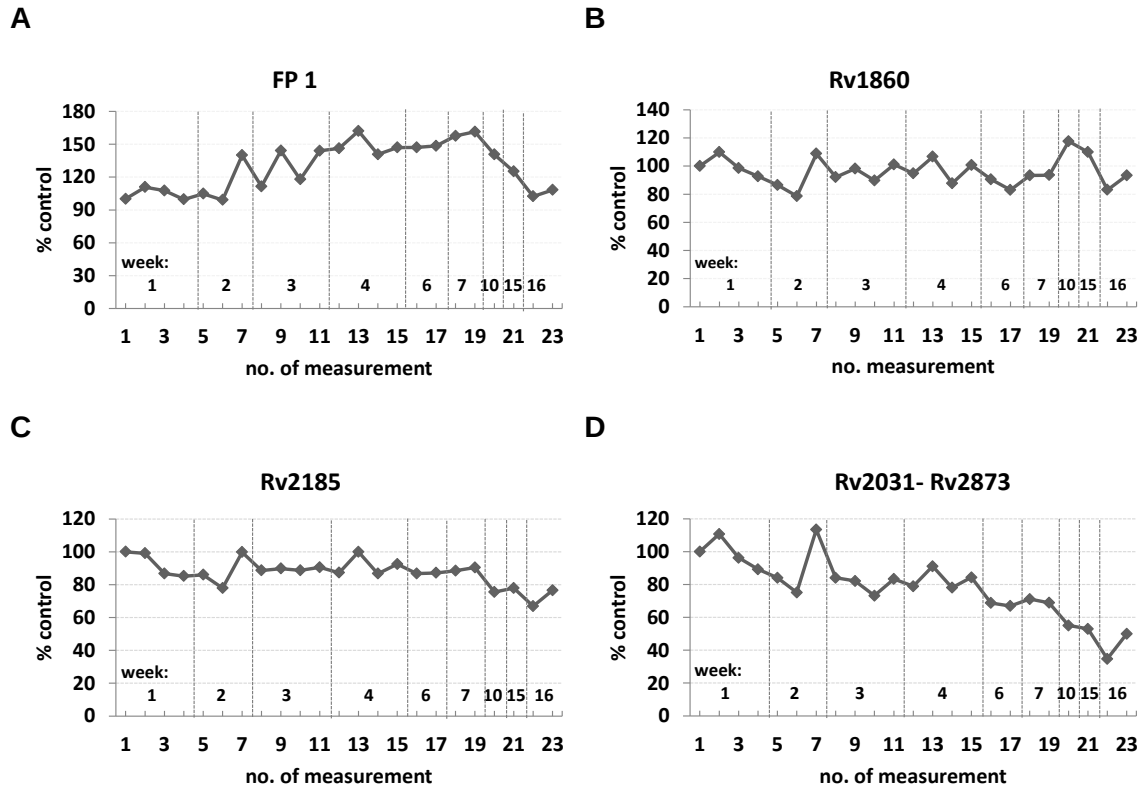
	sample no.								
CV [%]	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
mean antigens									
supplier 1	25	27	22	54	15	25	29	27	26
supplier 2	20	27	13	47	5	18	28	26	28
supplier 3	21	27	17	50	6	23	25	23	27
control beads	7	3	3	6	3	3	3	3	3
range antigens [%]									
supplier 1	8 - 34	15 - 37	4 - 50	42 - 74	2 - 27	4 - 49	10 - 55	4 - 42	9 - 47
supplier 2	10 - 28	14 - 34	5 - 30	41 - 58	2 - 9	7 - 44	9 - 52	9 - 40	12 - 43
supplier 3	13 - 33	17 - 44	10 - 27	40 - 70	2 - 14	11 - 45	14 - 37	10 - 51	16 - 42
no. of antigens < 10%									
supplier 1	4	0	6	0	10	3	1	2	1
supplier 2	1	0	10	0	21	7	1	1	0
supplier 3	0	0	1	0	12	0	0	0	0
no. of antigens 10% - 20%									
supplier 1	11	9	21	0	53	15	5	4	17
supplier 2	10	3	7	0	0	8	4	3	4
supplier 3	6	2	8	0	1	5	5	6	4
no. of antigens > 20%									
supplier 1	56	62	44	71	8	53	65	65	53
supplier 2	10	18	4	21	0	6	16	17	17
supplier 3	7	11	4	13	0	8	8	7	9

Stability of coupled antigens

For large sample screenings it is desirable to have antigen-coupled bead sets that are stable over a longer period of time to achieve comparable results.

Recombinant His-tagged proteins immobilized on beads were visualized in each assay with a mouse-α-His antibody followed by the detection using an RPE-labeled goat-α-mouse antibody. The signals obtained in the first measurement were taken as the reference signal for all consecutive measurements. The relative signal intensities for six representative antigens (two of each of the suppliers 1-4) are shown over a time period of 15 weeks (Figure 4-7). Table 4-5 summarizes the relative signal intensities for four antigen suppliers over a time period of seven weeks.

For antigens from supplier 1, the range was 36% - 283% (mean 92%), for supplier 2 65% - 162% (mean 94%), for supplier 3 64% - 127% (mean 86%) and for supplier 4 70% - 108% (mean 92%). For most antigens the remaining signal intensity was between 70% and 130% compared with the first measurement. Immobilized antigens on beads were stable over at least a time period of seven weeks. A considerable signal loss due to lower reactivity of the human IgGs against the TB antigens was seen after ten weeks. No correlation was found between the overall MFI signal intensity generated with the anti-His antibody and the stability of the antigens.



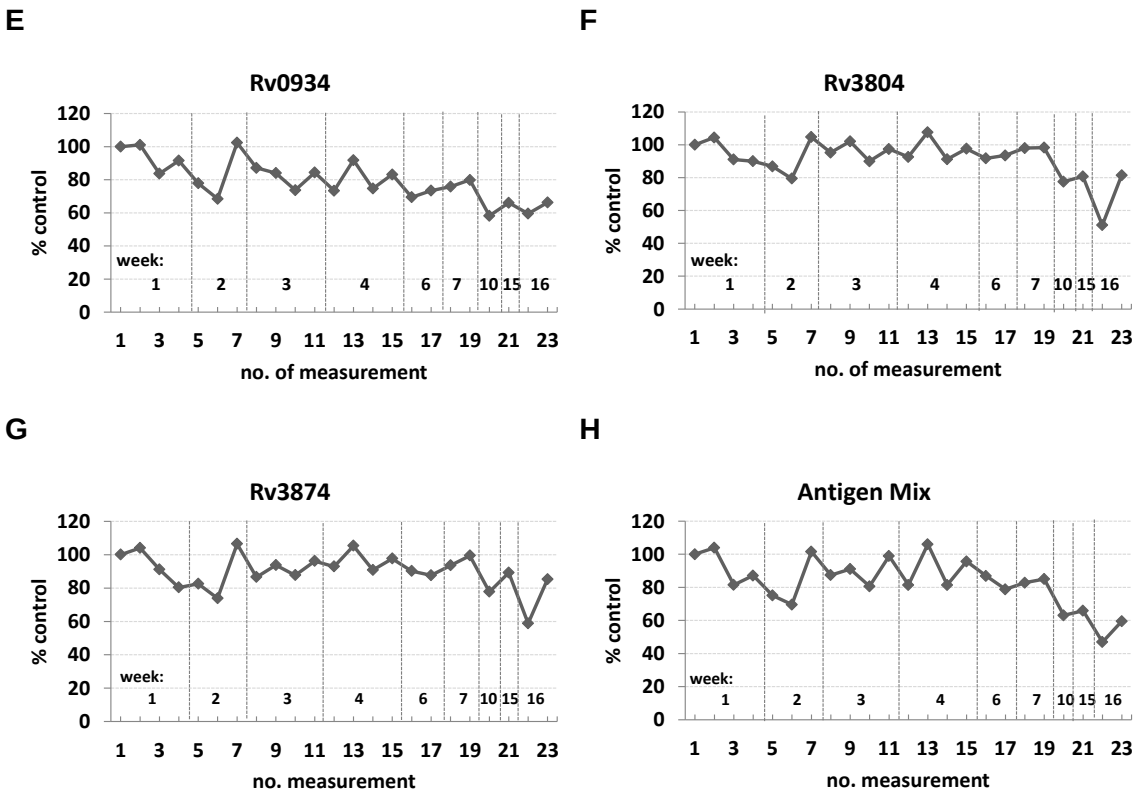


Figure 4-7: Stability of antigens immobilized on beads – control with anti-his antibody.

***A**, **B** supplier 2; **C**, **D** supplier 3; **E**, **F** supplier 1 and **G**, **H** supplier 4. Antigens Rv2185, Rv3874, Rv3804 and Rv1860 were stable over 16 weeks. For Rv0934, Antigen Mix and Rv2031-Rv2873 a constant signal decrease over the time period is observed. The fusion protein 1 delivered unstable results throughout the 16 weeks.*

Table 4-5: Summary of antigen stability over seven weeks.

	supplier 1	**supplier 2**	**supplier 3**	**supplier 4**
mean [%]	92	94	86	92
range [%]	36 - 283	65 - 162	64 - 127	70 - 108

Depending on the antigen, the signal intensities generated with serum samples varied over a time period of seven weeks. For example, in week six the values gained with most samples were much lower for most antigens compared with the values of other measurements. On the other hand the antigen mix showed increased signals for certain measurements. These effects are not attributed to a decrease in antigen stability but rather an assay specific effect (see chapter Precision).

For a graphical depiction MFI values instead of percentages were indicated to see the range (low or high) of MFI values. In general, these fluctuations do not change the classification as positive or negative sample.

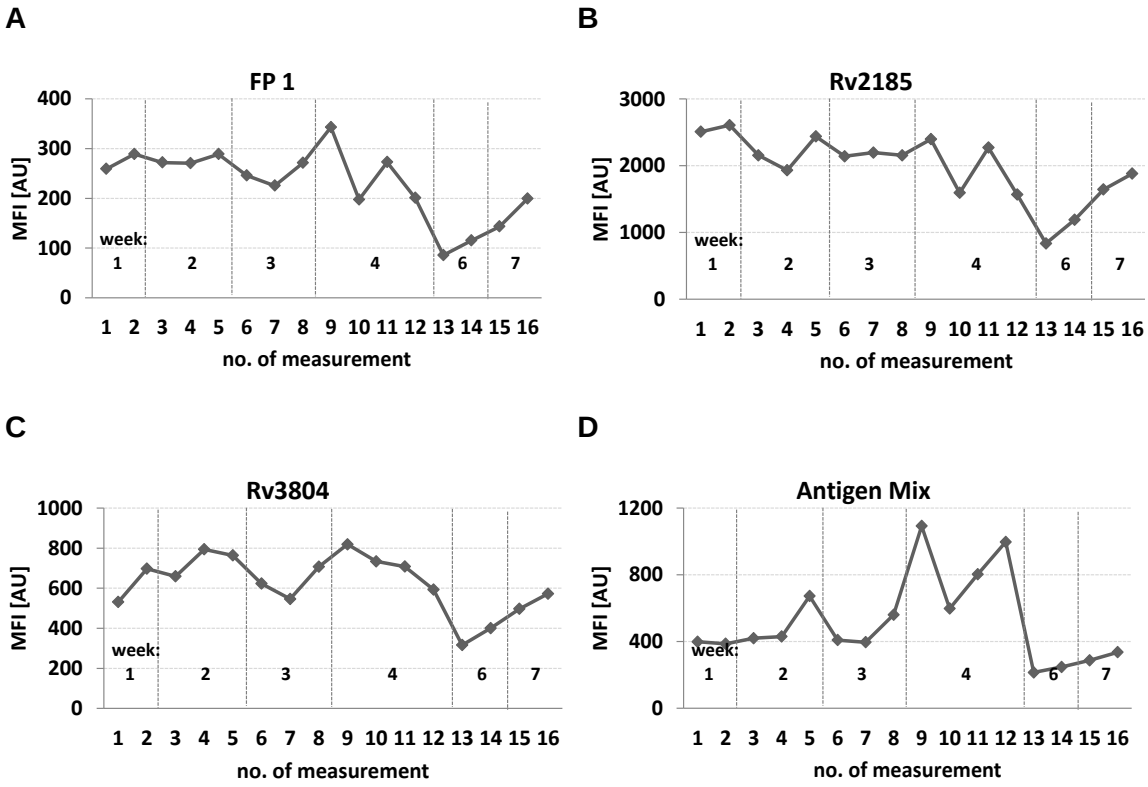


Figure 4-8. Stability of antigens immobilized on beads – IgG detection in one serum sample.

***A** supplier 2; **B** supplier 3; **C** supplier 1 and **D** supplier 4. No signal decrease is observed over the time period of seven weeks except of the values gained in week six. The antigen mix revealed increased values for some measurements.*

Assay reproducibility

Prerequisite for the successful implementation of a research and diagnostic assay is to ensure the reproducibility of an assay. The training sample set consisting of 160 active TB patients and 17 negative controls provided by FIND for assay development and validation was measured with two independently coupled bead sets with a time lag of approximately eight months.

Correlation coefficients were calculated for all immobilized antigens used in both screenings. Graphs of six exemplarily chosen antigens from three different suppliers are depicted in Figure 4-9 and the coefficient of determination is indicated in each graph.

The MFI signals for the first screening and bead set were in general lower than the MFI values for the second screening and bead set.

The correlation coefficients are summarized as a histogram in Figure 4-10. All correlation coefficients varied between 0.09 and 1.0 and are shown in

Table 9-1 in the supplementary material.

In summary, the assay seems to be highly reproducible. More than 50 % of all antigens have r > 0.7 or 0.8 respectively. This proves a good correlation of the two measurements with the two different bead sets. Antigens that cannot be coupled reproducible should be excluded from further studies or additional optimization will be required.

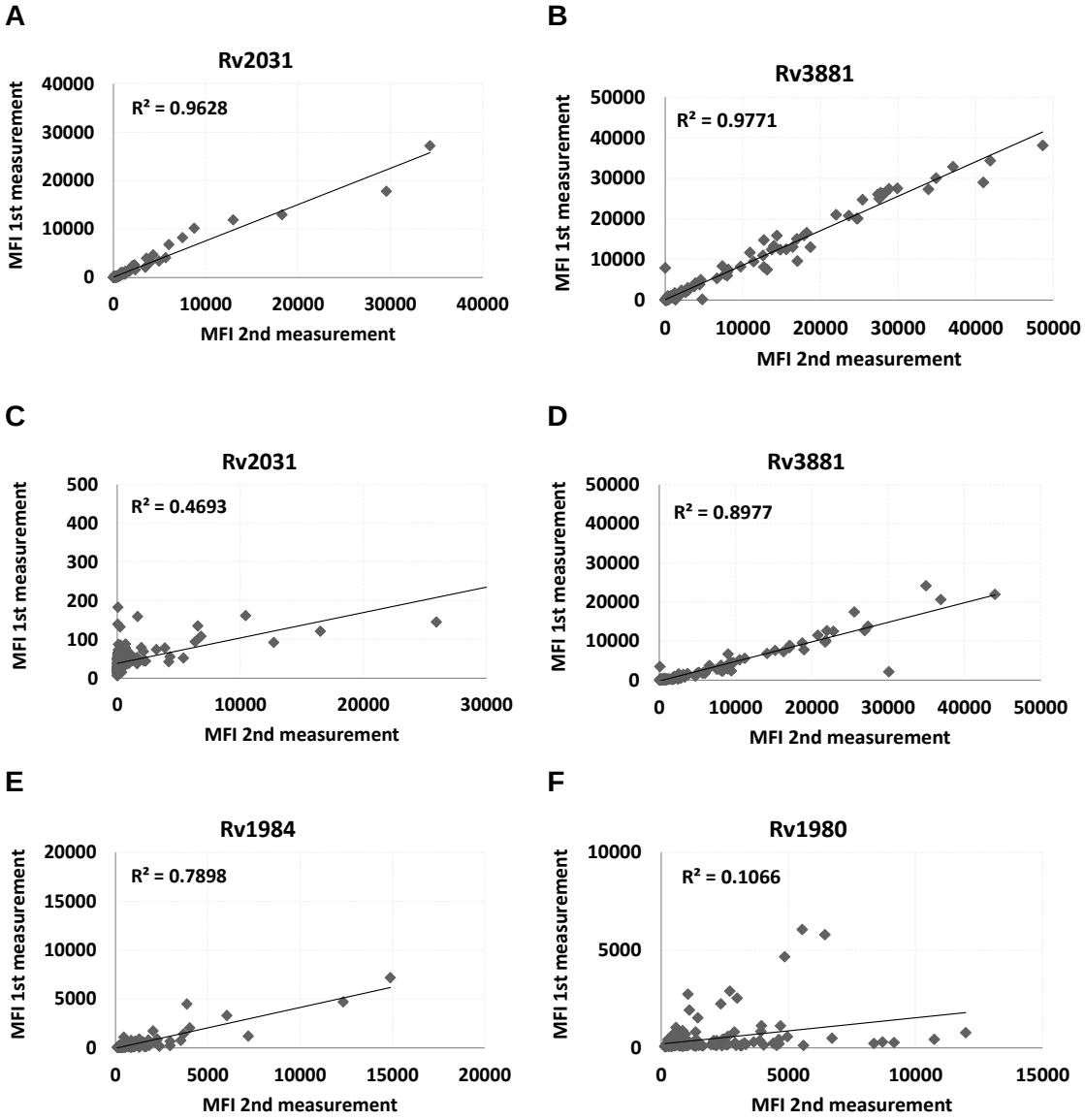


Figure 4-9: Correlation of MFI signal intensities of the same sample set with two independently coupled bead sets.

***A**, **B** antigens from supplier 1; **C**, **D** antigens from supplier 2; **E**, **F** antigens from supplier 3. Rv2031 (supplier 1, r=0.98), both Rv3881 (r=0.99 and 0.95, respectively) and Rv1984 (r=0.89) show good correlation coefficients. A poor correlation is seen for Rv2031 (supplier 2, r=0.69) and Rv1980 (r=0.33).*

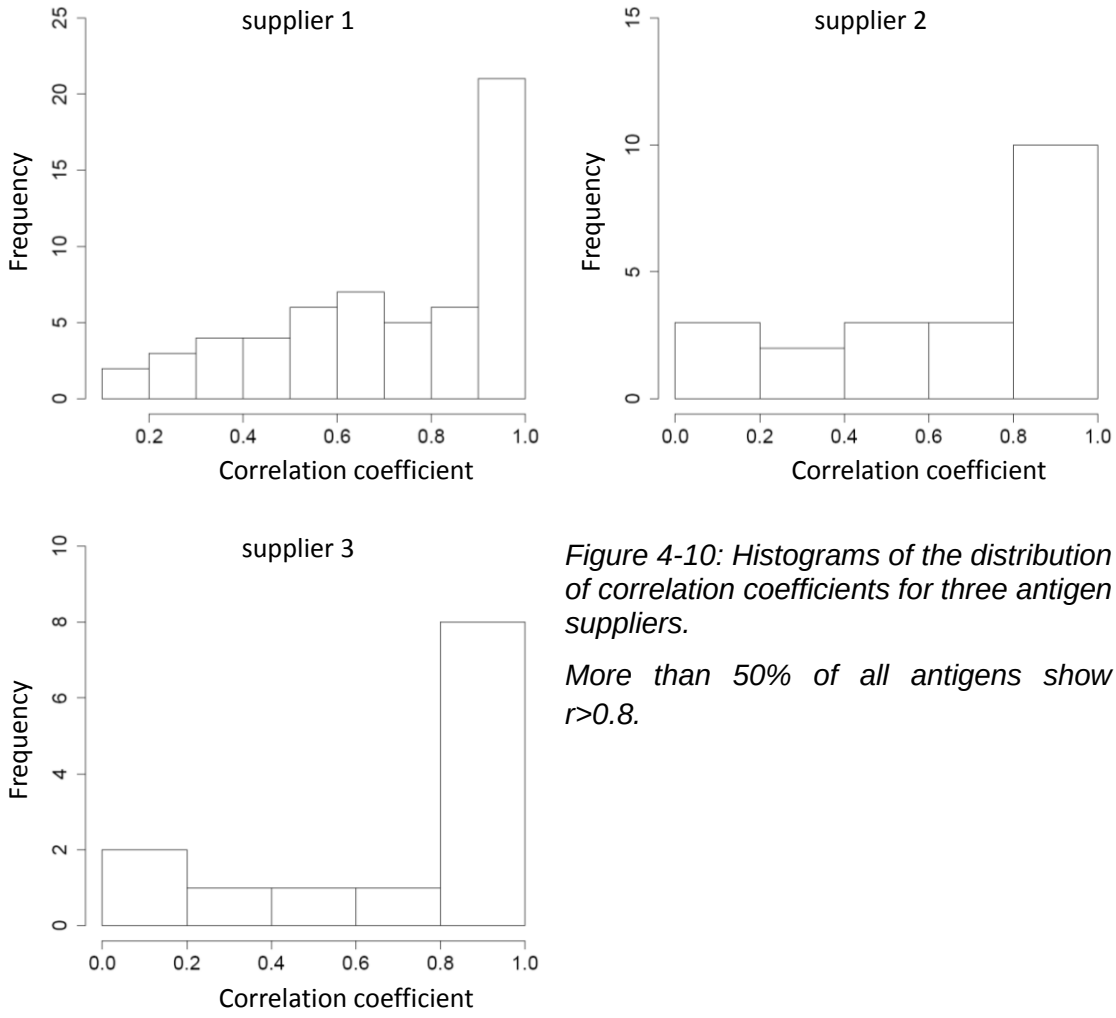


Figure 4-10: Histograms of the distribution of correlation coefficients for three antigen suppliers.

More than 50% of all antigens show r>0.8.

Comparison of antigens from different suppliers

Recombinant antigens were provided by different suppliers (see Table 4-6). The seroreactivity of 160 samples from active TB patients and 17 control sera to these recombinant antigens were analyzed.

Table 4-6: Overview of antigens provided by different suppliers

antigen	supplier
Rv0934	1, 2
Rv1860	1, 2, 4
Rv1886	1, 2
Rv1980	1 (2x), 2, 3
Rv2031	1, 2
Rv2873	1, 2
Rv2875	1, 2
Rv3804	1, 2

Rv3874	1, 2, 4
Rv3881	1, 2
Rv0577	3, 5

Correlation coefficients were determined for all antigens that were available from at least two different suppliers. The r values are summarized in Table 4-7 and vary from 0.22 for Rv1984 from suppliers 1 and 2 to 1.00 for Rv2031 from suppliers 1 and 2.

A poor linearity was seen when one of the recombinant protein showed no reactivity with almost all provided TB positive serum samples.

It was also possible to compare the performance of the antigen Rv0577 out of *M. tuberculosis* and *M. bovis* BCG strain from supplier 5 with the *M. tuberculosis* antigen of supplier 4. Good linearity was only observed when comparing the *M. tuberculosis* antigens from both suppliers (r = 0.92 vs. 0.47 and 0.41 respectively for the *M. tuberculosis* correlation with *M. bovis* BCG strain).

In total, the quality and functionality for certain antigens differ depending on the antigen itself and on the antigen supplier.

Table 4-7: Correlation coefficients for antigens from different suppliers measured with the same serum sample set.

antigen (suppliers)	**r**	**antigen (suppliers)**	**r**
Rv0934 (2 / 1)	0.99	**Rv1984 (1.1 / 1.2)**	0.70
Rv1860 (2 / 1)	0.96	**Rv2031 (2 / 1)**	1.00
Rv1860 (2 / 4)	0.97	**Rv2873 (2 / 1)**	0.94
Rv1860 (1 / 4)	0.97	**Rv2875 (2 / 1)**	0.87
Rv1886 (2 / 1)	0.36	**Rv3804 (2 / 1)**	0.77
Rv1980 (2 / 3)	0.46	**Rv3874 (2 / 1)**	0.99
Rv1980 (2 / 1)	0.87	**Rv3874 (4 / 2)**	0.98
Rv1980 (3 / 1)	0.29	**Rv3874 (4 / 1)**	0.97
Rv1984 (2 / 3)	0.90	**Rv3881 (2 / 1)**	0.94
Rv1984 (2 / 1.1)	0.22	**Rv0577 (5 *Mtb* / *M. bovis* BCG)**	0.47
Rv1984 (2 / 1.2)	0.29	**Rv0577 (3 *Mtb*/ 5 *Mtb*)**	0.92
Rv1984 (3 / 1.1)	0.23	**Rv0577 (3 / *M. bovis* BCG 5)**	0.41
Rv1984 (3 / 1.2)	0.30		

4.3 Generation of control samples

Quality control (QC) samples for serological assays are a prerequisite to verify the performance and to enable a comparison between different runs.

Negative control sample

A negative control sample is necessary to define a cut off value for the classification of positive or negative samples. Finding one single serum sample that is negative for all antigens is not trivial as there is always certain cross-reactivity with some antigens from other non-infectious mycobacteria species or vaccination.

For the definition of a negative control sample four serum samples of healthy lab members were tested. The sample with lowest signals over all antigens (donor no. 4) was chosen to serve as negative control in further sample screenings. However, this sample revealed high antibody reactivity towards three TB antigens (two of them were the same antigen from different suppliers). For the definition of a relative cut off, these values were replaced by true negative values observed in another donor.

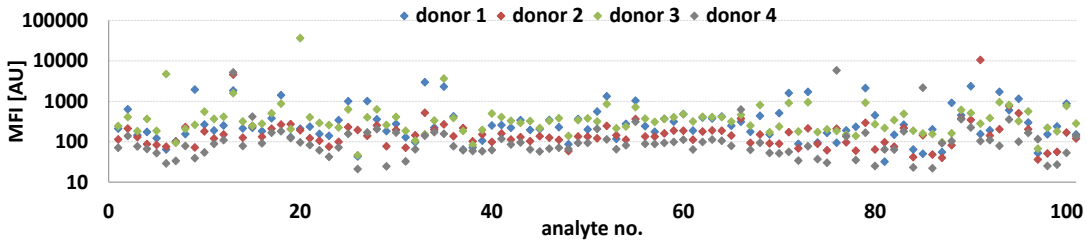


Figure 4-11: Distribution of MFI values for all antigens using four negative control sample donors.

Donor no. 4 (grey squares) was chosen to serve as negative control in further sample screenings because this sample revealed the lowest signals over all antigens.

Positive control sample

For a multiplexed serological assay, the availability of QC samples is not trivial, as a single sample does not contain antibodies against all antigens.

Thus, the idea was to generate an artificial positive control sample by pooling a certain number of positive serum samples, which differ in their serological activity but their combination should enable to cover all relevant antigens.

In order to reduce the amount of sample and working steps, operations research techniques can be applied to make an educated decision to identify the relevant samples whose combined reactivity could be used as a positive control. Based on already existing screening data on the multiplex assay response of a defined set of 160 serum samples, it was possible to calculate an optimal sample pool.
A mathematical model for predicting the outcome is the basis for any optimization task. In this case a model is required to predict which MFIs will be measured in a pool of samples for the multiplexed assay. The underlying hypothesis is that those values can be deduced from the MFIs measured in the single samples, which are considered to be pooled.

The purpose of the first experiment was to investigate the effect of sample pooling. Pools were created by subsequently adding randomly chosen samples in the scheme $S_1,\ S_1 + S_2,\ S_1 + S_2 + S_3$, ... up to a pool consisting of six samples.

The first result revealed that the effect of pooling samples is additive. The MFI signal of the pooled serum samples highly corresponds to the sum of MFIs of each single serum sample divided by the number of samples (see Table 4-8 for two examples).

$$MFI\ \ Pool\ of\ n\ samples\ \approx MFI(\frac{S_1 + S_2 + \ldots + S_n}{n})$$

Table 4-8: Calculation of the sample pool values depicted with two exemplarily chosen antigens.

sample no.	**Rv1980 [MFI]**	**Rv3804 [MFI]**
1	513	11224
2	945	21773
3	192	1530
4	125	542
5	17535	19492
6	611	25690
measured pool $(\mathbf{1+2})$	645	15347
calculated pool $(\frac{\mathbf{1+2}}{\mathbf{2}})$	729	16499
measured pool $(\mathbf{1+2+3})$	629	14421
calculated pool $(\frac{\mathbf{1+2+3}}{\mathbf{3}})$	550	11509
measured pool $(\mathbf{1+2+3+4})$	484	12339
calculated pool $(\frac{\mathbf{1+2+3+4}}{\mathbf{4}})$	444	8767
measured pool $\mathbf{1+2+3+4+5}$	4865	13880
calculated pool $(\frac{\mathbf{1+2+3+4+5}}{\mathbf{5}})$	3862	10912
measured pool $(\mathbf{1+2+3+4+5+6})$	3896	15404
calculated pool $(\frac{\mathbf{1+2+3+4+5+6+}}{\mathbf{6}})$	3320	13375

As shown in Figure 4-12 a correlation of R > 0.98 between the values of all 112 antigens predicted from single sample screenings (designated as calculated pool = $MFI\ \frac{S_1+S_2+\cdots+S_n}{n}$) to the MFIs measured in the pool (measured = $MFI\ \ Pool\ of\ n\ samples$) was observed. The slope of the linear regression is 1.03 for the least complex pool, and decreases to 0.57 for the pool of five samples.

In general, the signals generated by the respective pools were higher the more different samples were combined.

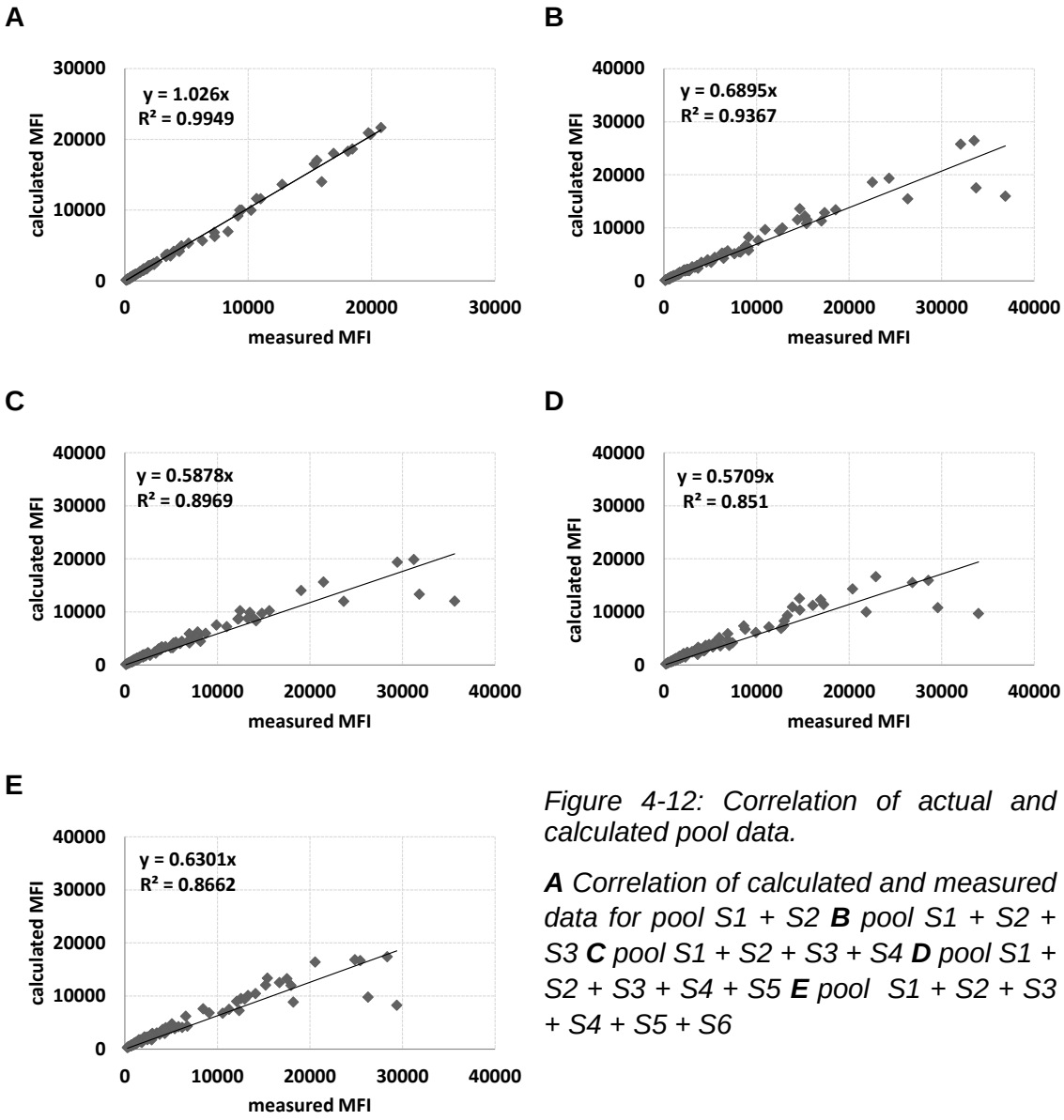


Figure 4-12: Correlation of actual and calculated pool data.

***A** Correlation of calculated and measured data for pool S1 + S2 **B** pool S1 + S2 + S3 **C** pool S1 + S2 + S3 + S4 **D** pool S1 + S2 + S3 + S4 + S5 **E** pool S1 + S2 + S3 + S4 + S5 + S6*

In a next experiment, the algorithm[3] described in the methods section was applied to find optimal pools for a panel of 112 antigens. Data for 170 previously screened samples were used as input. The allowed range for dilution of a single sample within the pool was 1:200 to 1:2000 and the final dilution of the pool was set to 1:200. In order to define whether a signal is positive or not, an arbitrary threshold was defined by multiplying the MFIs of the negative control by four separately for each antigen.

[3] The algorithm was created by Hannes Planatscher.

The output of the algorithm revealed ten pools consisting of up to ten different serum samples. It showed not to be possible to cover all antigens by a single sample or by pooling of two samples. It was predicted, that at least three serum samples have to be combined to show a combined reactivity to all 112 antigens.

In the next experiment the pool suggested by the algorithm (consisting of three samples) was tested.

The first result was, that the measured MFI values of the pool correlated with the calculated MFI values with a correlation coefficient of 0.96 (see Figure 4-13 A). The correlation of the measured MFI data between the pooled samples and the three single samples was comparatively low (Figure 4-13 B-D). This showed that no sample dominates the reactivity of the pool and that the signal pattern is indeed the result of the combined reactivity of all three samples.

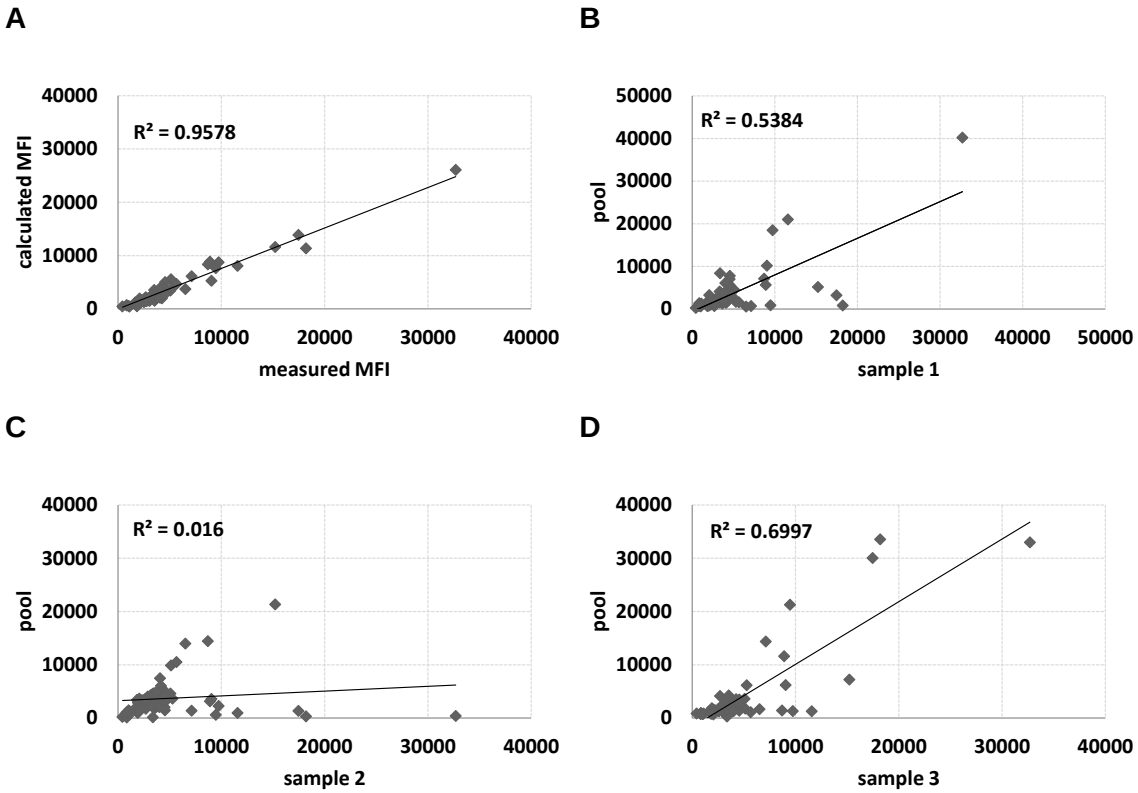


Figure 4-13: Correlation of the calculated sample pool MFI values and the measured pool MFI values.

*While in **A** the correlation of the measured MFI values of the sample pool with the calculated MFI values was quite high, the measured MFIs of the single samples did not correlate with the measured MFI values of the pool **B**-**D**. This showed that the unique coverage characteristic of the pool was due to the combination of the three samples.*

The second result was that the MFI values of the measured serum sample pool were above the artificially determined threshold of four times the negative control for 109 of the 112 antigens. For the three antigens that did not fulfill this condition, the negative control must be redefined.

For one antigen (Rv2462 from supplier 3), the signals of the three single positive samples and the resulting sample pool (517 MFI) were comparably low and could not exceed the artificially defined threshold of four times the negative control which had a comparable signal (~300 MFI) with that of other antigens. Another antigen (Rv3362 from supplier 1) revealed a high signal of more than 2500 MFI with the negative control. This antigen was also available from another batch which revealed a much lower signal (175 MFI) in the negative control. The values for the three measured serum samples were in a comparable range for both antigens and thus the value of the antigen with the lower negative control signal can be used as negative control value.

For the remaining third antigen (Rv3762 from supplier 1), the value of the negative control (870 MFI) could be replaced by a lower value from another negative control sample e.g. that of donor no. 1 (~300 MFI, see section negative control sample).

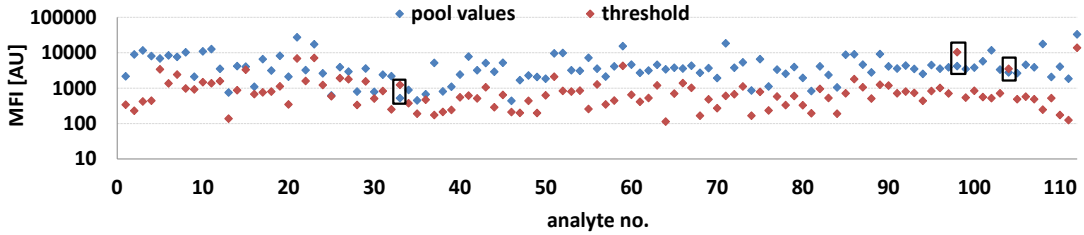


Figure 4-14: Test of serum reactivity of the sample pool.

For definition of a positive signal an arbitrary cut off was set at four times the MFI value of the negative control for each antigen. All MFI values of the pooled serum samples should exceed this defined threshold. Three antigens were not covered by the serum sample pool and are highlighted with a black box. In order to gain a positive signal on these three antigens the negative control must be redefined.

4.3.1 Freeze and thaw stability of the positive control pool

Freeze and thaw cycles (f/t) of the serum pool were performed to test the stability of the huIgGs in the serum samples. The freeze cycles were performed by placing the sample in liquid nitrogen. Samples were thawed at room temperature.

The MFI values generated with the pool of the first freeze and thaw cycle was set to 100% and all further f/t cycle values were calculated relatively to that value. The mean values for the remaining signals for the sample pool on all antigens were 93% for 2 f/t, 93% for 3 f/t, 91% for 4 f/t and 96% for 5 f/t. The antibodies in the sample pool showed > 80% stability within the five freeze and thaw cycles on at least 94% of the antigens (see Table 4-9). Thus, the antibodies in the sample pool seem to be stable over at least five freeze/thaw cycles.

Table 4-9: Mean values and range of antibody stability on all antigens.

no. of freeze/thaw cycle	**mean (range) [%]**	**% of antigens >80%**
2	93 (61 - 110)	98
3	93 (63 - 114)	94
4	91 (65 - 107)	96
5	96 (71 - 112)	98

4.4 Antibody profiles in different tuberculosis infection states

Miniaturized and multiplexed assays are an optimal solution for applications in which several parameters of a single sample have to be analyzed in parallel. Microtiter plate-based ELISA assays are robust and easy to use, but can only measure a single antigen. Multiplex immunoassays are excellent tools to perform serological assays.

A total of 644 non-TB, 592 active TB (343 S+C+ and 249 S-C+) and 85 latently infected TB patient serum samples were screened for their serological response of huIgGs to 112 recombinant tuberculosis antigens (some antigens appear twice or three times because they were delivered from different suppliers). For each sample a single measurement in one dilution was performed.

4.4.1 Determination of assay cut off values and calculation of diagnostic sensitivity and specificity

The goal of diagnostic testing is to accurately classify serum samples as positive or negative for a given infectious agent. Unfortunately, there is an overlapping continuum of assay responses from negative to positive. It is possible to measure how well an assay classifies serum samples as positive or negative in a given population by calculating the diagnostic sensitivity and diagnostic specificity of the assay. Sensitivity is the probability that an assay correctly identifies positive patients, and specificity is the probability that an assay correctly identifies negative patients. Sensitivity and specificity of an assay vary according to the assay cut off. Changing the assay cut off to increase sensitivity results in a similar decrease in specificity and also the other way round.

There are a variety of ways to determine the cut off. In this case, for each antigen an artificial score was chosen. The measured MFI value of the negative control was multiplied by three (weak positive), five (medium positive) and ten (strong positive). All samples below three times the MFI value of the negative control were considered as negative and hence all samples above three times the negative control were considered as positive with different grades.

The sensitivity of all antigens ranges between 5% and 82% and the specificity between 28% and 97%. The data are summarized in
Table 9-2 in the supplementary material. The cut off was set at three times the MFI value of the negative control.

All 112 proteins were recognized by serum from at least one patient. Most reactive proteins were predominantly recognized by active TB samples (see Figure 4-15). The percentage of reactive samples in active and non-TB patients for each antigen is graphically depicted in the supplementary material in Figure 9-2.

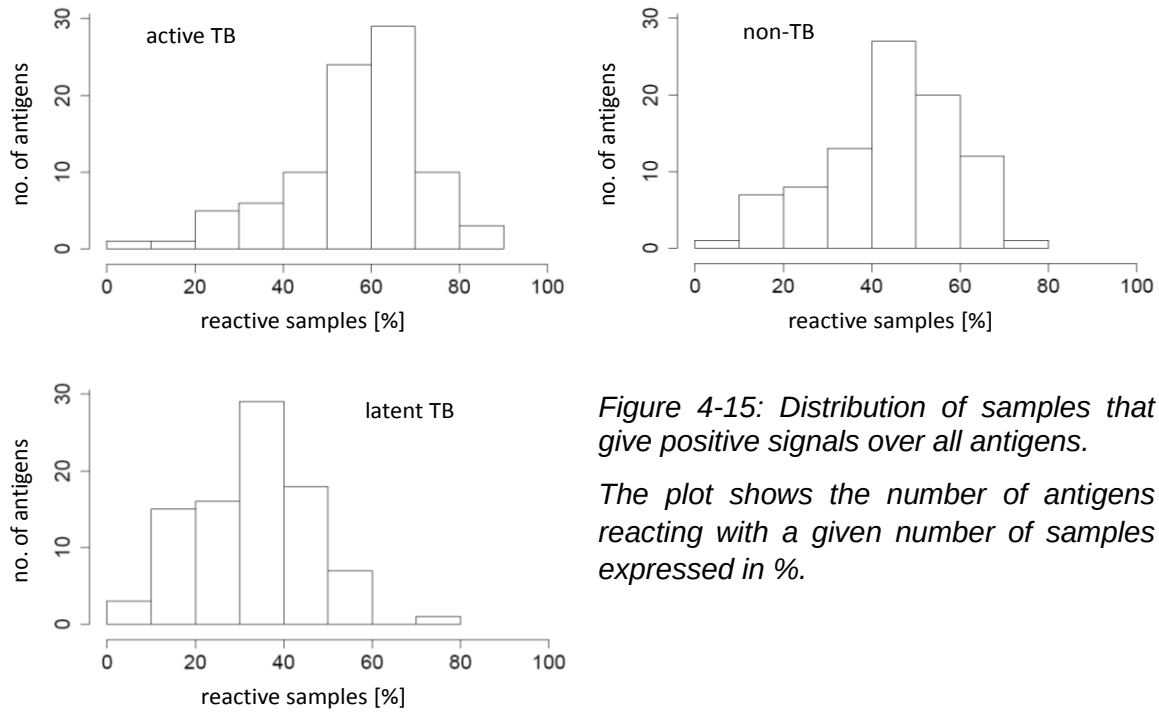


Figure 4-15: Distribution of samples that give positive signals over all antigens.

The plot shows the number of antigens reacting with a given number of samples expressed in %.

4.4.2 Profile of human serum IgG response

Box diagrams were chosen to show the distribution of measured data. Ten TB antigens were plotted to display the distribution levels of reactive antibodies in active and non-TB patients.

All 1321 measured serum samples were reactive towards at least several TB antigens. The reactivity of antibodies in non-TB samples was significantly lower compared with active TB patients. No antigen was found where the antibody reactivity was higher in non-TB controls compared with active TB patients. In

general, non-TB patients showed more outliers (data values outside the 95th percentile) compared with active TB patients.

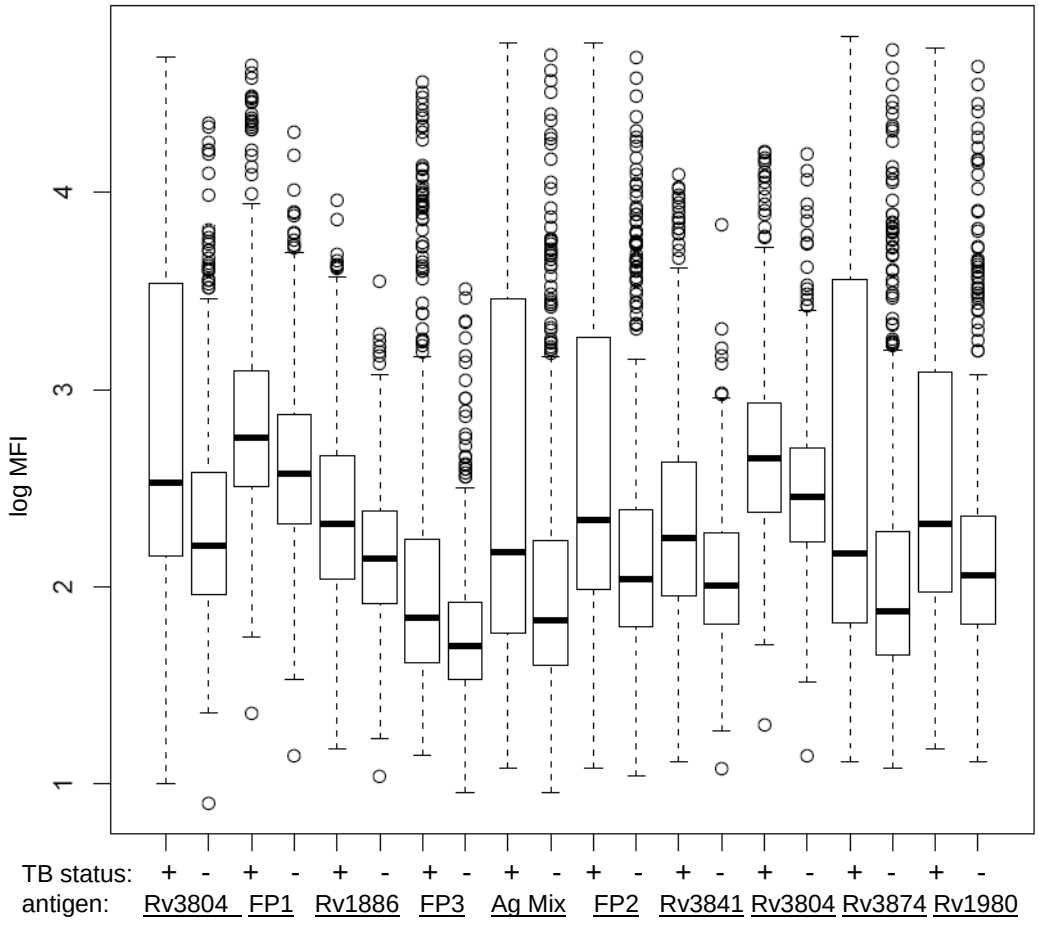


Figure 4-16: Distribution of reactive antibodies in the serum of active TB and non-TB groups.

Boxplots were chosen to graphically depict the distribution of antibody reactivity towards TB antigens. Statistical analysis revealed 41 antigens that can significantly discriminate between active- and non-TB patients. Exemplarily, the top ten antigens (see Table 4-10) with the best p-values are shown.

Although the boxplots graphically depict the key values and the distribution of the data, the deduction of the significance of differences between groups remains complicated. The Shapiro-test was used to test the frequency distribution of each group of patients. Subsequently, a Mann-Whitney U-test (non-parametric test) was performed to analyze the data and to enable a discrimination of the TB disease status.

The test results are summarized in Table 4-10 (active TB vs. non-TB) and Table 4-11 (latent TB vs. non-TB). In this study, the p-value of 0.005 was set as a predetermined cut off point. However, the empty bead alone revealed a p-value of 3.4E-07 in the case of active/non-TB discrimination and 6.9E-05 in the case

of latent/non-TB discrimination. This was set as a further cut off and only antigens with a p-value below this threshold were assigned.

When the test value was below the pre-determined threshold, known as the significance level, the examined antigen or parameter was regarded as significant. The median difference of active or latent and non-TB revealed the relative amount between the groups. Antibodies in active TB serum samples revealed a statistically significant higher reactivity towards all 41 antigens, shown in Table 4-10, than in non-TB samples. In contrast, antibodies in latent TB patients revealed a statistically significant lower reactivity towards all 18 antigens (listed in Table 4-11) compared with non-TB controls.

Table 4-10: Median MFI values (minimum – maximum) of 41 antigens in active TB and non-TB serum samples.

The Mann-Whitney U-test was performed to evaluate the significance of antibody reactivity between active and non-TB groups. The p-values were reported in the analysis. A cut off was set at the empty bead and all p-values below are displayed.

	median (min - max) [MFI]		
antigen_supplier	**active TB**	**non-TB**	**p-value**
Rv3804_1	337 (10 - 48123)	161 (8 - 22541)	4.9253E-26
FP1_2	150 (12 - 57048)	68 (9 - 49927)	2.427E-23
Rv1886_2	176 (13 - 12272)	102 (12 - 6850)	5.6154E-22
FP3_2	148 (13 - 61609)	75 (12 - 52624)	5.7385E-21
Antigen Mix_4	208 (15 - 53862)	114 (13 - 43731)	1.0953E-20
FP2_2	217 (12 - 57025)	109 (11 - 48344)	2.7892E-19
Rv3841_2	448 (20 - 16081)	283 (14 - 15651)	1.0936E-16
Rv3804_2	208 (15 - 9177)	138 (11 - 3555)	5.696E-16
Rv3874_2	69 (14 - 36362)	49 (9 - 3248)	8.0683E-16
Rv1980_2	572 (23 - 44098)	376 (14 - 20184)	1.2084E-15
Rv1860_2	115 (13 - 17545)	81 (10 - 11774)	1.2963E-15
Rv2031-Rv2873_3	182 (22 - 40066)	109 (20 - 39887)	4.8949E-15
Rv2875_2	424 (20 - 38712)	263 (20 - 22679)	9.2162E-14
Rv3874-Rv3875_2	82 (12 - 51179)	60 (10 - 6985)	1.2612E-13
Rv3881_2	107 (9 - 31480)	66 (8 - 37658)	4.2644E-13
Rv3874_4	101 (14 - 33899)	74 (11 - 3349)	6.0636E-13
Rv1860_4	359 (27 - 40821)	229 (19 - 36307)	7.5935E-13
Rv3874_1	95 (8 - 51604)	64 (6 - 6522)	1.1283E-12
Rv3875_4	133 (15 - 51729)	95 (12 - 16086)	1.2407E-12
Rv3881_1	167 (11 - 50126)	102 (8 - 24043)	2.5943E-12
Rv0934_2	149 (12 - 57426)	98 (10 - 48861)	4.5257E-12
Rv1196_1	150 (8 - 9301)	111 (5 - 6685)	1.2093E-11
Rv2873_2	187 (11 - 24825)	118 (15 - 9708)	1.248E-11
Rv2031_1	98 (7 - 45026)	72 (4 - 44867)	1.6792E-11
Rv0798_1	409 (21 - 12098)	289 (13 - 15095)	9.9923E-11
Rv1980_1	547 (32 - 61013)	324 (14 - 45005)	1.0583E-10
Rv3628_1	214 (16 - 2294)	153 (15 - 2057)	1.2204E-10
Rv0934_1	203 (13 - 53679)	139 (12 - 41746)	1.4964E-10
Rv3763_1	276 (14 - 3093)	203 (10 - 3169)	2.5385E-10
Rv0302_1	195 (11 - 48395)	136 (4 - 54097)	1.3545E-09
Rv0831_2	478 (52 - 8978)	335 (21 - 6956)	1.102E-08

Rv1980_3	256 (81 - 17832)	208 (81 - 24143)	1.2017E-08
Rv1411_1	49 (6 - 54870)	40 (4 - 8671)	1.2128E-08
Rv1860_1	229 (8 - 52831)	156 (5 - 48149)	3.0557E-08
Rv3326_1	281 (15 - 4021)	208 (11 - 2392)	7.9153E-08
Rv1284_1	498 (64 - 33645)	374 (17 - 12288)	1.0491E-07
Rv2927_1	486 (8 - 32226)	328 (7 - 30736)	1.1238E-07
Rv3248_1	244 (12 - 13180)	182 (9 - 5847)	2.1403E-07
Rv1926_1	179 (14 - 49301)	141 (12 - 19138)	2.4684E-07
Rv1837_1	294 (13 - 18038)	226 (10 - 10580)	2.7221E-07
Rv0583_1	1121 (10 - 55858)	838 (5 - 49893)	3.2573E-07
Empty bead	73 (19 - 1843)	59 (15 - 1181)	3.3908E-07
E. coli bead	148 (15 - 2272)	120 (14 - 2031)	1.0648E-06

Table 4-11: Median MFI values (minimum – maximum) of 18 antigens in latent TB and non-TB serum samples.

The Mann-Whitney U-test was performed to evaluate the significance of antibody reactivity between latent and non-TB groups. The p-values were reported in the analysis. A cut off was set at the empty bead and all p-values below are displayed.

	median (min - max) [MFI]		
antigen_supplier	**latent TB**	**non-TB**	**p-value**
Rv0272_1	85 (21 - 452)	146 (9 - 9883)	3.5938E-13
Rv0632_1	98 (16 - 603)	128 (6 - 13626)	2.2338E-08
Rv3878_1	52 (12 - 227)	60 (10 - 21515)	3.2719E-08
Rv3775_1	102 (22 - 473)	148 (10 - 3004)	1.1843E-07
Rv0583_1	439 (28 - 30644)	838 (5 - 49893)	1.2411E-07
Rv0984_1	141 (21 - 1000)	197 (8 - 14267)	1.339E-06
Rv2544_1	136 (16 - 715)	168 (8 - 7673)	1.4965E-06
Rv3319_1	179 (24 - 746)	239 (12 - 4838)	2.3964E-06
Rv1636_3	59 (17 - 236)	76 (5 - 1977)	2.489E-06
Rv2870_1	143 (20 - 15291)	239 (8 - 6764)	7.4497E-06
Rv2282_1	94 (16 - 433)	131 (7 - 4848)	9.9393E-06
Rv3762_1	203 (22 - 1617)	319 (13 - 11587)	1.5838E-05
Rv2185_3	164 (12 - 3040)	232 (6 - 31641)	2.0339E-05
Rv2984_1	285 (28 - 2753)	408 (11 - 18048)	3.972E-05
Rv0801_1	270 (27 - 1404)	382 (10 - 13248)	4.1863E-05
Rv2873_2	87 (14 - 5085)	118 (15 - 9708)	4.4801E-05
Rv1566_1	184 (15 - 37845)	289 (7 - 35010)	6.1011E-05
Rv2151_1	89 (13 - 427)	124 (8 - 17305)	6.1679E-05
Empty bead	44 (21 - 374)	59 (15 - 1181)	6.9344E-05
E. coli bead	88 (20 - 544)	120 (14 - 2031)	0.0005

Receiver operating characteristics (ROC) curves were chosen to graphically depict the sensitivity (true positive rate) and the specificity (true negative rate) of single antigens. The highest ranked TB antigens for diagnosis of active TB had an area under the curve (AUC) of not more than 0.68 which is not sufficient for a diagnostic assay (see Figure 4-17 for the highest ranked antigens of active TB).

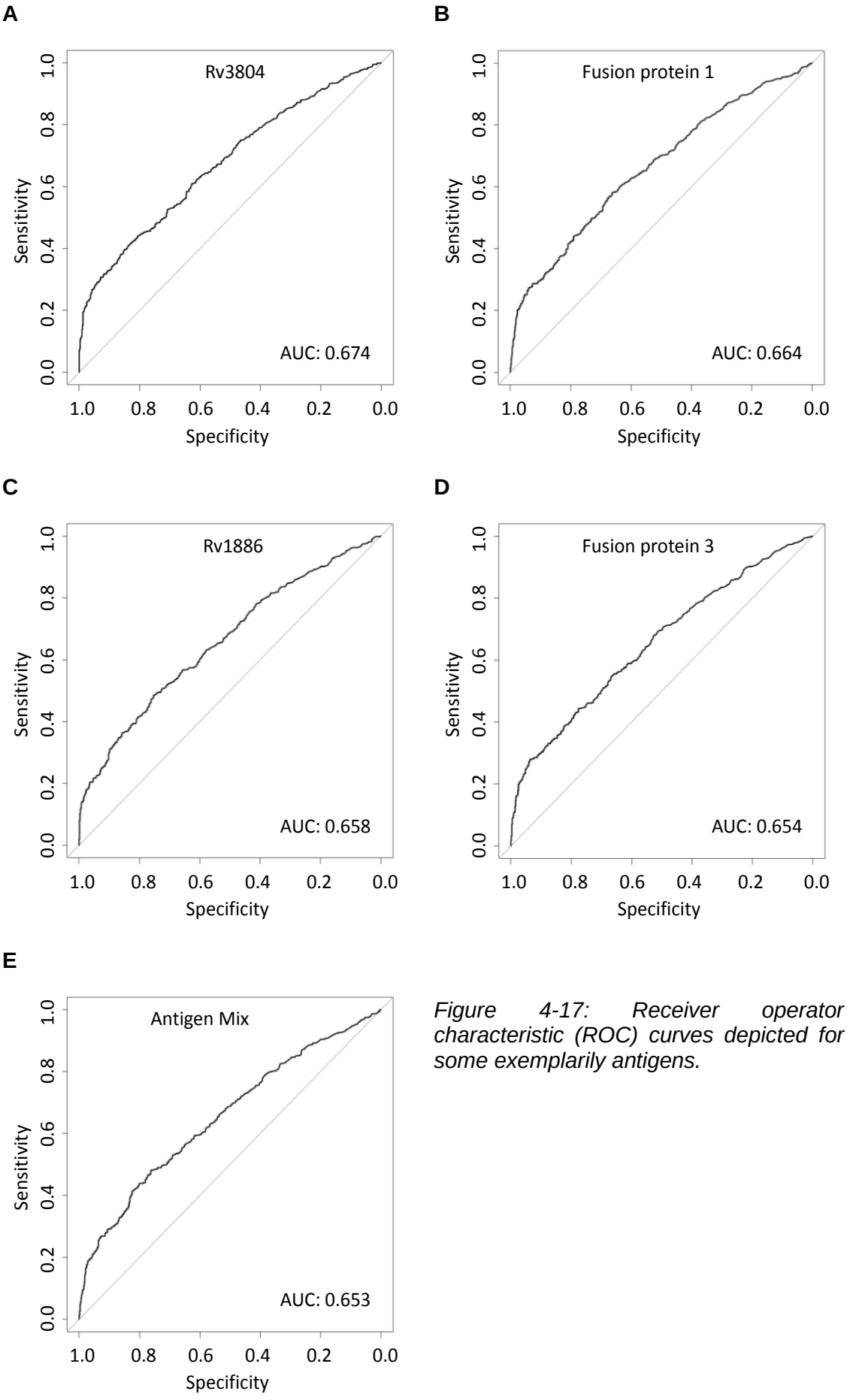


Figure 4-17: Receiver operator characteristic (ROC) curves depicted for some exemplarily antigens.

4.4.3 Characteristics of immunoreactive proteins

The subcellular localization of the proteins in the immunoproteome was investigated. This is a key property of bacterial antigens because it affects accessibility by the immune response as well as antibody function.

The more frequently reacting proteins that were associated with active TB (see Table 4-10) were found predominantly in the extracellular fraction of *Mycobacterium tuberculosis* and are thus secreted proteins (Table 4-12). Reactivity to membrane-associated TB proteins was mostly found in latent TB samples (Table 4-13).

Table 4-12: Proteins associated with active TB.

antigen	annotation
Rv3804	Secreted antigen 85A / mycolyltransferase
Rv1886	Secreted antigen 85B / mycolyltransferase
Rv3841	Probable bacterioferritin BfrB
Rv3874	Secreted antigen CFP10
Rv1980	Secreted antigen MPT64
Rv1860	Secreted glycoprotein MPT32
Rv2875	Secreted antigen MPT70
Rv2873	Surface lipoprotein antigen MPT83
Rv0934	Glycolipoprotein 38 kDa antigen / PstS1
Rv2031	16 kDa antigen / alpha-crystallin
Rv1411	Lipoprotein LprG

Table 4-13: Proteins associated with latent TB.

antigen	annotation
Rv0272	Membrane bound / hypothetical protein
Rv0632	Membrane bound / probable Enoyl-CoA hydratase
Rv3878	Membrane-bound / ESX-1 secretion-associated protein
Rv3775	Membrane-bound / probable lipase
Rv0583	Membrane-bound / probable conserved lipoprotein
Rv0984	Membrane-bound / possible moaB2, pterin-4-alpha-carbinolamine dehydratase
Rv2544	Secreted antigen / probable conserved lipoprotein
Rv3319	Membrane-bound / Probable succinate dehydrogenase
Rv1636	Secreted antigen / TB15.3

4.4.4 Antibody response and characteristics of non-TB patients

The serological response of non-TB patients to the 112 different TB antigens was investigated. One possibility was, that non-TB seroreactivity was caused by exposure to non-tuberculous mycobacteria which produce proteins cross-reactive with *M. tuberculosis*. It was found that 43 of the 644 (7%) non-TB patients had past TB history (see Table 3-4).

A Wilcox-test revealed five antigens with significant differences in serological response between patients with and without past TB history in non-TB samples. Antigens Rv3804, Rv1860 (suppliers 1 and 2) and Rv3881 showed significant increased antibody reactivity in patients with past TB history. In contrast, antigen Rv1926 revealed a significantly lower reactivity (see Table 4-14 and Figure 4-18).

Table 4-14: Past TB history of non-TB patients.

	median (min - max) [MFI]		
antigen	**yes**	**no**	**p-value**
Rv3804	265 (44 - 16504)	125 (28 - 9727)	7.87115E-05
Rv1860_1	151 (24 - 12741)	103 (22 - 8284)	0.03849
Rv1860_2	85 (21 - 2613)	61 (19 - 438)	0.00642
Rv3881	81 (19 - 8810)	62 (19 - 5254)	0.03996
Rv1926	91 (30 - 2055)	145 (28 - 4331)	0.00116

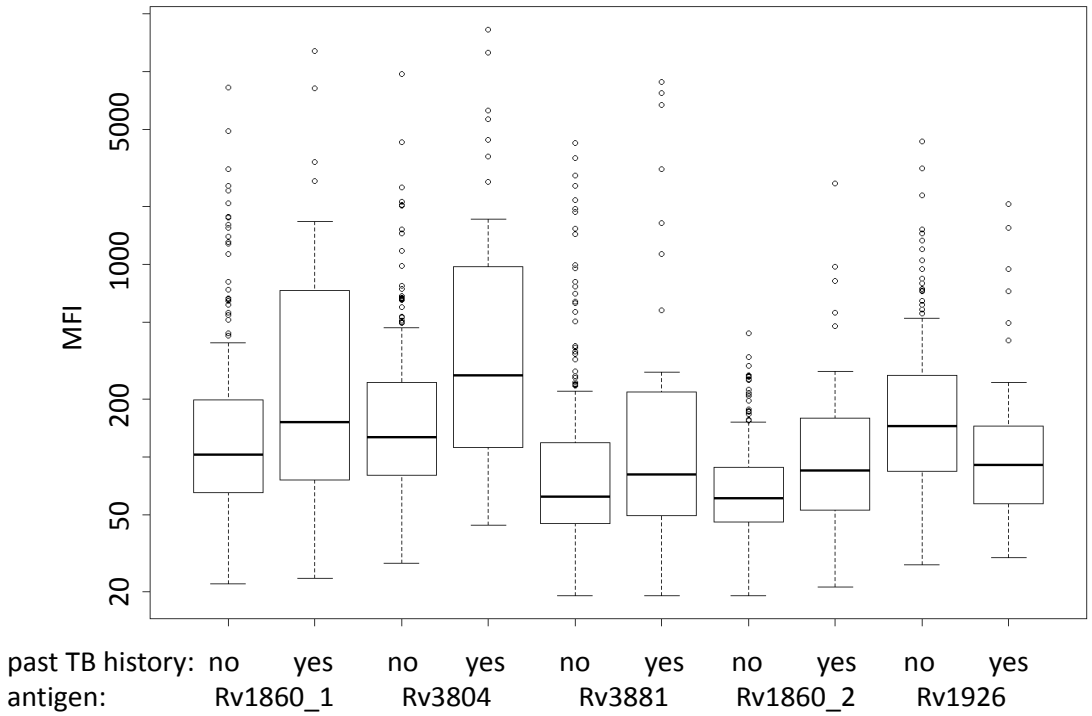


Figure 4-18: Influence of past TB history on the serological response towards different antigens.

Increased antibody reactivity towards three antigens (Rv1860, Rv3804 and Rv3881) in patients with past TB history was identified. Decreased antibody reactivity against Rv1926 in patients with past TB history was observed.

4.4.5 Clinical and demographic characteristics of active TB patients

The relationship between antibody response and clinical and demographic characteristics of active TB patients was assessed.

Smear grade results were only moderately associated with the antibody response. Only three antigens were identified, to which the sera showed a significantly increased reactivity derived from smear positive samples from active TB patients compared with smear negative samples (Table 4-15).

Table 4-15: Influence of smear grade of active TB patients on antibody reactivity.

	median (min - max) [MFI]		
antigen	**negative**	**positive**	**p-value**
Rv0577_3	184 (20 - 2227)	272 (39 - 9170)	0.04343
Rv1860_1	135 (27 - 15416)	159 (27 - 40090)	0.04129
Rv3874_1	53 (13 - 1100)	74 (21 - 21197)	0.00364

Next, the relative preference for each of the 41 identified antigens for individual patients' sera were assessed. Log MFI values of all active TB sera were shown in a heat map (Figure 4-19).

Almost all antigens seem to be reactive with a plethora of serum samples from active TB patients. Very diverse antibody reactivity (serological response) was observed. Thus, within the pool of TB-associated proteins, target preference varied among patients.

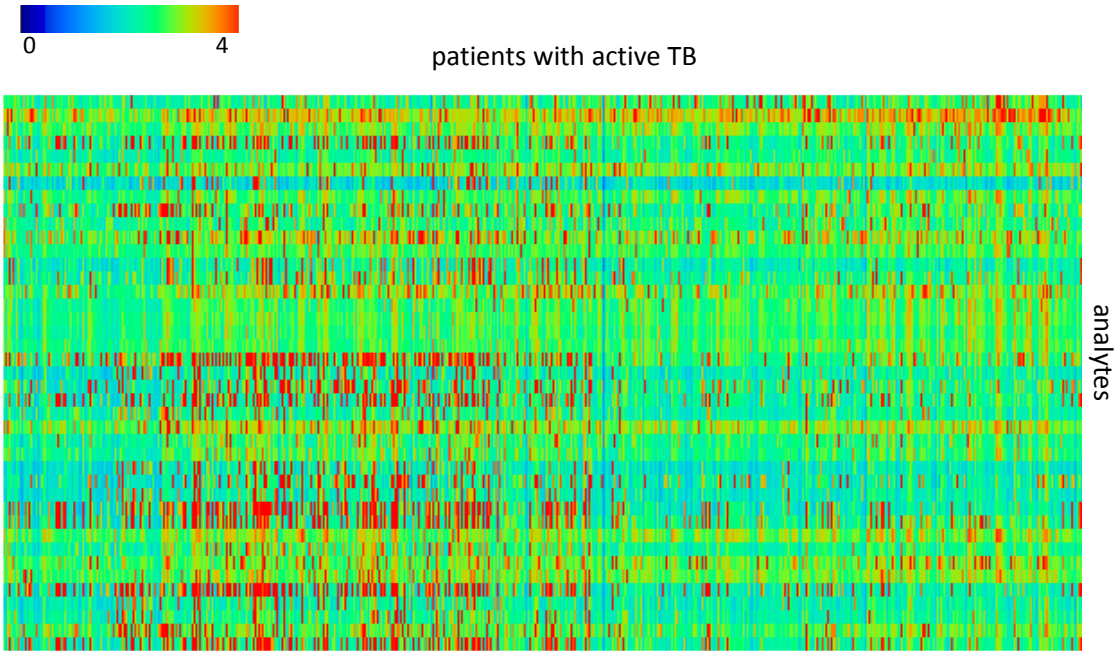


Figure 4-19: Heterogeneous antibody response pattern of active TB patients.

The heat map shows the reactivity of sera from active TB patients to 41 TB-associated proteins. Each column represents one serum and each row represents one TB-associated protein. Normalized signal intensities are visualized as a color spectrum as indicated.

4.4.6 Antibody response pattern generation

Since tuberculosis is a highly heterogeneous disease, it is generally recognized that information about the occurrence of a single serological marker does not reflect the status of the disease. A panel of markers could have the potential to be more appropriate in clinical practice. However, it is difficult to select and combine the potential candidates from the complex data sets generated in screening experiments (this is not a general issue but this is true in this case).

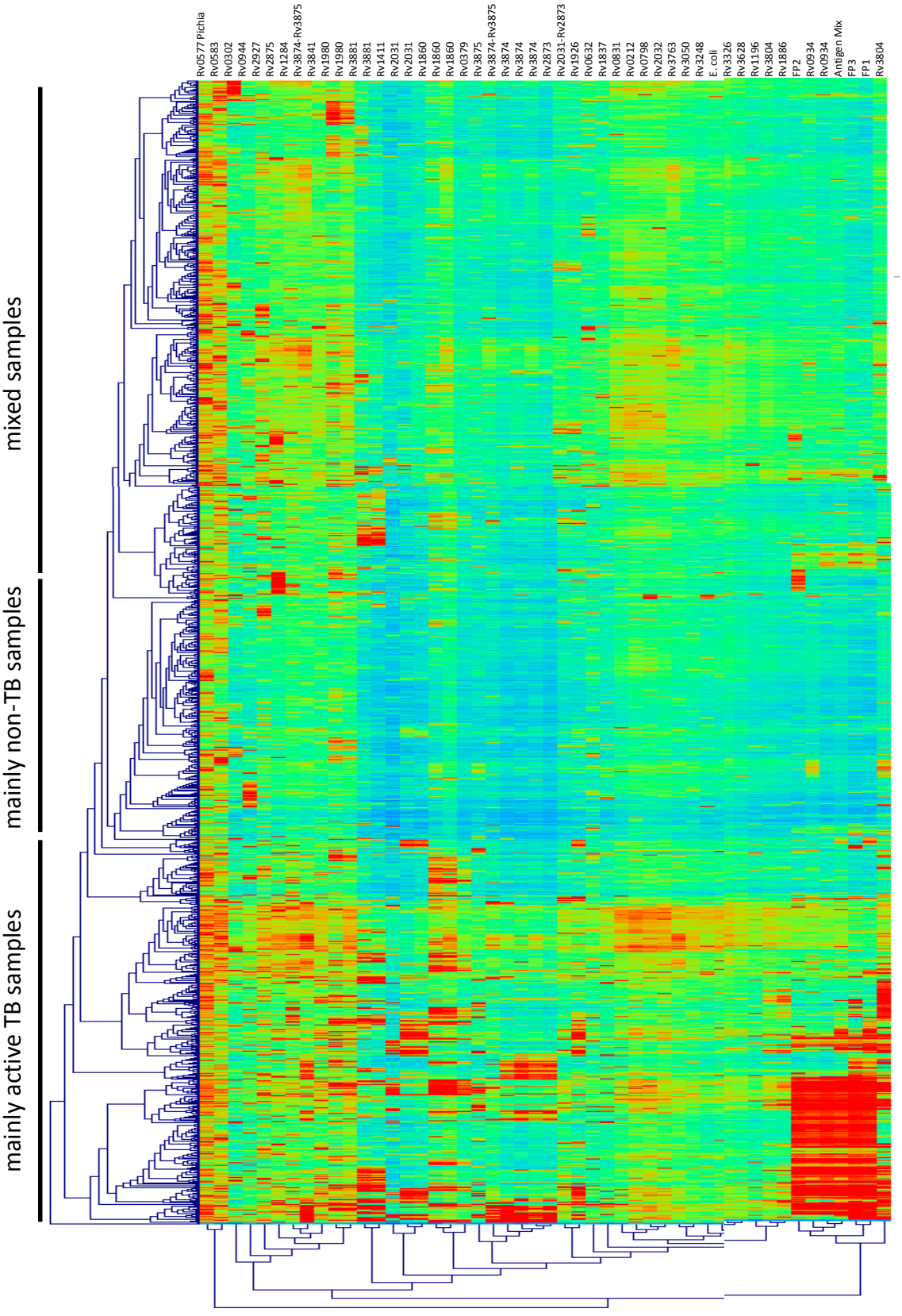
Screening of 644 non-TB, 592 active TB and 85 latent TB serum samples revealed a high number of significant antigens when a p-value of 0.005 was defined in the Mann-Whitney U-test (see Table 4-10 and Table 4-11).

Afterwards, it was necessary to assess the quality of these antigens for their suitability to differentiate between different tuberculosis states. The aim of the following data analysis was to identify patterns which allow the discrimination of i) non-TB patients from active TB patients, ii) non-TB patients from latent TB patients. Hierarchical clustering was applied to the generated datasets, which allowed the open-ended exploration and a simplified visualization of relevant data sets of a large number of samples.

Cluster analysis was performed using the set of antigens identified in the last section (4.4.2) in order to discover response patterns of diagnostic relevance. In this analysis, a two-class unpaired SAM (significance analysis of microarray) analysis was performed to identify differences in the levels of antibody response.

For the discrimination of active and non-TB disease state a panel of 50 antigens was identified by the significance analysis. In the left part of the heat map a clustering of mainly active TB samples was observed, whereas in a small part in the middle more non-TB samples are clustered. More than half of the samples are mixed in the left part of the heat map (Figure 4-20 A). A panel of 35 antigens was identified that should correctly cluster latent and non-TB samples (Figure 4-20 B). However, the small number (n=85) of samples derived from latently infected TB patients were distributed randomly between the non- TB samples (n=644, latent TB samples are marked with a black line at the right side of the heat map). In both cases a clear distinction between the respective patient groups was not possible with the identified combinations of antigens.

A


Rv0577 Pichia
Rv0583
Rv0302
Rv0944
Rv2927
Rv2875
Rv1284
Rv3874-Rv3875
Rv3841
Rv1980
Rv1980
Rv3881
Rv3881
Rv1411
Rv2031
Rv2031
Rv1860
Rv1860
Rv1860
Rv0379
Rv3875
Rv3874-Rv3875
Rv3874
Rv3874
Rv3874
Rv2873
Rv2031-Rv2873
Rv1926
Rv0632
Rv1837
Rv0831
Rv0212
Rv0798
Rv2032
Rv3763
Rv3050
Rv3248
E. coli
Rv3326
Rv3628
Rv1196
Rv3804
Rv1886
FP2
Rv0934
Rv0934
Antigen Mix
FP3
FP1
Rv3804
mixed samples
mainly non-TB samples
mainly active TB samples

B

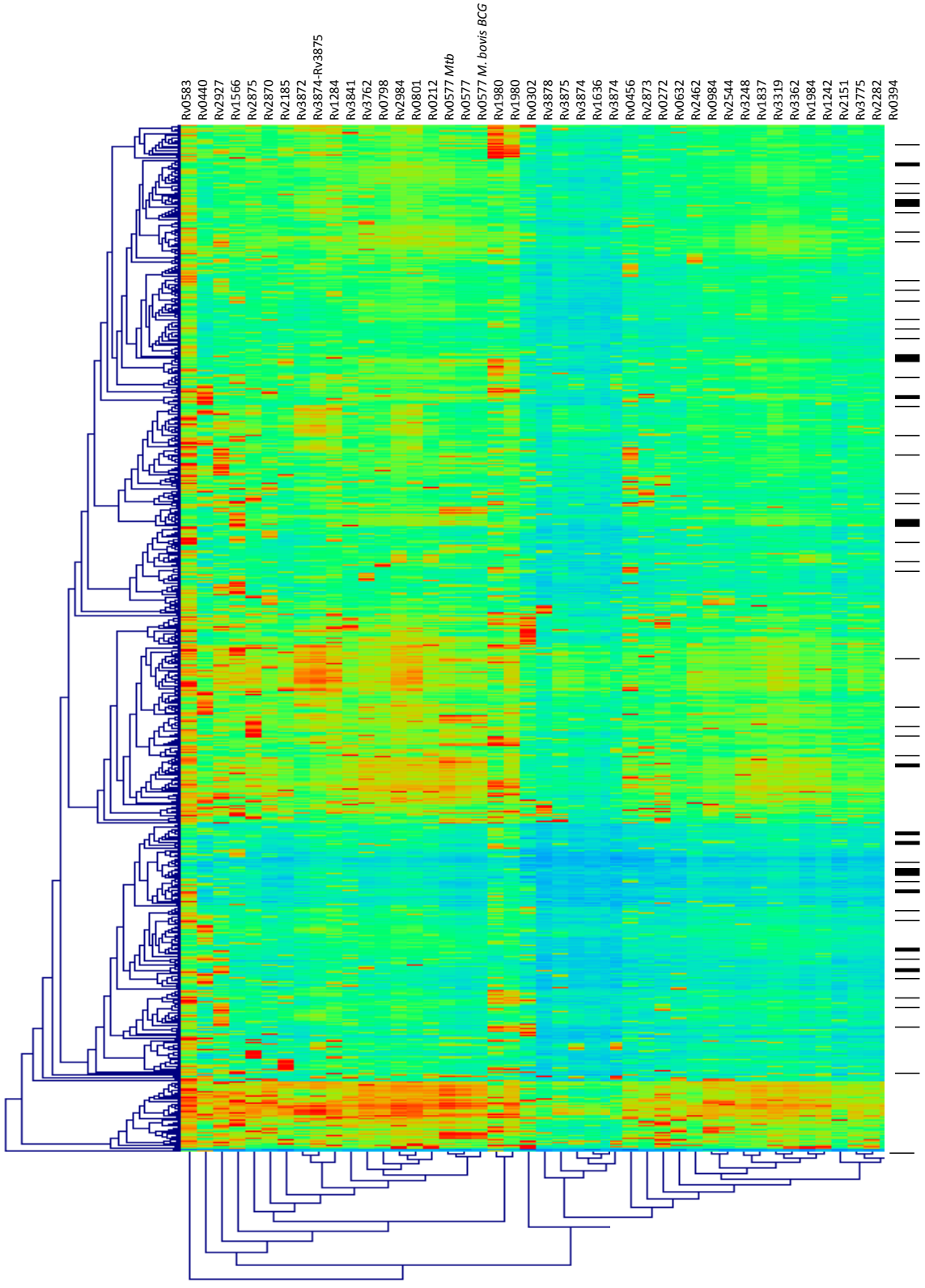


*Figure 4-20: Heat map of the candidate antigens that were suggested to distinguish active from non-TB (**A**) and latent from non-TB (**B**).*

In this analysis, samples were clinically assigned into the respective two groups A two-class unpaired SAM (significance analysis of microarray) analysis was carried out to identify differentially reactive antibodies against the antigens. A seroresponse to 50 distinct antigens was identified in the case of active/non-TB and 35 antigens in the

case of latent/non-TB. However, the discrimination between the respective groups is poor.

The results obtained with the significance analysis were analyzed with another bioinformatics approach. The machine learning program WEKA was used to identify a panel of antigens that classifies the serum samples according to their actual clinical status. The ability of these antigens for a correct classification was evaluated. The attribute selection mode "weka.attributeSelection.InfoGainAttributeEval" was used and a 10-fold cross-validation with all antigens was performed. Cross-validation is a technique for assessing how the results of a statistical analysis will generalize to an independent data set. In this case, the data set was divided into k=10 sub sets, and the holdout method was repeated 10 times. Each time, one of the 10 sub sets was used as the test set and the other k-1 (9) subsets were put together to form a training set. The average error across all 10 trials was computed [98].

Table 9-3 in the supplementary material shows a panel of 50 antigens that are ranked above the empty bead and that are differentially expressed in active and non-TB states. For the discrimination of latent and non-TB a panel of 35 antigens was identified (Table 9-4).

Next, the ability of the identified panels of antigens for a correct classification of the disease states was investigated. Therefore, the WEKA "random forest" classification algorithm was used. Random forest is a classifier that consists of many decision trees and outputs the class that is the mode of the classes output by individual trees [99].

Sixty three percent of the active tuberculosis and 65% of the non-TB samples were correctly classified (Table 4-16) by the 50 antigens identified by the multivariate analysis (Table 9-3). Seven percent of the latent tuberculosis infected patient samples and 98% of the non-TB patient samples were correctly classified (Table 4-17) by the 35 antigens listed in Table 9-4.

Table 4-16: Classification of active / non-TB.

	active TB	**non-TB**
classified as	**n (%)**	**n (%)**
active TB	375 (63%)	217 (37%)
non-TB	223 (35%)	421 (65%)

Table 4-17: Classification of latent / non-TB.

	latent TB	**non-TB**
classified as	**n (%)**	**n (%)**
latent TB	6 (7%)	79 (93%)
non-TB	15 (2%)	629 (98%)

5 Discussion

Tuberculosis is a global disease with an annual death toll of about 2 million people. To date, no easy to use and low cost test for the correct diagnosis of tuberculosis disease states with acceptable sensitivity and specificity is available. The serological response of TB patients is heterogeneous and it is anticipated that a large number of TB antigens would be required to develop a serodiagnostic test for TB [100, 101]. For the analysis of the serological response to a large number of TB antigens, sensitive, accurate and reliable methods are required. This is usually accomplished using solid phase antigen assays such as ELISA [94, 101-105]. However, these methods are characterized by low sample throughput, high reagent consumption, and the analysis of a single antigen per experiment. In contrast, suspension bead arrays are perfectly suited to perform serological assays testing for multiple antigens. Such a set up allows measuring of hundreds of parameters in parallel while at the same time does only requiring a minimum amount of sample and reagent material.

The purposes of this study were:

i) technology transfer from planar microarray technology to the multiplexed bead-based Luminex platform to allow high throughput screening of human samples.
ii) compare and validate the data generated by Kunnath-Velayudhan *et al.,* 2010 [81] with the data gained in this thesis.
iii) generate quality control samples.
iv) characterize the antibody response of different TB patient groups.
v) establish a serological assay consisting of a panel of TB-associated antigens to distinguish between the different TB patient groups.

For these purposes FIND provided the following material:

i) 112 recombinant tuberculosis antigens.

ii) 160 active TB samples and 17 negative controls for assay development and validation.
iii) 644 non-TB, 592 active TB and 85 latent TB human serum samples for marker identification.

5.1 Antigen assay development

In this thesis a multiplexed serological assay on the Luminex platform was developed for the detection of human IgGs in human serum samples derived from active TB, latent TB and non-TB patients.

Appropriate capture antigens are required to set up a multiplexed antigen assay. The performance of the assay is highly dependent on the availability and quality of these components. In addition, the assay is crucially dependent on the properties of the serum samples. The presence of antibodies against *E. coli* bacteria in human serum samples as well as unspecific binding and matrix effects in the tested specimen can lead to false positive or false negative results.

In the first step of the TB-assay development proper assay conditions were defined including an optimal serum sample dilution and a suitable detection antibody concentration. In addition, an assay buffer (Low Cross Buffer + PBS + 0.5% BSA) supplemented with 10% *E. coli* lysate was identified to improve results by pre-adsorption of human serum samples and blocking of unspecific bindings [81, 94, 106].

The coupling of antigens on Luminex MagPlex microspheres using a KingFisher 96 is highly reproducible and allows the coupling of beads for approx. 30 plates (corresponds to 2500 – 3000 samples) within one coupling procedure.
The performance of immobilized antigens in serum sample screenings differed between the antigens and suppliers. Correlation of the same antigens from different coupling batches measured within the same sample set was between 0.22 and 1.00. A failed coupling procedure for distinct antigens within one

coupled bead batch could lead to poor correlation coefficients, which makes a quality control procedure essential.

In a next step, the miniaturized multiplexed assay was validated for its analytical performance. Considering several criteria such as intra- and inter-assay variation and stability, the assay needs to be qualified for further screening tasks [107].

Assay precision was strongly dependent on the sample, antigen and MFI signal intensity. Intra- and inter-assay coefficients of variation were in a wide range between 10% and 70%. CVs on control beads were always <10%. A satisfying precision of the assay is crucial for reliable sample screenings and data analysis. Thus antigens with intra-assay CVs >25% should not be included in a final antigen panel.

Antigens immobilized on beads were stable over a time period of at least seven and for some antigens for up to 15 weeks. Again, antigens that did not show a sufficient stability should be excluded from an optimized panel or must be coupled again.

Further improvements in the performance of the assays may be achieved by i) selecting more stable antigens and ii) by optimizing the matrix conditions, bead batches, and by additional buffer optimization (e.g. adding stabilizers).

In summary, the developed serological assay is well suited for the detection of antibody reactivity towards a huge set of recombinant tuberculosis antigens in human serum samples. The ability to determine multiple analytes in parallel with minimal sample consumption enables the efficient screening of thousands of samples with drastically reduced hands on time and constant quality. Therefore, this technology represents a perfect tool for the rapid and efficient antibody profiling in human body fluids such as serum.

5.2 Quality controls

In order to assure that a test run is valid and results are reliable, quality control (QC) samples should be used in the performance of each assay. The QC samples should be treated in exactly the same manner as the test samples and are used to validate the test run.

In contrast to sandwich immunoassays, where reference samples can be generated by spike-in of a recombinant target antigen, any serological assay is based on the availability of human antibodies specific for the selected antigens. For singleplex assays, e.g. ELISAs [108, 109], normally one high-signal sample is used per antigen as QC sample for each assay plate. For a multiplex assay a sample is required that reacts with all targeted antigens. A commonly used, but not well documented, approach is to create pools from multiple sera in order to cover all targeted antigens [110, 111].

For the first time, a reference sample pool was generated not only by randomly pooling samples but by a systematical and bioinformatics approach using an already measured sample set and integer linear programming.

An algorithm[4] was created for the prediction and calculation of an optimal serum sample pool composition. The decisive factor is to apply as less sample material as possible and to fulfill the same criteria (e.g. dilution of the final sample pool) as for the serum samples to be screened.

Assigning a positive signal of a sample for a distinct antigen with 1 and a negative signal with 0, there are at least two possibilities of the effect of randomly pooling samples:

i) Pooling two samples that are negative for a distinct antigen result in a negative signal of the pooled samples for the same antigen. Having one positive and one negative sample should result in a positive pool signal as well as having two positive samples for one antigen or expressed mathematically: 0 + 0 = 0, 0 + 1 = 1 and 1 + 1 = 1.

[4] The algorithm was created by Hannes Planatscher.

ii) The other possibility would be to observe a kind of additive effect, which means that 0 + 0 is not necessarily 0 or 0 + 1 inconclusively 1.

Considering these possibilities the first task was to investigate the effect of pooling randomly chosen serum samples. This was implemented by adding stepwise one sample to another ending with a pool consisting of six randomly chosen samples. The result of the experiment shown in Table 4-8 and Figure 4-12 confirmed the second hypothesis, that is to say an additive effect of pooling serum samples. This means that the value for a distinct sample pool out of already existing data can be estimated by calculating the mean of the values of the samples that will be used in the pool.
Saturation effects in the input data could explain the increased signals of the measured pool values compared with the calculated values. A larger number of different antibodies for the same antigen originating from different individuals could also explain this observation.

For three antigens the signals of the pooled samples was lower than the artificially calculated threshold. In two cases the signal of the negative control was comparably high. This could be due to a certain reactivity of the negative sample donor towards these few antigens and the sample pool was not able to cover these antigens with a sufficient high signal. To include these antigens as well it may be necessary to pool more than three samples or to replace the high signals of the negative control for these antigens by lower signals from other negative controls. But in this case coverage of 97% of all antigens was more than adequate.

The observed results suggest that the mathematical model for sample pools makes reasonable predictions. With this information it is possible to calculate and predict an optimal pool of serum samples that covers as much antigens as possible in a multiplexed assay. This approach can be easily adapted to other assay formats and constraints.
This pooling strategy provides sufficient material for more than 10000 assay controls with less than 100 µL of each single sample and serves as a pre-aliquoted ready-to-use control.

5.3 Serodiagnostic antibody profiling

Serodiagnosis is considered as a promising approach for tuberculosis diagnostic research. A serological test would be the ideal method for worldwide implementation due to its rapidity, simplicity and low cost. To replace the current gold standard of TB diagnostic, culturing of *M. tuberculosis* from specimen taken from a patient, a serological test should reveal a clinical sensitivity above 80% and test specificities above 95%, according to the recommendations of the World Health Organization (WHO) [112]. To date, no test fulfills these requirements although researchers have been looking for a suitable test for more than a century. This implicates that it is very difficult to identify the “ideal” test.

First, the sensitivity (positive samples correctly classified as positive) and specificity (negative samples correctly classified as negative) was assessed. Serum from a healthy lab member was used as a negative control. An artificial cut off was defined at three times (weak positive), five times (medium positive) and ten times (strong positive) the negative control for each antigen separately. The sensitivity of all antigens ranged between 5% and 82% and the specificity between 28% and 97%. For most antigens the sensitivity was higher than the specificity. Interestingly, the fusion proteins all revealed a higher specificity than sensitivity. However, antigens with adequate sensitivity revealed no adequate specificity (and also vice versa). In summary, all tested antigens lack either satisfactory sensitivity or specificity.

Next, the number of positive samples per antigen was assessed. Positive was defined as three times the value of the negative control. Most positive samples were found in active TB samples, followed by non-TB and latent TB samples. Interestingly, healthy persons also revealed relatively high antibody reactivity to certain TB antigens. This was not expected in such an extent.
Investigating the serological response in latently infected TB patients, the antibody response was significantly lower in this group of patients compared with healthy people. This result is in accordance with the notion that latent infection per se fails to provide sufficient antigenic stimulus to elicit a strong antibody response [113-115].

Using multiplexed bead-based antigen assays combined with univariate analysis, 41 antigens (Table 4-10) were found to be associated with active TB. Most of these proteins were also identified by Kunnath-Velayudhan *et al.*, 2010 [81]. 18 antigens (Table 4-11) are found to be able to discriminate between latent and non-TB. The targets of the antibody response during latent infection differ from those occurring during active disease.
Given an overall preference of the immune response for membrane-bound over soluble antigens, membrane-associated proteins are more frequently targeted during latent TB infection. Membrane-bound antigens might derive from low numbers of live bacilli, dead bacilli, or be found in macrophage-secreted exosomes [116]. However, in this work the identified membrane-bound marker candidates reveal a lower serological antibody response in people infected with latent tuberculosis than in healthy people and thus are not suitable as markers to discriminate between two groups.

Altogether, the results presented here suggest that each tuberculosis state is characterized by a distinct bacterial antigen signature (absence or presence of distinct markers).

Based on the results described in the previous section, the reason for the increased antibody reactivity in non-TB samples was investigated. Previous exposure to *M. tuberculosis*, *M. bovis* or other *Mycobacterium* strains cannot be excluded and can lead to cross-recognition in TB negative persons. 43 from 242 (18%) non-TB patients had past history of TB. These non-TB patients with past TB history revealed increased antibody reactivity to three antigens out of the 41 antigens (Table 4-10), Rv3804, Rv3881 and Rv1860 from suppliers 1 and 2. This finding makes it difficult to distinguish healthy persons from active TB infected patients as these antigens are under the most promising candidate markers for active TB (see Table 4-10).

Smear grade was only modestly connected to antibody reactivity in active TB patients. However, one could assume that as bacillary burden increases with disease, metabolically active bacilli secrete proteins, which elucidate a B cell specific immune response. Thus a stronger association between the response

to the TB-associated proteins and sputum smear grade, which is a surrogate of bacillary burden, is expected. This also suggests a dynamic antibody response that reflects the progression of infection [81].

5.4 Analysis of the key antigen markers

Several antigens were identified with a strong potential to serve as potential tuberculosis markers. It is important that the identified antigens have clinical relevance and support the characteristics of the disease.

Some of the most promising markers are the proteins of the antigen 85 (Ag85) complex. It consists out of four TB protein, three abundantly secreted proteins, Ag85A (Rv3804), Ag85B (Rv1886) and Ag85C (Rv0129) which is a major protein component of the *Mycobacterium tuberculosis* cell wall. Each protein exhibits a mycosyltransferase activity [146] that is required for the biogenesis of trehalose dimycolate (cord factor), a dominant structure necessary for maintaining cell wall integrity [147, 148]. The proteins play a key role in the pathogenesis of tuberculosis as their expression is needed for intracellular survival within macrophages [148]. Additionally, the complex is very immunogenic and vaccination of mice or nonhuman primates with the Ag85 complex stimulates strong humoral and T-cell mediated immune responses and contributes to a significant protection against living *M. tuberculosis* H37Rv [149-151].

Another two well performing markers are Rv3874 and Rv3875, which are encoded in the RD1 region of *Mycobacterium tuberculosis* that is absent in all bacillus Calmette-Guérin strains [152, 153]. Rv1980 (and Rv1984) are found in the RD2 region and are therefore also absent in some bacillus Calmette-Guérin strains [26, 154]. The BCG vaccination status is known for 382 of 1321 patients. A Wilcox-test revealed no significant differences in antibody reactivity between vaccinated and non-vaccinated people (data not shown).

Rv0272 is a hypothetical protein with unknown function that was identified by mass spectrometry in Triton X-114 extracts and culture filtrates of *M. tuberculosis* [155, 156].

Rv3878 is also a member of the RD1 region of *M. tuberculosis* and elicits a prominent immune response in human tuberculosis patients [157]. No references were found in the context of latent TB.

5.5 Heterogeneity of antigen recognition by serum antibodies

Screening of 592 active TB serum samples (343 smear positive / culture positive and 249 smear negative / culture positive) revealed a heterogeneous antibody response pattern in active TB patients among the 41 identified marker candidates. The finding that the antibody response to antigens in tuberculosis is highly heterogeneous is consistent with the results of recent studies of other groups [101, 102]. The varied antibody profiles seen in TB patients make it difficult to exclude host characteristics. However, it is likely that the correlation between antigen load and serum reactivity heavily marks each patients' antibody profile. High variability might be introduced by bacillary load and metabolic state of the patient at the time of testing [117]. Other possibilities are the relative antigen expression of the infecting bacterial strain [118], and relative immunodominance of each protein. The effect of antigen load on frequency of sampling will be greater for some antigens than for others depending on relative antibody avidity. Due to the chronic nature of tuberculosis the above events occur over a considerable period of time. Thus, cross-sectional analyses of those events may accentuate the variation. In general, variation in antigen recognition by the sera is a typically observation in chronic infections [119-122].

In contrast, Samanich *et al.*, 2001 [100] fractionated native culture filtrate proteins from *Mycobacterium tuberculosis* by one-dimensional and two-dimensional polyacrylamide gel electrophoresis and probed the Western Blots with sera from TB patients. They showed that the antibody response in TB patients revealed a significant homogeneity in being directed against a well-defined subset of antigens. Because of the immunogenicity of secreted proteins, culture filtrates are substantially enriched with secreted proteins [10, 123, 124]. In general, using native culture filtrate proteins might be the better choice than recombinant proteins expressed in *E. coli* because of posttranslational modifications, e.g. glycosylation and lipid modifications [104].

Using these filtrate proteins has the limitation that serological reactions only to secreted proteins can be investigated.

In addition, an issue might be cross reactivity of antibodies with crude antigen preparations. Healthy individuals possess antibodies, elicited by exposure to commensal bacteria, environmental bacteria, and vaccinations, that cross react with several *M. tuberculosis* antigens [125-128]. Several antigens from *M. tuberculosis* have been isolated and characterized during the last decades [129] and a majority of them possess significant homology with proteins from other mycobacterial and non-mycobacterial species [130-134]. Thus, healthy and diseased people have antibodies to epitopes against conserved regions of these proteins which lead to a considerable cross-reactivity with crude antigen preparations [135-138].

In infections caused by certain intracellular pathogens, the antibody repertoire is highly diverse. Sera from different human beings react with different antigens. Examples for this include Crohn's disease [139], chlamydial infections [140] and cryptococcosis [141]. It is also assumed that the person-to-person antibody heterogeneity of antigen recognition is a key attribute to the antibody response in tuberculosis rather than the recognition of a limited number of particular antigens. Therefore, in order to improve the performance quality of an antibody detection assay, it is important to combine the responses against a number of different mycobacterial antigens [101]. The fact that antibody responses to *M. tuberculosis* infections are heterogeneous means that the desired sensitivity and specificity in serodiagnostic tests may be achieved only by a combination or fusion of several antigens.

It is striking that four of the top six markers for active tuberculosis identified in this work are either fusion proteins or an antigen cocktail. The three fusion proteins consist of three TB proteins respectively and the antigen mix is a mixture of four single proteins. Several examples of serological assays prove the advantage of fusion proteins and antigen mixes compared with single antigens. Houghton *et al.*, 2002 [142] developed an ELISA with a fusion protein of four antigens, CFP10 (Rv3874), MTB8, MTB48 and the 38 kDa antigen

(Rv0934) together with two additional antigens, DPEP (MPT32) and MTB81 to detect antibodies in TB patients. In samples from HIV negative individuals, the ELISA detected antibodies in >80% of sputum smear positive TB patients and >60% sputum smear negative TB patients, with a specificity of 98%. The antigen combination detected a significant number of TB/HIV co-infected cases, as well as antibodies in patients with extrapulmonary tuberculosis.

Ireton *et al.*, 2010 [143] used selected seroreactive proteins to design new poly-epitope fusion proteins and characterized them by multi-antigen print immunoassay (MAPIA). The vast majority of the TB patients produced antibody responses against these antigens.

Gennaro, M. L., 2000 [144] showed that assays utilizing cocktails of antigens effectively cover the diversity of immune responses and provide more accurate tools for immunologic diagnosis of tuberculosis. These antigen cocktails can also be used for the evaluation of new anti-tuberculosis vaccines.

The antigen cocktail strategy was also pursued by Lyashchenko *et al.,* 2000 [145]. They describe a multi-antigen print immunoassay for cocktail-based serological diagnosis which revealed a better diagnostic sensitivity compared with conventional ELISA.

The results of the above mentioned examples confirmed that an optimal multi-antigen cocktail or fusion proteins should be designed to cover the heterogeneity of antibody responses and thus improve the test sensitivity and specificity.

For a multivariate analysis approach, the cluster function of the program TIGR MeV was used to identify a panel of antigens that may be of diagnostic relevance. The heat map visually depicts the ability of the antigens for a correct classification. A panel of 50 antigens for the discrimination of active and non-TB was identified as well as a set of 35 antigens in the case of latent/non-TB discrimination. However, these identified panels of antigens did not provide sufficient sensitivity and specificity.

Next, the random forest classifier of the machine learning software WEKA was used for the ability of the antigens for a correct classification of the samples according to their clinical status. The random forest was preferred over other

classification methods such as support vector machines (SVM) because the random forest classifier is very fast and efficient in processing huge amounts of data.

Using this approach, only ~64% of the samples were correctly classified according to their clinical status. This is not specific enough to design a diagnostic test.
The allegedly almost perfect classification of 98% is misleading because the group of latent tuberculosis patients (n=85) is relatively small compared with the healthy control group (n=644). An issue of the algorithm is that it prefers to correlate all samples to the larger group because of a lower error rate. The smaller the group is, the smaller the chance of a correct classification.

In summary, this study shows, that a larger panel of antigens did not reveal satisfactory sensitivity and specificity to serve as an assay for a reliable diagnosis of different tuberculosis disease states. The reasons for the poor classifications were discussed above and include increased antibody amounts in non-TB patients with past TB history, cross-recognition with other TB species, or cross-reactivity with crude antigen preparations.

6 Summary and Outlook

In this thesis, a bead based multiplexed serological antigen assay was generated, optimized and validated for the detection of human antibodies in serum of patients infected with tuberculosis.

To ensure reliable and comparable results of the developed assay over a longer period of time quality control samples are always necessary. For the first time, a systematic bioinformatics approach was described to create a pool of serum samples that can be used as quality and positive controls in a multiplexed assay setup to cover all analytes with pre-calculated signal intensities. Using this approach, reference samples for several thousand screenings could be generated with only minimal consumption of sample material.

Screening of more than thousand human serum samples revealed a set of antigens which was able to discriminate different tuberculosis disease states. However, statistical analysis showed a poor sensitivity and specificity of the identified markers. A combination of these antigens did not lead to an improved sensitivity and specificity.

In summary, this study shows that the developed assay based on the available material and considering the complex nature of the disease is not able to clearly discriminate different tuberculosis disease states. However, the identified candidate antigens are promising for future work (e.g. the results generated by Kunnath-Velayudhan *et al.*, 2010 [81] were confirmed) but further studies are necessary. The assay described here can be used as an excellent research tool to gain further insight into the endemic and demographic properties of the disease. Thus, in the near future, endemic studies are planned to get an overview if different patterns of antibody reactivity can be observed. Detection of childhood tuberculosis is a very challenging task, as the results with current diagnostics (sputum smear microscopy and culture) often show false-negative results. Therefore, a rapid and specific serological test would also improve the medical care of tuberculosis in children. This opens an additional research area

where the assay developed during this thesis can serve as a promising research tool.

7 Zusammenfassung

In dieser Arbeit wurde ein Bead-basierter multiplexer serologischer Assay zur Detektion von humanen Antikörpern in Serum von mit Tuberkulose infizierten Patienten entwickelt, optimiert und validiert.

Um mit dem entwickelten Assay über einen längeren Zeitraum zuverlässige und vergleichbare Ergebnisse gewährleisten zu können, ist der Einsatz von Qualitätskontrollen notwendig. Erstmalig wurde ein systematischer bioinformatischer Ansatz zur Herstellung eines Probenpools beschrieben, der als Qualitäts- und Positivkontrolle in diesem multiplexen Assay verwendet werden kann. Hierdurch war es möglich Referenzproben für mehrere tausend Messungen herzustellen, für die nur sehr geringe Mengen an Probenmaterial eingesetzt werden mussten.

Die Analyse von mehr als eintausend humanen Serumproben ergab ein Set von Antigenen, das zwischen verschiedenen Krankheitszuständen bei Tuberkulose unterscheiden kann. Jedoch zeigte eine statistische Analyse der Daten eine schlechte Sensitivität und Spezifität der identifizierten Marker. Auch eine Kombination dieser identifizierten Antigene führte zu keiner verbesserten Sensitivität und Spezifität des Tests.

Zusammenfassend zeigt diese Studie, dass der entwickelte Assay mit dem verwendeten Material und aufgrund der komplexen Natur der Krankheit nicht zwischen verschiedenen Krankheitszuständen der Tuberkulose unterscheiden kann. Trotzdem erwiesen sich die identifizierten Antigene als vielversprechend (so konnten zum Beispiel die Ergebnisse von Kunnath-Velayudhan *et al.*, 2010 [81] bestätigt werden), jedoch sind hierzu weitere Studien notwendig. Der hier entwickelte Test kann als Forschungswerkzeug für weitere Studien herangezogen werden, um einen besseren Einblick in die endemischen und demographischen Unterschiede dieser Krankheit zu bekommen. In naher Zukunft sind hierfür endemische Studien geplant um eine Übersicht über verschiedene Muster der Immunantwort zu bekommen. Auch die Detektion von Tuberkulose bei Kindern ist eine große Herausforderung, da die aktuellen

diagnostischen Verfahren (Mikroskopie und Kultur) oft falsch-negative Ergebnisse liefern. Ein schneller und spezifischer serologischer Test würde somit auch die medizinische Versorgung bei mit Tuberkulose infizierten Kindern erheblich verbessern.

8 Literature

1. Netto, E.M., C. Dye, and M.C. Raviglione, *Progress in global tuberculosis control 1995-1996, with emphasis on 22 high-incidence countries. Global Monitoring and Surveillance Project.* Int J Tuberc Lung Dis, 1999. **3**(4): p. 310-20.

2. Raviglione, M.C., D.E. Snider, Jr., and A. Kochi, *Global epidemiology of tuberculosis. Morbidity and mortality of a worldwide epidemic.* JAMA, 1995. **273**(3): p. 220-6.

3. Bloom, B.R. and C.J. Murray, *Tuberculosis: commentary on a reemergent killer.* Science, 1992. **257**(5073): p. 1055-64.

4. Kochi, A., *The global tuberculosis situation and the new control strategy of the World Health Organization.* Tubercle, 1991. **72**(1): p. 1-6.

5. WHO, *The global plan to stop TB 2011-2015*, 2010: Geneva, Switzerland.

6. Zwerling, A., et al., *The BCG World Atlas: a database of global BCG vaccination policies and practices.* PLoS Med. **8**(3): p. e1001012.

7. Cole, S.T., et al., *Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence.* Nature, 1998. **393**(6685): p. 537-44.

8. *TubercuList*. Version 2.5, 24.12.2011 27.02.2012]; Available from: http://tuberculist.epfl.ch/.

9. Andersen, A.B., *Mycobacterium tuberculosis proteins. Structure, function, and immunological relevance.* Dan Med Bull, 1994. **41**(2): p. 205-15.

10. Nagai, S., et al., *Isolation and partial characterization of major protein antigens in the culture fluid of Mycobacterium tuberculosis.* Infect Immun, 1991. **59**(1): p. 372-82.

11. Sorensen, A.L., et al., *Purification and characterization of a low-molecular-mass T-cell antigen secreted by Mycobacterium tuberculosis.* Infect Immun, 1995. **63**(5): p. 1710-7.

12. Sonnenberg, M.G. and J.T. Belisle, *Definition of Mycobacterium tuberculosis culture filtrate proteins by two-dimensional polyacrylamide gel electrophoresis, N-terminal amino acid sequencing, and electrospray mass spectrometry.* Infect Immun, 1997. **65**(11): p. 4515-24.

13. Rosenkrands, I., et al., *Mapping and identification of Mycobacterium tuberculosis proteins by two-dimensional gel electrophoresis, microsequencing and immunodetection.* Electrophoresis, 2000. **21**(5): p. 935-48.

14. *TubercuList World Wide Web Server*. Version 3.1, 01.10.2008 27.02.2012]; Available from: http://genolist.pasteur.fr/TubercuList/.

15. Henderson, R.A., S.C. Watkins, and J.L. Flynn, *Activation of human dendritic cells following infection with Mycobacterium tuberculosis.* J Immunol, 1997. **159**(2): p. 635-43.

16. Thurnher, M., et al., *Bacillus Calmette-Guerin mycobacteria stimulate human blood dendritic cells.* Int J Cancer, 1997. **70**(1): p. 128-34.

17. Holt, P.G. and M.A. Schon-Hegrad, *Localization of T cells, macrophages and dendritic cells in rat respiratory tract tissue: implications for immune function studies.* Immunology, 1987. **62**(3): p. 349-56.

18. Oppmann, B., et al., *Novel p19 protein engages IL-12p40 to form a cytokine, IL-23, with biological activities similar as well as distinct from IL-12.* Immunity, 2000. **13**(5): p. 715-25.

19. Dinarello, C.A., *Biologic basis for interleukin-1 in disease.* Blood, 1996. **87**(6): p. 2095-147.

20. Tsenova, L., et al., *Tumor necrosis factor alpha is a determinant of pathogenesis and disease progression in mycobacterial infection in the central nervous system.* Proc Natl Acad Sci U S A, 1999. **96**(10): p. 5657-62.

21. Boussiotis, V.A., et al., *IL-10-producing T cells suppress immune responses in anergic tuberculosis patients.* J Clin Invest, 2000. **105**(9): p. 1317-25.

22. Hirsch, C.S., et al., *In vitro restoration of T cell responses in tuberculosis and augmentation of monocyte effector function against Mycobacterium tuberculosis by natural inhibitors of transforming growth factor beta.* Proc Natl Acad Sci U S A, 1997. **94**(8): p. 3926-31.

23. Toossi, Z. and J.J. Ellner, *The role of TGF beta in the pathogenesis of human tuberculosis.* Clin Immunol Immunopathol, 1998. **87**(2): p. 107-14.

24. van Crevel, R., T.H. Ottenhoff, and J.W. van der Meer, *Innate immunity to Mycobacterium tuberculosis.* Clin Microbiol Rev, 2002. **15**(2): p. 294-309.

25. Immunotech, O. *http://www.oxfordimmunotec.com/Tuberculosis_International*. 2010 13.12.2011].

26. Behr, M.A., et al., *Comparative genomics of BCG vaccines by whole-genome DNA microarray.* Science, 1999. **284**(5419): p. 1520-3.

27. Brewer, T.F. and G.A. Colditz, *Relationship between bacille Calmette-Guerin (BCG) strains and the efficacy of BCG vaccine in the prevention of tuberculosis.* Clin Infect Dis, 1995. **20**(1): p. 126-35.

28. WHO, *Treatment of tuberculosis Guidlines 4th edition*, 2010.

29. (CDC), C.f.D.C.a.P. *The Difference Between Latent TB Infection and Active TB Disease*. 2010; Available from: www.cdc.gov/tb.

30. Selwyn, P.A., et al., *A prospective study of the risk of tuberculosis among intravenous drug users with human immunodeficiency virus infection.* N Engl J Med, 1989. **320**(9): p. 545-50.

31. Dinnes, J., et al., *A systematic review of rapid diagnostic tests for the detection of tuberculosis infection.* Health Technol Assess, 2007. **11**(3): p. 1-196.

32. Huebner, R.E., et al., *Delayed-type hypersensitivity anergy in human immunodeficiency virus-infected persons screened for infection with Mycobacterium tuberculosis.* Clin Infect Dis, 1994. **19**(1): p. 26-32.

33. Huebner, R.E., M.E. Villarino, and D.E. Snider, Jr., *Tuberculin skin testing and the HIV epidemic.* JAMA, 1992. **267**(3): p. 409-10.

34. Caiaffa, W.T., et al., *Instability of delayed-type hypersensitivity skin test anergy in human immunodeficiency virus infection.* Arch Intern Med, 1995. **155**(19): p. 2111-7.

35. Brock, I., et al., *Comparison of tuberculin skin test and new specific blood test in tuberculosis contacts.* Am J Respir Crit Care Med, 2004. **170**(1): p. 65-9.

36. Detjen, A.K., et al., *Interferon-gamma release assays improve the diagnosis of tuberculosis and nontuberculous mycobacterial disease in children in a country with a low incidence of tuberculosis.* Clin Infect Dis, 2007. **45**(3): p. 322-8.

37. Pai, M., L.W. Riley, and J.M. Colford, Jr., *Interferon-gamma assays in the immunodiagnosis of tuberculosis: a systematic review.* Lancet Infect Dis, 2004. **4**(12): p. 761-76.

38. Whittaker, E., A. Gordon, and B. Kampmann, *Is IP-10 a better biomarker for active and latent tuberculosis in children than IFNgamma?* PLoS One, 2008. **3**(12): p. e3901.

39. Mori, T., et al., *Specific detection of tuberculosis infection: an interferon-gamma-based assay using new antigens.* Am J Respir Crit Care Med, 2004. **170**(1): p. 59-64.

40. Kang, Y.A., et al., *Discrepancy between the tuberculin skin test and the whole-blood interferon gamma assay for the diagnosis of latent tuberculosis infection in an intermediate tuberculosis-burden country.* JAMA, 2005. **293**(22): p. 2756-61.

41. Chapman, A.L., et al., *Rapid detection of active and latent tuberculosis infection in HIV-positive individuals by enumeration of Mycobacterium tuberculosis-specific T cells.* AIDS, 2002. **16**(17): p. 2285-93.

42. Lalvani, A. and K.A. Millington, *T cell-based diagnosis of childhood tuberculosis infection.* Curr Opin Infect Dis, 2007. **20**(3): p. 264-71.

43. Aabye, M.G., et al., *The impact of HIV infection and CD4 cell count on the performance of an interferon gamma release assay in patients with pulmonary tuberculosis.* PLoS One, 2009. **4**(1): p. e4220.

44. Griffith, D.E. and C.M. Kerr, *Tuberculosis: disease of the past, disease of the present.* J Perianesth Nurs, 1996. **11**(4): p. 240-5.

45. Behr, M.A., et al., *Transmission of Mycobacterium tuberculosis from patients smear-negative for acid-fast bacilli.* Lancet, 1999. **353**(9151): p. 444-9.

46. Nicas, M., W.W. Nazaroff, and A. Hubbard, *Toward understanding the risk of secondary airborne infection: emission of respirable pathogens.* J Occup Environ Hyg, 2005. **2**(3): p. 143-54.

47. Raviglione, M.C., J.P. Narain, and A. Kochi, *HIV-associated tuberculosis in developing countries: clinical features, diagnosis, and treatment.* Bull World Health Organ, 1992. **70**(4): p. 515-26.

48. Shafer, R.W. and B.R. Edlin, *Tuberculosis in patients infected with human immunodeficiency virus: perspective on the past decade.* Clin Infect Dis, 1996. **22**(4): p. 683-704.

49. Barnes, P.F., et al., *Tuberculosis in patients with human immunodeficiency virus infection.* N Engl J Med, 1991. **324**(23): p. 1644-50.

50. Selwyn, P.A., et al., *Prospective study of human immunodeficiency virus infection and pregnancy outcomes in intravenous drug users.* JAMA, 1989. **261**(9): p. 1289-94.

51. Miller, L.G., et al., *A population-based survey of tuberculosis symptoms: how atypical are atypical presentations?* Clin Infect Dis, 2000. **30**(2): p. 293-9.

52. Sherman, L.F., et al., *Patient and health care system delays in the diagnosis and treatment of tuberculosis.* Int J Tuberc Lung Dis, 1999. **3**(12): p. 1088-95.

53. WHO, *WHO Tuberculosis Diagnostics Workshop -product development guidelines-*, 1997: Cleveland, Ohio.

54. Welch, D.F., et al., *Timely culture for mycobacteria which utilizes a microcolony method.* J Clin Microbiol, 1993. **31**(8): p. 2178-84.

55. Steingart, K.R., et al., *Sputum processing methods to improve the sensitivity of smear microscopy for tuberculosis: a systematic review.* Lancet Infect Dis, 2006. **6**(10): p. 664-74.

56. Elliott, A.M., et al., *Negative sputum smear results in HIV-positive patients with pulmonary tuberculosis in Lusaka, Zambia.* Tuber Lung Dis, 1993. **74**(3): p. 191-4.

57. Johnson, J.L., et al., *Impact of human immunodeficiency virus type-1 infection on the initial bacteriologic and radiographic manifestations of pulmonary tuberculosis in Uganda. Makerere University-Case Western Reserve University Research Collaboration.* Int J Tuberc Lung Dis, 1998. **2**(5): p. 397-404.

58. Shingadia, D. and V. Novelli, *Diagnosis and treatment of tuberculosis in children.* Lancet Infect Dis, 2003. **3**(10): p. 624-32.

59. Perkins, M.D., G. Roscigno, and A. Zumla, *Progress towards improved tuberculosis diagnostics for developing countries.* Lancet, 2006. **367**(9514): p. 942-3.

60. Small, P.M. and M.D. Perkins, *More rigour needed in trials of new diagnostic agents for tuberculosis.* Lancet, 2000. **356**(9235): p. 1048-9.

61. Walsh, A. and R. McNerney, *Guidelines for establishing trials of new tests to diagnose tuberculosis in endemic countries.* Int J Tuberc Lung Dis, 2004. **8**(5): p. 609-13.

62. Boehme, C., et al., *Detection of mycobacterial lipoarabinomannan with an antigen-capture ELISA in unprocessed urine of Tanzanian patients with suspected tuberculosis.* Trans R Soc Trop Med Hyg, 2005. **99**(12): p. 893-900.

63. Tessema, T.A., et al., *Diagnostic evaluation of urinary lipoarabinomannan at an Ethiopian tuberculosis centre.* Scand J Infect Dis, 2001. **33**(4): p. 279-84.

64. Kunter, E., et al., *The value of pleural fluid anti-A60 IgM in BCG-vaccinated tuberculous pleurisy patients.* Clin Microbiol Infect, 2003. **9**(3): p. 212-20.

65. Tessema, T.A., et al., *Clinical and radiological features in relation to urinary excretion of lipoarabinomannan in Ethiopian tuberculosis patients.* Scand J Infect Dis, 2002. **34**(3): p. 167-71.

66. Schmidt, R., et al., *Single-molecule detection on a protein-array assay platform for the exposure of a tuberculosis antigen.* J Proteome Res, 2011. **10**(3): p. 1316-22.

67. Napolitano, D.R., et al., *Identification of Mycobacterium tuberculosis ornithine carboamyltransferase in urine as a possible molecular marker of active pulmonary tuberculosis.* Clin Vaccine Immunol, 2008. **15**(4): p. 638-43.

68. Singh, K.K., et al., *Combined use of serum and urinary antibody for diagnosis of tuberculosis.* J Infect Dis, 2003. **188**(3): p. 371-7.

69. Choudhry, V. and R.K. Saxena, *Detection of Mycobacterium tuberculosis antigens in urinary proteins of tuberculosis patients.* Eur J Clin Microbiol Infect Dis, 2002. **21**(1): p. 1-5.

70. Kashino, S.S., et al., *Identification and characterization of Mycobacterium tuberculosis antigens in urine of patients with active pulmonary tuberculosis: an innovative and alternative approach of antigen discovery of useful microbial molecules.* Clin Exp Immunol, 2008. **153**(1): p. 56-62.

71. Menzies, D., M. Pai, and G. Comstock, *Meta-analysis: new tests for the diagnosis of latent tuberculosis infection: areas of uncertainty and recommendations for research.* Ann Intern Med, 2007. **146**(5): p. 340-54.

72. Brock, I., et al., *Latent tuberculosis in HIV positive, diagnosed by the M. tuberculosis specific interferon-gamma test.* Respir Res, 2006. **7**: p. 56.

73. Pai, M., A. Zwerling, and D. Menzies, *Systematic review: T-cell-based assays for the diagnosis of latent tuberculosis infection: an update.* Ann Intern Med, 2008. **149**(3): p. 177-84.

74. Pai, M., et al., *Sensitivity of a whole-blood interferon-gamma assay among patients with pulmonary tuberculosis and variations in T-cell responses during anti-tuberculosis treatment.* Infection, 2007. **35**(2): p. 98-103.

75. Tsiouris, S.J., et al., *Sensitivity analysis and potential uses of a novel gamma interferon release assay for diagnosis of tuberculosis.* J Clin Microbiol, 2006. **44**(8): p. 2844-50.

76. Ferrara, G., et al., *Use in routine clinical practice of two commercial blood tests for diagnosis of infection with Mycobacterium tuberculosis: a prospective study.* Lancet, 2006. **367**(9519): p. 1328-34.

77. Luetkemeyer, A.F., et al., *Comparison of an interferon-gamma release assay with tuberculin skin testing in HIV-infected individuals.* Am J Respir Crit Care Med, 2007. **175**(7): p. 737-42.

78. Ruhwald, M., et al., *IP-10, MCP-1, MCP-2, MCP-3, and IL-1RA hold promise as biomarkers for infection with M. tuberculosis in a whole blood based T-cell assay.* BMC Res Notes, 2009. **2**: p. 19.

79. Ruhwald, M., et al., *Evaluating the potential of IP-10 and MCP-2 as biomarkers for the diagnosis of tuberculosis.* Eur Respir J, 2008. **32**(6): p. 1607-15.

80. Gaseitsiwe, S., et al., *Pattern recognition in pulmonary tuberculosis defined by high content peptide microarray chip analysis representing 61 proteins from M. tuberculosis.* PLoS One, 2008. **3**(12): p. e3840.

81. Kunnath-Velayudhan, S., et al., *Dynamic antibody responses to the Mycobacterium tuberculosis proteome.* Proc Natl Acad Sci U S A. **107**(33): p. 14703-8.

82. LuminexCorporation. *http://www.luminexcorp.com/*. 2011.

83. Dasso, J., et al., *A comparison of ELISA and flow microsphere-based assays for quantification of immunoglobulins.* J Immunol Methods, 2002. **263**(1-2): p. 23-33.

84. Dupont, N.C., et al., *Validation and comparison of luminex multiplex cytokine analysis kits with ELISA: determinations of a panel of nine cytokines in clinical sample culture supernatants.* J Reprod Immunol, 2005. **66**(2): p. 175-91.

85. Albeck, J.G., et al., *Collecting and organizing systematic sets of protein data.* Nat Rev Mol Cell Biol, 2006. **7**(11): p. 803-12.

86. Eisen, M.B., et al., *Cluster analysis and display of genome-wide expression patterns.* Proc Natl Acad Sci U S A, 1998. **95**(25): p. 14863-8.

87. Cobb, J.P., et al., *Application of genome-wide expression analysis to human health and disease.* Proc Natl Acad Sci U S A, 2005. **102**(13): p. 4801-6.

88. Frank, E., et al., *Data mining in bioinformatics using Weka.* Bioinformatics, 2004. **20**(15): p. 2479-81.

89. Witten, I.H., Eibe, F., *Data mining: practical machine learning tools and techniques*. 2nd ed2005, San Francisco: Elsevier.

90. Saeed, A.I., et al., *TM4: a free, open-source system for microarray data management and analysis.* Biotechniques, 2003. **34**(2): p. 374-8.

91. Saeed, A.I., et al., *TM4 microarray software suite.* Methods Enzymol, 2006. **411**: p. 134-93.

92. Tusher, V.G., R. Tibshirani, and G. Chu, *Significance analysis of microarrays applied to the ionizing radiation response.* Proc Natl Acad Sci U S A, 2001. **98**(9): p. 5116-21.

93. Invitrogen. *ProtoArray*. [cited 2011 23.08.2011]; Available from: http://www.invitrogen.com/site/us/en/home/Products-and-Services/Applications/Protein-Expression-and-Analysis/Biomarker-Discovery/ProtoArray.html?CID=fl-protoarray.

94. Laal, S., et al., *Human humoral responses to antigens of Mycobacterium tuberculosis: immunodominance of high-molecular-mass antigens.* Clin Diagn Lab Immunol, 1997. **4**(1): p. 49-56.

95. Williamson, D.F., R.A. Parker, and J.S. Kendrick, *The box plot: a simple visual method to interpret data.* Ann Intern Med, 1989. **110**(11): p. 916-21.

96. Mann, H.B., Whitney, D.R., *On a test of whether one of 2 random variables is stochastically larger than the other.* Annals of Mathematical Statistics, 1947. **18**: p. 50-60.

97. Henkel, R.E., *Tests of Significance*1976, CA: Sage Publications.

98. McLachlan, G.J., Do, Kim-Anh, Ambroise, Christophe, *Analyzing Microarray Gene Expression Data*, ed. I. John Wiley & Sons. Vol. 1. 2004, New Jersey.

99. Breiman, L., *Machine Learning*. Vol. 45. 2001, The Netherlands: Kluwer Academic Publishers.

100. Samanich, K., J.T. Belisle, and S. Laal, *Homogeneity of antibody responses in tuberculosis patients.* Infect Immun, 2001. **69**(7): p. 4600-9.

101. Lyashchenko, K., et al., *Heterogeneous antibody responses in tuberculosis.* Infect Immun, 1998. **66**(8): p. 3936-40.

102. Wu, X., et al., *Humoral immune responses against the Mycobacterium tuberculosis 38-kilodalton, MTB48, and CFP-10/ESAT-6 antigens in tuberculosis.* Clin Vaccine Immunol. **17**(3): p. 372-5.

103. Davidow, A., et al., *Antibody profiles characteristic of Mycobacterium tuberculosis infection state.* Infect Immun, 2005. **73**(10): p. 6846-51.

104. Julian, E., et al., *Comparison of antibody responses to a potential combination of specific glycolipids and proteins for test sensitivity improvement in tuberculosis serodiagnosis.* Clin Diagn Lab Immunol, 2004. **11**(1): p. 70-6.

105. Laal, S., et al., *Surrogate marker of preclinical tuberculosis in human immunodeficiency virus infection: antibodies to an 88-kDa secreted antigen of Mycobacterium tuberculosis.* J Infect Dis, 1997. **176**(1): p. 133-43.

106. Davies, D.H., et al., *Profiling the humoral immune response to infection by using proteome microarrays: high-throughput vaccine and diagnostic antigen discovery.* Proc Natl Acad Sci U S A, 2005. **102**(3): p. 547-52.

107. Findlay, J.W., et al., *Validation of immunoassays for bioanalysis: a pharmaceutical industry perspective.* J Pharm Biomed Anal, 2000. **21**(6): p. 1249-73.

108. Jacobson, R.H., *Validation of serological assays for diagnosis of infectious diseases.* Rev Sci Tech, 1998. **17**(2): p. 469-526.

109. Simmons, J.H., *Development, application, and quality control of serology assays used for diagnostic monitoring of laboratory nonhuman primates.* ILAR J, 2008. **49**(2): p. 157-69.

110. Wong, S.J., et al., *Detection of human anti-flavivirus antibodies with a west nile virus recombinant antigen microsphere immunoassay.* J Clin Microbiol, 2004. **42**(1): p. 65-72.

111. Cooley, G., et al., *High throughput selection of effective serodiagnostics for Trypanosoma cruzi infection.* PLoS Negl Trop Dis, 2008. **2**(10): p. e316.

112. WHO, *Tuberculosis Diagnostics Workshop: product development guidelines.* 1997: p. 1-27.

113. Bothamley, G.H., *Serological diagnosis of tuberculosis.* Eur Respir J Suppl, 1995. **20**: p. 676s-688s.

114. Kaplan, M.H. and M.W. Chase, *Antibodies to mycobacteria in human tuberculosis. I. Development of antibodies before and after antimicrobial therapy.* J Infect Dis, 1980. **142**(6): p. 825-34.

115. Turneer, M., et al., *Humoral immune response in human tuberculosis: immunoglobulins G, A, and M directed against the purified P32 protein antigen of Mycobacterium bovis bacillus Calmette-Guerin.* J Clin Microbiol, 1988. **26**(9): p. 1714-9.

116. Russell, D.G., *Who puts the tubercle in tuberculosis?* Nat Rev Microbiol, 2007. **5**(1): p. 39-47.

117. Garton, N.J., et al., *Cytological and transcript analyses reveal fat and lazy persister-like bacilli in tuberculous sputum.* PLoS Med, 2008. **5**(4): p. e75.

118. Pheiffer, C., et al., *Protein expression in Mycobacterium tuberculosis differs with growth stage and strain type.* Clin Chem Lab Med, 2002. **40**(9): p. 869-75.

119. Eberhardt, C., et al., *Proteomic analysis of the bacterial pathogen Bartonella henselae and identification of immunogenic proteins for serodiagnosis.* Proteomics, 2009. **9**(7): p. 1967-81.

120. Haas, G., et al., *Immunoproteomics of Helicobacter pylori infection and relation to gastric disease.* Proteomics, 2002. **2**(3): p. 313-24.

121. Wehmhoner, D., et al., *Inter- and intraclonal diversity of the Pseudomonas aeruginosa proteome manifests within the secretome.* J Bacteriol, 2003. **185**(19): p. 5807-14.

122. Beare, P.A., et al., *Candidate antigens for Q fever serodiagnosis revealed by immunoscreening of a Coxiella burnetii protein microarray.* Clin Vaccine Immunol, 2008. **15**(12): p. 1771-9.

123. Malen, H., et al., *Comprehensive analysis of exported proteins from Mycobacterium tuberculosis H37Rv.* Proteomics, 2007. **7**(10): p. 1702-18.

124. Malen, H., T. Softeland, and H.G. Wiker, *Antigen analysis of Mycobacterium tuberculosis H37Rv culture filtrate proteins.* Scand J Immunol, 2008. **67**(3): p. 245-52.

125. Bardana, E.J., Jr., et al., *Universal occurrence of antibodies to tubercle bacilli in sera from non-tuberculous and tuberculous individuals.* Clin Exp Immunol, 1973. **13**(1): p. 65-77.

126. Das, S., et al., *The pattern of mycobacterial antigen recognition in sera from Mantoux-negative individuals is essentially unaffected by bacille Calmette-Guerin (BCG) vaccination in either south India or London.* Clin Exp Immunol, 1992. **89**(3): p. 402-6.

127. Havlir, D.V., et al., *Human immune response to Mycobacterium tuberculosis antigens.* Infect Immun, 1991. **59**(2): p. 665-70.

128. Verbon, A., et al., *Antigens in culture supernatant of Mycobacterium tuberculosis: epitopes defined by monoclonal and human antibodies.* J Gen Microbiol, 1990. **136**(5): p. 955-64.

129. Young, D.B., et al., *Mycobacterial protein antigens: a compilation.* Mol Microbiol, 1992. **6**(2): p. 133-45.

130. Andersen, A.B., P. Andersen, and L. Ljungqvist, *Structure and function of a 40,000-molecular-weight protein antigen of Mycobacterium tuberculosis.* Infect Immun, 1992. **60**(6): p. 2317-23.

131. Andersen, A.B. and E.B. Hansen, *Structure and mapping of antigenic domains of protein antigen b, a 38,000-molecular-weight protein of Mycobacterium tuberculosis.* Infect Immun, 1989. **57**(8): p. 2481-8.

132. Garsia, R.J., et al., *Homology of the 70-kilodalton antigens from Mycobacterium leprae and Mycobacterium bovis with the Mycobacterium tuberculosis 71-kilodalton antigen and with the conserved heat shock protein 70 of eucaryotes.* Infect Immun, 1989. **57**(1): p. 204-12.

133. Carlin, N.I., S. Lofdahl, and M. Magnusson, *Monoclonal antibodies specific for elongation factor Tu and complete nucleotide sequence of the tuf gene in Mycobacterium tuberculosis.* Infect Immun, 1992. **60**(8): p. 3136-42.

134. Espitia, C., et al., *Antigenic and structural similarities between Mycobacterium tuberculosis 50- to 55-kilodalton and Mycobacterium bovis BCG 45- to 47-kilodalton antigens.* Infect Immun, 1995. **63**(2): p. 580-4.

135. Davenport, M.P., et al., *The variable C-terminal region of the Mycobacterium leprae 70-kilodalton heat shock protein is the target for humoral immune responses.* Infect Immun, 1992. **60**(3): p. 1170-7.

136. Grandia, A.A., et al., *Natural antibodies to 65 kD mycobacterial heat shock protein in rats do not correlate with susceptibility for Mycobacterium tuberculosis induced adjuvant arthritis.* Immunobiology, 1991. **182**(2): p. 127-34.

137. Meeker, H.C., et al., *Analysis of human antibody epitopes on the 65-kilodalton protein of Mycobacterium leprae by using synthetic peptides.* Infect Immun, 1989. **57**(12): p. 3689-94.

138. Thole, J.E., et al., *Characterization, sequence determination, and immunogenicity of a 64-kilodalton protein of Mycobacterium bovis BCG expressed in escherichia coli K-12.* Infect Immun, 1987. **55**(6): p. 1466-75.

139. Elsaghier, A., et al., *Antibodies to Mycobacterium paratuberculosis-specific protein antigens in Crohn's disease.* Clin Exp Immunol, 1992. **90**(3): p. 503-8.

140. Essig, A., et al., *Analysis of the humoral immune response to Chlamydia pneumoniae by immunoblotting and immunoprecipitation.* Clin Diagn Lab Immunol, 1999. **6**(6): p. 819-25.

141. Chen, L.C., et al., *Antibody response to Cryptococcus neoformans proteins in rodents and humans.* Infect Immun, 1999. **67**(5): p. 2218-24.

142. Houghton, R.L., et al., *Use of multiepitope polyproteins in serodiagnosis of active tuberculosis.* Clin Diagn Lab Immunol, 2002. **9**(4): p. 883-91.

143. Ireton, G.C., et al., *Identification of Mycobacterium tuberculosis antigens of high serodiagnostic value.* Clin Vaccine Immunol, 2010. **17**(10): p. 1539-47.

144. Gennaro, M.L., *Immunologic diagnosis of tuberculosis.* Clin Infect Dis, 2000. **30 Suppl 3**: p. S243-6.

145. Lyashchenko, K.P., et al., *A multi-antigen print immunoassay for the development of serological diagnosis of infectious diseases.* J Immunol Methods, 2000. **242**(1-2): p. 91-100.

146. Kremer, L., et al., *The M. tuberculosis antigen 85 complex and mycolyltransferase activity.* Lett Appl Microbiol, 2002. **34**(4): p. 233-7.

147. Ronning, D.R., et al., *Crystal structure of the secreted form of antigen 85C reveals potential targets for mycobacterial drugs and vaccines.* Nat Struct Biol, 2000. **7**(2): p. 141-6.

148. Armitige, L.Y., et al., *Disruption of the genes encoding antigen 85A and antigen 85B of Mycobacterium tuberculosis H37Rv: effect on growth in culture and in macrophages.* Infect Immun, 2000. **68**(2): p. 767-78.

149. Huygen, K., et al., *Immunogenicity and protective efficacy of a tuberculosis DNA vaccine.* Nat Med, 1996. **2**(8): p. 893-8.

150. Kamath, A.T., et al., *Differential protective efficacy of DNA vaccines expressing secreted proteins of Mycobacterium tuberculosis.* Infect Immun, 1999. **67**(4): p. 1702-7.

151. Baldwin, S.L., et al., *Immunogenicity and protective efficacy of DNA vaccines encoding secreted and non-secreted forms of Mycobacterium tuberculosis Ag85A.* Tuber Lung Dis, 1999. **79**(4): p. 251-9.

152. van Ingen, J., et al., *Region of difference 1 in nontuberculous Mycobacterium species adds a phylogenetic and taxonomical character.* J Bacteriol, 2009. **191**(18): p. 5865-7.

153. Soman, S., et al., *Presence of region of difference 1 among clinical isolates of Mycobacterium tuberculosis from India.* J Clin Microbiol, 2007. **45**(10): p. 3480-1.

154. Kozak, R.A., et al., *Region of difference 2 contributes to virulence of Mycobacterium tuberculosis.* Infect Immun, 2011. **79**(1): p. 59-66.

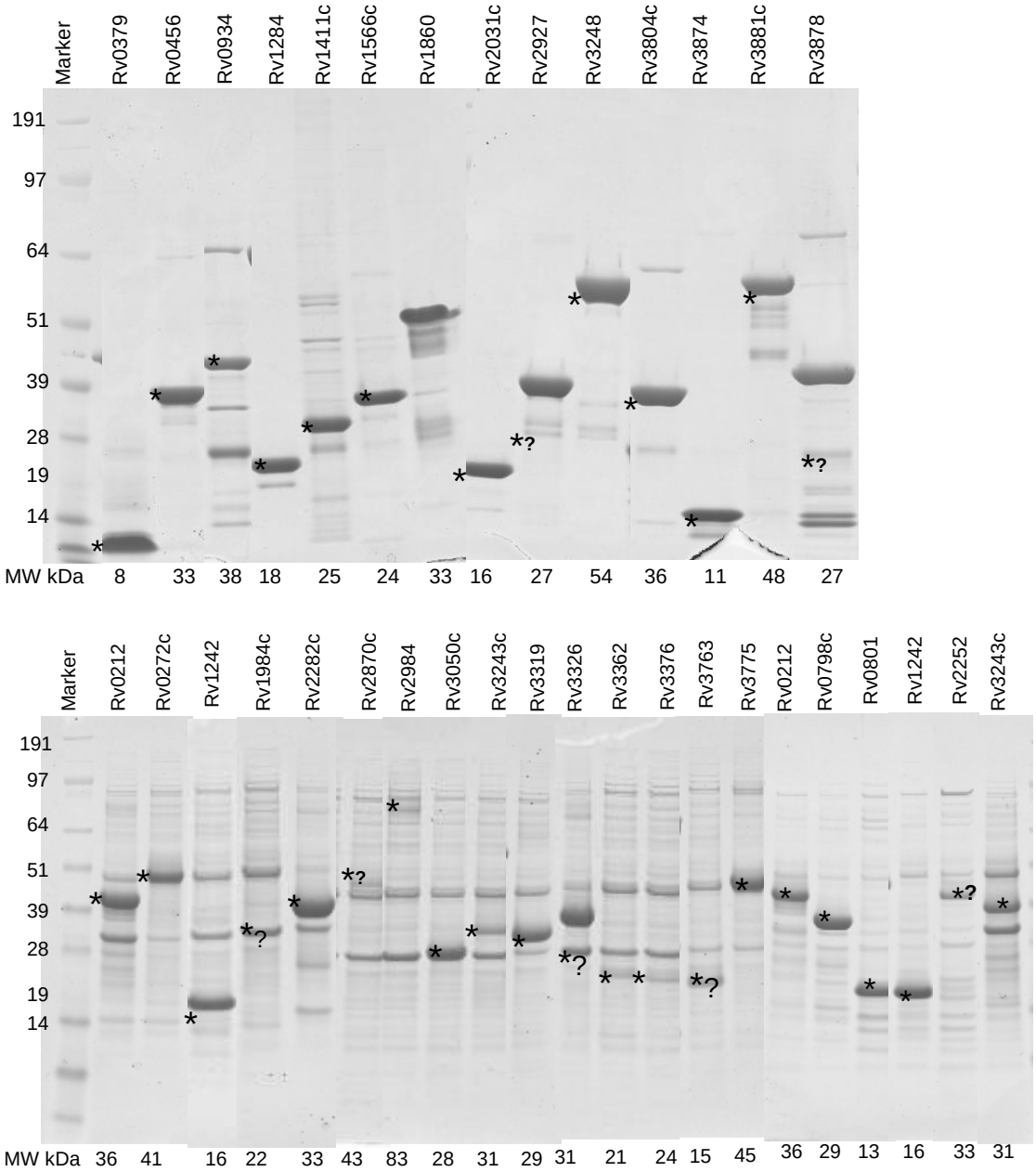
155. Malen, H., et al., *Definition of novel cell envelope associated proteins in Triton X-114 extracts of Mycobacterium tuberculosis H37Rv.* BMC Microbiol, 2010. **10**: p. 132.

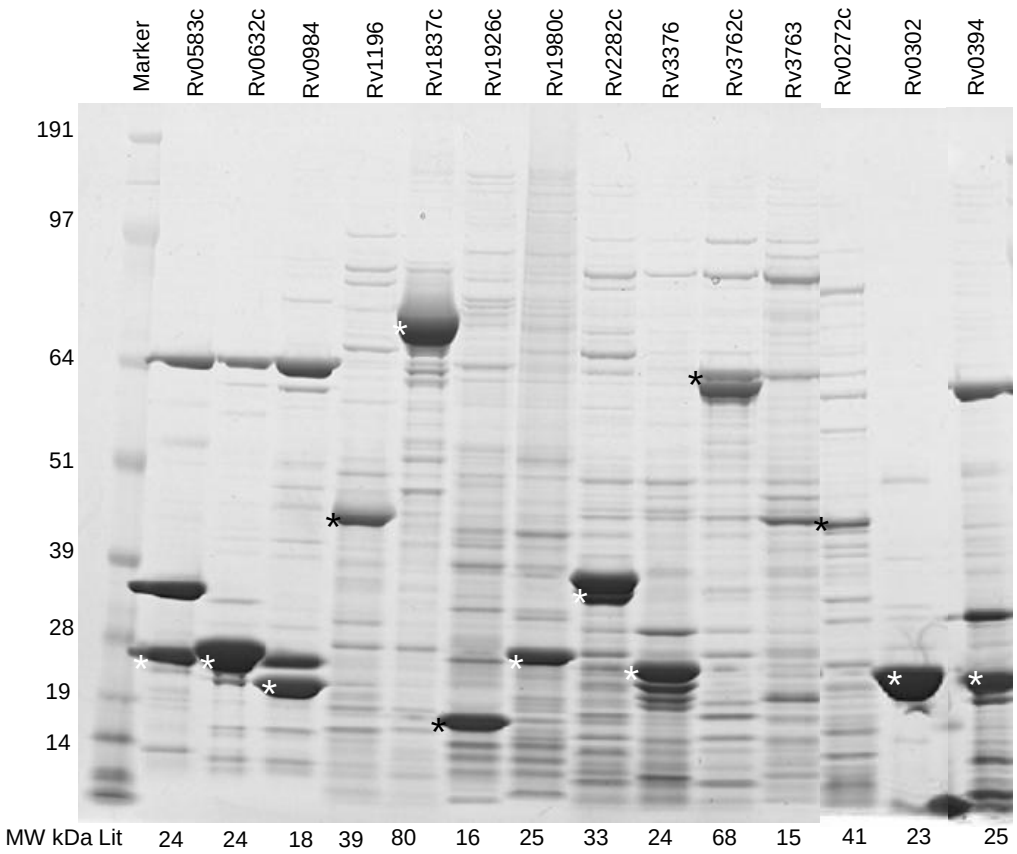
156. de Souza, G.A., et al., *Bacterial proteins with cleaved or uncleaved signal peptides of the general secretory pathway.* J Proteomics, 2011. **75**(2): p. 502-10.

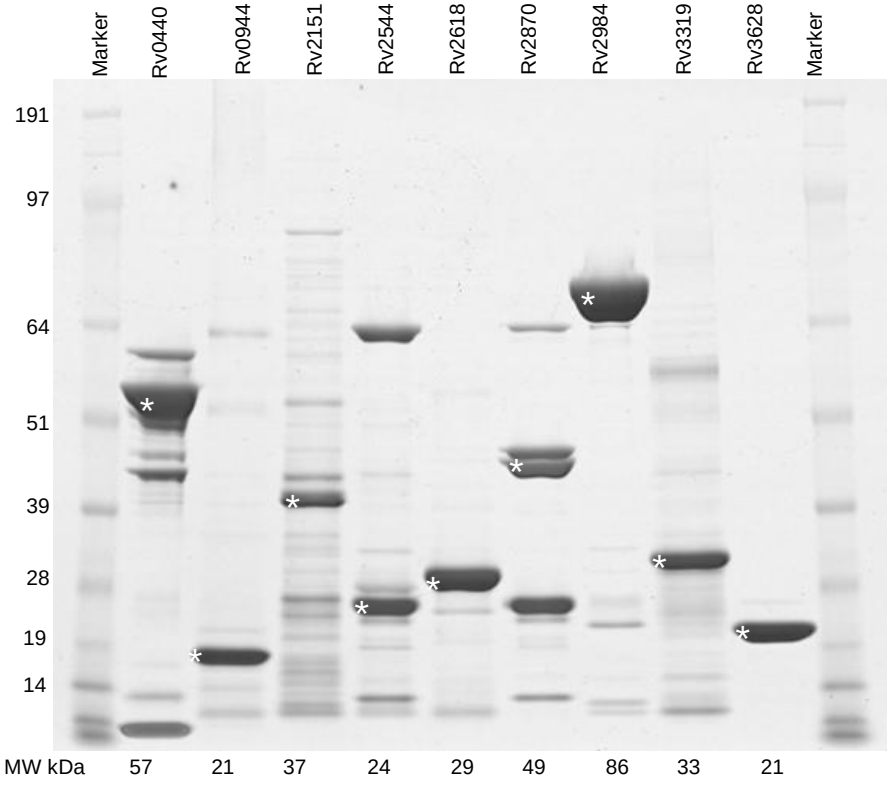
157. Agger, E.M., et al., *Human T-cell responses to the RD1-encoded protein TB27.4 (Rv3878) from Mycobacterium tuberculosis.* Immunology, 2003. **110**(4): p. 507-12.

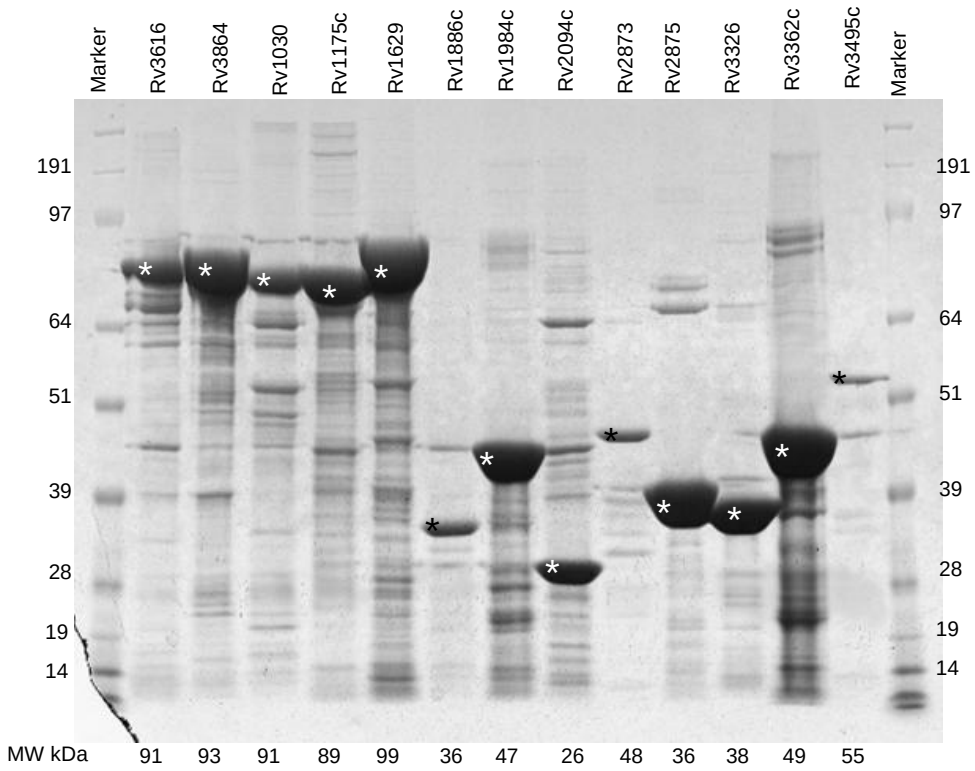
9 Supplementary material

A

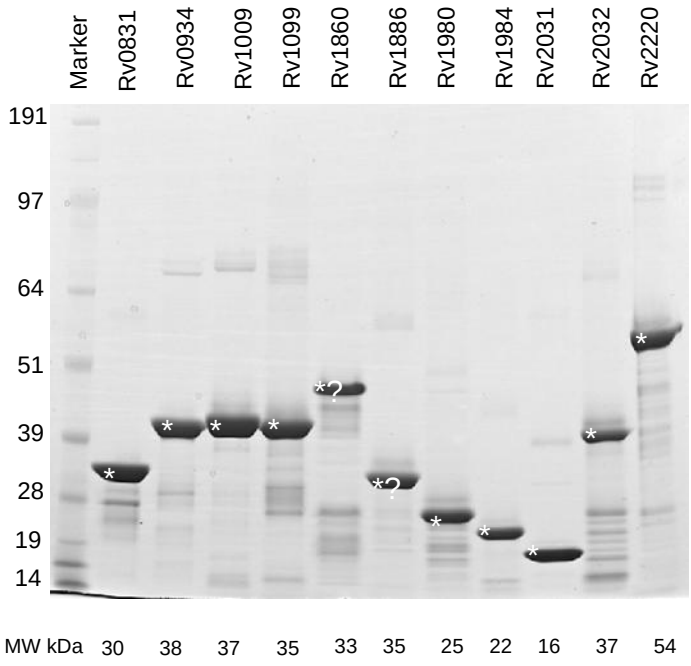


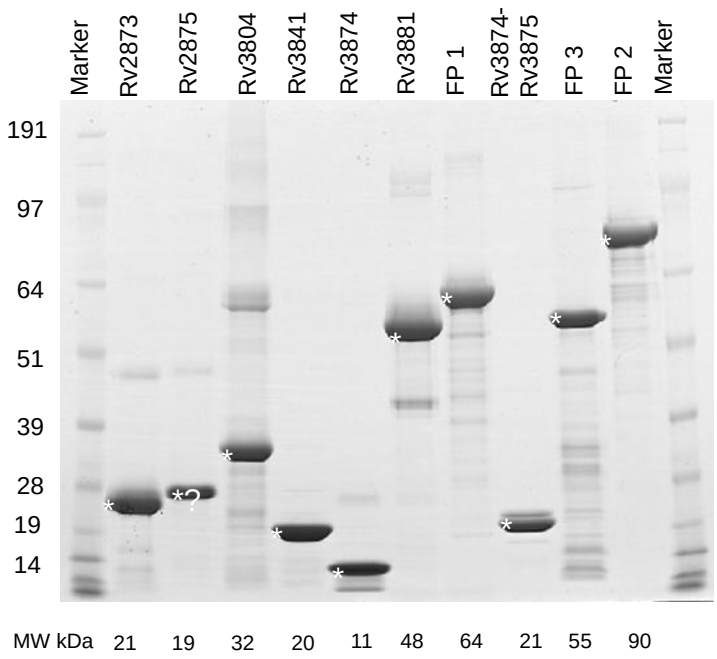

Marker
Rv0583c
Rv0632c
Rv0984
Rv1196
Rv1837c
Rv1926c
Rv1980c
Rv2282c
Rv3376
Rv3762c
Rv3763
Rv0272c
Rv0302
Rv0394
191
97
64
51
39
28
19
14
MW kDa Lit 24 24 18 39 80 16 25 33 24 68 15 41 23 25


Marker
Rv0440
Rv0944
Rv2151
Rv2544
Rv2618
Rv2870
Rv2984
Rv3319
Rv3628
Marker
191
97
64
51
39
28
19
14
MW kDa 57 21 37 24 29 49 86 33 21

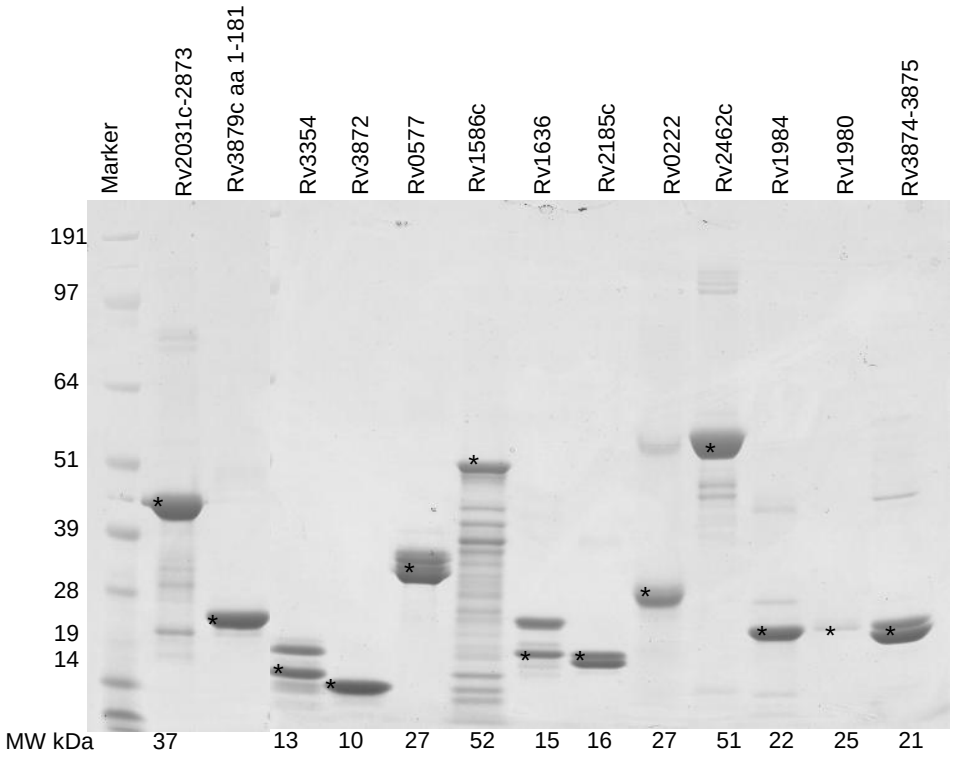

Marker
Rv3616
Rv3864
Rv1030
Rv1175c
Rv1629
Rv1886c
Rv1984c
Rv2094c
Rv2873
Rv2875
Rv3326
Rv3362c
Rv3495c
Marker
191
97
64
51
39
28
19
14
MW kDa
91
93
91
89
99
36
47
26
48
36
38
49
55

B


Marker
Rv0831
Rv0934
Rv1009
Rv1099
Rv1860
Rv1886
Rv1980
Rv1984
Rv2031
Rv2032
Rv2220
191
97
64
51
39
28
19
14
MW kDa
30
38
37
35
33
35
25
22
16
37
54


Marker
Rv2873
Rv2875
Rv3804
Rv3841
Rv3874
Rv3881
FP 1
Rv3874-Rv3875
FP 3
FP 2
Marker
191
97
64
51
39
28
19
14
MW kDa 21 19 32 20 11 48 64 21 55 90

C


Marker
Rv2031c-2873
Rv3879c aa 1-181
Rv3354
Rv3872
Rv0577
Rv1586c
Rv1636
Rv2185c
Rv0222
Rv2462c
Rv1984
Rv1980
Rv3874-3875
191
97
64
51
39
28
19
14
MW kDa 37 13 10 27 52 15 16 27 51 22 25 21

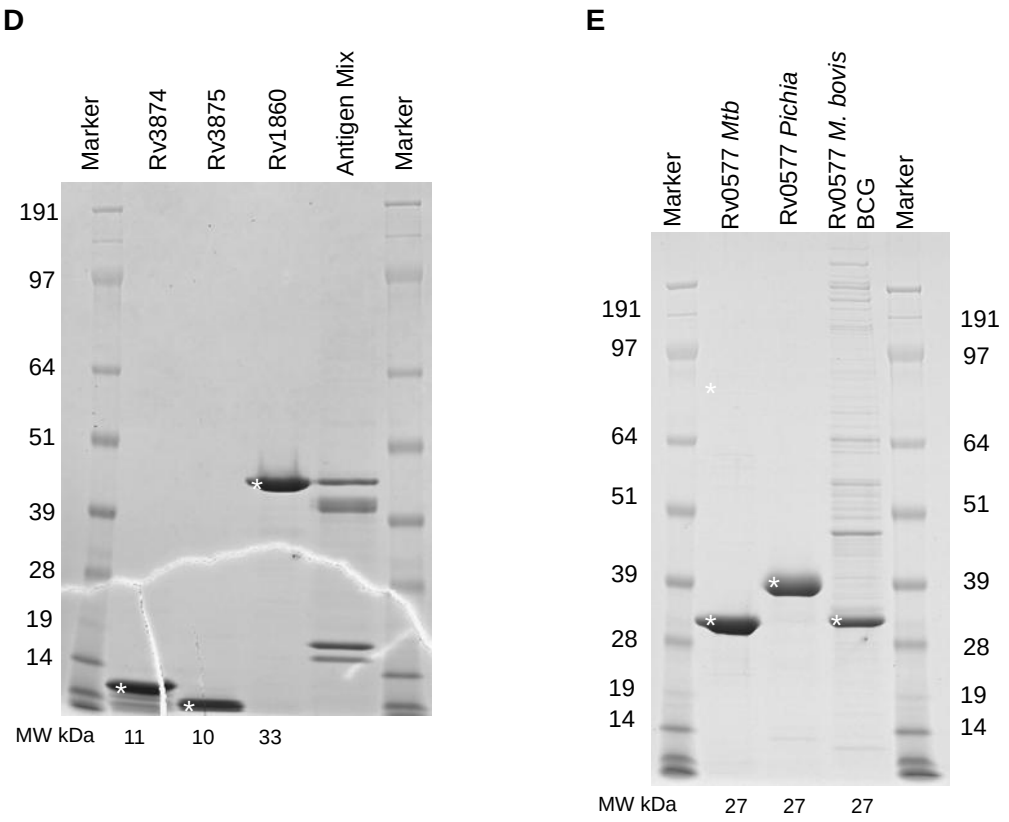


Figure 9-1: SDS-PAGE analysis of purified recombinant antigens.

Individual M. tuberculosis *proteins are listed by their H37Rv gene numbers. Five µg of each antigen were applied on a 4-12% SDS-PAGE gel and stained with Coomassie Brilliant blue to determine the relative purity. Not all antigens showed bands at the correct molecular weight size. The expected MW indicated below each lane was taken from literature. Antigens with a correct size are marked with an asterisk and antigens with an incorrect size are marked with a question mark.* ***A*** *supplier 1,* ***B*** *supplier 2,* ***C*** *supplier 3,* ***D*** *supplier 4: the antigen mix consists of four different antigens,* ***E*** *supplier 5: Rv0577 expressed in* Pichia *reveals a higher MW compared with* M. tuberculosis *and* M. bovis *BCG strain antigens expressed in E. coli because of posttranslational modifications.*

Table 9-1: Correlation coefficients (r) for all antigens.

The same sample set was measured with two individually coupled bead sets.

antigen	r	antigen	r	antigen	r
Rv0212	0.55	**Rv2544**	0.37	**Rv1009**	0.49
Rv0212	0.63	**Rv2618**	0.67	**Rv1099**	0.24
Rv0272	0.79	**Rv2870**	0.56	**Rv1860**	0.97
Rv0272	0.98	**Rv2870**	0.47	**Rv1886**	0.78
Rv0302	0.99	**Rv2927**	0.80	**Rv1980**	0.81
Rv0379	0.99	**Rv2984**	0.66	**Rv1984**	0.35
Rv0394	0.86	**Rv2984**	0.52	**Rv2031**	0.69
Rv0440	0.90	**Rv3050**	0.30	**Rv2032**	0.59
Rv0456	0.92	**Rv3243**	0.57	**Rv2220**	0.16
Rv0583	0.78	**Rv3243**	0.53	**Rv2873**	0.94
Rv0632	1.00	**Rv3248**	0.90	**Rv2875**	0.44
Rv0798	0.77	**Rv3319**	0.46	**Rv3804**	0.13
Rv0801	0.96	**Rv3319**	0.28	**Rv3841**	0.11
Rv0934	0.99	**Rv3326**	0.36	**Rv3874**	0.97
Rv0944	0.93	**Rv3362**	0.41	**Rv3881**	0.95
Rv0984	0.75	**Rv3376**	0.88	**Rv0577 *Mtb***	0.72
Rv1196	0.97	**Rv3376**	0.51	**Rv0577 *Pichia***	0.94
Rv1242	0.24	**Rv3628**	0.67	**Rv0577 *M. bovis* BCG**	0.77
Rv1242	0.35	**Rv3762**	0.67	**Rv0222**	0.86
Rv1284	0.86	**Rv3763**	0.79	**Rv0577**	0.61
Rv1411	0.99	**Rv3763**	0.64	**Rv1586**	0.57
Rv1566	0.97	**Rv3775**	0.36	**Rv1636**	0.11
Rv1837	0.97	**Rv3804**	0.99	**Rv1984**	0.89
Rv1860	0.98	**Rv3874**	0.99	**Rv2185**	0.84
Rv1926	0.99	**Rv3878**	0.99	**Rv1980**	0.33
Rv1980	1.00	**Rv3881**	0.99	**Rv2031-Rv2873**	0.99
Rv1984	0.15	**Rv3874-Rv3875**	0.97	**Rv2462**	0.90
Rv2031	0.98	**FP1**	0.98	**Rv3354**	0.09
Rv2151	0.81	**FP2**	0.92	**Rv3872**	0.86
Rv2252	0.64	**FP3**	0.97	**Rv3874-Rv3875**	0.99
Rv2282	0.16	**Rv0831**	0.63	**Rv3879 aa 1-181**	0.99
Rv2282	0.48	**Rv0934**	0.98		

Table 9-2: Sensitivity and specificity of all antigens.

antigen	Sensitivity [%]	Specificity [%]	antigen	Sensitivity [%]	Specificity [%]
Rv0212_1	60	48	**Rv3050_1**	41	68
Rv0222_3	30	82	**Rv3243_1**	46	66
Rv0272_1	66	46	**Rv3248_1**	58	56
Rv0302_1	65	50	**Rv3319_1**	62	51
Rv0379_1	55	55	**Rv3326_1**	68	45
Rv0394_1	51	62	**Rv3354_3**	69	41
Rv0440_1	75	31	**Rv3362_1**	45	67
Rv0456_1	70	36	**Rv3376_1**	54	57
Rv0577_3	63	47	**Rv3628_1**	63	53
Rv0577 *Mtb*_5	64	45	**Rv3762_1**	16	88
Rv0577 *Pichia*_5	5	97	**Rv3763_1**	67	50
Rv0577 *M. bovis BCG*_5	66	43	**Rv3775_1**	58	56
Rv0583_1	22	87	**Rv3804_1**	79	42
Rv0632_1	41	70	**Rv3872_3**	77	29
Rv0798_1	63	55	**Rv3874_1**	72	43
Rv0801_1	63	42	**Rv3874-Rv3875_3**	78	32
Rv0934_1	36	77	**Rv3878_1**	53	57
Rv0944_1	52	60	**Rv3879 aa 1-181_3**	67	43
Rv0984_1	68	44	**Rv3881_1**	51	64
Rv1196_1	51	67	**Rv0934_2**	36	78
Rv1242_1	57	54	**Rv1860_2**	64	56
Rv1284_1	37	74	**Rv1980_2**	81	38
Rv1411_1	45	65	**Rv1984_2**	62	49
Rv1566_1	53	58	**Rv2031_2**	22	86
Rv1586_3	68	43	**Rv2873_2**	71	45
Rv1636_3	67	45	**Rv3804_2**	60	55
Rv1837_1	55	58	**Rv3874_2**	51	67
Rv1860_1	70	44	**Rv3881_2**	41	73
Rv1926_1	54	58	**Rv3874-Rv3875_2**	61	56
Rv1980_1	56	59	**FP1_2**	44	76
Rv1980_3	27	81	**FP2_2**	52	68
Rv1984_3	31	74	**Rv0831_2**	55	61
Rv1984_1	57	52	**Rv1009_2**	48	63

Rv2031_1	43	71	**Rv1099_2**	69	43
Rv2031-Rv2873_3	61	55	**Rv1886_2**	64	56
Rv2151_1	53	55	**Rv2032_2**	82	28
Rv2185_3	75	32	**Rv2220_2**	55	54
Rv2252_1	59	52	**Rv2875_2**	79	39
Rv2282_1	59	53	**Rv3841_2**	66	53
Rv2462_3	55	52	**FP3_2**	39	81
Rv2544_1	68	39	**Rv3874_4**	66	49
Rv2618_1	36	73	**Rv3875_4**	71	45
Rv2870_1	42	67	**Rv1860_4**	81	33
Rv2927_1	23	84	**Antigen Mix_4**	53	69
Rv2984_1	68	38			

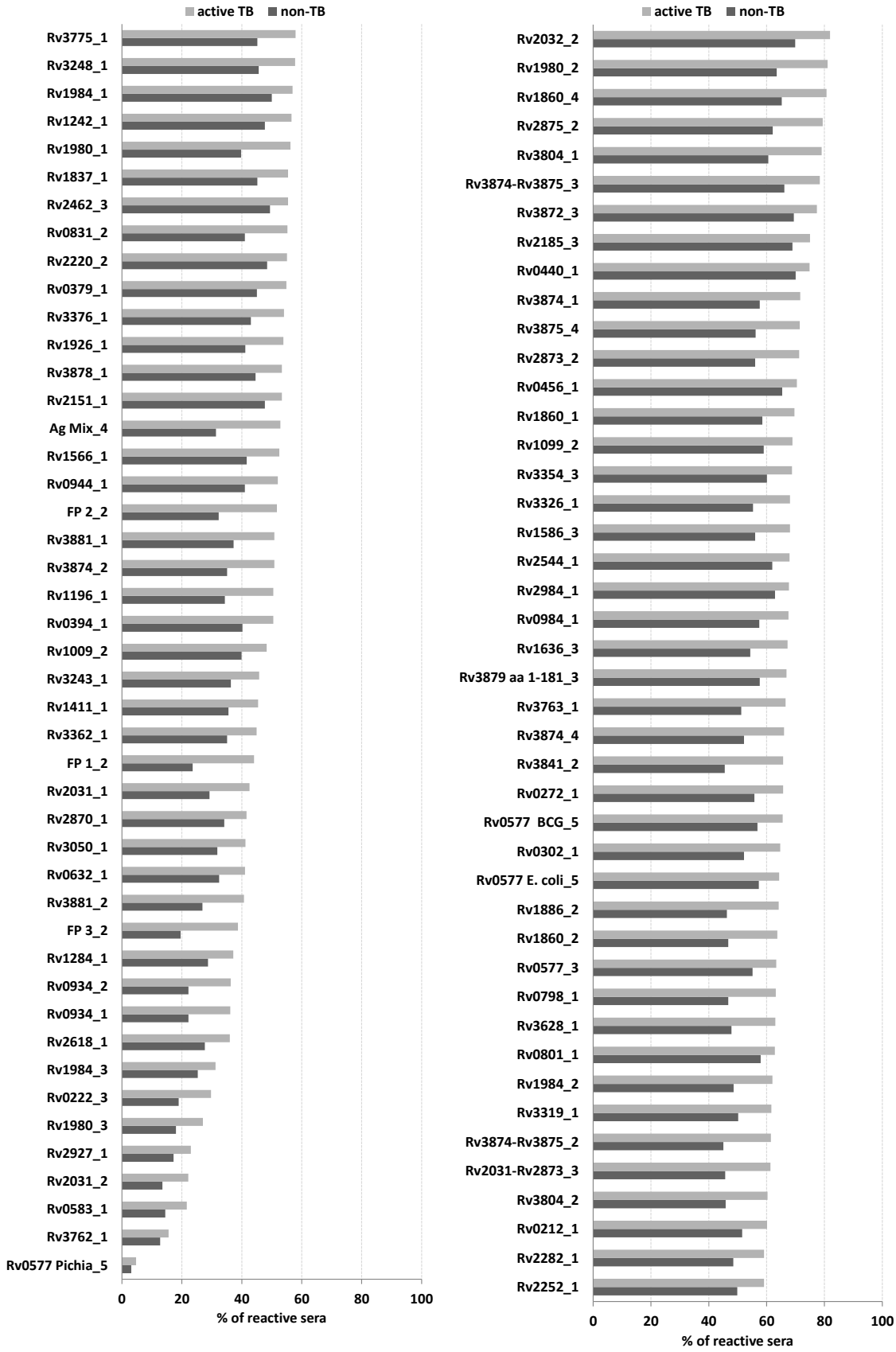


Figure 9-2: Percentage of reactive sera in active and non-TB patients.

Table 9-3: List of 50 antigens that discriminate active from non-TB based on multivariate analysis.

The average merit indicates how important the attribute is, averaged over the folds of the cross-validation. The average rank is the average of its ranking throughout the cross-validation. The numbers following +- are the standard deviations.

antigen_supplier	average merit	average rank
Rv3804_1	0.094 +- 0.006	1 +- 0
FP1_2	0.081 +- 0.006	2.1 +- 0.3
Rv1886_2	0.075 +- 0.004	3.3 +- 0.78
FP3_2	0.069 +- 0.006	4.7 +- 1.19
Antigen Mix_4	0.066 +- 0.009	5.7 +- 2.76
Rv3874_2	0.062 +- 0.006	6.5 +- 1.5
FP2_2	0.063 +- 0.007	6.7 +- 2.83
Rv3881_2	0.054 +- 0.003	8.7 +- 1.55
Rv3874_1	0.053 +- 0.003	9.7 +- 2.05
Rv3881_1	0.051 +- 0.005	11.2 +- 2.93
Rv1860_4	0.05 +- 0.006	11.4 +- 3.72
Rv3874_4	0.049 +- 0.007	12.9 +- 4.59
Rv1860_2	0.047 +- 0.007	14.2 +- 4.77
Rv1980_2	0.047 +- 0.007	14.8 +- 3.89
Rv3874-Rv3875_2	0.047 +- 0.005	14.8 +- 3.46
Rv0934_1	0.046 +- 0.002	15 +- 1.1
Rv0934_2	0.042 +- 0.002	17.4 +- 1.85
Rv3804_2	0.041 +- 0.003	18 +- 1.48
Rv0379_1	0.041 +- 0.004	19 +- 2.41
Rv1411_1	0.038 +- 0.007	20 +- 4.36
Rv3841_2	0.035 +- 0.003	22.4 +- 2.29
Rv2031-Rv2873_3	0.035 +- 0.007	22.9 +- 4.68
Rv3875_4	0.036 +- 0.009	22.9 +- 7.65
Rv2875_2	0.034 +- 0.003	23.5 +- 2.01
Rv1860_1	0.031 +- 0.002	25.6 +- 2.42
Rv1980_1	0.031 +- 0.005	25.8 +- 4.98
Rv2031_1	0.031 +- 0.007	26.5 +- 5.82
Rv1196_1	0.029 +- 0.003	27.1 +- 1.5
Rv3874-Rv3875_3	0.029 +- 0.004	27.4 +- 2.69
Rv0798_1	0.027 +- 0.003	28.5 +- 1.2
Rv2873_2	0.025 +- 0.002	31 +- 1.67
Rv3628_1	0.025 +- 0.002	31.9 +- 1.58

Rv2031_2	0.023 +- 0.003	34.3 +- 3.41
Rv0831_2	0.023 +- 0.002	34.3 +- 1.62
Rv3763_1	0.022 +- 0.003	34.8 +- 2.64
Rv0302_1	0.022 +- 0.001	35.2 +- 2.32
Rv0583_1	0.02 +- 0.003	37.9 +- 2.91
Rv1980_3	0.019 +- 0.002	38.8 +- 1.72
Rv1926_1	0.019 +- 0.003	38.9 +- 3.62
Rv2927_1	0.018 +- 0.002	41 +- 3.29
Rv1284_1	0.018 +- 0.004	41.8 +- 6.4
Rv3326_1	0.017 +- 0.003	42.5 +- 3.91
Rv0212_1	0.017 +- 0.002	44.1 +- 3.94
Rv0577 *Pichia*_5	0.017 +- 0.001	44.4 +- 3.56
Rv3248_1	0.017 +- 0.007	45.2 +-15
Rv0632_1	0.014 +- 0.005	48.2 +- 5.04
Rv1636_3	0.01 +- 0.008	50.3 +- 3.98
Rv0394_1	0.012 +- 0.006	51.5 +- 5.84
Rv2032_2	0.013 +- 0.005	52 +- 7.18
Rv1586_3	0.011 +- 0.007	52.8 +-10.97

Table 9-4: List of 35 antigens that discriminate latent from non-TB based on multivariate analysis.

antigen_supplier	average merit	average rank
Rv0583_1	0.033 +- 0.006	2.2 +- 0.98
Rv0272_1	0.032 +- 0.003	2.6 +- 1.36
Rv0632_1	0.031 +- 0.008	5.1 +- 2.81
Rv3762_1	0.026 +- 0.003	6.9 +- 4.23
Rv2185_3	0.023 +- 0.009	9.1 +- 6.04
Rv3872_3	0.025 +- 0.003	9.2 +- 7.21
Rv3319_1	0.023 +- 0.002	11.2 +- 3.87
Rv3874-Rv3875_3	0.023 +- 0.003	12.9 +- 9.13
Rv2870_1	0.024 +- 0.008	13.4 +-22.27
Rv2873_2	0.026 +- 0.013	13.5 +-15.03
Rv0302_1	0.022 +- 0.013	13.8 +-13.7
Rv2544_1	0.022 +- 0.008	16.9 +-20.55
Rv1284_1	0.017 +- 0.011	20 +-14.53
Rv2031-Rv2873_3	0.007 +- 0.01	23.4 +- 5.46
Rv2282_1	0.002 +- 0.007	24.3 +- 3.55
Rv1984_1	0.004 +- 0.009	25.3 +- 4.82
Rv3878_1	0.017 +- 0.006	25.3 +-22.42
Rv0379_1	0.009 +- 0.012	28.1 +-12.73
Rv2462_3	0.015 +- 0.01	29.1 +-25.15
Rv2151_1	0.002 +- 0.005	29.7 +-16.94
Rv0577 *Mtb*_5	0.004 +- 0.008	32.1 +- 8.37
Rv0984_1	0.012 +- 0.012	32.1 +-18.54
Rv0577 *M. bovis* BCG_5	0.004 +- 0.009	34.6 +- 9
Rv0440_1	0.002 +- 0.007	36.5 +- 9.31
Rv0212_1	0.002 +- 0.006	36.7 +- 6.12
Rv0394_1	0.004 +- 0.008	36.7 +- 5.81
Rv3874-Rv3875_2	0.01 +- 0.01	37.6 +-21.85
Rv3775_1	0.014 +- 0.01	38 +-32.52
Rv2031_2	0.013 +- 0.011	39.7 +-31.17
Rv1566c_1	0.004 +- 0.008	40.5 +- 8.85
Rv2032_2	0.008 +- 0.01	41.2 +-22.3
Rv3804_2	0.009 +- 0.011	42.6 +-24.74
Rv0801_1	0.002 +- 0.006	43.2 +- 9.67
FP2_2	0.006 +- 0.01	46 +-17

FP1_2	0.002 +- 0.006	55.9 +- 9.25

10 List of publications

Schmohl, M., Rimmele, S., Pötz, O., Kloog, Y., Gierschik, P., Joos, T.O. and Schneiderhan-Marra, N., Protein-protein-interactions in a multiplexed, miniaturized format, a functional analysis of Rho GTPase activation and inhibition. Proteomics, 2010; 10(8):1716-20.

Rimmele, S., Gierschik, P., Joos, T.O. and Schneiderhan-Marra, N., Bead-based protein–protein interactionassays for the analysis of Rho GTPase signaling. J. Mol. Recognit., 2010; 23(6): 543–550.

Schmohl, M., Rimmele, S., Gierschik, P., Joos, T.O. and Schneiderhan-Marra, N., Functional Analysis of Rho GTPase Activation and Inhibition in a Bead-Based Miniaturized Format. Meth. Mol. Biol., 2012; 827:271-82.

Rimmele, S., Planatscher, H., Michel, G., Schneiderhan-Marra, N. and Joos, T.O., A multiplexed bead-based assay for the detection of human IgGs directed against *Mycobacterium tuberculosis* proteins. Clin. Chem., *in preparation*

Planatscher, H., Rimmele, S., Michel, G., Schneiderhan-Marra, N. and Joos, T.O.; Generation of controls for multiplexed serological assays. Nat. Meth., *in preparation*

11 Curriculum vitae

Personal details:

Name:	Stefanie Rimmele
Date of birth:	October 17, 1983
Place of birth:	Reutlingen

Education:

01/2009 - 05/2012	Dissertation at the NMI, biochemistry department, Reutlingen Title: *Mycobacterium tuberculosis* Antigen Arrays for the in depth Characterization of the Serological Response of Infected Patient Cohorts. Supervisors: Prof. Dr. Stefan Stevanović, Prof. Dr. Ulrich Rothbauer, Dr. Thomas Joos and Dr. Nicole Schneiderhan-Marra
02/2008 - 10/2008	Diploma thesis at the NMI, biochemistry department, Reutlingen Title: Funktionelle Analyse von Rho GTPase Aktivierung und Inhibierung in einem multiplexen und miniaturisierten Ansatz. Supervisors: Dr. Thomas Joos and Dr. Nicole Schneiderhan-Marra
10/2003 - 10/2008	Studies of Biochemistry, Eberhard Karls University, Tübingen
09/1994 - 06/2003	Isolde-Kurz-Gymnasium, Reutlingen (Allgemeine Hochschulreife)