Identification of differentially recognized T cell epitopes in the spectrum of *Mtb* infection

Sudhasini Panda[1], Jeffrey Morgan[1], Catherine Cheng[1], Mayuko Saito[2], Robert H. Gilman[3,4], Nelly Ciobanu[5], Valeriu Crudu[5], Donald G Catanzaro[6], Antonino Catanzaro[7], Timothy Rodwell[7], Judy S.B. Perera[8], Teshan Chathuranga[8], Bandu Gunasena[9], Aruna D. DeSilva[1,8], Bjoern Peters[1,7], Alessandro Sette[1,7], Cecilia S. Lindestam Arlehamn[1,*]

1 Center for Infectious Disease and Vaccine Research, La Jolla Institute for Immunology, La Jolla, CA, USA
2 Department of Virology, Tohoku University Graduate School of Medicine, Sendai, Japan
3 Johns Hopkins School of Public Health, Baltimore, MD, USA
4 Universidad Peruana Cayetano Heredia, Lima, Peru
5 Phthisiopneumology Institute, Chisinau, Republic of Moldova
6 Department of Biological Sciences, University of Arkansas, Fayetteville, AR, USA
7 Department of Medicine, University of California San Diego, La Jolla, CA, USA
8 Faculty of Medicine, General Sir John Kotelawala Defense University, Ratmalana, Sri Lanka.
9 National Hospital for Respiratory Diseases, Welisara, Sri Lanka

*Corresponding author
Cecilia S. Lindestam Arlehamn
cecilia@lji.org

Keywords: Mycobacterium tuberculosis, T cell epitopes, CD4, Active TB,
Short title: Response against *Mtb*-specific epitopes in active TB infection

Abstract

Tuberculosis caused by *Mycobacterium tuberculosis* is one of the leading causes of death from a single infectious agent. Identifying dominant epitopes and comparing their reactivity in different tuberculosis (TB) infection states can help design diagnostics and vaccines. We performed a proteome-wide screen of 20,610 *Mtb* derived peptides in 21 Active TB (ATB) patients 3-4 months post-diagnosis of pulmonary TB (mid-treatment) using an IFNγ and IL-17 Fluorospot assay. Responses were mediated exclusively by IFNγ and identified a total of 137 unique epitopes, with each patient recognizing, on average, 8 individual epitopes and 22 epitopes (16%) recognized by 2 or more participants. Responses were predominantly directed against antigens part of the cell wall and cell processes category. Testing 517 peptides spanning TB vaccine candidates and ESAT-6 and CFP10 antigens also revealed differential recognition between ATB participants mid-treatment and healthy IGRA+ participants of several vaccine antigens. An ATB-specific peptide pool consisting of epitopes exclusively recognized by participants mid-treatment, allowed distinguishing participants with active pulmonary TB from healthy interferon-gamma release assay (IGRA)+/- participants from diverse geographical locations. Analysis of longitudinal samples indicated decreased reactivity during treatment for pulmonary TB. Together, these results show that a proteome-wide screen of T cell reactivity identifies epitopes and antigens that are differentially recognized depending on the *Mtb* infection stage. These have potential use in developing diagnostics and vaccine candidates and measuring correlates of protection.

Introduction

Tuberculosis is the ninth leading cause of death worldwide and the leading cause from a single infectious agent, ranking above HIV/AIDS. The World Health Organization (WHO) estimates that approximately one-quarter of the world's population (1.7 billion total) is infected with *Mycobacterium tuberculosis* (*Mtb*). In 2021 *Mtb* was responsible for 1.6 million deaths and ~10 million new infections (1)

Infection with *Mtb* manifests as a spectrum of diseases ranging from asymptomatic subclinical infection (latent infection, LTBI) to active disease (ATB) (2,3). To date, tuberculin skin test (TST) and interferon-gamma release assays (IGRA), which detect an immunological response against *Mtb*, are diagnostic tests for latent infection (4,5). Importantly, the classically used tuberculin skin test can be positive in both LTBI and BCG vaccinated participants and thus cannot distinguish them (6,7). A major advance in the diagnosis of *Mtb* infection was represented by the introduction of IGRA tests that consist of ex vivo analysis of peripheral blood cells for a cytokine (IFNγ) response to peptide pools spanning the ESAT-6 and CFP10 antigens recognized by T cells (8). As these antigens are absent from *M. bovis* BCG and most nontuberculous mycobacteria (NTMs), such responses can distinguish prior or current *Mtb* infection from BCG vaccination and NTM diseases. However, these tests cannot reliably discriminate between active TB and LTBI (9). These observations indicate the need for a more extensive search for a panel of antigens that can distinguish LTBI and active disease.

T cells are critical in the host immune response against *Mtb* infection. The responses against *Mtb* infection involve classically restricted CD4 and CD8 αβ T cells (10,11) and non-classically restricted T cells such as NKT (CD1), MAIT (MR1), and γδ T cells (12–14). Among these, CD4 T cells represent a major component of T cell response against *Mtb* infection (15). Individuals with low levels of CD4 T cells, such as those who are HIV positive, are more vulnerable to both primary and reactivation of TB (16). Multiple studies have revealed the recognition of T-cell epitopes from *Mtb* among diverse populations, including individuals with TB disease, LTBI, BCG vaccination, or exposure to *Mtb* and/or NTM (17–20), also reviewed in (21). T cell responses at the antigen and epitope levels are highly complex, particularly in complex organisms. They can involve multiple antigens and hundreds of epitopes (22–25), with broader patterns of immunodominance observed

in humans compared to genetically homogeneous murine model (26–28). While the mechanisms underlying T cell immunodominance and breadth have been extensively studied in murine models and, to some extent, in humans (22,29–31)(32), a quantitative assessment of the complexity of responses during natural infections at the population level is currently lacking. Most immuno-profiling studies have focused on individual antigens or a limited set of epitopes, assuming they represent the entire pathogen-specific response (33,34). It remains uncertain to what extent underestimating the true complexity of these responses could affect the outcomes of immuno-profiling studies. Therefore, a comprehensive understanding of *Mtb* epitopes across TB disease states is crucial. *Mtb* expresses different proteins during various stages of infection, resulting in stage-specific immune responses. Thus, identifying novel CD4 T cell epitopes at different stages of TB could aid in discovering new diagnostic and vaccine candidates against TB. Previously, our team conducted a proteome-wide screen to detect HLA class II restricted epitopes and antigens recognized in healthy IGRA+ participants without any signs of active TB (22). This screen included 4000 ORFs of the *Mtb* genome and approximately 20,000 predicted HLA class II epitopes. T cell reactivity against these epitopes was measured using ELISPOT to detect IFNγ response ex vivo. It is equally important to screen these epitopes in ATB disease. Identifying the ATB-specific antigens will be important as boosting immune responses against antigens expressed during the active phase of infection might help translate into reduced incidence or reactivation of infection.

This study aimed to explore the possibility that distinct *Mtb* disease states (i.e., healthy IGRA+ participants and patients with active pulmonary TB) respond to differential *Mtb*-derived epitopes and/or elicit different *Mtb*-specific T cell responses. We hypothesized that this differential reactivity could be attributed to variations in response magnitude, specificity (i.e., which antigens are recognized), breadth (i.e., how diverse the response is), and the phenotype of the specific T cells. By conducting a comprehensive proteome-wide screening for *Mtb*-derived T cell epitopes, we identified and characterized epitopes specifically recognized by participants with active pulmonary TB, suggesting their potential utility as diagnostic markers and vaccination candidates.

RESULTS

Breadth of responses to a comprehensive library of predicted HLA Class II epitopes in participants with ATB

Detailed knowledge regarding the antigens recognized in TB disease and associated phenotypes is relevant for understanding TB immunopathology and for vaccine and diagnostic applications. alike. We previously used a proteome-wide peptide library of 20,610 *Mtb*-derived 15-mer peptides predicted promiscuous HLA class II binders to define HLA class II restricted epitopes and associated antigens in healthy IGRA+ participants (i.e., LTBI) from San Diego, US (22). Here, we used the same comprehensive library to determine the repertoire of T cell antigens and epitopes recognized in ATB.

The proteome-wide peptide library was arranged into 1036 peptide pools of 20 peptides each. T cell reactivity against the 1036 peptide pools was measured by ex vivo production of IFNγ and IL-17 using Fluorospot assays and PBMCs from 21 participants mid-treatment for their ATB infection (3-4 months post diagnosis) recruited from the Universidad Peruana Cayetano Heredia site (Peru). A total of 78 unique peptide pools were selected for deconvolution, corresponding to the ten peptide pools with the highest response magnitude per participant, and/or that were recognized by at least two different participants. A total of 137 individual epitopes were identified (Table S1). Each participant recognized an average of 8 unique epitopes (range 1-27, median 7), underlining the breadth of responses to *Mtb* (Figure 1A). Among the 137 individual epitopes, 22 (16%) were recognized by multiple participants (Figure 1B). When epitopes were ranked based on magnitude, the top 55 epitopes accounted for 80% of the total response (Figure 1C). In conclusion, the breadth of responses detected in ATB is broad, although narrower ($p=0.002$, two-tailed unpaired Student's t-test) compared to the previous results in healthy IGRA+ participants who recognized, on average, 24 epitopes (22).

Figure 1:

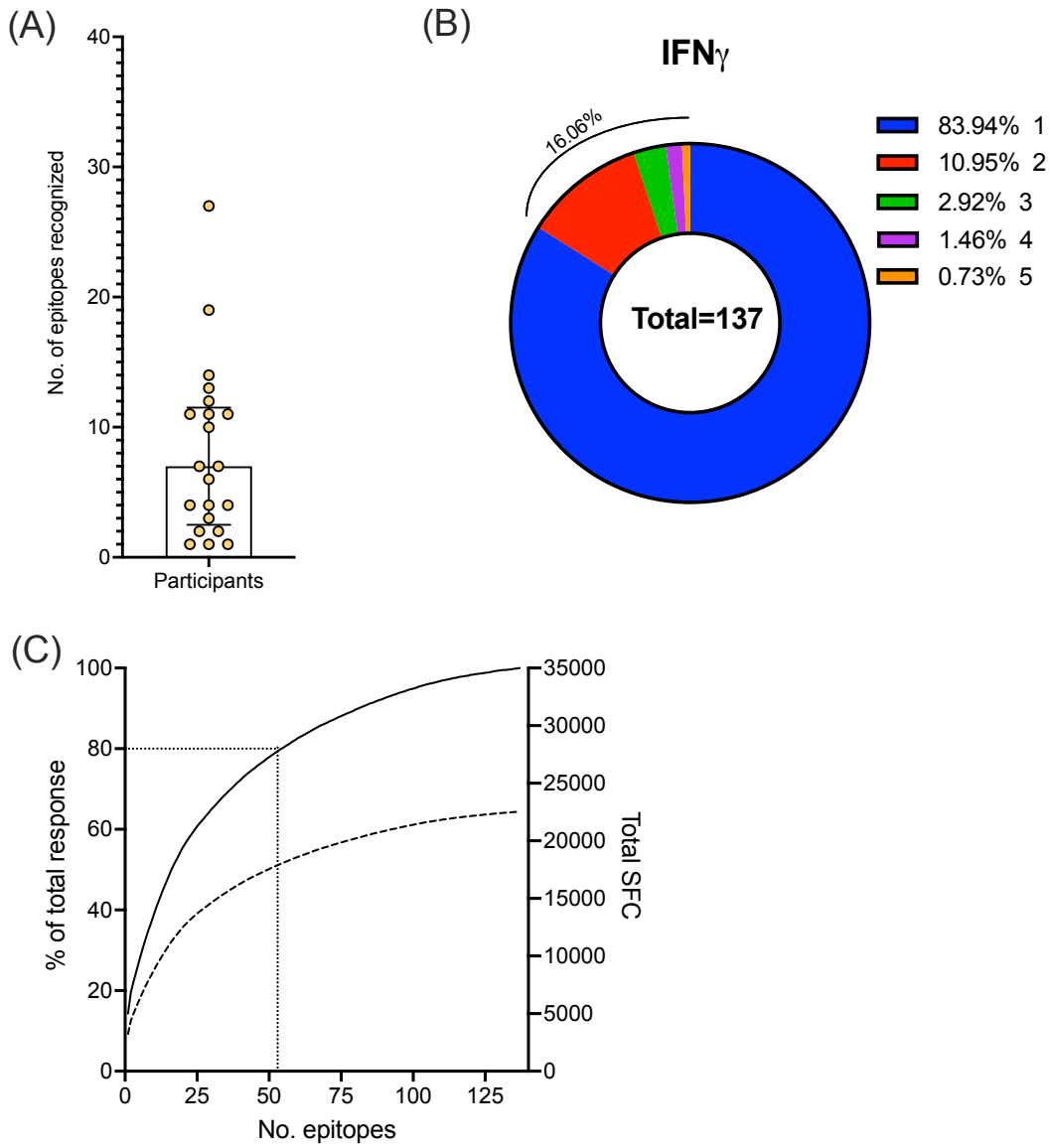


Figure 1: Breadth and dominance of epitopes in mid-treatment ATB participants. A) Number of epitopes recognized by each participant. Each dot is one participant, n=21; median ± interquartile range is shown. B) Distribution of recognized epitopes by the number of participants recognizing each epitope. (C) Epitopes ranked based on the magnitude of response (solid line - % of total response, dotted line – total SFC). Black dotted lines indicate the top 55 epitopes.

Immunodominant antigens in cellular responses in ATB participants

The epitopes were mapped to individual *Mtb* ORFs using H37Rv as a reference genome. A total of 97 ORFs were recognized, with each participant recognizing, on average, 6 ORFs (median 5, Figure 2A). As expected, the well-known antigens Rv0288 (TB10.4), Rv3875 (ESAT-6), Rv3874 (CFP10), and Rv3615c (included in the "ESAT-6 free" IGRA test (35) were the most frequently recognized. However, nine novel antigens, which have not been previously described as antigens for *Mtb*, were also identified (Table S1).

Using the Mycobrowser tool (36), we next determined the protein categories to which the identified antigenic ORFs belonged. As previously observed for IGRA+ participants (37), essentially every protein category was represented (Figure 2B). However, the ORFs antigenic in ATB were predominantly found in the cell wall and cell processes categories, followed by conserved hypotheticals, PE/PPE, and intermediary metabolism and respiration categories (Figure 2B). Compared to the H37Rv genome, the antigenic ORFs were overrepresented in the cell wall, cell processes category, and the PE/PPE categories and underrepresented in conserved hypotheticals (Figure 2C). The overrepresentation of the cell wall and cell processes category appears to be specific for the ATB cohort, because these antigen categories were not overrepresented in the previous screen of IGRA+ participants (22).

Reactivity in healthy IGRA+ participants (22) previously identified three "antigenic islands" of clustered antigen genes that comprised secreted and non-secreted *Mtb* proteins involved in type 7 secretion systems. Indeed, all three antigenic islands were also identified in the present screen of ATB participants. (Figure 2D). In contrast to IGRA+ participants in the previous screen, the breadth of responses to antigens targeting the antigenic islands was narrower in the participants with ATB (Figure 2E). For example, for island 1, only 3 antigens represented in the predicted peptide library of 20,610 peptides were recognized by participants with ATB, compared to 9 recognized by healthy IGRA+ participants. The same is true for antigens in islands 2 and 3. In conclusion, the data suggest that different patterns of antigenic ORFs might be associated with ATB vs. healthy IGRA+ participants.

Figure 2:

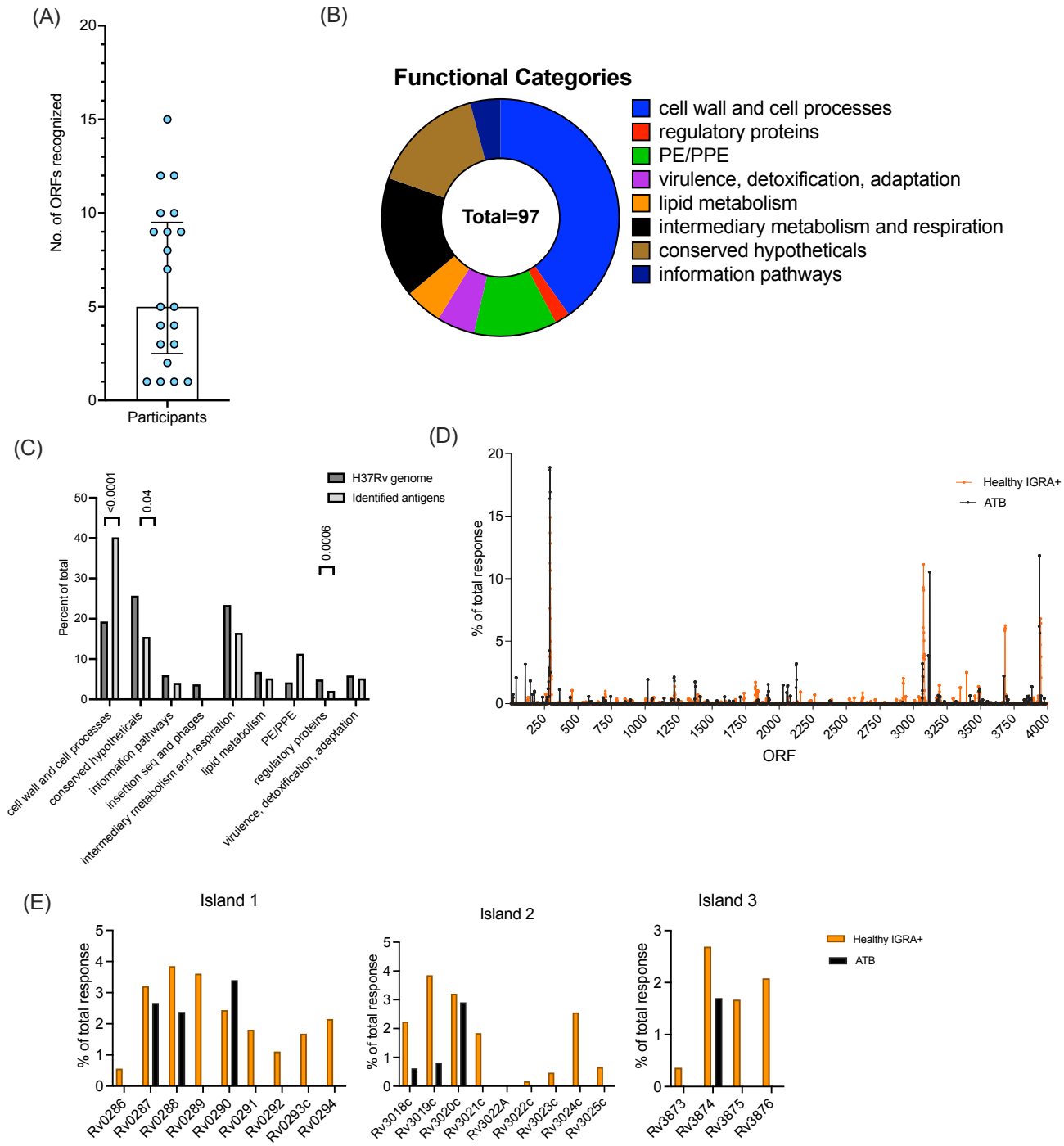


Figure 2. Immunodominant antigens in ATB. (A) Number of ORFs corresponding to recognized epitopes by each participant. Each dot represents one participant, n=21, median ± interquartile range is shown. (B) Distribution of recognized ORFs per protein category. (C). The identified antigens (black bars) were divided into protein categories (Mycobrowser) and compared to the H37Rv genome (grey bars). Chi-square test. (D) Antigenic islands identified by a 5-gene

window spanning the H37Rv genome. ATB (black) compared to healthy IGRA+ (orange, (22)). (E) Proteins within each antigenic island, % of total response per antigen across the proteome-wide screen. ATB (black bars) and healthy IGRA+ (orange bars).

Hierarchy in reactivity against TB Vaccine and IGRA antigens

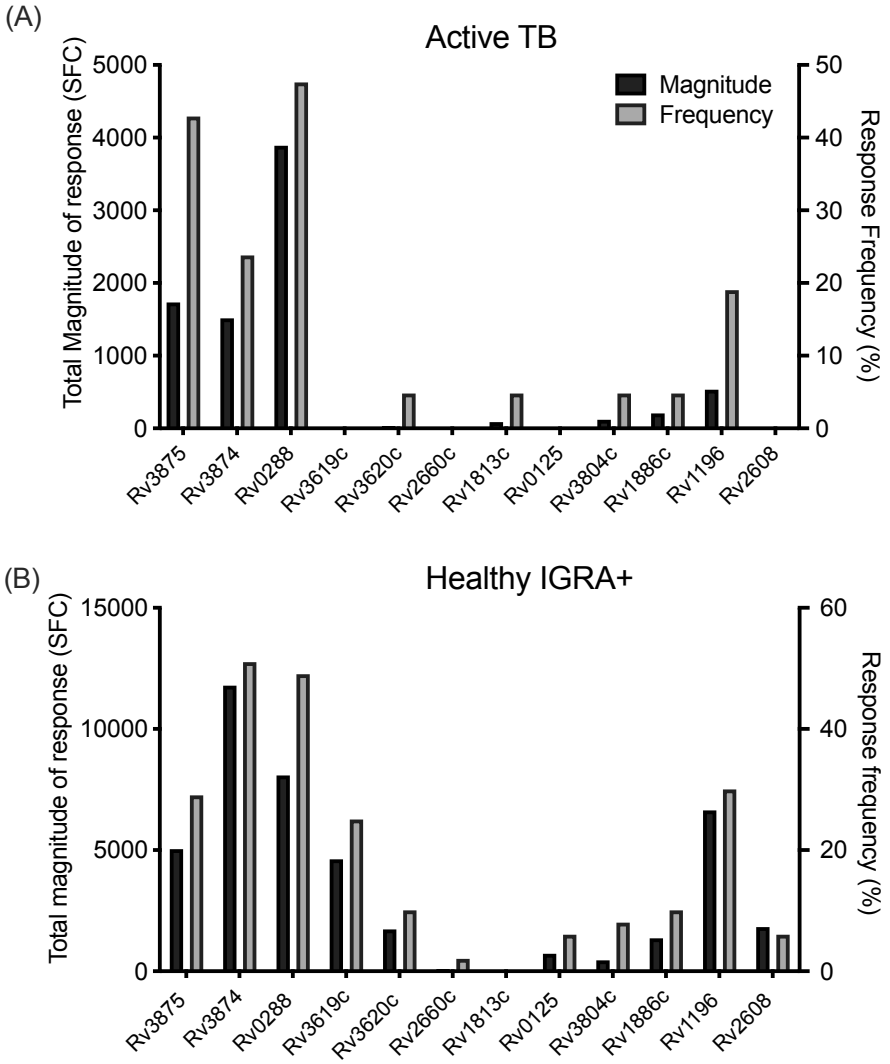
Previous studies showed that the proteome-wide library of predicted promiscuous HLA class II binders captures about 50% of the total reactivity (22,38). To further evaluate T cell responses against TB vaccine candidate and IGRA antigens (see methods), we tested 517 15-mer peptides overlapping by 10 amino acids spanning these antigens. Positive pools were deconvoluted to identify individual T cell epitopes. Overall, 67% of the ATB participants recognized epitopes from at least one antigen; on average, these participants recognized 2 different antigens (range 1-4). This is similar to our previous reports highlighting the inter-individual variability of epitope-specific responses (37).

We next compared the magnitude and frequency of the response for these antigens in the ATB participants with what was previously observed in healthy IGRA+ participants from South Africa (37). The most frequently recognized antigens in ATB and IGRA+ alike were Rv0288 (TB10.4), Rv3875 (ESAT-6), Rv3874 (CFP10), and Rv1196 (PPE18) (figure 3A, B). However, some antigens were differentially recognized in the two cohorts. Specifically, Rv3875 was more frequently recognized than Rv3874 in ATB vs. IGRA+ participants. The Rv1813c antigen was reactive in participants with ATB and not in IGRA+. Finally, the Rv3619c (EsxV), Rv2660c, Rv0125 (Mtb32a), and Rv2608 (PPE42) antigens were reactive in IGRA+ and completely unreactive in participants with ATB (Figure 3A, B).

The proteome-wide screen detected the two most reactive vaccine antigens, Rv0288 and Rv3874. Rv3875 was not detected in the proteome-wide screen, likely due to its small size, with only 2 peptides representing it in the proteome-wide library. The results confirm that the antigens detected in the proteome-wide screen are the most frequently recognized together with Rv3875 and that the other vaccine antigens account for a small fraction of the response in this ATB cohort. In conclusion, the screen of vaccine and IGRA antigens, together with the proteome-wide screen, identified a total of 174 epitopes, which were next investigated for functionality and differential reactivity in further experiments.

.

204 **Figure 3:**



205

206 **Figure 3. Hierarchy in T cell reactivity against TB vaccine and IGRA antigens.** Magnitude of
207 response, expressed as the total magnitude of response (black bars, left y-axis) or frequency of
208 participants responding (grey bars, right y-axis), amongst the participants. (A) ATB, n=21. (B)
209 Healthy IGRA+, n=63 (37), for comparison purposes. Rv number for each antigen are indicated
210 on the x-axis.

211

212 **Functionality of T cell responses specific for different protein antigen categories in ATB**

213 The responses associated with the 174 identified epitopes were characterized in more detail by
214 intracellular cytokine secretion assays to characterize T cell responses specific to antigens from
215 the different protein categories. One epitope pool (PC85) corresponded to epitopes from antigens

in the cell wall and cell processes category, and a separate pool encompassed epitopes from other protein categories (PC71) (Table S1). As a comparator, we included the previously described MTB300 pool, based on epitopes recognized in IGRA+ participants (22), which includes peptides from all functional categories.

As expected, based on the peptide library design to bind HLA class II alleles, the majority of the response was meditated by CD4 T cells (Figure 4A). Among them, the frequency of TNFα+ CD4 cells was relatively higher than IFNγ ($p = 0.0006$ for PC71) or IL-2 ($p = 0.04$ for PC85). In addition, some responses were detected in CD8+ T cells (figure 4B), likely due to nested HLA class I binding epitopes within the 15-mer peptides.

Frequencies of cytokine-expressing CD4 T cells in response to both pools were similar to MTB300 (IFNγ, $p=0.98$ for PC85 and PC71; IL-2, $p=0.31$ and 0.97 for PC85 and PC71 respectively; TNFα, $p=0.78$ and 0.51 for PC85 and PC71 respectively) (Figure 4A, gating strategy shown in Figure S1). The vast majority of cytokine producing CD4 T cells expressed a single cytokine TNFα or IFNγ, followed by TNFα+IL2+ and TNFα+IFNγ+ dual cytokine expressed cells with none to very few IFNγ+IL2+ cells and triple cytokine producing cells (Figure 4C and 4D). A higher frequency of single TNFα-producing T cells was present compared to other single (IFNγ $p=0.01$ (PC71), IL-2 $p=0.01$ (PC85)), dual (IFNγ+IL-2+ $p=0.0006$ (PC71), $p=0.0002$ (PC85), TNFα+IFNγ+ $p=0.002$ (PC71), $p=0.01$ (PC85), TNFα+IL2+ $p=0.02$ (PC71), $p=0.008$ (PC85)), or triple ($p=0.0008$ (PC71), $p=0.0003$ (PC85) cytokine producing cells (Figure 4C).

Next, we characterized which T cell memory subset was responsible for the reactivity. Memory subset phenotypes were determined using antibodies to CD45RA and CCR7. The majority of the cytokine producing CD4 T cells were effector memory (CD45RA-CCR7-), and as expected few were naïve (CD45RA+CCR7+) cells (Figure 4E-F). For IFNγ there was no difference between the CD4 memory subsets comparing the different peptide pools (Figure 4E). For IL-2 there was a significantly higher frequency of effector memory T cells ($p=0.0007$ and 0.03) and a lower frequency of T_{EMRA} (CD45RA+CCR7+) T cells ($p=0.001$ and 0.02) in response to PC85 compared to the PC71 and MTB300 respectively (Figure 4F). The PC85 stimulation also resulted in a significantly higher frequency of effector memory TNFα-producing cells ($p=0.0003$ and 0.008) and a lower frequency of central memory cells ($p=0.04$ and 0.02) than PC71 and MTB300

respectively (Figure 4G). This finding suggests distinct differentiation of the CD4 memory subsets in response to epitope pools representing different functional categories.

Figure 4:

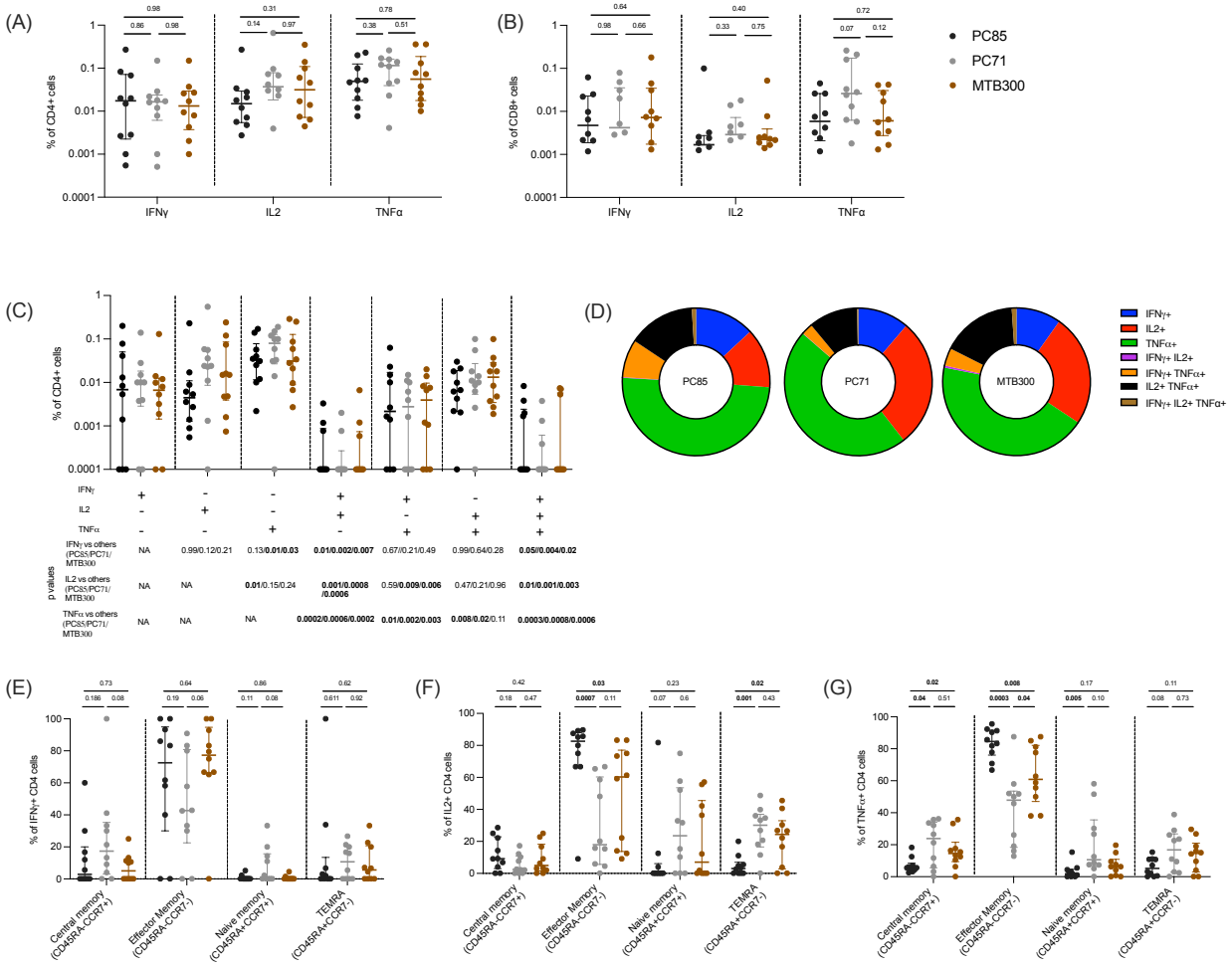


Figure 4. T cell responses specific for different protein categories. (A) Frequency of cytokine-producing, IFNγ, TNFα, and IL-2, CD4 T cells in response to PC85, PC71, and MTB300. (B) Frequency of cytokine-producing, IFNγ, TNFα, and IL-2, CD8+ T cells in response to PC85, PC71, and MTB300. (C) Percentage pool-specific IFNγ, TNFα, and IL-2 production by CD4 T cells expressing each of the seven possible combinations. (D) Pie charts representing single, dual and triple cytokine producing CD4 T cells. Each section of the pie chart represents a specific combination of cytokines, as indicated by the color. (E-G) Proportion of CCR7+CD45RA- (central memory), CCR7-CD45RA- (effector memory), CCR7+CD45RA+ (naïve), and CCR7-CD45RA+

(T_{EMRA}) T cells for each peptide pool. E) IFNγ, F) IL-2, G) TNFα. (A-C, E-G) Each point represents one participant, n=10, median ± interquartile range is shown. Two-tailed Mann-Whitney test.

Differential T cell response against disease-specific peptide pools

The data above suggests that certain antigens and epitopes are differentially recognized in ATB vs. healthy IGRA+ individuals. Accordingly, we next tested whether reactivity to these "ATB-specific epitopes," could differentiate individuals with ATB at diagnosis, as a cohort more relevant for diagnostic purposes vs. IGRA+ and IGRA- healthy controls. Peptides that were exclusively recognized by participants with ATB (mid-treatment) as compared to our previous studies in IGRA+ participants were pooled into an ATB-pool (ATB116) consisting of 116 peptides (Table S1). Reactivity to MTB300 was utilized as a comparator.

IFNγ response was determined in a Sri Lankan cohort, including 24 ATB patients (recruited at diagnosis), 25 IGRA+, and 43 IGRA- participants. ATB116 stimulation resulted in significantly higher IFNγ in patients with ATB compared to both IGRA+ and IGRA- controls (Figure 5A), with 63% of ATB responding compared to 8% of IGRA+ and 16% of IGRA- participants. The frequency of participants responding to MTB300 was similar between the cohorts. MTB300 could not discriminate between ATB and IGRA+, but showed significantly higher magnitude of reactivity in IGRA+ compared to IGRA- (Figure 5A). MTB300 contains epitopes that are conserved in nontuberculous mycobacteria, which have been shown to correlate with reactivity observed in IGRA- participants (39). Similar results were obtained when responses were measured by ICS rather than Fluorospot (Figure 5B). The IFNγ, IL-2, and TNFα cytokine production was determined in 9 participants with ATB (at diagnosis) and 9 IGRA- participants from Sri Lanka. The frequency of IFNγ and TNFα producing CD4 cells against the ATB116 was significantly higher in ATB compared to IGRA- participants (Figure 5B). No differences were observed between these cohorts in response to MTB300.

These results were independently confirmed in an independent cohort from the Republic of Moldova of ATB (at diagnosis and mid-treatment), IGRA+ and IGRA- participants (household contacts of an ATB index case). Similar to what was observed in the case of the Sri Lankan cohort, ATB116 stimulation resulted in a significantly higher magnitude of response in patients with ATB, both at diagnosis and mid-treatment (100% responders), compared to IGRA+ and IGRA- controls

(35% vs. 30% responders; Figure 5C). MTB300 did not distinguish between the different cohorts (Figure 5C).

Next, we calculated the sensitivity and specificity for ATB116 to explore its diagnostic potential. The sensitivity was calculated as the percentage of the participants with ATB responding to ATB116 out of the total ATB cohort. The specificity was calculated as the percentage of IGRA+/- who did not respond to ATB116 out of the total IGRA+/- participants. ATB116 demonstrated a high sensitivity of 62.5% and specificity of over 80% in distinguishing ATB patients from those who were IGRA+ and IGRA- in the Sri Lankan cohort. Similarly, the sensitivity was 100%, and the specificity was over 60% in the Moldovan cohort differentiating ATB from IGRA+ and IGRA- participants. (Table 1).

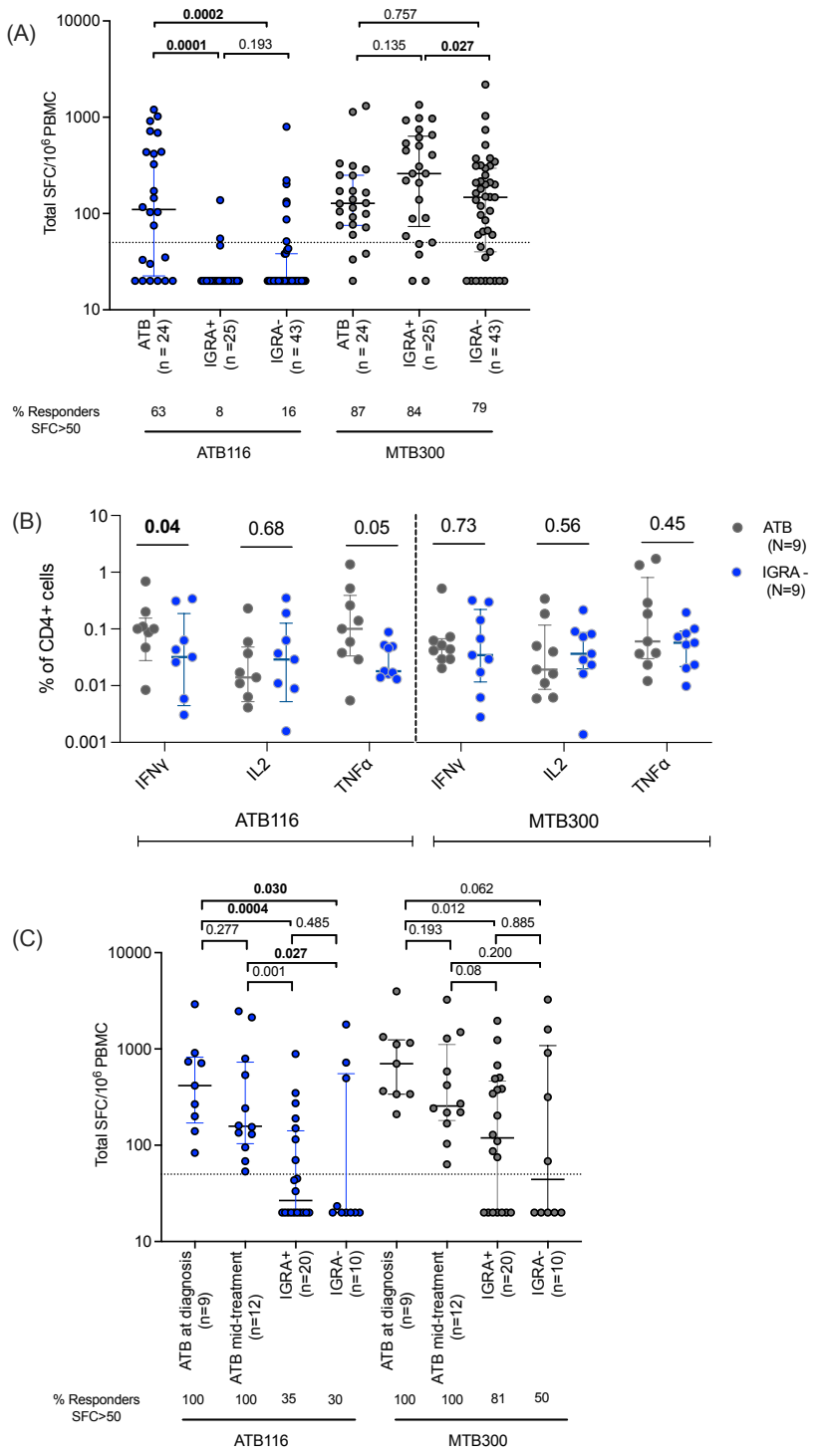


Figure 5. Cytokine response against the ATB-specific peptide pool, ATB116. (A) Magnitude of response (total SFC for IFNγ) against ATB116 and MTB300 in ATB (at diagnosis; n=24), IGRA+ (n=25), and IGRA- (n=43) from Sri Lanka. (B) Frequency of cytokine-producing, IFNγ, TNFα, and IL-2, CD4 T cells in response to ATB116 and MTB300 in ATB (at diagnosis, n=9)

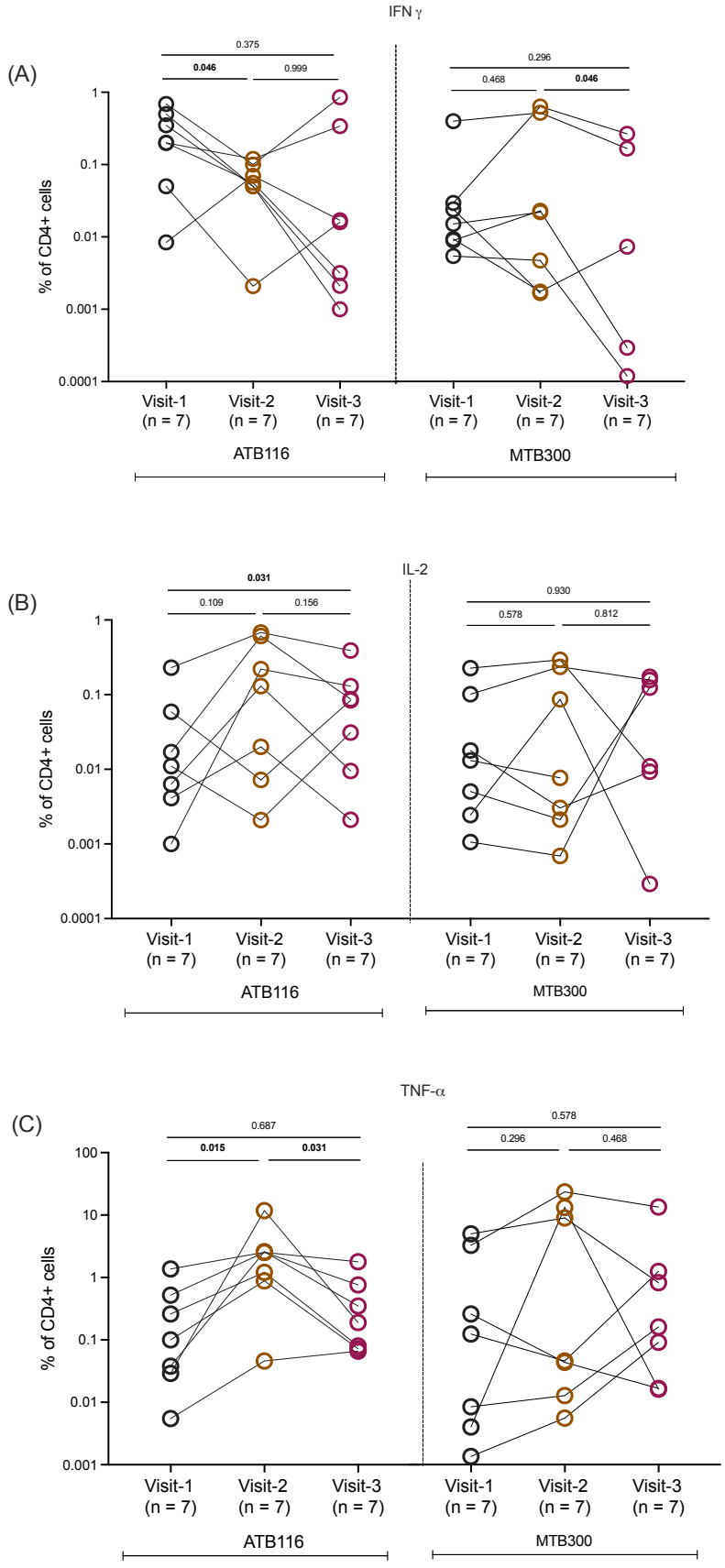
and IGRA- (n=9) participants. (C) Magnitude of response (total SFC for IFNγ) against ATB116 and MTB300 in ATB (at diagnosis; n=9), ATB (mid-treatment; n=12), IGRA+ (n=20), and IGRA- (n=10) household contacts from Moldova. (A-C) Each dot represents one participant, median ± interquartile range is shown. Two-tailed Mann Whitney test. (A, C) The dashed line indicates the cut-off used for a positive response (50 SFC/10^6 PBMCs).

Table 1. Diagnostic potential of ATB116 distinguishing patients with ATB from IGRA+ and IGRA- participants.

		Sensitivity (95% CI)	Specificity (95% CI)
Sri Lanka	ATB vs IGRA+	62.5 (42.71 to 78.84)	92 (75.03 to 98.58)
	ATB vs IGRA-		83.7 (70.0 to 91.8)
Moldova	ATB vs IGRA+	100 (84.5-100)	65 (43.2-81.8)
	ATB vs IGRA-		70 (39.6-89.2)

Longitudinal changes in epitope reactivity

Finally, to investigate the changes in T cell response during treatment, we followed 7 participants in Sri Lanka from diagnosis to mid-treatment and end of treatment. We analyzed the proportion of cytokine-producing CD4 T cells after pool stimulation. In this longitudinal cohort, we observed a significant decrease in ATB116-specific IFNγ responses between visit 1 (at diagnosis) and visit 2 (mid-treatment) (Figure 6A). At visit 3 (end of treatment), about half the cohort had a further decrease in their response, whereas in the other half the response against ATB116 increased again. The response against MTB300 did not change between visits 1 and 2 but was significantly decreased comparing visits 2 and 3 (Figure 6A). The ATB116-specific IL-2 response increased between visits 1 and 3 (Figure 6B). The response for MTB300 remained the same. Finally, the ATB116-specific TNFα response showed an increase at visit 2 compared to both visits 1 and 3 (Figure 6C). Again, there was no difference in the response against MTB300. These results indicate that the response against ATB116 changes during treatment for ATB and is differentially affected depending on the specific cytokine measured.



330

Figure 6. Characterization of ATB116- and MTB300-specific CD4 T cell responses. A-C) Frequency of cytokine-producing, IFNγ (A), IL-2 (B), and TNFα (C), CD4 T cells in response to ATB116 and MTB300 in longitudinal samples from ATB patients (visit 1; at diagnosis, visit 2; mid-treatment, visit 3; end of treatment, n=7). Each point represents one participant, median ± interquartile range is shown. Wilcoxon signed rank test.

Discussion

Here, we report the first proteome-wide identification of *Mtb*-derived T cell epitopes in a cohort of ATB patients. We have defined the epitopes from a library of over 21,000 peptides, with an in-depth investigation of the antigens in TB vaccine candidates. The lack of understanding of the broad range of *Mtb* antigens that may elicit a differential T cell response between individuals with various stages of *Mtb* infection is a bottleneck in developing diagnostic assays and vaccines. Many studies have focused on the identification of novel antigens which can be utilized for better diagnostic assays and vaccine development (21,40–43). Despite many efforts, only few antigens, <20, of the over 4,000 ORFs in the *Mtb* genome are included in subunit vaccine candidates (44–48).

Comparing the results described here with those from our previous studies in healthy IGRA+ participants (22), revealed differentially recognized epitopes and antigens between ATB and IGRA+ participants. This differential recognition of antigens can be explained by *Mtb*'s expression of infection stage-specific antigens, as previously reported (49). For instance, during *Mtb* infection, *Mtb* alters its metabolic state from active replication to slow or nonreplicating, accompanied by changes in the gene expression profile and thus protein expression and antigens available to the immune system (50). The differential expression and availability of antigens at distinct stages of TB infection might exhibit distinct patterns of differentiation and restricted capacities to mediate protective immunity (51).

The proteome-wide screen in IGRA+ participants revealed three antigenic islands that were immunodominant and related to type 7 secretion systems, with secreted and secretion apparatus proteins being recognized as antigens (22). Here, we found reactivity against antigens that are part of the antigenic islands, albeit with a more restricted recognition of primarily secreted proteins. The underlying cause of this is unclear, but it could be due to *Mtb*'s differential expression of proteins in different stages of TB infection. Our studies also highlighted a hierarchy of responses for the vaccine candidate and IGRA antigens in the two cohorts, with the well-studied antigens, Rv0288, Rv3875, Rv3874, and Rv1196, being the most frequently recognized. This highlights the complexity of *Mtb*-specific T cell responses and suggests the need for vaccine strategies targeting many different antigens.

The previous work identifying epitopes from IGRA+ participants led to the development of a peptide "megapool", MTB300, which allows for capturing diverse *Mtb*-specific responses irrespective of the HLA alleles expressed in the population (37). Using the same multi-epitope approach here, we defined a new megapool, ATB116, which contains epitopes that are not included in MTB300 and have only been found to be reactive in participants with ATB. Specifically, ATB116 could discriminate ATB at diagnosis and mid-treatment from control participants who were IGRA+ and IGRA-, in geographically diverse cohorts from Sri Lanka and Moldova. There was a higher frequency of reactivity against ATB116 in IGRA+ and IGRA- controls from Moldova compared to Sri Lanka. This could be because the Moldovan cohort of IGRA+/- are household contacts within 6 months of a primary ATB index case and have thus been recently exposed to *Mtb*. Furthermore, ATB116-specific reactivity in longitudinal samples from patients with ATB followed from diagnosis and during treatment revealed a decreasing IFNγ production during treatment, suggesting that antigen expression levels change during treatment as it takes effect. We have previously described a gene signature in individuals with latent TB at risk of developing active TB, which overlaps with a signature from ATB (52). This gene signature approach, together with ATB116-specific reactivity, may be useful for identifying individuals at risk of developing active TB. Taken together, further exploration of the ATB116 pool for its diagnostic potential across the spectrum of *Mtb* infection is warranted. Ideally, prospective studies to investigate whether reactivity to ATB116 can predict progression to ATB. In addition, ATB116, or versions thereof, may be useful as a whole blood assay in areas with high levels of IGRA positivity where there is a need to selectively identify ATB.

We were unable to detect a functionally distinct immune response in terms of multifunctionality. The majority of the responding T cells, irrespective of the functional categories, produced a single cytokine, with the majority producing TNFα. Some differences were observed regarding memory phenotype, where the cell wall and cell processes category had a higher frequency of effector memory T cells than the other categories. This suggests that T cell responses against these epitopes result in increased differentiation of the responding memory subset (53). Overall, we observed that epitope-specific T cells were predominantly CD45RA-CCR7- effector memory cells, followed by CD45RA- CCR7+ , which agrees with previous studies in ATB patients (54–57). Taken together, it has previously been shown that differentiation of T cells towards later-

397 stage effector memory during ongoing antigen expression primarily favors the expression of IFNγ
398 and/or TNFα (53,58), as observed here. Ongoing studies are investigating the ATB116- and other
399 epitope-specific T cell phenotypes and subsets in more detail, including the involvement of Th1*
400 (22,52,59,60).

401 In conclusion, this study provides a comprehensive identification and characterization of *Mtb*-
402 specific antigens and epitopes recognized by individuals with ATB. As a result, an active-specific
403 epitope megapool was defined, which was able to differentiate individuals with ATB from IGRA+
404 and IGRA- participants in two unrelated study populations and, therefore, could be potentially
405 useful for diagnostic development.

406

Methods

Study approval

All participants provided written informed consent for participation in the study. Ethical approval was obtained from the institutional review boards at La Jolla Institute for Immunology (LJI; Protocol Numbers: VD-090, VD-143, VD-175), Universidad Peruana Cayetano Heredia (66754), Phthisiopneumology Institute (CE-3/2018), University of California San Diego (180068), and University of Colombo (EC18-122, EC15-094). Results for IGRA+ individuals from San Diego, USA (Figure 2D) and the Western Cape region, South Africa (Figure 3B) are included here for comparison purposes. They have been reported previously; San Diego cohort (22), and South African cohort (37).

Study participants

We recruited 164 participants over 18 years of age for this study from UPCH in Peru, the General Sir John Kotelawala Defense University in Sri Lanka, and the Phthisiopneumology Institute in the Republic of Moldova. From Peru, 21 participants with ATB who were mid-treatment (3-4 months post diagnosis) were recruited. From Sri Lanka, a cohort of patients with ATB (at diagnosis, n=24), IGRA+ individuals (n=25), and IGRA- individuals (n=43) were recruited. A subset of the patients with ATB (n=7) was followed longitudinally from the time of diagnosis until the end of treatment. They provided blood samples at diagnosis, 2 months post-diagnosis, and 6 months post-diagnosis. From Moldova, a cohort of patients with ATB (at diagnosis, n=9), ATB (mid-treatment, n=12), IGRA+ individuals (n=20), and IGRA- individuals (n=10) were recruited. The IGRA+ and IGRA- individuals were household contacts of a patient with active TB (e.g., an "index case"). They provided blood samples up to 6 months after the index case received their ATB diagnosis.

Healthy participants were classified into IGRA+ (i.e., Latent TB infection) and IGRA- groups based on IGRA tests (QuantiFERON-TB Gold Plus, Cellestis and/or T-spot.TB, Oxford Immunotec). Individuals with ATB had symptomatic pulmonary TB, diagnosed by a positive GenXpert (Cepheid, Inc.), positive sputum smear, and/or a positive culture.

Participants with pulmonary TB recruited in Peru were aged 18-50, had documented culture confirmed TB from sputum, and were currently 3-4 months post-diagnosis of ATB. Individuals with diagnosed HIV, HBV or HCV infection were excluded, as well as patients with significant

systemic diseases, including, for example, diabetes, renal disease, liver disease, uncontrolled hypertension, and malignancy. They provided 100 ml leukapheresis samples.

PBMC isolation and thawing

PBMCs were obtained by density gradient centrifugation (Ficoll-Hypaque, Amersham Biosciences) from leukapheresis or whole blood samples, according to the manufacturer's instructions. The PBMC processing at the site in the Republic of Moldova used SepMate tubes (StemCell). Cells were resuspended in FBS (Gemini Bio-Products) containing 10% DMSO (v/v, Sigma) and cryopreserved in liquid nitrogen.

Cryopreserved PBMC were quickly thawed by incubating each cryovial at 37°C for 2 min, and cells transferred to cold medium (RPMI 1640 with L-glutamin and 25 mM HEPES; Omega Scientific), supplemented with 5% human AB serum (GemCell), 1% penicillin streptomycin (Life Technologies), 1% glutamax (Life Technologies) and 20 U/ml benzonase nuclease (MilliporeSigma). Cells were centrifuged and resuspended in complete RPMI medium to determine cell concentration and viability using trypan blue.

Peptide screening and peptide pool preparation

The present study screened a total of 21,220 peptides. These peptides include the same library that was screened in IGRA+ participants of 20,610 *Mtb* peptides (2 to 10 per ORF, average 5), including 1,660 variants not totally conserved amongst the selected *Mtb* genomes: five complete *Mtb* genomes (CDC1551, F11, H37Ra, H37Rv and KZN 1435) and sixteen draft assemblies (22). Along with these peptides, 610 peptides were also selected, which include 93 peptides that are not found in *Mtb* but present in the *Mycobacterium bovis* BCG strains Mexico, Tokyo 172, and Pasteur 1173P2, and 517 peptides that are 15-mers overlapping by ten amino acids spanning the entire sequence of 12 TB vaccine candidate and IGRA antigens. The vaccine candidates included ID93: GLA-SE (Rv3619c, Rv3620c, Rv1813, and Rv2608), H1:IC31 (ESAT-6 and Ag85B), H4:IC31 (Ag85B, TB10.4), H56:IC31 (Ag85B, ESAT-6, and Rv2660c), M72/AS01E (Mtb32A, PPE18), and three candidates with Ag85A alone (Ad5 Ag85A, ChAdOx1-85A/MVA85A, and MVA85A). The IGRA antigens include ESAT-6, which is also a vaccine candidate antigen, and CFP10.

The peptides were synthesized as crude material on a small (1mg) scale by Mimotopes (Australia). The peptides were solubilized using DMSO at a concentration of 20mg/ml. The peptides were pooled into peptide pools. The previous peptide library was pooled into 1036 peptide pools of about 20 peptides each (average 19.9 ± 0.5). The 93 non-*Mtb* peptides were pooled into 5

peptide pools. The 15-mer peptides overlapping by 10 amino acids spanning the TB vaccine candidate and IGRA antigens were pooled into 26 peptide pools. Thus, from the 21,220 total peptides, a total of 1,067 peptide pools were made.

Individual peptides were mixed in equal amounts after being dissolved in DMSO for the megapools (PC85, PC71, ATB116, and MTB300), as described previously (37). Each peptide pool was then placed in a lyophilizing flask and subjected to lyophilization for 24 hours. The resulting semi-solid product was dissolved in water, frozen, and lyophilized again until only solid product remained. This process was repeated several times until only solid product remained after lyophilization. Finally, the peptide pool was re-suspended in DMSO at a higher concentration per peptide (0.7 mg/ml per peptide) than before lyophilization, to reduce the concentration of DMSO in the assays.

Ex vivo IFNγ and IL-17 fluorospot assay

IFNγ and IL-17 production was measured by a Fluorospot assay with all antibodies and reagents from Mabtech (Nacka Strand, Sweden). Plates were coated overnight at 4°C with an antibody mixture containing mouse anti-human IFNγ (clone 1-D1K) and mouse anti-human IL-17 (clone MT44.6). Briefly, $2x10^5$ PBMCs were added to each well of pre-coated Immobilion-FL PVDF 96-well plate in the presence of peptide pools at a concentration of 2μg/ml, individual peptides at 5μg/ml, PHA at 10μg/ml (positive control) and media containing DMSO (amount corresponding to percent DMSO in the pools/peptides, as a negative control). Plates were incubated at 37°C in a humidified CO_2 incubator for 20-24 hours. All conditions were tested in triplicate, except the negative control, which was tested in six individual wells. After incubation, plates were developed according to the manufacturer's instructions. Briefly, cells were removed and wells were washed with PBS/0.05% Tween 20 using an automated plate washer. After washing, an antibody mixture containing anti-IFNγ (7-B6-1-FS-BAM) and anti-IL-17 (MT504-WASP) prepared in PBS with 0.1% BSA was added to each well and plates were incubated for 2 hours at room temperature. The plates were again washed and incubated with diluted fluorophore-conjugated anti-BAM-490 and anti-WASP-640 antibody for 1 hour at room temperature. Finally the plates were washed and incubated with a fluorescence enhancer for 15 min, blotted dry and fluorescent spots were counted by computer assisted image analysis (IRIS Fluorospot reader, Mabtech, Sweden).

Each pool or peptide was considered positive compared to the background that had an equivalent amount of DMSO based on the following criteria: (i) 20 or more spot-forming cells (SFC) per 10^6 PBMC after background subtraction, (ii) a greater than 2-fold increase compared to the background, and (iii) $p<0.05$ by student's t-test or Poisson distribution test when comparing the peptide or pool triplicates with the negative control. The response frequency was calculated by dividing the number of participants responding by the no. of participants tested. The magnitude of response (total SFC) was calculated by summation of SFC from responding participants.

Intracellular cytokine staining assay

PBMC at $1x10^6$ per condition were stimulated with peptide pools (2µg/ml) for 18-20h in complete RPMI medium at 37°C with 5% CO_2. PBMCs incubated with DMSO at the percentage corresponding to the amount in the peptide pools were used as a negative control to assess nonspecific or background cytokine production, and anti-CD3/CD28 (1µg/ml; OKT3 and CD28.2) stimulation was used as a positive control. For Chemokine receptor staining, antibodies were added during the stimulation. After 18 hours, 2.5µg/ml each of BFA and monensin was added for an additional 5h at 37°C. Cells were then harvested and incubated in a blocking buffer containing 10%FBS and 1µg/mL Human Fc block (BD Biosciences, USA) for 20 minutes at 4°C. Next, cells were stained using fixable live/dead stain for 20 minutes at room temperature and then stained with surface-expressed antibodies (Table S2) diluted in FACS buffer and 1X Brilliant Stain Buffer (BD Biosciences, USA) for 20 minutes at room temperature. Cells were permeabilized and fixed for intracellular cytokine staining using cytoperm fixation buffer (Biolegend) for 20 minutes at room temperature. After incubation, cells were stained for cytokines (IFNγ, IL-2 and TNFα) for 20 minutes at room temperature. Samples were acquired on a ZE5 cell analyzer (BioRad). Frequencies of CD4 or CD8 T cells responding to each peptide pool were quantified by determining the total number of gated CD4 or CD8 and cytokine-producing cells and background values subtracted (as determined from the negative control) using FlowJo X Software. Combinations of cytokine-producing cells were determined using Boolean gating. The lower limit of detection for the frequency of cytokine-producing CD4 T cells after background subtraction was set to 0.0001%.

Statistical analysis

Statistical analyses were performed using GraphPad Prism software (GraphPad Software, Inc., San Diego, CA, USA, version 9.2). Data is shown as median with interquartile range. Non-

parametric test was applied after checking for normality. A two-tailed Mann-Whitney U test was used for comparison between two groups. Wilcoxon signed-rank tests were used for the comparison of cytokine-producing cells in longitudinal samples from ATB patients. p value ≤ 0.05 was considered significant.

References:

1. Global tuberculosis report 2021 [Internet]. [cited 2023 Mar 29]. Available from: https://www.who.int/publications-detail-redirect/9789240037021

2. Barry CE, Boshoff HI, Dartois V, Dick T, Ehrt S, Flynn J, et al. The spectrum of latent tuberculosis: rethinking the biology and intervention strategies. Nat Rev Microbiol. 2009 Dec;7(12):845–55.

3. Lewinsohn DM, Lewinsohn DA. New Concepts in Tuberculosis Host Defense. Clin Chest Med. 2019 Dec;40(4):703–19.

4. Pai M, Denkinger CM, Kik SV, Rangaka MX, Zwerling A, Oxlade O, et al. Gamma interferon release assays for detection of Mycobacterium tuberculosis infection. Clin Microbiol Rev. 2014 Jan;27(1):3–20.

5. Yang H, Kruh-Garcia NA, Dobos KM. Purified protein derivatives of tuberculin — past, present, and future. FEMS Immunol Med Microbiol. 2012 Dec 1;66(3):273–80.

6. Huebner RE, Schein MF, Bass JB. The tuberculin skin test. Clin Infect Dis Off Publ Infect Dis Soc Am. 1993 Dec;17(6):968–75.

7. Farhat M, Greenaway C, Pai M, Menzies D. False-positive tuberculin skin tests: what is the absolute effect of BCG and non-tuberculous mycobacteria? Int J Tuberc Lung Dis Off J Int Union Tuberc Lung Dis. 2006 Nov;10(11):1192–204.

8. Hill PC, Jackson-Sillah D, Fox A, Franken KLMC, Lugos MD, Jeffries DJ, et al. ESAT-6/CFP-10 fusion protein and peptides for optimal diagnosis of mycobacterium tuberculosis infection by ex vivo enzyme-linked immunospot assay in the Gambia. J Clin Microbiol. 2005 May;43(5):2070–4.

9. Chegou NN, Black GF, Kidd M, van Helden PD, Walzl G. Host markers in Quantiferon supernatants differentiate active TB from latent TB infection: preliminary report. BMC Pulm Med. 2009 May 16;9:21.

10. Lin PL, Flynn JL. CD8 T cells and Mycobacterium tuberculosis infection. Semin Immunopathol. 2015 May;37(3):239–49.

11. Mayer-Barber KD, Barber DL. Innate and Adaptive Cellular Immune Responses to Mycobacterium tuberculosis Infection. Cold Spring Harb Perspect Med. 2015 Jul 17;5(12):a018424.

12. Boom WH. Gammadelta T cells and Mycobacterium tuberculosis. Microbes Infect. 1999 Mar;1(3):187–95.

13. Huang S. Targeting Innate-Like T Cells in Tuberculosis. Front Immunol. 2016;7:594.

14. Van Rhijn I, Moody DB. Donor Unrestricted T Cells: A Shared Human T Cell Response. J Immunol Baltim Md 1950. 2015 Sep 1;195(5):1927–32.

15. Morgan J, Muskat K, Tippalagama R, Sette A, Burel J, Lindestam Arlehamn CS. Classical CD4 T cells as the cornerstone of antimycobacterial immunity. Immunol Rev. 2021 May;301(1):10–29.

16. Havlir DV, Barnes PF. Tuberculosis in Patients with Human Immunodeficiency Virus Infection. N Engl J Med. 1999 Feb 4;340(5):367–73.

17. Arlehamn CSL, Sidney J, Henderson R, Greenbaum JA, James EA, Moutaftsi M, et al. Dissecting mechanisms of immunodominance to the common tuberculosis antigens ESAT-6, CFP10, Rv2031c (hspX), Rv2654c (TB7.7), and Rv1038c (EsxJ). J Immunol Baltim Md 1950. 2012 May 15;188(10):5020–31.

18. Vordermeier M, Whelan AO, Hewinson RG. Recognition of Mycobacterial Epitopes by T Cells across Mammalian Species and Use of a Program That Predicts Human HLA-DR Binding Peptides To Predict Bovine Epitopes. Infect Immun. 2003 Apr;71(4):1980–7.

19. Hammond AS, Klein MR, Corrah T, Fox A, Jaye A, McAdam KP, et al. Mycobacterium tuberculosis genome-wide screen exposes multiple CD8+ T cell epitopes. Clin Exp Immunol. 2005 Apr 1;140(1):109–16.

20. McMurry J, Sbai H, Gennaro ML, Carter EJ, Martin W, De Groot AS. Analyzing Mycobacterium tuberculosis proteomes for candidate vaccine epitopes. Tuberculosis. 2005 Jan 1;85(1):95–105.

21. Lindestam Arlehamn CS, Lewinsohn D, Sette A, Lewinsohn D. Antigens for CD4 and CD8 T cells in tuberculosis. Cold Spring Harb Perspect Med. 2014 May 22;4(7):a018465.

22. Lindestam Arlehamn CS, Gerasimova A, Mele F, Henderson R, Swann J, Greenbaum JA, et al. Memory T cells in latent Mycobacterium tuberculosis infection are directed against three antigenic islands and largely contained in a CXCR3+CCR6+ Th1 subset. PLoS Pathog. 2013 Jan;9(1):e1003130.

23. Weiskopf D, Angelo MA, de Azeredo EL, Sidney J, Greenbaum JA, Fernando AN, et al. Comprehensive analysis of dengue virus-specific responses supports an HLA-linked protective role for CD8+ T cells. Proc Natl Acad Sci U S A. 2013 May 28;110(22):E2046-2053.

24. Hinz D, Oseroff C, Pham J, Sidney J, Peters B, Sette A. Definition of a pool of epitopes that recapitulates the T cell reactivity against major house dust mite allergens. Clin Exp Allergy J Br Soc Allergy Clin Immunol. 2015 Oct;45(10):1601–12.

25. Schulten V, Greenbaum JA, Hauser M, McKinney DM, Sidney J, Kolla R, et al. Previously undescribed grass pollen antigens are the major inducers of T helper 2 cytokine-producing T cells in allergic individuals. Proc Natl Acad Sci U S A. 2013 Feb 26;110(9):3459–64.

26. Commandeur S, van Meijgaarden KE, Prins C, Pichugin AV, Dijkman K, van den Eeden SJF, et al. An Unbiased Genome-Wide Mycobacterium tuberculosis Gene Expression

606 Approach To Discover Antigens Targeted by Human T Cells Expressed during Pulmonary
607 Infection. J Immunol. 2013 Feb 15;190(4):1659–71.

608 27. Coppola M, Jurion F, van den Eeden SJF, Tima HG, Franken KLMC, Geluk A, et al. In-vivo
609 expressed Mycobacterium tuberculosis antigens recognised in three mouse strains after
610 infection and BCG vaccination. NPJ Vaccines. 2021 Jun 3;6:81.

611 28. Aagaard CS, Hoang TTKT, Vingsbo-Lundberg C, Dietrich J, Andersen P. Quality and
612 vaccine efficacy of CD4+ T cell responses directed to dominant and subdominant epitopes in
613 ESAT-6 from Mycobacterium tuberculosis. J Immunol Baltim Md 1950. 2009 Aug
614 15;183(4):2659–68.

615 29. Wambre E, DeLong JH, James EA, Torres-Chinn N, Pfützner W, Möbs C, et al. Specific
616 immunotherapy modifies allergen-specific CD4+ T-cell responses in an epitope-dependent
617 manner. J Allergy Clin Immunol. 2014 Mar 1;133(3):872-879.e7.

618 30. Grifoni A, Zhang Y, Tarke A, Sidney J, Rubiro P, Reina-Campos M, et al. Defining antigen
619 targets to dissect vaccinia virus and monkeypox virus-specific T cell responses in humans.
620 Cell Host Microbe. 2022 Dec 14;30(12):1662-1670.e4.

621 31. Tarke A, Sidney J, Kidd CK, Dan JM, Ramirez SI, Yu ED, et al. Comprehensive analysis of
622 T cell immunodominance and immunoprevalence of SARS-CoV-2 epitopes in COVID-19
623 cases. Cell Rep Med. 2021 Feb 16;2(2):100204.

624 32. Lindestam Arlehamn C, Sette A. Definition of CD4 Immunosignatures Associated with
625 MTB. Front Immunol [Internet]. 2014 [cited 2023 Mar 30];5. Available from:
626 https://www.frontiersin.org/articles/10.3389/fimmu.2014.00124

627 33. Arif S, Akhter M, Khaliq A, Akhtar MW. Fusion peptide constructs from antigens of M.
628 tuberculosis producing high T-cell mediated immune response. PloS One.
629 2022;17(9):e0271126.

630 34. Wang X, Chen S, Xu Y, Zheng H, Xiao T, Li Y, et al. Identification and evaluation of the
631 novel immunodominant antigen Rv2351c from Mycobacterium tuberculosis. Emerg
632 Microbes Infect. 2017 Jan 1;6(1):1–8.

633 35. Ruhwald M, de Thurah L, Kuchaka D, Zaher MR, Salman AM, Abdel-Ghaffar AR, et al.
634 Introducing the ESAT-6 free IGRA, a companion diagnostic for TB vaccines based on
635 ESAT-6. Sci Rep. 2017 Apr 7;7(1):45969.

636 36. Lew JM, Kapopoulou A, Jones LM, Cole ST. TubercuList--10 years after. Tuberc Edinb
637 Scotl. 2011 Jan;91(1):1–7.

638 37. Arlehamn CSL, McKinney DM, Carpenter C, Paul S, Rozot V, Makgotlho E, et al. A
639 Quantitative Analysis of Complexity of Human Pathogen-Specific CD4 T Cell Responses in
640 Healthy M. tuberculosis Infected South Africans. PLOS Pathog. 2016 Jul
641 13;12(7):e1005760.

642 38. Oseroff C, Sidney J, Kotturi MF, Kolla R, Alam R, Broide DH, et al. Molecular determinants
643 of T cell epitope recognition to the common Timothy grass allergen. J Immunol Baltim Md
644 1950. 2010 Jul 15;185(2):943–55.

645 39. Lindestam Arlehamn CS, Paul S, Mele F, Huang C, Greenbaum JA, Vita R, et al.
646 Immunological consequences of intragenus conservation of Mycobacterium tuberculosis T-
647 cell epitopes. Proc Natl Acad Sci. 2015 Jan 13;112(2):E147–55.

648 40. Bettencourt P, Müller J, Nicastri A, Cantillon D, Madhavan M, Charles PD, et al.
649 Identification of antigens presented by MHC for vaccines against tuberculosis. NPJ
650 Vaccines. 2020;5(1):2.

651 41. Wang Y, Li Z, Wu S, Fleming J, Li C, Zhu G, et al. Systematic Evaluation of
652 Mycobacterium tuberculosis Proteins for Antigenic Properties Identifies Rv1485 and
653 Rv1705c as Potential Protective Subunit Vaccine Candidates. Infect Immun. 2021 Feb
654 16;89(3):e00585-20.

655 42. Lewinsohn DM, Swarbrick GM, Cansler ME, Null MD, Rajaraman V, Frieder MM, et al.
656 Human Mycobacterium tuberculosis CD8 T Cell Antigens/Epitopes Identified by a
657 Proteomic Peptide Library. PloS One. 2013;8(6):e67016.

658 43. Meier NR, Jacobsen M, Ottenhoff THM, Ritz N. A Systematic Review on Novel
659 Mycobacterium tuberculosis Antigens and Their Discriminatory Potential for the Diagnosis
660 of Latent and Active Tuberculosis. Front Immunol. 2018;9:2476.

661 44. Woodworth JS, Clemmensen HS, Battey H, Dijkman K, Lindenstrøm T, Laureano RS, et al.
662 A Mycobacterium tuberculosis-specific subunit vaccine that provides synergistic immunity
663 upon co-administration with Bacillus Calmette-Guérin. Nat Commun. 2021 Nov 18;12(1):1–
664 13.

665 45. Rodo MJ, Rozot V, Nemes E, Dintwe O, Hatherill M, Little F, et al. A comparison of
666 antigen-specific T cell responses induced by six novel tuberculosis vaccine candidates. PLoS
667 Pathog. 2019 Mar;15(3):e1007643.

668 46. Stylianou E, Harrington-Kandt R, Beglov J, Bull N, Pinpathomrat N, Swarbrick GM, et al.
669 Identification and Evaluation of Novel Protective Antigens for the Development of a
670 Candidate Tuberculosis Subunit Vaccine. Infect Immun. 2018 Jun 21;86(7):e00014-18.

671 47. Nemes E, Geldenhuys H, Rozot V, Rutkowski KT, Ratangee F, Bilek N, et al. Prevention of
672 M. tuberculosis Infection with H4:IC31 Vaccine or BCG Revaccination. N Engl J Med. 2018
673 Jul 12;379(2):138–49.

674 48. Tait DR, Hatherill M, Van Der Meeren O, Ginsberg AM, Van Brakel E, Salaun B, et al.
675 Final Analysis of a Trial of M72/AS01E Vaccine to Prevent Tuberculosis. N Engl J Med.
676 2019 Dec 19;381(25):2429–39.

677 49. Scriba TJ, Carpenter C, Pro SC, Sidney J, Musvosvi M, Rozot V, et al. Differential
678 Recognition of Mycobacterium tuberculosis–Specific Epitopes as a Function of Tuberculosis
679 Disease History. Am J Respir Crit Care Med. 2017 Sep 15;196(6):772–81.

680 50. Schnappinger D, Ehrt S, Voskuil MI, Liu Y, Mangan JA, Monahan IM, et al. Transcriptional
681 Adaptation of Mycobacterium tuberculosis within Macrophages: Insights into the
682 Phagosomal Environment. J Exp Med. 2003 Sep 1;198(5):693–704.

683 51. Moguche AO, Musvosvi M, Penn-Nicholson A, Plumlee CR, Mearns H, Geldenhuys H, et
684 al. Antigen Availability Shapes T Cell Differentiation and Function during Tuberculosis.
685 Cell Host Microbe. 2017 Jun 14;21(6):695-706.e5.

686 52. Burel JG, Singhania A, Dubelko P, Muller J, Tanner R, Parizotto E, et al. Distinct blood
687 transcriptomic signature of treatment in latent tuberculosis infected individuals at risk of
688 developing active disease. Tuberc Edinb Scotl. 2021 Dec;131:102127.

689 53. Andersen P, Scriba TJ. Moving tuberculosis vaccines from theory to practice. Nat Rev
690 Immunol. 2019 Sep;19(9):550–62.

691 54. Arrigucci R, Lakehal K, Vir P, Handler D, Davidow AL, Herrera R, et al. Active
692 Tuberculosis Is Characterized by Highly Differentiated Effector Memory Th1 Cells. Front
693 Immunol [Internet]. 2018 [cited 2023 Mar 30];9. Available from:
694 https://www.frontiersin.org/articles/10.3389/fimmu.2018.02127

695 55. Petruccioli E, Petrone L, Vanini V, Sampaolesi A, Gualano G, Girardi E, et al. IFNγ/TNFα
696 specific-cells and effector memory phenotype associate with active tuberculosis. J Infect.
697 2013 Jun;66(6):475–86.

698 56. El Fenniri L, Toossi Z, Aung H, El Iraki G, Bourkkadi J, Benamor J, et al. Polyfunctional
699 Mycobacterium tuberculosis-specific effector memory CD4+ T cells at sites of pleural TB.
700 Tuberc Edinb Scotl. 2011 May;91(3):224–30.

701 57. Pollock KM, Whitworth HS, Montamat-Sicotte DJ, Grass L, Cooke GS, Kapembwa MS, et
702 al. T-cell immunophenotyping distinguishes active from latent tuberculosis. J Infect Dis.
703 2013 Sep;208(6):952–68.

704 58. Mahnke YD, Brodie TM, Sallusto F, Roederer M, Lugli E. The who's who of T-cell
705 differentiation: Human memory T-cell subsets. Eur J Immunol. 2013;43(11):2797–809.

706 59. Strickland N, Müller TL, Berkowitz N, Goliath R, Carrington MN, Wilkinson RJ, et al.
707 Characterization of Mycobacterium tuberculosis-Specific Cells Using MHC Class II
708 Tetramers Reveals Phenotypic Differences Related to HIV Infection and Tuberculosis
709 Disease. J Immunol Baltim Md 1950. 2017 Aug 9;ji1700849.

710 60. Acosta-Rodriguez EV, Rivino L, Geginat J, Jarrossay D, Gattorno M, Lanzavecchia A, et al.
711 Surface phenotype and antigenic specificity of human interleukin 17-producing T helper
712 memory cells. Nat Immunol. 2007 Jun;8(6):639–46.

713

714

Acknowledgement

Study funding

This work was supported by the National Institutes of Health contract HHSN272200900044C (to A.S.), 75N93019C00067 (to B.P. and C.S.L.A.), and R01 AI137681 (to A.C.). The funders had no role in study design, data collection, analysis, decision to publish, or manuscript preparation.

Competing Interests

The authors have declared that no competing interests exist.

Author contributions

C.S.L.A., A.S., and B.P. participated in the design and direction of the study. S.P., J.M., C.C., and C.S.L.A. performed and analyzed the experiments. M.S., R.H.G., N.C., V.C., D.G.C., A.C., T.R., J.S.B.P., T.C., B.G., and A.D.DS, recruited participants, performed clinical evaluations, and isolated PBMCs. C.S.L.A. and S.P. wrote the manuscript. All authors read, edited, and approved the manuscript before submission.