EVALUATION OF MYCOBACTERIUM TUBERCULOSIS ANTIGEN-SPECIFIC CD8+ T CELL RESPONSES

By

Matthew D. Iles-Shih

A THESIS

Presented to the Department of Public Health and Preventive Medicine
of the Oregon Health & Science University School of Medicine
in partial fulfillment of the requirements for the degree of

Master of Public Health

June 2012

School of Medicine

Oregon Health & Science University

CERTIFICATE OF APPROVAL

This is to certify that the Master's thesis of

Matthew D. Iles-Shih

has been approved

Kevin Winthrop, MD MPH
Committee Chair

Melissa Nyendak, MD MHS
Project Mentor, Committee Member

Jodi Lapidus, PhD
Committee Member

TABLE OF CONTENTS

TABLE OF FIGURES:

ACKNOWLEDGEMENTS:

The author owes a profound debt of gratitude to the many individuals, groups, and institutions that, together, made this study possible. In addition to developing the original research hypothesis and providing access to their laboratories' data, David and Deborah Lewinsohn provided invaluable guidance throughout the study as senior project mentors. Melissa Nyendak served as the principal project mentor, lending her unique combination of expertise in epidemiology and TB immunology to this study and encouraging the author's development as an aspiring clinician-scholar. Her contributions cannot be overstated. So too, the guidance provided by members of the author's thesis committee—including Kevin Winthrop, Jodi Lapidus, and Melissa Nyendak—informed not only the organization of this text but, more fundamentally, critical aspects of analysis presented within it. Still, this study would not have been possible without the generous support of Henry Boom, Catherine Stein, and their colleagues at Case Western Reserve University's Tuberculosis Research Unit, who provided access to essential clinical and demographic data from their Kawempe Community Health Study. Equally important were the organizational skills, intellectual leadership, and technical expertise of Harriet Mayanja-Kizza, Sarah Zalwango, Roy Mugerwa, and their colleagues at Makerere University and Mulago Hospital in Kampala, Uganda. In addition, the extraordinary staff of both the Lewinsohn's laboratories at Oregon Health & Science University and Makerere University—particularly Gwendolyn Swarbrick, Meghan Cansler and Megan Null—were essential to the development and execution of the research designs, laboratory methods, and analytical approaches that produced the data utilized here. The author would also like to thank the Oregon Clinical & Translational Institute and the National Institutes of Health for providing funding for this research through a TL-1 translational research training grant (and David and Deborah Lewinsohn for supporting the fellowship application leading to this award.) Finally, and most importantly, the research presented here would not be possible without the willingness of individual patients in Uganda to contribute to research aimed at enhancing our collective understanding of and ability to prevent and treat tuberculosis.

ABSTRACT:

BACKGROUND: Recent research indicates that CD8+ T cells play a critical role in the management of Mycobacterium tuberculosis (Mtb) infections and do so, in part, through an ability to detect and respond to changes in bacillary burden. Consequently, the development of TB vaccines and diagnostic tests intended for use within high prevalence populations can be aided by the identification of CD8+ T cell immunodominant Mtb antigens and epitopes—particularly those that differentiate between active and latent TB phenotypes. **OBJECTIVE:** This study tests the hypothesis that, among TB positive subjects from a high prevalence region, CD8+ T cell IFN-γ response to novel CD8+ T cell-associated Mtb antigens will be positively associated with active disease. **DATA & METHODS:** Five novel antigens that were hypothesized to have LTBI specificity were selected via an IFN-γ CD8+ T cell screening process from among the Large-Scale Antigen Discovery Program's (LSADP) 38,617-element Mtb genomic peptide library. Candidates were then screened in a cross-sectional clinical validation study wherein antigen-induced IFN-γ CD8+ T cell responses were measured, via ELISPOT, for 52 ATB and 56 LTBI subjects in Kampala, Uganda. A previously validated conservative threshold for determining positive vs. negative IFN-γ CD8+ T cell response was defined, a priori, as: Mtb Antigen-elicited Spot Forming Units (SFUs) minus two times the background SFU standard deviation minus ten SFUs. TB phenotype was determined clinically, based on WHO TB diagnostic criteria. Data related to potential confounding and effect-modifying factors were extracted from household- and individual-level surveys and clinical records. Unadjusted and adjusted odds ratios for IFN-γ CD8+ T cell response by TB phenotype were estimated using univariate and multivariate regression modeling. **RESULTS:** Examination of univariate statistics showed that subjects did not differ by age or sex according to either TB phenotype or IFN-γ CD8+ T cell response. Using the predefined cut point, neither univariate nor multivariate analyses demonstrated a statistically significant association between IFN-γ CD8+ T cell response and TB phenotype for individual antigens or the five antigen set as a whole (OR=0.64, $p = 0.24$). However, quantitative antigen specific INF-γ production was significantly associated with active tuberculosis for four of five antigens. Specifically, a statistically significant increase in IFN-γ CD8+ T cell response of 24.8-27.1 spot forming units was observe among ATB vs. LTBI subjects ($0.001 \leq p \leq 0.015$). **CONCLUSIONS:** Continuous CD8+ T cell IFN-γ responses to four out of five novel Mtb antigens demonstrate a positive relationship between CD8+ T cell response and elevated bacillary burden.

Key Words: Mycobacterium tuberculosis; CD8-positive T-Lymphocytes; Antigens, CD8; Interferon-gamma Release Tests; Epidemiology.

INTRODUCTION & OVERVIEW:

Mycobacterium tuberculosis (Mtb) is one of humanity's oldest and most successful bacterial pathogens. Even today, an estimated one third of the earth's population lives with latent Mtb infection and nearly two million individuals die annually from the active form of the disease. (1) Thus, as we enter the second decade of the 21st century, TB remains a major challenge to global health that, if it is to be met in a timely and equitable manner, will be best addressed through a multifaceted approach to research, prevention, and treatment. (2; 3) For this reason, Comas and Gagneux, for example, claim that moving to a "systems epidemiology" paradigm, as depicted schematically below (Figure 1), is needed in order to understand and intervene in the complex, recursive dynamics driving the pandemic [ibid].

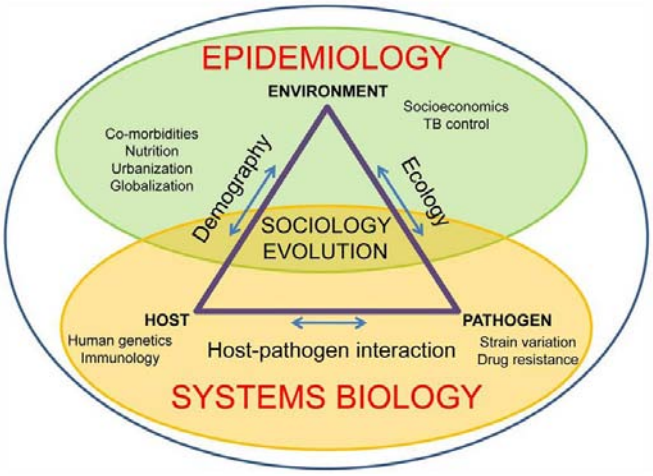


Figure 1 - A "Systems Epidemiology" Approach to TB Research (2)

In the clinical/biological domain of prevention strategies, the most important objective is the development of an effective TB vaccine, as well as the identification of biomarkers for TB phenotype (i.e., along the continuum of latent infection to active disease) that can assist in the targeting of treatment to those most at risk for progression to active disease. (4; 5; 6; 7; 8; 9) Progress in both of these areas will be dependent on the identification of Mtb antigens associated with, in particular, latent

illness [ibid]—progress now facilitated by recent research that is shedding light on the unique ability of CD8+ T cells to recognize and manage high intracellular Mtb burden, thereby modulating the complex host-pathogen dynamic governing movement along the continuum of latent TB infection (LTBI) and the potential for transitioning to active tuberculosis (ATB). (10; 11; 12; 13; 14; 15; 16; 17) Therefore, particularly in high LTBI prevalence settings, CD8+ T cell-associated Mtb antigens and epitopes would seem to hold special promise both as potential components of a composite vaccine and as the basis of diagnostic tests capable of more precisely locating patients along the LTBI-ATB continuum. Unfortunately, the interferon release assays (IGRAs) currently used in commercial TB diagnostics and antigen identification are based on CD4 T cell models that, by themselves,[i] have been unable to distinguish active and latent infection or differentiate between the antigens/epitopes associated with these states. (18; 19) Consequently, until only recently, the characterization of CD8+ T cell-associated immunodominant antigens and epitopes has not been undertaken in a comprehensive and systematic way.

The study described here is part of one such CD8+ T cell-driven effort, the Large-Scale CD8 Antigen Discovery Program (LSADP.) The LSADP has defined and continues to screen a 38,617-member genomic peptide library via a combination of computational and laboratory-based methods that are designed to identify TB phenotype-associated (LTBI vs. ATB) antigens. The initial round of screening identified five antigen candidates using an IFN-γ CD8+ T cell enzyme-linked immunospot (ELISPOT) assay and peripheral blood mononuclear cells (PBMC) samples from 20 TB-positive patients in Oregon. In 2008, these candidates were entered into a cross-sectional clinical validation study utilizing PBMC samples from 52 active and 56 latent Mtb-infected subjects in Kampala, Uganda. The analysis presented here utilizes the IFN-γ CD8+ T cell ELISPOT assay results from this Ugandan study, in combination with subject-level demographic, clinical, and environmental data, to determine whether IFN-γ CD8+ T cell response: (1) varies by TB phenotype; (2) is influenced by other clinical, environmental, nutritional, or behavioral factors; and (3) varies significantly between antigens.

To accomplish these objectives, unadjusted and adjusted odds ratios and linear regression parameter estimates for IFN-γ CD8+ T cell response and TB phenotype were developed through logistic and linear regression modeling. The first set of models examined TB phenotype as the outcome and IFN-γ CD8+ T cell response as the primary independent variable of interest whereas the second set sought to identify factors predictive of IFN-γ CD8+ T cell response. TB phenotype (LTBI vs. ATB), was determined clinically using WHO TB diagnostic protocol using evidence from sputum samples and chest radiography. IFN-γ CD8+ T cell response was examined in several forms. First, a binary variable was defined as positive or negative for each study subject, by antigen, based on whether the average number of spot-forming units per ELISPOT well greater than two standard deviations plus ten over the mean background SFU count. This outcome variable was modeled both for individual antigens and for the five antigen-set as a whole. In the latter case, a series of binary, tri-level ordinal, and count variables were defined based upon the number of antigen-specific positive responses. Second, IFN-γ CD8+ T cell response was examined as a continuous variable for each antigen, defined as the average number of SFUs per well minus mean background SFU count. Data related to additional predictor variables were extracted from household- and individual-level surveys and clinical records and included the following: Mtb Exposure Rates & TB-associated Symptoms (Hours per Month of Contact with ATB Case, Number of Cardinal Symptoms, and Extent of Lung Disease by Chest X-ray); General Nutritional Status (BMI); Basic Demographics (Sex, Age); and non-tuberculous Mycobacterium exposure (Water Source.)

While study design, analysis, and results will be presented in greater detail later in this paper, it will be useful to first review tuberculosis epidemiology and control strategies and to discuss the immunologic rationale for the approach taken to antigen discovery that has made the current study possible. In so doing, the significance and complexity of TB phenotype—in particular, latency's dynamic continuum—are highlighted and the role of certain socioeconomic, demographic, behavioral, and

general health factors are explored in relation to TB immunology more generally and IGRA-based diagnostic and vaccine development efforts, in particular.

BACKGROUND:

TUBERCULOSIS: SIGNIFICANCE, EPIDEMIOLOGY & APPROACHES TO CONTROL

An estimated third of the world's population is currently infected with Mycobacterium tuberculosis and an additional nine million people contract Mtb annually. While the majority of these individuals harbor the bacteria as a latent infection, approximately fourteen and a half million experience clinically active disease on an annual basis, resulting in approximately two million deaths per year. (1) Unfortunately, these high rates of prevalence and incidence are concentrated in those regions of the world with poorly-resourced health infrastructures and high rates of HIV infection—with co-infection occurring in an estimated 7% of TB cases but reaching > 50% in some regions (see Figure 2 and Figure 3, below.) Indeed, twelve of the fifteen countries with the highest TB incidence rates (>100/100,1000) are located in sub-Saharan Africa, while high rates are also found in China, India, portions of Southeast Asia and Micronesia while moderately high rates (25-99/100,000) have been observed in North Africa, Latin America, and Eastern Europe. Thus, even though the WHO's estimates suggest that global incidence may have peaked in 2003, TB remains a leading cause of infectious mortality worldwide [ibid].

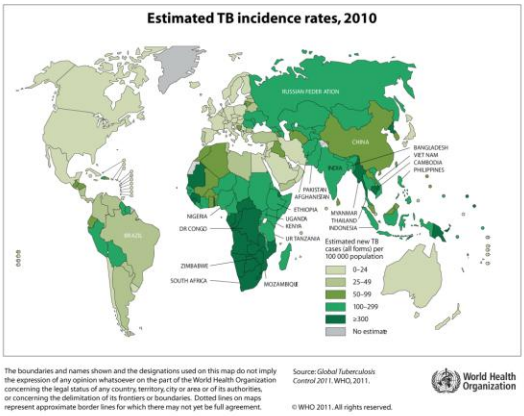


Figure 2 - TB Incidence Rates (2010)

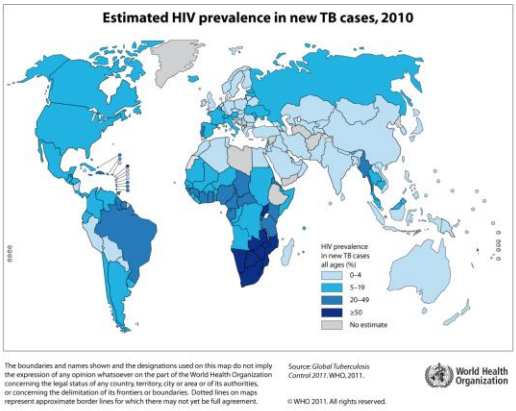


Figure 3 - HIV Prevalence in New TB Cases (2010)

Moreover, poor global leadership in TB control and treatment during much of the latter half of the 20th century, in the setting of insufficiently robust local treatment programs and practices and difficult therapy regimens (both in terms of side effect profiles and length of treatment), have led to an increasing prevalence of multi-drug resistant TB (MDR-TB) and extensively drug-resistant TB (XDR-TB) (20; 21) that threatens to render traditional treatment regimens and prevention strategies less effective. While the development and dissemination of "directly observed therapy, short-course" (DOTS) since the mid-1990s has undoubtedly helped to both improve cure rates (estimated to be between 75%-80% program-wide) and reduce antibiotic resistance, the continued emergence and circulation of resistant mycobacteria is a grave concern among global public health practitioners [ibid]. Leveraging the DOTS experience and incorporating TB-related Millennium Development goals, in 2006 the WHO initiated a more expansive program, the Global Plan to Stop TB, supported institutionally by the Stop TB Partnership. (22; 23) In so doing, it sought explicitly to build upon the successes of its predecessor as well as better address some of the challenges that it uncovered, including those associated with HIV co-infection, certain sociocultural and program-design issues relevant to community involvement and treatment acceptance, as well as other systems-level challenges related to articulating with local (and usually under-

resourced) health care infrastructures, promoting the development of new drugs and their equitable distribution, and supporting other innovations in prevention, diagnosis and treatment [ibid]. The goals and strategies set forth in these original documents have been refined and expanded upon since the program's initiation and continue to constitute an important set of conceptual and institutional frameworks for global TB prevention and treatment efforts. (24; 25) Although the picture is, at best, mixed and significant regional variation persists, data suggests that these efforts have met with some success. (26) Even so the achievement of two of the three TB Millennium Goals (i.e., achieving an incidence rate of ≤ 1:1 million and halving the 1990 prevalence rate) is unlikely under the proposed timeframe.

With high rates of latent infection, poorly tolerated treatment, increasingly complicated drug susceptibility profiles, and significant resource constraints, it is increasingly recognized that the mere extension of traditional prevention and treatment programs are important but insufficient strategies for controlling TB. For this reason, recent long-term TB control efforts have focused on two parallel strategies: targeted treatment and vaccine development. The former would help to make existent control strategies more effective by identifying those individuals with latent infections who are at risk for developing active disease so that they can be targeted for treatment prior to becoming infectious or suffering TB-associated morbidity/mortality. The latter would work, prophylactically and/or following exposure, by priming potential hosts' immune systems with the ability to recognize and respond to Mtb. (9) As they have evolved over the past decade, both strategies have been forced to grapple with the problematic distinction between latent and active states and are utilizing some of the same technologies in pursuit of their different goals. Because the research described in this paper, and the larger efforts in which it is embedded, are located at the intersection of these goals, technologies, and categories, it is worthwhile to provide a brief discussion of this emerging research paradigm.

To begin, we can summarize the general relationship between bacterial burden and clinical illness as follows: a paucibacillary state is associated with clinically asymptomatic "latent" tuberculosis infection (LTBI) whereas symptomatic, or "active" tuberculosis (ATB), is associated with higher bacillary burden (see diagram, below.) Consequently, the gold standard for diagnosis of ATB is the direct measurement and culture of Mtb bacilli (from sputum or tissue samples.) Alternatively, LTBI is inferred in asymptomatic individuals without radiographic evidence of disease and from tests that measure immunologic response to Mtb antigens either en vivo via the tuberculin skin test or the more recently developed—and more specific—ex vivo interferon gamma release assays (IGRAs.)

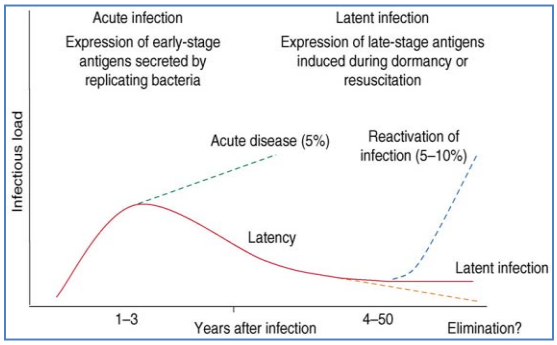


Figure 4 - Development of TB Infection & Disease (9)

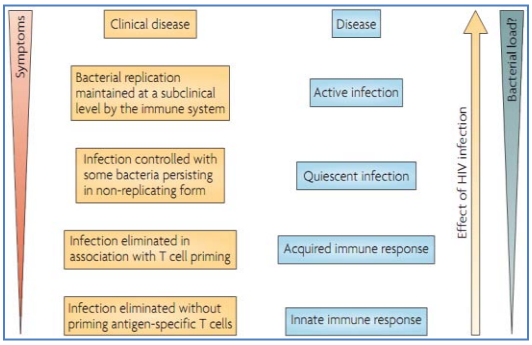


Figure 5 - Mtb Infection as a Spectrum (27)

This now common-sense LTBI/ATB ontology and the technologies used to diagnose and prevent TB's spread are, in some ways, inadequate as a means of describing the complex Mtb-host dynamic. Consequently, a number of scholars have suggested that it may be more useful to think of these two states as part of a dynamic continuum, or spectrum (see Figure 5, above.) (27; 28; 29; 30) At the same time, as Figure 4 indicates, it is crucial to also recognize that this immunologic continuum does have important inflection points. For example, while nearly a third of the human population may harbor Mtb, there is only a small subset for whom a crucial immunological/microbiologic

tipping point is reached, after which bacterial proliferation progresses to clinically recognizable disease.

Developing more nuanced understanding of such critical regions along the spectrum of Mtb infection is necessary for developing prophylactic and diagnostic tools capable of arresting this ongoing pandemic. Any attempt to paint such a detailed portrait of these regions must begin with the epidemiologic literature on tuberculosis, which provides insight into a number of factors that together help to determine an individual's location along the continuum of infection. Moreover, some subset of these factors may influence the outcomes of the kinds of immunologic assays used in the current study to identify TB phenotype-specific antigens, and so are particularly important to consider when conducting research at the intersection of TB immunology and epidemiology (see Figure 6, below.) For discussion purposes, these factors can be divided into three general categories: pathogen characteristics, host characteristics, and environmental factors.

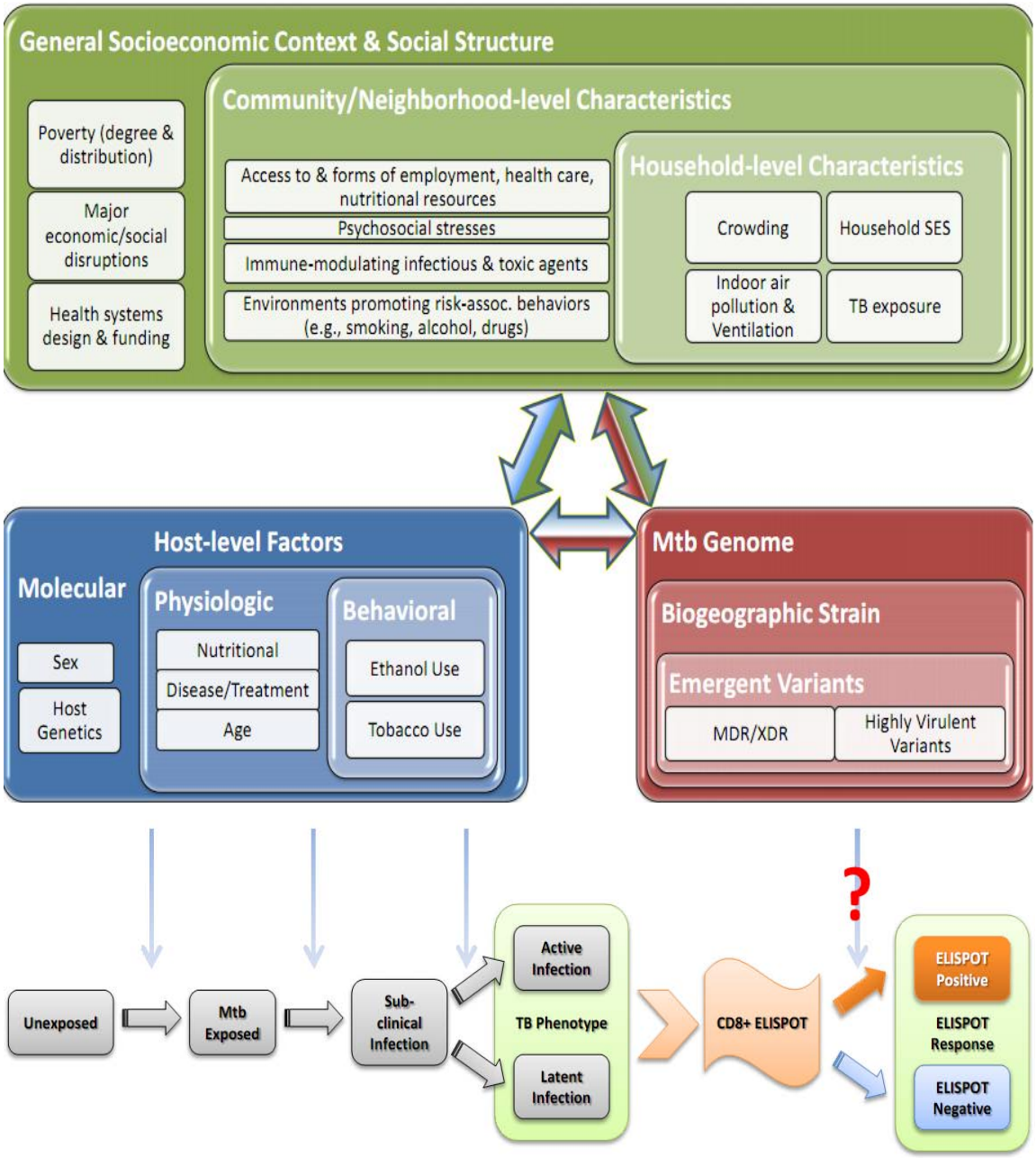


Figure 6 - Complex Associations between Mtb, Host, and Environment:

Potential Confounding/Modifying Factors Affecting the Relationship between TB Phenotype & IFN-γ CD8+ T Cell Response

With the exception of drug resistant Mtb strains' role in increasing rates of treatment failure in TB positive individuals, the extent and precise roll of diversity in Mtb genetics (Figure 7) in modulating the host-pathogen relationship is the least well

characterized of these three divisions. Recent research has shown significant genetic variation among individual clinical isolates and pronounced geographic differences in the distribution of Mtb strains (see Figure 8, below.) (31) (Incidentally, research suggests that the single Mycobacterium tuberculosis Uganda genotype accounts for nearly 100% of TB cases in Kampala.) (32) However, although several studies have suggested strain-related difference in virulence, significant differences in pathogenicity among these biogeographic strains have not been demonstrated. (33) On the other hand, increased virulence has been noted among specific Mtb variants isolated from TB outbreaks and, in most cases, has been linked to genetic modifications of genes coding either for enzymes known to alter cell wall composition or those associated with signaling pathways—altering host response to the pathogen or pathogen response to the environment, respectively—in ways that shift the host-pathogen dynamic in favor of Mtb. (34)

Figure 7 - Pathogen Characteristics associated with Mtb Infection & Disease

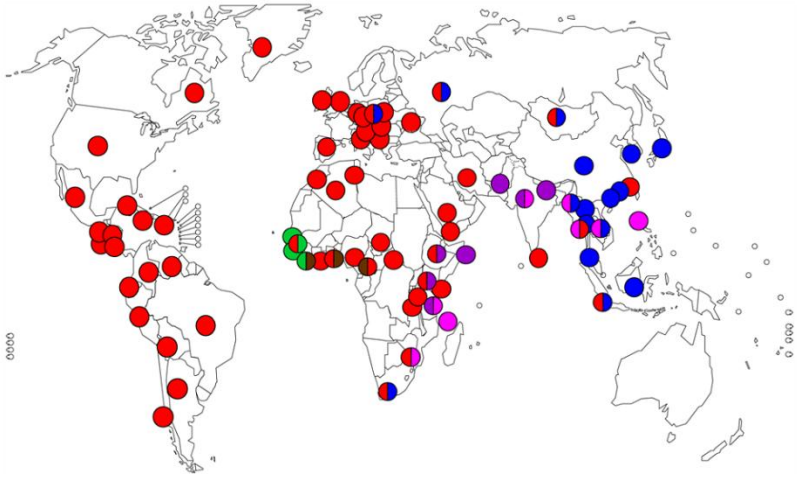


Figure 8 - Global Distribution of Six Major Mtb Strain Lineages (2)

Better characterized are a set of host characteristics (see Figure 9) that are known to be associated with infection and disease progression in specific contexts—often acting, it is believed, through modulation of the host's immune system. Host genetics, for example, have been implicated in a number of studies and estimates of heritability of disease concordance and immune response to Mtb antigens range as high as 71%. (35) While early interest in the role of HLA type have provided mixed results, recent linkage studies have repeatedly identified sites on chromosomes 5p15, 20p and 20q (and possibly 11p14) that impart resistance. Similarly, gene association studies have validated polymorphisms in INF-γ, NRAMP1, and NOS2A (with equivocal results for IL10, CCL2, DC-SIGN, P2RX7, VDR, TLR2, TLR9 and SP110.) [ibid] Among these studies, Stein et al conducted a genome-wide linkage scan that identified several sites (2q21-2q-24, 5p13-5q22, and 7p22-7p21) associated with different TB phenotypes using a Ugandan sample drawn from the population utilized in the current study. (36) Moving up from the molecular to the physiologic level, a number of nutritional, toxin/substance-related, and disease/treatment-related factors are known to increase conversion rates. For example, although being underweight (BMI ≤ 18.5) has long been known to increase risk of TB,[ii] recent research is beginning to elucidate some of the complex relationships between malnutrition and immune system dysfunction, including the potential relationship between protein malnutrition and T cell mediated Mtb control as well as between immune function and both nuclear hormone receptors and insulin receptors. (37; 38) Other nutritional perturbations, including Vitamin A, C and D deficiencies and elevated iron levels, are thought to affect TB rates by impeding macrophage activation and enhancing Mtb growth, respectively. (38; 39) Similarly, a number of systemic disorders are known to increase TB risk including: chronic renal disease, diabetes, hematologic and some solid tumor malignancies, celiac disease, hemophilia, silicosis, and HIV. Likewise, certain medical treatment—gastrectomy, solid organ transplantation, and immunosuppressive therapies (e.g., TNF-α inhibitors and glucocorticoids)—have been shown to increase risk.

Other host-level characteristics, such as age and sex, have also been shown to modulate risk. Among adults the risk of conversion to ATB demonstrates a complex pattern that—depending on the epidemiologic context (e.g., developed- vs. developing-nation status)—includes a rise in late adolescence/young adulthood and/or a increase in later adulthood—the former thought to be due to increased exposure and conversion rates among a younger population and the latter due to cumulatively increasing prevalence and, perhaps in elderly patients, immune system decline. Another set of interesting observations involve the sex of the patient. On the one hand, there is higher prevalence of infection among men —thought to be a function of increased exposure alone. However, for reasons that remain unclear, once infected, women tend to be more likely to develop active TB and, in particular, extra-pulmonary TB. (40; 41) In addition, certain behavioral practices have been shown to be associated with an increased risk for both Mtb infection and progression to active disease. In particular, the smoking of tobacco products (and even exposure to second-hand smoke) and consumption of alcohol are associated with significant increased risk (1.4-1.6 RR for conversion to ATB associated with smoking; 1.89-4.59 adjusted RR for active TB with alcohol consumption of ≥ 40g per day or diagnosis of alcohol use disorder.) (42; 43; 44; 45) Importantly, for both tobacco and alcohol use, elevated risk persists even after adjustment potential confounders.

Figure 9 - Host Characteristics associated with Mtb Infection & Disease

Finally, although long recognized as important, recent research and experience has only highlighted the centrality of environment and socioeconomic factors in determining patterns of TB prevalence and incidence and the distribution of TB-related morbidity

and mortality (see Figure 10.) (46; 40; 47; 48) As key distal and intermediate components of complex causal chains, factors such as residential crowding and working conditions amenable to transmission, indoor air pollution, access to health services, access to nutrition, sociocultural and economic environments that promote risk-enhancing health conditions and behaviors (such as diabetes, HIV infection, alcohol abuse), as well as psychosocial stresses that increase exposure to and/or the level of risk associated with more proximal casual factors.[ibid] For example, in a study of household contacts of ATB cases in Kampala, Uganda researchers found that co-prevalence of ATB in contacts was predicted by environmental factors such as hours per day of contact with the ATB case (adjusted OR = 2.39 for ≥18 hours/day, 95% CI: 1.23-4.63) and the type of housing in which subjects resided (adjusted OR = 2.12 for Muzigo vs. other, 95% CI: 1.13-3.95), as well as characteristics of the ATB case (such as presence of cavitary disease (adjusted OR = 2.23, 95% CI: 1.14-4.35) and chronic cough (adjusted OR =15.47, 95% CI: 5.21-45.95).) (49) [iii] In a second, community-wide cross-sectional study of an impoverished periurban area near Kampala, members of the same research group also found some evidence that ATB prevalence was positively associated with residential crowding (≥ 3 people per room) and the number of windows per person in the residence, although these associations were crude and did not reach the $\alpha < 0.05$ level of significance ($p = 0.061$ and 0.105, respectively.) (50) History also provides a number of compelling natural experiments that highlight the importance of general socioeconomic conditions relative to TB control. For example, the experience of Western developed nations over the course of the 19th and 20th centuries suggests that an increase in overall economic wellbeing and improved living conditions can lead to sharp and rapid declines in TB prevalence, even in the absence of significant investment in targeted prevention and treatment programs. Unfortunately, the more recent experience of economic and social decline among many former Soviet Block nations following the collapse of communism is thought to have greatly contributed to the re-emergence of TB (and the development of MDR TB strains) secondary to the erosion of

health infrastructures, poor nutrition, poor living conditions, increased rates of alcohol abuse, depression, and general declines in health status. (51; 52)

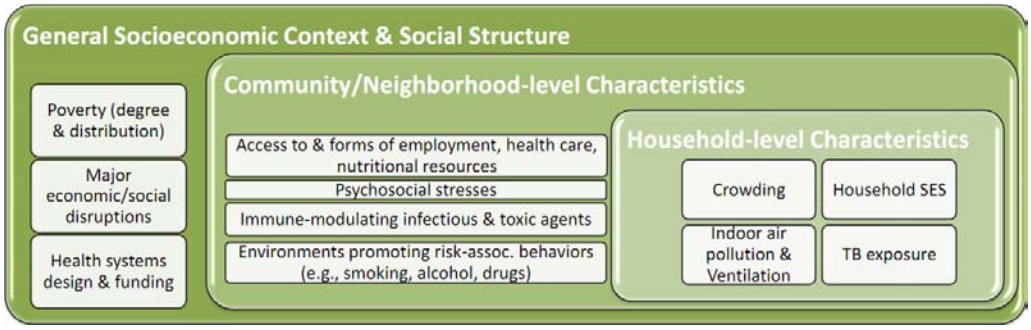


Figure 10 - Environmental Factors associated with Mtb Infection & Disease

Together, these multiple environmental, pathogen, and host characteristics form a complex network of risk factors that work to determine the probability of infection and disease progression (depicted graphically in Figure 6, page 9.) Their role in determining the location of the host-pathogen dynamic along the TB phenotype, however, has important implications for TB control. On the one hand, those factors that are modifiable offer an opportunity to reduce risk through strategic public health interventions. On the other, the number and complexity these relationships represents a challenge for research at the intersection of TB epidemiology and immunology—such as the research described here—when they also impact other diagnostic or therapeutic outcomes of interest. For example, the presence of a real relationship between TB phenotype and antigen-induced INF-γ CD8+ T cell response, as hypothesized in the current study, can be rendered less legible by confounding, interaction or the effect of strong independent determinants of cellular response.

TOWARDS RATIONAL DESIGN IN TB PREVENTION AND CONTROL

Having reviewed the background TB epidemiology we are now in a better position to define the central problem addressed in this paper. First, the array of clinical and

demographic factors identified above are known to correlate with Mtb infection and with the transition between latent infection and active disease and can, therefore, be useful in risk-profiling groups and individuals so as to target screening and various interventions. However, the proximal immunologic processes involved are far less well understood and characterizing the transitional dynamic at this level, it will be argued, is crucial to not only more precisely identifying those at risk for transitioning to active disease (and, therefore, to developing and targeting more effective prevention and treatment strategies) but to the very development of such prevention and treatment strategies. In particular, the development of simple laboratory-based diagnostic modalities capable of detecting this transition prior to the onset of clinical disease and, more importantly, the identification and inclusion of Mtb antigens associated with both latent and active portions of the infection continuum could improve the efficacy of composite vaccines may by providing both prophylactic and post-infectious immunity (see Figure 11 from Andersen (9), below.) In the next subsection, this potential is briefly explored.

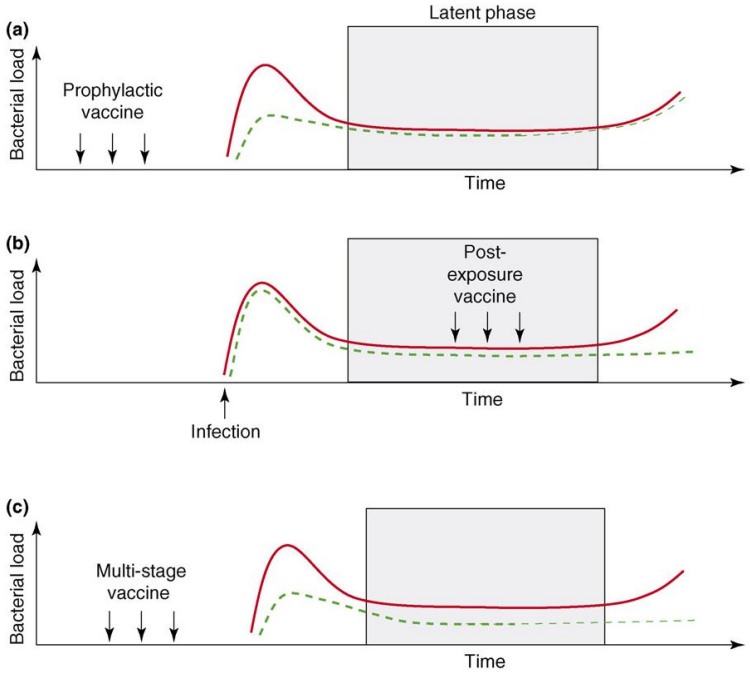


Figure 11 - Prophylactic, Post-infectious, & Combined Approaches to Vaccination (9)

TB Vaccines:

As discussed above, the eradication of TB in high-prevalence regions could be accomplished in several ways: though the redress of global socioeconomic inequalities, the diagnosis and treatment of vast populations of LTBI individuals, and/or the development of an effective vaccine. Unfortunately, the timeframe under which the elimination of global poverty and improvement in regional and national health infrastructures might be expected to take place would leave hundreds of millions of individuals at risk for TB-related morbidity and mortality for, potentially, many decades to come. The second strategy—mass diagnosis and treatment—would impose a massive financial burden on health systems that are, in most cases, already tragically under-resourced. Alternatively, the development of an effective vaccine would provide a practical, relatively low-cost solution that could be implemented over a reasonably short period of time. The Bacille-Calmette-Guerin (BCG) vaccine, first developed in 1928, remains the only viable vaccine option at this time and is, in fact, still used extensively in many TB-endemic areas. Unfortunately, its efficacy is limited with respect to both the prevention of pulmonary TB, in general, and several specific forms of pediatric TB, in particular. (53; 54; 55) Additionally, as a live vaccine that requires active replication within the host in order to induce protective memory, its effect depends on host naivety to mycobacterium. Moreover, its protective effect appears to wane over time (10-20 years), leaving previously immunized individuals susceptible to later infection and disease. Thus, in spite of decades of use, the BCG vaccine has failed to stem TB's rising tide. (56)

In many ways, this is not surprising. The BCG vaccine was developed through a non-selective process that utilized an attenuated strain of *Mycobacterium bovis* related to but differing in significant ways from Mtb. While this approach was revolutionary for its time, in the intervening century biomedicine has generated new understandings of pathogen-host dynamics and human immunology and has produced a new set of technologies and techniques that provide the foundation for a new set of vaccine design

approaches. This has led to the development of several promising new vaccine candidates, a number of which are currently under clinical trial (see, for example, the current Aeras Vaccine Portfolio depicted in Figure 12, below.)

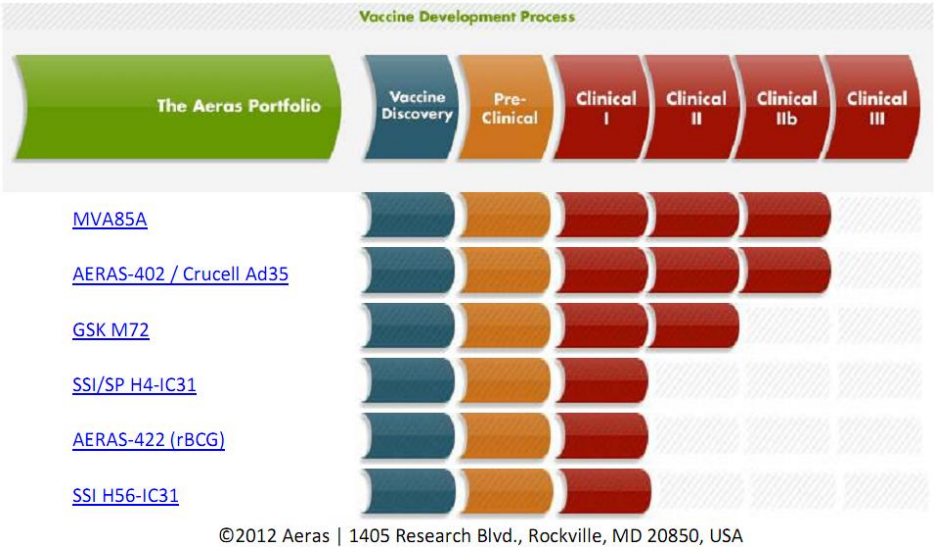


Figure 12 - TB Vaccines Currently Undergoing Clinical Trials (57)

In the remainder of this section, the most central of these developments—understandings of T cell mediated immunity and the methods used to detect and quantify T cell response—are reviewed and discussed in relation to the specific Mtb antigen discovery strategy employed to identify the five antigens used in this study.

T Cell-mediated Mtb Immunity and the Role of CD8+ T Cell Response Mtb Infection Management:

In 1882 Robert Koch identified *Mycobacterium tuberculosis*, an aerobic acid-fast Gram-positive bacillus, as the cause of tuberculosis. In the intervening 128 years, much has been learned about this organism and the pathologies for which it is responsible. It is now understood, for example, that Mtb's remarkable ability to evade the immune system relies in significant part on its capacity to arrest phagosome-lysosome fusion within macrophages and dendritic cells. Having successfully phagocytized bacteria that they are then unable to degrade, these antigen presenting cells (APCs) are unable to

efficiently present Mtb antigens to neighboring lymphocytes and may, themselves, come to serve as Mtb reservoirs. This dampening of the adaptive immune response allows the organism to persist in the body for many years as a latent TB infection. (58; 59) For this reason, many contemporary efforts to combat Mtb have focused on improving the immune system's capacity to effectively manage the infection by way of enhancing the synergistic interactions between Mtb antigen-presenting ACPs and T lymphocytes. (7; 60) To date, most of these efforts have focused on CD4+ T cells, which have been shown to be critical in controlling Mtb dissemination, reactivation, and the maintenance of latency through, primarily, MHC Class II and INF-γ-mediated activation of macrophages and cytolytic T lymphocytes. (60)

Recently, however, there has been an increasing interest in the role of CD8+ T cells in controlling disease reactivation and promoting the clearance of latent infection through cytotoxic T cell-mediated destruction of Mtb-infected macrophages and the inhibition of intracellular Mtb replication [ibid]. For example, there is evidence that cytotoxic T cells are more prevalent during early stages of infection while cytokine-producing CD8+ T cells are disproportionately represented during later stages of infection. Indeed, research has demonstrated the existence of Mtb-specific CD8+ T cell with both classical and non-classical HLA restriction, suggesting that CD8+ T cells not only augment CD4+ T cell function but also provide a basis for the immune regulation of intracellular infection. (59; 61) Moreover, Mtb-specific CD8+ T cells are found at significantly elevated numbers following infection with Mtb and preferentially recognize those APCs with higher Mtb burdens. (15) For example, in a recent pediatric study, Mtb-specific CD8+ T cell response to high bacillary load appeared to distinguish between children with TB disease and those children who were household contacts of adults with smear positive disease. (17) Finally, Mtb has been shown to achieve an altered metabolic state within macrophage phagosomes across the TB phenotype continuum, such that the array of antigens presented during latent and active phases of infection may well be different. (62; 27)

STUDY HYPOTHESIS AND OBJECTIVES:

Collectively, these recent findings in TB epidemiology and immunology indicate that:

1) **CD8+ T cells may be capable of distinguishing between infection phenotype based on bacillary burden and antigen repertoire;**
2) **These two characteristics of infection are likely correlated to an infected individual's location along the ATB-LTBI spectrum;**
3) **CD8+ T cell response may, therefore, be useful as a means of indexing disease phenotype (i.e., distinguishing latent from active infection);**
4) **A relationship between CD8+ T cell response and TB phenotype may be affected and/or obscured by clinical, environmental, nutritional, or demographic factors that modulate CD8+ T cell activity.**

Thus the characterization of immunodominant CD8+ T cell-associated Mtb antigens and epitopes—which are currently not well defined—may allow for the identification of TB phenotype specific candidates for new vaccines that are capable, in particular, of inducing host response in the setting of increasing bacillary burden. Moreover, as CD8+ T cells are able to preferentially recognize, activate, and lyse heavily infected APCs, they may be uniquely able to serve as a proxy for disease progression and resolution in the context of either therapy or drug/vaccine development. (59) The analysis presented here builds upon these recent finding and their implications for vaccine and diagnostics development by examining, ex vivo, CD8+ T cell response to five novel CD8+ T cell-associated Mtb antigens.

HYPOTHESIS: The authors hypothesize that CD8+ T cell-associated antigens will differentiate between ATB-LTBI status by eliciting higher IFN-γ responses in cells taken from individuals with ATB vs. LTBI.

OBJECTIVES: To test this hypothesis, the study examines data from a recent clinical validation study involving 108 TB-positive subjects in Kampala, Uganda to determine whether IFN-γ CD8+ T cell response:

1. ***Is positively associated with Active Tuberculosis;***
2. ***Is influenced by other clinical, environmental, nutritional, or demographic factors;***
3. ***Varies significantly between the five antigens.***[iv]

Together, these three study objectives will address the study hypothesis by examining the relationship between CD8+ T cell response and clinical disease states associated with differing levels of bacillary burdens (Objectives 1 & 2) and then by seeking to characterize the degree of TB phenotype specificity of within the antigen set (Objective 3.)

MATERIALS & METHODS

CD8+ ANTIGEN DISCOVERY AND VALIDATION:

In collaboration with an international network of colleagues, the laboratories of David and Deborah Lewinsohn at Oregon Health & Science University led a CD8+ *large scale antigen discovery program* based on a three-pronged approach—expression cloning, proteomics, and genomic peptide library development—to identify Mtb-associated CD8+ epitopes that were both common and immunogenic. This effort utilized a complex weighting schema[v] to select 422 genetic loci (approximately 10% of the Mtb proteome) from which 38,617 potential antigens were defined (each a 15 amino acid polymer that included an 11 amino acid overlap with each adjacent neighbor.)[vi] Subsequently, this peptide library was screened (directly, ex-vivo) against an ethnically diverse set of 20 TB-positive donors' CD8+ T cells (and several clones with unknown antigen specificity.)[vii] Screening was conducted using a CD8+ T cell INF-γ Release Assays (CD8_IGRAs.) Specifically, a CD8+ T cell-based Enzyme-Linked Immunosorbent Spot, or ELISPOT, was

used to detect and quantify Mtb-sensitized T cells' IFN-γ release following exposure to Mtb antigens. The process involved the following general steps: (1) T cells and antigen presenting cells (APCs) isolated from a subject's blood (standardized to 250,000 per well) were added, in combination with specific exogenous Mtb antigens, to an array of wells each pre-coated with IFN-γ-specific antibodies; (2) following incubation, wells were rinsed and a second, IFN-γ-specific conjugated antibody was added, forming a complex with any bound IFN-γ; (3) after a second rinse, a substrate was added each well that bound to and stained any antibody-IFN-γ-antibody complex that was present; (4) finally, the number of stained spots in each well was counted and compared to positive/negative controls in order to assess the magnitude of response. (63; 64; 8) [viii] Using this method, 47 promising antigens and 4 CD8+ HLA-B restricted epitope candidates were identified.

A two phase investigation was undertaken in an effort to characterize antigen/epitope recognition among Mtb-positive individuals and to evaluate antigen specificity for TB phenotype (ATB versus LTBI). In the diagnostic screening portion of this study ("Phase I"), 51 peptide pools and epitopes were screened against CD8+ T cells from 20 TB-positive donors to evaluate immunogenicity. To proceed to the next phase, an antigen needed to meet the minimum criteria of eliciting response from a least one of the twenty donors.[ix] Among these responders, candidates were ranked according to the robustness of elicited responses. In this manner, five strong antigens candidates were selected. In 2008, in collaboration with Case Western Reserve University's transnational Tuberculosis Research Unit (TBRU) and Makerere University, the antigens and epitopes identified through this process progressed on to a clinical validation study ("Phase II") in Kampala, Uganda. As in Phase I, at the time of enrollment in Phase II, the 108 TB-positive subjects provided blood samples from which cryopreserved peripheral blood mononuclear cells (PBMCs) were obtained. These samples were then analyzed using a CD8+ T cell ELISPOT assay at the Joint Clinical Research Center immunology laboratory (Kampala.) The study isolated CD8+ T cells and dendridic cells utilizing a technique involving magnetic bead-depletion of CD4 CD56 cells.[x]

STUDY SAMPLE:

Phase II of the antigen selection and validation process was designed as a cross sectional study involving 108 TB-positive subjects (52 active, 56 latent) to assess differential response among active and latent TB-positive individuals. HIV-negative individuals of greater than 15 years of age were recruited from an ongoing household contact study, the Kawempe Community Health Study (KCHS),[xi] then being conducted by the Tuberculosis Research Unit (TBRU) in Kampala. Individuals with ATB were identified through Mtb-positive sputum cultures and clinical and radiographic findings consistent with active TB. LTBI classification, alternatively, relied on a positive TB skin test (≥10 mm induration) in the absence of clinical or radiographic evidence of active TB. Subjects underwent clinical evaluations that included a physical exam and history (symptoms, behavioral factors, risk factors, nutritional status) using standard data collection forms, HIV EIA, chest radiographs, TSTs, sputum and/or blood AFB smear and culture. Table 1, below, provides a summary of the sample characteristics. In designing this portion (Phase II) of the study, an original sample size of 50 ATB and 50 LTBI subjects was planned in order to provide 89% power to detect a 50%-20% differential in positive response rate to each antigen, a significance level of $\alpha = 0.05$. In so doing, the prevalence of positive response among ATB and LTBI subjects was estimated separately, with a 14% margin of error (i.e., half the width of the 95% confidence interval). The results of this diagnostic validation investigation provide the outcome data for the current study.

DATA MANAGEMENT & VARIABLE SELECTION:[xii]

The demographic and clinical data obtained through the TBRU subject database and ELISPOT response data from OHSU were consolidated and then evaluated for completeness and consistency. Two of the original 110 subjects included in the ELISPOT dataset were not part of the TBRU study and were excluded from the analysis. All data were inspected for suspicious outliers and missing or irrational values. Where

appropriate, inconsistent string responses were standardized. For ordinal variables with unevenly distributed and/or large numbers of response values, an effort was made to rationally re-categorize the variable using fewer levels. For example, due to the presence of thirteen non-Gandan tribes each with small (1-5) cell counts, tribal affiliation was redefined as “Ganda” vs. “non-Ganda” (77:31). Similarly, smoking, alcohol use, and water source were re-categorized from multilevel and continuous variables to binary variables that signified: a history of vs. no history of smoking, consumption vs. no consumption of alcohol, and tap water vs. spring or other exposed water source, respectively. These re-categorizations of smoked tobacco- and alcohol-use were done in the setting of incomplete data regarding *current* smoking status (particularly among LTBI subjects) and well-documented imprecision of self-reported *quantified* estimates of stigma-associated substance-use. In so doing, however, an increased accuracy and completeness is accomplished in exchange for a reduction in precision; the *current* smoking status for an individual with a positive *history* of smoking is uncertain and the actual quantities of alcohol consumed by two individuals coded as positive for alcohol-use may differ markedly. The re-categorization of water source was done so as to consolidated variable levels representing similar risk of exposure to non-tuberculous Mycobacteria: tap water vs. open water sources (springs and wells.) BMI was computed using data on height and weight. Mean background response was calculated as the average number of SFUs from three ELISPOT wells to which no antigenic material had been added.

Variables that were potentially associated with or modified TB phenotype and/or INF-γ CD8+ T cell response were identified and grouped according to the following etiologic domains: Mtb Exposure Rates & TB-associated Symptoms, General Health Status, Basic Demographics, Health/Immune-modulating Behaviors and Environmental Exposures, TB-specific and general Immune Reactivity, and Socioeconomic Status. Prior to assessing an independent variable’s relationship to INF-γ CD8+ T cell response and TB phenotype, each was evaluated in terms of the percentage of missing data, conceptual validity (e.g., ensuring that, as a cut-point, BMI < 18.5 constituted a reasonable

threshold for identifying subjects with a degree of nutritional deficiency capable of altering immune function), and the existence of plausible alternatives. Where clear redundancies existed, those that were more conceptually sound and which resulted in less sample-size reduction were selected. At times, the effort to align variables with specific dimensions directly influenced data modification. For example, in re-categorizing occupation, the final three-level scheme ("Not Formally Employed" vs. "Service, Sales & Labor" vs. "White Collar & Technical Trade") is more likely to highlight possible socioeconomic differences between groups, as opposed to other possible schemes that might have differently emphasized inter-profession variations in sex, age, environmental exposures, general health status, and/or behavioral patterns. Similarly, while differences between Gandan and "non-Gandan" subjects within the Gandan kingdom (where Kampala is located) are most likely do to SES, there may be other associations with health-related behaviors, demographic differences, or even genetic (e.g., HLA variants).[xiii]

Following this selection process, the following variables, by etiologic category, were included in the descriptive analyses of the sample by TB Phenotype (Active/Latent Status); Mtb Exposure Rates & TB-associated Symptoms[xiv] (Hours per Month of Contact with ATB Case, Number of Cardinal Symptoms, and Extent of Lung Disease by Chest X-ray); General Health Status (BMI); Basic Demographics (Sex, Age); Health/Immune-modulating Behaviors and Environmental Exposures (Smoking, Alcohol Consumption, and Household Indoor Ventilation); Mycobacteria-specific and general Immune Reactivity (PPD, Water Source—a proxy for non-tuberculous mycobacterium exposure—and Mean Background Response); Socioeconomic Status (Tribal Affiliation, Occupation, Residential Crowding.) Of these, a smaller subset of variables were identified as combining greater ontological precision, biological plausibility, and causal proximity to TB phenotype and/or INF-γ CD8+ T cell response. This subset (see Table 1, below) was selected for inclusion in multivariate models of INF-γ CD8+ T cell response. Of note, four of these variables were specific to either LTBI or ATB subjects (Sputum Smear result, TB-associated Symptomatology, and Extent of Radiographic Evidence of Lung Disease) and

were analyzed by TB phenotype-specific sample subsets. Another variable, Water Source, was included as a possible marker of exposure to non-tuberculous Mycobacteria species and expected to be associated with INF-γ CD8+ T cell response, secondary to cross-reactivity, but not TB phenotype. All other variables were considered to be potential confounders of the TB phenotype-INF-γ response relationship, in addition to being possible independent main effects.

Table 1 –Independent Variables Evaluated for Inclusion in Multivariate Models of INF-γ CD8+ T Cell Response

Conceptual Domain	Construct Category	Variable
Host-level Characteristics	Demographics	**Sex**
		Age
	TB Infection Characteristics	**TB Phenotype**
		Sputum Smear (*ATB only*)‡
		Cardinal Symptoms (*ATB only)‡*
		Extent Lung Ds. (*ATB only*)‡
	Nutritional	**Underweight**
Environmental Factors	Mtb Exposure	TB+ Contact Exposure (*LTBI only*)‡
	Non-tuberculous Mycobacteria Exposure	**Water Source**

‡ Analyses that included TB phenotype-specific variables were subset for ATB/LTBI, as appropriate.

Finally, several INF-γ CD8+ T Cell Response variables were defined (see Table 2, below.) First, for each antigen, two variables were developed: (1) antigen-associated INF-γ CD8+ T Cell response above mean background response was calculated as a continuous variable and (2) a binary variable was defined as "positive" when antigen-associated INF-γ CD8+ T Cell response was at least two standard deviations plus ten above the background SFU-count.[xv] Second, for the set of five antigens as a whole, additional variables were defined as "positive" when an antigen-associated INF-γ CD8+ T Cell response was at least two standard deviations plus ten above the background for (a) one or more individual antigens, (b) ≥ 4 vs. 0 individual antigens, and (c) 0 vs. 1-2 vs. 3-5 individual antigens. Finally, a simple count variable with six possible values (0-5 positive antigens) was defined.

Table 2 – Definitions of INF-γ CD8+ T Cell Response Variables

Antigens	Variable Type	Definition
Individual Antigens: #2, 13, 14, 19, & 25	Binary	Positive Result = [Antigen-associated INF-γ CD8+ T Cell Response (SFUs)] > [Mean Background INF-γ Response + 2 Std Deviations + 10]
	Continuous	[Antigen-associated INF-γ CD8+ T Cell Response (SFUs)] - [Mean Background INF-γ Response (SFUs)]
Five-Antigen Set, Collectively	Binary	Positive Result = Positive INF-γ CD8+ T Cell Response (see above) in ≥ 1 Antigen
	Binary	Positive Result = Positive INF-γ CD8+ T Cell Response (see above) in ≥ 4 vs. 0 Antigens
	Three-level	Positive Response (as defined above) for 0, 1-2, & 3-5 Antigens
	Count (0-5)	No. of Antigens (0-5) with Positive INF-γ CD8+ T Cell Response (as defined above)

STATISTICAL METHODS:

The study's central hypothesis was tested through the development of regression models for the fourteen INF-γ CD8+ T Cell response outcome variables, based on Hosmer and Lemeshow's multistep analytic process. (70) First, descriptive statistics for independent variable distributions (within the sample as a whole and by TB-status and INF-γ CD8+ T Cell response) were computed. Prior efforts to render discrete independent variables more efficient via a re-categorization were re-assessed at this time and modifications were implemented where opportunities existed to further enhance or preserve biological plausibility, clinical rationality, and the integrity of independent variable-outcome relationships. Second, the relationship between each outcome variable and each independent variable was characterized using either (1) likelihood ratio chi-squared test statistics (or Fisher's exact test in cases where expected categorical variable cell counts were less than five) for nominal, ordinal, and interval variables or (2) univariate logistic regression for continuous variables. Additionally, agreement in INF-γ CD8+ T Cell response among the five antigens was assessed using Pearson correlation coefficients. Third, independent main effects were selected for inclusion in a multivariable logistic and linear regression model using a combination of criteria, including: an assessment of the association observed in step one (inclusion criteria: $p \leq 0.25$), Mallows C-statistic, stepwise and best subsets selection procedures (with entry and retention thresholds of 0.15 and 0.25, respectively), as well as a priori suspicion of independent association or confounding. Fourth, the appropriateness of each variable in a proposed model was assessed in relation to (1) the variable's Wald

chi-square or F statistic, (2) its effect on the coefficients and odds rations of fellow independent variables, particularly the primary variables of interest, INF-γ CD8+ T Cell response or TB phenotype (for study objectives 1 and 2, respectively), and (3) changes in model G-statistics and R-squared values (for logistic and linear regression models, respectively.) Fifth, scaling of continuous variables was assessed using the Lowess procedure. Where appropriate, (i) non-normally distributed variable were re-categorized as ordinal variables and (ii) extreme outliers were excluded to improve the linearity of relationships. (This latter approach was especially useful in improving the linearity of the relationship between TB phenotype and continuous INF-γ CD8+ T cell response, where the exclusion of two outliers with consistently extreme INF-γ responses generated inflection points for all antigens.) Sixth, the potential for both interaction and confounding within the preliminary main effects model were examined, based on β-associated p-value of < 0.10, for interaction, and a $> 10\%$ change in main effect OR or β_x (in logistic and linear regression models, respectively), for confounding. Seventh, the utility of including higher order covariate terms to improve model precision was explored.[xvi] Eight, the final model's goodness of fit was evaluated using, R-squared, deviance, Hosmer & Lemeshow χ^2 test,[xvii] and ROC curves, as appropriate. Finally, model diagnostics—particularly changes in χ^2 and deviance residuals—were reviewed in an effort to identify and exclude influential outliers. (All analysis was conducted using SAS 9.2.)

Using this approach, individual antigens' INF-γ responses (continuous and binary) and antigen set-wide responses were modeled against TB phenotype, as the dependent and independent variables, respectively. To address the second study objective, separate models were defined and fit for both set-wide and antigen-specific binary and continuous INF-γ CD8+ T Cell response. In these models, TB phenotype as well as other clinical, environmental and demographic factors that met the criteria outlined above were included as potential determinant of CD8+ T cell response. For the five models with a continuous dependent variable, Box Cox procedures were utilized to determine

the optimal λ and transformations were conducted, where appropriate, to improve linearity.

RESULTS:

Using a series of descriptive measures, Table 3 provides an overview of the Ugandan sample—stratified by TB phenotype—relative to key demographic, clinical, behavioral and environmental factors. Variables demonstrating a statically significant association with TB phenotype, included being underweight, alcohol use, occupation,[xviii] and crowding (in addition to TB status-specific variables, including smear positive status, having cardinal symptoms of TB infection, and evidence of lung involvement by chest x-ray), at the $\alpha = 0.05$ level, while three additional variables (sex, media background response, and indoor ventilation) were significant at the $\alpha = 0.25$ level. Re-categorization schemes for nominal and ordinal variables were re-evaluated to ensure efficiency and preservation of any associations. (The reclassification schemes presented here for continuous variables were informed by assessments of linearity that occurred later in the model-building process.)

Table 3 - Sample Characteristics by TB Phenotype

Sample Characteristics by TB Phenotype		TB Status				
		Active (52)		Latent (56)		p-value
Variable	Reference Value [units]	No./ Mean	Percentage / Range; SD	No. / Mean	Percentage / Range; SD	
Sex	Female	26	43%	35	57%	0.1917
	Male	26	55%	21	45%	--
AGE	[Years]	27	18-67; 10.54	24	17-80; 14.9	0.5942
Sputum Smear	Positive	48	92%	--	--	
	Negative	4	8%	--	--	
No. Cardinal Signs/ Symptoms	[Range: 0-7]	4	0-7; 1.36	0	0-3	<0.0001
Extent of Lung Involvement [n=107]	Normal CXR	3	5%	52	95%	<0.0001
	Minimal Disease	3	60%	2	40%	
	Moderate Disease	17	94%	1	6%	
	Advanced Disease	29	100%	0	0%	
Underweight	BMI ≤ 18.5	22	73%	8	27%	0.0018
	BMI > 18.6	30	38%	48	62%	--
PPD	[No. mm]	17	0-25.2; 5.12	15.7	10.4-22.9; 3.29	0.5834
Mean Media SFU Count	[No. SFUs]	20	0-190; 36.46	10	0-326; 57.08	0.2617
	≥ 40 SFUs	10	36%	18	64%	0.1294
	< 40 SFUs	42	53%	38	48%	--
Smokes	Yes	10	59%	7	41%	0.3304
	No	42	46%	49	54%	--
Drinks EtOH	Yes	11	32%	23	68%	0.0282
	No	41	55%	33	45%	--
Food Preparation Method (n=96)	Wood/ Charcoal	42	48%	46	52%	0.429
	Gas/Electric/Other	5	63%	3	38%	--
Occupation (n=99)	White Collar / Tech Trade	12	80%	3	20%	0.0072
	Sales, Service, & Labor	28	51%	27	49%	0.1531
	Not Formally Employed	10	34%	19	66%	--
Tribe (Ganda vs. Other)	Ganda	38	49%	39	51%	0.6936
	Non-Ganda	14	45%	17	55%	--
House Type (Muzigo vs. other) [n=69]	Muzigo	28	53%	25	47%	0.2649
	Semi-detached/ Free-standing House	11	69%	5	31%	--
No. Sharing Subject's Bedroom	[No. Co-habitants]	2	0-8; 1.53	2	0-11; 2.04	0.5376
	≥ 2	42	55%	34	44.7%	0.0249
	0-1	10	31%	22	68.8%	--
Indoor Ventilation [n= 107]	[# per Room]	1.6	0-4; 0.72	1.4	0.5-3; 0.57	0.1071
TB+ Contact Exposure	[Hrs/ month]			60.5	18-85; 16.17	
Watersource	Tap Water	38	49%	40	51%	0.8486
	Spring or Other Source	14	47%	16	53%	--

UNIVARIATE ANALYSIS OF INF-γ CD8+ T CELL RESPONSE AND TB PHENOTYPE (*STUDY OBJECTIVE 1*)

To address the first aim of the study—testing the hypothesis that ATB status is predictive of a positive/elevated INF-γ CD8+ T cell response—univariate logistic

regression models were fit for thirteen CD8+ T cell response variables, with TB phenotype serving as the common independent variable. The results are presented in Table 4 (a-c), below. The relationship between INF-γ CD8+ T cell response and the ATB phenotype was found to be positive but not statistically significant for all variants of INF-γ CD8+ T cell response that relied upon the predefined, conservative cut-point. Alternatively, those models examining INF-γ CD8+ T cell response minus background response for individual antigens demonstrated a positive, statistically significant association for antigens 13, 14, and 19 and a near-significant (p = 0.052) association for Antigen 2. Specifically, INF-γ CD8+ T cell response to these antigens estimated to be 20.4 - 22.5 SFUs higher among ATB relative to LTBI subjects. Still, a large minority of LTBI as well as ATB subjects demonstrated antigen-associated CD8+ T cell response levels that met the conservative threshold for positivity established by the study's predefined cut-point: 25-32% for individual antigens and 46% for the five-antigen set as a whole. Moreover, TB phenotype was found to account for only 4% to 8% of variability within the continuous response data (see Table 10), suggesting that one or more other factors contributed significantly to the observed INF-γ CD8+ T cell response rates.

Table 4(a-c) - INF-γ CD8+ T Cell Response by TB Phenotype

Table 2a	Outcome = INF-γ Response - Background (SFUs) ‡						
Antigen	Sample		Latent TB		Active TB		p-value
	Mean	Range; Std.Dev.	Mean*	Range; Std.Dev.	Mean	Range; Std.Dev.	
Antigen #2	12	-196, 249; 57	1	-196, 173; 55	23	116, 249; 58	0.052
Antigen #13	6	-110, 161; 42	-4	-110, 134; 41	16	-106, 161; 41	0.011
Antigen #14	4	-139, 93; 39	-7	-139, 75; 40	15	-109, 93; 34	0.003
Antigen #19	10	-104, 138; 42	0	-104, 108; 39	21	-88, 138; 42	0.011
Antigen #25	22	-48, 240; 45	21	-48, 240; 46	23	-23, 209; 44	0.813

‡ Two observations with large-magnitude INF-γ CD8+ T cell responses excluded to preserve linearity.
* Note: negative values indicate mean antigen-associated response was less than background response.

Table 2b	Outcome = INF-γ Response - (Background + 2(Std.Dev.) + 10)						
Antigen	Sample		Latent TB		Active TB		p-value
	No.	(%)	No.	(%)	No.	(%)	
Antigen #2	38	35%	18	32%	20	38%	0.493
Antigen #13	31	29%	14	25%	17	33%	0.378
Antigen #14	33	31%	15	27%	18	35%	0.378
Antigen #19	35	32%	17	30%	18	35%	0.637
Antigen #25	29	27%	16	29%	13	25%	0.676

Table 2c	Outcome: Antigen Set-wide INF-γ Responses Variables							
Antigen	Variable Definition	Sample		Latent TB		Active TB		p-value
		No.	(%)	No.	(%)	No.	(%)	
Five-Antigen Set	INF-γ Response Pos. in ≥ 1 vs. 0 Antigens	56	52%	26	46%	30	58%	0.243
	INF-γ Response Pos. in ≥ 3 vs. 1-2 vs. 0 Antigens	34, 22, 52	32, 20, 48%	17, 9, 30	30, 16, 54%	17, 13, 22	33, 25, 42%	0.421
	INF-γ Response Pos. in ≥ 4 vs. 0 Antigens	33	31%	14	25%	19	37%	0.258

ASSESSING THE ROLE OF CLINICAL, ENVIRONMENTAL, & DEMOGRAPHIC FACTORS IN DETERMINING INF-γ CD8+ T CELL RESPONSE (*STUDY OBJECTIVE 2*)

In addition to TB phenotype, a set of eight other factors with the potential for contributing to or modulating INF-γ CD8+ T cell response patterns were identified, based upon biological plausibility and data availability. This set included: sex, age, low BMI, non-tuberculous Mtb exposure (exposed water sources), and several TB phenotype-specific factors (such as Mtb exposure level and number of symptoms and signs of active disease.) Univariate assessments of relationships between these variables and INF-γ CD8+ T cell response were conducted for both individual antigens and for the five-antigen set, as a whole. The latter is summarized in Table 5, below, for the binary variable "INF-γ CD8+ T cell response ≥ 1 of 5 Antigens" and is representative of findings for other set-wide variables. (Note: this variable was found to be most predictive of TB phenotype and is, therefore, used as the default variable for modeling of antigen set-wide INF-γ CD8+ T cell response in the remainder of the paper.) Of note, only water source was found to be associated with INF-γ CD8+ T cell response at the $p < 0.05$ level ($OR_{TapWater} = 0.28$), while TB phenotype and being underweight were the only variables

meeting the $p < 0.25$ threshold ($OR_{Active} = 1.57$ with $CI_{95\%} = 0.74\text{-}3.37$ and $OR_{underweight} = 0.52$ with $CI_{95\%}$: 0.22-1.21, respectively.)

Table 5 - Sample Characteristics by Set-wide INF-γ CD8+ T Cell Response: Descriptive & Univariate Statistics

Variable	Reference Value [units]	Pos. in ≥ 1 Antigens (56)		Pos. in 0 Antigens (52)		p-value
		No./ Mean	Percentage / Range; SD	No./ Mean	Percentage / Range; SD	
Sex	Female	34	56%	27	44%	0.358
	Male	22	47%	25	53%	--
Age	Years Old (continuous)	26	17-75; 12.58	25	18-80; 13.45	0.896
Sputum Smear (ATB only)	Positive	28	58%	20	42%	0.747
	Negative	2	50%	2	50%	--
Cardinal Symptoms (ATB only)	0-7 Symptoms	3.87	0-7; 1.55	4.14	1-6; 1.08	0.477
Extent Lung Ds. (ATB only)	Normal/Min/Mod/Adv.	Adv.	Norm.-Adv.	Adv.	Norm.-Adv.	0.399
Underweight	BMI ≤ 18.5	12	40%	18	60%	0.129
	BMI > 18.6	44	56%	34	44%	--
TB+ Contact Exposure	[Hrs/ month]	61.3	18-85; 19.57	59.8	28-79; 13.0	0.732
Watersource	Tap Water	34	44%	44	56%	0.007
	Spring or Other Source	22	73%	8	27%	--

Univariate statistics describing the association between the ten antigen-specific INF-γ CD8+ T cell response variables based on (a) antigen-associated INF-γ CD8+ T cell response minus mean background response (continuous) and (b) "positive" INF-γ CD8+ T cell response (binary) were calculated for each antigen separately and are summarized in Table 10, in the Appendix. The only variable shown to be consistently associated with INF-γ CD8+ T cell response at the $p = 0.05$ level was water source, which demonstrated a positive relationship between INF-γ response and the use of water from a non-tap source. In other words, CD8+ T cells from subjects whose households obtained water from open sources, such as springs or local wells, were found to have high response rates to Mtb antigens then those utilizing tap water (74% vs. 44%, respectively.) Finally, to assess correlations among covariates, Pearson correlation coefficients were calculated for the ten independent variables of interest. Beyond the relationships already noted between TB phenotype and several independent variables (BMI and

cardinal sign/symptoms of ATB), all additional associations among independent variables were found to be, at most, mild/weak in magnitude and did not present any concerns for colinearity (a conclusion reconfirmed later in the model building process by examining variable inflation factor statistics.)

Employing the model development process described in the methods section, logistic and linear regression models for INF-γ CD8+ T cell response variables were fit for each antigen (Table 6 and Table 7) and for the collection of antigens, as a whole (Table 8). In all models, the variable "Underweight" was included as a negative confounder of the TB phenotype-INF-γ response relationship (i.e., ORs for TB phenotype decreased by at least ten percent when Underweight/Not Underweight was not included in the model.) Additionally, living in a residence with a spring, well, or other untreated non-tap "Watersource" was shown to be a significant independent predictor of INF-γ CD8+ T cell response. No additional factors met the criteria for inclusion. Linearity for each of the five models with a continuous outcome variable (antigen-associated INF-γ response minus mean background response) was assessed using the Box Cox procedure and the outcome variables transformed using the optimal λ value (where $\lambda \neq 0$). However, because these transformations did not significantly alter model performance but do limit interpretability of β_i values, the statistics presented in Table 6, below, are for non-transformed antigen-specific INF-γ CD8+ T cell response outcome variables.

As was seen in the univariate analysis, no statistically significant ORs were observed in models utilizing the predefined cut-point. However, INF-γ CD8+ T cell response minus background was predicted by active TB disease in four out of five antigens at a statistically significant level. Specifically, among antigens 2, 13, 14, and 19, ATB status was associated with an estimated increase in response rate of between 24.8 and 27.1 SFUs ($0.001 \leq p \leq 0.015$). However, even with these three variables included, models were unable to account for the majority of variability found within the INF-γ CD8+ T cell response data. For example, R-squared values for models of continuous antigen response fell between 0.03 and 0.16 (and between 0.14 and 0.16 among the

four higher-performing antigens.) Similarly, the area under ROC curves for binary CD8+ T cell response variables were found to range between 0.64 and 0.71 for individual antigens and was estimated to be 0.67 when examining zero vs. greater than or equal to one positive response among the five antigens. Thus, additional unknown factor(s) continue to account for a majority of variability in the outcome data.

Table 6 - Linear Regression Models for Continuous Antigen-specific INF-γ CD8+ T Cell Responses, Minus Background.

Antigen Candidate	Independent Variable	β Est.	p-value	R-squared
Antigen 2	TB Phenotype (Active vs. Latent TB)	27.1	0.015	0.14
	Water Source (Spring vs. Tap Water)	34.0	0.005	
	Underweight (BMI ≤ 18.5)	-19.1	0.128	
Antigen 13	TB Phenotype (Active vs. Latent TB)	24.8	0.003	0.16
	Water Source (Spring vs. Tap Water)	23.3	0.009	
	Underweight (BMI ≤ 18.5)	-14.2	0.123	
Antigen 14	TB Phenotype (Active vs. Latent TB)	26.7	0.001	0.16
	Water Source (Spring vs. Tap Water)	17.4	0.034	
	Underweight (BMI ≤ 18.5)	-14.9	0.079	
Antigen 19	TB Phenotype (Active vs. Latent TB)	25.2	0.002	0.14
	Water Source (Spring vs. Tap Water)	18.2	0.039	
	Underweight (BMI ≤ 18.5)	-17.0	0.063	
Antigen 25	TB Phenotype (Active vs. Latent TB)	5.4	0.557	0.03
	Water Source (Spring vs. Tap Water)	9.5	0.342	
	Underweight (BMI ≤ 18.5)	-11.8	0.258	

Table 7 - Logistic Regression Models for Binary Antigen-specific INF-γ CD8+ T Cell Responses‡

Antigen Candidate	Independent Variable	OR Est.	95% Wald Confidence Limits	
Antigen 2	TB Phenotype (Active vs. Latent TB)	1.99	0.80	4.95
	Water Source (Spring vs. Tap Water)	3.34	1.35	8.26
	Underweight (BMI ≤ 18.5)	0.37	0.12	1.13
Antigen 13	TB Phenotype (Active vs. Latent TB)	2.17	0.84	5.62
	Water Source (Spring vs. Tap Water)	3.04	1.20	7.69
	Underweight (BMI ≤ 18.5)	0.40	0.12	1.30
Antigen 14	TB Phenotype (Active vs. Latent TB)	2.00	0.81	4.95
	Water Source (Spring vs. Tap Water)	1.71	0.68	4.27
	Underweight (BMI ≤ 18.5)	0.44	0.15	1.33
Antigen 19	TB Phenotype (Active vs. Latent TB)	1.58	0.64	3.88
	Water Source (Spring vs. Tap Water)	3.16	1.28	7.81
	Underweight (BMI ≤ 18.5)	0.44	0.15	1.33
Antigen 25	TB Phenotype (Active vs. Latent TB)	0.91	0.37	2.29
	Water Source (Spring vs. Tap Water)	2.26	0.89	5.71
	Underweight (BMI ≤ 18.5)	0.73	0.24	2.20

‡ Estimates exclude one outlier with large one-step difference in Pearson chisquare statistic.

Table 8 - Logistic Regression Model for Antigen Set-wide INF-γ CD8+ T Cell Response Positive in ≥ 1 Antigen‡

Independent Variable	OR Estimate	95% Wald Confidence Limits	
TB Phenotype (Active vs. Latent TB)	2.08	0.88	4.93
Water Source (Spring vs. Tap Water)	3.25	1.25	8.40
Underweight (BMI ≤ 18.5)	0.48	0.18	1.26

‡ Estimates exclude two outliers with large one-step differences in the Pearson chisquare statistic.

EVALUATION OF INTER-ANTIGEN CORRELATION (STUDY OBJECTIVE 3)

IFN-γ CD8+ T cell response for each of the five antigens and for the set as a whole (see Figure 13, below) show approximately equal response rates for each antigen with subject-to-subject response variability leading to a higher cumulative response for ELISPOT-positive in ≥ 1 antigen. Specifically, the inclusion of each additional antigen resulted in an average increase of 5.2% in sample capture (above the common mean individual antigen response rate of 30.2%.) Additionally, the number of antigen-specific

ELISPOT-positive responses per subject was relatively constant across the 1-5 range (Figure 14); the percentage of the sample who responded to only one or two antigens (using the predefined cut-point) was similar to the percentage who responded to three or more (10.1% vs. 10.3%, respectively).

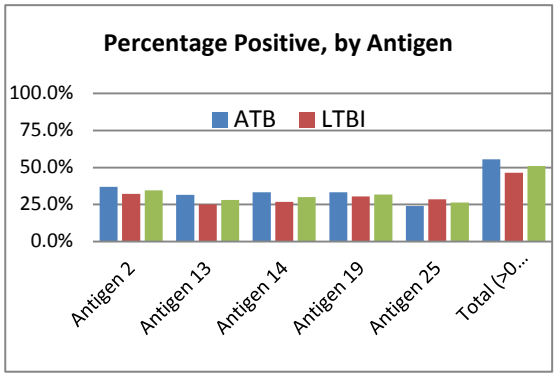


Figure 13 - INF-γ CD8+ T Cell Response by Antigen

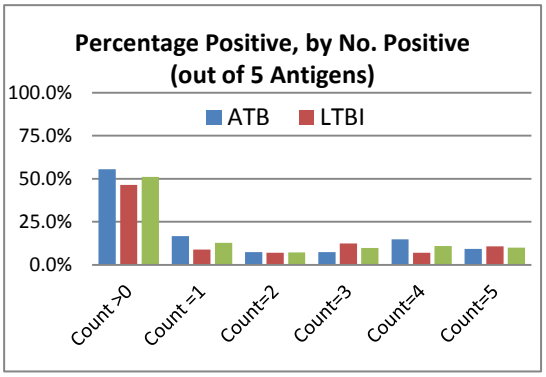


Figure 14 - INF-γ CD8+ T Cell Response by No. Positive

Correlation among the five antigens with respect to a positive INF-γ CD8+ T cell response was assessed formally by calculating Pearson correlation coefficients, which are presented in Table 9, below. Individual inter-antigen coefficients among antigens 2, 13, 14, and 19 were found to be greater than or equal to 0.88, while those for antigen 25 ranged between 0.68 and 0.75 in magnitude. These patterns would suggest a high degree of colinearity among antigens and, therefore, that the inclusion of INF-γ CD8+ T cell response variables from multiple antigens would not improve the performance of multivariate models for predicting TB phenotype. This prediction was, in fact, confirmed in the model-development process.

Table 9 - Agreement and Correlation among Antigens

Pearson Correlation Coefficients (Note: all p < 0.0001)					
Antigen	Antigen 2	Antigen 13	Antigen 14	Antigen 19	Antigen 25
Antigen #2: Mean SFUs	1	0.95	0.89	0.88	0.68
Antigen #13: Mean SFUs	0.95	1	0.88	0.91	0.70
Antigen #14: Mean SFUs	0.89	0.88	1	0.95	0.76
Antigen #19: Mean SFUs	0.88	0.91	0.95	1	0.75
Antigen #25: Mean SFUs	0.68	0.70	0.76	0.75	1

DISCUSSION:

ASSESSING EVIDENCE FOR AN ASSOCIATION BETWEEN CD8+ T CELL RESPONSE & TB PHENOTYPE:

Based on prior research indicating that CD8+ T cell response varies with APC intracellular Mtb burden and evidence for TB phenotype-specific antigen repertoires, this study sought to determine whether such cells, when presented with the appropriate Mtb antigen, could distinguish between Active TB Disease and Latent Mtb Infection. To do so, the authors utilized data from a recent cross-sectional clinical validation study in which CD4+ T cell-depleted PBMC samples from 52 active and 56 latent Mtb-infected subjects in Kampala, Uganda were exposed to five novel CD8+ T cell-associated Mtb candidates via an IFN-γ CD8+ T cell ELISPOT assay. A threshold for test positivity was defined, a priori, based on a conservative cut-point developed for the antigen-discovery process. These data, in combination with additional subject-level demographic, clinical, and environmental data, were used to determine whether IFN-γ CD8+ T cell response: (i) was predicted by TB phenotype; (ii) was influenced by other clinical, environmental, or demographic factors; and (iii) varied significantly by antigen. The following paragraphs provide an assessment of the evidence related to each of these study aims.

Objective 1: Characterizing the Relationship between IFN-γ CD8+ T cell response and TB Phenotype

Based on the observation that IFN-γ CD8+ T cell response frequency varies directly with Mtb bacillary burden, this paper hypothesized that CD8 T cells from ATB subjects would produce an elevated Mtb antigen-elicited IFN-γ CD8+ T cell response, relative to their LTBI counterparts. The analysis presented above shows; (1) the five-antigen set elicited a robust response in only 52% of study subjects; (2) higher positive response rates were seen among subjects with active TB (e.g, 46.4% of LTBI and 57.7% of ATB subjects tested positive for at least one antigen); (3) those models using continuous response variables from individual antigens showed linear, statistically significant relationships between TB phenotype and INF-γ CD8+ T cell responses in four out of five antigens, with an average 26 SFU increase in response among ATB vs. LTBI subjects; and (4) no statistically significant association between IFN-γ CD8+ T cell response and TB phenotype was demonstrated using the predetermined cut-point in either univariate or multivariate models ($OR_{ATB\ vs.\ LTBI}$ = 1.57, p=0.24).

A graphical representation of the third finding—i.e., the detection of statistically significant association between TB phenotype and continuous INF-γ CD8+ T cell response to four of the five antigens—is provided in

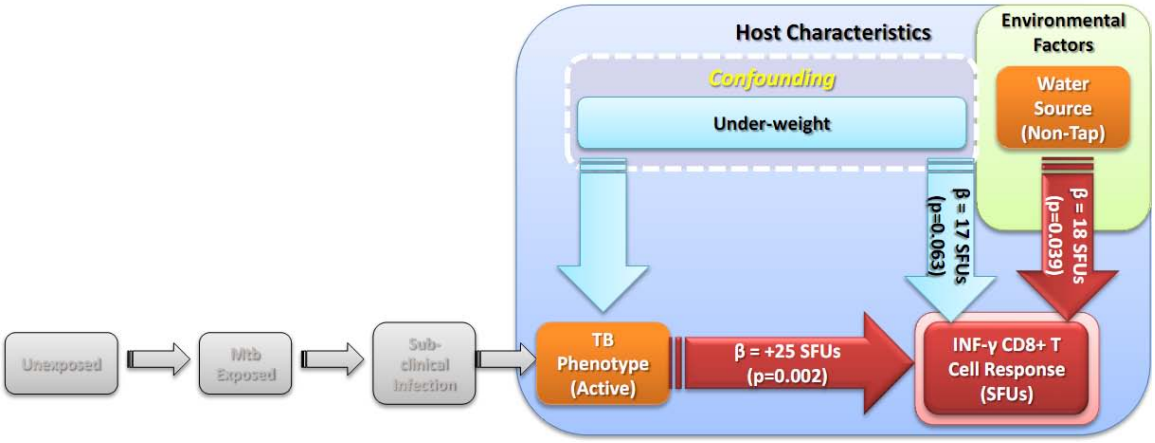


Figure 15, below. It depicts the magnitude and precision of estimates of association between TB phenotype, water source, low BMI, and CD8+ T cell response (over background) with respect to Antigen 19.

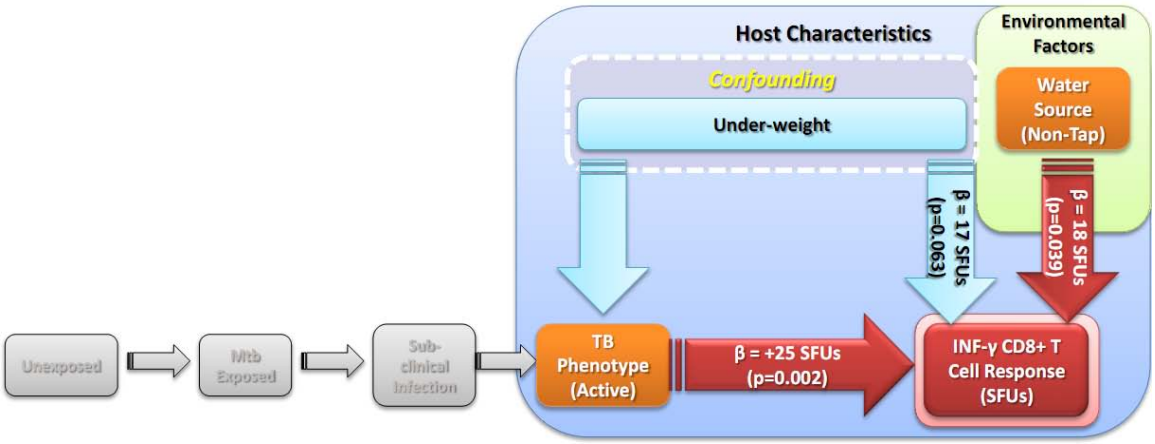


Figure 15 – Determinants of Continuous INF-γ CD8+ T cell Response to Antigen 19.

Similarly, a graphic representation of the fourth finding—i.e., no statistically significant association between TB phenotype and INF-γ CD8+ T cell response when employing a conservative cut-point to distinguish positive vs. negative INF-γ response—is provided in Figure 16, below. It depicts the magnitude and precision of estimates of association between TB phenotype, water source, low BMI, and CD8+ T cell response for the five antigen set when a positive screening result is defined as achieving an antigen-elicited response greater than two standard deviations plus ten SFUs over background for at least one of five antigens.

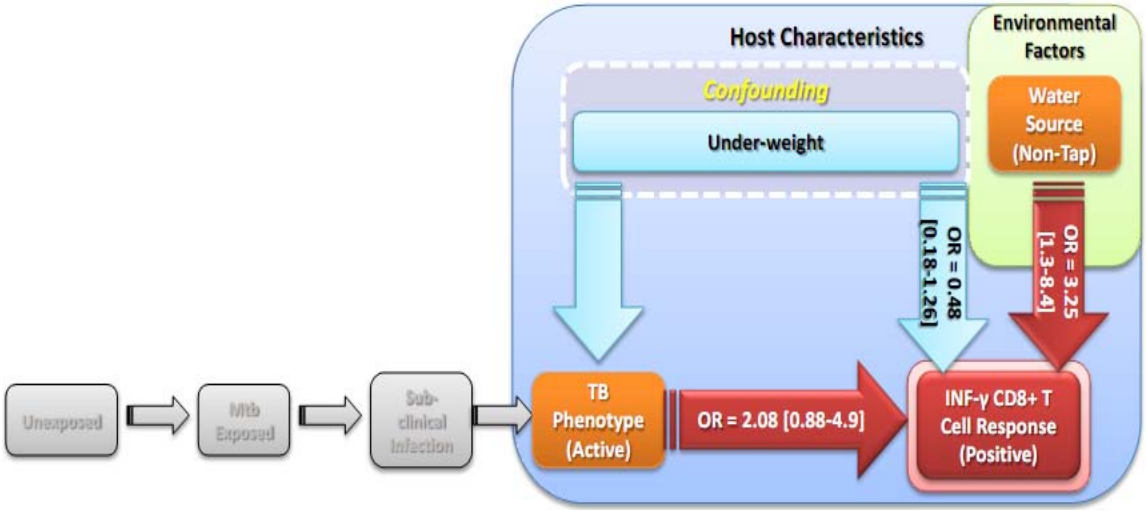


Figure 16 – Determinants of Positive INF-γ CD8+ T cell response in at Least One Antigen.

Consistently, such models suggested an increase in the odds of observing positive INF-γ CD8+ T cell response among ATB subjects. Active disease, as was described earlier, is characterized by high bacillary burden. Given this study's use of a CD8+ T cell model and antigen-screening strategies, it was hypothesized that an elevated INF-γ CD8+ T cell response would be observed among those subjects with active disease rather than those with the latent TB phenotype. This is, in fact, what was observed for four or five antigens in models utilizing continuous CD8+ T cell responses. However, no antigen exhibited a pronounced association with TB phenotype using the predefined cut-point, with differences in individual antigens' response rates—between ATB and LTBI subjects—of only 5% to 8%. As such, while this study provides support for a relationship between Mtb antigen-associated CD8+ T cell response and elevated bacillary burden, no ATB- or LTBI-specific antigens were identified.[xix] Moreover, the analyses showed that TB phenotype was only able to account for a small percentage of variability in INF-γ CD8+ T cell response, suggesting other factors shared responsibility for driving CD8+ T cell response rates.

Objective 2: Assessing the Role of Additional Clinical, Environmental, & Demographic Factors in Determining INF-γ CD8+ T Cell Response

As depicted in Figures 15 and 16, above, two of the eight factors investigated as potential confounders, effect modifiers, or independent determinants of INF-γ CD8+ T cell response were found to contribute to antigen-elicited CD8+ T cell response. First, being underweight (BMI < 18.5) accounted for a 25% reduction in the odds of observing a positive INF-γ CD8+ T cell response in ATB vs. LTBI subjects in unadjusted models—a finding that proved consistent across both antigen-specific and set-wide models. As was discussed in the Background section, the relationship between TB phenotype and low BMI is well established: malnourishment leads to immune compromise and the development of active disease while active disease can lead to wasting and malnourishment. The authors hypothesized that the relationship between INF-γ CD8+ T cell response and low BMI would be negative, secondary to immune suppression/dysregulation. Indeed, in this study low BMI was found to be positively associated with active disease and negatively associated with antigen-elicited INF-γ CD8+ T cell responses (though not at a statistically significant level, in the latter case.)

Analysis also indicated that the use of open water sources was associated with a significant increase in the odds of observing a positive/elevated INF-γ CD8+ T cell response. For example, the odds of responding positively to one or more antigens was 3.25 [95% CI:1.3-8.4] among those whose households obtained water from a spring or other open source vs. those from households utilizing tap water. This consistent finding was unaffected by subjects' TB phenotype and is hypothesized to result from increased exposure to environmental non-tuberculous Mycobacteria (NTM) which, though known to be present in both tap and non-tap water, may be more highly concentrated in open, untreated water sources that are likely to be frequently exposed to surface soils and other contaminants harboring endemic Mycobacteria. These elevated NTM exposure rates might then result in elevated antigen-elicited INF-γ CD8+ T cell response via to cross-reactivity. It could also be argued that within Kampala's urban context, not having

access to the city's water system may be associated with either living in an impoverished/underdeveloped neighborhood and/or direct exposure to infectious or other waterborne contaminants capable to producing poor health. Indeed, being underweight was weakly correlated with being water source (PCC = 0.20, p=0.038) and the former served as a negative confounder of the latter in multivariate models. However, the effect of water source was not significantly altered by adjustment for low BMI. It must be considered, then, that the strong association between INF-γ CD8+ T cell response and Water Source and lack of association between Water Source and TB phenotype introduces the possibility that an existing relationship between TB phenotype and INF-γ CD8+ T cell response is being obscured by a non-differential effect of antigen cross-reactivity.

Even in these models (that included TB phenotype, water source, and underweight status), R-squared values for models of continuous antigen response fell between 0.03 and 0.16 (and between 0.14 and 0.16 among only the four higher-performing antigens. Similarly, the area under ROC curves for binary CD8+ T cell response variables were found to range between 0.64 and 0.71 for individual antigens and was estimated to be 0.67 when examining zero vs. greater than or equal to one positive response among the five antigens. Thus, an additional unknown factor or combination of unknown factors account for a majority of variability in the outcome data.

Objective 3: Assessing Colinearity among Antigen-elicited INF-γ CD8+ T Cell Responses

Lastly, this study provided a means of characterizing the degree of heterogeneity among antigens with respect to INF-γ CD8+ T cell response and in relation to TB phenotype. The inter-antigen Pearson correlation coefficient matrix presented in the Results section indicated a high degree of colinearity among the five antigens. Even Antigen 25, a consistent outlier within the set, exhibited relatively high correlation (Pearson coefficients = 0.68-0.76) while all other antigen pairs had coefficients greater than 0.87 (including three pairs—2/13, 13/19, and 14/19—with coefficients greater than 0.90).

However, given the small number of antigens investigated in the study (relative to the number that might be included, for example, in a composite vaccine), the overall response rate can be seen as encouraging if it can be determined that response rates improve with the inclusion of additional antigens. Here, the average response rate for a single antigen was 30.2% and the inclusion of each additional antigen resulted in an average increase of 5.2% in sample capture. Thus, while the data available from the current study is limited, one can hypothesize that the inclusion of multiple antigens eliciting less highly correlated but still robust response patterns might eventually allow for complete sample capture. (It should be said, however, that without the inclusion of a TB negative control group in this study the false positive rate for the current set of antigens is unknown and, consequently, there is no way to precisely estimate an expected change in Mtb infection-associated response rate.)

STUDY LIMITATIONS AND VALIDITY ASSESSMENT:

This study was able to detect a significant positive association between INF-γ CD8+ T cell response and Active vs. Latent TB phenotype—in both crude and adjusted models—when CD8+ T cell response was examined as a continuous variable. However, this association was not significant when a predefined conservative threshold for CD8+ T cell response positivity was employed. In the setting of a small sample size (n=108), odds ratio estimates exhibited moderately wide 95% confidence intervals. Additionally, as with any non-randomized study there is concern that some factor(s) associated with the selection and/or classification processes might have artificially concentrated/diluted events around the putative exposures of interest. Regarding, first, the potential for selection bias, one must critically question whether the selection process in some way favored (1) the exclusion of subjects who were predisposed to generating high INF-γ CD8+ T cell response to specific Mtb antigens and happen to have active TB (and/or utilized a non-tap water source.) It should be noted, however, that the assessment of bias is, in this case, complicated by the novelty of the CD8+ T cell assay itself; as this is the first attempt to identify response predictors within such an assay, the characteristics

of a high or low response subject are not known. Nonetheless, there are certain elements of the subject selection process that should be highlighted as potential sources of bias. Recall, for example, that sample members were drawn from a cross sectional TB case-contact study in which new subjects were recruited through the identification of (1) active TB cases presenting at a large public Kampala hospital and (2) additional active cases and LBTI subjects located via follow-up with household members and close contacts. Thus, the assumption of complete independence is questionable as there are, in some cases, subject-pairs who are related to each other by place of residence and, perhaps, even blood. As these pairs would be biased towards mixed TB phenotype (i.e., one active TB case and a contact, usually LTBI) and would thus exert a non-differential bias with respect to this variable. Alternatively, one might expect that such a lack of independence—were it to link individuals predisposed to a particular pattern of INF-γ CD8+ T cell response —would exert a differential bias with respect to water source.

It should also be noted that participation was subject to explicit exclusion criteria, including a lower age limit (17 years of age) and negative HIV status. Additionally, the logistical requirements associated with the study's clinical follow-up procedures most likely generated a disincentive to participation among individuals outside of the greater Kampala metropolitan area which, as has been mentioned, is a large urban predominantly Gandan area.[xx] Moreover, incentive to participate in the study was provided in the form of free access to treatment, which may have differentially affected participation among members of different socioeconomic strata—a patient mix that may already be biased towards lower SES groups on account of population served by the public hospital. Each of these issues had implications for the generalizability of results and may also contribute to selection bias (though to extents and with directionalities that would be difficult to predict.)

A final issue, with respect to selection, was noted earlier regarding potential differences between (1) the small sample of TB-positive subjects in Oregon who were used to screen the Mtb antigen library (leading to the identification of the five antigens

tested here) and (2) the Ugandan sample used in the present study. For example, it was suggested that perhaps certain components of intracellular mycobacterial metabolism/reproduction—and subsequent host cell presentation/recognition of specific, associated antigens—might be stable in latent infection, increase near the LTBI-ATB transition point, and then fall off again during prolonged active disease. If the antigens tested in this study were part of such a dynamic, cellular response within the assay among those ATB individuals who only recently transitioned to active disease would be disproportionately responsive. Such individuals might then be disproportionately represented in the study sample—i.e., identified and subsequently enrolled because of a recent conversion. Under such a scenario, the original sample from Oregon might have been so constituted as to (unintentionally) include a mix of LTBI and ATB subjects that maximized differences between TB phenotype relative to the relevant characteristic—a distribution that is, in the Ugandan sample, reversed. A similar mechanism could introduce bias if the five test antigens' pattern of immunogenicity was, alternatively, affected by the recentness of first exposure/re-exposure, was elevated in the setting of TB reactivation relative to primary ATB, or some other scenario that included a disproportionate LTBI-ATB mix of subjects (relative to the relevant characteristic(s).)

In addition to subject selection, it is also important to consider the potential for systematic misclassification with respect to exposure/outcome. As in any study of this kind, the potential for information bias arises at several distinct points in the collection, processing, and management of data and materials. With respect to the attainment and management of the clinical, demographic, and household-level data, the use of trained clinical staff in the administration of standardized, validated data collection instruments with centralized, post-collection data deposition and management by the TBRU provides a measure of protection against bias arising from poor instrument design, variability in instrument administration, mis-coding, and data loss or corruption. In spite of these protections, the following features of the data development and management process need to be acknowledged as potential sources of information bias: (1) interviewers were

not blinded to subjects' case/contact status and the information that they reported on questionnaires was not subject to serial or duplicate validation; (2) participants themselves were not blinded to their own status as TB case/household contact, were required to report some information from memory, and were required to provide information through face-to-face interviews with staff; (3) subjects varied in terms of education, social status, age, gender, native language and other factors that could modify their understanding of questions and/or their ability/willingness to seek clarification or provide fully transparent information; (4) data recorded on paper forms were transcribed by TBRU staff into a digital database without secondary verification procedures; (5) the measurement, interpretation, and recording of clinical data (height, weight, PPD measurement, chest x-ray interpretation, sputum histology) were conducted in the context of clinical practice and are subject to the kinds and degrees of inter-observer variability and other sources of error common to such settings. In addition to issues of precision, discussed below, one might be concerned that the accuracy of reporting could be related to a patients knowledge/suspicion of a diagnosis. Even so, it is difficult to envision how rates of under/over reporting would be systematically related to INF-γ CD8+ T cell response when response rates were not associated with ATB-related symptoms, levels of contact with known ATB cases, or similar factors that might guide differential reporting.

Another potentially challenging multistep data development process involved the collection, transport, and analysis of subjects' blood samples. However, all samples demonstrated appropriate background activity and/or response to Mtb antigens (including PHA) by CD8+ ELISPOT assay and were analyzed by an automated digital ELISPOT reader with over-reading by blinded laboratory technicians in an order that was unrelated to TB phenotype. While this does not rule out the possibility of CD8+ T cell reactivity modification through storage/processing, it does suggest that any variability within these procedures was neither severe enough to lead to complete cell dysfunction nor would it be expected to affect a particular independent variable-associated subgroup disproportionately. [xxi] Thus, as with the potential sources of bias related to

clinical and demographic data enumerated above (excluding the first two, dealing with blinding[xxii]), any validity-related issues with these biological samples would be expected to introduce non-differential bias.

Additional factors to consider with respect to the study results' validity include variable specificity and residual confounding and interaction. As was discussed earlier, with respect to the former issue, the problematic use of a LTBI/ATB dichotomy to describe TB phenotype remains a concern that will only be addressed through the identification of appropriate biomarkers. Regarding the potential for residual confounding or interaction, there exist a number of potentially relevant factors for which data were either not collected as part of the KCHS, appeared incomplete in the KCHS dataset, or were not generated in subsequent laboratory analyses of subjects' serum samples. Among these were pathogen characteristics, serum solute levels (cytokines, iron, or Vitamin D), detailed exposure histories, host genetic factors, and other relevant neighborhood/workplace, individual, and/or household characteristics. While additional information regarding these concepts might be useful for further characterizing the complex, multifactorial character of TB-associated immunoreactivity, it is difficult to predict whether and in what ways the addition of this information would change the study outcomes.

Of those factors for which data was available, mean media INF-γ background response represents, perhaps, the most intriguing candidate for possible residual confounding. In spite of attempts to control for its effects via defining all INF-γ response outcome variables relative to its parameters, both univariate and multivariate models suggest that levels of high background immune reactivity are associated with a lower probability of INF-γ positive responses (Figure 18). One possible explanation is that the higher background response levels—which had larger standard deviations given that zero served as an absolute lower limit—were not accompanied by equally exaggerated CD8+ T cell response to Mtb antigens and that this mismatch resulted in an artificially elevated threshold for INF-γ response positivity. For example, one could envision that a large viral-associated clonal proliferation of reactive lymphocytes could coexist with a

stable population of Mtb-responsive lymphocytes whose absolute INF response to test antigens is appropriate and robust but, nonetheless, insufficient to surpass the threshold for positivity in the setting of a perturbed background signal. However, this hypothesis would not explain a similar relationship between TB phenotype and background response (Figure 17), which would have to emerge by way of an independent process. There are, then, a disproportionate number of LTBI subjects with high mean background signals and a corresponding group of subjects with high backgrounds and low positive INF-γ CD8+ T cell response rates. Modeling the relationship between TB phenotype and extreme mean background (defined as Background SFUs > 85) showed a 2.7 point estimate odds ratio favoring LTBI, but with a non-significant parameter estimate [95% CI: 0.68 – 10.9, p=0.157].

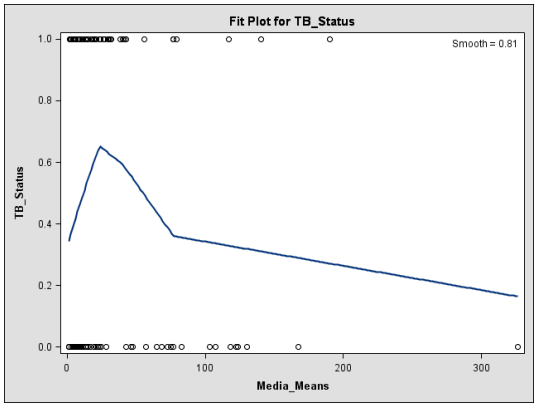


Figure 17 - TB Phenotype Probability (ATB = 1) vs. Mean Background INF-γ (via LOESS Method)

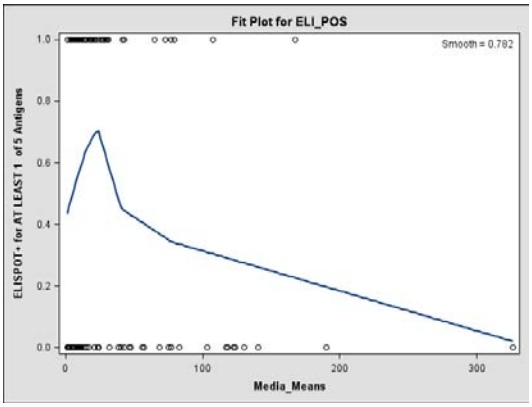


Figure 18 - Probability of ≥ 1 Positive Antigen Response vs. Mean Background INF-γ (via LOESS Method)

The linkage between high background CD8+ T cell reactivity and negative antigen-elicited INF-γ response is unclear but may, in fact, be related to another unusual finding: a significant minority of INF-γ CD8+ T cell response-negative subjects (n=22; 13 LTBI, 9 ATB) exhibited INF-γ response to one or more test antigens that fell at least two standard deviations *below* the mean background (see Figure 19 and Figure 20, below.) This subset was not only PHA test positive but was also associated with higher

background response levels (see Figure 21, below) and cannot, therefore, be easily explained by cell sample failure or by global CD8+ T cell anergy. Moreover the existence of some anomaly leading to cell failure in individual wells is possible but very unlikely given the high rate of multiple extreme low SFU values seen across several antigens in the same individuals. Another possibility is that there are a subset of TB-positive individuals who develop an Mtb antigen-specific anergy. This would explain the differing rates of extreme low values among the five test antigens (with Antigen 25 being a market outlier in Figure 19.) While this latter notion is an intriguing one, it remains highly speculative without additional study.

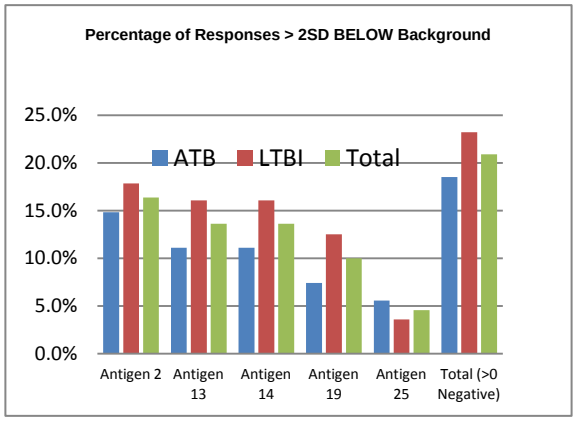


Figure 19 - Percentage of INF-y CD8 T Cell Responses ≥ Two Stnd. Deviations below Mean Background (by Antigen)

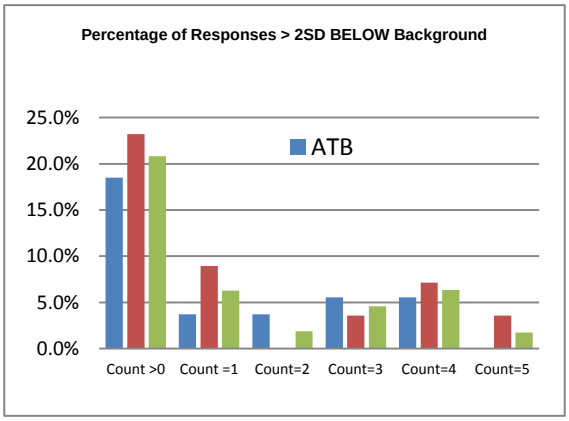


Figure 20 - Percentage of INF-y CD8 T Cell Responses ≥ Two Stnd. Deviations below Mean Background (by No. of Antigens Meeting Criteria)

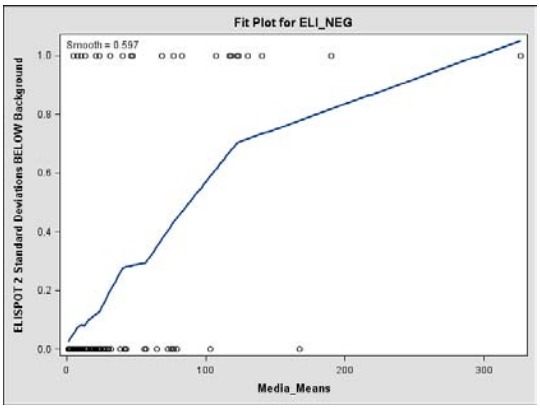


Figure 21 - Probability of ≥ 1 Antigen Response at Least Two Standard Deviations Below Background vs. Mean Background INF-γ (via LOESS Method)

Among the remaining components of validity assessment, biologic plausibility and, where appropriate, dose response relationships relevant to the proposed models were discussed, above. Finally, whereas the external validity of the results presented here is not directly assessable at the present time (given the lack of external data on LTBI-associated, CD8+ T cell-activating Mtb antigens), evidence for consistency of association within the study was provided by parallel analyses of individual antigens.

CONCLUSIONS & FUTURE DIRECTIONS:

The study's results allow for a preliminary assessment of the five novel antigen candidates in relation to both their potential diagnostic value and utility with respect to vaccine development efforts. **First, the ability to discriminate between higher vs. lower bacillary burden states based on antigen-elicited INF-γ CD8+ T cell responses was established for four of five novel CD8+ T cell-associated antigens**. Specifically, using Active TB and Latent Mtb Infection status as a proxy for Mtb burden, continuous antigen-elicited INF-γ CD8+ T cell response was found to be associated with Mtb burden in a linear, positive fashion for Antigens 2, 13, 14, and 19. This association was also observed, though not at a statistically significant level ($p = 0.08$), when a high predefined threshold for positive vs. negative response was employed.

Second, large pluralities of both ATB and LTBI subjects exhibited robust INF-γ CD8+ T cell responses, such that no antigen could be described as strictly ATB- or LTBI-specific. This was found to be true when considering both response rates to individual antigen and the performance of the five-antigen set, as a whole. Antigen set-wide data, for example, showed that 58% of ATB and 46% of LTBI subjects were found to respond positively (i.e., based on a conservative predefined threshold) to at least one antigen and 25% of LTBI and 37% of ATB subjects were found to be "high responders," demonstrating positive responses to at least four antigens.

Third, INF-γ CD8+ T cell response was found to be affected by two factors other than TB phenotype: being malnourished (BMI < 18.5) and exposure to non-tuberculous Mycobacterium species (via use of open water sources.) Specifically, low BMI was found to function as a negative confounder of the relationship between TB phenotype and INF-γ CD8+ T cell response such that failure to adjust for subjects' underweight status resulted in a 25% reduction in the odds of observing a positive response among ATB vs. LTBI subjects. Water source, however, served as an independent determinant of INF-γ CD8+ T cell response, with subjects using open water sources, such as well and springs, demonstrating a 3.6 increase in the odds of a positive response to one or more Mtb antigen.

Fourth, the models presented here were only able to account for a small percentage of variability in INF-γ CD8+ T cell response, with R-squared values for multivariate models of continuous antigen response ranging between 0.14 and 0.16 (among the four higher-performing antigens) while areas under ROC curves for binary CD8+ T cell response variables fell between 0.64 and 0.71 for individual antigens and set-wide models. Moreover, by itself, TB phenotype accounted for between only 4% and 8% of continuous individual antigen response variability. Together, these first four findings suggest that the association between TB phenotype and INF-γ CD8+ T cell response is either relatively weak (especially in the setting of relatively low power provided by the sample size and when using a very rigorous screening threshold),

obscured by the non-differential effects of other factors, or masked by residual/unmeasured confounding.

Fifth, although antigen-specific responses were highly correlated (Pearson coefficients: 0.68 – 0.95), for each addition antigen included in the screen, there was an average observed increase in sample-capture of 5.2% (above the mean single-antigen response rate of 30.2%.) This would at least suggest the possibility of approaching complete sample-capture through the inclusion of a sufficnetly large and diverse number of antigens, particularly if a less highly correlated subset could be identified.

Several possible dynamics could contribute to these findings. First, these antigenic protein fragments might be recognized because they represent antigenic materials previously presented to CD8+ T cells by APCs in the setting of subjects' Mtb infections. This is the paper's working hypothesis and is supported by the linear relationship between continuous INF-γ CD8+ T cell response and TB phenotype. However, cross-reactivity involving CD8+ T cells sensitized to NTM may also being contributing to INF-γ response, as suggested by the strong association between water source and CD8 T cell response. Given the independence of effects for both TB phenotype and water source found in this study, together with the low levels of variation in response that the models were able to account for, it seems likely that multiple drivers of INF-γ CD8+ T cell response contribute simultaneously to observed patterns of INF-γ CD8+ T cell response. To the extent that TB phenotype contributes to this, and given the moderate heterogeneity in inter-antigen response, we might anticipate that the inclusion of additional antigens may improve the overall response rate. Unfortunately, without access to data from an uninfected control group it is impossible to accurately determine the relative contributions of Mtb infection and other potential drivers of CD8+ T cell response. Thus, at a minimum, these results suggest that additional research is needed to more precisely characterize the immunogenicity—both in magnitude and specificity—of the five antigens included in this study, as well as a

number of additional, more recently-identified immnodominant CD8+ T cell-associated antigens and epitopes.

More generally, the potential for various host and environmental factors in determining IFN-γ CD8+ T cell response presents a particular problem with implications for not only future antigen assessment efforts but for our understanding of TB immunology more generally. It might be, for example, that exposure/re-exposure patterns within TB-endemic regions, particular forms of compromised health, certain environmental exposures, and/or other factors not fully accounted for in the current analysis lead to enhanced CD8+ T cell reactivity among those with active TB with respect to antigens that, in other contexts, are preferentially expressed in latent infections. In this way, the presence of multiple confounders, interactions, and non-TB-specific CD8+ T cell response determinants might partially obscure our view of a universally optimal antigen-set. On the other hand, it might be that the composition of an "optimal" antigen set is, to some extent, dependent upon the particular biosocial contexts of a given population (on account of immunomodulatory effects of various environmental, behavioral and health factors. The former possibility would demand additional, larger studies to render the TB phenotype/CD8+ T cell response relationship more legible. The latter would require the study of multiple populations and a more inclusive approach to final antigen selection.

Thus, while the results of the current study represent an important first step in the assessment of recent efforts to identify TB phenotype-specific antigens, they also point to the need for significant additional research. Given the small, relatively homogenous sample without Mtb-negative controls, single point timeframe, and limited set of test antigens employed in this study, confirmation or refutation of its results should be undertaken in a study utilizing a larger, controlled, diverse sample that is followed over time and employs a larger set of novel LTBI-associated antigens from among the expanding candidate library. Doing so would better enable researchers to discern (1) whether there exist significant differences—with respect to the INF-γ CD8+ T cell response—between certain sub-groups (based on, for example, HLA type,

health/nutritional status, patterns of Mtb and NTM and other environmental exposures); (2) whether these represent phenomena that are (a) stable across different populations, (b) more or less specific to particular sets of antigens, and/or (c) from the standpoint of further antigen testing or diagnostics development, requires modification of cut-off points, either globally or for certain populations; and (3) whether antigen-elicited CD8+ T cell response varying along the natural course of Mtb infection in such a way that particular antigen subgroups are found to be associated with specific regions along the TB phenotype spectrum.

Regarding the first two points, there may be some utility to assessing INF-γ CD8+ T cell response in relation to factors such as: specific Mtb strains or variants of interest, serum levels of immunity-related solutes (cytokines, Mtb proteins, iron, or Vitamin D), measures of general inflammation (e.g., ESR), more detailed exposure histories, host genetic factors, host health status (e.g., immune-compromised states, HIV+ individuals, pediatric populations) or even specific neighborhood/workplace characteristics and/or measures of SES and psychosocial stress. This is particularly important given the goal of identifying candidates for a composite vaccine in a setting where the target population is extremely diverse and the financial and biologic economies governing vaccine production favor the use of fewer, immunologically more potent antigen/epitopes over more inclusive but less specific shotgun approaches to component identification.

With respect to the third point, the simultaneous assessment of multiple antigens/epitopes and a range of immuomodulating host characteristics may help to better characterize the TB Phenotype Continuum, perhaps even assisting in the identification of important tipping points—and associated biomarkers—within the disease life-course that could lead to not only more effective vaccine design but also clinical tests capable of identifying those at greatest risk for transition to active disease so that treatment could be targeted in an efficient, precise fashion. Longitudinal studies that follow INF-γ CD8+ T cell response over time would be extremely helpful in characterizing infection-associated global changes in response rates (as well as distinguishing these trends from normal background stochastic variability) and,

potentially, would provide a mechanism for mapping individual antigens to specific sites along the course of disease. Such serial testing would also allow for the monitoring of response during clinical therapeutics and vaccine trials.

WORKS CITED

1. **WHO.** Global Tuberculosis Control: Surveillance, Planning, Financing. *Global Tuberculosis Control: Surveillance, Planning, Financing.* Geneva : s.n., 2007. Document no. WHO/HTM/TB/2006.362.

2. *The past and future of tuberculosis research.* **Comas, Iñaki and Gagneux, Sebastien.** 10, Oct 2009, PLoS Pathog, Vol. 5, p. e1000600. DOI: 10.1371/journal.ppat.1000600.

3. *Systems biology of persistent infection: tuberculosis as a case study.* **Young, Douglas, Stark, Jaroslav and Kirschner, Denise.** 7, Jul 2008, Nat Rev Microbiol, Vol. 6, pp. 520-528. DOI: 10.1038/nrmicro1919.

4. *Tuberculosis: what we don't know can, and does, hurt us.* **Russell, David G., Barry,, Clifton E and Flynn, JoAnne L.** 5980, May 2010, Science, Vol. 328, pp. 852-856. DOI: 10.1126/science.1184784.

5. *The population dynamics and control of tuberculosis.* **Dye, Christopher and Williams, Brian G.** 5980, May 2010, Science, Vol. 328, pp. 856-861. DOI: 10.1126/science.1185449.

6. *New tools and emerging technologies for the diagnosis of tuberculosis: part I. Latent tuberculosis.* **Pai, Madhukar, Kalantri, Shriprakash and Dheda, Keertan.** 3, May 2006, Expert Rev Mol Diagn, Vol. 6, pp. 413-422. DOI: 10.1586/14737159.6.3.413.

7. *Confronting the scientific obstacles to global control of tuberculosis.* **Young, Douglas B., et al.** 4, Apr 2008, J Clin Invest, Vol. 118, pp. 1255-1265. DOI: 10.1172/JCI34614.

8. *Tuberculosis in Africa: learning from pathogenesis for biomarker identification.* **Kaufmann, S H and Parida, S K.** 3, Sep 2008, Cell Host Microbe, Vol. 4, pp. 219-228. DOI: 10.1016/j.chom.2008.08.002.

9. *Vaccine strategies against latent tuberculosis infection.* **Andersen, Peter.** 1, Jan 2007, Trends Microbiol, Vol. 15, pp. 7-13. DOI: 10.1016/j.tim.2006.11.008.

10. *Major histocompatibility complex class I-restricted T cells are required for resistance to Mycobacterium tuberculosis infection.* **Flynn, J. L., et al.** 24, Dec 1992, Proc Natl Acad Sci U S A, Vol. 89, pp. 12013-12017.

11. *Human alveolar T lymphocyte responses to Mycobacterium tuberculosis antigens: role for CD4+ and CD8+ cytotoxic T cells and relative resistance of alveolar macrophages to lysis.* **Tan, J. S., et al.** 1, Jul 1997, J Immunol, Vol. 159, pp. 290-297.

12. *Human cytolytic and interferon gamma-secreting CD8+ T lymphocytes specific for Mycobacterium tuberculosis.* **Lalvani, A., et al.** 1, Jan 1998, Proc Natl Acad Sci U S A, Vol. 95, pp. 270-275.

13. *Characterization of human CD8+ T cells reactive with Mycobacterium tuberculosis-infected antigen-presenting cells.* **Lewinsohn, D. M., et al.** 10, May 1998, J Exp Med, Vol. 187, pp. 1633-1640.

14. *Antimicrobial activity of MHC class I-restricted CD8+ T cells in human tuberculosis.* **Cho, S., et al.** 22, Oct 2000, Proc Natl Acad Sci U S A, Vol. 97, pp. 12210-12215. DOI: 10.1073/pnas.210391497.

15. *Mycobacterium tuberculosis-specific CD8+ T cells preferentially recognize heavily infected cells.* **Lewinsohn, Deborah A., et al.** 11, Dec 2003, Am J Respir Crit Care Med, Vol. 168, pp. 1346-1352. DOI: 10.1164/rccm.200306-837OC.

16. *Phenotypical and functional analysis of memory and effector human CD8 T cells specific for mycobacterial antigens.* **Caccamo, Nadia, et al.** 3, Aug 2006, J Immunol, Vol. 177, pp. 1780-1785.

17. *CD8+ T cells provide an immunologic signature of tuberculosis in young children.* **Lancioni, Christina, et al.** 2, Jan 2012, Am J Respir Crit Care Med, Vol. 185, pp. 206-212.

18. *Comparison of T-cell interferon-gamma release assays for Mycobacterium tuberculosis-specific antigens in patients with active and latent tuberculosis.* **Kobashi,**

Yoshihiro, et al. 4, Aug 2010, Lung, Vol. 188, pp. 283-287. DOI: 10.1007/s00408-010-9238-3.

19. *Interferon-gamma release assay test characteristics depend upon the prevalence of active tuberculosis.* **Davidow, A. L.** 11, Nov 2009, Int J Tuberc Lung Dis, Vol. 13, pp. 1411-1415.

20. *Infectious disease. Extensively drug-resistant TB gets foothold in South Africa.* **Cohen, Jon.** 5793, Sep 2006, Science, Vol. 313, p. 1554. DOI: 10.1126/science.313.5793.1554a.

21. *Extensively drug-resistant tuberculosis: 2 years of surveillance in Iran.* **Masjedi, Mohamad Reza, et al.** 7, Oct 2006, Clin Infect Dis, Vol. 43, pp. 841-847. DOI: 10.1086/507542.

22. *The Global Plan to STOP TB, 2006-2015.* 2006. **The Stop TB Partnership.** (Accessed January 2012).

23. *The Stop TB Strategy: Building on and enhancing DOTS to meet the TB-related Millennium Development Goals.* ***WHO*** 2006.

24. *The Global Plan to STOP TB, 2011-2015.* 2011. **The Stop TB Partnership** (Accessed January 2012).

25. *Roadmap For MDR-TB Scale Up: The Global Drug Facility (GDF).* **The Stop TB Partnership** 2010.

26. *The Global Plan to STOP TB, 2006-2015: Progress Report 2006-2008.* 2009. **The Stop TB Partnership** (Accessed January 2012).

27. *The spectrum of latent tuberculosis: rethinking the biology and intervention strategies.* **Barry,, Clifton E, et al.** 12, Dec 2009, Nat Rev Microbiol, Vol. 7, pp. 845-855. DOI: 10.1038/nrmicro2236.

28. *Eliminating latent tuberculosis.* **Young, Douglas B., Gideon, Hannah P. and Wilkinson, Robert J.** 5, May 2009, Trends Microbiol, Vol. 17, pp. 183-188. DOI: 10.1016/j.tim.2009.02.005.

29. *The equilibria that allow bacterial persistence in human hosts.* **Blaser, Martin J. and Kirschner, Denise.** 7164, Oct 2007, Nature, Vol. 449, pp. 843-849. DOI: 10.1038/nature06198.

30. *Interaction of Mycobacterium tuberculosis with the host: consequences for vaccine development.* **Dietrich, Jes and Doherty, T Mark.** 5-6, May 2009, APMIS, Vol. 117, pp. 440-457. DOI: 10.1111/j.1600-0463.2009.02458.x.

31. **Gagneux, S.** Mycobacterium: Genomics and Molecular Biology. [ed.] Paris T and Brown A. s.l. : Caister Academic Press: Norfolk, UK, 2009, Strain Variation and Evolution.

32. *Mycobacterium tuberculosis Uganda genotype is the predominant cause of TB in Kampala, Uganda.* **Asiimwe, B. B., et al.** 4, Apr 2008, Int J Tuberc Lung Dis, Vol. 12, pp. 386-391.

33. *Global phylogeography of Mycobacterium tuberculosis and implications for tuberculosis product development.* **Gagneux, Sebastien and Small, Peter M.** 5, May 2007, Lancet Infect Dis, Vol. 7, pp. 328-337. DOI: 10.1016/S1473-3099(07)70108-1.

34. **Casali, N.** Mycobacterium: Genomics and Molecular Biology. [ed.] Paris T and Brown A. s.l. : Caister Academic Press: Norfolk, UK, 2009, Hypervirulent Mycobacterium tuberculosis.

35. *Current findings, challenges and novel approaches in human genetic susceptibility to tuberculosis.* **Möller, Marlo and Hoal, Eileen G.** 2, Mar 2010, Tuberculosis (Edinb), Vol. 90, pp. 71-83. DOI: 10.1016/j.tube.2010.02.002.

36. *Genome scan of M. tuberculosis infection and disease in Ugandans.* **Stein, Catherine M., et al.** 12, 2008, PLoS One, Vol. 3, p. e4094. DOI: 10.1371/journal.pone.0004094.

37. *The relationship between malnutrition and tuberculosis: evidence from studies in humans and experimental animals.* **Cegielski, J. P. and McMurray, D. N.** 3, Mar 2004, Int J Tuberc Lung Dis, Vol. 8, pp. 286-298.

38. *Tuberculosis and host metabolism: ancient associations, fresh insights.* **Blumenthal, Antje, Isovski, Flonza and Rhee, Kyu Y.** 1, Jul 2009, Transl Res, Vol. 154, pp. 7-14. DOI: 10.1016/j.trsl.2009.04.004.

39. *The effect of the host's iron status on tuberculosis.* **Boelaert, Johan R., et al.** 12, Jun 2007, J Infect Dis, Vol. 195, pp. 1745-1753. DOI: 10.1086/518040.

40. **Reider, H.** *Epidemiologic Basis of Tuberculosis Control, First Edition.* International Union Again Tuberculosis and Lung Disease. 1999. Tech. rep.

41. *Does the risk of tuberculosis increase in old age?* **Stead, W. W. and Lofgren, J. P.** 5, May 1983, J Infect Dis, Vol. 147, pp. 951-955.

42. *Risk of tuberculosis from exposure to tobacco smoke: a systematic review and meta-analysis.* **Bates, Michael N., et al.** 4, Feb 2007, Arch Intern Med, Vol. 167, pp. 335-342. DOI: 10.1001/archinte.167.4.335.

43. *Tobacco and tuberculosis: a qualitative systematic review and meta-analysis.* **Slama, K., et al.** 10, Oct 2007, Int J Tuberc Lung Dis, Vol. 11, pp. 1049-1061.

44. *Passive smoking and tuberculosis.* **Leung, Chi C., et al.** 3, Feb 2010, Arch Intern Med, Vol. 170, pp. 287-292. DOI: 10.1001/archinternmed.2009.506.

45. *Alcohol use as a risk factor for tuberculosis - a systematic review.* **Lönnroth, Knut, et al.** 2008, BMC Public Health, Vol. 8, p. 289. DOI: 10.1186/1471-2458-8-289.

46. *Drivers of tuberculosis epidemics: the role of risk factors and social determinants.* **Lönnroth, Knut, et al.** 12, Jun 2009, Soc Sci Med, Vol. 68, pp. 2240-2246. DOI: 10.1016/j.socscimed.2009.03.041.

47. *Factors affecting time delay to treatment in a tuberculosis control programme in a sub-Saharan African country: the experience of The Gambia.* **Lienhardt, C., et al.** 3, Mar 2001, Int J Tuberc Lung Dis, Vol. 5, pp. 233-239.

48. *Investigation of the risk factors for tuberculosis: a case-control study in three countries in West Africa.* **Lienhardt, C., et al.** 4, Aug 2005, Int J Epidemiol, Vol. 34, pp. 914-923. DOI: 10.1093/ije/dyi100.

49. *Tuberculosis in household contacts of infectious cases in Kampala, Uganda.* **Guwatudde, D., et al.** 9, Nov 2003, Am J Epidemiol, Vol. 158, pp. 887-898.

50. *Burden of tuberculosis in Kampala, Uganda.* **Guwatudde, David, et al.** 11, 2003, Bull World Health Organ, Vol. 81, pp. 799-805.

51. *Huge variation in Russian mortality rates 1984-94: artefact, alcohol, or what?* **Leon, D. A., et al.** 9075, Aug 1997, Lancet, Vol. 350, pp. 383-388. DOI: 10.1016/S0140-6736(97)03360-6.

52. *Economic Change, Crime, and Mortality Crisis in Russia: Regional Analysis.* **Walberg, Peder, et al.** 7154, s.l. : BMJ Publishing Group, 1998, BMJ: British Medical Journal, Vol. 317, pp. pp. 312-318. ISSN: 09598138.

53. *Long-term efficacy of BCG vaccine in American Indians and Alaska Natives: A 60-year follow-up study.* **Aronson, Naomi E., et al.** 17, May 2004, JAMA, Vol. 291, pp. 2086-2091. DOI: 10.1001/jama.291.17.2086.

54. *Efficacy of BCG vaccine in the prevention of tuberculosis. Meta-analysis of the published literature.* **Colditz, G. A., et al.** 9, Mar 1994, JAMA, Vol. 271, pp. 698-702.

55. *BCG: the challenge continues.* **Fine, P. E.** 4, 2001, Scand J Infect Dis, Vol. 33, pp. 243-245.

56. *TB vaccines: current status and future perspectives.* **Aagaard, Claus, et al.** 4, 2009, Immunol Cell Biol, Vol. 87, pp. 279-286. DOI: 10.1038/icb.2009.14.

57. **Aeras.** Vaccine Development Process. *Vaccine Development Process.* 1405 Research Blvd., Rockville, MD 20850, USA : s.n.

58. *Human genetics of intracellular infectious diseases: molecular and cellular immunity against mycobacteria and salmonellae.* **van de Vosse, E, Hoeve M.A., and Ottenhoff T.H.M.** 12, Dec 2004, Lancet Infect Dis, Vol. 4, pp. 739-749. DOI: 10.1016/S1473-3099(04)01203-4.

59. *Role of CD8+ T lymphocytes in control of Mycobacterium tuberculosis infection.* **Grotzke, Jeff E. and Lewinsohn, David M.** 4, Apr 2005, Microbes Infect, Vol. 7, pp. 776-788. DOI: 10.1016/j.micinf.2005.03.001.

60. *Tuberculosis vaccine development: goals, immunological design, and evaluation.* **Hoft, Daniel F.** 9633, Jul 2008, Lancet, Vol. 372, pp. 164-175. DOI: 10.1016/S0140-6736(08)61036-3.

61. *Secreted proteins from Mycobacterium tuberculosis gain access to the cytosolic MHC class-I antigen-processing pathway.* **Lewinsohn, David M., et al.** 1, Jul 2006, J Immunol, Vol. 177, pp. 437-442.

62. *Hypoxia: a window into Mycobacterium tuberculosis latency.* **Rustad, Tige R., et al.** 8, Aug 2009, Cell Microbiol, Vol. 11, pp. 1151-1159. DOI: 10.1111/j.1462-5822.2009.01325.x.

63. *T-cell assays for the diagnosis of latent tuberculosis infection: moving the research agenda forward.* **Pai, Madhukar, et al.** 6, Jun 2007, Lancet Infect Dis, Vol. 7, pp. 428-438. DOI: 10.1016/S1473-3099(07)70086-5.

64. *Systematic review: T-cell-based assays for the diagnosis of latent tuberculosis infection: an update.* **Pai, Madhukar, Zwerling, Alice and Menzies, Dick.** 3, Aug 2008, Ann Intern Med, Vol. 149, pp. 177-184.

65. *T-cell epitope mapping in Mycobacterium tuberculosis using pepmixes created by micro-scale SPOT- synthesis.* **Frieder, Marisa and Lewinsohn, David M.** 2009, Methods Mol Biol, Vol. 524, pp. 369-382. DOI: 10.1007/978-1-59745-450-6_27.

66. *Immune correlates of acute Mycobacterium tuberculosis infection in household contacts in Kampala, Uganda.* **Whalen, Christopher C., et al.** 1, Jul 2006, Am J Trop Med Hyg, Vol. 75, pp. 55-61.

67. *Replacing the tuberculin skin test with a specific blood test.* **Weldingh, Karin and Andersen, Peter.** 8, Aug 2005, Kekkaku, Vol. 80, pp. 581-585.

68. *The measurement of observer agreement for categorical data.* **Landis, J. R. and Koch, G. G.** 1, Mar 1977, Biometrics, Vol. 33, pp. 159-174.

69. *Detection of interleukin-2 in addition to interferon-gamma discriminates active tuberculosis patients, latently infected individuals, and controls.* **Biselli, R., et al.** 8, Aug 2010, Clin Microbiol Infect, Vol. 16, pp. 1282-1284. DOI: 10.1111/j.1469-0691.2009.03104.x.

70. **Hosmer, DW and Lemeshow S**. *Applied Logistic Regression, Second Edition*. John Wiley & Sons. 2000

APPENDIX:

Table 10 - Antigen-specific CD8+ T Cell Response: Univariate Statistics for Key Independent Variables

	Sample Characteristics by INF-γ CD8+ T Cell Response		Contin. Outcome = (Mean Antigen-associated SFU) - (Background SFU) *			Binary Outcome = (Mean Antigen-associated SFU Count) > (2*SD of Background SFU +10)	
Antigen	**Variable**	**Reference Value [units]**	**β**	**p-value**	**R^2**	**Odds Ratio**	**p-value**
Antigen 2	Sex	Female vs. Male	5.5	0.628	0.002	1.53	0.304
	Age	Years Old (continuous)	0.6	0.189	0.017	1.00	0.887
	TB Phenotype	Active vs. Latent TB	21.5	0.052	0.036	1.32	0.493
	Sputum Smear (ATB only)‡	Positive vs. Negative	0.5	0.986	0.000	1.40	0.747
	Cardinal Symptoms (ATB only)‡	0-7 Symptoms	-8.3	0.166	0.038	0.75	0.477
	Extent Lung Ds. (ATB only)‡	Normal/Min/Mod/Adv.	-15.5	0.107	0.051	0.74	0.399
	Underweight	BMI ≤ 18.5 vs. > 18.5	-16.6	0.180	0.017	0.46	0.114
	TB+ Contact Exposure (LTBI only)‡	[Hrs/ month]	0.6	0.186	0.034	1.02	0.280
	Watersource	Spring/Other vs. Tap	37.5	0.002	0.086	3.55	0.005
Antigen 13	Sex	Female vs. Male	1.3	0.876	0.000	1.60	0.287
	Age	Years Old (continuous)	0.5	0.112	0.024	1.01	0.687
	TB Phenotype	Active vs. Latent TB	20.7	0.011	0.061	1.46	0.378
	Sputum Smear (ATB only)‡	Positive vs. Negative	5.5	0.801	0.001	1.40	0.747
	Cardinal Symptoms (ATB only)‡	0-7 Symptoms	-1.8	0.679	0.003	0.75	0.477
	Extent Lung Ds. (ATB only)‡	Normal/Min/Mod/Adv.	-9.0	0.189	0.034	0.74	0.399
	Underweight	BMI ≤ 18.5 vs. > 18.5	-10.3	0.259	0.012	0.53	0.219
	TB+ Contact Exposure (LTBI only)‡	[Hrs/ month]	0.4	0.210	0.031	1.03	0.258
	Watersource	Spring/Other vs. Tap	25.9	0.004	0.075	3.14	0.012
Antigen 14	Sex	Female vs. Male	-6.0	0.439	0.006	0.89	0.788
	Age	Years Old (continuous)	0.3	0.260	0.012	0.99	0.563
	TB Phenotype	Active vs. Latent TB	22.5	0.003	0.083	1.45	0.378
	Sputum Smear (ATB only)‡	Positive vs. Negative	5.0	0.783	0.002	1.40	0.747
	Cardinal Symptoms (ATB only)‡	0-7 Symptoms	-2.4	0.496	0.009	0.75	0.477
	Extent Lung Ds. (ATB only)‡	Normal/Min/Mod/Adv.	-4.6	0.427	0.013	0.74	0.399
	Underweight	BMI ≤ 18.5 vs. > 18.5	-9.3	0.274	0.011	0.61	0.315
	TB+ Contact Exposure (LTBI only)‡	[Hrs/ month]	0.3	0.361	0.016	1.00	0.879
	Watersource	Spring/Other vs. Tap	20.0	0.018	0.053	1.81	0.189
Antigen 19	Sex	Female vs. Male	-7.3	0.375	0.008	1.77	0.183
	Age	Years Old (continuous)	0.8	0.007	0.067	1.03	0.100
	TB Phenotype	Active vs. Latent TB	20.4	0.011	0.061	1.22	0.637
	Sputum Smear (ATB only)‡	Positive vs. Negative	13.7	0.533	0.008	1.40	0.747
	Cardinal Symptoms (ATB only)‡	0-7 Symptoms	-1.2	0.776	0.002	0.75	0.477
	Extent Lung Ds. (ATB only)‡	Normal/Min/Mod/Adv.	-1.5	0.831	0.001	0.74	0.399
	Underweight	BMI ≤ 18.5 vs. > 18.5	-12.0	0.181	0.017	0.42	0.093
	TB+ Contact Exposure (LTBI only)‡	[Hrs/ month]	0.3	0.452	0.011	1.01	0.686
	Watersource	Spring/Other vs. Tap	21.3	0.018	0.053	3.55	0.005
Antigen 25	Sex	Female vs. Male	7.9	0.368	0.008	1.13	0.786
	Age	Years Old (continuous)	0.6	0.056	0.035	1.02	0.140
	TB Phenotype	Active vs. Latent TB	2.1	0.813	0.001	0.83	0.676
	Sputum Smear (ATB only)‡	Positive vs. Negative	17.5	0.449	0.012	1.00	1.000
	Cardinal Symptoms (ATB only)‡	0-7 Symptoms	-1.4	0.762	0.002	0.80	0.646
	Extent Lung Ds. (ATB only)‡	Normal/Min/Mod/Adv.	1.8	0.809	0.001	1.15	0.702
	Underweight	BMI ≤ 18.5 vs. > 18.5	-11.8	0.223	0.014	0.60	0.322
	TB+ Contact Exposure (LTBI only)‡	[Hrs/ month]	0.4	0.232	0.028	1.00	0.782
	Watersource	Spring/Other vs. Tap	11.8	0.227	0.014	2.39	0.059

‡ TB phenotype-specific variables; analyses were subset for the appropriate ATB/LTBI status.

* Two outlier observations with large-magnitude INF-γ CD8+ T cell responses were excluded to ensure model linearity.

ENDNOTES:

[i] Some have reported that the combination of CD4+ T cell-based assays and cytokine levels (e.g., IL-2) provide a mechanism for distinguishing ATB and LTBI (e.g., see Biselli et al., 2010 (69))

[ii] Including historical patterns of TB-associated mortality in Europe during the first half of the 20th century that appear to relate directly to food shortages. (40)

[iii] This cross-sectional household contact study was one that directly preceded and informed the subsequent Kawempe Community Health Study from which the current investigation's subjects' were derived.

[iv] Particularly in the setting of sub-optimal crude assay specificity it will be useful to determine the degree of inter-antigen agreement. In other words, with each antigen eliciting positive INF-γ CD8+ T cell responses among a minority of subjects, the degree to which each is identifying the same subset of subjects (and, conversely, consistently ignoring another) will provide some evidence of the potential for increasing collective response rates and, potentially, vaccine-induced immunity (1) through the accumulation of additional antigens identified through the same selection process or (2) whether capturing a distinct, non-responsive subset might require the development of a separate parallel selection process based on other selection criteria.

[v] Weighting utilized a complex scoring system utilizing proteonomic data, Mtb transcriptional data from macrophage models, BCG-non-inclusion, and multiple functional attributes. In addition, 34 genes were selected based on a priori interest (including 10 with prior evidence of CD8+ T cell recognition, 9 known CD4+ antigens, and15 ESAT-like peptides.) The complete criteria are outlined elsewhere (65)

[vi] A paper, currently under preparation by David Lewinsohn and colleagues, will provide a detailed description of this antigen discovery process and its results, to date.

[vii] Seventy-two TB+ individuals were screened but only twenty went on to be leukophoresed. This set of donors included five ATB and fifteen LTBI individuals, with the latter group evenly divided into thirds by ancestry (Eastasian, African, European.) Additionally, each individual was required to meet the following three criteria prior to aphoresis: 1) negative for infection with HIV, HBV, and HCV; 2) low ELISPOT background INF-γ levels, and 3) $CD4^+$ T cells demonstrated a positive response to either CFP10 or

Esat-6 peptide pools (as these antigens are absent from BCG and thus may be used to distinguish true Mtb infection from exposure to BCG or atypical mycobacteria.)

[viii] This approach offers several advantages that render it useful across a range of TB-related research efforts. First, the ELISPOT assay can be tailored—through the inclusion of specific antigens—to avoid cross-reactivity with prior bacillus Calmette-Guerin (BCG) vaccination (unlike, for example, the tuberculin skin test (TST).) Second, while some variability exists across patient subgroups, contemporary T cell-based ELISPOT assays appear to provide equivalent-to-improved sensitivity relative to TST and, arguably, even other IGRAs. Moreover, as is argued below, there is reason to believe that CD8+ T cell-based ELISPOT assays may provide significantly superior specificity. Third, the ability to vary the kinds and combinations of antigens used in the test may facilitate the identification of different subgroups of responders and improve the sensitivity and/or specificity of subsequent tests. Fourth, by providing a more precise means of assessing immune response, this technology has the potential to develop into a powerful prognostic tool. Finally, ELISPOT TB assays have tremendous potential as a mechanism for both identifying components for novel combined vaccine therapy and for testing promising vaccine and drug candidates.

[ix] This stopping rule set the probability of obtaining zero out of twenty responses, for an antigen with a true positive rate of 20%, at < 0.015

[x] This approach to cell isolation (which required as little as 7cc of blood) was taken in place of one utilizing peptide-pulsed dendridic cells, as the latter normally requires larger blood volumes to generate highly purified CD8+ T and dendridic cells.

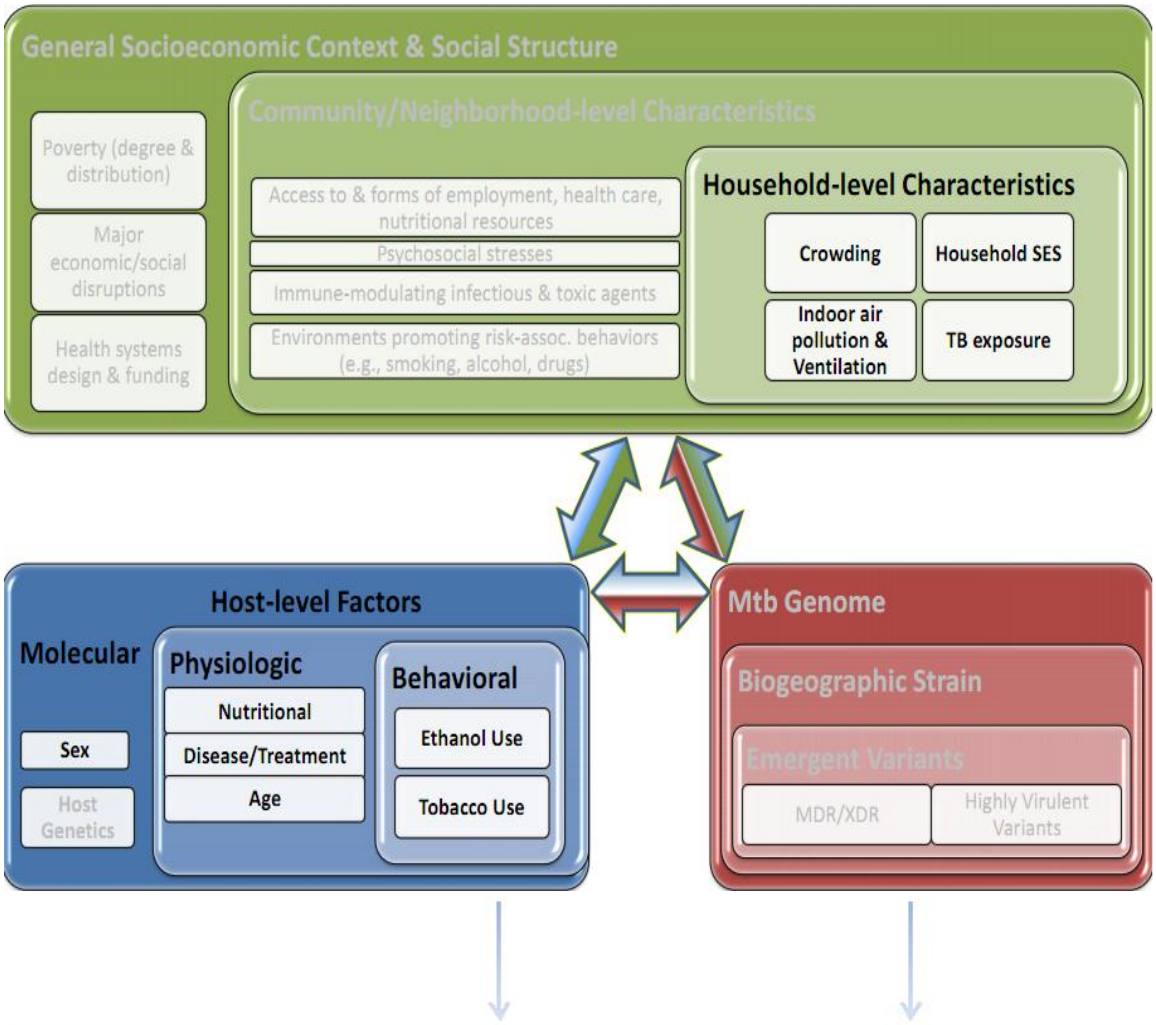
[xi] The Kawempe Community Health Study is an observational household-contact investigation, run by the TBRU, that was designed as a hybrid cross-sectional/cohort study. It was initiated in 2002 and enrolls approximately 150 ATB subjects and their household members per year. The study is intended to 1) “determine critical host factors associated with primary MTB infection, re-infection, reactivation, and progression of clinical disease [and 2)] to identify and track individual strains of MTB through Ugandan households and local community.”

[xii] As discussed in the Background section, there exist a number of factors related to Mtb infection and progression that may also, independently, effect the pathogen-host relationship, the character of Mtb-associated CD8+ T cells (in terms of specific antigen/epitope specificity, general reactivity, etc), or some other aspect of immunity that would influence the assay outcome. Among these were factors related specifically

to the pathogen (e.g., Mtb strain) and as well as those specific to the host. Included among the latter are factors related more or less directly to immune status as well as a range of demographic, behavioral, and other host attributes known to impact TB susceptibility and disease course, such as: sex, age, nutritional status, vitamin D and iron levels, HLA typology, degree of background systemic inflammation, exposure to non-tuberculous mycobacteria, and major systemic diseases/treatments (including diabetes, malignancy, renal disease, gastrectomy, celiac disease, silicosis, HIV infection, organ transplant, pharmacologic immune suppression.) Stepping back even further, the relationship between host and pathogen is influenced by yet additional layers of distal-to-intermediate links within increasingly complex causal chains that involve, for example, the magnitude and temporal character of Mtb exposure/re-exposures. However, the practical utility of examining distal and intermediate factors is limited and, instead, an effort will be made to identify and characterize the most proximal links within these causal chains (e.g., malnutrition, elevated rates of Mtb exposure) that more directly give rise to increased risk. Doing so will provide greater generalizability and precision with respect to both the assessment of the antigen-discovery process more generally and the diagnostic utility of the antigen set, specifically.

Data regarding a number of these factors were either not collected as part of the KCHS, appeared incomplete in the KCHS dataset, or were not generated in subsequent laboratory analyses of subjects' blood samples. This gap—i.e. the difference between the smaller set of potential confounding variables incorporated into this study and the larger set identified earlier in the Background section—is depicted graphically, below. Among these were pathogen characteristics, detailed exposure histories, host genetic factors, serum solute levels (cytokines, iron, or Vitamin D), and other potentially relevant neighborhood/workplace, individual, and/or household characteristics

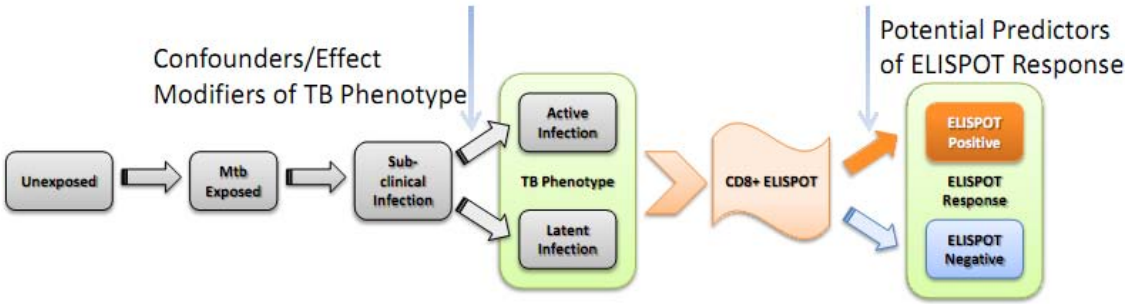
Selection of Host & Environmental Factors included in Multivariate Analyses



Variables Included in Inferential Analysis of INF-γ CD8+ T Cell Response (Based on Biological Plausibility/Proximity in Causal Chain):

Potential Secondary Independent and Confounding Variables		
Conceptual Domain	**Construct Category**	**Variable**
Host-level Characteristics	Demographics	**Sex**
		AGE
	TB Infection Characteristics	Sputum Smear (*ATB only*)‡
		Cardinal Symptoms (*ATB only)*‡
		Extent Lung Ds. (*ATB only*)‡
	Nutritional	**Underweight**
Environmental Factors	Mtb Exposure	TB+ Contact Exposure (*LTBI only*)‡
	Non-tuberculous Mycobacteria Exposure	**Water Source**

‡ Analyses that included TB phenotype-specific variables were subset for ATB/LTBI, as appropriate.



[xiii] While factors such as occupation and tribal affiliation are salient "social facts" that are likely to have real, if complex, etiologic relationships to TB phenotype (and, potentially, INF-γ CD8+ T cell response), in the face of such conceptual imprecision they are utilized in the descriptive analysis and bu excluded from inferential portions, where (1) interpretability would be highly problematic and (2) the inclusion of factors such as smoking and low BMI—both likely to be more proximal, direct risk factors within multiple causal chains—offer greater precision and generalizability.

[xiv] Each of these variables are associated with a specific TB phenotype—hours of MTB contact with LTBI and Number of Symptoms and Extent of Lung Disease with ATB subjects—and are therefore only included in phenotype-specific models.

[xv] This is a standard, validated definition utilized throughout the CD8+ T cell antigen discovery process.

[xvi] This was limited to interval and ratio-level variables and assessed relative to magnitude in the reduction in model deviance (i.e. -2*Log Likelihood.)

[xvii] Hosmer & Lemeshow χ^2 test was used in any model with a continuous variable to determine if the predicted distribution of events across risk deciles differs significantly from what was observed.

[xviii] Active TB cases were disproportionately from the student/white collar occupational strata. This unexpected finding is thought to be an artifact of the recruitment process and does not indicate an actual elevated risk for active disease among this occupational group.

[xix] Even if such a TB phenotype-specific antigen were to be included among those tested in this study, several challenges to identifying it are inherent in the study's population and design. First, the ATB subjects participating in this study were members of a population with high Mtb infection prevalence and incidence rates, many of whom may

have experienced extended periods of latent infection prior to the onset of the active disease episode that led to their identification and enrollment in the TBRU household contact study. As such, they would both have had the opportunity to develop clonal populations of CD8+ T cells primed for latency-associated antigens. Moreover, for these individuals, there would have been a period of high intracellular Mtb proliferation associated with the transition from latent infection to active disease that could further boost CD8+ T cell response. More generally, given that the boundaries between LTBI and ATB may be far more fluid than this common binary ontology would suggest, it may be that the character of immunologic overlap between these two clinical states is more or less pronounced in different settings. For example, it might be that exposure/re-exposure patterns within TB-endemic regions, duration of latency, particular forms of compromised health, certain environmental exposures, and/or other factors not fully accounted for in the current analysis lead to enhanced CD8+ T cell reactivity among those with active TB with respect to antigens that, in other contexts, are preferentially expressed in latent infections. In the context of the current investigation, the fact that subjects were drawn from a household contact study necessarily increases the likelihood of enrolling a particular kind of LTBI subject: one with a significant degree of sustained, recent direct exposure to a known ATB case (60.5 hours per month, on average.) Such individuals, despite their latent status, might exhibit immunologic dynamics more characteristic of regions along the TB phenotype spectrum associated with the ATB-LTBI transition point, in some respects sharing more in common with ATB individuals then, perhaps, other LTBI individuals with more distant exposures and a stable, classically latent immunologic dynamic.

[xx] The sample appeared to be predominantly urban; there was, for example, only one individual who identified himself as working in the primary sector (as a farmer.)

[xxi] The notable exception to this claim is the variable “mean background response.” However, one might reasonably expect diminished T cell activity to equally affect background and antigen-specific reactivity such that any fluctuation in activity would be accounted for in the outcome variable (which incorporates both measures.) Moreover, the analysis suggests that lower, not higher, background levels were associated positive ELISPOT response.

[xxii] However, direction bias resulting from non-blinding of subjects and interviewers would, in this study, not affect the validity of any association between ELISPOT outcome and the primary independent variable of interest, TB phenotype, but instead be limited to a misclassification of some other reported variable (e.g., smoking, alcohol use, length

of contact with TB positive individuals, or other factor perceived to be related to TB status.) Thus, the misclassification of exposure would not be based on outcome directly but only indirectly and only to the extent that that TB phenotype was related to ELISPOT response.