medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20100776; this version posted May 16, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

# Discerning divergent tuberculosis endotypes: A meta-analysis and systematic review of individual patient data

Andrew R. DiNardo*[1], Kimal Rajapakshe*[2,3], Tanmay Gandhi*[2], Sandra L. Grimm[2], Tomoki Nishiguchi[1], Jan Heyckendorf[4], Jaqueline Kahari[5], Qiniso Dlamini[5], Christoph Lange[6], Jeffrey D. Cirillo[7], Stefan HE Kaufmann[8,9], Reinout van Crevel[10], Mihai G. Netea[10, 11], Anna M. Mandalakas$[1], Cristian Coarfa$[2,3,12]

*Co-first authors contributing equally. $Co-senior authors contributing equally.

1 The Global Tuberculosis Program, Texas Children's Hospital, Immigrant and Global Health, Department of Pediatrics, Baylor College of Medicine, Houston, USA
2 Dan L Duncan Comprehensive Cancer Center, Baylor College of Medicine, Houston, USA
3 Molecular and Cellular Biology Department, Baylor College of Medicine, Houston, USA
4 Division of Clinical Infectious Diseases, Research Center Borstel; German Center for Infection Research (DZIF) Clinical Tuberculosis Unit, Borstel, Germany; Respiratory Medicine &International Health, University of Lübeck, Lübeck, Germany
5 Baylor-Swaziland Children's Foundation, Mbabane, Swaziland
6 Department of Medicine, Karolinska Institute, Stockholm, Sweden
7 Department of Microbial Pathogenesis and Immunology, Texas A&M Health Science Center, Bryan, TX, USA
8 Max Planck Institute for Infection Biology, Berlin, Germany
9 Hagler Institute for Advanced Study at Texas A&M University, College Station, TX, USA
10 Department of Internal Medicine and Radboud Center for Infectious Diseases, Radboud University Medical Center, Nijmegen, Netherlands
11 Genomics and Immunoregulation, Life & Medical Sciences Institute, University of Bonn, 53115 Bonn, Germany
12 Center for Precision Environmental Health, Baylor College of Medicine, Houston, USA

## ABSTRACT

**Rationale**: Host response is a critical factor determining susceptibility to tuberculosis (TB). A delicate balance should be maintained between intracellular immunity against *Mycobacterium tuberculosis (Mtb)* and minimizing detrimental immunopathology. Studies have identified incongruous immune responses that can lead to a similar TB disease phenotype. Instead of envisioning that susceptibility to TB follows a singular path, we propose the hypothesis that varied host endotypes exist within the TB clinical phenotype.

**Methods and Results**: Unbiased clustering analysis from 12 publicly available gene expression datasets consisting of data from 717 TB patients and 527 controls, identified 4 TB patient endotypes with distinct immune responses. The two largest endotypes exhibit divergent metabolic, epigenetic and immune pathways. TB patient endotype A, comprising 333 TB patients (46.4%), is characterized by increased expression of genes important for i) glycolysis, ii) IL-2-STAT5, IL-6-STAT3, Type I and II Interferon IFN-γ and TNF signaling and iii) epigenetic-modifying genes. In contrast, TB patient endotype B, comprising 313 TB patients (43.6%), is characterized by i) upregulated NFAT and hormone metabolism, and ii) decreased glycolysis, IFN-γ and TNF signaling. *In silico* evaluation suggests therapies beneficial for endotype A could be detrimental to endotype B, and vice versa. Multiplex ELISA completed from an external validation cohort confirmed a TB patient sub-group with decreased immune upregulation.

**Conclusions**: Host immunity to TB is heterogenous. Unbiased clustering analysis identified distinct TB endotypes with divergent metabolic, epigenetic and immune gene expression profiles that may enable stratified or personalized treatment management in the future.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20100776; this version posted May 16, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

## INTRODUCTION:

For almost half a century, treatment recommendations for patients with drug-susceptible tuberculosis (TB) have not changed. With existing drug combination therapy lasting at least six months, there has been interest in identifying host-directed therapies (HDT) that could both improve the efficacy of and shorten existing treatment regimens. To date, a singular immune correlate of protection against TB has not been identified. In contrast, asthma, COPD and most cancers are managed based on their endotype or molecular sub-classification [1-3]. For example, individuals with lung cancer are sub-categorized by histopathology (small-cell vs. non-small cell), by genotype (EGFR Exon 18-21 mutations, ALK translocations, ROS1 translocations, BRAF V600 mutations, NTRK fusion products), by PDL-1 expression and stage (localized vs. metastatic) [3]. Treatment depends on these sub-categorizations. A classification system for guiding TB therapy has not yet been devised; however new data suggests that simple clinical information may help stratify the duration of TB therapy [4]. Identification of host immune endotypes in TB would further guide a stratified based approach for host-directed therapeutics.

Immune studies of TB have identified incongruous immune responses that can lead to a similar TB disease phenotype [5-9]. Therefore, instead of envisioning TB as an infectious disease with a stereotypical immune response, we sought to assess whether diverse host immune responses could result in a similar clinical phenotype. Previous studies have identified a narrow host therapeutic window with both suppressed immunity or exuberant immunity resulting in decreased *Mtb* killing capacity [8,10,11]. Anti-mycobacterial immunity is complex, requiring a multicomponent approach including multiple cell types. For example, upstream defects in the IL-12- IFN-γ signaling pathway result in decreased IFN-γ production and susceptibility to TB. In contrast, downstream defects in the IL-12- IFN-γ pathway result in excess IFN-γ but decreased IFN-γ-induced gene expression and increased susceptibility to TB [7,12]. Similarly complex is the immune response to TNF: decreased TNF results in intracellular mycobacterial survival, while excess TNF induces immune cell death via necroptosis and extracellular mycobacterial survival [5,11,13]. A narrow therapeutic window of host mycobacterial immunity has also been described for the MyD88-adaptor-like protein (MAL) and immune checkpoint inhibitors [6,8,10]. We also reasoned that some individuals would progress to TB despite an apparently well-functioning host immune response if they had been infected with a high bacillary burden, experienced prolonged exposure, or were infected with a hyper-virulent strain of *Mtb*. Therefore, to identify potential endotypes of host immune response during TB, we implemented an unbiased clustering of publicly available microarray gene expression datasets.

## METHODS

### Study inclusion

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20100776; this version posted May 16, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

We implemented a systematic review and individual participant meta-analysis according to the PRISMA-IPD guidelines. An a priori hypothesis was established and shared with co-authors prior to implementation of bioinformatics. Publicly available data was identified using PubMed, and the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) repository. Participants without microbiologic confirmation were not included. Studies that did not describe the methods of microbiologic confirmation, or evaluated cells other than whole blood, were excluded. Studies using microarray were used for a discovery cohort and genome-wide (RNA-seq) data sets were used for a validation cohort (Table 1). The Remove Unwanted Variation (RUV) method was used to normalize the data.

**Table 1**: Summary of the gene expression studies included.

**Microarray Discovery Cohort**

| Reference | GEO Identifier | Platform | Healthy Controls used | TB Patients used | Countries included | Age Range (years) | Method of Microbiologic Confirmation | Inclusion of HIV co-infected individuals |
|---|---|---|---|---|---|---|---|---|
| Berry 2010 | GSE19491 | GPL6947 | 45 | 103 | UK and South Africa | >18 | Culture | No |
| Bloom 2012 | GSE40553 | GPL10558 | 38 | 166 | UK and South Africa | >17 | Culture | No |
| Cliff 2012 | GSE31348, GSE36238 | GPL570 | 0 | 36 | South Africa | 18 – 65 | Smear and culture | No |
| Maertzdorf 2012 | GSE34608 | GPL6480 | 0 | 8 | Germany | 21 – 65 | ??? | No |
| Bloom 2013 | GSE42834 | GPL10558 | 113 | 35 | UK | > 18 | Culture | No |
| Verhagen 2013 | GSE41055 | GPL5175 | 0 | 2 | Venezuela | 1.1 – 14.6 | Culture | No |
| Anderson 2014 | GSE39939, GSE39940 | GPL10558 | 68 | 35 | South Africa, Malawi, Kenya | < 15 | Culture | **Yes** |
| Kaforou 2014 | GSE37250 | GPL10558 | 167 | 195 | South Africa, Malawi, Kenya | 19 – 53 | Culture | **Yes** |
| Lai 2015 | GSE58411 | GPL10558 | 0 | 31 | South Africa | > 18 | Smear or culture | **Yes** |
| Tientcheu 2015 | GSE62147 | GPL6480 | 0 | 26 | The Gambia and UK | 15 – 79 | Culture | No |
| Blankley 2016 | GSE83456 | GPL10558 | 61 | 45 | UK | 20 – 80 | Culture | No |
| Walter 2016 | GSE73408 | GPL11532 | 35 | 35 | United States | 18 – 65 | Smear and culture | No |

**RNA-Seq Validation Cohort**

| Reference | GEO Identifier | Platform | Healthy Controls used | TB Patients used | Countries included | Age Range (years) | Method of Microbiologic Confirmation | Inclusion of HIV co-infected individuals |
|---|---|---|---|---|---|---|---|---|
| Suliman 2018 | GSE94438 | GPL11154 | 314 | 0 | The Gambia, South Africa, Ethiopia, Uganda | 10 - 18 | Symptoms and culture | No |
| Leong 2018 | GSE101705 | GPL18573 | 16 | 28 | India | 16-65 | Culture | No |
| Singhania 2018 | GSE107995 | GPL20301 | 192 | 90 | UK | 17-82 | Culture | No |
| Turner 2020 | E-MTAB-8290. (Array Express) | | 0 | 54 | South Africa | 27-48 | Culture and Xpert | Yes |

## Clustering

To identify distinct TB endotypes, we used the Louvain network-based clustering algorithm implemented in the single-cell RNA-seq analysis package Seurat. Clusters were detected at a wide range of resolutions, from R = 0 to R = 1.2. Transcriptomes of identified clusters were compared to healthy controls; genes with an FDR-adjusted p-value<0.05 and fold change exceeding 2x were considered significant. Pathway enrichment analysis was implemented using Gene Set Enrichment Analysis (GSEA). Chemical compounds associated with the gene signatures were determined using the Library of Integrated Network-based Cellular Signatures (LINCS).

## Gene classifier

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20100776; this version posted May 16, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Given the fundamental difference between the microarray and RNA-seq profiling platforms, we normalized independently the microarray data and the RNA-seq data and used the former as the discovery cohort and the latter as the validation cohort. Binary classifier for endotype A and B, which consisted of 89% of the discovery cohort, was built using Logistic Regression.

### Pseudotime

Transcriptome profile trajectory analysis (pseudotime) was implemented on the 1244 TB patients and healthy controls using Monocle with data for selected gene groups plotted in the R statistical package.

### External Multiplex ELISA Validation Cohort

In order to validate the gene expression clustering findings, a cohort of TB patients and their asymptomatic household contacts from Eswatini were evaluated (n = 79) [14]. The study protocol was reviewed by appropriate ethical and institutional review boards and all participation was voluntary and in concordance with the Declaration of Helsinki. Pulmonary TB patients were defined by both the presence of symptoms and all included TB cases had microbiologic confirmation. Twenty-two (55%) and nine (23%) of TB cases and household contacts were HIV co-infected. Included household contacts remained asymptomatic for 12-months after enrollment. Forty-eight percent of controls had a positive Quantiferon test (Qiagen, Hilden, Germany).

### Statistics

Fisher's exact test assessed incidence of clinical variables between endotypes. Rank-sum of the cytokines/chemokines from the ELISA validation was used to stratify the top and bottom 50% of TB patients. Differences between ELISA sub-groups were analyzed using Mann-Whitney rank sum test.

## RESULTS

### Systematic selection of TB patient cohorts with whole blood transcriptomic profiles

Twenty-two gene expression studies were identified that included microbiologically confirmed pulmonary TB from whole blood. Studies evaluating less than 10,000 genes were excluded. Twelve studies applied microarrays that included 12,468 commonly evaluated genes, including 717 individuals with microbiologically confirmed TB and 527 asymptomatic healthy controls [15-25]. These 12 studies were used to complete clustering identification of TB-endotypes. Four additional studies used RNA-seq transcriptome profiling and included 172 TB cases and 522 controls. These four studies were reserved for creation of a validation cohort [26-29].

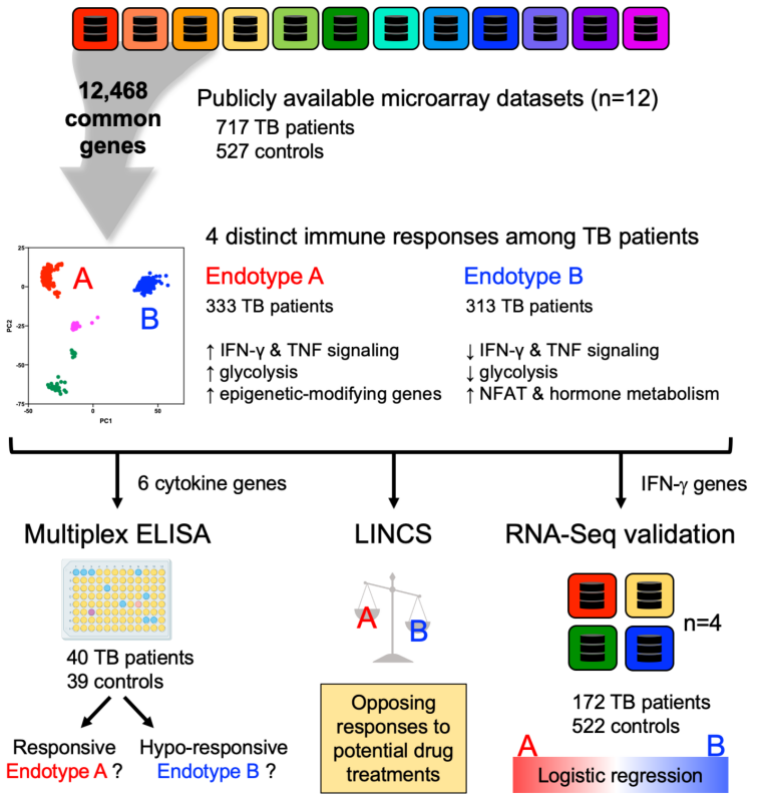
### Identification and clinical data characterization of TB endotypes

In order to identify potential TB endotypes, unbiased clustering of the 717 patients with microarray gene expression data was implemented (Figure 1). To better evaluate the structure of the TB transcriptomes, Louvain network-based clustering was run using a range of

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20100776; this version posted May 16, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

resolutions spanning from R=0 to R=1.2, and a clustering tree was generated to identify the number of potential sub-clusters. Using resolution R=0 four clusters were identified, at R=0.1-0.6 five clusters were identified, whereas at R=1.2 seven clusters were identified (Figure 2A). To maximize the detection and characterization power, the resolution R=0 was selected for further analysis. At resolution 0, the largest endotypes included 333 (46.4% of all TB patients) and 313 individuals (43.6%), respectively (Figure 2B). The identified endotypes were evaluated by available epidemiologic data such as sex, age, HIV status and country of origin (Supplemental Figure 1). There were no statistical differences in the sex of the different endotypes. In the discovery cohort, only three studies included individuals co-infected with HIV (GSE 37250, GSE39939 and GSE58411; Supplemental Figure 1). All individuals co-infected with HIV clustered in endotype A (32.4% of endotype A). At resolution R=0.6, endotype A splits into endotype A1 and endotype A2 and, interestingly, all HIV patients clustered into endotype A1 (49% of endotype A2). Deconvolution of whole-blood cell types was performed using CIBERSORT on the TB transcriptomes; endotype A had a relative increase in neutrophils compared to endotype B (Supplemental Figure 3). Only one study included children under 15 years of age (GSE39939), with all children clustering into endotype A (Supplemental Figure 1).

**Figure 1**: **Experimental design.** Microarray data from TB patients (n = 717) was first normalized using RUV, then clustered using the Louvain network-based unbiased clustering algorithm as implemented in the Seurat package. The gene expression profiles of distinct TB patient endotypes were compared to healthy controls (n = 527); pathway analysis identified divergent immune, metabolic and epigenetic pathway enrichment. Multiplex ELISA from an external validation cohort confirmed there is a sub-group of TB patients with decreased immune activation. An *in silico* screen of chemical compounds using the LINCS platform suggests that the two main divergent TB endotypes will respond differentially to candidate host-directed therapies. Logistic regression classification was developed for the two major endotypes and further validated on a transcriptomics (RNA-Seq) datasets of TB patients (n = 172) and controls (n = 522).




## Characterizing TB endotypes

We posited that while patients with pulmonary TB have a relatively similar clinical presentation, the host immune response would consist of potentially distinct and divergent endotypes. To test this hypothesis, differentially expressed genes from each cluster were used to determine pathway enrichment, assessed by Gene Set Enrichment Analysis (GSEA) using the Hallmark, REACTOME, GO Biological Process, and KEGG pathway compendia. Hierarchical clustering of GSEA Normalized Enrichment Scores (NES) results showed that endotypes A, C and D had

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20100776; this version posted May 16, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

similar gene expression responses defined by upregulation of type I and type II interferon, TNF, IL-2-STAT5 signaling, glycolysis, oxidative phosphorylation, and Myc targets (Figure 2C). In contrast, these biologic pathways were downregulated in endotype B. The TNF and IFN-γ signaling pathways are necessary, but not sufficient, components of the anti-mycobacterial immune response [7]. However both pathways have been shown to have narrow therapeutic windows, with either too little or too exuberant TNF and IFN-γ resulting in detrimental immunopathology [6,10,11,13].

Individuals in TB endotypes A, C and D had upregulation of the IFN-γ signal pathway, while this was downregulated in TB endotype B (Figure 3A). IFN-γ itself was not differentially expressed between the clusters, however

**Figure 2**: **Unbiased clustering identifies unique TB endotypes.** (A) Network-based unbiased clustering using the Louvain method identifies 4 endotypes of TB at resolution R=0 and increasingly up 7 endotypes at resolution R=1.2. (B) Principal Component Analysis (PCA) shows a robust separation between the four distinct clusters identified at resolution R=0. (C) TB endotypes A-D were compared to healthy controls, then pathway enrichment via Gene Set Enrichment Analysis (GSEA) was carried out against the Hallmark pathway compendium; hierarchical clustering based on the Normalized Enrichment Scores (NES) reveals striking differences between the endotypes.

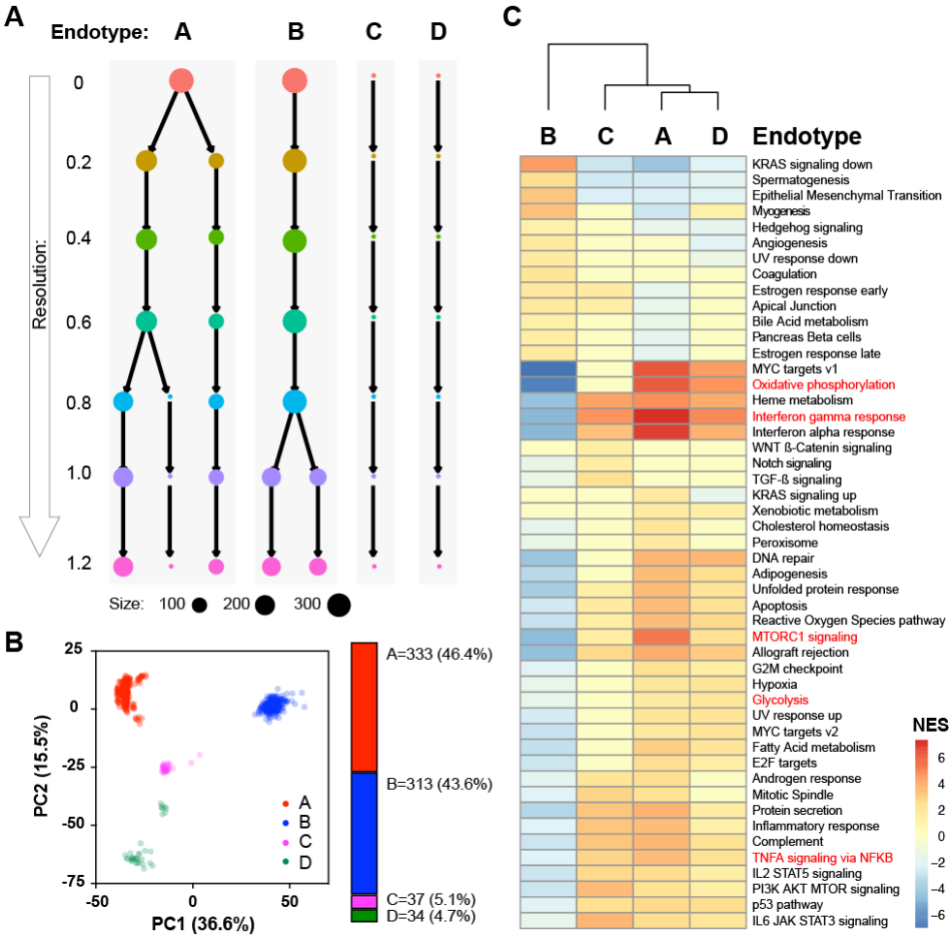
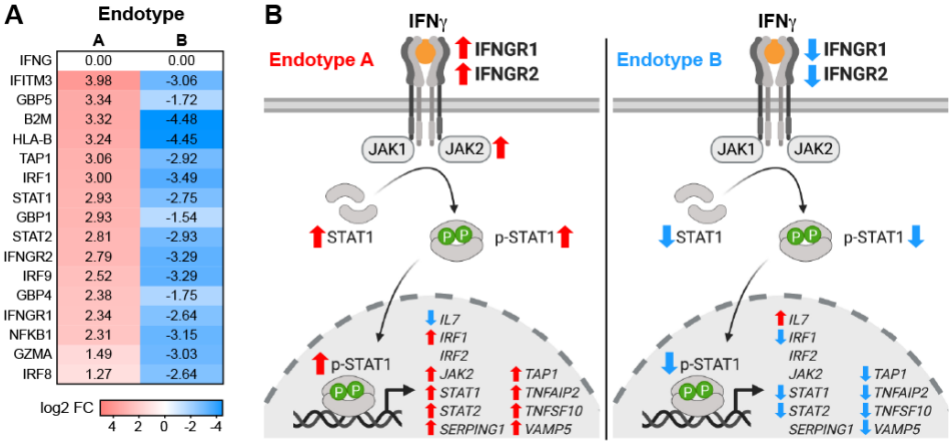



**Figure 3: Interferon-gamma signaling pathway in TB patient endotypes**. (A) Heatmap of differentially expressed genes in endotypes A and B compared to healthy controls (log2 fold change, FDR-adjusted p < 0.05). (B) Mapping differential gene expression between endotypes A and B compared to healthy controls onto the interferon gamma signaling pathway (over expressed, red up arrows; under expressed, blue down arrows).



medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20100776; this version posted May 16, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

endotype B had decreased expression of the IFN-γ signaling pathway, akin to a downstream Mendelian susceptibility to mycobacterial diseases (MSMD) genetic defect (Figure 3B). Similar to IFN-γ, TNF is necessary, but not sufficient, and TB in patients taking TNF antagonists are at increased risk of disease progression [30]. In contrast to TB patient endotype A, endotype B had both decreased expression of TNF as well as decreased TNF via NFKB signaling (Figure 4A,B).

**Figure 4**: **Tumor necrosis factor signaling pathway in TB patient endotypes.** (A) Heatmap of differentially expressed genes in endotypes A and B compared to healthy controls (log2 fold change, FDR-adjusted p < 0.05). (B) Mapping differential gene expression between endotypes A and B compared to healthy controls onto the interferon gamma signaling pathway (over expressed, red up arrows; under expressed, blue down arrows).

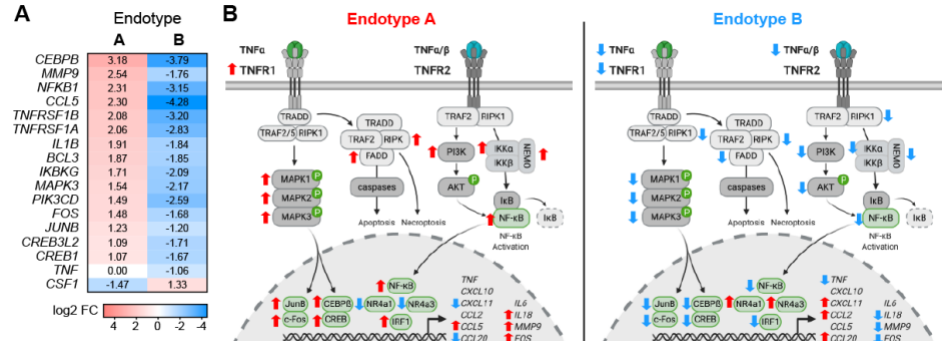
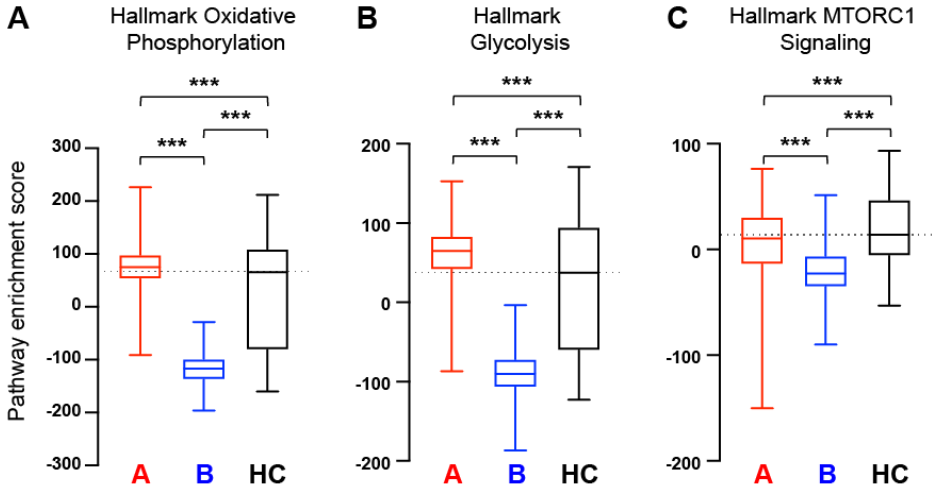



**Figure 5: Metabolic pathways show robust and distinct gene expression profiles in TB patient endotypes.** Pathway enrichment scores for Hallmark oxidative phosphorylation (A), glycolysis (B), and mTOR signaling (C) in endotypes A and B compared to healthy controls. *** denotes p<0.001 in Mann-Whitney unpaired non-parametric two-tailed test. Whiskers denote minimum-maximum and dotted line represents the healthy control median.

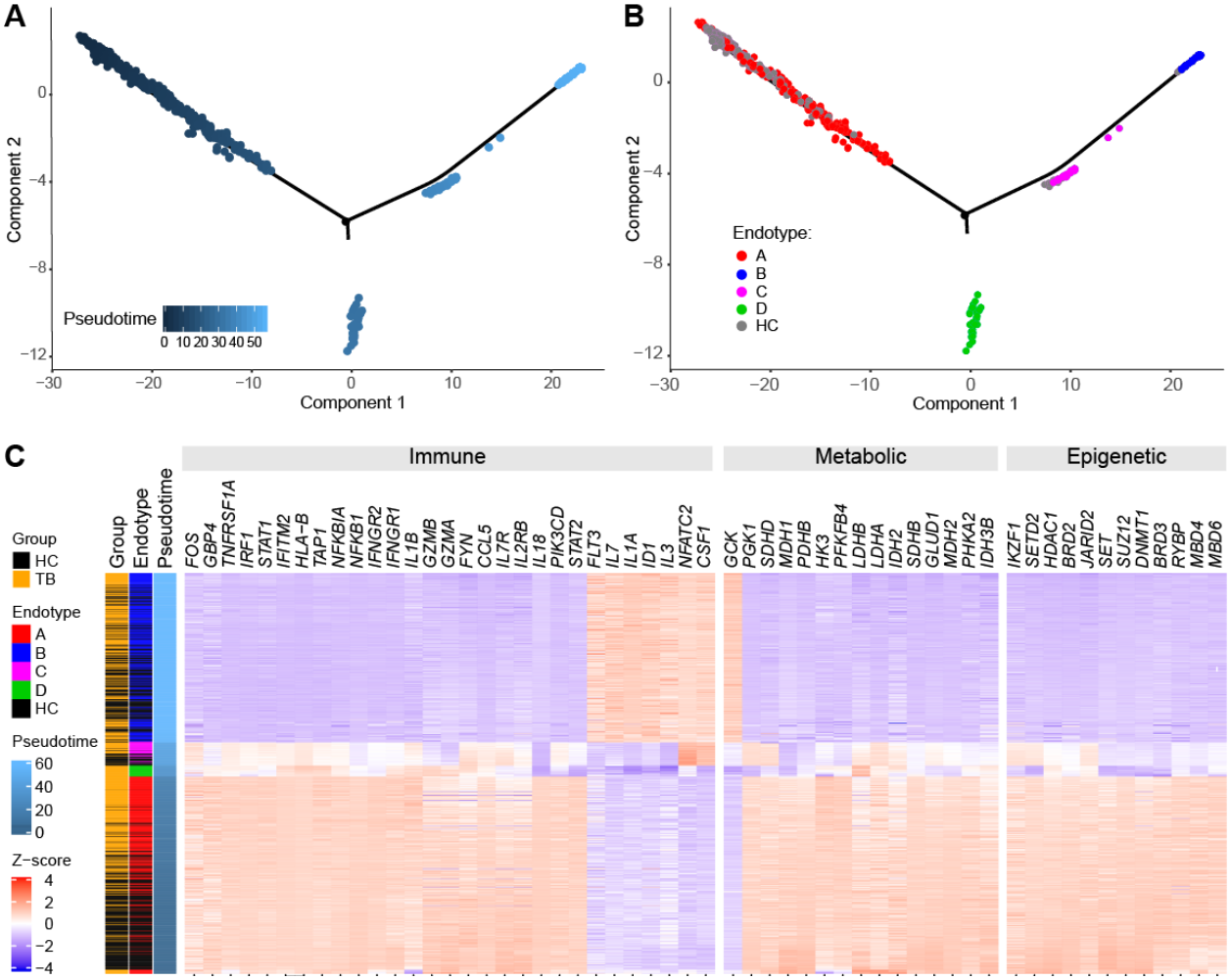



After activation, to increase metabolic demands for proliferation and effector functions, immune cells undergo metabolic shifts including increases in their metabolic activity via transitions towards glycolysis, termed the "Warburg effect" [31,32]. Using pathway activity scores, TB patient endotype A had increased oxidative phosphorylation, glycolysis, TCA cycle activation and glucose metabolism, while in contrast these pathways were significantly downregulated in endotype B (Figure 5, Supplemental Figure 2). After brief or chronic antigenic stimulation, this shift in glycolysis induces activation of epigenetic enzymes, such as DNMT and EZH2, and long-lasting chromatin conformation changes [33-38]. Therefore, to simultaneously evaluate metabolic, epigenetic and immune gene expression, we implemented trajectory analysis (pseudotime) using Monocle. Pseudotime demonstrated that endotype A expressed simultaneous upregulation of genes related to immunity (*IL1B, IL12RB, GZMA, IRF1*, *STAT1*), metabolism and epigenetics, while in contrast endotype B downregulated these same genes (Figure 6).

**Multiplex ELISA external cohort validation**

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20100776; this version posted May 16, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

**Figure 6: Immune, metabolic, and epigenetic trajectories of different TB endotypes.** Pseudotime representation of the trajectory of all 12,468 commonly evaluated genes (A) in TB endotypes (B). C) Pseudotime heatmap of the differential expression of select immune, metabolic, and epigenetic genes based on TB endotype and healthy controls, ordered along the inferred pseudotime.




To further evaluate the presence of different endotypes within TB patients, multiplex protein data was re-analyzed from an independent cohort of 40 TB patients and 39 healthy controls from Eswatini. Rank-sum analysis was implemented to divide the patients into immune hypo-responsive versus responsive groups based on their production of cytokines and chemokines (IFN-γ, TNF, IL-1β, IL-6, CXCL9, and CXCL10) after whole blood stimulation to mitogen (phytohemagglutinin) (Figure 7A-B). The TB patient sub-groups (top 50% versus bottom 50%) were compared to healthy controls. The hypo-responsive group demonstrated statistically lower levels of IFN-γ, TNF, IL-1β, IL-6, CXCL9, and CXCL10 induction, while the immune-responsive group was statistically similar to the healthy controls (Figure 7C; Mann-Whitney < 0.007).

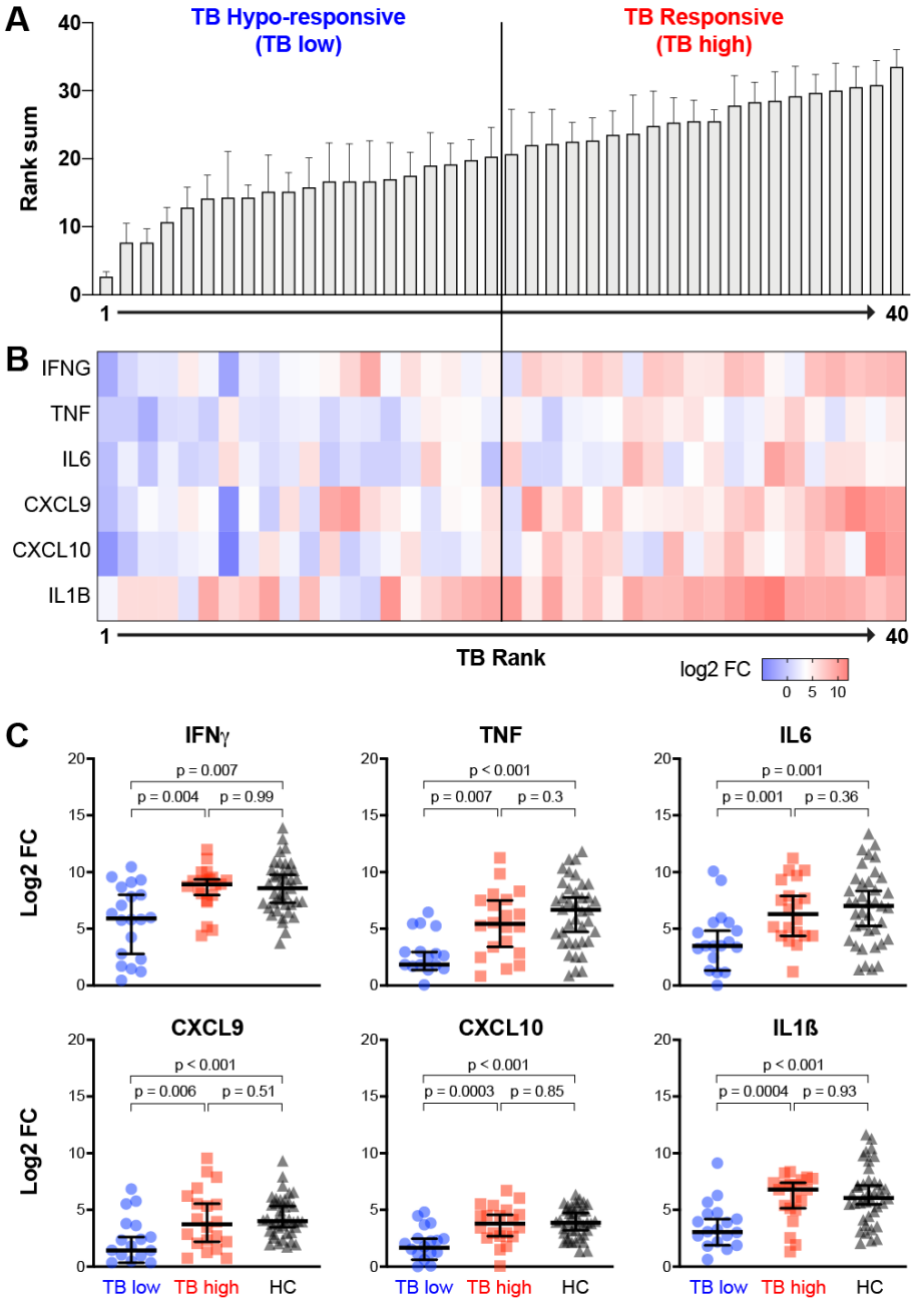
## Analysis of chemical compounds signatures for TB endotypes

A strategy to determine potentially effective drugs against a transcriptome signature is the use of the Library of Integrated Network-based Cellular Signatures (LINCS). Using the gene signatures against healthy controls for endotypes A and B, we searched LINCS ranked lists of

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20100776; this version posted May 16, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

over 5,000 chemical compounds to see which were similar or dissimilar. Similar to the contradictory metabolic, epigenetic and immune signatures, exogenous cytokines IL-2 and IFN-γ demonstrated contradictory connectivity scores between endotypes A and B. Similarly, other previously identified candidates for host-directed therapy (HDT), Vitamin D, glucocorticoids, non-steroidal anti-inflammatory drugs (NSAIDS), cyclooxygenase inhibitors, retinoids and metformin, demonstrated contradictory connectivity scores between TB patient endotypes A and B (Figure 8).

**Figure 7**: **Differential cytokine production by TB patient endotypes.** Whole blood from TB patients (n = 40) and healthy controls (n = 39) was stimulated overnight with or without mitogen (PHA), followed by measurement of cytokines and chemokines. (A) Samples were ranked for expression of six cytokines to determine an overall rank sum (1 is lowest, with 40 highest). Using the rank sum value, TB patients were then split in half into "hypo-responsive" and "responsive" groups. (B) Heatmap of cytokine expression as log2 fold change relative to controls. (C) Cytokine protein expression (log2 fold change) is graphed for each sub-group, with significance determined by Mann-Whitney non-parametric test.



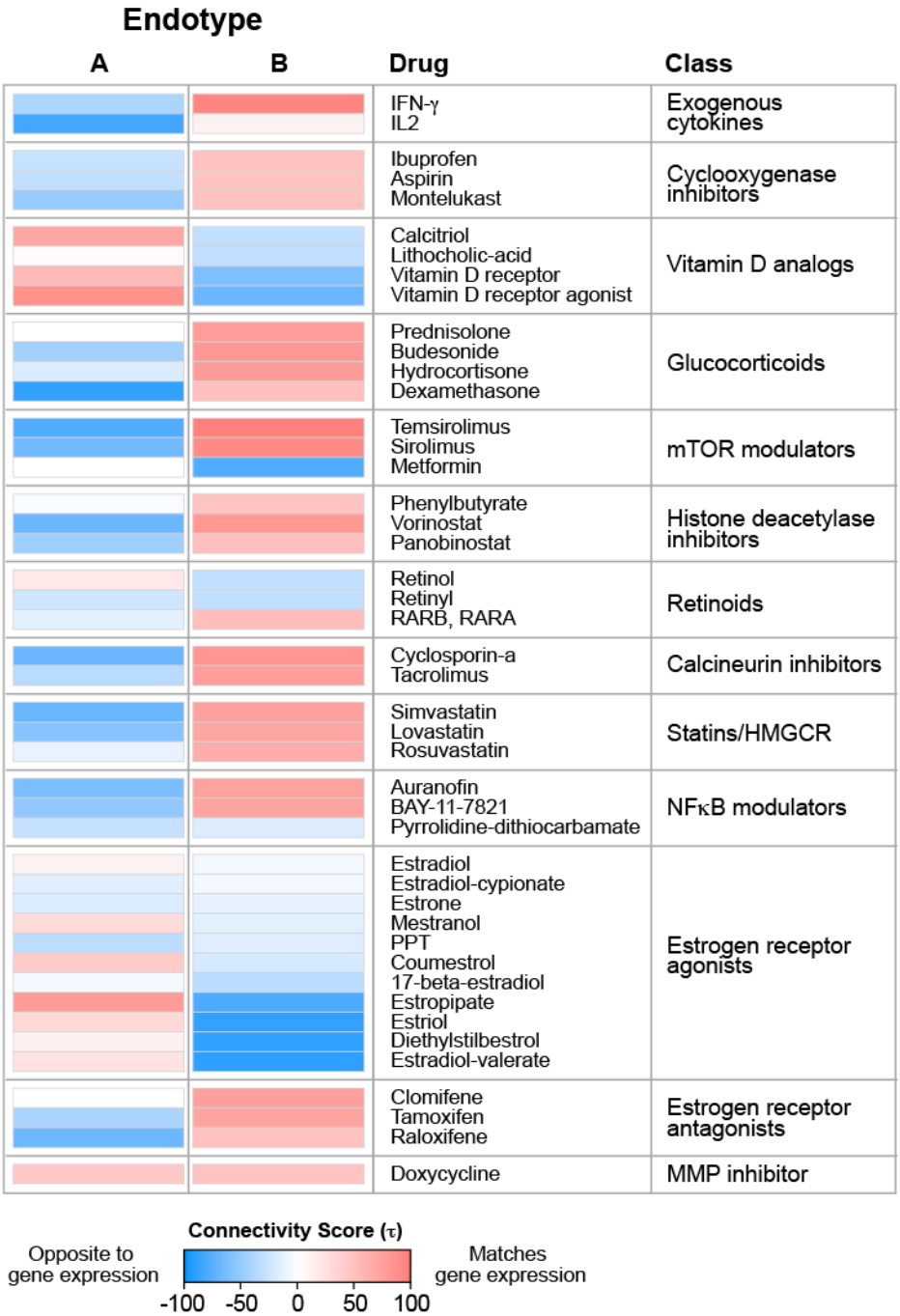

## Validation of endotypes A and B in TB blood RNA-seq cohorts

Given the importance of the IFN-γ pathway distinguishing endotype A from endotype B, we attempted to infer a classifier using the 200 Hallmark IFN-γ pathway genes. We applied the endotype A vs B IFN-γ gene classifier on an RNA-Seq validation cohort comprised of 172 TB patients and 522 healthy controls (Table 1). We employed the logistic regression method, and applying a 10-fold cross validation in the discovery cohort led to >99%

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20100776; this version posted May 16, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

accuracy. Given the discovery process of endotypes A and B, i.e. using the Louvain network community detection methods, such effective separations of the endotypes via machine learning are expected. Logistic regression classified 152 TB samples as endotype A and 20 as endotype B. We computed pathway enrichment using GSEA between each of the predicted validation subgroups and control samples. Logistic Regression captured the significant divergent gene expression enrichments for immune related pathways (Supplemental Figure 4). IL2/STAT5, IL6/JAK/STAT3, inflammatory response, IFN-α signaling, IFN-γ signaling, and allograft rejection were positively enriched for endotype A, and they were negatively enriched in endotype B in the discovery cohort and in the validation cohort for both classifier methods.

**Figure 8**. Heatmap of connectivity score for select chemical compounds within the TB endotypes A and B based on the Library of Integrated Network-based Cellular Signatures (LINCS). Positive connectivity scores represent chemical inducing gene expression profiles similar to the endotype, while negative connectivity scores represent chemical compounds inducing gene expression profiles antithetical to the endotype.



medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20100776; this version posted May 16, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

## DISCUSSION

In order to identify different distinct molecular sub-categories (endotypes) of host immune response against *Mtb*, we explored publicly available gene expression datasets to identify at least two distinct endotypes of host gene expression in TB with opposing immune, metabolic, and epigenetic characteristics. The two most predominant TB endotypes display contradictory responses to pathways previously thought to be required for mycobacterial immunity. Endotype A has upregulation of the IFN-γ and TNF, while both of these cytokine-signaling pathways are downregulated in endotype B.

The most commonly evaluated candidates for correlates of protection against *Mtb* are IFN-γ and TNF and their corresponding signaling pathways. Therefore, it is significant that they are contradictory between the two most commonly observed endotypes. Deficiencies in IFN-γ and TNF increase the risk for TB progression. However, animal models have demonstrated that both IFN-γ and TNF have narrow therapeutic windows with exuberant responses resulting in immune-mediated pathology [5,6,10,11,13]. Determining appropriateness of the host immune response identified in the two largest divergent endotypes is unclear with the limited available clinical information and lack of quantified bacillary burden. Theoretically, some individuals with endotype A TB have appropriate upregulation of IFN-γ and TNF targeting effector mechanisms against *Mtb*. In contrast, theoretically, some individuals with endotype A have detrimental, excessive IFN-γ and TNF inducing pathology. Combining gene expression analysis with functional immunology and quantitative measures of *Mtb* bacillary burden will clarify the appropriateness of the immune response between endotypes.

The link between metabolism, particularly glycolysis, and immune function has been appreciated for at least 93 years [39-41]. More recently, metabolism has been demonstrated as a mediator of the epigenetic mechanisms driving immune function [33-35,42]. Therefore, it is of particular interest that endotype A displayed upregulation of genes related to metabolism, epigenetics and the IFN-γ and TNF signaling, while in contrast endotype B displayed downregulation of these pathways. Many candidates for host-directed therapies target these pathways. For example, metformin mediates the AKT-mTOR pathway, blunting cellular glycolysis leading to inhibition of chromatin conformational changes that ultimately drive antigen-induced immune function [33,34]. These data suggest that drugs modulating metabolic and epigenetic mechanisms may need to be evaluated in an endotype-specific manner. Previously identified candidates for HDTs include exogenous IFN-γ, exogenous GM-CSF, exogenous TNF, TNF inhibitors, NSAIDS, Vitamin D, glucocorticoids, mTOR modulators (rapamycin, metformin), retinoids, and statins [43]. The divergent gene expression profiles identified between the two largest TB endotypes suggest a “one-size fits all” HDT approach may not be feasible. The *in silico* analysis demonstrated that previously identified HDTs will have contradictory predicted responses in endotype A versus endotype B. If functional studies validate one endotype to have decreased immune responsiveness, then vitamin D or exogenous recombinant IFN-γ would be appropriate HDT. In contrast, if future validation

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20100776; this version posted May 16, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

studies demonstrate one endotype to have pathologic, exuberant immunity, then an NSAID, TNF inhibitor or glucocorticoid would be appropriate. Evaluation of candidate HDTs will likely need to be endotype-specific and animal and *in vitro* models are needed to recapitulate the clinically relevant endotypes to better evaluate the appropriateness of candidate HDTs.
All included studies evaluated host gene expression at baseline, without antigenic or other stimulation. TB is a chronic infection and induces host immune suppression. While many genes down-stream of IFN-γ are elevated in TB patients at baseline [22,25,44,45], the TB-induced immune exhaustion decreases antigen-induced immune upregulation [14,46-48]. HIV suppresses host immunity. Therefore, the fact that all HIV participants clustered into endotype A1 (resolution 0.2 and below; Supplemental Figure 1) highlights the trouble with over-inferring immune function based on baseline, non-stimulated gene expression measurements. Therefore, future studies will need to evaluate the functional capacity of each endotype upon antigenic stimulation.

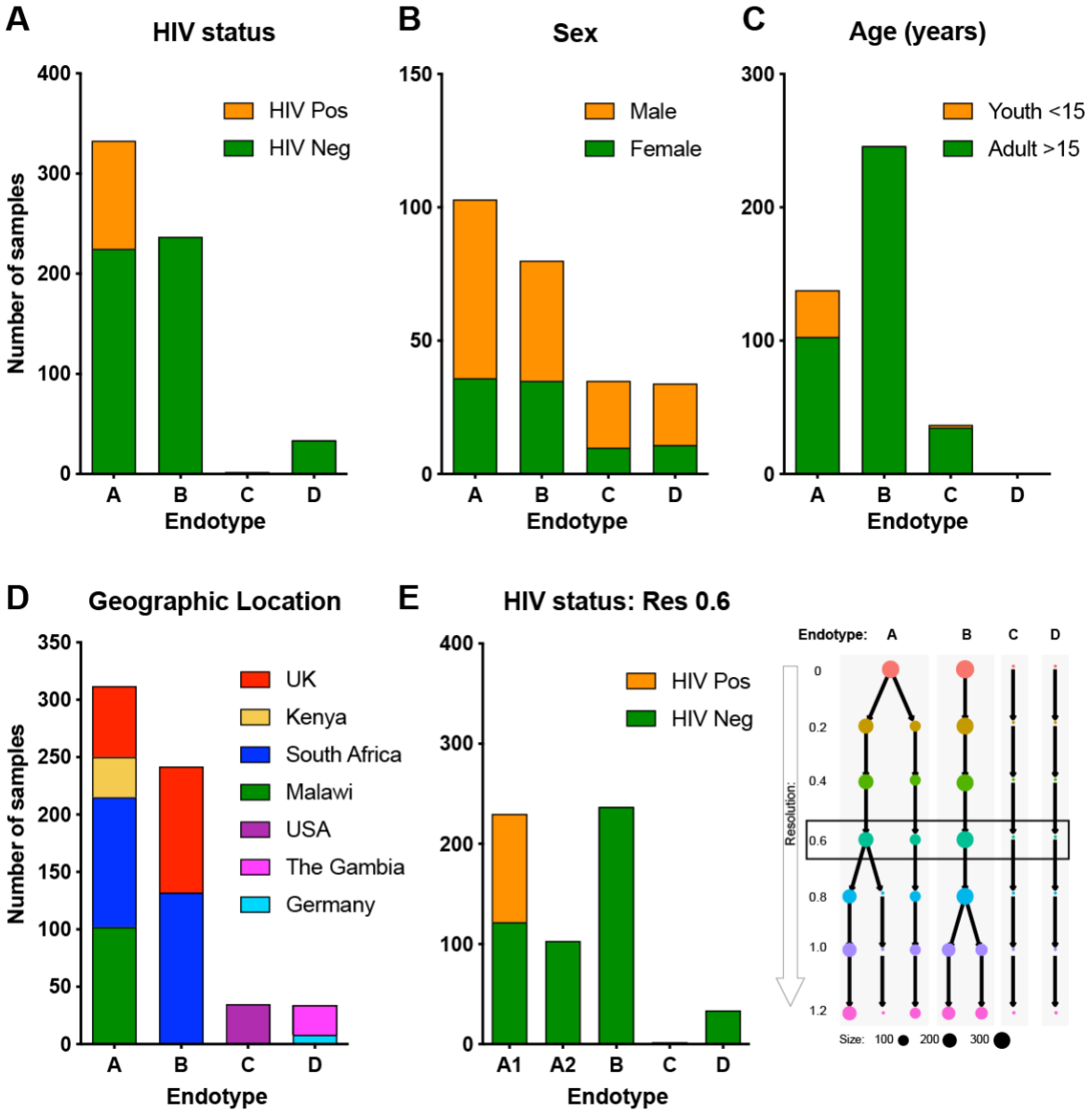
This study was predominantly derived from publicly deposited data and therefore has multiple inherent limitations. Progression to TB is related to the interaction between host, pathogen and environmental factors. The progression to a specific endotype of TB is also likely related to similar host, pathogen and environment interactions. However, few pathogen characteristics and very limited epidemiology are available in existing public data repositories. Epidemiologic predispositions likely to drive the divergent host gene expression endotypes include malnutrition, HIV, helminth, tobacco use and or indoor biomass fuel exposure. For example, despite successful deworming, previous schistosomiasis infection ablates mycobacterial immunity and leaves a long-lasting detrimental epigenetic-mediated immune exhaustion [49]. Another limitation is the lack of cell-specific data. Endotype A had increased neutrophils compared to endotype B, one potential explanation why all HIV cases clustered into Endotype A1. The defining feature of HIV infection is decreased cell-mediated immunity[50], particularly decreased IFN-γ production from lymphocytes. Considering previous evidence that neutrophils drive the gene expression signatures of TB patients [15], this suggests that cell-specific gene expression evaluations of endotype A1 would identify decreased lymphocyte IFN-γ signaling, but increased non-lymphocyte IFN-γ signaling. Based on the current data, it is unclear if the identified endotypes represent dichotomous groups or a continuous spectrum of host immunity. Similarly, the current data in unable to address if individuals are able to transition from one endotype to another during the disease process.

In summary, unbiased *in silico* analysis of publicly available data on markers of host immunity in patients with TB provides detailed evidence about the heterogeneity of immune responses in TB patients. Specifically, host gene expression in TB consists of at least two major, distinct endotypes with opposing immune, metabolic and epigenetic transcriptomic signatures. These observations suggest that despite similar clinical phenotypes, different TB patient endotypes display contradictory responses that likely have clinical and pathologic relevance. These unique endotypes are likely to differentially benefit from stratified host-directed therapies.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20100776; this version posted May 16, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

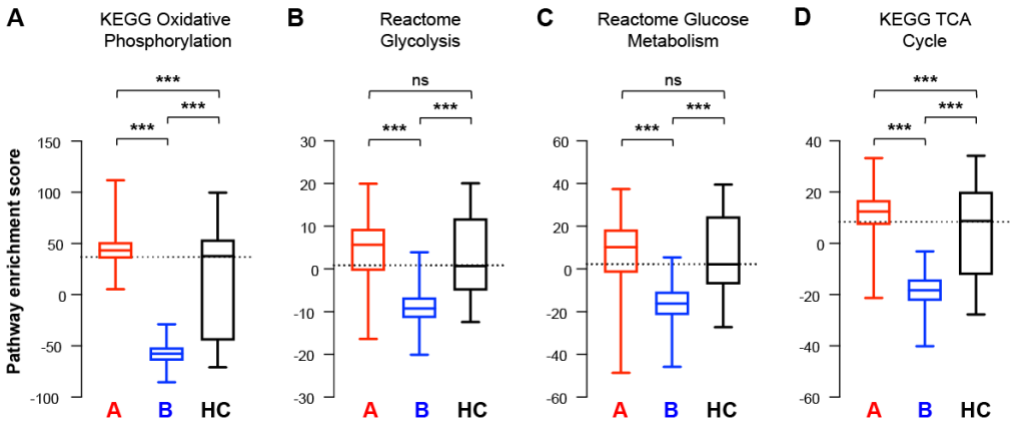
medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20100776; this version posted May 16, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

**Supplemental Figure 1**. Evaluation of HIV status (A), sex (B), age (C), geographic location (D) by endotype status at resolution 0. (E) HIV status is also shown at Resolution 0.6.

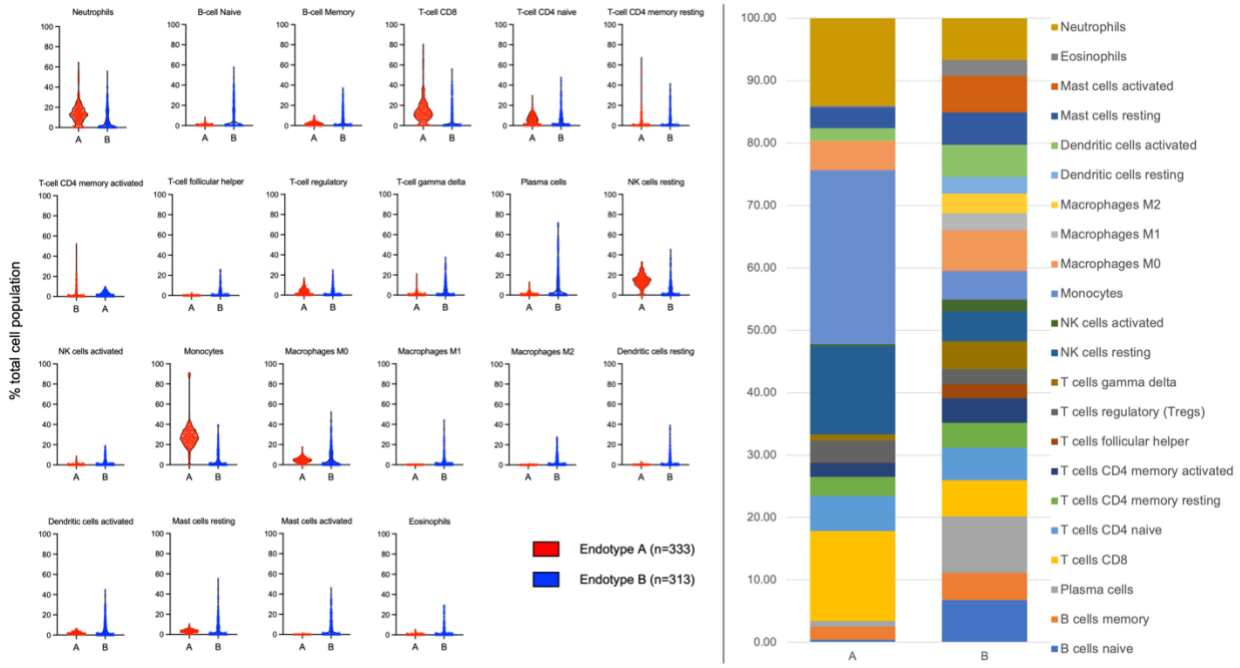


medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20100776; this version posted May 16, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

**Supplemental Figure 2**. **Metabolic pathways show robust and distinct gene expression profiles in TB patient endotypes.** Pathway enrichment scores for KEGG oxidative phosphorylation (A), Reactome glycolysis (B), Reactome glucose metabolism (C), and KEGG TCA cycle (D) in endotypes A and B compared to healthy controls. *** denotes p<0.001 in Mann-Whitney unpaired non-parametric two-tailed test. Whiskers denote minimum-maximum and dotted line represents the healthy control median.

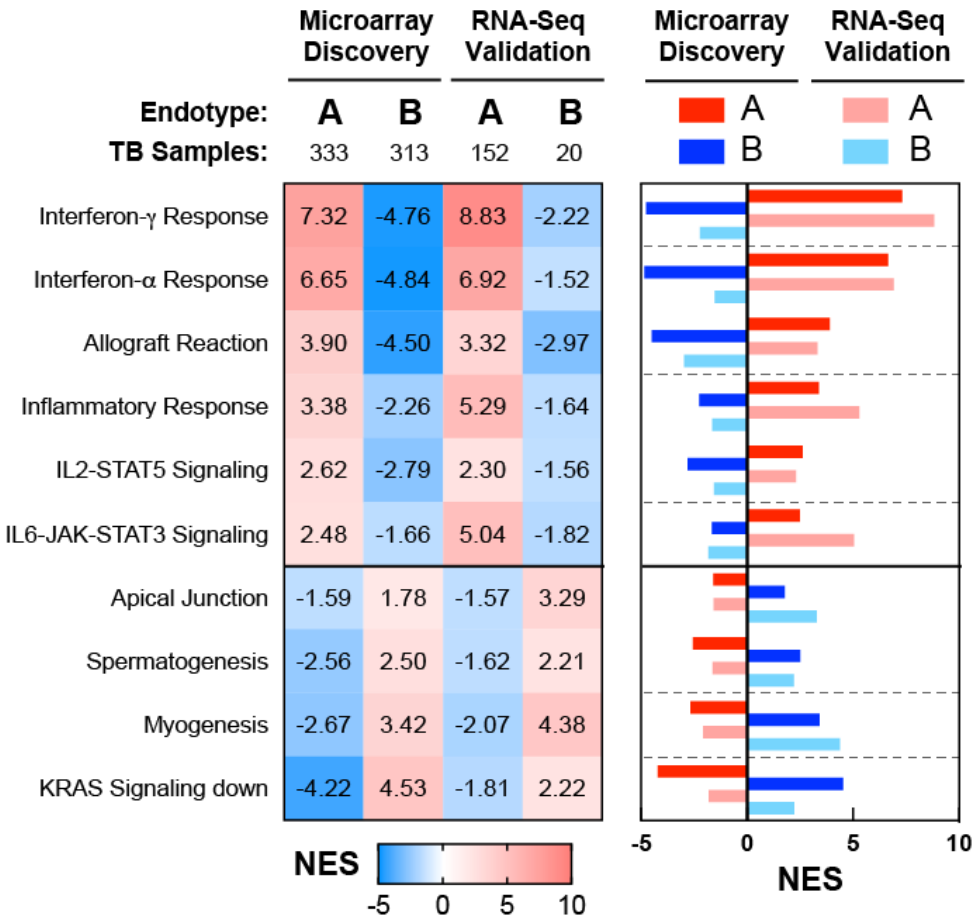



**Supplemental Figure 3**. Deconvolution of whole-blood cell types using CIBERSORT.



medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20100776; this version posted May 16, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

**Supplemental Figure 4.** Enrichment of Hallmark Immune Pathways in the validation RNA-Seq cohort as compared to the discovery microarray cohort.



medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20100776; this version posted May 16, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

# REFERENCES


1. Barnes PJ. Inflammatory endotypes in COPD. Allergy 2019;74(7):1249-56.
2. Svenningsen S, Nair P. Asthma Endotypes and an Overview of Targeted Therapy for Asthma. Front Med (Lausanne) 2017;4:158.
3. Dong J, Li B, Lin D, Zhou Q, Huang D. Advances in Targeted Therapy and Immunotherapy for Non-small Cell Lung Cancer Based on Accurate Molecular Typing. Front Pharmacol 2019;10:230.
4. Imperial MZ, Nahid P, Phillips PPJ, Davies GR, Fielding K, Hanna D, Hermann D, Wallis RS, Johnson JL, Lienhardt C, et al. A patient-level pooled analysis of treatment-shortening regimens for drug-susceptible pulmonary tuberculosis. Nat Med 2018;24(11):1708-15.
5. Tobin DM, Roca FJ, Oh SF, McFarland R, Vickery TW, Ray JP, Ko DC, Zou Y, Bang ND, Chau TT, et al. Host genotype-specific therapies can optimize the inflammatory response to mycobacterial infections. Cell 2012;148(3):434-46.
6. Sakai S, Kauffman KD, Sallin MA, Sharpe AH, Young HA, Ganusov VV, Barber DL. CD4 T Cell-Derived IFN-gamma Plays a Minimal Role in Control of Pulmonary Mycobacterium tuberculosis Infection and Must Be Actively Repressed by PD-1 to Prevent Lethal Disease. PLoS Pathog 2016;12(5):e1005667.
7. Bustamante J, Boisson-Dupuis S, Abel L, Casanova JL. Mendelian susceptibility to mycobacterial disease: genetic, immunological, and clinical features of inborn errors of IFN-gamma immunity. Semin Immunol 2014;26(6):454-70.
8. Ni Cheallaigh C, Sheedy FJ, Harris J, Munoz-Wolf N, Lee J, West K, McDermott EP, Smyth A, Gleeson LE, Coleman M, et al. A Common Variant in the Adaptor Mal Regulates Interferon Gamma Signaling. Immunity 2016;44(2):368-79.
9. Andrade BB, Pavan Kumar N, Amaral EP, Riteau N, Mayer-Barber KD, Tosh KW, Maier N, Conceicao EL, Kubler A, Sridhar R, et al. Heme Oxygenase-1 Regulation of Matrix Metalloproteinase-1 Expression Underlies Distinct Disease Profiles in Tuberculosis. J Immunol 2015;195(6):2763-73.
10. Barber DL, Mayer-Barber KD, Feng CG, Sharpe AH, Sher A. CD4 T cells promote rather than control tuberculosis in the absence of PD-1-mediated inhibition. J Immunol 2011;186(3):1598-607.
11. Tobin DM, Vary JC, Jr., Ray JP, Walsh GS, Dunstan SJ, Bang ND, Hagge DA, Khadge S, King MC, Hawn TR, et al. The lta4h locus modulates susceptibility to mycobacterial infection in zebrafish and humans. Cell 2010;140(5):717-30.
12. Orange JS, Levy O, Brodeur SR, Krzewski K, Roy RM, Niemela JE, Fleisher TA, Bonilla FA, Geha RS. Human nuclear factor kappa B essential modulator mutation can result in immunodeficiency without ectodermal dysplasia. J Allergy Clin Immunol 2004;114(3):650-6.
13. Roca FJ, Ramakrishnan L. TNF dually mediates resistance and susceptibility to mycobacteria via mitochondrial reactive oxygen species. Cell 2013;153(3):521-34.
14. DiNardo A, Rajapakshe K, Nishiguchi T, Mtetwa G, Grimm SL, Dlamini Q, Kahari J, Mahapatra S, Kay AW, Maphalala G, et al. DNA hyper-methylation during Tuberculosis dampens host immune responsiveness. J Clin Invest 2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20100776; this version posted May 16, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

15. Berry MP, Graham CM, McNab FW, Xu Z, Bloch SA, Oni T, Wilkinson KA, Banchereau R, Skinner J, Wilkinson RJ, et al. An interferon-inducible neutrophil-driven blood transcriptional signature in human tuberculosis. Nature 2010;466(7309):973-7.
16. Cliff JM, Lee JS, Constantinou N, Cho JE, Clark TG, Ronacher K, King EC, Lukey PT, Duncan K, Van Helden PD, et al. Distinct phases of blood gene expression pattern through tuberculosis treatment reflect modulation of the humoral immune response. J Infect Dis 2013;207(1):18-29.
17. Maertzdorf J, Weiner J, 3rd, Mollenkopf HJ, Network TB, Bauer T, Prasse A, Muller-Quernheim J, Kaufmann SH. Common patterns and disease-related signatures in tuberculosis and sarcoidosis. Proc Natl Acad Sci U S A 2012;109(20):7853-8.
18. Bloom CI, Graham CM, Berry MP, Wilkinson KA, Oni T, Rozakeas F, Xu Z, Rossello-Urgell J, Chaussabel D, Banchereau J, et al. Detectable changes in the blood transcriptome are present after two weeks of antituberculosis therapy. PLoS One 2012;7(10):e46191.
19. Bloom CI, Graham CM, Berry MP, Rozakeas F, Redford PS, Wang Y, Xu Z, Wilkinson KA, Wilkinson RJ, Kendrick Y, et al. Transcriptional blood signatures distinguish pulmonary tuberculosis, pulmonary sarcoidosis, pneumonias and lung cancers. PLoS One 2013;8(8):e70630.
20. Walter ND, Miller MA, Vasquez J, Weiner M, Chapman A, Engle M, Higgins M, Quinones AM, Rosselli V, Canono E, et al. Blood Transcriptional Biomarkers for Active Tuberculosis among Patients in the United States: a Case-Control Study with Systematic Cross-Classifier Evaluation. J Clin Microbiol 2016;54(2):274-82.
21. Verhagen LM, Zomer A, Maes M, Villalba JA, Del Nogal B, Eleveld M, van Hijum SA, de Waard JH, Hermans PW. A predictive signature gene set for discriminating active from latent tuberculosis in Warao Amerindian children. BMC Genomics 2013;14:74.
22. Anderson ST, Kaforou M, Brent AJ, Wright VJ, Banwell CM, Chagaluka G, Crampin AC, Dockrell HM, French N, Hamilton MS, et al. Diagnosis of childhood tuberculosis and host RNA expression in Africa. N Engl J Med 2014;370(18):1712-23.
23. Kaforou M, Wright VJ, Oni T, French N, Anderson ST, Bangani N, Banwell CM, Brent AJ, Crampin AC, Dockrell HM, et al. Detection of tuberculosis in HIV-infected and -uninfected African adults using whole blood RNA expression signatures: a case-control study. PLoS Med 2013;10(10):e1001538.
24. Tientcheu LD, Maertzdorf J, Weiner J, Adetifa IM, Mollenkopf HJ, Sutherland JS, Donkor S, Kampmann B, Kaufmann SH, Dockrell HM, et al. Differential transcriptomic and metabolic profiles of M. africanum- and M. tuberculosis-infected patients after, but not before, drug treatment. Genes Immun 2015;16(5):347-55.
25. Blankley S, Graham CM, Turner J, Berry MP, Bloom CI, Xu Z, Pascual V, Banchereau J, Chaussabel D, Breen R, et al. The Transcriptional Signature of Active Tuberculosis Reflects Symptom Status in Extra-Pulmonary and Pulmonary Tuberculosis. PLoS One 2016;11(10):e0162220.
26. Suliman S, Thompson E, Sutherland J, Weiner Rd J, Ota MOC, Shankar S, Penn-Nicholson A, Thiel B, Erasmus M, Maertzdorf J, et al. Four-gene Pan-African Blood Signature Predicts Progression to Tuberculosis. Am J Respir Crit Care Med 2018.
27. Leong S, Zhao Y, Joseph NM, Hochberg NS, Sarkar S, Pleskunas J, Hom D, Lakshminarayanan S, Horsburgh CR, Jr., Roy G, et al. Existing blood transcriptional classifiers

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20100776; this version posted May 16, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

accurately discriminate active tuberculosis from latent infection in individuals from south India. Tuberculosis (Edinb) 2018;109:41-51.

28. Singhania A, Verma R, Graham CM, Lee J, Tran T, Richardson M, Lecine P, Leissner P, Berry MPR, Wilkinson RJ, et al. A modular transcriptional signature identifies phenotypic heterogeneity of human tuberculosis infection. Nat Commun 2018;9(1):2308.
29. Turner CT, Gupta RK, Tsaliki E, Roe JK, Mondal P, Nyawo GR, Palmer Z, Miller RF, Reeve BW, Theron G, et al. Blood transcriptional biomarkers for active pulmonary tuberculosis in a high-burden setting: a prospective, observational, diagnostic accuracy study. Lancet Respir Med 2020;8(4):407-19.
30. Ehlers S. Role of tumour necrosis factor (TNF) in host defence against tuberculosis: implications for immunotherapies targeting TNF. Ann Rheum Dis 2003;62 Suppl 2:ii37-42.
31. Palsson-McDermott EM, Curtis AM, Goel G, Lauterbach MAR, Sheedy FJ, Gleeson LE, van den Bosch MWM, Quinn SR, Domingo-Fernandez R, Johnston DGW, et al. Pyruvate Kinase M2 Regulates Hif-1alpha Activity and IL-1beta Induction and Is a Critical Determinant of the Warburg Effect in LPS-Activated Macrophages. Cell Metab 2015;21(2):347.
32. Buck MD, Sowell RT, Kaech SM, Pearce EL. Metabolic Instruction of Immunity. Cell 2017;169(4):570-86.
33. Arts RJW, Carvalho A, La Rocca C, Palma C, Rodrigues F, Silvestre R, Kleinnijenhuis J, Lachmandas E, Goncalves LG, Belinha A, et al. Immunometabolic Pathways in BCG-Induced Trained Immunity. Cell Rep 2016;17(10):2562-71.
34. Cheng SC, Quintin J, Cramer RA, Shepardson KM, Saeed S, Kumar V, Giamarellos-Bourboulis EJ, Martens JH, Rao NA, Aghajanirefah A, et al. mTOR- and HIF-1alpha-mediated aerobic glycolysis as metabolic basis for trained immunity. Science 2014;345(6204):1250684.
35. Bengsch B, Johnson AL, Kurachi M, Odorizzi PM, Pauken KE, Attanasio J, Stelekati E, McLane LM, Paley MA, Delgoffe GM, et al. Bioenergetic Insufficiencies Due to Metabolic Alterations Regulated by the Inhibitory Receptor PD-1 Are an Early Driver of CD8(+) T Cell Exhaustion. Immunity 2016;45(2):358-73.
36. Vire E, Brenner C, Deplus R, Blanchon L, Fraga M, Didelot C, Morey L, Van Eynde A, Bernard D, Vanderwinden JM, et al. The Polycomb group protein EZH2 directly controls DNA methylation. Nature 2006;439(7078):871-4.
37. Schietinger A, Philip M, Krisnawan VE, Chiu EY, Delrow JJ, Basom RS, Lauer P, Brockstedt DG, Knoblaugh SE, Hammerling GJ, et al. Tumor-Specific T Cell Dysfunction Is a Dynamic Antigen-Driven Differentiation Program Initiated Early during Tumorigenesis. Immunity 2016;45(2):389-401.
38. Ghoneim HE, Fan Y, Moustaki A, Abdelsamed HA, Dash P, Dogra P, Carter R, Awad W, Neale G, Thomas PG, et al. De Novo Epigenetic Programs Inhibit PD-1 Blockade-Mediated T Cell Rejuvenation. Cell 2017;170(1):142-57 e19.
39. Ratzan KR, Musher DM, Keusch GT, Weinstein L. Correlation of increased metabolic activity, resistance to infection, enhanced phagocytosis, and inhibition of bacterial growth by macrophages from Listeria- and BCG-infected mice. Infect Immun 1972;5(4):499-504.
40. Kelly B, O'Neill LA. Metabolic reprogramming in macrophages and dendritic cells in innate immunity. Cell Res 2015;25(7):771-84.
41. Warburg O, Wind F, Negelein E. The Metabolism of Tumors in the Body. J Gen Physiol 1927;8(6):519-30.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20100776; this version posted May 16, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

42. Scharping NE, Menk AV, Moreci RS, Whetstone RD, Dadey RE, Watkins SC, Ferris RL, Delgoffe GM. The Tumor Microenvironment Represses T Cell Mitochondrial Biogenesis to Drive Intratumoral T Cell Metabolic Insufficiency and Dysfunction. Immunity 2016;45(2):374-88.
43. Kaufmann SHE, Dorhoi A, Hotchkiss RS, Bartenschlager R. Host-directed therapies for bacterial and viral infections. Nat Rev Drug Discov 2018;17(1):35-56.
44. Zak DE, Penn-Nicholson A, Scriba TJ, Thompson E, Suliman S, Amon LM, Mahomed H, Erasmus M, Whatney W, Hussey GD, et al. A blood RNA signature for tuberculosis disease risk: a prospective cohort study. Lancet 2016;387(10035):2312-22.
45. Gupta RK, Turner CT, Venturini C, Esmail H, Rangaka MX, Copas A, Lipman M, Abubakar I, Noursadeghi M. Concise whole blood transcriptional signatures for incipient tuberculosis: a systematic review and patient-level pooled meta-analysis. Lancet Respir Med 2020.
46. Harari A, Rozot V, Enders FB, Perreau M, Stalder JM, Nicod LP, Cavassini M, Calandra T, Blanchet CL, Jaton K, et al. Dominant TNF-alpha+ Mycobacterium tuberculosis-specific CD4+ T cell responses discriminate between latent infection and active disease. Nat Med 2011;17(3):372-6.
47. Sahiratmadja E, Alisjahbana B, de Boer T, Adnan I, Maya A, Danusantoso H, Nelwan RH, Marzuki S, van der Meer JW, van Crevel R, et al. Dynamic changes in pro- and anti-inflammatory cytokine profiles and gamma interferon receptor signaling integrity correlate with tuberculosis disease activity and response to curative treatment. Infect Immun 2007;75(2):820-9.
48. Singh A, Mohan A, Dey AB, Mitra DK. Inhibiting the programmed death 1 pathway rescues Mycobacterium tuberculosis-specific interferon gamma-producing T cells from apoptosis in patients with pulmonary tuberculosis. J Infect Dis 2013;208(4):603-15.
49. DiNardo AR, Nishiguchi T, Mace EM, Rajapakshe K, Mtetwa G, Kay A, Maphalala G, Secor WE, Mejia R, Orange JS, et al. Schistosomiasis Induces Persistent DNA Methylation and Tuberculosis-Specific Immune Changes. J Immunol 2018.
50. Aaron L, Saadoun D, Calatroni I, Launay O, Memain N, Vincent V, Marchal G, Dupont B, Bouchaud O, Valeyre D, et al. Tuberculosis in HIV-infected patients: a comprehensive review. Clin Microbiol Infect 2004;10(5):388-98.