Gene expression signatures identify biologically and clinically distinct tuberculosis endotypes

Andrew R. DiNardo*[#1], Tanmay Gandhi[#2,3], Jan Heyckendorf[#4,5], Sandra L. Grimm[2#], Kimal Rajapakshe[2,3], Tomoki Nishiguchi[1], Maja Reimann[4], Jaqueline Kahari[6], Qiniso Dlamini[5], Christoph Lange[1,4,5], Torsten Goldmann[4], Sebastian Marwitz[4], DZIF-TB cohort study group[13], Jeffrey D. Cirillo[7], Stefan HE Kaufmann[8,9], Mihai G. Netea[10, 11], Reinout van Crevel[10], Anna M. Mandalakas*[$1], Cristian Coarfa*[$2,3,12]

*Corresponding authors. [#]Co-first authors contributing equally. [$]Co-senior authors contributing equally, [13]Members of the DZIF-TB cohort study group are listed below.

[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;
[5] Respiratory Medicine & International Health, University of Lübeck, Lübeck, Germany
[6] Baylor-Swaziland Children's Foundation, Mbabane, Swaziland
[7] Department of Microbial Pathogenesis and Immunology, Texas A&M Health Science Center, Bryan, TX, USA
[8] Max Planck Institute for Infection Biology, Berlin, and Max Planck Institute for Biophysical Chemistry, Göttingen, 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

Corresponding authors: Andrew DiNardo, MD: dinardo@bcm.edu
Anna M. Mandalakas: anna.mandalakas@bcm.edu
Cristian Coarfa PhD: coarfa@bcm.edu

Running head: Distinct host gene expression endotypes in TB

Word count: 3183 words; 7 Figures;

The authors have declared that no conflict of interest exists.

Funding Sources: ARD is supported by NIAID K23 AI141681-02. CC is supported by the Cancer Prevention Institute of Texas (CPRIT) RP170005, NIH P30 shared resource grant CA125123, and NIEHS grants 1P30ES030285 and 1P42ES0327725. JH and CL are supported by the German Center for Infection Research (DZIF). JDC is funded in part from funds provided by the Texas A&M University System and National Institutes of Health grant AI104960. MGN is supported by an ERC Advanced Grant (#833247) and a Spinoza grant of the Netherlands Organization for Scientific Research. Rvc is supported by National Institute of Health (R01AI145781).AMM is supported by NIH R01AI137527, U01GH002278 and DoD W81XWH1910026. The authors report no competing or conflicts of interest.

×**Members of the DZIF-TB cohort study group:** Jan Heyckendorf[1,2,3], Sebastian Marwitz[4,5], Maja Reimann[1,2,3], Korkut Avsar[6], Andrew DiNardo[7], Gunar Günther[8,9], Michael Hoelscher[10,11], Elmira Ibraim[12], Barbara Kalsdorf[1,2,3], Stefan H.E. Kaufmann[13,14,15], Irina Kontsevaya[1,2,3], Frank van Leth[16,17], Anna Maria Mandalakas[18], Florian P. Maurer[19,20], Marius Müller[21], Dörte Nitschkowski[4,5], Ioana D. Olaru[22,23], Cristina Popa[12], Andrea Rachow[10,11], Thierry Rolling[24,25], Jan Rybniker[26,27,28], Helmut J. F. Salzer[29], Patricia Sanchez-Carballo[1,2,3], Maren Schuhmann[30], Dagmar Schaub[1,2,3], Victor Spinu[12], Isabelle Suárez[27], Elena Terhalle[1,2,3], Markus Unnewehr[31,32], January Weiner 3rd [33], Torsten Goldmann[4,5], Christoph Lange[1,2,3,18]

Affiliations of the members of the DZIF-TB cohort study group: [1]Division of Clinical Infectious Diseases, Research Center Borstel, Borstel, Germany; [2]German Center for Infection Research (DZIF), Germany; [3]International Health/Infectious Diseases, University of Lübeck, Lübeck, Germany; [4]Pathology of the Universal Medical Center Schleswig-Holstein (UKSH) and the Research Center Borstel, Campus Borstel, Airway Research Center North (ARCN); [5]German Center for Lung Research (DZL), Germany; [6]Asklepios Fachkliniken München-Gauting, Munich, Germany; [7]The Global Tuberculosis Program, Texas Children's Hospital, Immigrant and Global Health, Department of Pediatrics, Baylor College of Medicine, Houston, USA; [8]Department of Medicine, University of Namibia School of Medicine, Windhoek, Namibia; [9]Inselspital Bern, Department of Pulmonology, Bern, Switzerland; [10]Division of Infectious Diseases and Tropical Medicine, University Hospital, LMU Munich, Munich, Germany; [11]German Center for Infection Research (DZIF), partner site Munich, Germany; [12]Institutul de Pneumoftiziologie "Marius Nasta", MDR-TB Research Department, Bucharest, Romania; [13]Max Planck Institute for Infection Biology, Berlin, Germany; [14]Max Planck Institute for Biophysical Chemistry, Göttingen, Germany; [15]Hagler Institute for Advanced Study, Texas A&M University, College Station, USA; [16]Department of Global Health, Amsterdam University Medical Centres, Location AMC, Amsterdam, The Netherlands; [17]Amsterdam Institute for Global Health and Development, Amsterdam, The Netherlands; [18]The Global Tuberculosis Program, Texas Children's Hospital, Immigrant and Global Health, Department of Pediatrics, Baylor College of Medicine, Houston, USA; [19] National and WHO Supranational Reference Center for Mycobacteria, Research Center Borstel, Borstel, Germany; [20] Institute of Medical Microbiology, Virology and Hygiene, University Medical Center Hamburg-Eppendorf, Hamburg, Germany; [21]Sankt Katharinen-Krankenhaus, Frankfurt, Germany; [22]London School of Hygiene and Tropical Medicine, London, United Kingdom; [23]Biomedical Research and Training Institute, Harare, Zimbabwe; [24]Division of Infectious Diseases, I. Department of Internal Medicine, German Center for Infection Research (DZIF); University Medical Centre Hamburg-Eppendorf, Hamburg, Germany, [25]Department of Clinical Immunology of Infectious Diseases, Bernhard-Nocht-Institute for Tropical Medicine, Hamburg, Germany; [26]Department I of Internal Medicine, Division of Infectious Diseases, University of Cologne, Cologne, Germany; [27]German Center for Infection Research (DZIF), Partner Site Bonn-Cologne, Cologne, Germany; [28]Center for Molecular Medicine Cologne, University of Cologne, Cologne, Germany; [29]Department of Pulmonology, Kepler University Hospital, Linz, Austria; [30]Universitäts Thoraxklinik-Heidelberg, Heidelberg, Germany; [31]Department of Respiratory Medicine and Infectious Diseases, St. Barbara-Klinik, Hamm, Germany; [32]University of Witten-Herdecke, Witten, Germany; [33]Berlin Institute of HealthCUBI (Core Unit Bioinformatics), Berlin, Germany

ABSTRACT:

Rather than a uniform response, we hypothesized that tuberculosis (TB) patients would have distinct molecular pathways and pathologies (endotypes). Unbiased clustering identified separate TB endotypes with distinctive gene expression patterns and different clinical outcomes. A discovery cohort that included 443 TB patients and 527 asymptomatic controls identified two TB endotypes exhibiting distinct gene expression in proliferation, metabolism and immune pathways. Specifically, TB patient endotype A is characterized by increased expression of genes related to inflammation and immunity and decreased metabolism and proliferation. The endotype signatures were validated using an independent RNA-seq validation cohort consisting of two studies that included 118 TB patients and 208 controls. A second validation cohort confirmed differential pathway activity between TB endotypes and demonstrated that TB patients in endotype A had slower time to culture conversion, increased risk of death, and increased risk of poor clinical outcomes. These observations suggest that the endotypes impact functional immunity; this postulate was further supported by the observation that a hyporesponsive cytokine and chemokine responses occurred among TB patients with the hyperinflammatory endotype A compared to healthy controls. Endotype-specific host-directed therapies (HDT) might be an appropriate strategy; therefore, we examined the Library of Integrated Network Connections (LINCs) to identify drugs under consideration for TB that could be beneficial for patients with endotype A, however would be inconsequential or even harmful for patients with endotype B. Collectively, these findings suggest that metabolic and immune profiling could inform and optimize targeted endotype-specific HDT for TB.

INTRODUCTION:

With existing drug combination therapy for tuberculosis (TB) lasting at least six months, there is interest in identifying host-directed therapies (HDT) that could improve the efficacy, shorten the duration of treatment regimens, or ameliorate TB-induced lung pathology. Many HDTs have been evaluated(1), but to date, no single HDT has not shown significant benefit for all forms of TB. The identification of biologically distinct disease subtypes has revolutionized care for cancer and reactive airway disease. In these fields, biologically distinct subtypes, or endotypes, are now treated depending on the specific biologic pathways that are disturbed, with endotype-specific therapies improving clinical outcomes(2-5). For example, individuals with lung cancer are 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), and by PDL-1 expression(4). Treatment is guided by endotype categorizations and distinct molecular properties. A similar categorization system for guiding TB therapy has not yet been devised; however, new data suggests that simple clinical information may help to stratify the duration of TB therapy (6). Identification of TB endotypes should better guide a stratified based approach to HDT.

In vitro and animal studies of TB have identified incongruous immune responses that can lead to a similar TB disease phenotype (7-12). Therefore, instead of envisioning TB as an infectious disease with a single stereotypical immune response, we sought to assess the diverse host immune responses that result in a similar clinical phenotype. Immunity to *Mtb* requires a delicate homeostatic regulation, with previous studies identifying a narrow host therapeutic window with either suppressed immunity or exuberant immunity resulting in decreased *Mtb* killing capacity(10, 13, 14). Anti-mycobacterial immunity is complex, requiring a balanced, well-regulated response from 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(9, 15). Equally complex is the immune response to TNFα: decreased TNFα results in intracellular mycobacterial survival, while excess TNFα induces immune cell death and extracellular mycobacterial survival(7, 14, 16). Therefore, to identify potential endotypes of host immune response during TB, we implemented an unbiased clustering of publicly available gene expression datasets, then validated the results on two external cohorts, ultimately identifying that a hyperinflammatory TB endotype associated with worse clinical outcomes.

RESULTS

Systematic selection of TB patient cohorts with whole blood transcriptomic profiles

Seven studies, with both TB cases and controls, employed microarray gene expression analysis profiling from whole blood of 443 individuals with microbiologically confirmed TB and 527 asymptomatic healthy controls(17-23). Available TB micro-array transcriptomics studies that either did not profile whole blood or that did not include both cases and controls were not considered. These 7 studies were used to perform unbiased clustering identification of TB endotypes using 12,468 commonly evaluated genes (Supplemental Table 1). Two additional studies with both cases and controls used RNA-seq transcriptome profiling and included 118 TB patients and 208 controls(24, 25). These studies were reserved for creation of an RNA-seq validation cohort (Supplemental Table 1)(24-26). An additional cohort that included extensive clinical outcome data (Table 1) from Germany and Romania, was evaluated as an additional external validation(27).

Identification and clinical data characterization of TB endotypes

In order to identify potential TB endotypes, unbiased clustering of the 443 patients with microarray gene expression data was performed (Figure 1A-C). Endotype A consisted of 243 TB patients (54.8% of all TB cases in the microarray discovery cohort). Endotype B consisted of 200 TB

patients (45.1% of all TB cases). Patients from South Africa, Kenya, Malawi, United Kingdom and the United States were represented in both endotypes (Supplemental Table 1). In the discovery cohort, only two studies included individuals co-infected with HIV (GSE 37250, and GSE39939; Supplemental Table 1). The 108 Individuals living with HIV clustered evenly into endotype A and endotype B, with 54 in each. Only one study included children under 15 years of age (GSE39939), with 26 children clustering into endotype A and 9 clustering into endotype B (Supplemental Table 1).

The top 50 discriminatory genes (Supplemental Table 2) from the discovery cohort were used to infer a classifier to distinguish endotype A and B and this was applied to a validation cohort comprised of 118 TB patients and 208 controls from two previously published RNA-seq datasets(24, 25). By employing random forest classification with a 10-fold cross validation in the discovery cohort there was over 99% 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. In the RNA-seq validation cohort, random forest classified 42 TB samples as endotype A (35%) and 76 as endotype B (65%; Figure 1D). We computed pathway enrichment using GSEA between each of the predicted endotypes and the control samples, and also for the comparison between the two endotypes. The Random forest classification captured the significant divergent gene expression enrichments for immune related pathways. Similar to the discovery cohort, Endotype A enriched for inflammation, IFN-γ signaling, TNFα, and IL-6 pathways, while Endotype B enriched for E2F, MYC, oxidative phosphorylation and mTOR pathways (Figure 1F).

Differential clinical outcome between TB endotypes

The 50 gene random forest classification system was applied to a cohort of TB patients, (further termed the Borstel TB cohort), from a prospective, multicenter trial that enrolled patients from Germany and Romania. The Borstel TB cohort contained information on baseline bacillary burden, time to culture conversion, and clinical outcomes following the TBNET criteria (Table 1)(28, 29). Of

121 TB patients, 54 classified as endotype A (45%), while 67 classified as endotype B (55%; Figure 1E). Similar to the RNA-seq validation cohort, Endotype A and B demonstrated distinct gene expression signatures (Figure 1E-1F). The two Endotypes had similar rates of cavitary disease and initial bacillary burden (as measured by baseline time to positivity by liquid culture; Table 1). Although the initial bacterial load as measured by time to positivity (Endotype A 15 days vs. Endotype B 17.5 days, Mann Whitney p= 0.4435) and the rate of cavitary disease (Endotype A 76.3% vs. Endotype B 68.5%, Fischer exact p= 0.485) were similar in both groups, Endotype A displayed a slower time to culture clearance (63.5 days versus 49.0 days; Mann Whitney p = 0.017; Figure. 2A; Table 1) and decreased rates of favorable outcomes (70.5% versus 93.1%; Fisher's exact test p = 0.027; Figure 2B; Table 1). The poor prognosis was increased in Endotype A despite lower rates of multi-drug resistance in Endotype B (Endotype A 44.0% vs. Endotype B 74.6%, Fischer's exact test p = 0.0011). All deaths occurred in endotype A (11.7% versus 0%; Fischer's exact test p = 0.117). Since this cohort contained extensive clinical data, we were able to assess the association between pathway activity scores and time to culture conversion (TCC) for all participants with TB, as well as in each of the predicted endotypes (Figure 2C). Interestingly, in Endotype A, but not Endotype B, there was a positive and significant correlation between TCC and pathway activity scores for Hallmark Inflammatory response and TNFα signaling, respectively (Figure 2C-D).

Characterizing TB endotype gene expression signatures

Cell trajectory (also termed pseudotime trajectory) based on transcriptomic profiles is widely used to assess development and disease(30). We employed the pseudotime analytical method Monocle and identified an increase in trajectory score from healthy controls to Endotype B to Endotype A (Figure 3); this result led to a search for specific molecular properties, such as enriched pathways, that follow the predicted disease trajectory. We computed pathway activity scores for each patient in the discovery cohort. The results demonstrate that pathways related to inflammation and immunity increase in a monotonic manner (continuing in one direction) from healthy controls to endotype B to

endotype A (Figure 4), thus corroborating the cell trajectory analysis. Upon immune activation, metabolic shifts drive cellular proliferation through MYC and mTOR signaling; however, if cellular activation persists, cells become metabolically exhausted and proliferation decreases(31-33). Pathways related to metabolism and proliferation were non-monotonic: between healthy controls, Endotype B demonstrated upregulation of pathways related to metabolism and proliferation (oxidative phosphorylation, glycolysis, mTOR and the E2F pathways; Figure 4); however Endotype A patients exhibit decreased activity scores for pathways related to cellular proliferation (MYC and E2F; $p < 0.0001$) and related to metabolism (oxidative phosphorylation, glycolysis, the TCA cycle and the electron transport chain; $p < 0.001$; Figure 4). The endotypes were evaluated against two previously published scores that identified risk for TB disease severity(34, 35) and six scores that identified risk of treatment failure(34-39). Endotype A exhibited higher scores for risk of treatment failure and more severe disease, compared to Endotype B (Figure 5), validating the worse clinical outcomes observed in endotype A from the Borstel cohort (Figure 2).

Hyperinflammatory, hyporesponsive TB endotype

TB Endotype A, at the transcriptome level, was characterized by increased baseline inflammation and increased interferon, TNFα and IL-6 signaling in non-stimulated blood. To evaluate the functional differences of different endotypes, an independent cohort of 40 TB patients and 39 healthy controls from Eswatini with multiplex protein profiles was analyzed by ELISA. 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, TNFα, IL-1β, IL-6, CXCL9, and CXCL10) after whole blood stimulation to mitogen (phytohemagglutinin) (Figure 6A-B). The two TB patient sub-groups were compared to healthy controls. The hypo-responsive group demonstrated a baseline hyperinflammatory condition, similar to Endotype A, with statistically elevated baseline IFN-γ, TNFα, CXCL9, and CXCL10, but statistically decreased capacity to up-regulate IFN-γ, TNFα, IL-1β, IL-6, CXCL9, and CXCL10 upon stimulation with mitogen (Figure 6C-6D; Mann-Whitney < 0.007).

The immune-responsive group was statistically similar to the healthy controls with regard to the capacity to increase cytokine responses following stimulation. Of the 40 TB patients, there were 4 deaths, 3 in the hyperinflammatory, hypo-responsive group and 1 in the immune responsive group (chi-square 1.1, $p = 0.27$).

Comparison of chemical compounds signatures to TB endotype signatures

The Library of Integrated Network-based Cellular Signatures (LINCS) allows for the identification of putative, potentially effective drugs against a transcriptome signature. Using the gene signatures against healthy controls for endotypes A and B against healthy controls, we obtained LINCS ranked lists of over 5,000 chemical compounds for each endotype and performed a comparative analysis. Previously identified candidates for host-directed therapy (HDT), including Vitamin D, glucocorticoids, non-steroidal anti-inflammatory drugs (NSAIDS), cyclooxygenase inhibitors, and retinoids, demonstrated connectivity scores that suggest a putative benefit for one TB endotype and either an inconsequential or contradictory response in the other endotype (Figure 7). For example, HDAC inhibitors, such as vorinostat and phenylbutyrate, demonstrated transcriptomic signatures similar to endotype A, but dissimilar to endotype B.

DISCUSSION

By identifying disease subtypes based on specific molecular pathologies, or endotypes, the fields of cancer, asthma, and COPD have improved clinical outcomes through the development of endotype-specific therapies. In this study, by using unbiased clustering of unstimulated blood transcriptomes, we were able to identify TB subgroups characterized by unique gene expression signatures that have different clinical outcomes. Our analysis identified two distinct TB endotypes, termed endotype A and endotype B, both with elevated basal gene expression of the IFN-γ and TNFα pathways; but endotype A had notably higher basal levels of IFN-γ and TNFα signaling and decreased pathways related to cellular metabolism and proliferation. Endotype A patients, defined by

elevated basal hyperinflammation by gene expression, when stimulated, were hyporesponsive and experienced worse clinical outcomes.

TB is currently treated in a uniform manner; however, this data suggests that HDT options beneficial for one endotype would be inconsequential or could even be detrimental for another endotype. The most commonly evaluated candidates for correlates of protection against *Mtb* are IFN-γ and TNFα and their corresponding signaling pathways. Therefore, it is notable that while both endotypes exhibit elevated IFN-γ and TNFα at baseline compared to healthy controls, Endotype A, which is more elevated at baseline, has the least ability to upregulate IFN-γ and TNFα upon stimulation. Provided that future pair-wise transcriptomic and immune function studies demonstrate that Endotype A displays characteristics of immune exhaustion (hyperinflammatory but hyporesponsive), defined as decreased effector function and antigen-induced proliferative capacity, this TB classification could help guide endotype-specific HDTs.

The pseudotime trajectory analysis suggests immunity and inflammatory pathways have a monotonic progression from healthy controls to endotype B to endotype A, while in contrast, pathways related to metabolism and cellular proliferation display a non-monotonic progression, first increasing in endotype B, and then decreasing in endotype A. This is a pattern similar to murine models of TB and other chronic infections(33, 40) and therefore suggests a stage-specific intervention could prevent the progression to terminal immune exhaustion.

Our study is limited in its capacity to determine appropriateness of the host immune response due to limited metadata and because there are sub-optimal mechanisms for quantifying bacillary burden. Host immunity to *Mtb* must be balanced in order to avoid exuberantly induced collateral damage to host organs(7, 13, 41). Animal models have demonstrated that both IFN-γ and TNFα have narrow therapeutic windows with deficient responses allowing disease progression and exuberant responses resulting in immune-mediated pathology(7, 8, 13, 14, 16, 41, 42). The independent analysis from the Borstel TB cohort, comprised of German and Romanian patients, uses the best

currently available measurement of baseline bacillary burden (liquid culture time to positivity), and generated preliminary evidence suggesting that endotype A has a hyperinflammatory, hyporesponsive immune response with slower culture conversion and worse clinical outcomes. Prospective studies will need to combine gene expression analysis with functional immunology and quantitative measures of bacillary burden in order to clarify the appropriateness of the immune response unique to each endotype.

The link between metabolism, particularly glycolysis, and immune function has been appreciated for at least 93 years(41, 43-45). Metabolism is mediator of the epigenetic mechanisms driving immune function(40, 46-49). Therefore, it is of particular interest that the two endotypes displayed incongruent regulation of genes and pathways related to metabolism, proliferation and immune response. Many HDT candidates target these pathways. For example, metformin mediates the AKT-mTOR pathway, blunting cellular glycolysis and the TCA cycle leading to inhibition of chromatin conformational changes that ultimately drive antigen-induced immune function(46, 47, 50). These data suggest that attempts to modulate metabolic and immune mechanisms should be applied in an endotype-specific manner.

Previously identified candidates for HDTs include exogenous IFN-γ, exogenous GM-CSF, exogenous TNFα, TNFα inhibitors, NSAIDS, Vitamin D, glucocorticoids, HDAC inhibitors, mTOR modulators (rapamycin, metformin), retinoids, and statins(1). The *in-silico* analysis demonstrated that previously identified HDTs might perform better if applied in endotype-specific manners. If functional studies validate one endotype to have decreased immune responsiveness, then vitamin D or exogenous recombinant IFN-γ may be an appropriate HDT option. In contrast, if future validation studies demonstrate one endotype to have pathologic, exuberant immunity, then an NSAID, TNFα inhibitor or glucocorticoid would be appropriate. The pseudotime trajectory analysis suggests blunting endotype B progression to endotype A may be beneficial, but in-depth studies that evaluate trajectory

while considering therapy are needed. Animal and *in vitro* models are needed that 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(21, 23, 51, 52), TB-induced immune exhaustion decreases antigen-induced immune upregulation(53-56). The multiplex ELISA data highlight the limitations of inferring immune function based on baseline, non-stimulated gene expression measurements; this limitation is highlighted by the observation that the group with elevated baseline cytokine and chemokines (hyperinflammatory) was hyporesponsive upon stimulation. Therefore, future studies will need to simultaneously evaluate gene expression and the functional capacity of each endotype upon stimulation.

This study was predominantly derived from publicly available TB blood transcriptome 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 yet unappreciated host, pathogen and environment interactions. However, unlike the Cancer Genome Atlas (TCGA), very limited epidemiology is available in existing public data repositories. Epidemiologic predispositions likely driving the divergent host gene expression endotypes include malnutrition, HIV, helminth, tobacco use and or indoor biomass fuel exposure(42). For example, despite successful deworming, previous schistosomiasis infection ablates mycobacterial immunity and leaves a long-lasting detrimental epigenetic-mediated immune exhaustion(57). It is likely that individuals with pre-existing infections and preexisting immune exhaustion progress to a TB endotype distinct from the previously healthy individual who develops TB after high inoculum.

Another limitation of this research is the lack of cell-specific data. A defining feature of HIV infection is decreased cell-mediated immunity(58), particularly decreased IFN-γ production from

lymphocytes. Considering previous evidence that neutrophils drive the gene expression signatures of TB patients (17), it is likely that cell-specific gene expression evaluations of endotype A would identify decreased lymphocyte IFN-γ responsiveness, but might identify retained non-lymphocyte IFN-γ signaling. The current trajectory data suggests that TB patients progress from healthy controls to endotype B to endotype A. Considering the long duration of TB symptoms, it can be hypothesized that endotype states are dynamic and the trajectory analysis supports transition from endotype B to A and then back to B upon successful therapy.

In summary, unbiased clustering analysis of transcriptome and protein data from TB patients provides strong evidence for biologically distinct TB endotypes that differentially affect clinical outcomes. Specifically, host gene expression in TB patient clusters into two major, distinct endotypes with differential immune and metabolic transcriptomic signatures. These observations suggest that despite similar clinical phenotypes, different TB patient endotypes display responses that are likely to have clinical and pathologic relevance. Similar to endotype-specific therapy for asthma, COPD, and cancer (2-5), patients belonging to each TB endotype are likely to benefit from distinctive host-directed therapies.

METHODS

Additional details on the methods is provided in an online supplement.

Study inclusion

We implemented a systematic review and individual participant meta-analysis. The study is reported according to the Preferred Reporting Items for a Review and meta-analysis of individual patient data (PRISMA-IPD). An a priori set of hypotheses were established 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. Twenty-two gene expression studies from whole blood were identified that included participants with microbiologically confirmed pulmonary TB. Studies that included only cases without controls and those evaluating less than 10,000 genes were excluded. Studies using microarray were used for a discovery cohort and genome-wide (RNA-seq) data sets were used for a validation cohort (Supplemental Table 1). The ComBat method was used to normalize the data(59). An additional micro-array validation cohort was identified, with extensive clinical annotations (Table 1; Supplemental Table 1)(27).

Clustering

To identify distinct TB endotypes, we used the Louvain network-based clustering algorithm implemented in the single-cell RNA-seq analysis package Seurat. 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

Given the fundamental difference between the microarray and RNA-seq profiling platforms, we normalized independently the discovery cohort microarray data, the RNA-seq data, and the Borstel micro-array data, and used the former as the training cohort and the last two as validation cohorts. A binary classifier for endotype A and B was built using RandomForest.

Pseudotime

Transcriptome profile trajectory analysis (pseudotime) was implemented on the 1244 TB patients and healthy controls using the Monocle algorithm.

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)(56). 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 the TB cases and household contacts were HIV co-infected. 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 TB patients. Differences between sub-groups were analyzed using Mann-Whitney rank sum test.

REFERENCES

1. Kaufmann SHE, Dorhoi A, Hotchkiss RS, and Bartenschlager R. Host-directed therapies for bacterial and viral infections. *Nat Rev Drug Discov.* 2018;17(1):35-56.
2. Barnes PJ. Inflammatory endotypes in COPD. *Allergy.* 2019;74(7):1249-56.
3. Svenningsen S, and Nair P. Asthma Endotypes and an Overview of Targeted Therapy for Asthma. *Front Med (Lausanne).* 2017;4:158.
4. Dong J, Li B, Lin D, Zhou Q, and Huang D. Advances in Targeted Therapy and Immunotherapy for Non-small Cell Lung Cancer Based on Accurate Molecular Typing. *Front Pharmacol.* 2019;10:230.
5. Figueroa ME, Lugthart S, Li Y, Erpelinck-Verschueren C, Deng X, Christos PJ, et al. DNA methylation signatures identify biologically distinct subtypes in acute myeloid leukemia. *Cancer Cell.* 2010;17(1):13-27.
6. Imperial MZ, Nahid P, Phillips PPJ, Davies GR, Fielding K, Hanna D, et al. A patient-level pooled analysis of treatment-shortening regimens for drug-susceptible pulmonary tuberculosis. *Nature medicine.* 2018;24(11):1708-15.
7. Tobin DM, Roca FJ, Oh SF, McFarland R, Vickery TW, Ray JP, et al. Host genotype-specific therapies can optimize the inflammatory response to mycobacterial infections. *Cell.* 2012;148(3):434-46.
8. Sakai S, Kauffman KD, Sallin MA, Sharpe AH, Young HA, Ganusov VV, et al. 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 pathogens.* 2016;12(5):e1005667.

9. Bustamante J, Boisson-Dupuis S, Abel L, and 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.
10. Ni Cheallaigh C, Sheedy FJ, Harris J, Munoz-Wolf N, Lee J, West K, et al. A Common Variant in the Adaptor Mal Regulates Interferon Gamma Signaling. *Immunity.* 2016;44(2):368-79.
11. Andrade BB, Pavan Kumar N, Amaral EP, Riteau N, Mayer-Barber KD, Tosh KW, et al. Heme Oxygenase-1 Regulation of Matrix Metalloproteinase-1 Expression Underlies Distinct Disease Profiles in Tuberculosis. *Journal of immunology.* 2015;195(6):2763-73.
12. DiNardo AR NT, Grimm SL, Schlesinger LS, Graviss EA, Cirillo JD, Coarfa C, Mandalakas AM, Heyckendorf, Kaufmann SHE, Lange C, Netea MG, van Crevel R. Tuberculosis Endotypes to Guide Stratified Host-Directed Therapy. *Med.* 2021.
13. Barber DL, Mayer-Barber KD, Feng CG, Sharpe AH, and Sher A. CD4 T cells promote rather than control tuberculosis in the absence of PD-1-mediated inhibition. *Journal of immunology.* 2011;186(3):1598-607.
14. Tobin DM, Vary JC, Jr., Ray JP, Walsh GS, Dunstan SJ, Bang ND, et al. The lta4h locus modulates susceptibility to mycobacterial infection in zebrafish and humans. *Cell.* 2010;140(5):717-30.
15. Orange JS, Levy O, Brodeur SR, Krzewski K, Roy RM, Niemela JE, et al. Human nuclear factor kappa B essential modulator mutation can result in immunodeficiency without ectodermal dysplasia. *The Journal of allergy and clinical immunology.* 2004;114(3):650-6.
16. Roca FJ, and Ramakrishnan L. TNF dually mediates resistance and susceptibility to mycobacteria via mitochondrial reactive oxygen species. *Cell.* 2013;153(3):521-34.
17. Berry MP, Graham CM, McNab FW, Xu Z, Bloch SA, Oni T, et al. An interferon-inducible neutrophil-driven blood transcriptional signature in human tuberculosis. *Nature.* 2010;466(7309):973-7.
18. Bloom CI, Graham CM, Berry MP, Wilkinson KA, Oni T, Rozakeas F, 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, 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, et al. Blood Transcriptional Biomarkers for Active Tuberculosis among Patients in the United States: a Case-Control Study with Systematic Cross-Classifier Evaluation. *Journal of clinical microbiology.* 2016;54(2):274-82.
21. Anderson ST, Kaforou M, Brent AJ, Wright VJ, Banwell CM, Chagaluka G, et al. Diagnosis of childhood tuberculosis and host RNA expression in Africa. *The New England journal of medicine.* 2014;370(18):1712-23.
22. Kaforou M, Wright VJ, Oni T, French N, Anderson ST, Bangani N, et al. Detection of tuberculosis in HIV-infected and -uninfected African adults using whole blood RNA expression signatures: a case-control study. *PLoS medicine.* 2013;10(10):e1001538.
23. Blankley S, Graham CM, Turner J, Berry MP, Bloom CI, Xu Z, et al. The Transcriptional Signature of Active Tuberculosis Reflects Symptom Status in Extra-Pulmonary and Pulmonary Tuberculosis. *PloS one.* 2016;11(10):e0162220.
24. Singhania A, Verma R, Graham CM, Lee J, Tran T, Richardson M, et al. A modular transcriptional signature identifies phenotypic heterogeneity of human tuberculosis infection. *Nat Commun.* 2018;9(1):2308.
25. Leong S, Zhao Y, Joseph NM, Hochberg NS, Sarkar S, Pleskunas J, et al. Existing blood transcriptional classifiers accurately discriminate active tuberculosis from latent infection in individuals from south India. *Tuberculosis.* 2018;109:41-51.
26. Suliman S, Thompson E, Sutherland J, Weiner Rd J, Ota MOC, Shankar S, et al. Four-gene Pan-African Blood Signature Predicts Progression to Tuberculosis. *American journal of respiratory and critical care medicine.* 2018.
27. Heyckendorf J, Marwitz S, Reimann M, Avsar K, DiNardo AR, Guenther G, et al. A 22-gene transcriptomic model indicating individual therapy durations in multidrug-resistant tuberculosis. *European Respiratory Journal.* 2021;In Press.

28. Heyckendorf J, Marwitz S, Reimann M, Avsar K, DiNardo AR, Guenther G, et al. A 22-gene transcriptomic model indicating individual therapy durations in multidrug-resistant tuberculosis. *medRxiv.* 2020:2020.08.21.20177238.
29. Gunther G, Lange C, Alexandru S, Altet N, Avsar K, Bang D, et al. Treatment Outcomes in Multidrug-Resistant Tuberculosis. *The New England journal of medicine.* 2016;375(11):1103-5.
30. Mukherjee S, Heath L, Preuss C, Jayadev S, Garden GA, Greenwood AK, et al. Molecular estimation of neurodegeneration pseudotime in older brains. *Nat Commun.* 2020;11(1):5781.
31. Palsson-McDermott EM, Curtis AM, Goel G, Lauterbach MAR, Sheedy FJ, Gleeson LE, 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, and Pearce EL. Metabolic Instruction of Immunity. *Cell.* 2017;169(4):570-86.
33. Russell SL, Lamprecht DA, Mandizvo T, Jones TT, Naidoo V, Addicott KW, et al. Compromised Metabolic Reprogramming Is an Early Indicator of CD8(+) T Cell Dysfunction during Chronic Mycobacterium tuberculosis Infection. *Cell Rep.* 2019;29(11):3564-79 e5.
34. Penn-Nicholson A, Mbandi SK, Thompson E, Mendelsohn SC, Suliman S, Chegou NN, et al. RISK6, a 6-gene transcriptomic signature of TB disease risk, diagnosis and treatment response. *Sci Rep.* 2020;10(1):8629.
35. Warsinske HC, Rao AM, Moreira FMF, Santos PCP, Liu AB, Scott M, et al. Assessment of Validity of a Blood-Based 3-Gene Signature Score for Progression and Diagnosis of Tuberculosis, Disease Severity, and Treatment Response. *JAMA Netw Open.* 2018;1(6):e183779.
36. Ronacher K, Chegou NN, Kleynhans L, Djoba Siawaya JF, du Plessis N, Loxton AG, et al. Distinct serum biosignatures are associated with different tuberculosis treatment outcomes. *Tuberculosis.* 2019;118:101859.
37. Sivakumaran D, Jenum S, Vaz M, Selvam S, Ottenhoff THM, Haks MC, et al. Combining host-derived biomarkers with patient characteristics improves signature performance in predicting tuberculosis treatment outcomes. *Commun Biol.* 2020;3(1):359.
38. Thompson EG, Du Y, Malherbe ST, Shankar S, Braun J, Valvo J, et al. Host blood RNA signatures predict the outcome of tuberculosis treatment. *Tuberculosis.* 2017;107:48-58.
39. Kumar NP, Moideen K, Nancy A, Viswanathan V, Thiruvengadam K, Nair D, et al. Plasma chemokines are baseline predictors of unfavorable treatment outcomes in pulmonary tuberculosis. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America.* 2020.
40. Bengsch B, Johnson AL, Kurachi M, Odorizzi PM, Pauken KE, Attanasio J, 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.
41. Nair S, Huynh JP, Lampropoulou V, Loginicheva E, Esaulova E, Gounder AP, et al. Irg1 expression in myeloid cells prevents immunopathology during M. tuberculosis infection. *The Journal of experimental medicine.* 2018;215(4):1035-45.
42. DiNardo AR NT, Grimm SL, Schlesinger LS, Edward A. Graviss, Jeffrey D. Cirillo, Cristian Coarfa, Anna M. Mandalakas, Jan Heyckendorf, Stefan H.E. Kaufmann,Christoph Lange, Mihai G. Netea, and Reinout Van Crevel. Tuberculosis Endotypes to Guide Stratified Host-Directed Therapy, . *Med, Cell.* 2021.
43. Ratzan KR, Musher DM, Keusch GT, and 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. *Infection and immunity.* 1972;5(4):499-504.
44. Kelly B, and O'Neill LA. Metabolic reprogramming in macrophages and dendritic cells in innate immunity. *Cell Res.* 2015;25(7):771-84.
45. Warburg O, Wind F, and Negelein E. The Metabolism of Tumors in the Body. *J Gen Physiol.* 1927;8(6):519-30.
46. Cheng SC, Quintin J, Cramer RA, Shepardson KM, Saeed S, Kumar V, et al. mTOR- and HIF-1alpha-mediated aerobic glycolysis as metabolic basis for trained immunity. *Science.* 2014;345(6204):1250684.
47. Arts RJW, Carvalho A, La Rocca C, Palma C, Rodrigues F, Silvestre R, et al. Immunometabolic Pathways in BCG-Induced Trained Immunity. *Cell Rep.* 2016;17(10):2562-71.

48. Scharping NE, Menk AV, Moreci RS, Whetstone RD, Dadey RE, Watkins SC, et al. The Tumor Microenvironment Represses T Cell Mitochondrial Biogenesis to Drive Intratumoral T Cell Metabolic Insufficiency and Dysfunction. *Immunity.* 2016;45(2):374-88.
49. Barili V, Fisicaro P, Montanini B, Acerbi G, Filippi A, Forleo G, et al. Targeting p53 and histone methyltransferases restores exhausted CD8+ T cells in HCV infection. *Nat Commun.* 2020;11(1):604.
50. Bohme J, Martinez N, Li S, Lee A, Marzuki M, Tizazu AM, et al. Metformin enhances anti-mycobacterial responses by educating CD8+ T-cell immunometabolic circuits. *Nat Commun.* 2020;11(1):5225.
51. Zak DE, Penn-Nicholson A, Scriba TJ, Thompson E, Suliman S, Amon LM, et al. A blood RNA signature for tuberculosis disease risk: a prospective cohort study. *Lancet.* 2016;387(10035):2312-22.
52. Gupta RK, Turner CT, Venturini C, Esmail H, Rangaka MX, Copas A, et al. Concise whole blood transcriptional signatures for incipient tuberculosis: a systematic review and patient-level pooled meta-analysis. *The Lancet Respiratory medicine.* 2020.
53. Harari A, Rozot V, Enders FB, Perreau M, Stalder JM, Nicod LP, et al. Dominant TNF-alpha+ Mycobacterium tuberculosis-specific CD4+ T cell responses discriminate between latent infection and active disease. *Nature medicine.* 2011;17(3):372-6.
54. Sahiratmadja E, Alisjahbana B, de Boer T, Adnan I, Maya A, Danusantoso H, 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. *Infection and immunity.* 2007;75(2):820-9.
55. Singh A, Mohan A, Dey AB, and Mitra DK. Inhibiting the programmed death 1 pathway rescues Mycobacterium tuberculosis-specific interferon gamma-producing T cells from apoptosis in patients with pulmonary tuberculosis. *The Journal of infectious diseases.* 2013;208(4):603-15.
56. DiNardo A, Rajapakshe K, Nishiguchi T, Mtetwa G, Grimm SL, Dlamini Q, et al. DNA hyper-methylation during Tuberculosis dampens host immune responsiveness. *The Journal of clinical investigation.* 2020.
57. DiNardo AR, Nishiguchi T, Mace EM, Rajapakshe K, Mtetwa G, Kay A, et al. Schistosomiasis Induces Persistent DNA Methylation and Tuberculosis-Specific Immune Changes. *Journal of immunology.* 2018.
58. Aaron L, Saadoun D, Calatroni I, Launay O, Memain N, Vincent V, et al. Tuberculosis in HIV-infected patients: a comprehensive review. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases.* 2004;10(5):388-98.
59. Johnson WE, Li C, and Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics.* 2007;8(1):118-27.

Graphic Abstract:

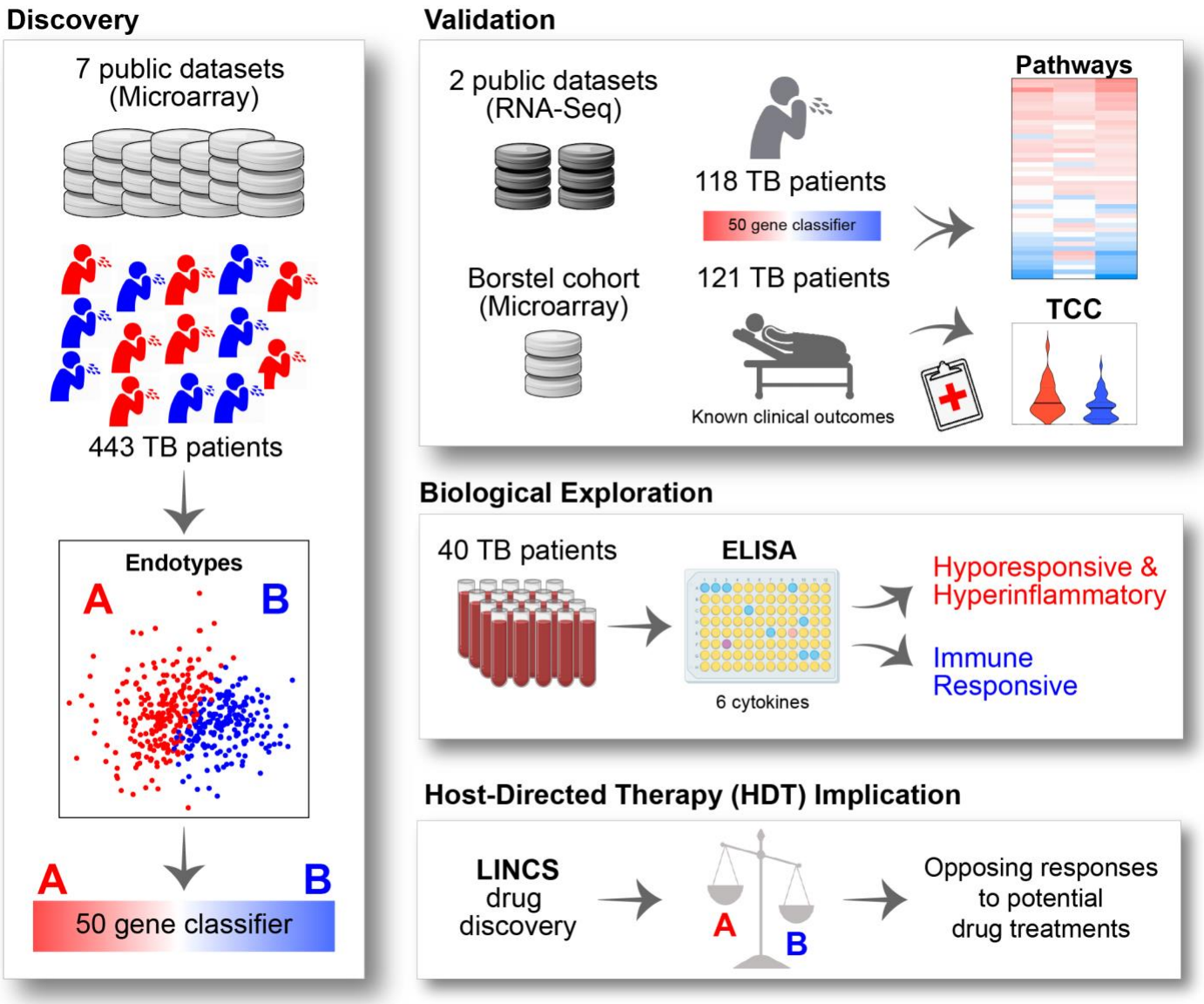

Figure 1: Unbiased clustering identifies unique TB endotypes. (A) Unbiased clustering was implemented on discovery cohort of 7 studies, identifying two major endotypes; next, a 50 gene classifier was developed and applied to two external validation cohorts. (B) Network-based unbiased clustering using the Louvain method identifies 2 major endotypes of TB. (C) Distribution of individual studies into Endotype A or B. (D-E) Clustering of the validation cohorts by endotype A or B. (F) TB endotypes were compared to healthy controls and against each other, then pathway enrichment via Gene Set Enrichment Analysis (GSEA) was carried out against the Hallmark pathway compendium.

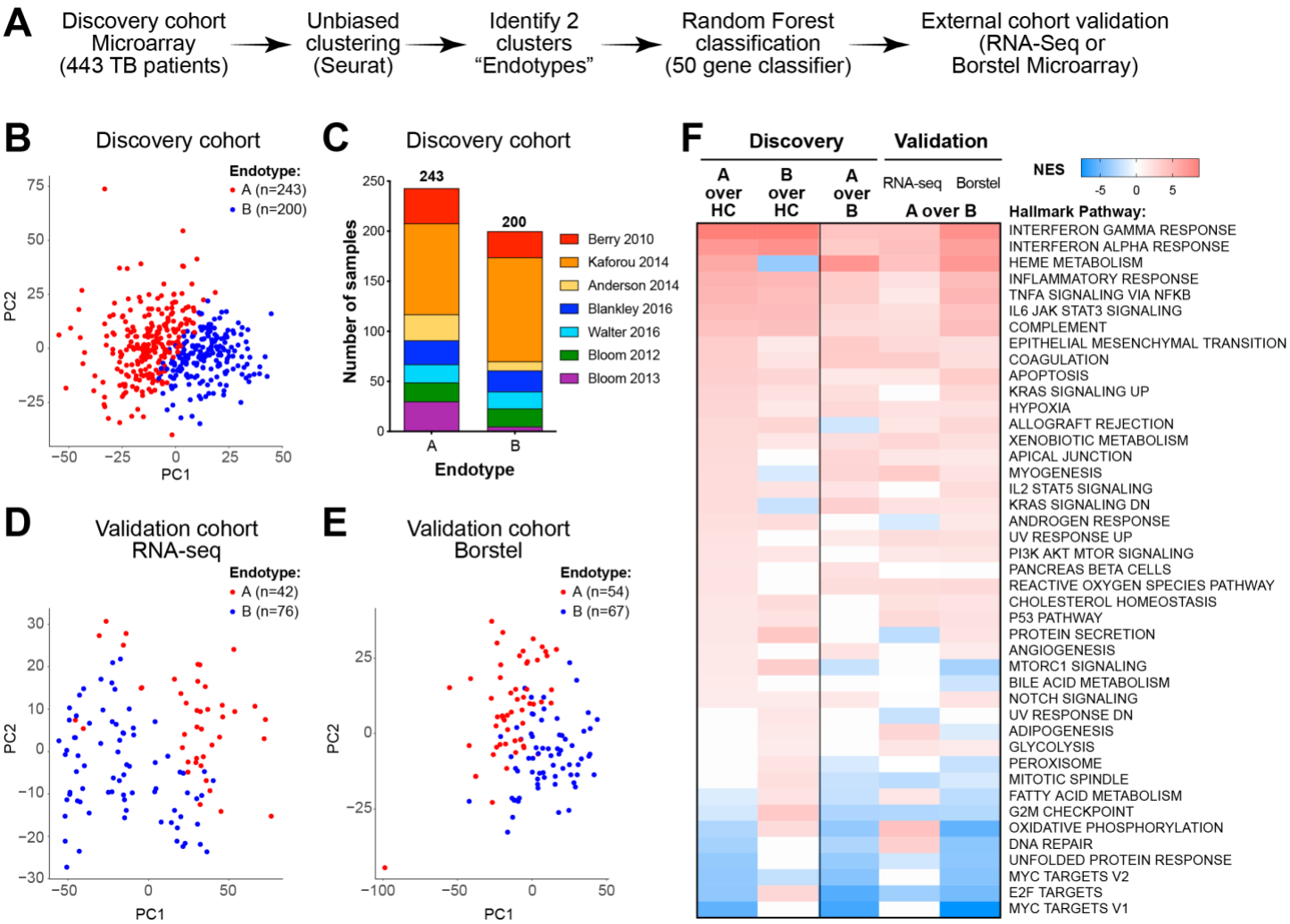

Figure 2: **Endotype evaluation of TB clinical outcomes**. Using the clinical annotations of the Borstel TB cohort, outcome differences between endotypes and association of pathway scores with outcomes were evaluated. (A) Time to culture conversion (TCC) in TB patients identified as endotype A or B (p = 0.017 by Mann Whitney). (B) Rates of treatment failure in TB patients identified as endotype A or B (p = 0.027 by Fisher's exact test). (C). Heatmap depicting significant Pearson correlation coefficients between clinical outcomes and pathway activity score. (D) Scatter plots of the Inflammatory and TNF signaling pathway activity score and TCC by endotype A and B.

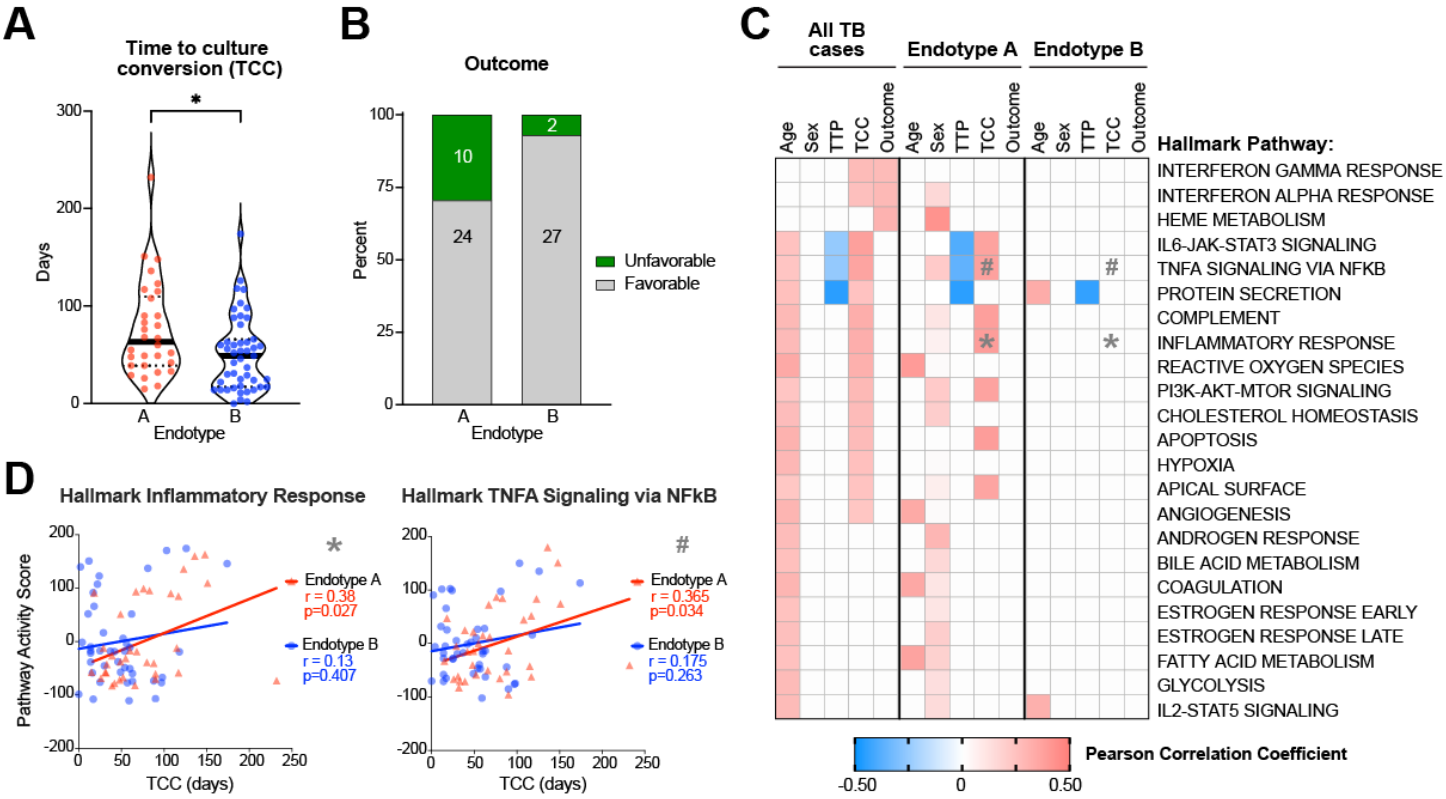

Figure 3: Pseudotime trajectory analysis of healthy controls and TB endotypes. (A)-(B) Pseudotime analysis stratifies the patients and indicates a progression from healthy controls (HC, grey) to the TB endotype B (blue) and then to the TB endotype A (red). C) Cell trajectory analysis of the 1244 TB cases and controls using the 12,468 commonly evaluated genes assigned a pseudotime to each patient.

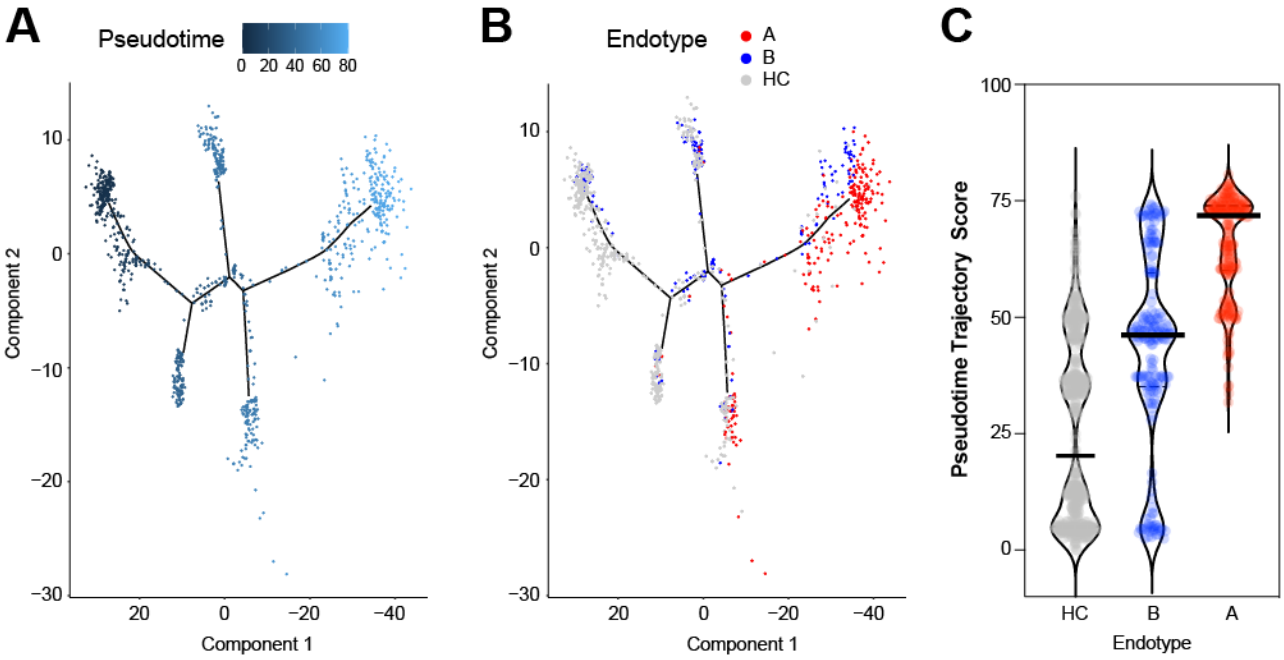

Figure 4: TB endotypes display distinct immune and metabolic gene expression activity scores. Pathway activity scores were evaluated between healthy controls (HC), endotype B (B) and endotype A (A). Inflammation and immunity pathways (A), Metabolic pathways (B), and Proliferation pathways (C). Mann-Whitney test was used; **** $p < 0.0001$, *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, ns = non-significant. Specific gene changes are available in Supplemental Table 2.

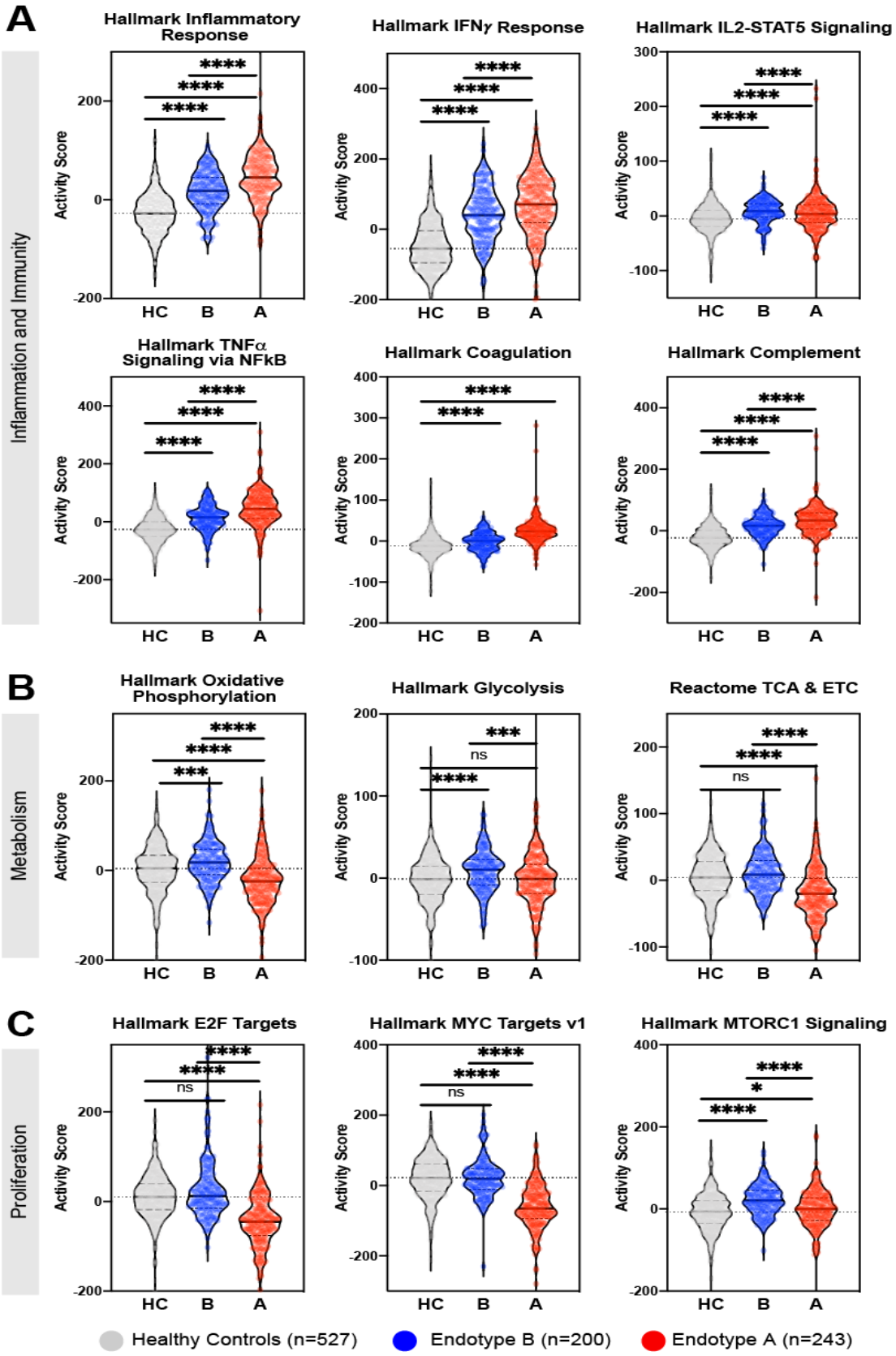

Figure 5: Evaluation of TB risk signatures over the TB endotypes. Previous studies identified gene signatures associated with risk of treatment failure or with disease severity and treatment failure. Activity scores were computed for each risk signatures across healthy controls (HC) and Endotype A (A) and Endotype B (B). Mann-Whitney scores with **** $p < 0.0001$, *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, ns = non-significant.

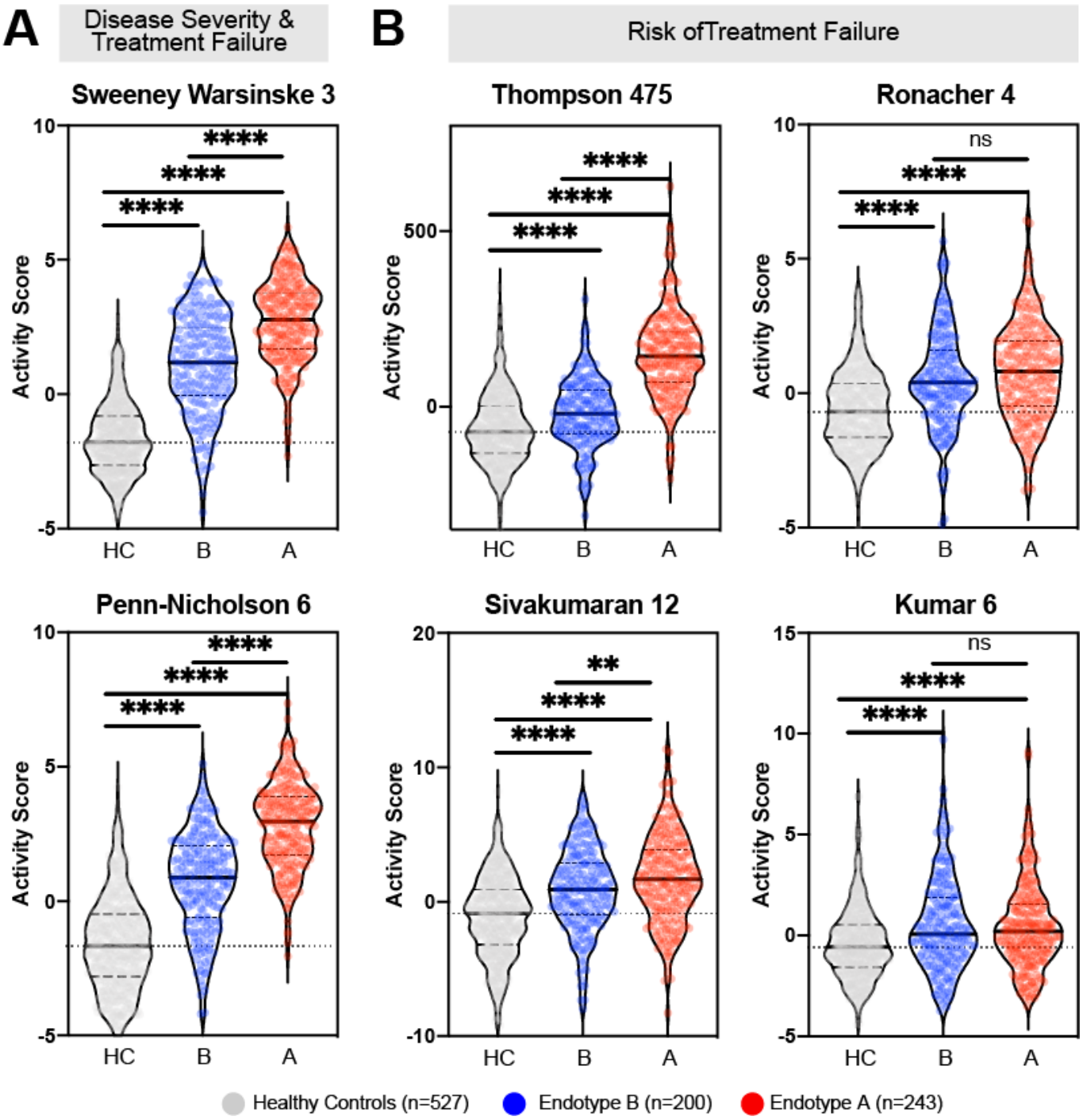

Figure 6: Identification of hyper-inflammatory, hyporesponsive cytokine production in TB patient endotypes. Whole blood from TB patients (n = 40) and healthy controls (n = 39) was stimulated overnight with or without mitogen (PHA), followed by measurement of cytokines and chemokines. (A) Samples were ranked for up-regulation 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" (red) and "responsive" (blue) groups. (B) Heatmap of cytokine expression as log2 fold change relative to controls. (C) Absolute protein expression with nil (N) or mitogen stimulation (S) expressed in pg/mL. (D) Cytokine protein expression (log2 fold change) is graphed for each sub-group. Significance determined by Mann-Whitney non-parametric test.

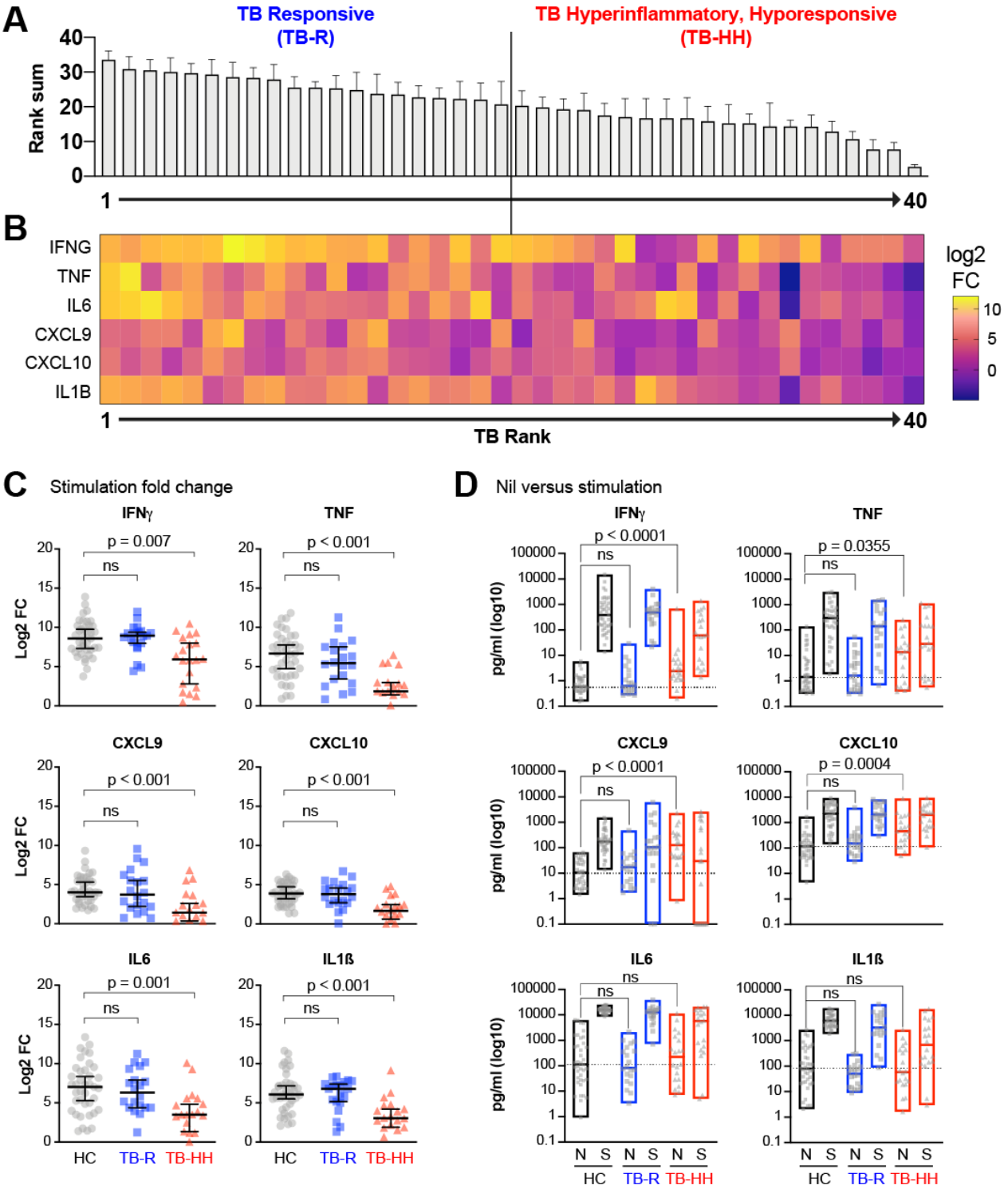

Figure 7. Heatmap of connectivity score for select chemical compounds within the TB endotypes A and B based on the Library of Integrated Network-based Cellular Signatures (LINCS). Positive connectivity scores represent chemical inducing gene expression profiles similar to the endotype, while negative connectivity scores represent chemical compounds inducing gene expression profiles antithetical to the endotype.

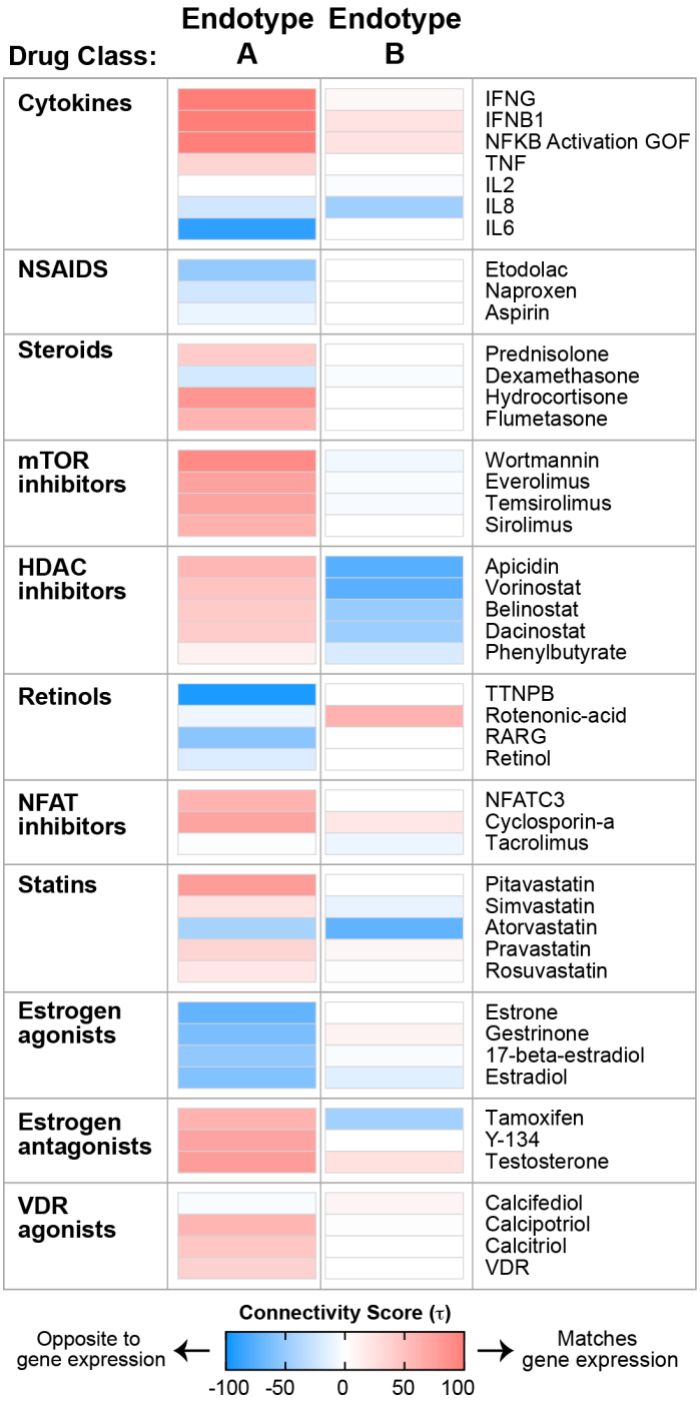

Table 1: Epidemiologic characteristics of the German and Romanian validation cohort. (MDR = multi-drug resistance; TTP = time to positivity; BMI = body mass index; TCC = time to culture conversion). (Not all data was available for all participants.) Mann-Whitney test performed for continuous variables and fisher's exact test evaluated for differences in populations. Not all data was available for all participants.

Table 1. Epidemiologic characteristics of the German and Romanian validation cohort.

	All	Endotype A	Endotype B	p-value
Cavitary Disease	71.7% (66/92)	76.3% (29/39)	68.5% (37/54)	0.485
MDR TB	61.9% (75/121)	44% (22/50)	74.6% (53/71)	0.001
Baseline TTP (days)	17	15	17.5	0.443
Current or previous smoker	53.2% (33/62)	41.1% (7/17)	57.7% (26/45)	0.268
Median BMI	20.7	18.8	21.2	0.144
% Male	62.2% (56/90)	65.7% (23/35)	60.0% (33/55)	0.658
Median age	37.9	42.9	37.6	0.253
Median TCC (days)	54	63.5	49	0.017
Favorable outcome	80.9% (51/63)	70.5% (24/34)	93.1% (27/29)	0.027
Mortality	6.3% (4/63)	11.7% (4/34)	0% (0/29)	0.117

MDR, multi-drug resistance; TTP, time to positivity; BMI, body mass index; TCC, time to culture conversion. (Not all data was available for all participants).

Supplemental Tables:

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

Supplemental Table 2. A: Differential gene expression of genes with 2-fold log change and $p < 0.05$. B: Differentially enriched Hallmark pathways. C: 50 gene classifier.

Supplemental Table 3. A: Differential pathway analysis from Hallmark, KEGG and Reactome.