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], Abhimanyu[1], 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: Cristian Coarfa PhD: coarfa@bcm.edu

Running head: Distinct host gene expression endotypes in TB

Word count: 2996 words; 6 Figures; 1 Table

The authors have declared that no conflict of interest exists.

Funding Sources: ARD is supported by NIAID K23 AI141681-02. SLG and CC are supported by the Cancer Prevention Institute of Texas (CPRIT) RP170005, NIH P30 shared resource grant CA125123, and NIEHS grants P30 ES030285 and P42 ES027725. 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.

[x]**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:

Background*: In vitro*, animal model, and clinical evidence suggests that tuberculosis is not a monomorphic disease, and that host response to tuberculosis is protean with multiple distinct molecular pathways and pathologies (endotypes). We applied unbiased clustering to identify separate tuberculosis endotypes with classifiable gene expression patterns and clinical outcomes.

Methods: A cohort comprised of microarray gene expression data from microbiologically confirmed tuberculosis patients were used to identify putative endotypes. One microarray cohort with longitudinal clinical outcomes was reserved for validation, as was one RNA-seq cohorts. Finally, a separate cohort of tuberculosis patients with functional immune results was evaluated to clarify stimulated from unstimulated immune responses.

Results: A discovery cohort, including 435 tuberculosis patients and 533 asymptomatic controls, identified two tuberculosis endotypes. Tuberculosis patient endotype A is characterized by increased expression of genes related to inflammation and immunity and decreased metabolism and proliferation; in contrast, endotype B increased activity of metabolism and proliferation pathways. An independent RNA-seq validation cohort, including 118 tuberculosis patients and 179 controls, validated the discovery results. Gene expression signatures for treatment failure were elevated in endotype A in the discovery cohort, and a separate validation cohort confirmed that endotype A patients had slower time to culture conversion, and a reduced incidence of cure. These observations suggest that endotypes reflect functional immunity, supported by the observation that tuberculosis patients with a hyperinflammatory endotype have less responsive cytokine production upon stimulation.

Conclusion: These findings provide evidence that metabolic and immune profiling could inform optimization of endotype-specific host-directed therapies for tuberculosis.

INTRODUCTION:

Host-directed therapies (HDT) could improve the efficacy, shorten the duration of treatment regimens, or ameliorate tuberculosis (TB)-induced lung pathology. The fields of asthma, COPD, and most cancers have identified biological endotypes, distinct cellular pathology. Consequently, treatment for these diseases depends on the specific pathways that are disturbed, with endotype-specific therapies improving clinical outcomes[1]. For example, asthma endotypes can generally be divided into neutrophilic versus eosinophilic disease, with the former being more responsive to macrolide therapy and the latter being corticosteroid responsive [1]. Leprosy is also treated based on immune endotypes, with the cell-mediated and paucibacillary form requiring less antibiotics for a shorter duration, while the anergic and multibacillary endotype requires more antibiotics for a longer duration. No similar categorization system is available to guide TB HDT.

Studies have identified incongruous immune responses that can lead to TB[2-7]. To further test the premise that there is not a single stereotypical immune response to TB, we sought to provide evidence for the diversity of host responses during TB. Mycobacterial immunity requires a balanced, well-regulated response from multiple cell types. For example, immune control of *Mycobacterium tuberculosis* (*Mtb)* requires tightly regulated IFN-γ and TNF□ with decreased IFN-γ or TNF□ resulting in decreased *Mtb* intracellular killing capacity[4], but with exuberant IFN-γ or TNF□ inducing macrophage and tissue necrosis with extracellular *Mtb* survival[2, 3]. To better characterize TB endotypes, we implemented an unbiased clustering of publicly available gene expression data, then validated the results using two external cohorts, eventually identifying a hyperinflammatory TB endotype associated with worse clinical outcomes.

METHODS:

Study inclusion

A systematic review and meta-analysis were implemented according to PRISMA guidelines (described in detail in supplemental methods). Publicly available data was identified using PubMed and the NCBI Gene Expression Omnibus (GEO) repository. Studies without microbiologic confirmation, without description of the methods of microbiologic confirmation, or without evaluation of whole blood were excluded. Twenty-two gene expression studies from whole blood were identified that included participants with microbiologically confirmed pulmonary TB. Studies or datasets that included only cases without controls or evaluated fewer than 10,000 genes were excluded. Studies using microarray were used for a discovery cohort and RNA-seq datasets for a validation cohort (Supplemental Table 1). One microarray cohort (GSE147689, GSE147691) was reserved for validation because it contained longitudinal clinical outcomes (Table 1; Supplemental Table 1). Normalization, processing, clustering, and development of a gene classifier is described in full in the supplemental methods. Clinical outcomes for this cohort[8] were defined using the TBNET criteria: cure is defined as culture-negative at 6 months, with no positive cultures thereafter and no disease relapse within a year after treatment completion. Treatment failure was based on a positive culture 6 months after treatment initiation or disease relapse within one year after treatment completion.

Immunology Validation Cohort

Multiplex ELISA (LegendPlex) was implemented with and without overnight mitogen stimulation in a cohort of pulmonary TB patients (n = 40) or their asymptomatic household contacts (n = 39) from Eswatini. The study protocol was reviewed by institutional review boards and all participation was voluntary and in concordance with the Declaration of Helsinki. TB patients were defined by both the presence of symptoms and microbiologic confirmation by culture or Gene Xpert. Twenty-two (55%) of the TB cases and nine (23%) of the household contacts were HIV co-infected, respectively.

Statistics

Chi-squared test with a one-sided tail assessed incidence of clinical variables between endotypes. Rank-sum of the cytokines/chemokines from the ELISA validation was used to stratify TB patients. Differences between sub-groups were analyzed using Mann-Whitney rank sum test.

RESULTS

Systematic selection of TB patient cohorts with whole blood transcriptomic profiles

Seven studies profiled whole blood and included both cases and controls by microarray gene expression analysis, profiling 435 individuals with microbiologically confirmed TB and 533 healthy controls[9-15]. These 7 studies, comprising 12 datasets in NCBI GEO, were used for unbiased clustering to identify TB endotypes, based on 12,468 commonly evaluated genes (Figure 1; Supplemental Table 1). Two additional studies with both cases and controls used RNA-seq transcriptome profiling and included 118 TB patients and 179 controls[16, 17]. These studies were reserved for creation of an RNA-seq validation cohort (Supplemental Table 1). A cohort from Germany and Romania including 121 TB patients and 14 healthy controls using microarrays was reserved as an additional external validation dataset because it included clinical outcome data (Table 1)[18]. A fourth cohort of TB patients and healthy controls with multiplex ELISA at baseline and phytohemagglutinin stimulated whole blood (Figure 1).

Identification and clinical data characterization of TB endotypes

To identify potential TB endotypes, unbiased clustering of the microarray transcriptome of 435 TB patients was performed (Figure 2A-C). Clustering was evaluated at various resolutions (Supplemental Figure 1) with final analysis being performed at resolution 0.4 due to limited discriminatory capacity at higher resolutions. At resolution 0.4, two distinct endotypes were identified. Endotype A consisted of 269 TB patients (54.8%), and endotype B included 166 TB patients (45.1%). Patients from each country and each study were well distributed in both endotype clusters (Figure 2C; Supplemental Table 1). In the discovery cohort, only two datasets included individuals co-infected with HIV (GSE37250, and GSE39939; Supplemental Table 1). The 108 TB patients living with HIV clustered into endotype A and endotype B, with 57 and 51 in each respectively. Only one study included children under 15 years of age (GSE39939), with 23 children clustering into endotype A and 12 clustering into endotype B (Supplemental Table 1).

Using the random forest machine learning algorithm, a classifier was derived to categorize endotype A vs endotype B based on the discovery cohort. Genes were ranked by their individual classification score, then accuracy of a random forest gene classifier was evaluated across a spectrum of genes using the measured classification error (Supplementary methods; Supplemental Figure 2A). A 40-gene classifier had both a low misclassification rate and was comprised of a low gene count. We applied the endotype classifier to a validation cohort comprised of 118 TB patients and 179 controls from two previously published RNA-seq studies (Supplemental Table 1). In the RNA-seq validation cohort, 65 TB samples classified as endotype A (55%) and 53 as endotype B (45%; Supplemental Table 1). We computed pathway enrichment using GSEA between each of the predicted endotypes and the control samples, and for the comparison between the endotypes. Random forest classification captured the significant divergent gene expression enrichments for immune related pathways (Figure 2D, Supplemental Figure 2B). For one study, microarray profiles (GSE19442 and GSE19444) were used in the discovery dataset whereas RNA-seq profiles (GSE107991) for the same samples were also used in the RNA-Seq validation cohort. Our classifier achieved good concordance between the microarray discovery cohort endotypes and the predicted RNA-seq endotypes (Supplemental Figure 3). Similar to the discovery cohort, endotype A enriched for inflammation, IFN-γ signaling, TNF□, and heme metabolism, while endotype B enriched for pathways related to cellular proliferation including E2F, G2M and mitosis (Figure 2D). Compared to healthy controls, gene targets of the transcription factors E2F, ELK1, and NRF1 were increased in endotype B, but were decreased in endotype A (Supplemental Table 3; Supplemental Figure 4). The endotypes were evaluated against six previously published scores that identified risk of treatment failure, with two of the scores also identifying risk for TB disease severity (Supplemental methods). Endotype A exhibited higher scores for risk of treatment failure and more severe disease compared to healthy controls (Figure 3).

Differential clinical outcome between TB endotypes

The classifier was applied to a cohort of TB patients assayed using microarray from a prospective, multicenter trial in Germany and Romania[18]. This cohort contained information on baseline bacillary burden, time to culture conversion, and clinical outcomes defined by the TBNET criteria (Table 1). Of 121 TB patients, 64 classified as endotype A (53%), while 57 classified as endotype B (47%). Similar to the RNA-seq validation cohort, endotype A and B demonstrated distinct enriched pathway profiles (Figure 2D, Supplemental Figure 2B). Based on the increased predicted treatment failure and disease severity signatures (Figure 3), we hypothesized that endotype A patients will display worse clinical outcome compared to the endotype B patients. Whereas endotype A had a lower rate of multi-drug resistance (54.7% vs 70.2%), Endotype A had slightly increased rates of cavitary disease (endotype A 76.5% vs. endotype B 65.9%, p= 0.1305), higher initial bacterial load (14 versus 21 days; p= 0.0169), slower times to culture clearance (64.6 days versus 33.5 days, p = 0.0005; Figure 4A; Table 1) and decreased rates of cure outcomes (74.4% versus 91.7%, p = 0.0447; Figure 4B; Table 1). All deaths occurred in endotype A (11.7% versus 0%, p = 0.0525).

Characterizing transcriptome trajectories across endotypes

To understand the relationships between controls and endotype A and B, we employed Cell Trajectory (also termed pseudotime trajectory) based on transcriptomic profiles. The trajectory score increased from healthy controls to endotype B to endotype A (Figure 5A); this result led to a search for specific molecular properties that follow the predicted disease trajectory. We computed normalized pathway activity scores for each patient in the discovery cohort. The results demonstrate that, in general, pathways related to inflammation and immunity increase in a monotonic manner from healthy controls to endotype B to endotype A (Figure 5B). Upon acute infection, increased glycolysis, the tricarboxylic acid cycle (TCA) and one-carbon metabolism provide metabolites requisite to fuel cellular proliferation; however, if infection is chronic, cells become metabolically exhausted and proliferation decreases[19-22]. Compared to healthy controls, endotype B has increased expression of pathways related to metabolism and proliferation (oxidative phosphorylation, electron transport

chain (ETC), and G2M; Figure 5C-D; Supplemental Table 3). In contrast, endotype A patients exhibit decreased activity scores for pathways related to cellular proliferation (G2M, MYC and E2F; $p < 0.0001$) and metabolism (oxidative phosphorylation, the TCA cycle and the ETC; $p < 0.001$; Figure 5C-D).

Hyperinflammatory, hyporesponsive TB endotype

At the gene transcription level, TB endotype A is characterized by increased inflammation and increased interferon, TNF□, and IL-6 signaling in non-stimulated blood (Figure 2D). To evaluate the functional response upon stimulation, an independent cohort from Eswatini with mitogen-stimulated whole blood samples was analyzed by ELISA. Rank-sum analysis was implemented to stratify the patients into immune-responsive versus less-responsive groups based on their response to stimulation with 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, but decreased capacity to up-regulate IFN-γ, TNF□, IL-1β, IL-6, CXCL9, and CXCL10 upon stimulation with mitogen (Figure 6C-6D, $p < 0.007$). The immune-responsive group was similar to the healthy controls with regard to the capacity to respond to stimulation. Among the 40 TB patients, there were 4 deaths, 3 in the hyperinflammatory, hypo-responsive group and 1 in the immune responsive group (x^2 $p = 0.27$).

Comparison of chemical compound signatures to TB endotype signatures

The Library of Integrated Network-based Cellular Signatures (LINCS) enables the identification and prioritization of putative drugs to treat a pathologic condition by countering its transcriptome signature. Comparing gene signatures for endotypes A and B to healthy controls, we obtained 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), and retinoids, demonstrated connectivity scores that suggest a putative benefit for one endotype and either an inconsequential or contradictory response

for the other endotype (Supplemental Figure 5). For example, HDAC inhibitors, such as vorinostat and phenylbutyrate, demonstrated transcriptomic signatures similar to endotype A, but dissimilar to endotype B.

DISCUSSION

In the pre-antibiotic era, a fifth of humans with active tuberculosis survived more than 10-years[23], but present knowledge is inadequate to describe the underlying mechanisms of a sufficient immune response to overcome TB and contain *Mtb* infection. Further, to date, a single adjuvant host-directed therapy has not been identified, probably because the immune response to TB is protean and polymorphic. In this study, we identified clinically relevant TB endotypes by using unbiased clustering of unstimulated blood transcriptomes. Compared to controls, both endotypes displayed elevated gene expression related to pathways for inflammation and immunity, with higher levels among endotype A. In contrast, endotype B enriched for oxidative phosphorylation, the TCA cycle and pathways related to cellular proliferation, while endotype A demonstrated decreased pathways related to proliferation. Heme metabolism was upregulated in endotype A and downregulated in endotype B compared to controls, as described previously[6]. We derived a concise random forest classifier for TB endotypes, then used it to predicted endotypes in a validation cohort with richly annotated clinical outcomes; endotype A demonstrated slower times to bacterial clearance, and reduced incidence of disease cure.

Patients with TB are currently treated based on studies examining large heterogeneous groups. However, it is reasonable to hypothesize that subgroups exist within these large populations and that stratified and precision medicine strategies may improve outcomes[7]. These data provide support for individually stratified treatment approaches. Considering the animal model, *in vitro,* and human evidence, additional subtypes and endotypes will likely be identifiable when more robust epidemiology, strain characterizations, and functional immune analyses are integrated with transcriptomic results. It is notable that both endotypes exhibit elevated unstimulated gene expression levels of IFN-γ and TNF-□; however, in functional studies the TB patients with elevated basal IFN-γ and TNF-□ were less likely to upregulate IFN-γ and TNF-□ upon stimulation. Provided that future pair-wise transcriptomic and immune function studies confirm that endotype A displays characteristics

of immune exhaustion (hyperinflammatory but hyporesponsive), this classification should guide future endotype-specific HDTs.

The trajectory analysis suggest that pathways related to immunity and inflammation monotonically progress from healthy controls to first endotype B and subsequently to endotype A. In contrast, cellular proliferation and oxidative phosphorylation and the TCA cycle increase in endotype B, but decrease in endotype A. Similarly, pathways related to proliferation decrease from controls to endotype A. This is a pattern similar to murine models of TB and other chronic infections[19, 24, 25] and therefore suggests that a stage-specific intervention can prevent the progression to terminal immune exhaustion in tuberculosis.

This study is limited in its capacity to determine appropriateness of the host immune response due to limited metadata and sub-optimal means to quantify bacillary burden. Immunity to *Mtb* is tightly regulated to avoid pathologic inflammation[2, 22, 26, 27]. Animal models have demonstrated that both IFN-γ and TNF□ require delicate homeostatic regulation with deficient responses allowing disease progression and exuberant responses resulting in immune-mediated pathology [2, 3, 22, 26, 28]. The included validation cohort used the best available measurement of bacillary burden (liquid culture time to positivity) and suggests that endotype A has a hyperinflammatory response with delayed culture conversion. Prospective studies need to combine gene expression analysis with functional immunology and quantitative measures of bacillary burden to clarify the appropriateness of host immunity. We anticipate that once endotypes are analyzed using robust multi-omic platforms, effective and pragmatic classifiers could use a minimal complement of informative features. Capitalizing on this minimized complement, cost-effective diagnostics could be developed and deployed at point-of-care in TB high burden settings.

The link between metabolism, particularly glycolysis, and immune function has been appreciated for over 90 years [20-22]. Metabolism is a mediator of the epigenetic mechanisms driving immune function[24, 25, 29-31]. Therefore, it is interesting that 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 drive antigen-induced immune function[29, 32]. Another inhibitor of mTOR, everolimus decreased TB-induced lung damage[27]. *In silico* evidence indicates the most pronounced benefit of everolimus for endotype A.

Previously identified candidates for HDTs include IFN-γ, GM-CSF, TNF□, TNF□ inhibitors, NSAIDS, Vitamin D, glucocorticoids, HDAC inhibitors, mTOR modulators, retinoids, and statins[27, 33]. The *in silico* analysis demonstrated that previously identified HDTs would 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 NSAID, TNF□ inhibitor, or glucocorticoid treatment would be appropriate. Animal and *in vitro* models that recapitulate the clinically relevant endotypes are also needed to better evaluate endotype-specific HDTs.

All included studies evaluated unstimulated host gene expression. TB is a chronic infection resulting in immune suppression. While many genes downstream of IFN-γ are elevated in TB patients at baseline[13, 15, 34, 35], they have decreased antigen-induced immune upregulation[36-39]. The multiplex ELISA data highlight the limitations of inferring immune function based on non-stimulated gene expression measurements; in fact, the group with elevated baseline cytokines (hyperinflammatory) was hyporesponsive upon stimulation. Additional gene expression subclusters were visible at higher resolution; however, biological relevance was not readily obvious in these sub-clusters. We speculate that the integration of transcriptomics with functional immune analysis, more robust epidemiology, and strain characterization will identify more than two endotypes.

Progression to TB is related to interactions among host, pathogen, and environmental factors. Progression to a specific endotype of TB is likely similarly related to as-yet unappreciated interactions. Unlike The Cancer Genome Atlas (TCGA), very limited epidemiology is available in existing public data repositories. Epidemiologic predispositions likely drive the divergent endotypes, including malnutrition, HIV, helminth, tobacco use, and/or indoor biomass fuel exposure. For example, despite successful deworming, previous schistosomiasis infection ablates mycobacterial immunity, leaving long-lasting immune suppression. We speculate that individuals with pre-existing immune suppression progress rapidly to endotype A, in contrast to previously healthy individuals.

In conclusion, this unbiased clustering provides additional evidence that there are multiple molecular host pathways modulated during TB [2-7]. This analysis of transcriptome and protein data from TB patients provides additional evidence for biologically distinctive TB endotypes that differentially affect clinical outcomes. Specifically, host gene expression in TB patient clusters into at least two endotypes with differential immune and metabolic transcriptomic signatures. These observations suggest that different endotypes display responses that are likely to have clinical and pathologic relevance and provides the basis for studies to evaluate endotype-specific host-directed therapies.

REFERENCES

1. Svenningsen S, Nair P. Asthma Endotypes and an Overview of Targeted Therapy for Asthma. *Front Med (Lausanne)* 2017: 4: 158.

2. Tobin DM, Roca FJ, Oh SF, McFarland R, Vickery TW, Ray JP, Ko DC, Zou Y, Bang ND, Chau TT, Vary JC, Hawn TR, Dunstan SJ, Farrar JJ, Thwaites GE, King MC, Serhan CN, Ramakrishnan L. Host genotype-specific therapies can optimize the inflammatory response to mycobacterial infections. *Cell* 2012: 148(3): 434-446.

3. 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 pathogens* 2016: 12(5): e1005667.

4. 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): 454-470.

5. Ni Cheallaigh C, Sheedy FJ, Harris J, Munoz-Wolf N, Lee J, West K, McDermott EP, Smyth A, Gleeson LE, Coleman M, Martinez N, Hearnden CH, Tynan GA, Carroll EC, Jones SA, Corr SC, Bernard NJ, Hughes MM, Corcoran SE, O'Sullivan M, Fallon CM, Kornfeld H, Golenbock D, Gordon SV, O'Neill LA, Lavelle EC, Keane J. A Common Variant in the Adaptor Mal Regulates Interferon Gamma Signaling. *Immunity* 2016: 44(2): 368-379.

6. Andrade BB, Pavan Kumar N, Amaral EP, Riteau N, Mayer-Barber KD, Tosh KW, Maier N, Conceicao EL, Kubler A, Sridhar R, Banurekha VV, Jawahar MS, Barbosa T, Manganiello VC, Moss J, Fontana JR, Marciano BE, Sampaio EP, Olivier KN, Holland SM, Jackson SH, Moayeri M, Leppla S, Sereti I, Barber DL, Nutman TB, Babu S, Sher A. Heme Oxygenase-1 Regulation of Matrix Metalloproteinase-1 Expression Underlies Distinct Disease Profiles in Tuberculosis. *Journal of immunology* 2015: 195(6): 2763-2773.

7. 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.

8. Heyckendorf J, Marwitz S, Reimann M, Avsar K, DiNardo A, Gunther G, Hoelscher M, Ibraim E, Kalsdorf B, Kaufmann SHE, Kontsevaya I, van Leth F, Mandalakas AM, Maurer FP, Muller M, Nitschkowski D, Olaru ID, Popa C, Rachow A, Rolling T, Rybniker J, Salzer HJF, Sanchez-Carballo P, Schuhmann M, Schaub D, Spinu V, Suarez I, Terhalle E, Unnewehr M, Weiner J, 3rd, Goldmann T, Lange C. Prediction of anti-tuberculosis treatment duration based on a 22-gene transcriptomic model. *Eur Respir J* 2021.

9. Berry MP, Graham CM, McNab FW, Xu Z, Bloch SA, Oni T, Wilkinson KA, Banchereau R, Skinner J, Wilkinson RJ, Quinn C, Blankenship D, Dhawan R, Cush JJ, Mejias A, Ramilo O, Kon OM, Pascual V, Banchereau J, Chaussabel D, O'Garra A. An interferon-inducible neutrophil-driven blood transcriptional signature in human tuberculosis. *Nature* 2010: 466(7309): 973-977.

10. Bloom CI, Graham CM, Berry MP, Wilkinson KA, Oni T, Rozakeas F, Xu Z, Rossello-Urgell J, Chaussabel D, Banchereau J, Pascual V, Lipman M, Wilkinson RJ, O'Garra A. Detectable changes in the blood transcriptome are present after two weeks of antituberculosis therapy. *PloS one* 2012: 7(10): e46191.

11. Bloom CI, Graham CM, Berry MP, Rozakeas F, Redford PS, Wang Y, Xu Z, Wilkinson KA, Wilkinson RJ, Kendrick Y, Devouassoux G, Ferry T, Miyara M, Bouvry D, Valeyre D, Gorochov G, Blankenship D, Saadatian M, Vanhems P, Beynon H, Vancheeswaran R, Wickremasinghe M, Chaussabel D, Banchereau J, Pascual V, Ho LP, Lipman M, O'Garra A. Transcriptional blood signatures distinguish pulmonary tuberculosis, pulmonary sarcoidosis, pneumonias and lung cancers. *PloS one* 2013: 8(8): e70630.

12. Walter ND, Miller MA, Vasquez J, Weiner M, Chapman A, Engle M, Higgins M, Quinones AM, Rosselli V, Canono E, Yoon C, Cattamanchi A, Davis JL, Phang T, Stearman RS, Datta G, Garcia BJ, Daley CL, Strong M, Kechris K, Fingerlin TE, Reves R, Geraci MW. 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): 274-282.

13. Anderson ST, Kaforou M, Brent AJ, Wright VJ, Banwell CM, Chagaluka G, Crampin AC, Dockrell HM, French N, Hamilton MS, Hibberd ML, Kern F, Langford PR, Ling L, Mlotha R, Ottenhoff TH, Pienaar S, Pillay V, Scott JA, Twahir H, Wilkinson RJ, Coin LJ, Heyderman RS, Levin M, Eley B, Consortium I, Group KTS. Diagnosis of childhood tuberculosis and host RNA expression in Africa. *The New England journal of medicine* 2014: 370(18): 1712-1723.

14. Kaforou M, Wright VJ, Oni T, French N, Anderson ST, Bangani N, Banwell CM, Brent AJ, Crampin AC, Dockrell HM, Eley B, Heyderman RS, Hibberd ML, Kern F, Langford PR, Ling L, Mendelson M, Ottenhoff TH, Zgambo F, Wilkinson RJ, Coin LJ, Levin M. 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.
15. Blankley S, Graham CM, Turner J, Berry MP, Bloom CI, Xu Z, Pascual V, Banchereau J, Chaussabel D, Breen R, Santis G, Blankenship DM, Lipman M, O'Garra A. The Transcriptional Signature of Active Tuberculosis Reflects Symptom Status in Extra-Pulmonary and Pulmonary Tuberculosis. *PloS one* 2016: 11(10): e0162220.
16. Singhania A, Verma R, Graham CM, Lee J, Tran T, Richardson M, Lecine P, Leissner P, Berry MPR, Wilkinson RJ, Kaiser K, Rodrigue M, Woltmann G, Haldar P, O'Garra A. A modular transcriptional signature identifies phenotypic heterogeneity of human tuberculosis infection. *Nat Commun* 2018: 9(1): 2308.
17. Leong S, Zhao Y, Joseph NM, Hochberg NS, Sarkar S, Pleskunas J, Hom D, Lakshminarayanan S, Horsburgh CR, Jr., Roy G, Ellner JJ, Johnson WE, Salgame P. Existing blood transcriptional classifiers accurately discriminate active tuberculosis from latent infection in individuals from south India. *Tuberculosis* 2018: 109: 41-51.
18. Heyckendorf J, Marwitz S, Reimann M, Avsar K, DiNardo AR, Guenther G, Hoelscher M, Ibraim E, Kalsdorf B, Kaufmann SHE, Kontsevaja I, van Leth F, Mandalakas AM, Maurer F, Mueller M, Nitschkowski D, Olaru ID, Popa C, Rachow A, Rolling T, Rybniker J, Salzer HJF, Sanchez-Carballo P, Schuhmann M, Schaub D, Spinu V, Suarez I, Terhalle E, Unnewehr M, Weiner J, Goldmann T, Lange C. A 22-gene transcriptomic model indicating individual therapy durations in multidrug-resistant tuberculosis. *European Respiratory Journal* 2021: In Press.
19. Russell SL, Lamprecht DA, Mandizvo T, Jones TT, Naidoo V, Addicott KW, Moodley C, Ngcobo B, Crossman DK, Wells G, Steyn AJC. Compromised Metabolic Reprogramming Is an Early Indicator of CD8(+) T Cell Dysfunction during Chronic Mycobacterium tuberculosis Infection. *Cell Rep* 2019: 29(11): 3564-3579 e3565.
20. 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. *Infection and immunity* 1972: 5(4): 499-504.
21. Warburg O, Wind F, Negelein E. The Metabolism of Tumors in the Body. *J Gen Physiol* 1927: 8(6): 519-530.
22. Nair S, Huynh JP, Lampropoulou V, Loginicheva E, Esaulova E, Gounder AP, Boon ACM, Schwarzkopf EA, Bradstreet TR, Edelson BT, Artyomov MN, Stallings CL, Diamond MS. Irg1 expression in myeloid cells prevents immunopathology during M. tuberculosis infection. *The Journal of experimental medicine* 2018: 215(4): 1035-1045.
23. Tattersall WH. The survival of sputum-positive consumptives; a study of 1,192 cases in a county borough between 1914 and 1940. *Tubercle* 1947: 28(6): 107-114.
24. Bengsch B, Johnson AL, Kurachi M, Odorizzi PM, Pauken KE, Attanasio J, Stelekati E, McLane LM, Paley MA, Delgoffe GM, Wherry EJ. 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): 358-373.
25. 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): 374-388.
26. Ravimohan S, Auld SC, Maenetje P, Ratsela N, Mlotshwa M, Ncube I, Smith JP, Vangu MD, Sebe M, Kossenkov A, Weissman D, Wallis RS, Churchyard G, Kornfeld H, Bisson GP. Lung Injury on Antiretroviral Therapy in Adults With Human Immunodeficiency Virus/Tuberculosis. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* 2020: 70(9): 1845-1854.
27. Wallis RS, Ginindza S, Beattie T, Arjun N, Likoti M, Edward VA, Rassool M, Ahmed K, Fielding K, Ahidjo BA, Vangu MDT, Churchyard G. Adjunctive host-directed therapies for pulmonary tuberculosis: a prospective, open-label, phase 2, randomised controlled trial. *The Lancet Respiratory medicine* 2021.
28. Roca FJ, Ramakrishnan L. TNF dually mediates resistance and susceptibility to mycobacteria via mitochondrial reactive oxygen species. *Cell* 2013: 153(3): 521-534.
29. Arts RJW, Carvalho A, La Rocca C, Palma C, Rodrigues F, Silvestre R, Kleinnijenhuis J, Lachmandas E, Goncalves LG, Belinha A, Cunha C, Oosting M, Joosten LAB, Matarese G, van Crevel R, Netea MG. Immunometabolic Pathways in BCG-Induced Trained Immunity. *Cell Rep* 2016: 17(10): 2562-2571.

30. Barili V, Fisicaro P, Montanini B, Acerbi G, Filippi A, Forleo G, Romualdi C, Ferracin M, Guerrieri F, Pedrazzi G, Boni C, Rossi M, Vecchi A, Penna A, Zecca A, Mori C, Orlandini A, Negri E, Pesci M, Massari M, Missale G, Levrero M, Ottonello S, Ferrari C. Targeting p53 and histone methyltransferases restores exhausted CD8+ T cells in HCV infection. *Nat Commun* 2020: 11(1): 604.
31. Abhimanyu, Ontiveros CO, Guerra-Resendez RS, Nishiguchi T, Ladki M, Hilton IB, Schlesinger LS, DiNardo AR. Reversing Post-Infectious Epigenetic-Mediated Immune Suppression. *Frontiers in immunology* 2021: 12: 688132.
32. Bohme J, Martinez N, Li S, Lee A, Marzuki M, Tizazu AM, Ackart D, Frenkel JH, Todd A, Lachmandas E, Lum J, Shihui F, Ng TP, Lee B, Larbi A, Netea MG, Basaraba R, van Crevel R, Newell E, Kornfeld H, Singhal A. Metformin enhances anti-mycobacterial responses by educating CD8+ T-cell immunometabolic circuits. *Nat Commun* 2020: 11(1): 5225.
33. 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.
34. Zak DE, Penn-Nicholson A, Scriba TJ, Thompson E, Suliman S, Amon LM, Mahomed H, Erasmus M, Whatney W, Hussey GD, Abrahams D, Kafaar F, Hawkridge T, Verver S, Hughes EJ, Ota M, Sutherland J, Howe R, Dockrell HM, Boom WH, Thiel B, Ottenhoff TH, Mayanja-Kizza H, Crampin AC, Downing K, Hatherill M, Valvo J, Shankar S, Parida SK, Kaufmann SH, Walzl G, Aderem A, Hanekom WA, Acs, groups GCcs. A blood RNA signature for tuberculosis disease risk: a prospective cohort study. *Lancet* 2016: 387(10035): 2312-2322.
35. 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. *The Lancet Respiratory medicine* 2020.
36. Harari A, Rozot V, Enders FB, Perreau M, Stalder JM, Nicod LP, Cavassini M, Calandra T, Blanchet CL, Jaton K, Faouzi M, Day CL, Hanekom WA, Bart PA, Pantaleo G. Dominant TNF-alpha+ Mycobacterium tuberculosis-specific CD4+ T cell responses discriminate between latent infection and active disease. *Nature medicine* 2011: 17(3): 372-376.
37. Sahiratmadja E, Alisjahbana B, de Boer T, Adnan I, Maya A, Danusantoso H, Nelwan RH, Marzuki S, van der Meer JW, van Crevel R, van de Vosse E, Ottenhoff TH. 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): 820-829.
38. 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. *The Journal of infectious diseases* 2013: 208(4): 603-615.
39. DiNardo A, Rajapakshe K, Nishiguchi T, Mtetwa G, Grimm SL, Dlamini Q, Kahari J, Mahapatra S, Kay AW, Maphalala G, Mace EM, Makedonas G, Cirillo JD, Netea M, van Crevel R, Coarfa C, Mandalakas AM. DNA hyper-methylation during Tuberculosis dampens host immune responsiveness. *The Journal of clinical investigation* 2020.

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 one-sided chi-squared evaluated differences in populations. Not all data was available for all participants.

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

	All	Endotype A	Endotype B	p-value*
Median age	37.95	39.78	37.95	0.6948
% Male	62.2% (56/90)	62.5% (30/48)	61.9% (26/42)	0.4768
Median BMI	20.75	19.49	21.28	0.0009
Current or previous smoker	52.0% (51/98)	46.2% (24/52)	58.7% (27/46)	0.1074
Cavitary Disease	71.7% (66/92)	76.5% (39/51)	65.9% (27/41)	0.1305
MDR TB	62.0% (75/121)	54.7% (35/64)	70.2% (40/57)	0.0399
Baseline TTP (days)	17	14	21	0.0169
Median TCC (days)	54.0	64.5	33.5	0.0005
Cure	81.0% (51/63)	74.4% (29/39)	91.7% (22/24)	0.0447
Mortality	6.3% (4/63)	10.3% (4/39)	0% (0/24)	0.0525

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 U test or one-sided Chi-squared (A worse than B) were used to determine significance, as appropriate.*

Figure 1: Overview of study. Using unbiased clustering of TB patients from 12 publicly available microarray datasets a Random Forest gene classifier was derived to predict TB Endotype A versus B. This was validated on an RNA-seq cohort compiled over 4 publicly available datasets and on a microarray patient cohort that included longitudinal clinical outcomes data. Immunological validation was evaluated by multiplex ELISA using a separate cohort from Eswatini. Finally, similarity to endotype gene signatures was used to assess and rank previously evaluated host directed therapy candidates.

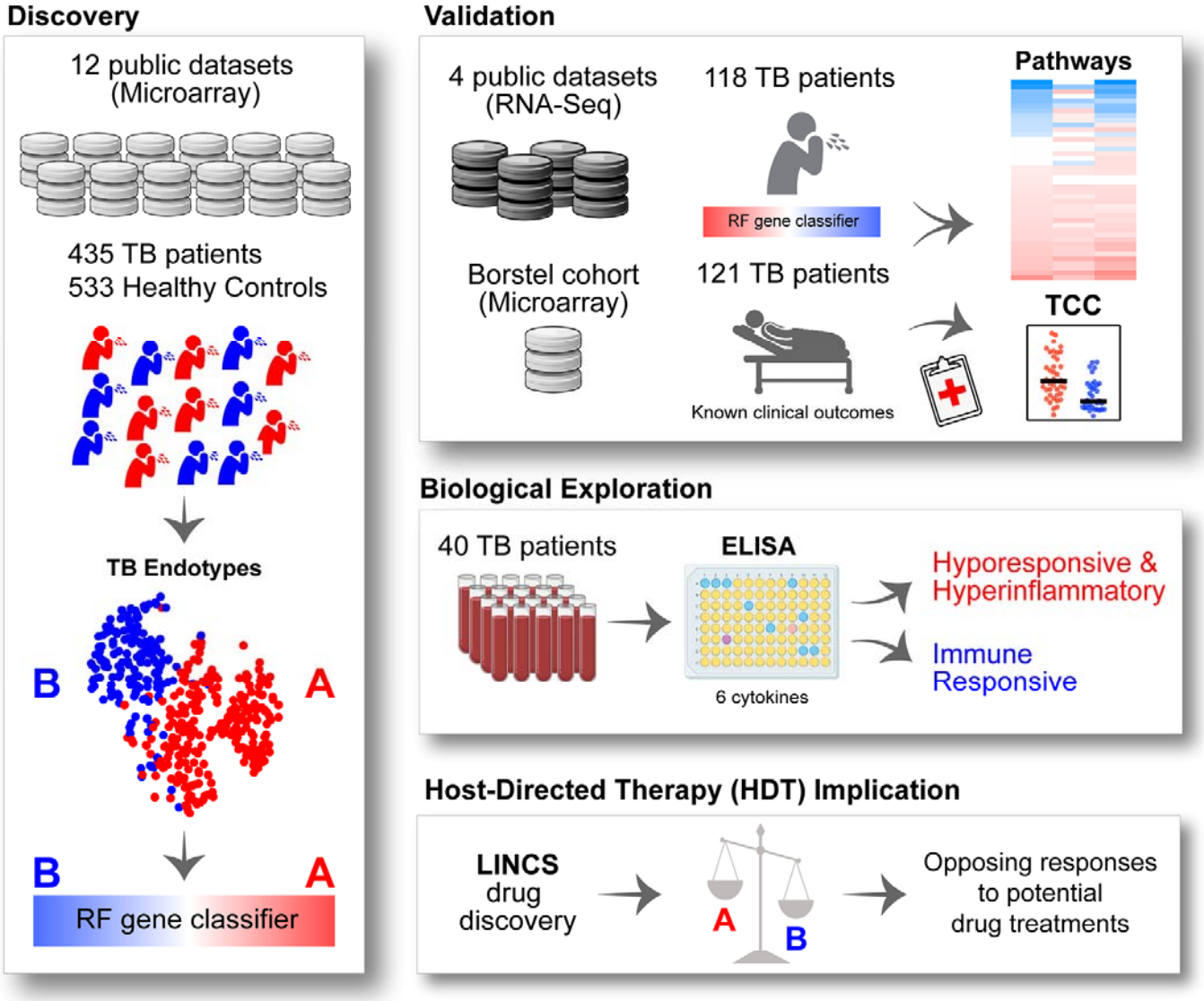


Figure 2: Unbiased clustering identifies unique TB endotypes. (A) Unbiased clustering was implemented on discovery cohort of 7 studies, identifying two major endotypes; next, a random forest gene classifier was developed and applied to two external validation datasets. (B) Network-based unbiased clustering using the Louvain method identifies 2 major endotypes of TB. (C) Distribution of individual studies into Endotype A or B. (D) 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.

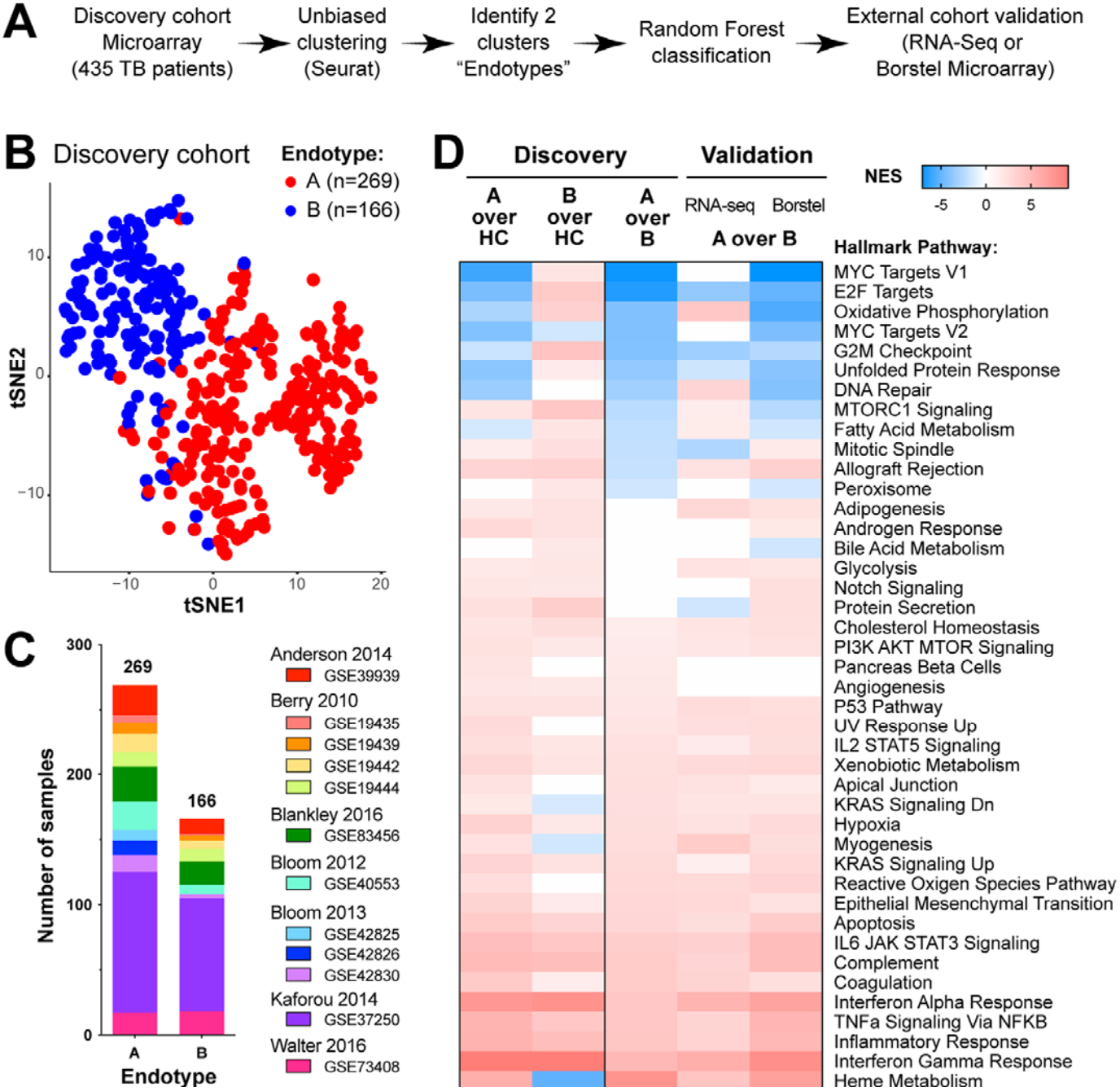


Figure 3: Evaluation of TB risk signatures over the TB endotypes. Previous studies identified gene signatures associated with A) disease severity and treatment failure or with B) risk of treatment failure. Activity scores were computed for each risk signature across healthy controls (HC, grey) and Endotype A (A, red) and Endotype B (B, blue). Mann-Whitney scores with **** $p < 0.0001$, *** $p < 0.0002$, ns = non-significant. Solid line at median, dashed lines at interquartile range, dotted line across at median of HC.

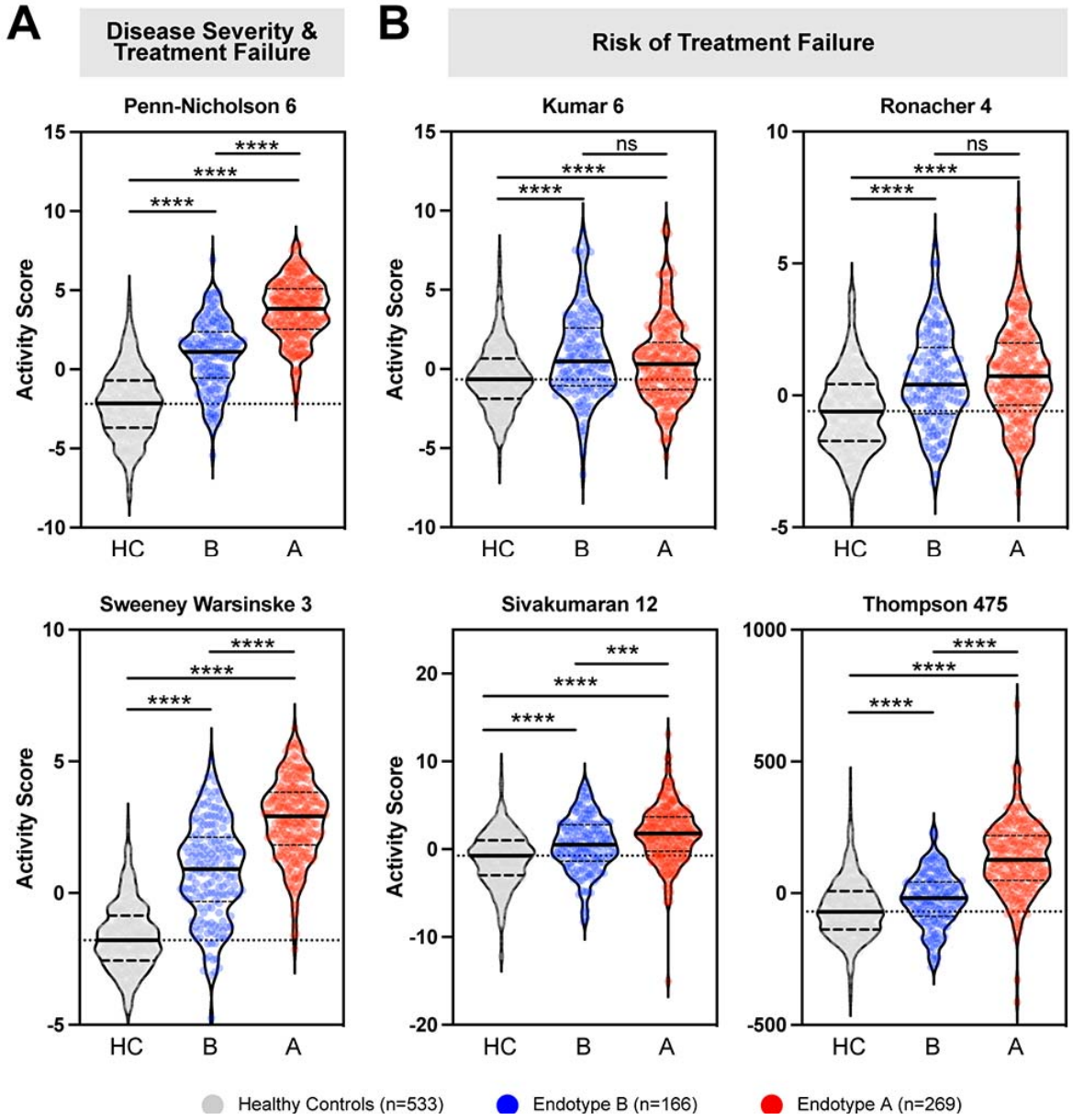


Figure 4: Endotype evaluation of TB clinical outcomes. Using the clinical annotations of the Borstel TB cohort, outcome differences between endotypes and association of pathway scores with outcomes were evaluated. (A) Time to culture conversion (TCC) in TB patients identified as endotype A or B (p = 0.0005 by Mann Whitney). (B) Rates of diseases cure in TB patients identified as endotype A or B (p = 0.0447 by one-sided Chi-squared test).

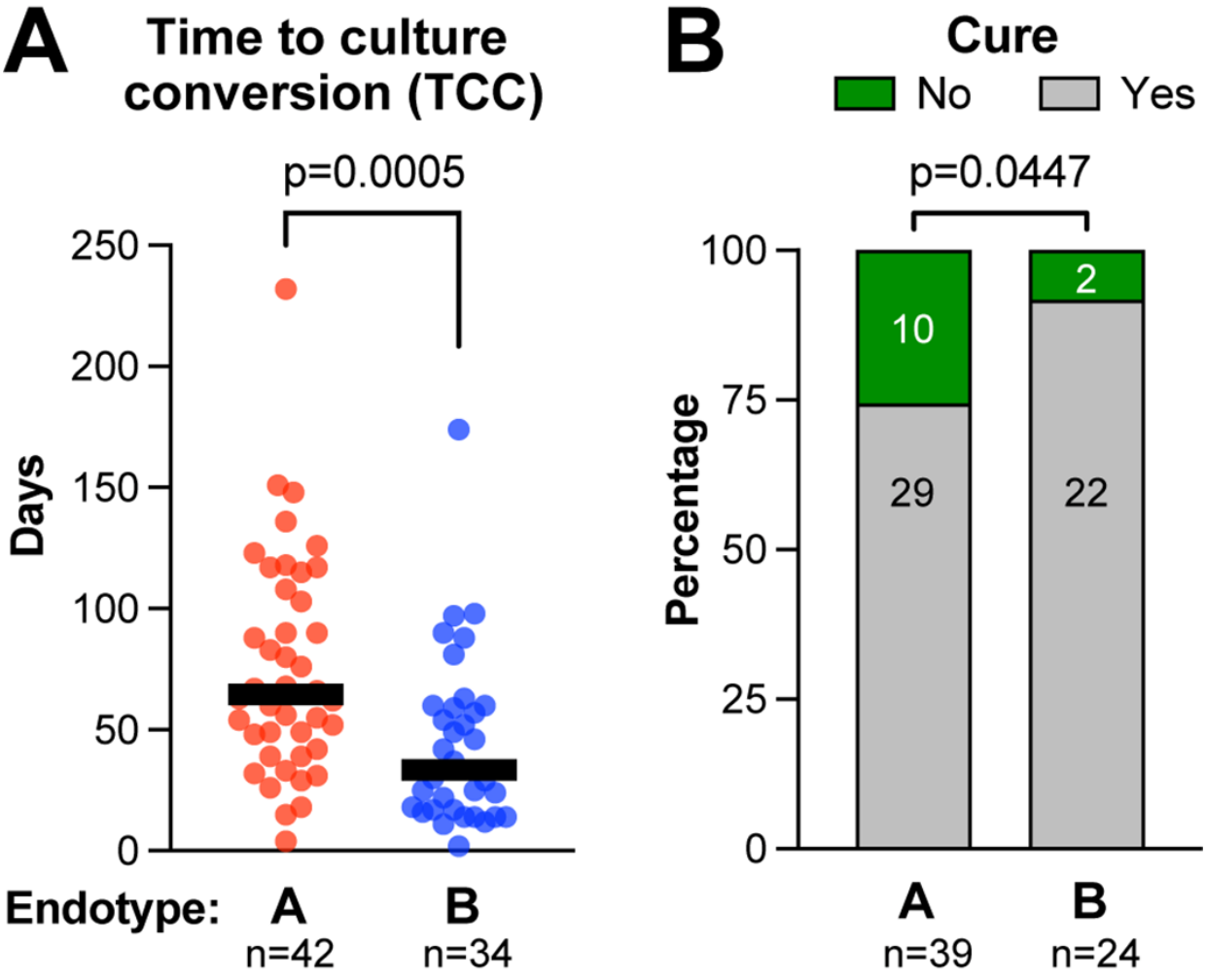


Figure 5: TB endotypes display distinct immune and metabolic gene expression activity scores. (A) Pseudotime TB trajectory score. Pathway activity scores were evaluated between healthy controls (HC, grey), endotype B (B, blue) and endotype A (A, red). Inflammation and immunity pathways (B), Metabolic pathways (C), and Proliferation pathways (D). Mann-Whitney test was used; **** $p < 0.0001$, *** $p < 0.0002$, ** $p < 0.0021$, ns = non-significant. Solid line at median, dashed lines indicate interquartile range, dotted line across HC median. Specific gene changes are available in Supplemental Table 2.

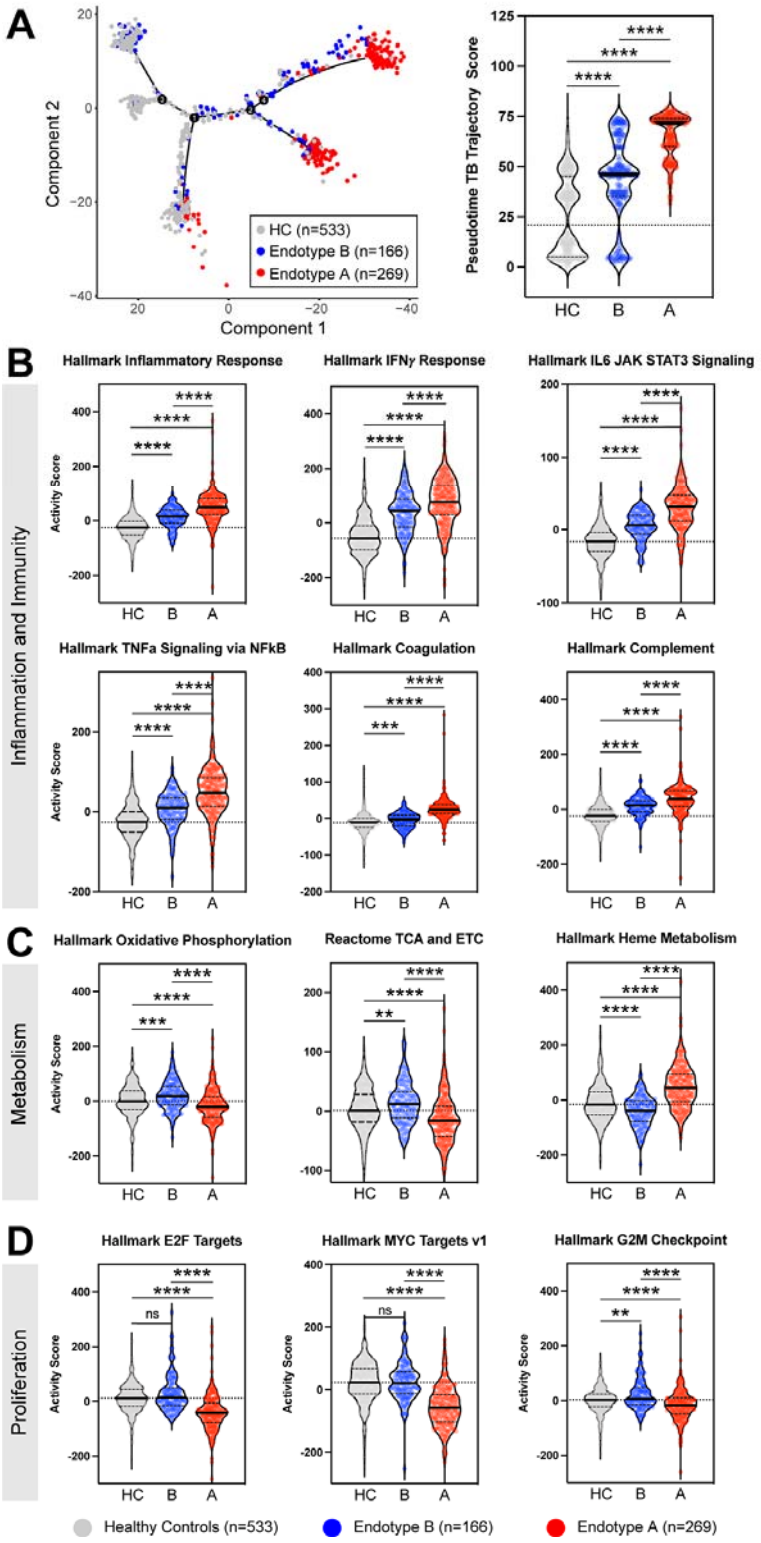
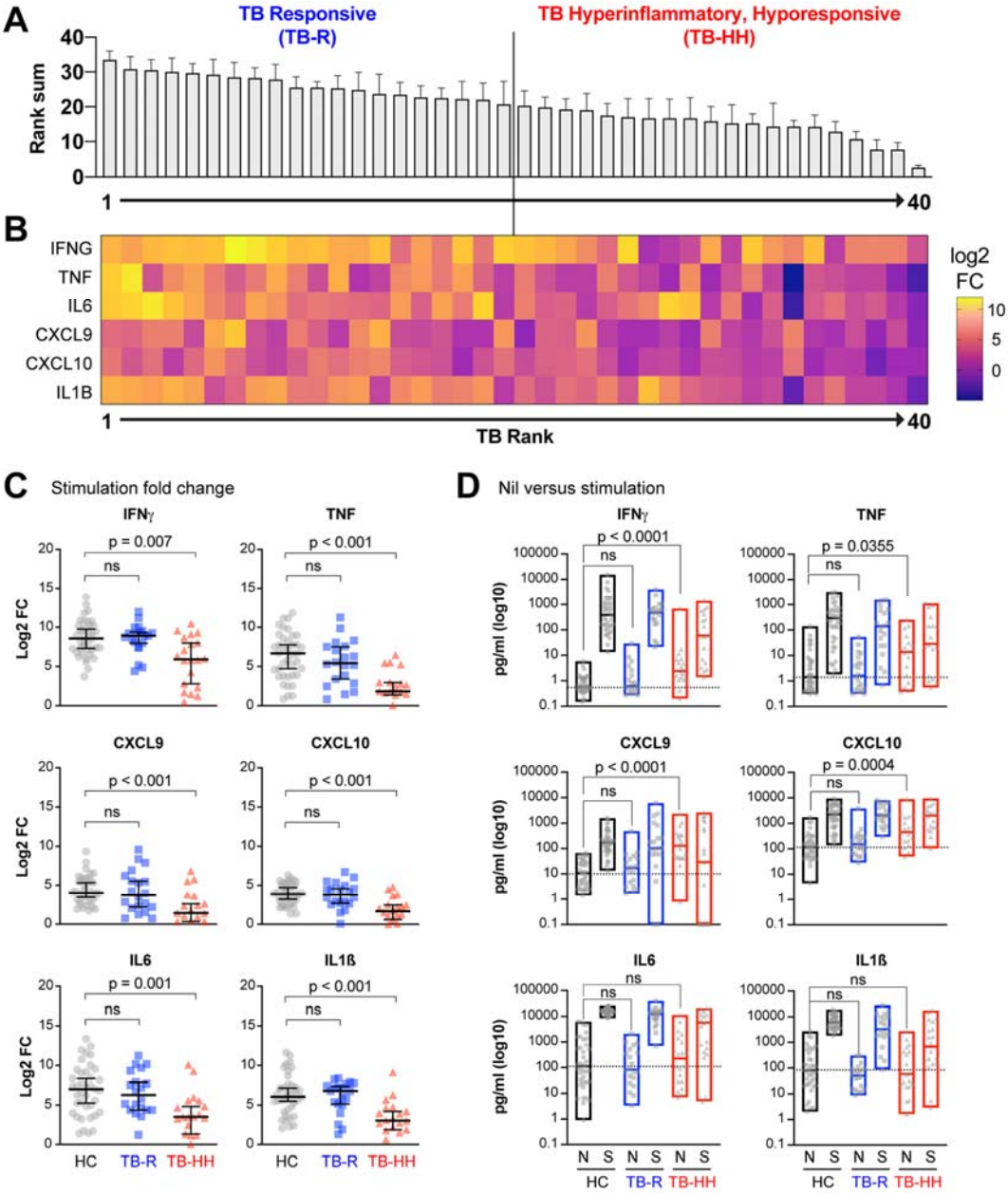
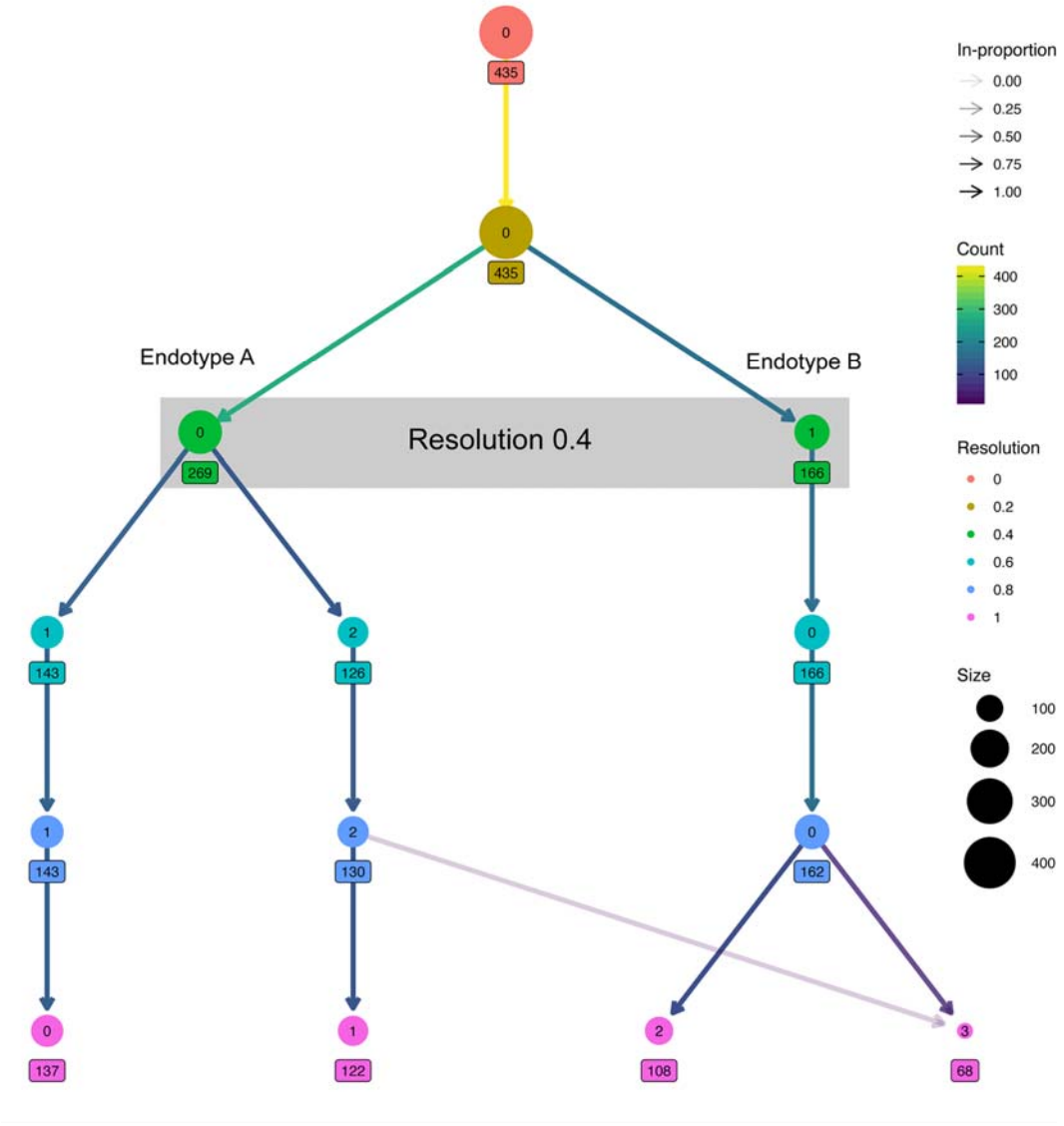


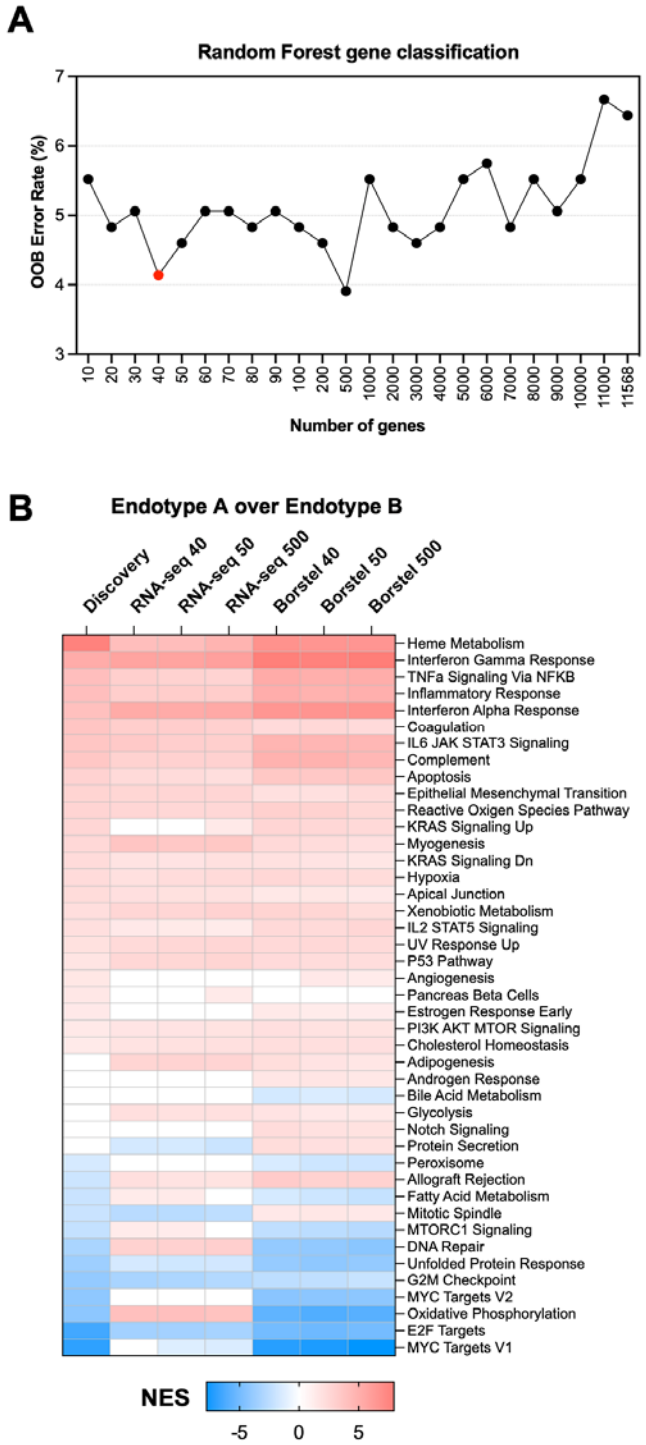
Figure 6: Identification of hyper-inflammatory, hyporesponsive cytokine production in TB patient endotypes. Whole blood from TB patients (n = 40) and healthy controls (n = 39) was stimulated overnight with or without mitogen (PHA), followed by measurement of cytokines and chemokines. (A) Samples were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



Supplemental Figure 1: Hierarchical cluster tree identifying different clusters based on resolution.

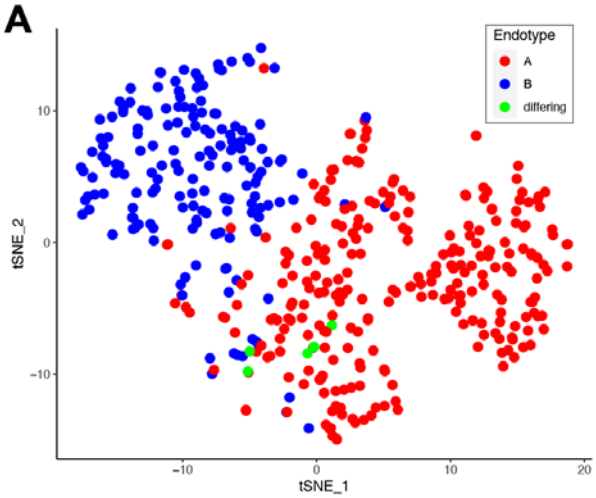


Supplemental Figure 2: A) Out of bag error rate across a spectrum of genes for applying the Random Forest classifier. B) Heatmap of GSEA Normalized Enrichment Scores (NES) for Hallmark pathways using the 40, 50 and 500 gene TB endotype classifiers over discovery and validation cohorts.



Supplemental Figure 3: Concordance of sample classification for the same patient samples using microarray data in Berry 2010, compared to RNA-seq in Singhania 2018.

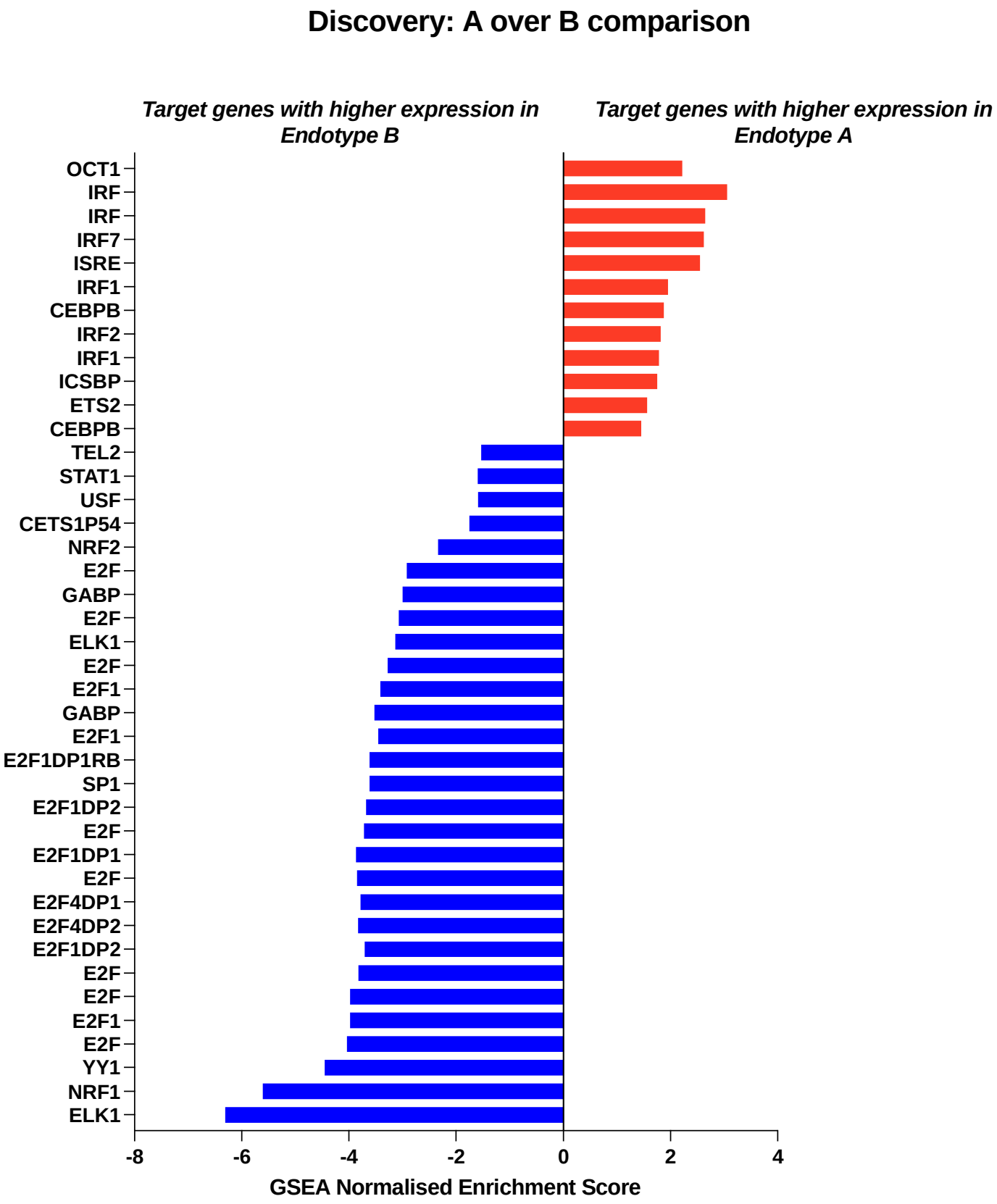
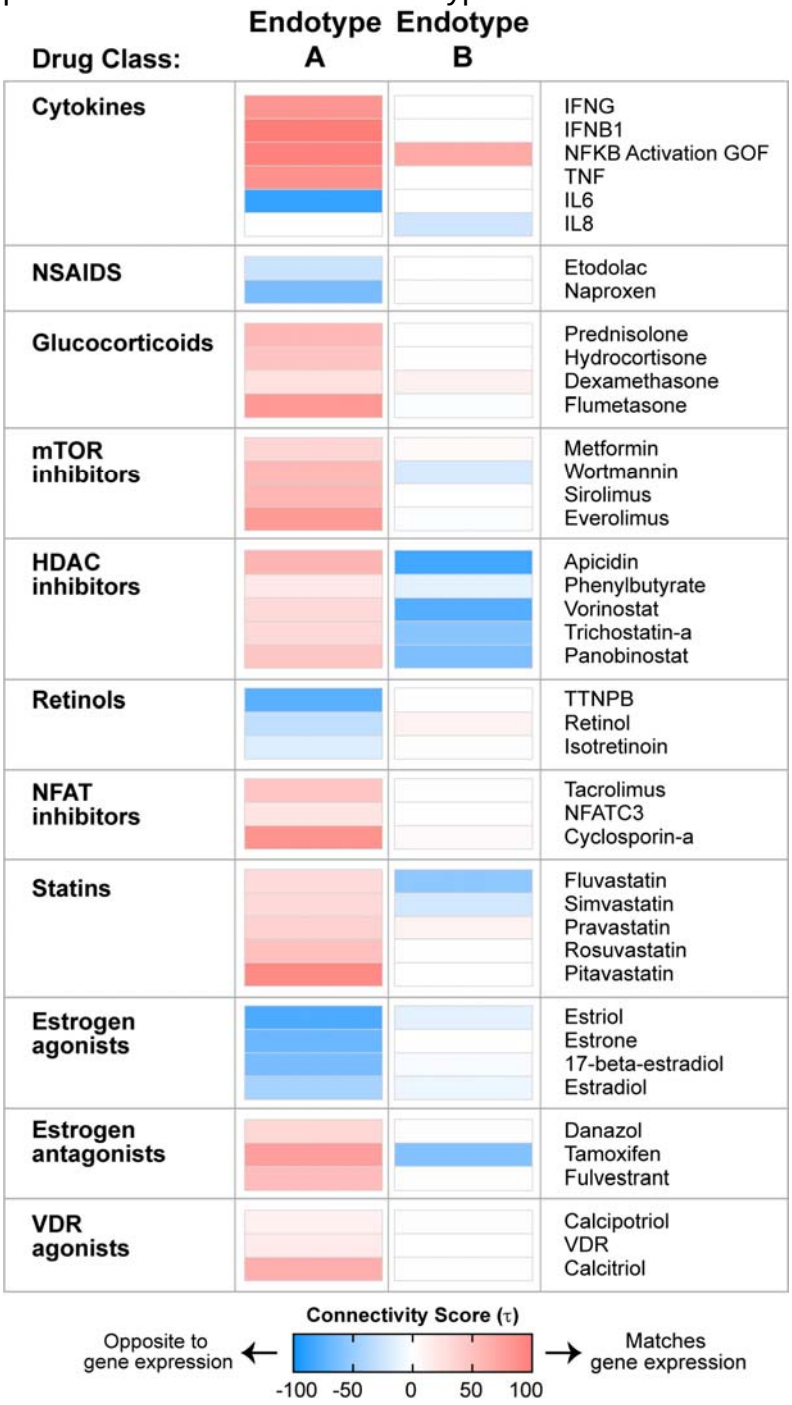
A



B

Microarray (Berry 2010)			RNA-Seq (Singhania 2018)		
GSE	sample	Endotype	GSE	sample	Endotype
GSE19442	GSM484500	A	GSE107992	GSM2886127	A
GSE19442	GSM484501	A	GSE107992	GSM2886128	B
GSE19442	GSM484502	A	GSE107992	GSM2886129	B
GSE19442	GSM484515	A	GSE107992	GSM2886130	A
GSE19442	GSM484524	B	GSE107992	GSM2886133	B
GSE19442	GSM484525	B	GSE107992	GSM2886134	B
GSE19442	GSM484526	A	GSE107992	GSM2886135	A
GSE19442	GSM484533	A	GSE107992	GSM2886093	A
GSE19442	GSM484534	A	GSE107992	GSM2886098	A
GSE19442	GSM484535	A	GSE107992	GSM2886099	A
GSE19442	GSM484536	A	GSE107992	GSM2886104	A
GSE19442	GSM484543	A	GSE107992	GSM2886105	A
GSE19442	GSM484544	A	GSE107992	GSM2886110	A
GSE19444	GSM484595	A	GSE107991	GSM2886035	B
GSE19444	GSM484596	A	GSE107991	GSM2886043	A
GSE19444	GSM484597	A	GSE107991	GSM2886061	B
GSE19444	GSM484598	A	GSE107991	GSM2886064	A
GSE19444	GSM484605	B	GSE107991	GSM2886041	B
GSE19444	GSM484606	A	GSE107991	GSM2886044	A
GSE19444	GSM484610	B	GSE107991	GSM2886049	B
GSE19444	GSM484611	A	GSE107991	GSM2886070	A
GSE19444	GSM484618	A	GSE107991	GSM2886068	B
GSE19444	GSM484619	A	GSE107991	GSM2886071	A
GSE19444	GSM484620	B	GSE107991	GSM2886051	B
GSE19444	GSM484621	A	GSE107991	GSM2886054	A
GSE19444	GSM484628	B	GSE107991	GSM2886075	B
GSE19444	GSM484629	B	GSE107991	GSM2886036	B
GSE19444	GSM484630	B	GSE107991	GSM2886077	B
GSE19444	GSM484631	A	GSE107991	GSM2886079	A
GSE19444	GSM484638	A	GSE107991	GSM2886057	A
GSE19444	GSM484639	B	GSE107991	GSM2886059	B
GSE19444	GSM484641	B	GSE107991	GSM2886080	B
GSE19444	GSM484642	B	GSE107991	GSM2886078	B
GSE19444	GSM484645	B	GSE107991	GSM2886081	B

Supplemental Figure 4. Differentially enriched Transcription Factor Targets pathways**.** Genes targeted by different transcription factors are divergent between endotype A and endotype B.



Supplemental Figure 5. Heatmap of connectivity scores 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 compounds inducing gene expression profiles similar to the endotype, while negative connectivity scores represent compounds inducing gene expression profiles antithetical to the endotype.