The Journal of Infectious Diseases

MAJOR ARTICLE

Use of a Contained *Mycobacterium tuberculosis* Mouse Infection Model to Predict Active Disease and Containment in Humans

Fergal J. Duffy,[1] Gregory S. Olson,[1] Elizabeth S. Gold,[1] Ana Jahn,[1] Alan Aderem,[1] John D. Aitchison,[1] Alissa C. Rothchild,[1,a] Alan H. Diercks,[1] and Johannes Nemeth[1,2]

[1]Seattle Children's Research Institute, Seattle, Washington, USA, and [2]Division of Infectious Diseases and Hospital Epidemiology, University Hospital Zurich, Zurich, Switzerland

Previous studies have identified whole-blood transcriptional risk and disease signatures for tuberculosis; however, several lines of evidence suggest that these signatures primarily reflect bacterial burden, which increases before symptomatic disease. We found that the peripheral blood transcriptome of mice with contained *Mycobacterium tuberculosis* infection (CMTI) has striking similarities to that of humans with active tuberculosis and that a signature derived from these mice predicts human disease with accuracy comparable to that of signatures derived directly from humans. A set of genes associated with immune defense are up-regulated in mice with CMTI but not in humans with active tuberculosis, suggesting that their up-regulation is associated with bacterial containment. A signature comprising these genes predicts both protection from tuberculosis disease and successful treatment at early time points where current signatures are not predictive. These results suggest that detailed study of the CMTI model may enable identification of biomarkers for human tuberculosis.

Keywords. Tuberculosis; signature of risk; cross-species; blood transcription.

Identification of biomarkers to diagnose stages of *Mycobacterium tuberculosis* infection and disease (active tuberculosis) remains an important clinical and public health goal. Pioneering work has shown that expression of interferon-inducible genes is strongly increased in the peripheral blood of patients with active tuberculosis [1]. Additional studies have derived signatures correlated with risk of progression to active disease for latently infected (latent tuberculosis–positive) individuals in endemic areas and household contacts of active tuberculosis cases [1–7]. Multiple lines of evidence suggest that these biomarkers reflect subclinical immune responses correlated with bacterial burden rather than identifying individuals predisposed to ineffective immune control of tuberculosis [8–11], and correlates of protective immunity to tuberculosis thus remain elusive. All of these signatures have been derived directly from clinical or observational studies of human populations. While human studies are essential for proving a biomarker's clinical utility, they are time consuming and expensive. In addition, they offer very limited opportunities to experimentally interrogate the target population, a significant challenge to deciphering the immune mechanisms that underlie a biomarker.

Several recent studies have identified strong similarities between the peripheral blood transcriptomes of *M. tuberculosis*–infected humans and those of *M. tuberculosis*–infected mice [8–10]. In particular, it has been demonstrated that whole-blood transcriptional signatures derived from mice infected by aerosol at the "conventional" dose (approximately 50–100 bacterial colony-forming units [CFUs]) or an "ultra-low" dose (ULD; approximately 1–3 CFUs) can predict the outcome of *M. tuberculosis* infection in humans as accurately as signatures derived directly from human studies [10].

These results suggest that many of the immune mechanisms that are activated by *M. tuberculosis* infection are shared between mice and humans. However, it has long been known that the course of tuberculosis disease differs significantly between these species. In humans, the overwhelming majority (90%) of *M. tuberculosis* infections do not progress to clinically diagnosable disease [3, 4, 12, 13]. Nonprogressors either clear the infection or successfully contain and control replication-competent bacteria for their lifetimes, often in lymphoid tissue, while remaining asymptomatic [14]. In addition, it has been well documented that humans who resist progression to clinical disease after exposure to *M. tuberculosis* have increased protection against subsequent *M. tuberculosis* exposure [15, 16].

Received 30 October 2020; editorial decision 4 March 2021; accepted 8 March 2021; published online March 10, 2021.

[a]Present affiliation: Department of Veterinary and Animal Sciences, University of Massachusetts Amherst.

Correspondence: Fergal J. Duffy, Center for Global Infectious Disease Research, Seattle Children's Research Institute, 307 Westlake Ave N #500, Seattle, WA 98109, USA (fergal.duffy@seattlechildrens.org).

The Journal of Infectious Diseases® 2022;225:1832–40

DOI: 10.1093/infdis/jiab130

In contrast, if detectable infection is established in a mouse, the animal will ultimately fail to control the bacteria [17]. It is therefore impossible to study the immune impact of contained *M. tuberculosis* infection (CMTI) on the peripheral blood transcriptome in mice after aerosol challenge because the infection is never contained. This is a significant limitation of using the mouse model for tuberculosis biomarker development as well as for mechanistic studies.

The development of a CMTI model [18, 19] has overcome many of these limitations. In mice, intradermal inoculation of the ear with approximately 10 000 CFUs of *M. tuberculosis* establishes an infection in the draining lymph node that is asymptomatic, stable for ≥1 year, and excluded from the lung [18, 19]. After the establishment of CMTI, mice are strongly protected against subsequent aerosol challenge, mimicking the well-established protective effect of prior *M. tuberculosis* infection in humans [19].

We hypothesized that CMTI might be a useful model for the development of biomarkers that predict outcomes of asymptomatic *M. tuberculosis* infection in humans. In concordance with studies of the ULD model [10], we found that peripheral blood transcriptional responses in mice with CMTI are similar to those measured in humans with active tuberculosis and are predictive of risk for progression to active disease. We also identified a set of genes strongly up-regulated in mice with CMTI the expression of which is not elevated in humans with active tuberculosis. When measured ≥18 months before diagnosis, elevated expression of these genes was associated with lower risk of progression to disease. In contrast, a previously identified correlate of risk (ACS-CoR) score [3] is not predictive of progression to active disease this early. Furthermore, when measured at the time of diagnosis, elevated expression of these genes was correlated with better treatment outcome.

METHODS

Establishment of CMTI in Mice

Intradermal infections to establish CMTI were performed as described elsewhere [18], with the following modifications: 10 000 CFUs of *M. tuberculosis* (H37Rv) in logarithmic phase growth in 10 μL of phosphate-buffered saline were injected intradermally with a 10-μL Hamilton syringe into mice anesthetized with ketamine.

CMTI Whole-Blood RNAseq

RNA isolation was performed using TRIzol (Invitrogen) with 2 sequential chloroform extractions, Glycoblue carrier (Thermo Fisher), isopropanol precipitation, and washes with 75% ethanol. RNA was quantified with the Bioanalyzer RNA 6000 Pico Kit (Agilent). Complementary DNA libraries for alveolar macrophages were constructed and amplified using the SMARTer Stranded Total RNA-Seq Kit version 2—Pico Input Mammalian (Clontech), according to the manufacturer's instructions. Complementary DNA for whole blood was prepared using the TruSeq Stranded messenger RNA kit (Illumina). Libraries were amplified and then sequenced on an Illumina NextSeq sequencer (2 × 75 base pairs; paired-end reads). Stranded paired-end reads of length 76 were preprocessed. For the Pico Input prep, the first 3 nucleotides of R2 (version 2 kit) were removed as described in the SMARTer Stranded Total RNA-Seq Kit—Pico Input Mammalian user manual (version 2: 063017); read ends consisting of ≥50 of the same nucleotide were removed. The remaining read pairs were aligned to the mouse genome (mm10) plus the *M. tuberculosis* H37Rv genome, using the GSNAP aligner [20] (version 2016-08-24), allowing for novel splicing. Concordantly mapping read pairs (average, 10–20 million per sample) that aligned uniquely were assigned to exons using the subread program (version 1.4.6.p4) [21] and gene definitions from Ensembl MusMusculus GRCm38.78 coding and noncoding genes. Differential expression was calculated using the edgeR [22] package from bioconductor.org [23]. The false discovery rate (FDR) was computed with the Benjamini-Hochberg algorithm. Raw and processed data are deposited in the National Center for Biotechnology Information Gene Expression Omnibus (GSE126355).

Preparation of Human Tuberculosis Risk and Disease Cohorts

The following data sets containing whole-blood transcriptomic measurements of human tuberculosis risk, disease, and treatment cohorts were downloaded from the National Center for Biotechnology Information Gene Expression Omnibus: the human tuberculosis disease cohorts GSE107991, GSE107992 [1, 24] and GSE37250 [2]; the Adolescent Cohort Study (ACS) and GC6-74 tuberculosis risk cohorts GSE79362 [3] and GSE94438 [4]; and the Catalysis tuberculosis treatment cohort GSE89403 [25]. RNA-seq data was normalized and differential expression calculated using the edgeR v3.3 package. Similarly, microarray data were background subtracted, quantile normalized, and $\log_2$-transformed and differential expression calculated using the R limma package. Signatures were translated across species by identification of gene homologues, using the Mouse Genome Informatics database ([http://www.informatics.jax.org/downloads/reports/HOM_MouseHumanSequence.rpt]).

Calculation of Signature Scores

CMTI signature score genes were selected as genes that were consistently highly expressed 28 and 42 days after CMTI establishment (FDR, <0.001; | log fold change | > 1.5) that were also detected as expressed in human data sets (Table 1). For each data set, CMTI-derived signature scores were calculated as the mean per-sample expression of all signature genes.

The ACS-CoR signature scores were calculated using the published pairwise support vector machine (SVM) parameters for each junction pair [3]. The signature was adapted for prediction on gene read counts (rather than junction read counts) by

substituting the corresponding gene counts for junction counts. In addition, instead of using a voting threshold of 0.5 to classify pair scores into votes, and calculating the average of pair votes, the average of pair scores was used as the output. This enabled calculation of signature scores on gene-wise RNAseq and microarray studies without reparameterization to account for differing sequencing depths. For RNAseq, signature predictions were made on gene counts, and for microarrays, probe intensity values were used. Areas under the receiver operating characteristic curve (AUCs) and accompanying 95% confidence intervals (CIs) were calculated using the R pROC v1.17 package [26].

RESULTS

CMTI in Mice Induces Sustained Blood Transcriptome Changes

To examine the effect of a contained mycobacterial infection on the whole-blood transcriptome in mice, we established CMTI, as described elsewhere [19]. We performed RNAseq analysis of blood obtained immediately before the establishment of CMTI (day 0), and 10, 28, and 42 days afterward. Sixty-eight genes were strongly differentially expressed for ≥1 time point compared with day 0 ($|\log_2(\text{fold change})| > 1.5$; FDR, <0.001) (Figure 1A–1C). These genes consisted of 2 broad groups: (1) a set of 22 genes that were transiently up-regulated at day 10 and (2) a set of 46 genes that were most strongly up-regulated at day 28 and remained highly up-regulated at day 42 (Figure 1D). We took the latter set of 46 genes to represent the long-term whole-blood transcriptome associated with low-level CMTI in mice.

Predicting Human Tuberculosis Outcomes with a Mouse-Derived CMTI Gene Expression Signature

We hypothesized that many of the systemic immune processes that are activated by CMTI in mice would be similar in humans harboring replicating bacteria and lead to similar transcriptional responses in the peripheral blood. Currently, there are several publicly available data sets containing whole-blood transcriptomic measurements of subjects with active tuberculosis (Table 1) (the cohorts reported by BCoR et al and Singhania al [1, 24] and by Kaforou et al 2], GSE107991, GSE107992 [1, 24]), as well as data sets comprising individuals at risk of progression to tuberculosis (the ACS and GC6-74 tuberculosis risk cohorts GSE79363 [3], GSE94438 [4]), and responding to tuberculosis treatment (the Catalysis tuberculosis treatment cohort GSE89403 [25]). Therefore, we constructed a gene expression signature from the 46 "long-term" genes and tested its ability to predict tuberculosis disease in humans. We determined the human homologues of each of the 46 long-term genes and retained those that were consistently expressed at a detectable level in all of the human data sets (23 genes in total; "consistently detected" genes in Figure 1D). The signature score for each subject, which we termed the *CMTI disease score* (CMTI-DS) was defined as the average expression of these 23 genes.

To test the ability of the CMTI-DS to identify tuberculosis disease in humans, we analyzed a set of 3 whole-blood transcriptional profiles from patients with active pulmonary tuberculosis, latent tuberculosis, and healthy controls [1, 24] (Table 1). Despite being derived from a mouse model, the CMTI-DS was highly effective (AUC, 0.92 [95% CI, .86–.97]) at discriminating individuals with active disease from those with latent tuberculosis and performed equivalently to the ACS-CoR (0.93 [.88–.99]) [3] (Figure 2A). We also tested the predictive power of the CMTI-DS using the data set reported by Kaforou et al [2], consisting of whole-blood expression profiles from both human immunodeficiency virus (HIV)–infected and HIV-uninfected individuals with active or latent tuberculosis in Malawi and South Africa (Table 1). Again, the CMTI-DS performed comparably to the ACS-CoR in discriminating active disease from latent tuberculosis in both HIV-infected and HIV-uninfected individuals (Figure 2B and 2C).

One of the most significant recent advances in tuberculosis biomarker development was the demonstration that the ACS-CoR score can predict the risk of progression to active

Table 1. Previously Published RNAseq Whole-Blood Profiles Used in Current Study

Cohort	Description	Samples Analyzed, No.
Berry et al [1] and Singhania et al[24]	Patients with active tuberculosis vs patients with latent tuberculosis and healthy controls	Tuberculosis: 37; latent tuberculosis: 64; healthy controls: 12
Kaforou et al [2]	Patients with active tuberculosis vs patients with latent tuberculosis, including HIV-infected and HIV-uninfected individuals	Malawi: HIV uninfected, 51 with tuberculosis and 35 with latent tuberculosis; HIV infected, 51 with tuberculosis and 36 with latent tuberculosis South Africa: HIV uninfected, 46 with tuberculosis and 48 with latent tuberculosis; HIV infected, 47 with tuberculosis and 48 with latent tuberculosis
ACS (Zak et al) [3]	Longitudinal study (2-y follow-up) of adolescents with latent tuberculosis, some of progressing to tuberculosis	Progressor samples: 94; control samples (nonprogressors): 245
GC6-74 (Suliman et al) [4]	Longitudinal study (2-y follow-up) of household contacts of tuberculosis index case patients, some progressing to tuberculosis	Progressor samples: 82; control samples (nonprogressors): 282
Catalysis Treatment Cohort (Thompson et al) [25]	Longitudinal study (24 wk) of tuberculosis treatment responses	Definite cures: 78; not cured: 7 (4 time points each: diagnosis, d 7, wk 4, and wk 24)

Abbreviation: HIV, human immunodeficiency virus; ACS: Adolescent Cohort Study.

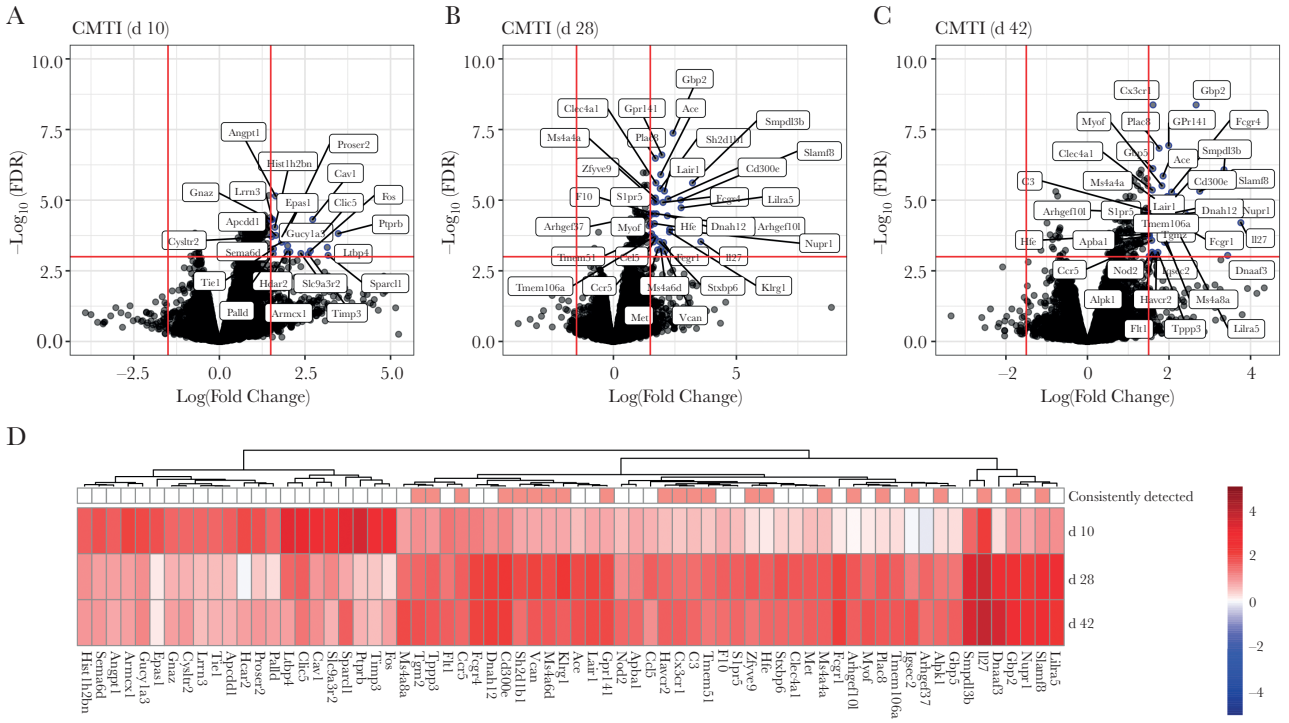

Figure 1. Establishment of mouse contained infection leads to alterations in blood transcriptional state. *A–C,* Volcano plots of differentially expressed genes comparing transcriptional states after contained *Mycobacterium tuberculosis* infection (CMTI) initiation (days 10 [*A*], 28 [*B*], and 42 [*C*]) to those before CMTI initiation (day 0). Abbreviation: FDR, false discovery rate. *D,* Heat map of significant differential expression changes shown in *A–C*. Genes showing high expression at days 28–42 and consistently detected in multiple human data sets are indicated by the "consistently detected" annotation bar above the heat map.

tuberculosis up to 18 months before clinical diagnosis. Given that the CMTI-DS performed equivalently to the ACS-CoR at identifying active disease, we tested whether it also predicted risk of progression. CMTI-DS is significantly predictive of tuberculosis progression in the ACS cohort (Table 1) (AUC, 0.65 [95% CI, .58– .71]), although with reduced performance versus ACS-CoR (0.8 [.74–.85]) (Figure 2D and Supplementary Figure 1). This reduced performance compared with ACS-CoR is not surprising, as ACS-CoR was originally developed and trained on these samples. Importantly, we found that this CMTI-DS signature had nearly equivalent predictive power compared with the ACS-CoR for predicting progression to active disease in the independent GC6-74 household contact risk cohort (Table 1) (AUCs for CMTI-DS vs ACS-CoR, 0.67 [95% CI, .6–.74] vs 0.71 [.64–.78]) (Figure 2D and Supplementary Figure 1).

Association of CMTI Disease Signature Genes with Bacterial Containment

The CMTI-DS signature was strongly associated with tuberculosis disease in humans. However, in mice, CMTI is asymptomatic and does not reduce lifespan [18, 19]. Moreover, mice with CMTI are protected against subsequent *M. tuberculosis* aerosol challenge [19]. Therefore, we hypothesized that in addition to a component that reflects the presence of replicating bacteria, the whole-blood transcriptome of mice with CMTI contains a component that reflects immune mechanisms associated with successful bacterial containment.

We sought to identify genes that differed in their expression pattern between CMTI in mice and active tuberculosis in humans and then set out to assess their association with improved tuberculosis outcomes in independent human tuberculosis risk cohorts. The expression profiles of the 23 mouse CMTI-DS genes were compared with those of their human homologues in healthy individuals or individuals with latent or active tuberculosis (Figure 3). We labeled as "concordant" genes significantly up-regulated in mice with CMTI versus naive mice and in humans with tuberculosis versus latent tuberculosis. Genes significantly up-regulated in mice with CMTI versus naive mice that also showed reduced expression in humans with tuberculosis versus latent tuberculosis were labeled as "opposite," and all other genes not fitting either of these patterns as "ambiguous."

Thus, the CMTI-DS genes were divided into groups whose expression profiles are concordant (12 of 23 genes), opposite (5 of 23), or ambiguous (6 of 23) between CMTI in mice and tuberculosis disease in humans. None of the 5 opposite genes (*CCR5, TPPP3, ZFYVE9, SH2D1B,* and *KLRG1*) are included in the ACS-CoR score, which primarily consists of type I/II interferon responsive genes that are up-regulated in active

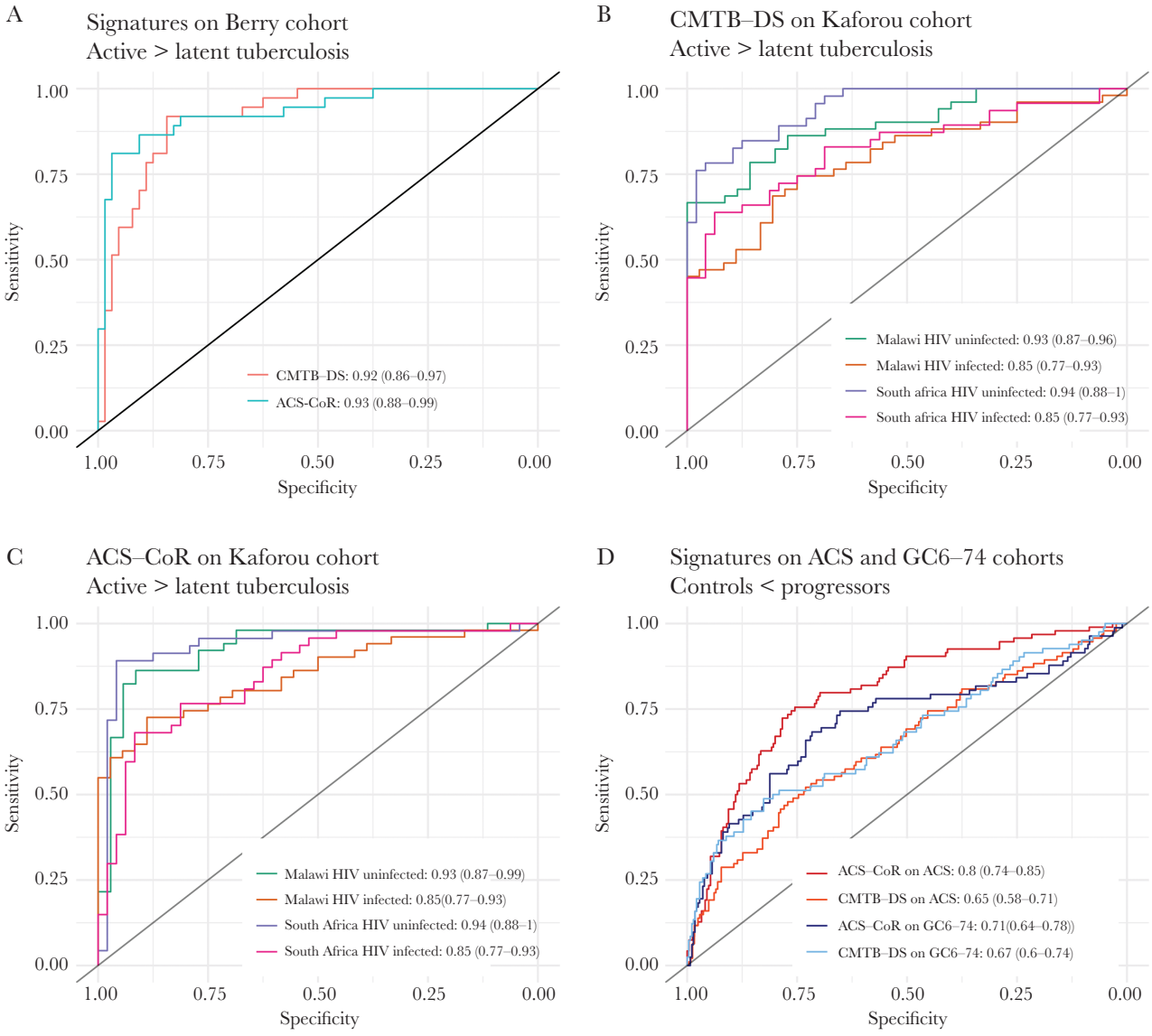

Figure 2. The contained *Mycobacterium tuberculosis* infection disease score (CMTI-DS) signature predicts human tuberculosis outcomes. *A-C*, Receiver operating characteristic (ROC) curves for the CMTI-DS and Adolescent Cohort Study–correlate of risk (ACS-CoR) signatures discriminating active from latent tuberculosis in the cohorts reported by Berry et al [1] (*A*) and Kaforou et al [2] (*B, C*). Abbreviation: HIV, human immunodeficiency virus. *D*, ROC curves for the CMTI-DS and ACS-CoR signatures tuberculosis progressors from nonprogressors in the ACS and CG6 cohorts Areas under the ROC curve and accompanying 95% confidence intervals are shown.

tuberculosis. Modular analysis showed that, while concordant genes were primarily annotated as members of inflammation, interferon, monocyte, and DC modules, none of the opposite genes fell into these modules and were instead associated with cytotoxic/natural killer (NK) cell and chemokine processes (Supplementary Table 1). We hypothesized that elevated expression of these 5 opposite genes in mice with CMTI was correlated with a protective immune response acting to contain infecting bacteria and sought to determine whether a signature defined by their average expression (CMTI containment or CMTI-CT; see Methods) could predict protection from tuberculosis disease in humans.

Association of CMTI-CT Signature With Long-Term Disease Containment and Effective Treatment in Longitudinal Human Cohorts

While it is not possible to experimentally test protective immunity in humans, longitudinal studies of high-risk cohorts allowed us to investigate CMTI-CT score variation over time in humans exposed to tuberculosis. The performance of the ACS-CoR signature has been shown to decline substantially as the time to diagnosis increases and it is unable to discriminate tuberculosis progressors from nonprogressors ≥18 months before tuberculosis diagnosis (Figure 4A). In contrast, CMTI-CT scores are elevated in nonprogressors compared with tuberculosis progressors ≥18 months from tuberculosis diagnosis in

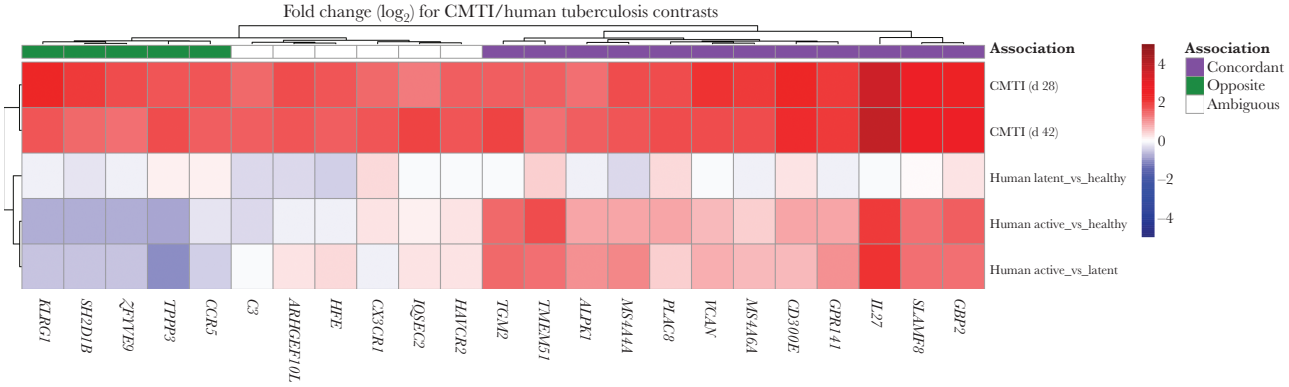

Figure 3. Contained *Mycobacterium tuberculosis* infection (CMTI) containment signature genes. Heat map shows hierarchical clustering of $\log_2$-transformed fold changes for contained infection signature genes (*columns*) for each comparison (*rows*). Positive fold changes are shown in red, and negative changes in blue. Genes are annotated as "concordant," "opposite," or "ambiguous," based on the consistency of differential expression between CMTI and active versus latent human tuberculosis.

the ACS cohort (AUC, 0.78 [95% CI, .59–.97]) (Figure 4A), and is significantly up-regulated in tuberculosis nonprogressors in the GC6-74 cohort (Supplementary Figure 2). Furthermore, CMTI-CT scores and time to diagnosis in ACS and GC6-74 tuberculosis progressors are only weakly correlated (ACS $r = 0.2$; $P = .09$) (Figure 4B and Supplementary Figure 2). In contrast, ACS-CoR scores were strongly correlated with time to disease diagnosis and showed a sharp increase beginning approximately 6 months before the onset of symptomatic disease ($r = -0.52$; $P < .001$) (Figure 4C). Thus, unlike ACS-CoR, CMTI-CT discriminates individuals likely to progress to tuberculosis from nonprogressors independent of subclinical disease processes. This is consistent with CMTI-CT's association with a more stable long-term phenotype, not a time-dependent response to tuberculosis infection.

Similarly, we hypothesized that tuberculosis treatment failure may be linked to reduced natural immunity to *M. tuberculosis* and used the data from the Catalysis tuberculosis treatment cohort [25] (Table 1) to investigate the relationship between CMTI-CT scores and treatment failure. The CMTI-CT score, measured at the time of diagnosis and before treatment initiation, is elevated in individuals who successfully respond to a 24-week course of treatment compared with controls (Supplementary Figure 3*A*) and is significantly predictive of treatment outcome (Figure 4D). In contrast, ACS-CoR is not predictive of treatment outcome at diagnosis, although it is significantly decreased in treatment successes 24 weeks after treatment initiation (Figure 4D and Supplementary Figure 3). Taken together, this points to CMTI-CT having predictive power independent of underlying bacterial burden in humans, thus consistent with natural immunity to tuberculosis.

DISCUSSION

In the current study, we have used a CMTI model in mice to develop a peripheral blood gene expression signature that is as predictive of human tuberculosis disease states as a signature derived directly from human data. The comparable performance of the CMTI-DS and ACS-CoR signatures supports our 2 core hypotheses: (1) the CMTI mouse model mimics at least a portion of the tuberculosis disease spectrum in humans, and (2) current peripheral-blood transcriptional signatures that predict tuberculosis disease state in humans reflect the immune response to live bacteria. Our results are consistent with findings of other studies that have uncovered strong similarities between the peripheral blood transcriptional response to *M. tuberculosis* infection in mice and humans [8–10]. In particular, the CMTI-DS signature is comparable to the recently published ULD signature [10], being a mouse-derived tuberculosis phenotype applied to human cohorts. While none of the ULD signature genes overlap with CMTI-DS, both ULD and CMTI-DS signatures contain several genes associated with interferon and inflammation (Supplementary Table 1). However, to our knowledge, this is the first study to specifically explore the predictive accuracy of a transcriptional signature derived from the CMTI model. This result also suggests that the increased experimental control afforded by a well-defined mouse model offsets the uncertainty that is introduced by translating between species.

It is generally assumed that *M. tuberculosis* infection is a spectrum of conditions ranging from asymptomatic exposure and sterilizing immunity to highly infectious pulmonary tuberculosis, and clinical data suggest that tuberculosis-infected individuals may move back and forth across the disease continuum [27]. Pioneering studies in nonhuman primates using positron emission tomography–computed tomography have shown that within the same host, some lesions heal while other lesions increase in size [28]. Similar processes have been observed in humans [14, 29]. Thus, it appears that within the same host, progression and containment can take place simultaneously. The expression profile of whole blood is therefore likely

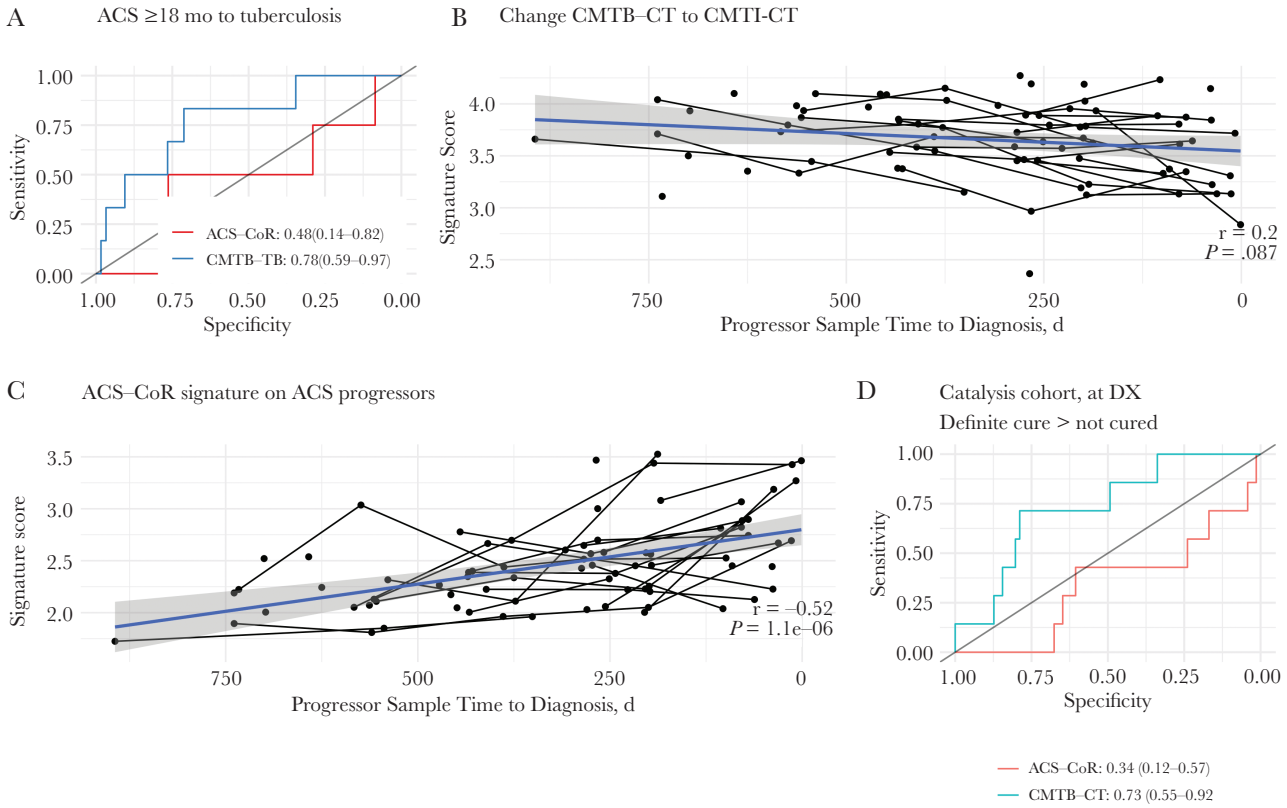

Figure 4. The contained *Mycobacterium tuberculosis* infection (CMTI) containment (CMTI-CT) signature is increased in nonprogressors in a time-independent manner. *A,* Receiver operating characteristic (ROC) curves for the CMTI-CT signature discriminating tuberculosis nonprogressors from tuberculosis progressors ≥18 months from tuberculosis diagnosis in the Adolescent Cohort Study. ROC curves were calculated assuming tuberculosis nonprogressor scores were greater than those of progressors for CMTI-CT, and tuberculosis progressor scores were greater than those of nonprogressors for ACS–correlate of risk (CoR). *B, C,* Signature scores by time to tuberculosis for ACS tuberculosis progressors for each CMTI-CT and ACS-CoR. Points connected by lines represent multiple samples from the same individual progressor, while individual points represent progressors from whom only a single sample was collected. *D,* ROC curves for the CMTI-CT and ACS-CoR signature discriminating treatment success (definite cures) from failures (not cured) at the time of tuberculosis diagnosis, before treatment initiation in the Catalysis tuberculosis treatment cohort.

to reflect a mixture of immune responses associated with both disease progression and containment, and any whole-blood signature will reflect the composition and activation of multiple cell types circulating in the periphery.

Indeed, despite the fact that mice with CMTI are protected against aerosol challenge, the CMTI-DS signature, which is composed of the most strongly differentially expressed genes in CMTI mice compared with controls, is elevated in humans with active tuberculosis. However, subsequent comparison with a human active tuberculosis cohort revealed that a subset of genes up-regulated in CMTI mice but not in human active tuberculosis. We combined these genes into a new signature, the CMTI-CT score that we could then test for association with improved outcomes in human tuberculosis risk and treatment cohorts. We confirmed that expression of the CMTI-CT gene signature is elevated in *M. tuberculosis*–infected individuals capable of controlling the bacteria compared with those who will progress to active tuberculosis. Unlike ACS-CoR, CMTI-CT significantly discriminated tuberculosis progressors from controls when measured >18 months before disease onset. Thus CMTI-CT appears to be correlated with a long-term protection or containment phenotype that is not associated with subclinical responses to replicating tuberculosis or bacterial burden.

It is tempting to speculate that the CMTI-CT signature is correlated with a natural disposition to tuberculosis control, potentially supporting an effective treatment response. However, these results are limited by the relatively small number of tuberculosis progressors ≥18 months from diagnosis in the ACS cohort (6 total), and the 2-year maximum follow-up time. The Vukuzazi study [30], an ongoing, population-based prospective study of a tuberculosis-endemic region in KwaZulu Natal, South Africa, will follow up enrolled individuals for 4 years and is powered to both validate and discover long-term correlates of tuberculosis immunity independent of short-term tuberculosis disease–associated responses.

The CMTI-CT signature includes 2 genes, *KLRG1* and *CCR5*, that have been investigated in the context of *M. tuberculosis* infection. Memory-like NKp46^{+}CD27^{+}KLRG1^{+} NK cells have been observed to be expanded after BCG vaccination and protective against *M. tuberculosis* challenge in a murine model

[31]. *CCR5* has been associated with control of bacterial migration to lymph nodes in mice [32]. The other 3 CMTI-CT genes (*TPPP3, ZFYVE9,* and *SH2D1B*) have not been investigated in the context of tuberculosis, however, they are known to play roles in immune and proliferative signaling:, *ZFYVE9* in the transforming growth factor β pathway [33], *TPPP3* in the AKT-STAT3 pathway [34], and *SH2D1B*, as a regulator of signal transduction in antigen-presenting cells [35]. This is in contrast with the ACS-CoR signature, which principally consists of genes associated with type I/II interferon responses.

In conclusion, the findings of the current study, as well as other recent work [10] demonstrate that well-defined and experimentally accessible mouse *M. tuberculosis* infection models can be used to generate whole-blood transcriptional signatures that are translatable to humans. Because much smaller human cohorts are required to test a proposed predictive signature than are required to discover the same signature, the current findings highlights an opportunity for more rapid and cost-effective development of biomarkers for tuberculosis disease and disease progression.

Supplementary Data

Supplementary materials are available at *The Journal of Infectious Diseases* online. Consisting of data provided by the authors to benefit the reader, the posted materials are not copyedited and are the sole responsibility of the authors, so questions or comments should be addressed to the corresponding author.

Notes

Disclaimer. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Financial support. This work was supported by the National Institute of Allergy and Infectious Diseases (grant U19AI135976), the National Institutes of Health (grants P41 GM109824, R01AI032972, U19AI35976, and IMPAc-TB BAA-NIAID-NIHAI201700104), and the Swiss National Science Foundation (grant P300PB_164742 to J. N.).

Potential conflicts of Interest. All authors: No reported conflicts. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

1. BCoR MP, Graham CM, McNab FW, et al. An interferon-inducible neutrophil-driven blood transcriptional signature in human tuberculosis. Nature **2010**; 466:973–7.
2. Kaforou M, Wright VJ, Oni T, et al. Detection of tuberculosis in HIV-infected and -uninfected African adults using whole blood RNA expression signatures: a case-control study. PLoS Med **2013**; 10:e1001538.
3. Zak DE, Penn-Nicholson A, Scriba TJ, et al; ACS and GC6-74 cohort study groups. A blood RNA signature for tuberculosis disease risk: a prospective cohort study. Lancet **2016**; 387:2312–22.
4. Suliman S, Thompson EG, Sutherland J, et al; GC6-74 cohort study team, the ACS cohort study team. Four-gene Pan-African blood signature predicts progression to tuberculosis. Am J Respir Crit Care Med **2018**; 197:1198–208.
5. Warsinske H, Vashisht R, Khatri P. Host-response-based gene signatures for tuberculosis diagnosis: A systematic comparison of 16 signatures. PLoS Med **2019**; 16:e1002786.
6. Sweeney TE, Braviak L, Tato CM, Khatri P. Genome-wide expression for diagnosis of pulmonary tuberculosis: a multicohort analysis. Lancet Respir Med **2016**; 4:213–24.
7. Duffy FJ, Thompson EG, Scriba TJ, Zak DE. Multinomial modelling of TB/HIV co-infection yields a robust predictive signature and generates hypotheses about the HIV+TB+ disease state. PLoS One **2019**; 14:e0219322.
8. Ahmed M, Thirunavukkarasu S, Rosa BA, et al. Immune correlates of tuberculosis disease and risk translate across species. Sci Transl Med **2020**; 12:eaay0233.
9. Moreira-Teixeira L, Tabone O, Graham CM, et al. Mouse transcriptome reveals potential signatures of protection and pathogenesis in human tuberculosis. Nat Immunol **2020**; 21:464–76.
10. Plumlee CR, Duffy FJ, Gern BH, et al. Ultra-low dose aerosol infection of mice with *Mycobacterium tuberculosis* more closely models human tuberculosis. Cell Host Microbe **2021**; 29:1–15.
11. Scriba TJ, Penn-Nicholson A, Shankar S, et al; other members of the ACS cohort study team. Sequential inflammatory processes define human progression from M. tuberculosis infection to tuberculosis disease. PLoS Pathog **2017**; 13:e1006687.
12. Houben RM, Dodd PJ. The global burden of latent tuberculosis infection: a re-estimation using mathematical modelling. PLoS Med **2016**; 13:e1002152.
13. Behr MA, Edelstein PH, Ramakrishnan L. Is *Mycobacterium tuberculosis* infection life long? BMJ **2019**; 367:l5770.
14. Cadena AM, Fortune SM, Flynn JL. Heterogeneity in tuberculosis. Nat Rev Immunol **2017**; 17:691–702.
15. Andrews JR, Noubary F, Walensky RP, Cerda R, Losina E, Horsburgh CR. Risk of progression to active tuberculosis following reinfection with *Mycobacterium tuberculosis.* Clin Infect Dis **2012**; 54:784–91.
16. Blaser N, Zahnd C, Hermans S, et al. Tuberculosis in Cape Town: an age-structured transmission model. Epidemics **2016**; 14:54–61.
17. Beamer GL, Turner J. Murine models of susceptibility to tuberculosis. Arch Immunol Ther Exp (Warsz) **2005**; 53:469–83.

18. Kupz A, Zedler U, Stäber M, Kaufmann SH. A Mouse model of latent tuberculosis infection to study intervention strategies to prevent reactivation. PLoS One **2016**; 11:e0158849.
19. Nemeth J, Olson GS, Rothchild AC, et al. Contained *Mycobacterium tuberculosis* infection induces concomitant and heterologous protection. PLoS Pathog **2020**; 16:e1008655.
20. Wu TD, Reeder J, Lawrence M, Becker G, Brauer MJ. GMAP and GSNAP for genomic sequence alignment: enhancements to speed, accuracy, and functionality. Methods Mol Biol **2016**; 1418:283–334.
21. Liao Y, Smyth GK, Shi W. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. Nucleic Acids Res **2019**; 47:e47.
22. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics **2010**; 26:139–40.
23. Reimers M, Carey VJ. Bioconductor: an open source framework for bioinformatics and computational biology. Methods Enzymol **2006**; 411:119–34.
24. Singhania A, Verma R, Graham CM, et al. A modular transcriptional signature identifies phenotypic heterogeneity of human tuberculosis infection. Nat Commun **2018**; 9:2308.
25. Thompson EG, Du Y, Malherbe ST, et al; Catalysis TB–Biomarker Consortium. Host blood RNA signatures predict the outcome of tuberculosis treatment. Tuberculosis (Edinb) **2017**; 107:48–58.
26. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics **2011**; 12:77.
27. Barry CE 3rd, Boshoff HI, Dartois V, et al. The spectrum of latent tuberculosis: rethinking the biology and intervention strategies. Nat Rev Microbiol **2009**; 7:845–55.
28. Lin PL, Ford CB, Coleman MT, et al. Sterilization of granulomas is common in active and latent tuberculosis despite within-host variability in bacterial killing. Nat Med **2014**; 20:75–9.
29. Malherbe ST, Shenai S, Ronacher K, et al; Catalysis TB–Biomarker Consortium. Persisting positron emission tomography lesion activity and *Mycobacterium tuberculosis* mRNA after tuberculosis cure. Nat Med **2016**; 22:1094–100.
30. Ngwenya N, Luthuli M, Gunda R, et al; Vukuzazi team. Participant understanding of informed consent in a multidisease community-based health screening and biobank platform in rural South Africa. Int Health **2020**; 12:560–6.
31. Venkatasubramanian S, Cheekatla S, Paidipally P, et al. IL-21-dependent expansion of memory-like NK cells enhances protective immune responses against *Mycobacterium tuberculosis.* Mucosal Immunol **2017**; 10:1031–42.
32. Algood HM, Flynn JL. CCR5-deficient mice control *Mycobacterium tuberculosis* infection despite increased pulmonary lymphocytic infiltration. J Immunol **2004**; 173:3287–96.
33. Lin HK, Bergmann S, Pandolfi PP. Cytoplasmic PML function in TGF-beta signalling. Nature **2004**; 431:205–11.
34. Zhou W, Li J, Wang X, Hu R. Stable knockdown of *TPPP3* by RNA interference in Lewis lung carcinoma cell inhibits tumor growth and metastasis. Mol Cell Biochem **2010**; 343:231–8.
35. Morra M, Lu J, Poy F, et al. Structural basis for the interaction of the free SH2 domain EAT-2 with SLAM receptors in hematopoietic cells. EMBO J **2001**; 20:5840–52.