UCSF
UC San Francisco Previously Published Works

## Title

A Novel, 5-Transcript, Whole-blood Gene-expression Signature for Tuberculosis Screening Among People Living With Human Immunodeficiency Virus.

## Permalink

https://escholarship.org/uc/item/4qx290mv

## Journal

Clinical Infectious Diseases, 69(1)

## ISSN

1058-4838

## Authors

Rajan, Jayant V
Semitala, Fred C
Mehta, Tejas
et al.

## Publication Date

2019-06-18

## DOI

10.1093/cid/ciy835

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Clinical Infectious Diseases
MAJOR ARTICLE

# A Novel, 5-Transcript, Whole-blood Gene-expression Signature for Tuberculosis Screening Among People Living With Human Immunodeficiency Virus

**Jayant V. Rajan,[1] Fred C. Semitala,[2] Tejas Mehta,[3] Mark Seielstad,[4] Lani Montalvo,[5] Alfred Andama,[6] Lucy Asege,[6] Martha Nakaye,[6] Jane Katende,[6] Sandra Mwebe,[6] Moses R. Kamya,[2] Christina Yoon,[3] and Adithya Cattamanchi[3]**

[1]Division of Experimental Medicine, Department of Medicine, Zuckerberg San Francisco General Hospital, University of California, San Francisco; [2]Department of Medicine, Makerere University School of Medicine, Kampala, Uganda; [3]Division of Pulmonary and Critical Care Medicine, Department of Medicine, Zuckerberg San Francisco General Hospital and [4]Institute for Human Genetics, Department of Laboratory Medicine, Department of Epidemiology and Biostatistics, University of California, San Francisco, and [5]Blood Systems Research Institute, San Francisco, California; and [6]Infectious Diseases Research Collaboration, Kampala, Uganda

***Background.*** Gene-expression profiles have been reported to distinguish between patients with and without active tuberculosis (TB), but no prior study has been conducted in the context of TB screening.

***Methods.*** We included all the patients (n = 40) with culture-confirmed TB and time-matched controls (n = 80) enrolled between July 2013 and April 2015 in a TB screening study among people living with human immunodeficiency virus (PLHIV) in Kampala, Uganda. We randomly split the patients into training (n = 80) and test (n = 40) datasets. We used the training dataset to derive candidate signatures that consisted of 1 to 5 differentially-expressed transcripts ($P \leq .10$) and compared the performance of our candidate signatures with 4 published TB gene-expression signatures, both on the independent test dataset and in 2 external datasets.

***Results.*** We identified a novel, 5-transcript signature that met the accuracy thresholds recommended for a TB screening test. On the independent test dataset, our signature had an area under the curve (AUC) of 0.87 (95% confidence interval [CI] 0.72–0.98), with sensitivity of 94% and specificity of 75%. None of the 4 published TB signatures achieved desired accuracy thresholds. Our novel signature performed well in external datasets from both high (AUC 0.81, 95% CI 0.74–0.88) and low (0.81, 95% CI 0.77–0.85) TB burden settings.

***Conclusions.*** We identified the first gene-expression signature for TB screening. Our signature has the potential to be translated into a point-of-care test to facilitate systematic TB screening among PLHIV and other high-risk populations.

**Keywords.** tuberculosis; HIV; screening.

The World Health Organization (WHO) has targeted a 90% reduction in tuberculosis (TB) incidences and a 95% reduction in TB-related mortality by 2035 [1, 2]. The current rates of decline are insufficient to achieve these goals, and prevalence surveys in multiple countries have shown a large burden of undiagnosed TB in their communities. Therefore, the WHO now recommends systematic screening for TB in high-risk groups [3]. In contrast to passive case-finding, systematic screening involves provider-initiated assessments of TB risks, followed by diagnostic testing of patients deemed to be at-risk. However, the uptake of systematic TB screening has been limited to date, due to the lack of an adequate risk assessment (ie, screening) strategy or test.

The WHO target product profile for a TB screening test recommends that it have at least 90% sensitivity and 70% specificity [4] in order to ensure that few patients with TB are missed and to limit unnecessary confirmatory diagnostic testing. The current, symptom-based TB screening strategy for people living with human immunodeficiency virus (PLHIV) has high sensitivity, but specificity has ranged from 5–33% in studies from sub-Saharan Africa [5]. Thus, symptom-based screening identifies nearly all PLHIV as being at-risk and requiring confirmatory diagnostic testing. Chest radiography improves specificity, but requires trained readers and infrastructure that is not routinely available in community settings [6]. Identification of a non-sputum, biomarker-based screening test is, therefore, among the highest priorities for TB diagnostics research.

Whole-blood gene-expression profiling is being evaluated as a potential biomarker for identifying patients with TB. A 2010 study reported that gene-expression signatures distinguished active TB from latent tuberculosis infection (393 genes) and active TB from other infectious and inflammatory conditions

Received 21 May 2018; editorial decision 20 September 2018; accepted 15 November 2018; published online November 20, 2018.

Correspondence: A. Cattamanchi, Division of Pulmonary and Critical Care Medicine, Department of Medicine, Zuckerberg San Francisco General Hospital, 1001 Potrero Avenue, San Francisco, CA 94110 (Adithya.Cattamanchi@ucsf.edu).

**Clinical Infectious Diseases® 2019;69(1):77–83**

© The Author(s) 2018. Published by Oxford University Press for the Infectious Diseases Society of America. All rights reserved. For permissions, e-mail: journals.permissions@oup.com.

DOI: 10.1093/cid/ciy835

(86 genes) [7]. Subsequent studies have evaluated gene expression in populations more representative of those patients undergoing TB evaluations in high-burden countries [8, 9]. These studies identified signatures that had high sensitivity (>90%) and moderate specificity (80–90%), and that outperformed the 2010 signatures. More recently, various groups have focused on identifying smaller gene signatures that could more readily be translated into point-of-care tests. Sweeney et al. combined data from 13 published gene-expression datasets, primarily from case-control studies, to identify a 3-gene signature (sensitivity 85%, specificity 93%) [10]. Another group derived a single-gene signature (sensitivity 96%, specificity 76%) from a cohort of patients in a low TB-incidence setting that outperformed all previously-reported signatures on external-validation datasets [11].

A common feature of published TB gene signatures is that they have high sensitivity but only moderate specificity. This suggests that gene-expression signatures are more likely to be useful for TB screening than for TB diagnosis, which requires high specificity (≥98%), such that clinicians can confidently treat patients for TB based on a positive test result. Here, we report the results of the first study to evaluate whole-blood gene expression in the context of systematic TB screening. We focus on PLHIV, a key high-risk group for which the WHO recommends systematic TB screening [3]. Our objective was to identify the best-performing small (≤ 5 transcripts) gene-expression signature, and to assess its accuracy in comparison to the accuracy thresholds recommended for a TB screening test, as well as previously-published TB gene-expression signatures.

## METHODS AND FINDINGS

### Study Design

We conducted a case-control study that was nested within a TB screening cohort of human immunodeficiency virus (HIV)-infected adults (age ≥ 18 yeas) initiating antiretroviral therapy at 2 prototypical, urban, outpatient HIV clinics in Kampala, Uganda [12]. From July 2013 to April 2015, the parent study recruited 665 consecutive PLHIV with CD4 counts ≤350 cells/µl. All participants underwent comprehensive TB testing, including liquid mycobacterial cultures on 2 sputum samples at baseline. Patients were classified as having active TB if *Mycobacterium tuberculosis* was isolated from at least 1 sputum culture. For this nested case-control study, we included all 40 participants with culture-confirmed TB as cases and selected 2 controls for each case, which were matched by study enrollment date.

### Ethics Statement

This work was approved by the University of California, San Francisco, Committee on Human Research and the Makerere University School of Medicine Research Ethics Committee. Informed consent was obtained from all study participants using an approved consent form.

### Sample Collection, Storage, and Processing

All participants had 2.5 mL of venous blood drawn into a PAXGene tube at enrollment. Tubes were stored at -80°C until RNA extraction and were processed in batches of 15, as previously described, with both cases and controls represented in each batch [13]. RNA samples were quantitated and quality-assessed by BioAnalyzer (Agilent), after which they were amplified using the TotalPrep RNA amplification kit (Illumina). The final complementary RNA (cRNA) was purified, quantified, and quality-tested, as described [13], and all samples were adjusted to a final concentration of 150 ng/uL.

### Microarray Gene-expression Profiling

We sent 10 uL cRNA samples to the University of California, San Francisco, Microarray core for hybridization using the HumanHT-12 v4.0 Beadchip platform (Illumina). Raw array data was processed using Genome Studio V2009.1 and Gene Expression Module V1.1.1 (Illumina). Array data were full-quantile normalized and log-transformed using the R *beadarray* package.

### Identification of Novel Tuberculosis Gene-expression Signatures

We randomly partitioned the 120 patients into training (n = 80) and test (n = 40) datasets. We first eliminated all transcripts that had a detection $P$ value of ≥.05 across ≥90% of training dataset samples. We then calculated a Student's $t$ statistic for each transcript in this reduced set of features, comparing the mean expression value of each transcript in patients with and without TB. $P$ values were corrected for multiple-hypothesis testing using the Benjamini and Hochberg method [14]. We used a corrected $P$-value threshold of ≤.10 to select transcripts to be used for training novel gene-expression signatures.

Next, we created nested sets of gene-expression signatures, consisting of 1–5 transcripts that were based on signed and unsigned sums of the expression levels of each transcript included in the signature. For signed sums, expression values were multiplied by -1 if they decreased in patients with TB. Signatures consisting of 2–5 transcripts were selected iteratively. For 2-transcript signatures, we identified a second transcript for each single transcript that maximized specificity at a sensitivity threshold of ≥90% on the training dataset. This threshold was chosen based on the WHO target product profile for a TB screening test [4]. The 3-transcript signatures were selected by adding a third transcript to all 2-transcript signatures, using the same procedure. This process resulted in a nested set of expression signatures, with each 1-transcript signature contained in each 2-transcript signature, each 2-transcript signature contained in each 3-transcript signature, and so on, for signatures consisting of up to 5 transcripts. We determined the minimum

number of transcripts to include in a signature and the performance of signed versus unsigned sums by examining the distribution of specificities of the nested set of 1- to 5-transcript signatures on the training dataset.

Last, we calculated receiver operator characteristic (ROC) curves on 500 bootstrapped validation sets, derived from the training dataset. For each candidate signature, we calculated the mean of the maximum specificities across all bootstrapped validation sets at a sensitivity threshold of ≥90%. Signatures that met a mean specificity threshold of ≥70% across the bootstrapped validation sets were retained as candidate signatures to be examined on the test dataset.

#### Performance of Novel Expression Signatures on Test Data

For the final set of candidate signatures, we calculated ROC curves on the independent test dataset. For those that met a specificity threshold of ≥70% at a set sensitivity threshold of 90%, we determined the area under the curve (AUC). For comparison, we calculated ROC curves and AUCs on the test dataset for the non-specific inflammatory marker C-reactive protein (CRP) [12, 15] and for previously-reported 1-transcript [11], 3-transcript [10], 44-transcript [8], and 144-transcript [9] signatures. Calculations used the methods described in the original publications. For all signatures, we calculated 95% confidence intervals (CI) for the AUC using 1000 bootstrapped test datasets.

#### External Validity of Novel Expression Signatures

From the National Center of Biotechnology Information's Genome Omnibus database, we obtained 2 previously-published expression-profiling datasets that were used to derive TB expression signatures (GSE37250 [8], GSE42834 [9]). We calculated ROC curves, AUC, and 95% CIs to assess the performance of our novel signature on these datasets, as already described.

## RESULTS

#### Patient Characteristics

There were 61 females (50.8%) and 59 males (49.2%) among the 120 patients included in this study. The median age was 33 years (interquartile range [IQR] 28–39) and the median CD4 count was 156 cells/ul (IQR 58–253).

#### Selection of Genes for Signature Training

Of the 47 323 transcripts represented on the HumanHT-12 v4.0 Beadchip platform, 23 136 were not detected across ≥90% of the training dataset samples and were excluded from further analyses. Of the remaining 24 187 transcripts, 337 (1.4%) had a mean expression value that was significantly different ($P < .10$) when comparing patients with and without TB. The median fold-change in the 337 differentially-expressed genes was 1.18 (IQR -0.80–1.37; Supplementary Figure 1). These 337 transcripts were used to derive gene-expression signatures.

#### Performance of Signed and Unsigned Sums on the Training Dataset

The median specificities at a set sensitivity threshold of ≥90% for signed sums consisting of 1–5 transcripts were 39.3% (1 transcript), 33.9% (2 transcripts), 35.7% (3 transcripts), 37.5% (4 transcripts), and 39.3% (5 transcripts; Figure 1, left panel). The median specificities for unsigned sums for the same numbers of transcripts were 37.5%, 62.5%, 75.0%, 82.1%, and 85.7%, respectively (Figure 1, right panel). Only unsigned sums consisting of 3 to 5 transcripts met the minimum specificity threshold for a TB screening test, and we thus focused on these to derive novel gene-expression signatures (Supplementary Figure 2).

#### Identification of Novel Gene-expression Signatures for Tuberculosis Screening

Across the 500 bootstrapped validation datasets derived from the training dataset, 417 signatures, consisting of 3 to 5 transcripts, had mean specificities of ≥70% at a set sensitivity threshold of 90%. This included 83 (24.6%) 3-transcript signatures, 152 (45.1%) 4-transcript signatures, and 182 (54%) 5-transcript signatures from among all 3–5 transcript combinations of the 337 transcripts selected for signature training (Figure 2).

Of these 417 signatures, only 1 met the required sensitivity and specificity benchmarks of 90% and 70%, respectively, for a TB screening test in the independent test dataset. This signature included 5 distinct transcripts: actin, alpha 2 (ACTA2); guanylate binding protein family member 6 (GBP6); glycogenin 1 (GYG1); mitochondrial translation factor 1 like (MTRF1L); RAB, member of RA oncogene family like 2A (RABL2A). The AUC for this novel signature was 0.87 (95% CI 0.72–0.98)

#### Comparison of Novel Signature to Other Reported Tuberculosis Gene-expression Signatures

None of the previously-described TB signatures evaluated achieved the sensitivity and specificity thresholds of ≥90% and ≥70%, respectively, required of a TB screening test in our independent test dataset (Figure 3) [8–11]. The AUCs ranged from 0.62 (0.41–0.83) for the 44-transcript signature described by Kaforou et al. to 0.82 (0.62–0.98) for the 144-transcript signature reported by Bloom et al. [8, 9]. At a threshold sensitivity of 90%, specificity ranged from 12.5% for the 144-transcript signature to 49.1% for the 44-transcript signature. In contrast, CRP had an AUC of 0.87, and came closest to achieving the target accuracy benchmarks (specificity of 83.4% at a maximum sensitivity of 87.5%).

#### External Validity of Novel, 5-Gene Signature

We measured the performance of our novel, 5-transcript signature in 2 previously-reported datasets, 1 representing a large cohort from Africa and 1 from the United Kingdom [8, 9]. Both included patients undergoing TB diagnostic testing rather than TB screening. Nonetheless, the AUC for our novel signature was 0.81 in both the African cohort (95% CI 0.74–0.88) and the United Kingdom cohort (0.77–0.85; Figure 4). At a threshold

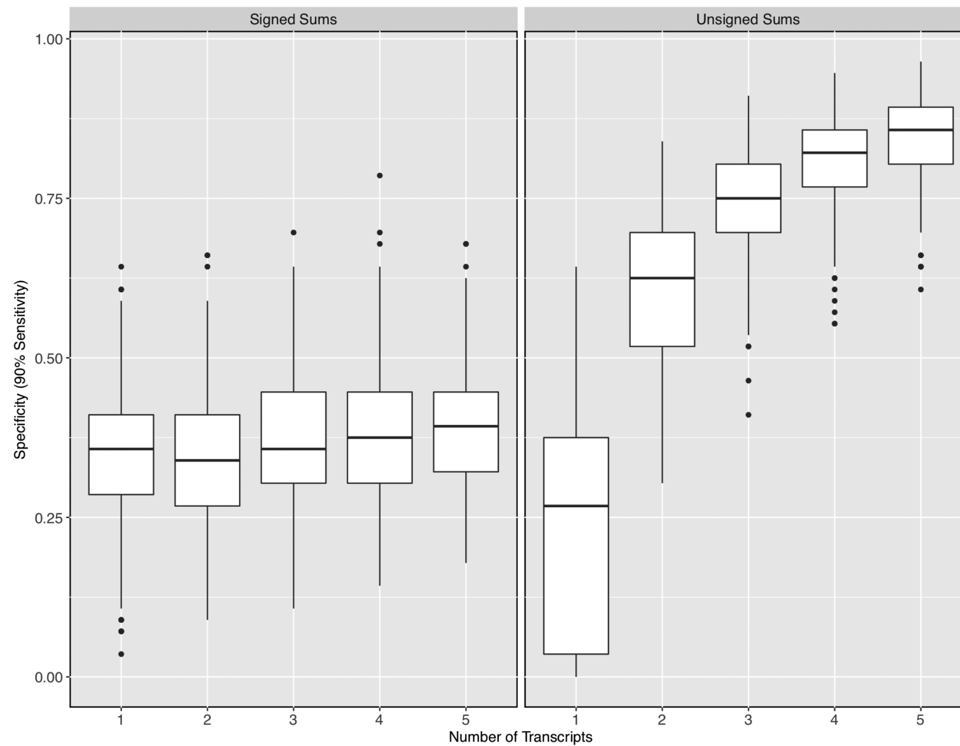


**Figure 1.** Unsigned sums of 3 to 5 transcripts met the specificity threshold for a screening test on the training data. We determined the best-performing set of 1–5 transcripts for each of the differentially-expressed transcripts in the final dataset (N = 337). This process resulted in a set of nested signatures. We computed a receiver operator characteristic curve for each of these signatures on the training dataset, and found that the only unsigned sums, consisting of 3 to 5 transcripts, performed well as groups, with a median specificity (at a sensitivity threshold of ≥90%) of ≥70%. Only these signatures were used to determine our final set of signatures.

sensitivity of ≥90%, the maximum specificity of our novel signature was 49.1% in the African cohort and 52.4% in the United Kingdom cohort.

## DISCUSSION

No current TB screening test meets the performance characteristics recommended by the WHO. Here, in the first study to evaluate gene-expression profiles in the context of TB screening, we identified a novel, 5-transcript signature that exceeded the minimum sensitivity and specificity thresholds recommended for a TB screening test in an independent test dataset. We verified that our novel signature outperformed 4 published TB gene-expression signatures in a screening cohort and, as expected, it maintained high sensitivity but had reduced specificity in external datasets that enrolled patients more likely to have systemic illnesses other than TB [8–11]. Our novel signature has strong potential to be developed into a point-of-care assay to enhance TB screening among high-risk groups.

We found that previously-published TB gene-expression signatures performed poorly in the context of TB screening [7–11]. The 4 signatures we examined here outperformed prior signatures, which included a 144-transcript signature derived among patients in low TB-prevalence settings (United Kingdom, France) [9]; a 44-transcript signature derived among patients who were self-reporting TB symptoms in 2 African countries [8]; a 3-transcript signature derived by combining data from multiple studies that enrolled either patients who self-reported TB symptoms or TB patients and healthy controls [10]; and a 1-gene signature derived by comparing TB patients pre-treatment and post-recovery [11]. Importantly, none of these signatures were derived or validated among patients undergoing TB screening, a context in which patients with TB are often identified at an earlier stage and patients without TB are less likely to have other systemic illnesses. Not surprisingly, each of the signatures performed considerably worse in our screening cohort than originally reported.

Our novel, parsimonious gene-expression signature was the only signature to achieve the minimum accuracy thresholds recommended for a TB screening test. In comparison to prior studies, we generally observed both a smaller number of differentially-expressed genes and smaller fold changes in gene expression. This finding could be related to a lower pathogen burden among patients undergoing TB screening rather than self-reporting TB symptoms, resulting in a weaker inflammatory response profile being reflected in gene-expression levels. The absence of large, inflammation-dependent changes in gene expression in our cohort could also explain why CRP, a

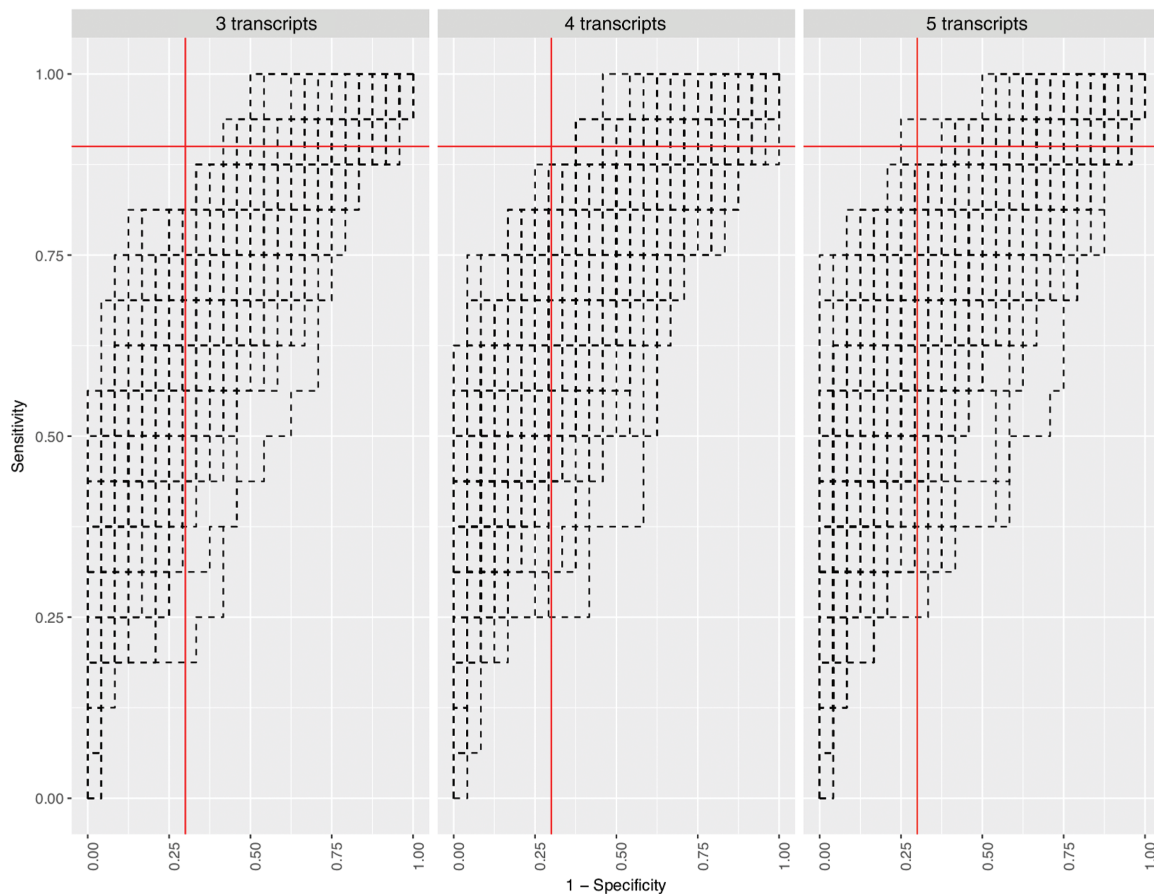


**Figure 2.** Performance of candidate 3–5 transcript signatures on the independent test dataset. Only a subset of 3–5 transcript signatures had mean specificities of ≥70% across 500 bootstrapped validation datasets (see Supplementary Figure 1). We examined the performance of each of these signatures on an independent test dataset by generating receiver operator characteristic curves. A single, 5-transcript signature met the minimum specificity threshold of ≥70% at a sensitivity threshold of ≥90%.

non-specific but sensitive marker of inflammation, performed comparatively well. Although the AUCs were the same for CRP and our novel signature, only our 5-transcript signature had a portion of the ROC curve that included points exceeding the desired sensitivity and specificity combinations for a TB screening test. Overall, our findings highlight the importance of deriving and validating gene-expression signatures for TB screening among the intended target population (ie, populations undergoing systematic TB screening).

Like other whole-blood gene-expression signature studies, the primary transcript signal in our signature appears to be from leukocytes. One of the genes (GBP6) is thought to have an undefined immune function and 2 (ACTA2 and RABL2A) are thought to play roles in cell-motility (ACTA2) and signal transduction (GBP6, RABL2A) [16–19]. GBP6 is present in the signatures of both Kaforou et al. and Bloom et al., and GYG1 and MTRF1L were also present in the signature of Bloom et al. [8, 9]. Further work is required to understand the biological basis of our novel signature. Our primary motivation here, however, was the development of a robust screening test that performs well in a real-world setting.

Our study has several important strengths. First, it was nested within a cohort specifically designed for the purpose of identifying novel tools for TB screening. In addition, our eligibility criteria were minimal and the nested case-control design helps ensure the sub-population examined here is reflective of the underlying study base (ie, PLHIV undergoing TB screening). Second, we directly optimized sums of gene-expression values for their predictive ability, rather than first identifying the best predictors of TB status and then summing gene-expression values of the best predictors to obtain the final signatures [8, 10, 11]. In doing so, we believe that we have generated platform-independent signatures that should be readily adaptable to technologies (such as quantitative real-time polymerase chain reaction) that can be used at the point of care. Last, we validated the performance of our signature and previously-published signatures in an independent test dataset.

Our study also has several limitations. It included only patients from a single, high-HIV/TB–burden country. However, we also examined the performance of our signature in 2 external datasets, though neither was from a TB screening cohort. A related limitation is that our signature is based on microarray gene-expression

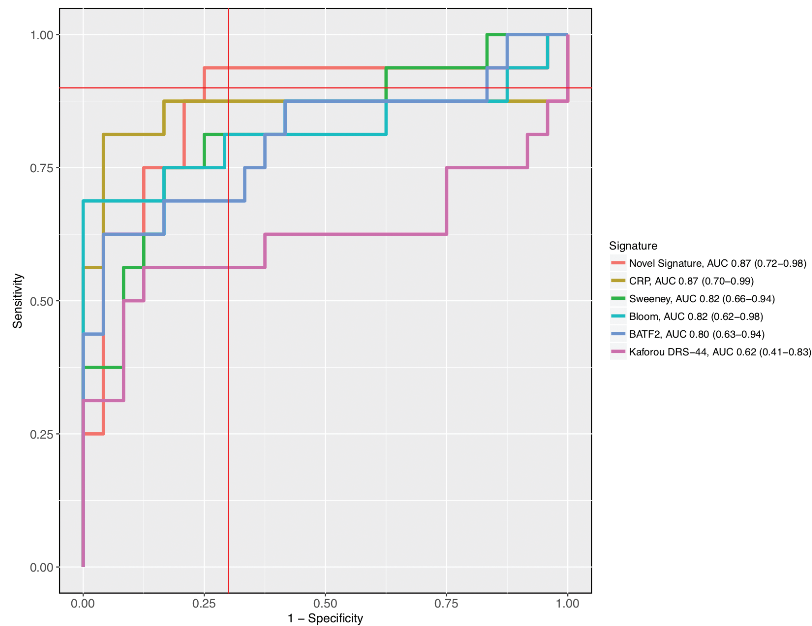


**Figure 3.** Comparison of performance of novel, 5-transcript TB screening signature with published signatures and CRP. We calculated receiver operator characteristic curves for our novel, 5-transcript signature; 4 previously-published TB gene-expression signatures; and the non-specific marker of inflammation, CRP, on our independent test dataset. Our signature outperformed each of the published signatures and CRP, as it was the only signature that met the minimum specificity threshold of ≥70% at a sensitivity threshold of ≥90%. Abbreviations: AUC, area under the curve; BATF2, basic leucine zipper transcription factor 2; CRP, C-reactive protein; DRS, disease risk score; TB, tuberculosis.

profiling. Hybridization-based technologies are known to be noisy and are not always reproducible when repeated using the technology that is most likely to be used to implement a signature: quantitative real-time polymerase chain reaction [20–23]. This lack of reproducibility is 1 of the factors that led to the use of large fold-change cutoffs in selecting differentially-expressed transcripts that can be used to derive a gene-expression signature. As we show here, however, this approach may not be practical in a screening setting, where large differences in transcript levels may not be present. Finally, repeat TB assessment was not part of

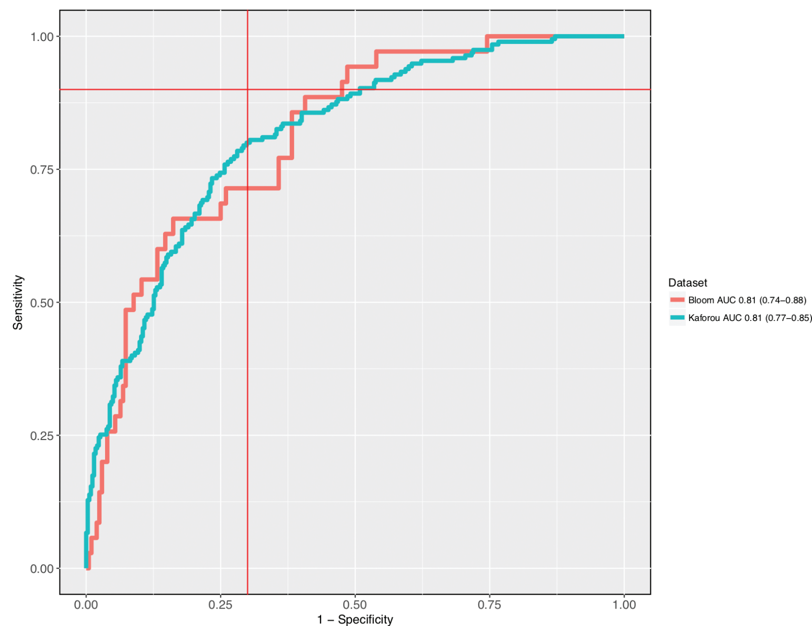


**Figure 4.** Performance of novel, 5-transcript, TB screening signature on 2 external TB diagnosis cohorts. We obtained data from 2 of the largest TB gene-signature cohorts published, from Bloom et al. and from Kaforou et al., representing non-endemic (Bloom) and endemic (Kaforou) settings [8, 9]. Our signature achieved the same AUC in each dataset, of 0.81, but had lower specificity (<70%). Abbreviations: AUC, area under the curve; TB, tuberculosis.

the study protocol at the time patients included in this sub-study were enrolled. We thus cannot rule out that some of the participants may have had sub-clinical TB, which has recently been shown to have its own distinct, transcriptional signature [24].

The novel, 5-gene signature that we describe here is an important contribution to the TB gene signature field. It is the only signature that has been derived from and tested in a TB screening setting and found to perform well. It is also a small signature, making it readily adaptable to point-of-care technologies. With further validation, it has the potential to substantially improve upon screening approaches that are currently used in low-income, high-HIV/TB–burden settings.

### Supplementary Data

Supplementary materials are available at *Clinical 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

A. C., C. Y., F. C. S., and M. R. K. designed the study. F. C. S. and A. A. managed patient enrollment and specimen collection. M. S. and L. M. processed the patient samples and generated the data. J. V. R. performed the analysis. J. V. R., F. C. S., T. M., and A. C. wrote the manuscript. All authors critically reviewed and revised the manuscript. All authors approved the final version of the manuscript.

***Acknowledgments.*** The authors thank both the clinic staff and patients in Uganda, without whom this study would not have been possible.

***Disclaimer.*** The funding organizations had no role in the design, collection, analysis, or interpretation of data, or in the writing of the manuscript.

***Financial support.*** This study was supported by the National Institutes of Health (NIH)/National Institute of Allergy and Infectious Diseases (NIAID; grant number K23 AI114363 to C. Y.); the University of California, San Francisco (UCSF), Gladstone Institute of Virology and Immunology (GIVI) Center for AIDS Research (CFAR; NIH/NIAID pilot award P30 AI027763 to C. Y.); the UCSF Nina Ireland Program for Lung Health (to C. Y.); and the NIH/NIAID President's Emergency Plan for AIDS Relief Administrative Supplement award for UCSF-GIVI CFAR (grant number P30 A120163 to A. C.).

***Potential conflicts of interest.*** J. V. R. and A. C. have a patent pending for GeneScreen-TB, a 5-transcript gene-expression signature for tuberculosis screening in people living with human immunodeficiency virus. All other authors report no potential 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. 2030 agenda for sustainable development. Available at: http://www.un.org/sustainabledevelopment/health/. Accessed 14 January 2017.
2. Global strategy and targets for tuberculosis prevention, care and control after 2015. Geneva, Switzerland: World Health Organization, **2013**.
3. Systematic screening for active tuberculosis: Principles and recommendations. Geneva, Switzerland: Organization WH, **2013**.
4. High-priority target product profiles for new tuberculosis diagnostics. Geneva, Switzerland: Organization WH, **2014**.
5. Getahun H, Kittikraisak W, Heilig CM, et al. Development of a standardized screening rule for tuberculosis in people living with HIV in resource-constrained settings: individual participant data meta-analysis of observational studies. PLoS Med **2011**; 8:e1000391.
6. Systematic screening for active tuberculosis: an operational guide. Geneva, Switzerland: Organization WH, **2014**.
7. Berry MP, Graham CM, McNab FW, et al. An interferon-inducible neutrophil-driven blood transcriptional signature in human tuberculosis. Nature **2010**; 466:973–7.
8. 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.
9. Bloom CI, Graham CM, Berry MP, et al. Transcriptional blood signatures distinguish pulmonary tuberculosis, pulmonary sarcoidosis, pneumonias and lung cancers. PLoS One **2013**; 8:e70630.
10. 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.
11. Roe JK, Thomas N, Gil E, et al. Blood transcriptomic diagnosis of pulmonary and extrapulmonary tuberculosis. JCI Insight **2016**; 1:e87238.
12. Yoon C, Semitala FC, Atuhumuza E, et al. Point-of-care C-reactive protein-based tuberculosis screening for people living with HIV: a diagnostic accuracy study. Lancet Infect Dis **2017**; 17:1285–92.
13. Ferreira LR, Ferreira FM, Nakaya HI, et al. Blood gene signatures of chagas cardiomyopathy with or without ventricular dysfunction. J Infect Dis **2017**; 215:387–95.
14. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B **1995**; 57:289–300.
15. Yoon C, Davis JL, Huang L, et al. Point-of-care C-reactive protein testing to facilitate implementation of isoniazid preventive therapy for people living with HIV. J Acquir Immune Defic Syndr **2014**; 65:551–6.
16. Morisaki H, Akutsu K, Ogino H, et al. Mutation of ACTA2 gene as an important cause of familial and nonfamilial nonsyndromatic thoracic aortic aneurysm and/or dissection (TAAD). Hum Mutat **2009**; 30:1406–11.
17. Olszewski MA, Gray J, Vestal DJ. In silico genomic analysis of the human and murine guanylate-binding protein (GBP) gene clusters. J Interferon Cytokine Res **2006**; 26:328–52.
18. Kim BH, Shenoy AR, Kumar P, Das R, Tiwari S, MacMicking JD. A family of IFN-γ-inducible 65-kD GTPases protects against bacterial infection. Science **2011**; 332:717–21.
19. Kanie T, Abbott KL, Mooney NA, Plowey ED, Demeter J, Jackson PK. The CEP19-RABL2 GTPase complex binds IFT-B to initiate intraflagellar transport at the ciliary base. Dev Cell **2017**; 42:22–36.e12.
20. Dalman MR, Deeter A, Nimishakavi G, Duan ZH. Fold change and p-value cutoffs significantly alter microarray interpretations. BMC Bioinformatics **2012**; 13(Suppl 2):S11.
21. Irizarry RA, Warren D, Spencer F, et al. Multiple-laboratory comparison of microarray platforms. Nat Methods **2005**; 2:345–50.
22. Larkin JE, Frank BC, Gavras H, Sultana R, Quackenbush J. Independence and reproducibility across microarray platforms. Nat Methods **2005**; 2:337–44.
23. Wang Y, Barbacioru C, Hyland F, et al. Large scale real-time PCR validation on gene expression measurements from two commercial long-oligonucleotide microarrays. BMC Genomics **2006**; 7:59.
24. 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.