Polymorphisms in Interferon Pathway Genes and Risk of Mycobacterium tuberculosis Infection in Contacts of Tuberculosis Cases in Brazil


Pii: S1201-9712(19)30486-2
DOI: https://doi.org/10.1016/j.ijid.2019.12.013
Reference: IJID 3866
To appear in: International Journal of Infectious Diseases

Received Date: 27 October 2019
Revised Date: 6 December 2019
Accepted Date: 9 December 2019

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2019 Published by Elsevier.
Polymorphisms in Interferon Pathway Genes and Risk of *Mycobacterium tuberculosis* Infection in Contacts of Tuberculosis Cases in Brazil

**Running Head:** IFN Pathway SNPs and TB risk.

Juan Manuel Cubillos-Angulo¹,²,³a, María B. Arriaga¹,²,³a, Mayla G. M. Melo⁴,⁵, Elisangela C. Silva⁴,⁶, Lucia Elena Alvarado-Arnez⁷,⁸, Alexandre S. de Almeida⁵, Milton O. Moraes⁷, Adriana S. R. Moreira⁵, Jose R. Lapa e Silva⁵, Kiyoshi F. Fukutani¹,³, Timothy R. Sterling⁹, Thomas R. Hawn¹⁰b, Afrânio L. Kritski⁵b, Martha M. Oliveira¹⁰b, Bruno B. Andrade¹,²,³,⁹,¹²,¹³,¹⁴b.

¹Instituto Gonçalo Moniz, Fundação Oswaldo Cruz, Salvador, Bahia, Brazil
²Faculdade de Medicina, Universidade Federal da Bahia, Salvador, Bahia, Brazil
³Multinational Organization Network Sponsoring Translational and Epidemiological Research (MONSTER) Initiative, Fundação José Silveira, Salvador, Bahia, Brazil
⁴Laboratório de Micobacteriologia Molecular, Centro de Pesquisas em Doenças Infecciosas e Parasitárias- CEPEDIP – Universidade Federal do Rio de Janeiro, Rio de janeiro, Brazil.
⁵Programa Acadêmico de Tuberculose, Faculdade de Medicina e Complexo Hospitalar HUCFF-IDT, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil
⁶Laboratório de Biologia do Reconhecer, Centro de Biociências e Biotecnologia, Universidade Federal do Norte Fluminense Darcy Ribeiro, Rio de Janeiro, Brazil.
⁷Laboratório de Hanseníase, Instituto Oswaldo Cruz, Fundação Oswaldo Cruz (Fiocruz), Rio de Janeiro, Brazil
⁸Coordinación Nacional de Investigación, Universidad Franz Tamayo (UNIFRANZ), La Paz, Bolivia
⁹Division of Infectious Diseases, Department of Medicine, Vanderbilt University School of Medicine, Nashville, Tennessee, USA
¹⁰Department of Medicine, University of Washington, Seattle, WA, USA.
¹¹Centro de Desenvolvimento Tecnológico em Saúde, Fundação Oswaldo Cruz, Rio de Janeiro, Brazil
¹²Wellcome Centre for Infectious Disease Research in Africa, Institute of Infectious Disease and Molecular Medicine, University of Cape Town, Cape Town, South Africa
¹³Universidade Salvador (UNIFACS), Laureate Universities, Salvador, Bahia, Brazil
¹⁴Escola Bahiana de Medicina e Saúde Pública, Salvador, Bahia, Brazil

ᵃJMC-A and MBA equally contributed to the work.
ᵇTRH, ALK, MMO and BBA equally contributed to the work.

**Corresponding author:** Bruno B. Andrade, Laboratório de Inflamação e Biomarcadores, Instituto Gonçalo Moniz, Fundação Oswaldo Cruz, Rua Waldemar Falcão, no. 121, Candeal, Salvador, Bahia 40269-710, Brazil (bruno.andrade@fiocruz.br).
Highlights
- High frequency of TST positivity in TB contacts from Brazil was observed
- SNPs in type-I IFN pathway, DNA and RNA sensing genes were screened in TB contacts
- SNPs in PYHIN1-IFI16-AIM2 and in IRF7 were linked to altered susceptibility to TB

ABSTRACT

Background: Host genetic polymorphisms may be important in determining susceptibility to *Mycobacterium tuberculosis* (Mt) infection, but their role is not fully understood. Detection of microbial DNA and activation of type I interferon (IFN) pathways regulate macrophage responses to Mt infection.

Methods: We examined whether seven candidate gene SNPs were associated with tuberculin skin test (TST) positivity in close contacts of microbiologically confirmed pulmonary TB patients in Brazil. Independent associations with TST positivity were tested using multivariable logistic regression (using genotypes and clinical variables) and genetic models.

Results: Among 482 contacts of 145 TB index cases, 296 contacts were TST positive. Multivariable regression analysis adjusted for population admixture, age, family relatedness, sex and clinical variables related to increased TB risk demonstrated that SNPs in *PYHIN1-IFI16-AIM2* rs1101998 (adjusted OR [aOR]: 3.72; 95%CI = 1.15–12.0; p=0.028) and in *PYHIN1-IFI16-AIM2* rs1633256 (aOR= 24.84; 95%CI = 2.26–272.95; p=0.009) were associated with TST positivity in a recessive model. Furthermore, an *IRF7* polymorphism (rs11246213) was associated with reduced odds of TST positivity in a dominant model (aOR: 0.50, 95%CI: 0.26-0.93; p=0.029).

Conclusions: Polymorphisms in *PYHIN1-IFI16-AIM2* rs1633256, rs1101998 and in *IRF7* rs11246213 were associated with altered susceptibility to Mt infection in this Brazilian cohort.

Keywords: single nucleotide polymorphism, tuberculin skin test, *Mycobacterium tuberculosis*. 

Introduction
Tuberculosis (TB) is the leading cause of death from a single infectious agent (WHO, 2018). Approximately one-quarter of the global population is infected with *Mycobacterium tuberculosis* (Mtb) (Houben and Dodd, 2016). Latent tuberculosis infection (LTBI) is defined by immunological sensitization to Mtb antigens in the absence of clinical symptoms of disease and the diagnosis is based on the tuberculin skin test (TST) and/or Interferon-γ (IFN-γ) release assay (IGRA) (Robertson et al., 2012). Nevertheless, these tests do not discriminate between active disease and LTBI and, more importantly, have a low predictive value for progression to active TB (Rangaka et al., 2012). Many risk factors for developing active TB have been described, including HIV co-infection, diabetes, young age and recently-acquired Mtb infection (Reid et al., 2019), but intriguingly some TB patients do not exhibit any known risk factors (Yan et al., 2015). TB occurs as the result of an intricate and dynamic relationship involving host genetics (van Tong et al., 2017) as well as immunological (Mahan et al., 2012, Tameris et al., 2013), and epidemiological (Shin et al., 2016) factors, in addition to characteristics of the Mtb strain itself (Koch and Mizrahi, 2018), that contribute to disease susceptibility (Pai et al., 2016).

Genetic factors are important for TB susceptibility, but the major genes involved remain unknown (van Tong et al., 2017). Candidate gene/pathway studies interrogate selected pathways that are important in the human host response to mycobacterial infection (Kinnear et al., 2017). Type I IFN pathways mediate an important role in TB pathogenesis. Whole blood RNA signatures dominated by Type I IFN-signaling identify individuals who will develop active disease (Berry et al., 2010). In Mtb-infected mice, increased expression of type I IFNs is deleterious for
survival in association with reduced Th1 immunity (Manca et al., 2005). The Type I 
IFN pathway is activated by DNA (e.g. *IFI16-PYHIN1-AIM2*, cGAS, STING) and RNA 
sensors (e.g. *IFIT1* and 5), and contains several important signaling molecules and 
transcription factors (e.g. *IRF* family). For example, the cytosolic DNA sensor cGAS 
regulates IFN production during Mtb infection of macrophages (Watson et al., 2015). 
Although these murine and cellular studies suggest an important role for Type I IFNs 
in TB pathogenesis, the human genetics of this pathway in the context of Mtb 
infection is poorly understood (Donovan et al., 2017).

In a longitudinal investigation examining TB contacts from Brazil, we recently found 
that polymorphisms in toll-like receptor 4 (*TLR4*) and tumor necrosis factor (*TNFA*) 
are associated with increased risk of TST conversion and development of active TB 
(Cubillos-Angulo et al., 2019). Here we investigated in this same cohort whether 
genetic variation of Type I IFN pathway genes were associated with susceptibility to 
Mtb infection by examining single nucleotide polymorphisms (SNPs) involved in 
DNA and RNA sensing: (rs1101998, rs1633256, rs866484 in *IFI16-PYHIN1-AIM2 
region*, rs59633641 and rs10887959 in *IFIT5*), rs304478 and rs304498 in *IFIT1* and 
the IFN signaling pathway (rs11246213 [*IRF7*]). The objective of this study was to 
identify potential genetic biomarkers of susceptibility to Mtb infection. We studied 
close contacts of microbiologically confirmed pulmonary TB patients to estimate 
factors associated with a positive versus negative TST.

**Methods**

**Study design**
The present study was based on analyses performed retrospectively on a cohort of contacts of pulmonary TB patients, recruited between November 1998 through March 2004. The parent study was reported previously (Cubillos-Angulo et al., 2019). The cases and controls were enrolled in the state of Rio de Janeiro, Brazil where the population is mostly white and brown (‘parda, mixed ethnic ancestries) (IBGE, 2012). Racial/ethnic background was self-reported used the definitions/approaches employed by the Brazilian government for race documentation. TB index cases were diagnosed by acid-fast bacilli (AFB) smear and/or culture, according to Brazilian Ministry of Health Guidelines ((Brasil), 2019). TB index case variables included cough, AFB sputum grade, and chest radiographs. TB contacts were defined as living in the same household or reporting contact with the TB index case for >20 hours weekly for 2 months (Cubillos-Angulo et al., 2019). In the analyses presented here, we used data from a subgroup of 482 individuals, which were selected by such criteria and included contacts with TST-positive or TST-negative results. Patients who developed active TB were excluded from the analysis. Additional details on inclusion and exclusion criteria as well as patient characteristics have been described previously (Cubillos-Angulo et al., 2019).

A standardized questionnaire was administered to obtain demographic and clinical data, including a history of risk factors for TB (e.g., HIV, diabetes, hematologic malignancies, and use of immunosuppressant drugs) and duration of contact with the index case. Consanguinity was considered if a contact was a grandparent, parent or sibling of the index case, whereas spouses or other relationships were not. At study baseline, a medical visit and chest radiograph were performed. BCG scar
was assessed and TST reading was performed 48-72 hours after administration at baseline, using 2 tuberculin units of the purified protein derivative RT 23 (Statens Serum Institute, Copenhagen, Denmark).

**TST interpretation and TB diagnosis**

A positive TST was defined as an induration larger than ≥5mm induration, according to the Brazilian Ministry of Health ((Brasil), 2019). Contacts with any TST ≥5 mm were not re-tested with TST. The Brazilian National TB Guidelines indicated that treatment of TST-positive individuals was systematically offered but implementation was not mandatory during the study period ((Brasil), 2019). For the index case, active TB was diagnosed when ≥1 specimen yielded a positive microbiologic (AFB smear or culture) result by AFB smear and/or culture in Lowenstein Jensen (LJ) medium (Cubillos-Angulo et al., 2019).

**Genotyping**

Genomic DNA was extracted from peripheral blood collected from TB contacts at study enrollment. DNA extraction and genotyping were performed using the FlexiGene kit (Qiagen, Germany). Genotypes of 8 gene polymorphisms were chosen for convenience since a RFLP assay was available: rs1101998 (IFI16-PYHIN1-AIM2), rs1633256 (IFI16-PYHIN1-AIM2), rs866484 (IFI16-PYHIN1-AIM2), rs304478 (IFIT1), rs304498 (IFIT1), rs11246213 (IRF7), rs59633641 (IFIT5) and rs10887959 (IFIT5) were detected using polymerase chain reaction restriction
fragment length polymorphism (RFLP) method (Cubillos-Angulo et al., 2019). The primer sequences are in Supplementary Table 1. The PCR products were digested by the enzymes EcoRI for rs1101998 (IFI16), AgsI for rs1633256 (IFI16), AgsI for rs866484 (IFI16), AarI for rs304478 (IFIT1), TfiI for rs304498 (IFIT1), BsaAI for rs11246213 (IRF7), Apol for rs59633641 (IFIT5) and AgsI for rs10887959 (IFIT5).

Hardy-Weinberg equilibrium was tested for each SNP. We did not find significant deviation from Hardy Weinberg equilibrium except in rs304498 (IFIT1), and thus this SNP was excluded from further analysis. Linkage disequilibrium coefficients were calculated using Package “LDheatmap” (Ji-Hyung Shin, 2006) in the stats package in R 3.5.2 and using an R² and D’ cutoff of 0.8. Haplotypes analysis were constructed in the stats package R 3.5.2 using the haplo.stats (version 1.6.0) R package (Jason P. Sinnwell; Daniel J. Schaid, 2018).

Data analysis

Categorical data were presented as proportions and continuous data as medians and interquartile ranges (IQR). For clinical characteristics, a Fisher’s exact test was used to perform 2x2 comparisons. Continuous variables were compared using the Mann-Whitney U test. For genetic analysis, a Cochrane-Armitage test for trend was used initially to examine the association of genotypes with TST positivity. SNPs were then evaluated with a Fisher’s exact test using dominant (00 vs 01/11) and recessive (00/01 vs 11) models. We also estimated significant associations between indicated SNPs and TST positivity using multivariable logistic regression adjusted for race/ethnicity, family relatedness, gender and age in both dominant and recessive models. Finally, we also performed additional investigations with dominant and
recessive models in a multivariable analysis with adjustment for age, gender, race/ethnicity, family relatedness, household contact status and characteristics of TB index case, such as cavities on chest x-ray, ≥2+ AFB sputum smear grade and positive sputum culture for Mtb. We also used the GTEx portal (https://gtexportal.org/home/) to evaluate the expression quantitative trait loci (eQTL) of the SNPs (Consortium, 2013). Furthermore, the likelihood of being a regulatory SNP was examined using the RegulomeDB dataset (http://www.regulomedb.org/snp/chr10/91150921) (Boyle et al., 2012).

Results

Characteristics of the study participants

We used a retrospective cohort study of contacts (N=482) of pulmonary TB index cases (N=145) to examine whether genetic variants of candidate genes were associated with TST positivity. Household contacts were more frequently observed in the group of individuals presenting with a positive TST result than in those with a negative TST (Table 1). Cavitary lesions as well as cough in the index TB cases were more frequent in participants who were TST positive compared to those who had negative results (p=0.04 and p=0.008, respectively). Other characteristics were similar between TST positive and TST negative individuals.

The study population was mostly female (n=321, 67%), with a high frequency of first degree relatives with the index case (n=229, 62%) (Table 1). In addition, the vast majority of participants were household contacts (n=434, 90%). There were low frequencies of HIV infection, illicit drug use, prior TB and use of immunosuppressive
drugs. Approximately 97% (n=141) of the index cases had TB confirmed by culture. TB index cases frequently reported cough for more than 4 weeks (80%) and had high bacterial loads in sputum (60% had AFB grade ≥ +2;). In addition, 100 index TB patients had cavitary lesions on chest radiograph.

**Association between polymorphisms and TST positivity**

Two of seven polymorphisms were associated with TST positivity (rs1633256 and rs59633641 with an unadjusted genotypic trend test, Table 2). *PYHIN1-IFI16-AIM2* SNPs rs1101998 allele C (p=0.01) and rs1633256 allele A (p=<0.01) were more common in TST positive participants and fit a recessive model (Table 2). *IFIT5* rs59633641 allele G (p=0.04) was more common in TST positive individuals (trend test p=0.04, Table 2). *IFIT1* rs304478 and *IFIT5* rs10887959 were also significantly associated with outcomes in recessive and dominant models, respectively.

In a multivariable model that included adjustment for race/ethnicity, family relatedness, gender, and age (Figure 1), we observed in the recessive model that *PYHIN1-IFI16-AIM2* rs1101998 (adjusted OR [aOR] =2.90; 95%CI = 1.24–6.78; p=0.014) and rs1633256 (aOR = 10.1; 95%CI = 2.20–46.28; p=0.003) were associated with an increased risk TST positivity. Moreover, in the dominant model, *IFIT5* rs10887959 (aOR = 0.49; 95%CI = 0.28–0.84; p=0.01) and *IRF7* rs11246213 (aOR = 0.60; 95% CI = 0.36–1.00; p=0.049) were also linked to a lower likelihood of positive TST.
We next used a multivariable regression analysis to adjust for household contact and characteristics of TB index case (cavities on chest x-ray, ≥2 AFB sputum smear and positive Mtb culture) as well as race/ethnicity, family relatedness, gender, and age (Figure 2). We confirmed in the recessive model that PYHIN1-IFI16-AIM2 rs1101998 (aOR = 3.72; 95%CI = 1.15–12.0; p = 0.028) and rs1633256 (aOR = 24.84; 95%CI = 2.26–272.95; p = <0.009) were independently associated with increased odds of a positive TST. In addition, in the dominant model, IRF7 rs11246213 was also independently associated with a lower likelihood of a positive TST (aOR: 0.50, 95%CI: 0.26-0.93; p = 0.029).

We next examined effects of linkage disequilibrium and SNP-SNP interactions in the PYHIN1-IFI16-AIM2 region on chromosome 1. PYHIN1-IFI16-AIM2 SNPs rs8666484, rs1101998 and rs1633256 were all in moderate to high linkage disequilibrium (Supplemental Figure 1). In a haplotype analysis of chromosome 1 adjusted for age, gender, race/ethnicity, family relatedness and household contact, the haplotypes containing allele C from rs1101998 and allele A from rs1633256 did not have a higher risk of TST positivity compared to single SNP analyses (Figure 3 compared to Figure 2).

Using an in silico approach with data from the GTEx portal tool (see Methods for details and also in (Consortium, 2013), we found that six polymorphisms (rs1101998, rs1633256, rs866484, rs304478, rs10887959 and rs11246213) were eQTLs in different tissues (Supplementary Table 2). Interestingly, three different SNPs were reported to be expressed in the spleen and/or lung, which are organs
commonly affected by TB (Figure 4). The findings indicated that the \textit{PYHIN1-IFI16-AIM2} rs1101998 genotype CC was linked to decreased expression of \textit{AIM2} in spleen (Figure 4). The \textit{PYHIN1-IFI16-AIM2} rs1633256 genotype AA was also associated with dampened expression of \textit{AIM2} in spleen tissue (Figure 4). The \textit{IFIT5} rs10887959 genotype CC was associated with lower expression of \textit{IFIT5} in spleen and lung tissues (Figure 4). Finally, using a different online tool, the RegulomeDB dataset, we observed that \textit{PYHIN1-IFI16-AIM2} rs1101998 exhibited high likelihood of being a regulatory SNP for a DNAase I hypersensitivity peak or transcription factor binding. Moreover, \textit{IFIT5} rs10887959 displayed high likelihood of being a regulatory SNP for transcription factor binding and a DNAase I hypersensitivity peak. Together, these data suggest that rs1101998, rs1633256, and rs10887959 are eQTLs.

**Discussion**

In the present study, we tested associations between SNPs from related genes in different pathways of DNA and RNA sensing and the type I IFN pathway in a large number of TB contacts. The notable finding was that \textit{PYHIN1-IFI16-AIM2} rs1633256 and rs1101998 were associated with an increased risk of TST positivity whereas \textit{IRF7} rs11246213 was associated with a lower probability of TST positivity. To our knowledge, SNPs in these genes have not previously been reported to be associated with the pathogenesis of Mtb infection in contacts.

Our results suggest that the \textit{PYHIN1-IFI16-AIM2} rs1633256 and rs1101998 polymorphisms are associated with increased susceptibility to Mtb infection (i.e., a positive TST). The two polymorphisms are in a 3-gene locus on chromosome
1q23.1; thus, it is not possible to know which specific gene is most likely to exert a functional effect related to these genetic variants. The gene encoding Interferon-γ-inducible protein 16 (IFI16) (Trapani et al., 1994) is a multifunctional and ubiquitous host protein (Trapani et al., 1992), and a member of the PyHIN (pyrin and HIN200 domain-containing) protein family that consists of four family members: PYHIN1 (alias IFIX), IFI16 (alias PYHIN2), MNDA (alias PYHIN3) and AIM2 (alias PYHIN4) (Thompson et al., 2011). During Mtb infection of macrophages, IFI16 is reported to be localized into the cytosolic compartment (Thompson et al., 2011) and Mtb DNA activates the cytosolic surveillance pathway. Mice genetically lacking IFI204 (a homolog gene of human IFI16) show reduced IFIT1 and IFN-β induction against Mtb infection (Manzanillo et al., 2012). Furthermore, mycobacterial infection of AIM2−/− (absent in melanoma 2) mice induces elevated IFN-γ and reduced IFN-γ responses, leading to higher infection burdens and more severe pathology (Yan et al., 2018). In addition, in vitro studies demonstrated that AIM2-deficient macrophages display impaired activation of the inflammasome and defective production of IL-1β and IL-18 upon Mtb infection, making such cells highly susceptible to bacterial proliferation and cell death (Saiga et al., 2012). To the best of our knowledge, there are no previously reported studies on the relationship of PYHIN1 and TB. PYHIN1 detects Herpes Simplex (HSV-1) DNA and contributes to the induction of interferon response in human fibroblasts (Diner et al., 2015). In the present study, the SNPs associated with TST positivity (rs1633256 and rs1101998) are part of a large locus; thus it is possible that at least these two SNPs could be associated with any one of the 3 genes described above (PYHIN1-IFI16-AIM2) and influence the detection of
Mtb DNA during infection. Future studies are warranted to directly elucidate the molecular mechanisms underlying these associations.

The human \textit{IRF7} gene is located on chromosome 11p15.5 and is a member of the interferon regulatory factor family of transcription factors, comprised of nine members (IRF1 to 9) (Ning et al., 2011). This family is recognized by the regulation of many facets of innate and adaptive immune responses (Tamura et al., 2008). \textit{IRF7} is the central transcription factor that induces \textit{IFNA/B} gene transcription in response to cytosolic viral DNA and RNA in host cells (McNab et al., 2015). In addition, \textit{IRF7} is produced by murine bone marrow–derived macrophage infected with Mtb (Cheng and Schorey, 2018, Leisching et al., 2017). In a recent meta-analysis, Mtb infection of THP-1 macrophages induced differential expression of \textit{IRF7} (Zhang et al., 2019). Excessive type I IFN expression has been linked to increased TB-associated immunopathology and susceptibility to severe TB (Mayer-Barber et al., 2011, Mayer-Barber et al., 2014). \textit{IRF7} SNPs have been reported to significant reduce IFNα production by plasmacytoid dendritic cells following stimulation with HIV-1 (Chang et al., 2011). The effect of \textit{IRF7} SNPs on reduced IFNα production, if present also in exposure to Mtb, could be a factor explaining the decreased susceptibility to Mtb infection reported here.

We also found that \textit{IFIT5} rs59633641 was less frequently observed in individuals with a positive TST whereas \textit{IFIT5} rs10887959 was more commonly detected in individuals with positive TST. \textit{IFIT5} (IFN-induced protein with tetratricopeptide repeats-5) is a member of an interferon-induced protein with tetratricopeptide
repeats (IFIT) family with five members (IFIT1, IFIT2, IFIT3, IFIT1B and IFIT5) localized in chromosome 10q23 (Diamond, 2014). The multivariable model with adjustment for race/ethnicity, family relatedness, gender and age demonstrated associations between the IFIT5 rs10887959 and increased chance of negative TST. It has been recently demonstrated that IFIT5 physically interacts with MAP3K7/TAK1 and IκB kinase (IKK) to activate the transcription factor NF-κB, which is a key regulator of the expression of genes involved in immune responses, inflammation, cell survival and cancers (Zheng et al., 2015). IFIT5 is one of the main genes upregulated in active TB patients (Ahmed et al., 2016). The IFN-induced proteins regulate immune response against viruses. For example, it has been recently shown that IFIT3 has a protective role in response to dengue virus infection of human lung epithelial cells (Hsu et al., 2013).

Our study has several strengths such as systematic TST testing (currently recommended as the diagnostic test for LTBI in most resource-restrained countries) and inclusion criteria that ensures microbiological confirmation of TB index cases. However, it is also important to highlight potential limitations of our investigation, such as the cross-sectional nature of the analyses, which are not able to establish causal relationships. We have not performed functional validation of the findings; however, we showed an analysis of gene expression data in silico. In addition, we considered a common Mtb strain to be responsible for infections within a household, but it is possible that that may not have always been true. In addition, LTBI was only measured by TST with no IGRA assessments. These two tests are not perfectly concordant, so the TST negative group could probably include some individuals with
positive IGRA results. Of note, IGRA was not available in Brazil at the time of the patient enrollment. Food and Drug Administration (FDA) approved IGRA in 2001, and this test was introduced in Brazil in 2014, 10 years after the data collection of the present study was finalized. Regardless, our results clearly indicate associations between polymorphisms in innate immune genes linked to interferon responses and odds of Mtb infection assessed by TST positivity. Further translational studies are required to delineate the molecular events behind these associations.

NOTES

Ethics Statement: The study was approved by the Clementino Fraga Filho University Hospital (HUCFF), Federal University of Rio de Janeiro Ethics Review Board. Written informed consent was obtained from all participants or their legally responsible guardians, and all clinical investigations were conducted according to the principles expressed in the Declaration of Helsinki. The anonymity of study subjects was preserved with a code created with a link to personal identifiers.

Contributions: Study design: BBA, ALK, MMO, JRLS. Data collection: ECS, LEAA, ASdA, MOM, ASRM. Data analysis: MGMM, ECS, JMCA, MBA, KFF, TRS, TRH, MMO, BBA. Writing: JMCA, MBA, TRS, TRH, BBA.

Acknowledgments: The authors acknowledge study participants and also the staff of the Clementino Fraga Filho University Hospital of the Federal University of Rio de Janeiro. This work was supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) / Instituto Nacional de Ciência e Tecnologia (INCT,
grant number: 421703/2017-2) and Fundação de Amparo à Pesquisa do Rio de Janeiro (FAPERJ, grant number: E-26/110.974/2011). BBA, JRLS, and AK are senior investigators from CNPq and AK and JRLS receive senior fellowships from FAPERJ. The work from BBA and KFF was supported by intramural research program from FIOCRUZ and from the National Institutes of Health (U01AI115940). JMC-A was supported by the Organization of American States - Partnerships Program for Education and Training (OAS-PAEC) and his study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brazil (CAPES) - Finance Code 001. MBA receives a fellowship from the Fundação de Amparo à Pesquisa da Bahia (FAPESB). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Potential conflicts of interest:** The authors declare that they have no conflicts of interest.
References


Donovan ML, Schultz TE, Duke TJ, Blumenthal A. Type I Interferons in the Pathogenesis of Tuberculosis: Molecular Drivers and Immunological Consequences. Front Immunol 2017;8:1633.


Figures:

Figure 1. Multivariable model of association between genetic variants and TST positivity

<table>
<thead>
<tr>
<th>SNP</th>
<th>Model</th>
<th>OR (95%CI)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PYHIN1-IFI16-AIM2 (T/C) rs1101998</td>
<td>dominant</td>
<td>1.08 (0.61-1.91)</td>
<td>0.784</td>
</tr>
<tr>
<td></td>
<td>recessive</td>
<td>2.90 (1.24-6.78)</td>
<td>0.014</td>
</tr>
<tr>
<td>PYHIN1-IFI16-AIM2 (G/A) rs1633256</td>
<td>dominant</td>
<td>0.90 (0.47-1.71)</td>
<td>0.738</td>
</tr>
<tr>
<td></td>
<td>recessive</td>
<td>10.10 (2.2-46.28)</td>
<td>0.003</td>
</tr>
<tr>
<td>PYHIN1-IFI16-AIM2 (C/G) rs666484</td>
<td>dominant</td>
<td>0.88 (0.52-1.47)</td>
<td>0.614</td>
</tr>
<tr>
<td></td>
<td>recessive</td>
<td>1.37 (0.60-3.14)</td>
<td>0.455</td>
</tr>
<tr>
<td>IFIT1 (T/G) rs304478</td>
<td>dominant</td>
<td>1.16 (0.69-1.94)</td>
<td>0.576</td>
</tr>
<tr>
<td></td>
<td>recessive</td>
<td>0.53 (0.24-1.18)</td>
<td>0.120</td>
</tr>
<tr>
<td>IFIT5 (C/T) rs10887959</td>
<td>dominant</td>
<td>0.49 (0.28-0.84)</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td>recessive</td>
<td>1.00 (0.35-2.87)</td>
<td>0.996</td>
</tr>
<tr>
<td>IRF7 (A/G) rs11246213</td>
<td>dominant</td>
<td>0.60 (0.36-1.00)</td>
<td>0.049</td>
</tr>
<tr>
<td></td>
<td>recessive</td>
<td>1.39 (0.68-2.85)</td>
<td>0.367</td>
</tr>
</tbody>
</table>

Analysis in all study participants. Data represent no. SNP: single-nucleotide polymorphism; OR: odds ratio, 95% CI: confidence interval; P-value represents comparison of genotype frequencies in a dominant and recessive model with adjustment for race/ethnicity, family relatedness, gender, and age. OR (Odds ratio) represents association of minor allele with risk of TST positivity.

Figure 2. Multivariable model of association between genetic variants and TST positivity including clinical variables.
Analysis in all study participants. Data represent no. SNP: single-nucleotide polymorphism; OR: odds ratio, 95% CI: confidence interval; P-value represents comparison of genotype frequencies in a dominant and recessive model with adjustment for age, gender, race/ethnicity, family relatedness, household contact and characteristics of TB index case: Cavities on chest x-ray, ≥2 AFB sputum smear and positive Mtb culture.

**Figure 3. Haplotype analysis chromosome 1.**

Haplotype analysis chromosome 1 of SNPs rs1101998, rs1633256, rs866484-PYHIN1-IFI16-AIM2. P-value represents comparison of haplotype frequencies with...
TST conversion in an unadjusted and adjusted model for age, gender, race/ethnicity, family relatedness and household contact.

**Figure 4. In silico** expression of SNPs rs1101998, rs1633256 and rs10887959 adapted from GTEx eQTL database.

Normalized expression values were obtained from the GTEx eQTL database and violin plots (with median and interquartile range values) were used to represent the trends in data variation between the different SNPs. The full list of the SNPs and tissues evaluated is described in the Supplementary Table 2. The Figures describe the SNPs that had publicly available data on expression in spleen and/or lungs, due to its importance in TB pathogenesis. Thus, data on the SNPs rs1101998, rs1633256 and rs10887959 are shown. Allele frequency was determined as the following: 00, homozygous common allele; 01, heterozygous allele; 11, homozygous rare allele. A summary of the results of the analysis from the present study testing the association between each indicated allele and a positive TST result is shown at the bottom of the graphs. A star denotes statistically significant associations with TST positivity in the following conditions: (i) univariate analysis: a comparison of genotype frequencies without adjustment for any covariates; (ii) Adjusted model 1: analysis in a dominant and recessive model with adjustment for race/ethnicity, family relatedness, sex, and age; and (iii) Adjusted model 2: analysis in a dominant and recessive model with adjustment for age, sex, race/ethnicity, family relatedness, household contact and characteristics of TB index case (cavity on chest x-ray, ≥2 AFB sputum smear and positive Mtb culture). SNP: single-nucleotide polymorphism.
Table 1. Clinical characteristics of the study participants & association with tuberculin skin test (TST) positivity.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>n</th>
<th>TST negative</th>
<th>TST positive</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>n=219</td>
<td>n=263</td>
<td></td>
</tr>
<tr>
<td>Age -median (IQR)</td>
<td>482</td>
<td>34 (23-50)</td>
<td>37 (24-49)</td>
<td>0.40</td>
</tr>
<tr>
<td>Male sex</td>
<td>482</td>
<td>74 (34)</td>
<td>87 (33)</td>
<td>0.92</td>
</tr>
<tr>
<td>First-degree relative of the index case</td>
<td>482</td>
<td>138 (63)</td>
<td>161 (61)</td>
<td>0.71</td>
</tr>
<tr>
<td>HIV infection</td>
<td>21</td>
<td>2 (67)</td>
<td>1 (6)</td>
<td>0.04</td>
</tr>
<tr>
<td>Race (% white)</td>
<td>462</td>
<td>113 (53)</td>
<td>144 (58)</td>
<td>0.40</td>
</tr>
<tr>
<td>Illicit drug usea</td>
<td>412</td>
<td>3 (2)</td>
<td>3 (1)</td>
<td>1.00</td>
</tr>
<tr>
<td>Prior tuberculosis</td>
<td>411</td>
<td>0 (0)</td>
<td>3 (1.4)</td>
<td>0.25</td>
</tr>
<tr>
<td>Household contact</td>
<td>480</td>
<td>190 (88)</td>
<td>244 (93)</td>
<td>0.06</td>
</tr>
<tr>
<td>Duration of contact (&gt;20 hours)</td>
<td>482</td>
<td>202 (92)</td>
<td>248 (94)</td>
<td>0.46</td>
</tr>
<tr>
<td>Comorbid conditionsb</td>
<td>459</td>
<td>53 (26)</td>
<td>65 (26)</td>
<td>1.00</td>
</tr>
<tr>
<td>Immunosuppressant drugsc</td>
<td>414</td>
<td>0 (0)</td>
<td>2 (1)</td>
<td>0.50</td>
</tr>
<tr>
<td>Cough (&gt; 4 weeks)</td>
<td>481</td>
<td>5 (2)</td>
<td>5 (2)</td>
<td>0.76</td>
</tr>
</tbody>
</table>

Characteristics of TB index case

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Cavities on chest x-ray</td>
<td>473</td>
<td>24 (11)</td>
<td>47 (18)</td>
<td>0.04</td>
</tr>
<tr>
<td>Cough (&gt; 4 weeks)</td>
<td>481</td>
<td>86 (39)</td>
<td>136 (52)</td>
<td>0.008</td>
</tr>
<tr>
<td>≥2 AFB sputum smear</td>
<td>443</td>
<td>77 (39)</td>
<td>103 (42)</td>
<td>0.44</td>
</tr>
<tr>
<td>Mtb positive culture</td>
<td>339</td>
<td>147 (94)</td>
<td>178 (97)</td>
<td>0.18</td>
</tr>
</tbody>
</table>

“n” is the number of TB contacts for whom such data were available, out of a total of 482 included in the study. Data represents no. (%) or median and interquartile range (IQR) and were compared using the Fisher’s exact test (categorical variables) or the Mann-Whitney U test (for age). TST: tuberculin skin test; AFB: acid-fast bacilli on sputum smear, CI: confidence interval; OR: odds ratio; a illicit drugs: cannabis, cocaine, or crack. b co-morbid conditions: diabetes mellitus, heart failure, chronic obstructive pulmonary disease, neoplasia, systemic lupus erythematosus and hepatitis. c Immunosuppressant drugs: corticosteroids, tumor necrosis factor blockers, calcineurin inhibitors, or interleukin inhibitors.
Table 2. Association between candidate gene polymorphisms and TST.

<table>
<thead>
<tr>
<th>SNP</th>
<th>Genotype Frequency in TST negative</th>
<th>Genotype Frequency in TST positive</th>
<th>P-value</th>
<th>HW E</th>
<th>Genotypic Trend†</th>
<th>Dominate 00 vs 01/11</th>
<th>Recessive 00/01 vs 11</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>00  01  11  Total</td>
<td>00  01  11  Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs1101998 - PYHIN1- IFLAG (T/C)</td>
<td>0.43 0.4 0.0 90</td>
<td>0.43 0.2 0.0 20</td>
<td>0.2 0 0.0 0.02</td>
<td>1</td>
<td>0.0 0.04 0.009 0.009</td>
<td>0.009 0.80</td>
<td></td>
</tr>
<tr>
<td>rs1633256 - PYHIN1- IFLAG (G/A)</td>
<td>0.56 0.4 0.0 80</td>
<td>0.53 0.2 0.0 20</td>
<td>0.2 0 0.0 0.02</td>
<td>1</td>
<td>0.0 0.04 0.009 0.009</td>
<td>0.009 0.80</td>
<td></td>
</tr>
<tr>
<td>rs866484 - PYHIN1- IFLAG (C/G)</td>
<td>0.55 0.3 0.0 20</td>
<td>0.49 0.2 0.0 10</td>
<td>0.2 0 0.0 0.02</td>
<td>1</td>
<td>0.0 0.04 0.009 0.009</td>
<td>0.009 0.80</td>
<td></td>
</tr>
<tr>
<td>rs304478 - IFLAG (T/G)</td>
<td>0.37 0.2 0.0 90</td>
<td>0.37 0.2 0.0 20</td>
<td>0.2 0 0.0 0.02</td>
<td>1</td>
<td>0.0 0.04 0.009 0.009</td>
<td>0.009 0.80</td>
<td></td>
</tr>
<tr>
<td>rs59633641 - IFLAG (C/G)</td>
<td>0.97 0.0 0.0 12</td>
<td>0.9 0.1 0 10</td>
<td>0.2 0 0.0 0.02</td>
<td>1</td>
<td>0.0 0.04 0.009 0.009</td>
<td>0.009 0.80</td>
<td></td>
</tr>
<tr>
<td>rs10887959 - IFLAG (C/T)</td>
<td>0.64 0.3 0.0 60</td>
<td>0.46 0.2 0.0 10</td>
<td>0.2 0 0.0 0.02</td>
<td>1</td>
<td>0.0 0.04 0.009 0.009</td>
<td>0.009 0.80</td>
<td></td>
</tr>
<tr>
<td>rs11246213 - IFLAG (A/G)</td>
<td>0.48 0.4 0.1 20</td>
<td>0.37 0.4 0.1 10</td>
<td>0.2 0 0.0 0.02</td>
<td>1</td>
<td>0.0 0.04 0.009 0.009</td>
<td>0.009 0.80</td>
<td></td>
</tr>
</tbody>
</table>

Data represent genotype frequency of SNP TST: tuberculin skin test; SNP: single-nucleotide polymorphism; 00, homozygous common allele; 01, heterozygous allele; 11, homozygous rare allele. HWE: Hardy Weinberg equilibrium. Data was analyzed using the Fisher’s exact test (2x2 comparisons) or the chi-square trend test (3x2 comparisons). * no uncommon homozygous mutation; in this particular case, the test employed for the genotypic analysis was based on 2x2 comparison. P-value represents comparison of genotype frequencies without adjustment for any covariates. †Cochrane-Armitage trend test.