Title

History of tuberculosis disease is associated with genetic regulatory variation in Peruvians.

Journal

PLoS Genetics, 20(6)

Authors

Nieto-Caballero, Victor
Reijneveld, Josephine
Ruvalcaba, Angel
et al.

Publication Date

2024-06-01

DOI

10.1371/journal.pgen.1011313

PLOS GENETICS

RESEARCH ARTICLE

History of tuberculosis disease is associated with genetic regulatory variation in Peruvians

Victor E. Nieto-Caballero[1,2,3☯], Josephine F. Reijneveld[4☯], Angel Ruvalcaba[4], Gabriel Innocenzi[5], Nalin Abeydeera[5], Samira Asgari[3,6,7,8,9,10], Kattya Lopez[6,11], Sarah K. Iwany[6], Yang Luo[1,6,7,8,9,12], Aparna Nathan[3,6,7,8,9], Daniela Fernandez-Salinas[1], Marcos Chiñas[1,3], Chuan-Chin Huang[13], Zibiao Zhang[13], Segundo R. León[11,14], Roger I. Calderon[11], Leonid Lecca[11], Jonathan M. Budzik[5], Megan Murray[13], Ildiko Van Rhijn[6,15], Soumya Raychaudhuri[3,6,7,8,9], D. Branch Moody[6], Sara Suliman[6,4,16,17]*‡, Maria Gutierrez-Arcelus[1,3,6,7,8,9]‡*

1 Division of Immunology, Department of Pediatrics, Boston Children's Hospital, Harvard Medical School, Boston, Massachusetts, United States of America, **2** Undergraduate Program in Genomic Sciences, Center for Genomic Sciences, Universidad Nacional Autónoma de México (UNAM), Morelos, Mexico, **3** Broad Institute of MIT and Harvard, Cambridge, Massachusetts, United States of America, **4** Zuckerberg San Francisco General Hospital, Division of Experimental Medicine, University of California San Francisco, San Francisco, California, United States of America, **5** Division of Pulmonary, Critical Care, Allergy and Sleep Medicine, Department of Medicine, University of California San Francisco, San Francisco, California, United States of America, **6** Division of Rheumatology, Inflammation and Immunity, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, United States of America, **7** Division of Genetics, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, United States of America, **8** Department of Biomedical Informatics, Harvard Medical School, Boston, Massachusetts, United States of America, **9** Center for Data Sciences, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, United States of America, **10** Institute for Genomic Health, Icahn School of Medicine at Mount Sinai, New York, New York, United States of America, **11** Socios En Salud Sucursal Peru, Lima, Peru, **12** Kennedy Institute of Rheumatology, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, United Kingdom, **13** Department of Global Health and Social Medicine, Harvard Medical School, Boston, Massachusetts, United States of America, **14** Medical Technology School and Global Health Research Institute, San Juan Bautista Private University, Lima, Perú, **15** Department of Infectious Diseases and Immunology, Faculty of Veterinary Medicine, Utrecht University, Utrecht, The Netherlands, **16** Gladstone-UCSF Institute of Genomic Immunology, University of California San Francisco, San Francisco, California, United States of America, **17** Chan Zuckerberg Initiative Biohub, San Francisco, California, United States of America

☯ Victor Nieto-Caballeri and Josephine Reijneveld contributed equally to this work.
‡ Sara Suliman and Maria Gutierrez-Arcelus contributed equally to this work.
* sara.suliman@ucsf.edu (SS); mgutierr@broadinstitute.org (MGA)

OPEN ACCESS

Citation: Nieto-Caballero VE, Reijneveld JF, Ruvalcaba A, Innocenzi G, Abeydeera N, Asgari S, et al. (2024) History of tuberculosis disease is associated with genetic regulatory variation in Peruvians. PLoS Genet 20(6): e1011313. https://doi.org/10.1371/journal.pgen.1011313

Editor: Erwin Schurr, McGill University, CANADA

Received: September 13, 2023

Accepted: May 21, 2024

Published: June 13, 2024

Data Availability Statement: We created a public interactive website for data visualization of (1) gene expression levels in DCs and MPs with differential expression analysis results, (2) cell-type dependent eQTLs, (3) all eQTLs found in each cell-type available at: https://jbcmultiomics.org/tbeqtls/. Gene expression matrices, meta data matrices, and raw RNA-seq data have been deposited in the Gene Expression Omnibus (GEO): GSE269009.

Funding: This study was funded by the National Institutes of Health (NIH) TB Research Unit Network (Grant U19 AI111224-01, DBM, MM), the

Abstract

A quarter of humanity is estimated to have been exposed to *Mycobacterium tuberculosis* (*Mtb*) with a 5–10% risk of developing tuberculosis (TB) disease. Variability in responses to *Mtb* infection could be due to host or pathogen heterogeneity. Here, we focused on host genetic variation in a Peruvian population and its associations with gene regulation in monocyte-derived macrophages and dendritic cells (DCs). We recruited former household contacts of TB patients who previously progressed to TB (cases, n = 63) or did not progress to TB (controls, n = 63). Transcriptomic profiling of monocyte-derived DCs and macrophages measured the impact of genetic variants on gene expression by identifying expression quantitative trait loci (eQTL). We identified 330 and 257 eQTL genes in DCs and macrophages (False Discovery Rate (FDR) < 0.05), respectively. Four genes in DCs showed interaction between eQTL variants and TB progression status. The top eQTL interaction for a protein-

National Institute of Allergy and Infectious Diseases (R01 AI049313, DBM, and R01 AI175614, SS). SS is also supported by an investigator grant from the Chan Zuckerberg Biohub. M-GA was supported by NIH/NIAMS P30AR070253, by NIH/NIAID R01AI175614, by Gilead Sciences, by a Global Team Science Award, a Lupus Innovation Award, and a Diversity in Lupus Career Development Award from the Lupus Research Alliance, and by the Arthritis National Research Foundation Vic Braden Family Fellowship. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

coding gene was with *FAH*, the gene encoding fumarylacetoacetate hydrolase, which mediates the last step in mammalian tyrosine catabolism. *FAH* expression was associated with genetic regulatory variation in cases but not controls. Using public transcriptomic and epigenomic data of *Mtb*-infected monocyte-derived dendritic cells, we found that *Mtb* infection results in *FAH* downregulation and DNA methylation changes in the locus. Overall, this study demonstrates effects of genetic variation on gene expression levels that are dependent on history of infectious disease and highlights a candidate pathogenic mechanism through pathogen-response genes. Furthermore, our results point to tyrosine metabolism and related candidate TB progression pathways for further investigation.

Author summary

Tuberculosis (TB), caused by *Mycobacterium tuberculosis* (*Mtb*), is a leading cause of mortality globally. *Mtb*-exposed individuals have heterogeneous outcomes following infection with *Mtb*, where some progress to TB, while many remain asymptomatic. We hypothesized that genetic variation could partly explain risk of TB, by regulating the expression of genes involved in the control of *Mtb* infection. In this study in Peru, we recruited former TB patients and *Mtb*-exposed household contacts of TB patients who remained asymptomatic to define how genetic variation regulates expression of genes in the innate immune system. We identified 4 genetic variants that have different effects on gene expression based on the TB history of the participant. The lead variant regulated expression of a key enzyme (Fumarylacetoacetate Hydrolase; *FAH*) in tyrosine catabolism. Knocking out the gene in myeloid cells increased susceptibility to *Mtb* infection. The results implicate *FAH* as a candidate host factor involved in TB progression.

Introduction

Tuberculosis disease (TB), caused by infection with *Mycobacterium tuberculosis* (*Mtb*), is a leading cause of death from infection globally [1]. Notably, only 5–10% of *Mtb*-infected individuals are estimated to develop active TB, and these individuals are at higher risk of recurrent TB, suggesting the existence of durable host factors that influence disease outcome [2]. Unbiased systems biology approaches have uncovered several host pathways associated with TB progression, including interferon signaling [3,4], metabolic dysregulation [5], and depletion of Th17 cells in the peripheral blood [6]. Genetic association studies from our group [7,8], and others [9–13], identified host genetic variants associated with higher risk of active TB. However, these studies did not systematically integrate the association between genetic variation and transcriptional profiles of myeloid cells, which are the first line of defense following *Mtb* infection.

Growing evidence supports critical roles for myeloid cells in progression to TB disease in human cohorts [3,4,14–17], where macrophages are the main target of infection [18], and dendritic cells present mycobacterial antigens to prime T cells [19]. Monocytes can be differentiated *in vitro* to generate monocyte-derived dendritic cells (DCs) [20] and macrophages [21], which have become useful tools to study innate responses to pathogens. For instance, monocyte-derived DCs from TB susceptible individuals showed elevation of autophagy associated genes, including Fasciculation And Elongation Protein Zeta 2 (*FEZ2*), a repressor of autophagosome maturation [22], and group I CD1 genes, antigen presenting molecules that mark DC

maturation [13,23]. Hence, these *in vitro* models can be used to study the impact of genetic variants associated with differential gene expression that may underlie TB susceptibility or disease-induced regulatory effects [24,25].

Multiple studies mapping genetic variation that influence gene expression levels as expression quantitative trait loci (eQTL) have highlighted the importance of genetic regulatory variation in disease [8,26]. These genetic regulatory effects are often context-dependent; that is, the same genetic variant can affect gene regulation to varying degrees depending on the cell type or cell state [27–31]. For example, eQTL can be present in some human tissues but not others [26,27]. They can also vary between subsets of a cell type, such as between naïve and memory B cells [32], or between cell states of a given cell type, such as resting and activated monocytes [33]. It is becoming more evident that identifying these cell state-dependent genetic regulation events are crucial for understanding disease mechanisms [34,35]. Multiple studies have shown that *in vitro* pathogen exposure changes cell states and consequently changes the regulatory landscape, revealing pathogen-dependent genetic regulatory effects [31,36–39]. Collectively, pathogens, such as *Mtb*, could exploit the cell type-specific impacts of host genetic variation on gene expression to establish virulence or mediate long term pathogenic effects in diverse populations [40]. Hence, integration of genetic variation with transcriptional differences in myeloid cells may provide insights into disease mechanisms, as supported by prior studies in monocyte-derived DCs [22]. These genetically mediated transcriptional changes may control host immune responses in ways that change *Mtb* infection outcomes, as well as susceptibility to long term sequalae of infection [24].

In this study, we re-enrolled patients in Peru that were rigorously defined as index TB patients and their *Mtb*-infected household-contacts who either progressed or did not progress to TB disease [6–8,41,42]. We profiled gene expression in monocyte-derived DCs and macrophages from these previously genotyped participants [7]. We identified genetic variants that differentially regulated transcription in DCs and macrophages in progressors compared to non-progressors. The top regulatory event was mediated by a variant on chromosome 15 that specifically regulated expression of fumarylacetoacetate hydrolase (*FAH*) in DCs from progressors, but not non-progressors. This cell type-specific regulation of *FAH* in recovered TB patients provides a biologically plausible metabolic determinant of TB disease outcome.

Methods

Ethics statement

The Institutional Review Board of the Harvard Faculty of Medicine and Partners Healthcare (protocol number IRB16-1173), and the Institutional Committee of Ethics in Research of the Peruvian Institutes of Health approved this study protocol. Written informed consent was provided by study participants.

Human participants

Participants were nested and re-recruited from the original prospective cohort of index TB patients and their household contacts [42] (**Fig 1A**). Briefly, household contacts in the original study were screened for symptoms of TB disease. All household contacts received an intradermal tuberculin skin test at enrollment, which was read 48–72 hours later. Cases in the current study were defined to include both former HIV-negative index TB patients, with pulmonary TB and microbiologically confirmed *Mtb* culture (primary cases), as well as household contacts of index TB patients who progressed to active tuberculosis disease within 12 months of contact (secondary cases). All cases received antibiotic treatment according to the standard of care per Peruvian national guidelines.

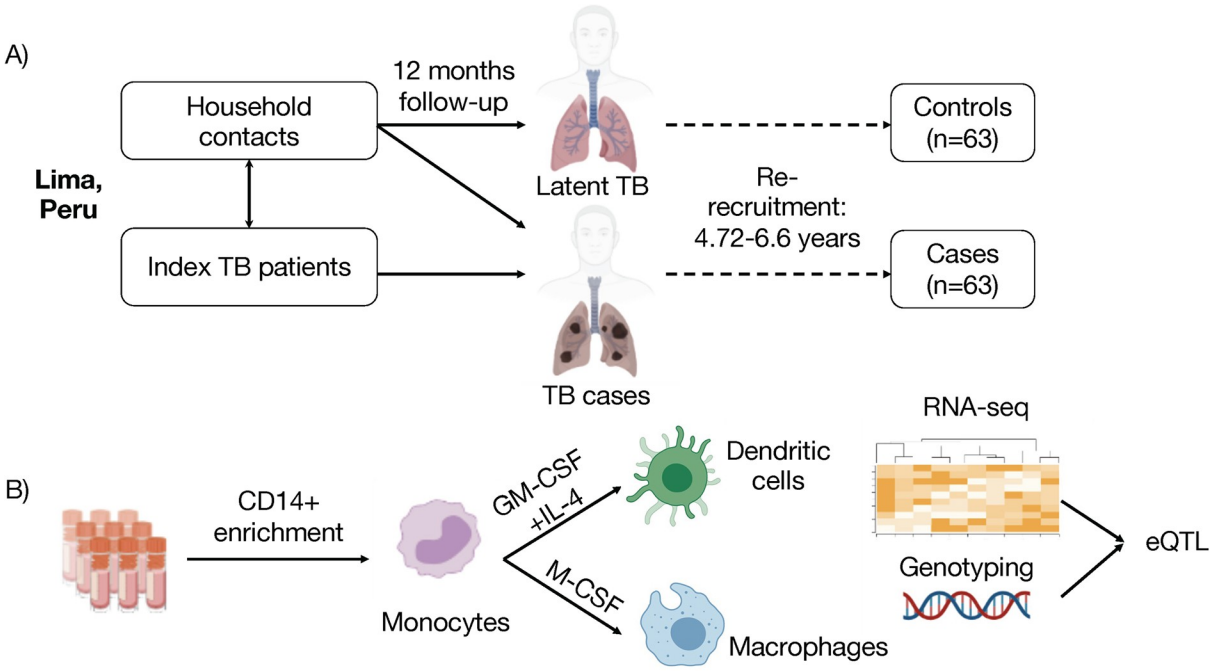

Fig 1. Myeloid cell transcriptomic study in a Peruvian cohort reveals genetic variants associated with *FAH* transcripts in dendritic cells for subjects with history of TB disease but not for household contacts exposed to *Mtb* that did not progress to active disease. (A) Cohort scheme: human participants included former household contacts of TB patients who either progressed (cases, progressors) or did not progress to active TB disease but have converted their tuberculin skin test within 12 months of exposure (non-progressors, controls). Genotyping data were previously obtained from all participants. **(B)** Transcriptomic study scheme: Monocytes were magnetically enriched from peripheral blood mononuclear cells and differentiated for 6 days into either monocyte-derived dendritic cells (GM-CSF + IL4) or macrophages (M-CSF). Samples were analyzed by low-input RNA-sequencing, and genotyping data were integrated with transcriptional profiles in an eQTL study. Icons were generated in Biorender.

https://doi.org/10.1371/journal.pgen.1011313.g001

Controls were defined as former household contacts of TB patients who had a positive tuberculin skin test result during 12 months of follow up. Importantly, this group did not develop active TB disease between the initial recruitment into the original longitudinal prospective study [7,42] and re-recruitment into the current nested case-control study. Genotyping data from all individuals in this study have been previously generated for a genome-wide association study [7] and integrated with transcriptomic data as outlined below.

Samples

Peripheral blood mononuclear cell (PBMC) samples were separated using standard Ficoll density gradient centrifugation from 50mL of venous blood and cryopreserved at $5X10^6$ cells/vial. Samples were shipped to the Brigham and Women's Hospital for storage in liquid nitrogen and subsequent experiments.

Differentiation of monocyte-derived dendritic cells and macrophages *in vitro*

PBMC samples were thawed, and CD14-expressing monocytes were magnetically sorted using the pan monocyte isolation kit, humans (Miltenyi) following the manufacturer's instructions in magnetic-associated cell sorting (MACS) buffer: 0.5% bovine serum albumin, 2 mM ethylene-diamine-tetra-acetic acid (EDTA) in 1X phosphate-buffered saline (PBS). To generate either monocyte-derived dendritic cells or macrophages, monocytes were washed and

resuspended in supplemented complete RPMI (5% fetal calf serum, 1 mM 2-mercaptoethanol, penicillin-streptomycin, L-Glutamine, HEPES buffer, both essential and non-essential amino acids, and sodium hydroxide). The media was either supplemented with 300 Units/mL of granulocyte-macrophage colony stimulating factor (GM-CSF) and 200 Units/mL of interleukin (IL)-4 to generate monocyte-derived dendritic cells (DCs), or 20 ng/mL of macrophage colony-stimulating factor (M-CSF) to generate monocyte-derived macrophages. For both cell types, monocytes were differentiated in 6-well tissue culture plates for 6 days, then split between analytical flow cytometry to test viability and differentiation, and cell sorting to generate pure populations for RNA extraction.

Cell sorting and flow cytometry analysis

Magnetically isolated CD14 cells treated with cytokines for 6 days were sorted by gating on forward and side scatter profiles after differentiation to remove subcellular debris. Cells were sorted directly into the TCL RNA lysis buffer (Qiagen). Only samples with at least 1000 cells were included in the final RNA-sequencing analysis. In parallel, to confirm viability and differentiation, cultured cells were stained first with a fixable blue viability cell stain (ThermoFisher Scientific) according to manufacturer's instructions, followed by different cocktails of fluorochrome-conjugated antibody cocktails for each cell type (**S1 Table**), and analyzed by analytical flow cytometry on a 5-laser Fortessa (BD Biosciences) to confirm viability and differentiation. Flow cytometry panels were optimized by titrating individual antibodies to find optimal signal-to-noise ratios, and minimal spillover from spectral overlap of emissions using fluorescence minus-one. Flow cytometry (FCS3) files were analyzed using FlowJo v9.9.

RNA sequencing

RNA-seq library preparation was performed at the Broad Technology Labs (Broad Institute) using a modified Smart-seq2 [43] protocol for low-input RNA-sequencing, which improves detection of full-length transcripts using template switching and pre-amplification. There were 3 plates of 96 samples each (**S2 Table**) including blank samples. Samples from cases and controls and the two cell-types were randomized across all three plates. Sequencing was performed in an Illumina NextSeq500 as paired-end 2 x 25 bp reads with an additional 8 cycles for each index.

Analysis of RNA-sequencing data

Read alignment was performed using STAR (v2.4.2) [44] with the following parameters:—twopassMode Basic,—alignIntronMax 1000000,—alignMatesGapMax 1000000,—sjdbScore 2,—quantMode TranscriptomeSAM,—sjdbOverhang 24. RSEM (v1.2.21) [45] was then used for gene quantification with paired-end mode. We used the hg38 reference genome (University of California Santa Cruz Genome Browser) and GENCODE annotation version 24 (Ensembl 83). Quality summary statistics were gathered using Picard tools. Data quality was evaluated using RNA-SeQC [46]. RNA-seq data analysis was performed with R (version 3.6.0) and R Bioconductor. Samples with low (<80%) proportion of common genes detected (common genes defined as those present in >75% of samples) or presenting as outliers in principal component analysis (performed on top 1000 variable genes based on standard deviation) were filtered out. Five samples that did not match their genotype information (based on proportion of both alleles seen in RNA-seq data over heterozygous sites) were filtered out. Differential expression of genes by TB status was determined using a multivariable linear regression model adjusting for plate (batch), age, and gender, using the R limma package [47] and P-values were adjusted using the Benjamini-Hochberg false discovery rate method.

Genotyping data

We used genotyping data from the customized Affymetrix LIMAArray genome-wide genotyping chip as described [7], where QC and imputation was performed. Briefly, individuals were filtered out if they presented higher than 5% missing genotyping data, or had excess of heterozygous genotypes, or had >40 years at age of diagnosis. Variants were filtered out if they fell in either of these criteria: call rate <95%, duplicated based on coordinates, presentation of batch effect, deviation from Hardy-Weinberg equilibrium (HWE, $P < 10^{-5}$), large missing rate differences between cases and controls. After these filters, the 677,232 SNPs left were used for imputation with SHAPEIT2 [48] and IMPUTE2 [49]. After requiring imputation score r2 > = 0.4 and re-filtering based on HWE and missing rate, there were 7,756,401 genetic variants. For eQTL analyses, we removed SNPs with minor allele frequency (MAF) <0.09 to ensure presence of at least 10 minor alleles in tested genes, and avoid problems due to outliers, especially for the interaction analyses. We reported all genomic positions using Grch37 coordinates.

Expression quantitative trait loci (eQTL) analysis

In total, 126 individuals (with 118 and 109 RNA-seq samples, corresponding to DCs and macrophages respectively) had high data quality for both genotype and gene expression for cis-eQTL mapping. We performed a separate analysis for each cell type. Genes expressed at low levels where $\log_2$(transcripts per million (TPM)+1) <1 in more than 30 samples, were filtered out in each cell subset. We restricted the analysis to SNPs that were within 1Mb of the gene start coordinate. QTLtools [50] was run in the permutation pass mode (1000 permutations for top SNP per gene) to identify associations between genetic variants and quantile normalized gene expression levels expressed as $\log_2$(TPM+1) using linear regression. To control for covariates, we included 5 genotyping principal components (PCs), 15 expression PCs, age, and sex.

To screen for potential interactions of genetic variants with TB status, we followed a two-step nested approach that was recently shown to improve the discovery of gene-by-environment interactions [51]. First, we selected SNP-gene pairs with a suggestive eQTL (p<0.001), where only top SNP per gene was selected. Next, we applied Levene's test (LT) to prioritize SNPs that were significantly associated with the variance in gene expression at 5% FDR using the qvalue R package [52]. Finally, we used these variance quantitative trait loci (vQTLs) to test for an interaction between the SNP and TB status on gene expression. Specifically, we tested for the SNP-TB status interaction effects by performing a likelihood ratio test between two nested models using the R anova function to conduct an analysis of variance (ANOVA) test. The null model estimates the effects of the SNP, TB status, and covariates evaluated using 5 genotyping PCs, 15 expression PCs, age, and sex) on a gene's $\log_2$(TPM+1) normalized expression. The alternative model has an additional SNP by TB status interaction term:

$$E_{i,j} = \theta + \beta_{geno} \cdot g_j + \beta_{tb} \cdot t_j + \beta_x \cdot t_j \cdot g_j + \sum_{l=1}^{15} \varphi_l \cdot pc_{i,l} + \sum_{m=1}^{5} \gamma_m \cdot pc_{j,m} + \beta_{age} \cdot a_j + \beta_{sex} \cdot s_j$$

where $E_{i,j}$ is the normalized gene expression for the ith sample form the jth subject, θ is the intercept, β_{geno} is the effect (eQTL) of the genotype for subject j (g_j), β_{tb} is the effect of the TB status (cases vs. controls) for subject j (g_j), and β_x is the effect of the TB status by genotype interaction ($t_j \cdot g_j$).

As before, we used the qvalue package to calculate the FDR across tests and called all events with q<0.1 (<10% FDR) as significant. Given our limited sample size, we sought to rigorously determine if the significant events were driven by potential outlier events, so we performed

permutations by shuffling the TB status for each individual 1000 times and re-computing the SNP by TB status interaction.

To identify cell-type specific eQTL interactions, we first identified the best SNP per gene for expression using both cell types together with QTLtools. For the covariates' matrix, we removed the first expression PC because it separated the samples according to cell type. We then followed the same two-step nested approach described above with the cell type as an interaction term to identify significant SNP-cell type interactions at 10% FDR.

Concordance between eQTLs

To verify the agreement between the mapped eQTLs in our study and the previous data, we selected the variant-gene pairs that were significant at FDR <0.05 in each cell type and compared the effect size magnitude (beta from QTLtools results) with the effect size in the other cell type. Additionally, we compared the beta of the most associated variant for each of the genes with an eQTL (FDR <0.05) in macrophages with the corresponding effect size of that variant-gene pair reported in the eQTL catalogue [53] from a previously published dataset [39].

Analysis of public RNA-seq data

We obtained gene expression levels from public RNA-seq samples of *Mtb*-infected and non-infected monocyte-derived DCs in 6 individuals [54]. We used a mixed-effects linear model to test for differential expression between uninfected and or infected DCs, along with heat-inactivated *Mtb* treated DCs, controlling for time as a fixed effect and donor as a random effect.

Analysis of public methylation data

We retrieved the CpG sites that showed significant differential methylation in dendritic cells in response to either live or heat-killed *Mtb* according to Pacis et al [54], following the methylation difference values at four time points post-infection (2, 18, 48, 72h) within a 250 kb window around the transcription start site of *FAH*.

Generating *Fah* knockout conditionally immortalized macrophage (CIM) progenitors

Bone marrow-derived progenitor Cas9-transgenic CIMs were provided by Jeffery S. Cox. Edited CIMs were generated as described by Roberts, et al [55]. In short: CRISPR guide RNA (gRNA) sequences targeting genes of interest were selected from the murine Brie guide library: Fah exon 2 anti-sense gRNA: CCGATGGCTACACCAATCCG, and non-targeting gRNA sequence GAACTCGTTAGGCCGTGAAG, as control. Oligonucleotides encoding the chosen gRNAs were cloned into lenti-gRNA hygro (Addgene plasmid #104991) and verified by Sanger sequencing using the human U6 sequencing primer. Lenti-X 293T cells were co-transfected with pLenti-Guide-Hygro containing gRNA, psPAX2, and pMD2.G using Lipofectamine and Optimem according to manufacturer's guidelines to generate lentiviral particles for transduction into Cas9-expressing CIM progenitors. For optimal transduction of Cas9-expressing CIM progenitors, $5.0x10^5$ cells/well in a 6-well plate were spinfected at 1000xg for 2hr at 32˚C in the presence of 10 mg/ml protamine sulfate. Two days post- transduction, we added 250 microgram/mL hygromycin B to CIMs and hygromycin resistant cells were maintained as polyclonal populations. Genomic DNA and RNA was extracted using the AllPrep DNA/RNA extraction kit (Qiagen, catalogue: 80204). We amplified genomic sites encompassing targeted regions by PCR using Q5 high-fidelity DNA polymerase (NEB) and sequenced, using the following

primers: Forward: TTCATGGGTCTGGGTCAAGC and Reverse: TCAGATCACAGGTGCTCAGG. The resulting PCR product was purified using the QIAquick PCR Purification Kit (Qiagen) according to manufacturer's protocol and sent for sequencing. Population-level genome editing was estimated using the Inference of CRISPR Edits (ICE) from Sanger sequencing trace data (Synthego) [56].

Differentiation of CIM progenitors into macrophages

We maintained suspension CIMs prior to differentiation in non-treated tissue culture flasks in RPMI (Gibco) supplemented with 10% FBS, 2 mM L-glutamine (Gibco), 1 mM sodium pyruvate (Gibco), 10 mM HEPES (Gibco), 43 uM β-mercaptoethanol, 2 mM β-estradiol (Sigma #E2758), and 2% GM-CSF supernatant produced by a B16 murine melanoma cell line or 20 ng/mL recombinant mouse GM-CSF (R&D Systems). For differentiation into macrophages, CIMs were washed twice in PBS + 1% FBS to fully remove β-estradiol, resuspended in complete macrophage media (DMEM, Gibco) supplemented with 10% FBS, 2 mM L-glutamine (Gibco), 1 mM sodium pyruvate (Gibco), and 10% M-CSF supernatant produced by 3T3-MCSF cells or 20 ng/mL recombinant mouse M-CSF (R&D systems), and seeded onto non-treated tissue culture plates or flasks. Fresh media was added on day 3 post-differentiation, cells were re-seeded into plates on day 7 and used for *Mtb* infection assays 8-days post-differentiation.

Mycobacterium tuberculosis growth assay

We seeded differentiated CIMs at 60,000 cells per well onto white, clear-bottom polystyrene 96-well plates (Corning) in the same macrophage media used for differentiation one day prior to infection. Macrophages were infected with *Mtb*-lux [57] as follows: H37Rv expressing the luxCDABE operon (PMV306hsp-LuxG13, Addgene #26161) was used for all infections. *Mtb* was cultured in 7H9 liquid media (BD) supplemented with 10% Middlebrook OADC (Sigma), 0.5% glycerol, 0.05% Tween-80 in roller bottles at 37˚C. Briefly, mid-log *Mtb* cultures were washed twice with PBS, gently sonicated to disperse clumps, and resuspended in phagocytosis infection media (DMEM supplemented with 10% horse serum). Macrophages were infected at a multiplicity of infection (MOI) of 1 by removing media from cells, followed by the addition of the bacterial suspensions in phagocytosis media and then spinfected at 1000 rpm for 10 min, after which the infection media was removed, and fresh macrophage media was added. Bacterial luminescence signal was measured at the time of infection (day 0) and every day starting 24 hr post-infection after daily media changes. All growth measurements are normalized to day 0 luminescence readings for each infected well and are presented as fold change in luminescence compared to day 0.

Results

Participants for this study were nested from a larger cohort of index Peruvian TB patients and their household contacts [42], from which we previously conducted a genome-wide association study [7]. We re-recruited a subset of these former household contacts who had been infected with *Mtb* and either progressed to active TB disease (progressors, cases) or remained TB-disease free until sampling (non-progressors, controls) after initial enrollment into the parent cohort [6] (**Table 1**). Due to sample availability, we restricted the analysis to high quality data for 63 progressors and 63 non-progressor controls, sampled a median of 6 years after exposure to TB in the household (**Fig 1A**). We standardized the protocol to differentiate monocyte-derived DCs and macrophages using magnetic enrichment of $CD14^+$ monocytes, and differentiation with a combination of interleukin-4 (IL4) and granulocyte macrophage-

Table 1. Cohort demographics of individuals included in the RNA-seq and eQTL analysis.

Demographics	Progressors (n = 63)	Controls (n = 63)	Overall (n = 126)
Age (Years)			
Mean (SD)	29 (13.59)	40 (16.86)	35 (16.22)
Median [Min, Max]	24 [16, 69]	40 [3, 77]	32 [3, 77]
Sex			
Female	27	37	64
Male	36	26	62

https://doi.org/10.1371/journal.pgen.1011313.t001

colony stimulating factor (GM-CSF) or macrophage-colony stimulating factor (M-CSF), respectively (**Fig 1B**).

We profiled the transcriptome of cells using low-input mRNA sequencing. Macrophages and DCs showed expression of expected genes that mark their cell type specificity (**Fig 2**). For example, genes characteristic of the DC lineage that are absent in macrophages, including group 1 CD1 genes (*CD1a*, *CD1b*, *CD1c* and *CD1e*) [20,21], were significantly up-regulated in DCs. Similarly, genes characteristic of macrophage differentiation, including V-Set And Immunoglobulin Domain Containing 1 (*VSIG*), *CD14*, Fc Gamma Receptor Ia (*FCGR1A*) and IIIa (*FCGR3A*) [21], were significantly upregulated in macrophages compared to DCs.

First, we asked whether we could detect gene expression differences directly associated with TB disease status of donors. Comparing cells from former progressors and controls, we did not identify any statistically significant differentially expressed genes (FDR < 0.05) in either cell type (**Fig 3A and 3B**). This outcome is in part expected given that RNA was extracted from *in vitro* differentiated monocytes that were collected a median of 6 years after cases had active disease.

However, we hypothesized that prior TB progression could be associated with changes in the regulatory landscape of differentiated myeloid cells, which could modulate the effects of genetic variants on gene expression. Therefore, we first identified genes whose expression levels were associated with genetic variants within 1Mb of the gene start (*cis* eQTLs), controlling for age, biological sex, library preparation batch, 15 gene expression and 5 genotyping principal components (PCs). We identified significant eQTLs for 257 genes in macrophages, and 330 genes in DCs (FDR <0.05). The top eQTL in DCs showed that the minor allele C at rs199659767 was associated with a significant change in the expression of Zinc Finger Protein 57 *ZPF57* ($p = 1.8 \text{ X } 10^{-31}$) (**S3 Table**). The most significant eQTL ($p = 3.4 \text{ X } 10^{-30}$) in macrophages showed a positive association between Glutathione S-Transferase Mu 1 (*GSTM1*) expression and the allele A at rs140584594 (**S4 Table**). Our results demonstrated a high level of concordance between the mapped eQTLs within our study (macrophages and DCs) and between our macrophage eQTLs and a previously published eQTL study in macrophages [39] whose summary statistics were made available by the eQTL Catalogue [53]. In all comparisons we observed that most of the significant variant-gene pairs (FDR <0.05) had a beta (eQTL effect) in a consistent direction and comparable magnitude (**Fig 4A–4C**), further supporting the strong agreement between our results and previous findings.

We then sought to identify genetic regulatory variants whose effects on gene expression were differentially associated with TB disease history. Previously, it has been observed that SNPs associated with the variance of a quantitative trait (vQTL) can increase the detection of gene by environment (GxE) interactions with large effects [51]. We followed this statistical framework to reduce the number of eQTLs tested for interaction. First, for genes with suggestive eQTLs ($p < 0.001$), we identified 105 vQTL genes for DCs and 96 for macrophages (FDR < 0.05). For each of these vQTLs, we then tested for TB status interaction (**Methods**). In the DCs, we observed five vQTLs with significant TB status-eQTL interaction effects

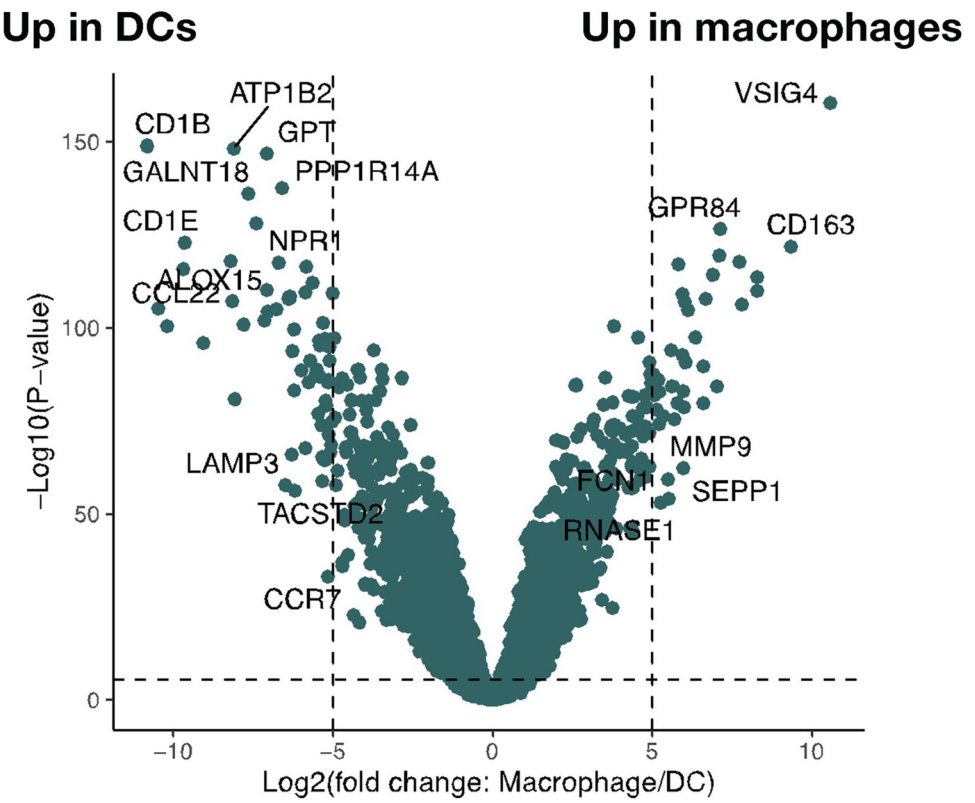

Fig 2. Volcano plot depicting differential gene expression between samples corresponding to monocyte-derived DCs and M2 macrophages using a multivariate linear regression analysis on $\log_2$(transcripts per million +1) adjusted for age, sex and batch (R limma package). Several genes were distinctly cell type-specific with more than 32-fold change and below an FDR-adjusted p-value threshold of 0.05.

https://doi.org/10.1371/journal.pgen.1011313.g002

(FDR < 0.1) (**S5A Table**). We conducted permutation analysis to rigorously confirm the validity of the eQTL interactions observed in DCs and macrophages by reassigning the TB status 1000 times across samples and retesting for TB status-eQTL interactions (**S1A Fig**). Four out of the five TB status eQTL interactions passed the test ($p_{permute} < 0.05$), further suggesting that genetic regulatory variation is associated with history of TB disease. These interactions found in DCs included two protein-coding genes: fumarylacetoacetate hydrolase (*FAH*) and N-ethylmaleimide sensitive factor (*NSF*), and two long non-coding RNA genes: *RP3-340B19.2* and *RP11-667K14.4*. We identified a single strong interaction event for macrophages at this threshold with Lysine Demethylase 6B (*KDM6B*), which did not pass the permutation p-value threshold of 0.05 (**S5B Table**, **S1B Fig**).

In DCs, we found that an indel on chromosome 15 was the top genetic variant associated with expression levels of *FAH*, a gene encoding Fumarylacetoacetate Hydrolase: the last enzyme in mammalian tyrosine catabolism [58]. The minor allelic variant of rs142312981 (AAG insertion) was associated with lower expression of *FAH* in former TB progressors but not in non-progressors (**Fig 5A**). The allele frequency in our cohort is 12%, compared to the estimated Peruvian allele frequency in the 1000 Genomes Project at 7%. Interestingly, this variant's effect on gene regulation was not observed in monocyte-derived macrophages from the same donors (**Fig 5B**), suggesting a disease and cell type-specific regulatory axis.

To validate these results, we first reanalyzed public RNA-seq and methylation dataset of *Mtb*-infected and non-infected monocyte-derived DCs in 6 individuals [54]. Stimulation with live

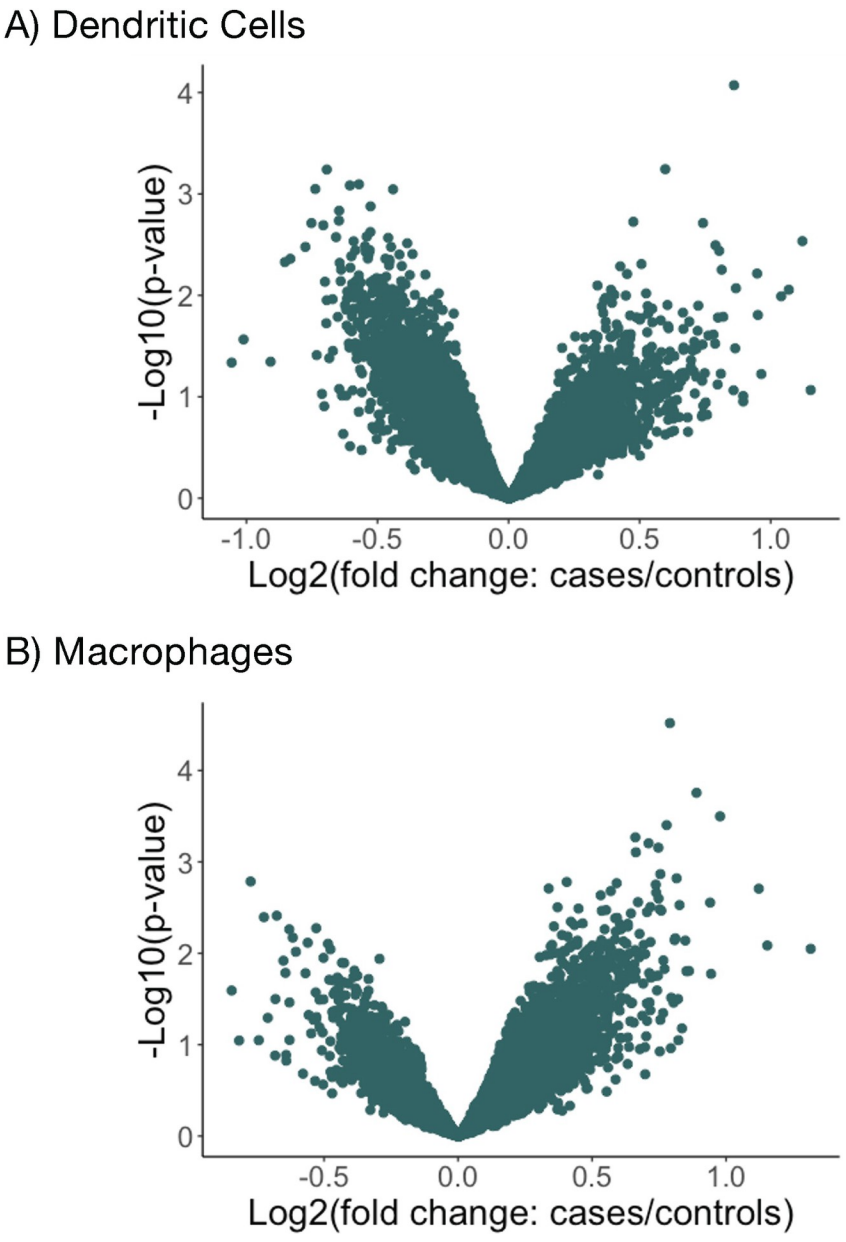

Fig 3. Differential gene expression in monocyte-derived DCs (A) or M2 macrophages (B) between former TB cases and controls, using a multivariate linear regression analysis adjusted for age, sex and batch (R limma package). There are no TB-associated differentially expressed genes in either cell type.

https://doi.org/10.1371/journal.pgen.1011313.g003

($p = 1.4X10^{-10}$) or heat inactivated ($p = 5.2X10^{-10}$) *Mtb* downregulated *FAH* expression (**Fig 5C**). In the same public dataset [54], we observed that CpG sites upstream of the *FAH* transcriptional start site and near rs142312981 were differentially methylated after *Mtb* infection (**Fig 5D**), consistently with an association between *Mtb* infection and epigenetic remodeling of the *FAH* locus.

To experimentally validate these cohort-based transcriptional results, we hypothesized that *FAH* deletion would increase susceptibility of myeloid cells to *Mtb* infection. To address this hypothesis, we leveraged conditionally immortalized macrophages (CIMs) from Cas9-transgenic mice to generate *Fah*-edited CIMs [55]. CIM progenitors ectopically express the estrogen

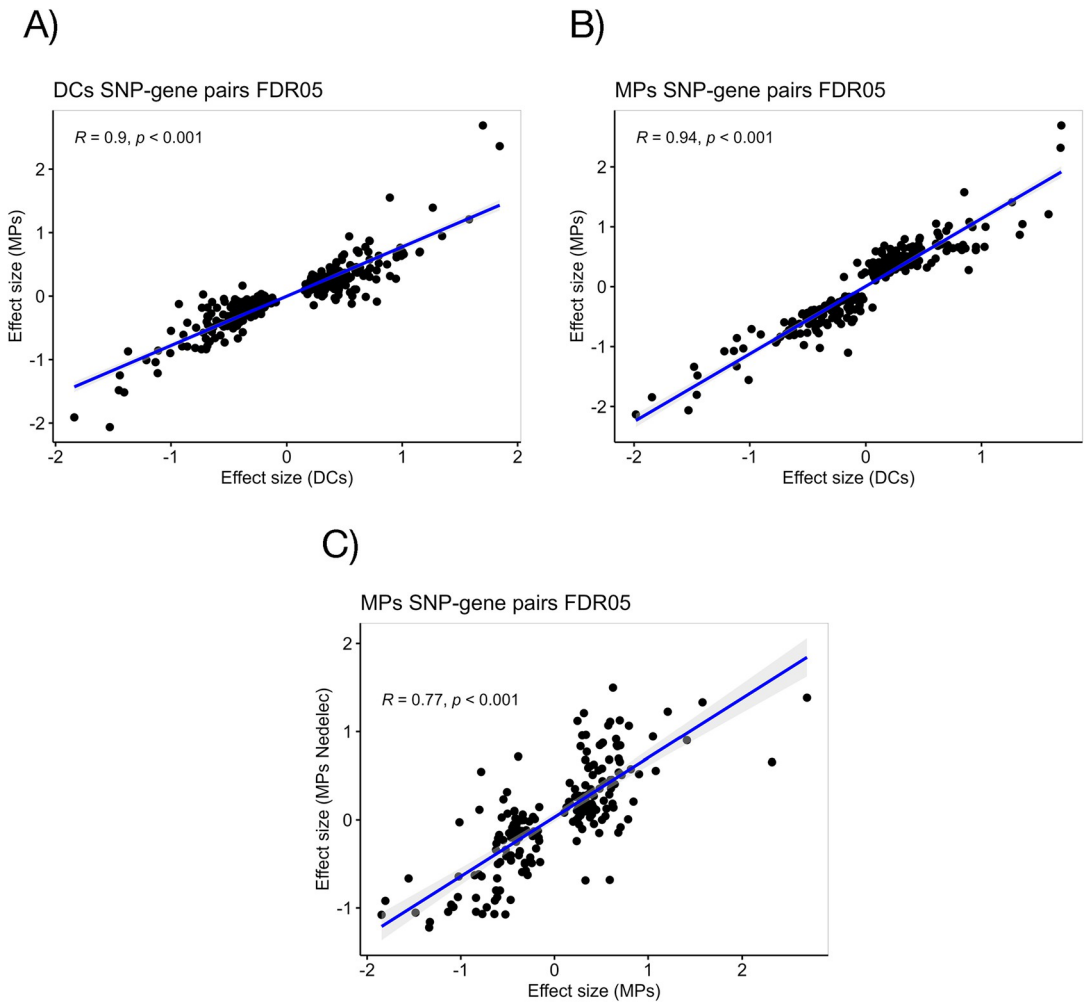

Fig 4. Concordance between effect sizes of significant DC-eQTL events (FDR<0.05) with the macrophage dataset (A), or significant macrophage (MP)-eQTL events (FDR<0.05) with the DC dataset (B). (C) Validation of significant macrophage eQTL using effect sizes of significant eQTL events from monocyte-derived macrophages from Nedelec et al [39] (PMID: 27768889) publicly available on the eQTL catalogue. Pearson correlation analysis results are displayed for A-C.

https://doi.org/10.1371/journal.pgen.1011313.g004

regulated homeobox transcription factor Hoxb8 (ER-Hoxb8), which maintains their self-renewal when cultured in the presence of β-estradiol [59]. We transduced Cas9-transgenic CIM progenitors with a lentivirus expressing guide RNA (gRNA) targeting the antisense strand of exon 2 of *Fah* (the mouse orthologue of human *FAH*) or a non-targeting scramble gRNA. We determined the editing efficiency of the resulting mixed population by sequencing PCR amplicons flanking the targeted region, followed by Inference of CRISPR Edits (ICE) analysis [56], which showed 95% editing efficiency (**Fig 6A**). We then differentiated these progenitors into macrophages by culturing them in β-estradiol-free media supplemented with M-CSF [55]. We infected control and *Fah*-edited CIMs with a luciferase-expressing *Mtb* strain (H37Rv-lux) at a multiplicity of infection (MOI) of 1:1 and quantified fold change in luminescence over 5 days as a reflection of *Mtb* growth. We detected significantly higher *Mtb* luminescence in *Fah*-edited CIMs compared to wildtype counterparts treated with scrambled gRNA at 5 days post-infection (p = 0.0044) (**Fig 6B**). These results are consistent with a protective role for *FAH* against *Mtb* infection.

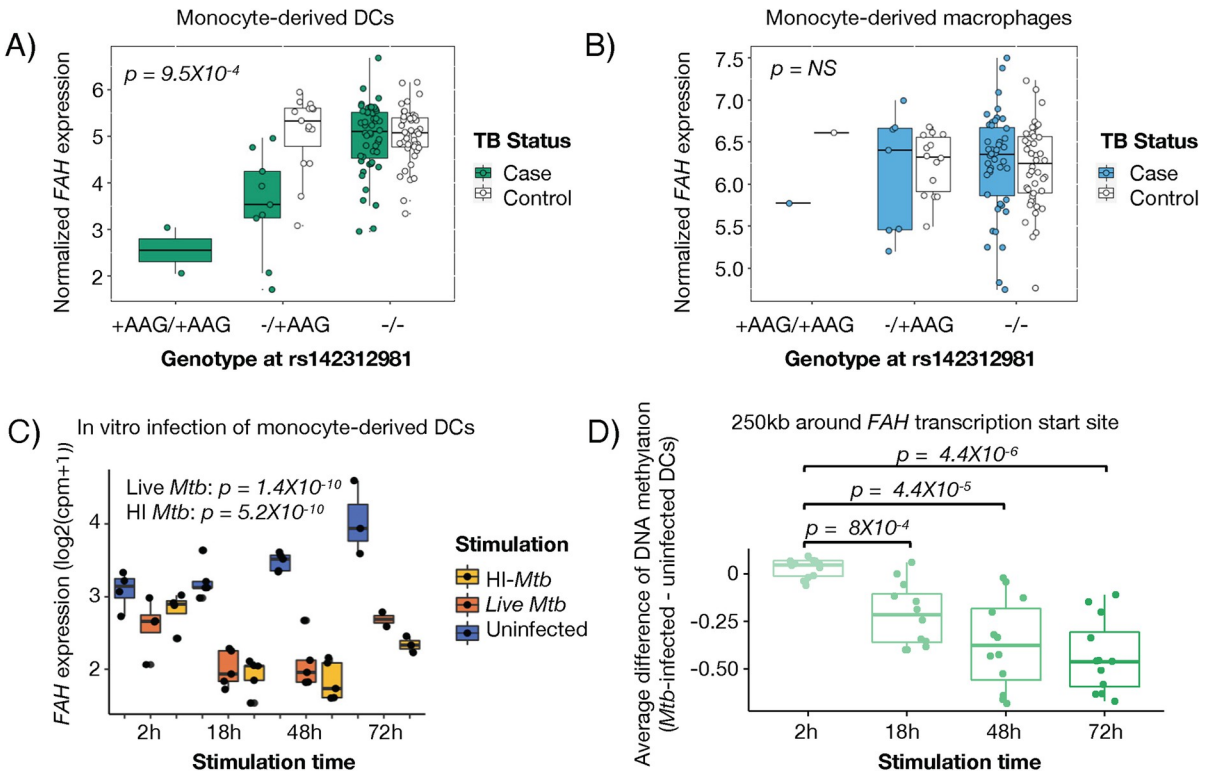

Fig 5. **(A-B)** *FAH* expression levels are stratified by genotype of indel rs142312981. P-value shown is for the TB status interaction with the genetic variant on their effects on gene expression (case = TB progressor, control = non-progressor) in DCs **(A)** and macrophages **(B)**, details of the linear regression and likelihood ratio test are described in Methods. **(C-D)** Figure shows re-analysis of public data from Pacis et al [54] (PMID: 30886108, GSE116405). **(D)** *FAH* expression levels are shown from RNA-seq data from uninfected (blue), *Mtb*-infected (red), heat-killed *Mtb* treated (green) monocyte derived DCs using a mixed-effects linear model. **(D)** Figure shows average mean change in DNA methylation levels between *Mtb*-infected and non-infected monocyte-derived DCs from 6 donors is shown. Data for CpG sites within 250kb (upstream and downstream) of *FAH* transcriptional start site (TSS) are shown.

https://doi.org/10.1371/journal.pgen.1011313.g005

Finally, several studies have demonstrated that gene expression can be regulated by genetic variants in a cell type-dependent manner [27]. Hence, we hypothesized that some genetic variants would regulate gene expression at varying degrees in the two cell types in an expression quantitative trait loci analysis. We identified 30 cell-type eQTL interactions at 10% FDR (**S6 Table**). For example, *CD1A* expression was influenced by the rs366316 SNP in DCs but not in macrophages (**Fig 7A**). In contrast, rs3907022 was strongly associated with *GSDMA* expression in macrophages, but not DCs (**Fig 7B**). We validated this association between rs366316 and CD1A levels in DCs at the protein level with flow cytometry ($p < 2X10^{-16}$, **Fig 7C**). These two cell type-dependent eQTL events were driven by high expression in one cell type and low expression in the other. However, we also observed other eQTL events in which a gene has similar levels of expression in both cell types but is regulated by a genetic variant in only one (e.g. *CTSH*, *DFNA5*, **Fig 7D and 7E**). Hence, our analysis captured different types of cell type-dependent genetic regulatory effects in two related myeloid cell types (macrophages and DCs).

Discussion

It remains unclear how host genetic heterogeneity impacts susceptibility to different *Mtb* infection outcomes [7,9,10,12,25,41,60]. In this study, we re-recruited Peruvians with a history of TB disease or asymptomatic infection with *Mtb*. We analyzed transcriptomes of monocyte-

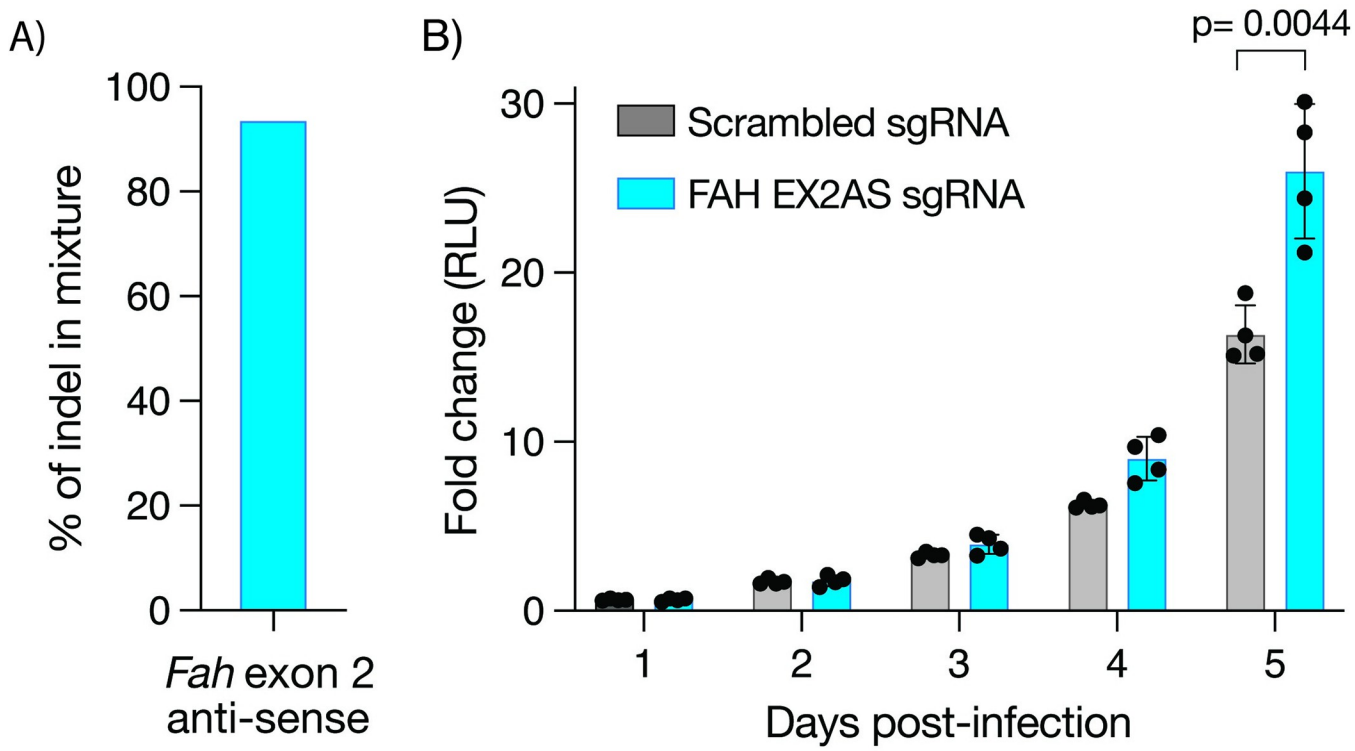

Fig 6. (A) Frequency of indels in *fah*-edited CIMs, consistent with editing efficiency as determined by ICE analysis on amplified region around target site. (B) *Mtb* bacterial growth assay of scramble control and *fah*-edited macrophage CIMs infected with the H37Rv-lux *Mtb* strain. Fold-change in luminescence is reported as *Mtb* growth relative to the average of day 0 for each condition. Data are representative of four independent experiments each performed in quadruplicates; error bar depicts mean ± standard deviation. P-value reflects student t-test of 4 replicates in each condition.

https://doi.org/10.1371/journal.pgen.1011313.g006

derived DCs and macrophages, which showed the expected cell type-specific transcriptional programs [21]. Genetic variants in differentiated myeloid cells are known to influence gene expression levels in a cell type and state-dependent manner [22,24,61]. Identification of cell-specific eQTLs has been described in several studies, even when these cell lineages are derived from a common progenitor, as reported for distinct lineages derived from common induced pluripotent stem cells (iPSCs) [30]. Here, we identified 30 cell type-dependent eQTLs, in an admixed Peruvian cohort, specifically comparing DCs with macrophages, which can serve as a resource for gene regulation in non-European populations. For instance, the regulatory effect of the rs366316 on CD1a, a protein that presents lipid antigens, was previously described in a European-ancestry cohort [62], and points to shared gene and protein expression QTL between Europeans and Peruvians. Interestingly, an independent study reported an intronic variant rs411089 in the CD1a locus in a Vietnamese cohort, which was associated with higher TB susceptibility [13]. However, the DC-specific rs366316-mediated regulation of CD1a expression was not associated with TB history, despite reported roles of CD1a in presenting mycobacterial antigens [63,64].

Noticeably, some of the genetic regulatory events in DCs were associated with history of TB disease. We discovered a clear interaction between TB status and a non-coding variant in their effects on expression of *FAH*, which encodes the Fumarylacetoacetate Hydrolase enzyme. The genetic regulatory effect was surprisingly restricted to monocyte-derived DC samples in individuals with TB disease history, but not the household contacts who remained healthy. We propose two possible explanations for this genetic regulatory effect associated with TB

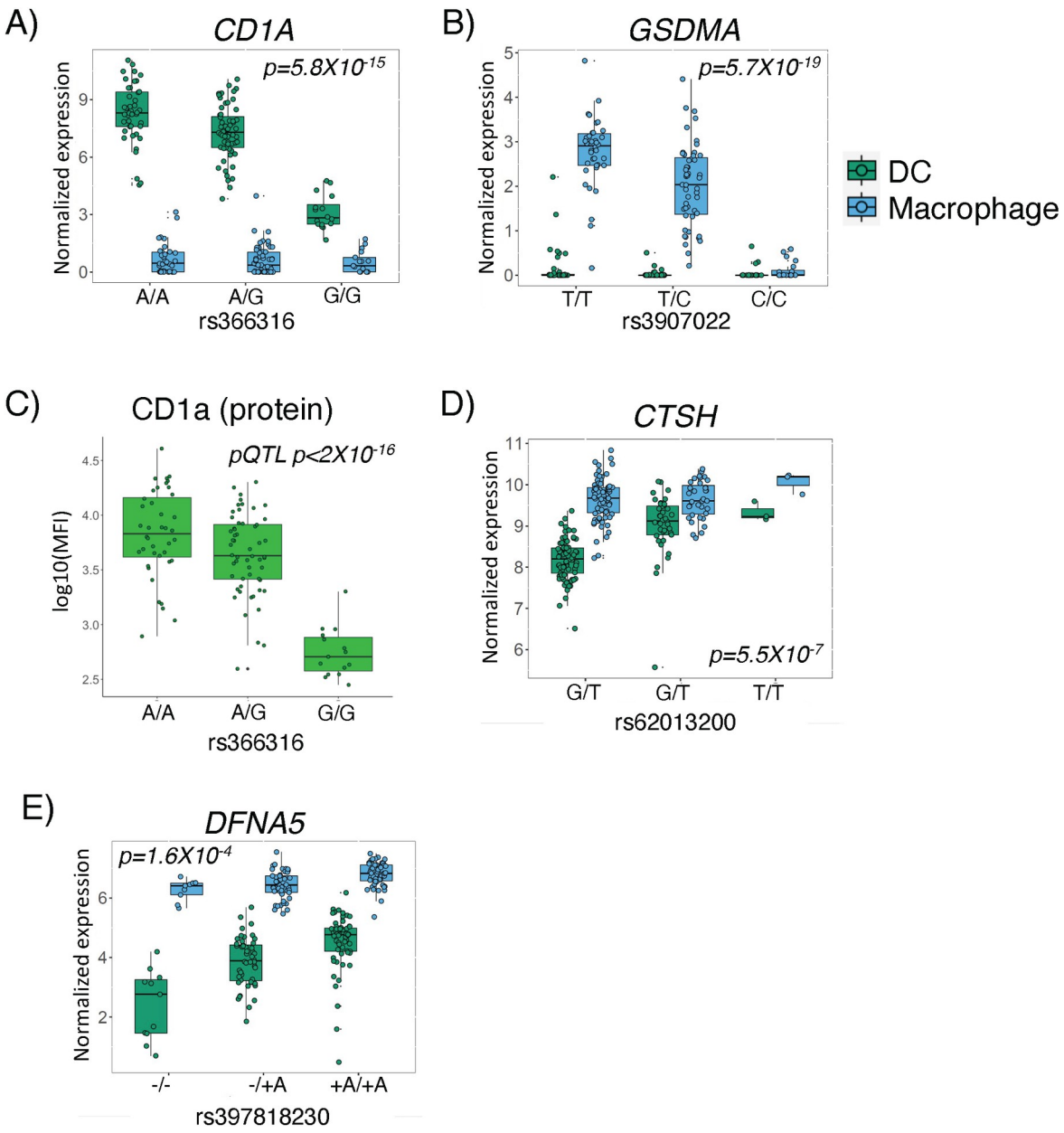

Fig 7. Cell type-dependent effects of genetic variation on gene expression: Top cell type-dependent eQTL in DCs **(A)** and macrophages **(B)** were seen when transcript was expressed at higher levels in one cell type compared to the other. **(C)** Validation of CD1a transcript eQTL at the protein level is shown by flow cytometry. P-value is derived from a univariate linear regression model testing the association between rs366316 genotype and CD1a protein expression by flow cytometry. **(D and E)** Figure shows examples of cell type-dependent eQTL events for genes expressed in both DCs **(D)** and macrophages **(E)**.

https://doi.org/10.1371/journal.pgen.1011313.g007

progression: (1) individuals with susceptibility to TB progression had pre-existing differences in their immune cell state which uncovered their *FAH* eQTL activity, which may have contributed to TB progression, or (2) the *FAH* eQTL appeared in progressors *after* active TB. The TB status interaction eQTL variant was not significantly associated with susceptibility to primary TB progression in the GWAS conducted on this same population (p = 0.49) [7], and we found evidence that the *FAH* locus changed epigenetically and transcriptionally after *Mtb* infection of DCs. Therefore, we believe the second scenario is more likely. Durable remodeling of host

immunity has classically been associated with acquired immunity in lymphocytes, mouse [65] and human [66]. However, studies provide precedent for durable regulation of myeloid cells through epigenetic mechanisms, a phenomenon known as trained immunity. Even though in our study, both progressors and non-progressors have been infected with *Mtb*, a high bacterial burden or strong stimulus occurring during disease could modify the cell state of the host's immune cells, for example through chromatin architecture, with a lasting effect years after disease and treatment. To our knowledge, pathogen-associated eQTLs have been described using *in vitro* infection models, but not after natural infectious disease in humans [24,38]. However, 'epigenetic scars' have been reported as long-term sequalae of TB and thought to be associated with a higher long-term risk of mortality and morbidity [67]. Therefore, our data raises the question of whether eQTLs activated following an infectious disease could contribute to risk of post-infection sequelae such as recurrent TB, akin to reported hypermethylation in the promoter region of interferon induced protein 44 Like (*IFI44L*) gene post-COVID-19 [68].

Several loss-of-function mutations in *FAH* are associated with hereditary tyrosinemia type I in humans, characterized by elevated blood tyrosine levels [69]. In contrast to a system-wide missense or antisense loss-of-function mutation, the variant we describe here is a regulatory variant that alters *FAH* expression in a cell type and state-dependent manner. In this metabolic pathway, tyrosine is converted into homogentisate, then into fumarylacetoacetate, which the FAH enzyme converts into fumarate and acetoacetate. Since infection of DCs with live *Mtb* results in down-regulation of *FAH* expression and infection of *Fah*-edited macrophages results in increased *Mtb* growth, the data raise the hypothesis that interference with tyrosine metabolism is associated with TB progression. Consistent with this hypothesis, a prior study of African household contacts of index TB patients reported that serum tyrosine and phenylalanine, a precursor to tyrosine, were higher in individuals who progressed to active TB than non-progressor counterparts, consistent with impaired tyrosine metabolism during TB progression [5]. The role of *FAH* and tyrosine metabolism in TB warrants detailed investigation in future studies.

Population level differences in TB prevalence highlight the importance of including underrepresented groups in genetic studies [70,71]. Despite successes in defining variants associated with TB disease susceptibility, many susceptibility loci have been reported to be population-specific [7,9,10,12,60]. Differences in genetic ancestry and population history effects have been shown to affect susceptibility to TB, especially in admixed populations [41,72–74]. The minor allele frequency (AAG insertion) of rs142312981 allele is 7% in Peruvian (12% in our cohort) compared to 22% for the European reference genomes (NCI, LDproxy tool [75]). Therefore, this regulatory allele, and others, may have higher impacts in populations where the allele frequencies are higher.

Limitations of this study include profiling the transcriptomes of monocyte-derived DCs and macrophages years after TB disease, where some of the regulatory changes may have reverted to baseline. Furthermore, the *in vitro* differentiation of DCs and macrophages may alter the epigenetic landscape in ways that are different from unmanipulated *ex vivo* cells. We also could not fully rule out whether *Mtb*-mediated downregulation of *FAH* was caused by stimulation of receptors for pattern-associated molecular patterns (PAMPs). However, one publication showed that stimulation of THP-1 cells with several *Mtb*-derived PAMPs did not change *FAH* expression [76]. Finally, the limited sample size may have reduced the power to detect additional significant eQTL genes and eQTL interactions. Hence, future replication studies with larger sample sizes and additional populations will be needed to validate the findings and discover new post-TB regulatory changes, as currently there are no studies with a similar experimental design with a sufficiently large sample size for QTL analysis with TB interaction.

In conclusion, our study identifies previously unknown genetic regulatory variation associated with history of active TB and with cell type of origin. Especially notable is the DC-specific downregulation of FAH because it is a biologically plausible metabolic gene that produces metabolites, fumarate, and acetoacetate, which have been previously proposed to have immune modulatory roles [77,78]. The data raise the question of whether these changes could then compromise subsequent immune responses to *Mtb* infection and contribute to the known higher risk of recurrence among individuals with prior TB disease. Subsequent studies should aim at defining mechanistic links to TB susceptibility and mapping these regulatory changes in different cell types and states, to better understand the long-term regulatory sequelae of TB disease.

Supporting information

S1 Fig. Permutation analysis to confirm the validity of the eQTL interactions observed in DCs. We reshuffled TB status (cases and controls) 1000 time across samples and retesting for TB status-eQTL interactions. P-value cutoffs based on this permutation analysis are shown for top TB:genotype interaction model in DCs (**A**) and macrophages (**B**).
(TIFF)

S1 Table. Antibody cocktails used for analysis of monocyte-derived DCs and macrophages.
(XLSX)

S2 Table. Metadata of RNA-seq samples, include participant characteristics and RNA-seq quality check metrics.
(XLSX)

S3 Table. Significant eQTLs in dendritic cells.
(XLSX)

S4 Table. Significant eQTLs in macrophages.
(XLSX)

S5 Table. TB-eQTL interactions in DCs at 10% FDR.
(XLSX)

S6 Table. Differential expression of TB-eQTL interaction genes in Mtb-infected DCs (public data set, Pacis et al [1]).
(XLSX)

S7 Table. Significant cell type-specific eQTL interactions.
(XLSX)

Author Contributions

Conceptualization: Leonid Lecca, Megan Murray, Ildiko Van Rhijn, Soumya Raychaudhuri, D. Branch Moody, Sara Suliman.

Data curation: Victor E. Nieto-Caballero, Josephine F. Reijneveld, Nalin Abeydeera, Samira Asgari, Sarah K. Iwany, Yang Luo, Aparna Nathan, Daniela Fernandez-Salinas, Marcos Chiñas, Chuan-Chin Huang, Zibiao Zhang, Leonid Lecca, Megan Murray, Ildiko Van Rhijn, Soumya Raychaudhuri, D. Branch Moody, Sara Suliman, Maria Gutierrez-Arcelus.

Formal analysis: Victor E. Nieto-Caballero, Josephine F. Reijneveld, Angel Ruvalcaba, Nalin Abeydeera, Samira Asgari, Sarah K. Iwany, Yang Luo, Aparna Nathan, Daniela Fernandez-Salinas, Marcos Chiñas, Sara Suliman, Maria Gutierrez-Arcelus.

Funding acquisition: Megan Murray, Ildiko Van Rhijn, Soumya Raychaudhuri, D. Branch Moody, Sara Suliman, Maria Gutierrez-Arcelus.

Investigation: Victor E. Nieto-Caballero, Josephine F. Reijneveld, Angel Ruvalcaba, Nalin Abeydeera, Samira Asgari, Kattya Lopez, Sarah K. Iwany, Yang Luo, Aparna Nathan, Daniela Fernandez-Salinas, Marcos Chiñas, Segundo R. León, Roger I. Calderon, Leonid Lecca, Jonathan M. Budzik, Megan Murray, Ildiko Van Rhijn, D. Branch Moody, Sara Suliman, Maria Gutierrez-Arcelus.

Methodology: Victor E. Nieto-Caballero, Josephine F. Reijneveld, Angel Ruvalcaba, Gabriel Innocenzi, Nalin Abeydeera, Kattya Lopez, Sarah K. Iwany, Segundo R. León, Roger I. Calderon, Leonid Lecca, Jonathan M. Budzik, Megan Murray, Ildiko Van Rhijn, Soumya Raychaudhuri, Sara Suliman, Maria Gutierrez-Arcelus.

Project administration: Kattya Lopez, Ildiko Van Rhijn, Soumya Raychaudhuri, D. Branch Moody, Sara Suliman, Maria Gutierrez-Arcelus.

Resources: Chuan-Chin Huang, Zibiao Zhang, Segundo R. León, Roger I. Calderon, Leonid Lecca, Jonathan M. Budzik, Megan Murray, Ildiko Van Rhijn, D. Branch Moody.

Software: Victor E. Nieto-Caballero, Samira Asgari, Yang Luo, Aparna Nathan, Daniela Fernandez-Salinas, Marcos Chiñas, Chuan-Chin Huang, Zibiao Zhang, Sara Suliman, Maria Gutierrez-Arcelus.

Supervision: Megan Murray, Ildiko Van Rhijn, Soumya Raychaudhuri, D. Branch Moody, Sara Suliman, Maria Gutierrez-Arcelus.

Validation: Victor E. Nieto-Caballero, Josephine F. Reijneveld, Angel Ruvalcaba, Gabriel Innocenzi, Nalin Abeydeera, Jonathan M. Budzik, Sara Suliman, Maria Gutierrez-Arcelus.

Visualization: Victor E. Nieto-Caballero, Josephine F. Reijneveld, Daniela Fernandez-Salinas, Marcos Chiñas, Sara Suliman, Maria Gutierrez-Arcelus.

Writing – original draft: Victor E. Nieto-Caballero, Josephine F. Reijneveld, Sara Suliman, Maria Gutierrez-Arcelus.

Writing – review & editing: Victor E. Nieto-Caballero, Josephine F. Reijneveld, Nalin Abeydeera, Samira Asgari, Yang Luo, Aparna Nathan, Jonathan M. Budzik, Megan Murray, Ildiko Van Rhijn, Soumya Raychaudhuri, D. Branch Moody, Sara Suliman, Maria Gutierrez-Arcelus.

References

1. WHO. Global Tuberculosis Report 2022. 2022.
2. Vega V, Rodriguez S, Van der Stuyft P, Seas C, Otero L. Recurrent TB: a systematic review and meta-analysis of the incidence rates and the proportions of relapses and reinfections. Thorax. 2021; 76 (5):494–502. Epub 2021/02/07. https://doi.org/10.1136/thoraxjnl-2020-215449 PMID: 33547088; PubMed Central PMCID: PMC8225554.
3. Berry MP, Graham CM, McNab FW, Xu Z, Bloch SA, Oni T, et al. An interferon-inducible neutrophil-driven blood transcriptional signature in human tuberculosis. Nature. 2010; 466(7309):973–7. Epub 2010/08/21. https://doi.org/10.1038/nature09247 PMID: 20725040; PubMed Central PMCID: PMC3492754.

4. Zak DE, Penn-Nicholson A, Scriba TJ, Thompson E, Suliman S, Amon LM, et al. A blood RNA signature for tuberculosis disease risk: a prospective cohort study. Lancet. 2016; 387(10035):2312–22. https://doi.org/10.1016/S0140-6736(15)01316-1 PMID: 27017310.
5. Weiner J 3rd, Maertzdorf J, Sutherland JS, Duffy FJ, Thompson E, Suliman S, et al. Metabolite changes in blood predict the onset of tuberculosis. Nat Commun. 2018; 9(1):5208. Epub 2018/12/13. https://doi.org/10.1038/s41467-018-07635-7 PMID: 30523338; PubMed Central PMCID: PMC6283869.
6. Nathan A, Beynor JI, Baglaenko Y, Suliman S, Ishigaki K, Asgari S, et al. Multimodally profiling memory T cells from a tuberculosis cohort identifies cell state associations with demographics, environment and disease. Nat Immunol. 2021; 22(6):781–93. Epub 2021/05/26. https://doi.org/10.1038/s41590-021-00933-1 PMID: 34031617; PubMed Central PMCID: PMC8162307.
7. Luo Y, Suliman S, Asgari S, Amariuta T, Baglaenko Y, Martinez-Bonet M, et al. Early progression to active tuberculosis is a highly heritable trait driven by 3q23 in Peruvians. Nat Commun. 2019; 10(1):3765. https://doi.org/10.1038/s41467-019-11664-1 PMID: 31434886; PubMed Central PMCID: PMC6704092.
8. Nathan A, Asgari S, Ishigaki K, Valencia C, Amariuta T, Luo Y, et al. Single-cell eQTL models reveal dynamic T cell state dependence of disease loci. Nature. 2022; 606(7912):120–8. Epub 2022/05/12. https://doi.org/10.1038/s41586-022-04713-1 PMID: 35545678.
9. Curtis J, Luo Y, Zenner HL, Cuchet-Lourenco D, Wu C, Lo K, et al. Susceptibility to tuberculosis is associated with variants in the ASAP1 gene encoding a regulator of dendritic cell migration. Nat Genet. 2015; 47(5):523–7. Epub 2015/03/17. https://doi.org/10.1038/ng.3248 PMID: 25774636; PubMed Central PMCID: PMC4414475.
10. Thye T, Vannberg FO, Wong SH, Owusu-Dabo E, Osei I, Gyapong J, et al. Genome-wide association analyses identifies a susceptibility locus for tuberculosis on chromosome 18q11.2. Nat Genet. 2010; 42(9):739–41. Epub 2010/08/10. https://doi.org/10.1038/ng.639 PMID: 20694014; PubMed Central PMCID: PMC4975513.
11. van Tong H, Velavan TP, Thye T, Meyer CG. Human genetic factors in tuberculosis: an update. Trop Med Int Health. 2017; 22(9):1063–71. Epub 2017/07/08. https://doi.org/10.1111/tmi.12923 PMID: 28685916.
12. Schurz H, Kinnear CJ, Gignoux C, Wojcik G, van Helden PD, Tromp G, et al. A Sex-Stratified Genome-Wide Association Study of Tuberculosis Using a Multi-Ethnic Genotyping Array. Front Genet. 2018; 9:678. https://doi.org/10.3389/fgene.2018.00678 PMID: 30713548; PubMed Central PMCID: PMC6346682.
13. Seshadri C, Thuong NT, Yen NT, Bang ND, Chau TT, Thwaites GE, et al. A polymorphism in human CD1A is associated with susceptibility to tuberculosis. Genes Immun. 2014; 15(3):195–8. https://doi.org/10.1038/gene.2014.5 PMID: 24500401; PubMed Central PMCID: PMC3998877.
14. Khader SA, Divangahi M, Hanekom W, Hill PC, Maeurer M, Makar KW, et al. Targeting innate immunity for tuberculosis vaccination. J Clin Invest. 2019; 129(9):3482–91. Epub 2019/09/04. https://doi.org/10.1172/JCI128877 PMID: 31478909; PubMed Central PMCID: PMC6715374.
15. Scriba TJ, Penn-Nicholson A, Shankar S, Hraha T, Thompson EG, Sterling D, et al. Sequential inflammatory processes define human progression from M. tuberculosis infection to tuberculosis disease. PLoS Pathog. 2017; 13(11):e1006687. https://doi.org/10.1371/journal.ppat.1006687 PMID: 29145483; PubMed Central PMCID: PMC5689825.
16. Suliman S, Thompson E, Sutherland J, Weiner Rd J, Ota MOC, Shankar S, et al. Four-gene Pan-African Blood Signature Predicts Progression to Tuberculosis. Am J Respir Crit Care Med. 2018. https://doi.org/10.1164/rccm.201711-2340OC PMID: 29624071; PubMed Central PMCID: PMC6019933.
17. Singhania A, Verma R, Graham CM, Lee J, Tran T, Richardson M, et al. A modular transcriptional signature identifies phenotypic heterogeneity of human tuberculosis infection. Nat Commun. 2018; 9(1):2308. https://doi.org/10.1038/s41467-018-04579-w PMID: 29921861; PubMed Central PMCID: PMC6008327.
18. VanderVen BC, Huang L, Rohde KH, Russell DG. The Minimal Unit of Infection: Mycobacterium tuberculosis in the Macrophage. Microbiol Spectr. 2016; 4(6). Epub 2017/01/14. https://doi.org/10.1128/microbiolspec.TBTB2-0025-2016 PMID: 28084213; PubMed Central PMCID: PMC5245711.
19. Rodrigues TS, Conti BJ, Fraga-Silva TFC, Almeida F, Bonato VLD. Interplay between alveolar epithelial and dendritic cells and Mycobacterium tuberculosis. J Leukoc Biol. 2020; 108(4):1139–56. Epub 2020/07/04. https://doi.org/10.1002/JLB.4MR0520-112R PMID: 32620048.
20. Sallusto F, Lanzavecchia A. Efficient presentation of soluble antigen by cultured human dendritic cells is maintained by granulocyte/macrophage colony-stimulating factor plus interleukin 4 and downregulated by tumor necrosis factor alpha. J Exp Med. 1994; 179(4):1109–18. https://doi.org/10.1084/jem.179.4.1109 PMID: 8145033; PubMed Central PMCID: PMC2191432.

21. Sander J, Schmidt SV, Cirovic B, McGovern N, Papantonopoulou O, Hardt AL, et al. Cellular Differentiation of Human Monocytes Is Regulated by Time-Dependent Interleukin-4 Signaling and the Transcriptional Regulator NCOR2. Immunity. 2017; 47(6):1051–66 e12. https://doi.org/10.1016/j.immuni.2017.11.024 PMID: 29262348; PubMed Central PMCID: PMC5772172.

22. Blischak JD, Tailleux L, Myrthil M, Charlois C, Bergot E, Dinh A, et al. Predicting susceptibility to tuberculosis based on gene expression profiling in dendritic cells. Sci Rep. 2017; 7(1):5702. Epub 2017/07/20. https://doi.org/10.1038/s41598-017-05878-w PMID: 28720766; PubMed Central PMCID: PMC5516010.

23. Roura-Mir C, Wang L, Cheng TY, Matsunaga I, Dascher CC, Peng SL, et al. Mycobacterium tuberculosis regulates CD1 antigen presentation pathways through TLR-2. J Immunol. 2005; 175(3):1758–66. https://doi.org/10.4049/jimmunol.175.3.1758 PMID: 16034117.

24. Barreiro LB, Tailleux L, Pai AA, Gicquel B, Marioni JC, Gilad Y. Deciphering the genetic architecture of variation in the immune response to Mycobacterium tuberculosis infection. Proc Natl Acad Sci U S A. 2012; 109(4):1204–9. Epub 2012/01/12. https://doi.org/10.1073/pnas.1115761109 PMID: 22233810; PubMed Central PMCID: PMC3268270.

25. Jiao L, Song J, Ding L, Liu T, Wu T, Zhang J, et al. A Novel Genetic Variation in NCF2, the Core Component of NADPH Oxidase, Contributes to the Susceptibility of Tuberculosis in Western Chinese Han Population. DNA Cell Biol. 2020; 39(1):57–62. Epub 2019/12/04. https://doi.org/10.1089/dna.2019.5082 PMID: 31794672.

26. Consortium GT, Laboratory DA, Coordinating Center -Analysis Working G, Statistical Methods groups-Analysis Working G, Enhancing Gg, Fund NIHC, et al. Genetic effects on gene expression across human tissues. Nature. 2017; 550(7675):204–13. https://doi.org/10.1038/nature24277 PMID: 29022597; PubMed Central PMCID: PMC5776756.

27. Consortium GT. The GTEx Consortium atlas of genetic regulatory effects across human tissues. Science. 2020; 369(6509):1318–30. Epub 2020/09/12. https://doi.org/10.1126/science.aaz1776 PMID: 32913098; PubMed Central PMCID: PMC7737656.

28. Gutierrez-Arcelus M, Ongen H, Lappalainen T, Montgomery SB, Buil A, Yurovsky A, et al. Tissue-specific effects of genetic and epigenetic variation on gene regulation and splicing. PLoS Genet. 2015; 11(1):e1004958. Epub 2015/01/31. https://doi.org/10.1371/journal.pgen.1004958 PMID: 25634236; PubMed Central PMCID: PMC4310612.

29. Soskic B, Cano-Gamez E, Smyth DJ, Ambridge K, Ke Z, Matte JC, et al. Immune disease risk variants regulate gene expression dynamics during CD4(+) T cell activation. Nat Genet. 2022; 54(6):817–26. Epub 2022/05/27. https://doi.org/10.1038/s41588-022-01066-3 PMID: 35618845; PubMed Central PMCID: PMC9197762.

30. Strober BJ, Elorbany R, Rhodes K, Krishnan N, Tayeb K, Battle A, et al. Dynamic genetic regulation of gene expression during cellular differentiation. Science. 2019; 364(6447):1287–90. Epub 2019/06/30. https://doi.org/10.1126/science.aaw0040 PMID: 31249060; PubMed Central PMCID: PMC6623972.

31. Randolph HE, Fiege JK, Thielen BK, Mickelson CK, Shiratori M, Barroso-Batista J, et al. Genetic ancestry effects on the response to viral infection are pervasive but cell type specific. Science. 2021; 374(6571):1127–33. Epub 2021/11/26. https://doi.org/10.1126/science.abg0928 PMID: 34822289; PubMed Central PMCID: PMC8957271.

32. Yazar S, Alquicira-Hernandez J, Wing K, Senabouth A, Gordon MG, Andersen S, et al. Single-cell eQTL mapping identifies cell type-specific genetic control of autoimmune disease. Science. 2022; 376(6589):eabf3041. Epub 20220408. https://doi.org/10.1126/science.abf3041 PMID: 35389779.

33. Fairfax BP, Humburg P, Makino S, Naranbhai V, Wong D, Lau E, et al. Innate immune activity conditions the effect of regulatory variants upon monocyte gene expression. Science. 2014; 343(6175):1246949. Epub 2014/03/08. https://doi.org/10.1126/science.1246949 PMID: 24604202; PubMed Central PMCID: PMC4064786.

34. Gutierrez-Arcelus M, Baglaenko Y, Arora J, Hannes S, Luo Y, Amariuta T, et al. Allele-specific expression changes dynamically during T cell activation in HLA and other autoimmune loci. Nat Genet. 2020; 52(3):247–53. Epub 20200217. https://doi.org/10.1038/s41588-020-0579-4 PMID: 32066938; PubMed Central PMCID: PMC7135372.

35. Umans BD, Battle A, Gilad Y. Where Are the Disease-Associated eQTLs? Trends Genet. 2021; 37(2):109–24. Epub 2020/09/12. https://doi.org/10.1016/j.tig.2020.08.009 PMID: 32912663.

36. Fairfax BP, Knight JC. Genetics of gene expression in immunity to infection. Curr Opin Immunol. 2014; 30:63–71. Epub 2014/08/01. https://doi.org/10.1016/j.coi.2014.07.001 PMID: 25078545; PubMed Central PMCID: PMC4426291.

37. Klunk J, Vilgalys TP, Demeure CE, Cheng X, Shiratori M, Madej J, et al. Evolution of immune genes is associated with the Black Death. Nature. 2022; 611(7935):312–9. Epub 2022/10/20. https://doi.org/10.1038/s41586-022-05349-x PMID: 36261521; PubMed Central PMCID: PMC9580435 Biosciences,

which provided the myBaits hybridization capture kits for this work. All other authors declare no competing interests.

38. Oelen R, de Vries DH, Brugge H, Gordon MG, Vochteloo M, single-cell e Qc, et al. Single-cell RNA-sequencing of peripheral blood mononuclear cells reveals widespread, context-specific gene expression regulation upon pathogenic exposure. Nat Commun. 2022; 13(1):3267. Epub 2022/06/08. https://doi.org/10.1038/s41467-022-30893-5 PMID: 35672358; PubMed Central PMCID: PMC9174272 Sciences and ImmunAI, a consultant for and hold equity in Maze Therapeutics, and a consultant for TReX Bio. C.J.Y. has received research support from Chan Zuckerberg Initiative, Chan Zuckerberg Biohub, and Genentech. The remaining authors declare no competing interests.

39. Nedelec Y, Sanz J, Baharian G, Szpiech ZA, Pacis A, Dumaine A, et al. Genetic Ancestry and Natural Selection Drive Population Differences in Immune Responses to Pathogens. Cell. 2016; 167(3):657–69 e21. https://doi.org/10.1016/j.cell.2016.09.025 PMID: 27768889.

40. Comas I, Coscolla M, Luo T, Borrell S, Holt KE, Kato-Maeda M, et al. Out-of-Africa migration and Neolithic coexpansion of Mycobacterium tuberculosis with modern humans. Nat Genet. 2013; 45(10):1176–82. https://doi.org/10.1038/ng.2744 PMID: 23995134; PubMed Central PMCID: PMC3800747.

41. Asgari S, Luo Y, Huang CC, Zhang Z, Calderon R, Jimenez J, et al. Higher native Peruvian genetic ancestry proportion is associated with tuberculosis progression risk. Cell Genom. 2022; 2(7). Epub 2022/07/26. https://doi.org/10.1016/j.xgen.2022.100151 PMID: 35873671; PubMed Central PMCID: PMC9306274.

42. Becerra MC, Huang CC, Lecca L, Bayona J, Contreras C, Calderon R, et al. Transmissibility and potential for disease progression of drug resistant Mycobacterium tuberculosis: prospective cohort study. BMJ. 2019; 367:l5894. Epub 2019/10/28. https://doi.org/10.1136/bmj.l5894 PMID: 31649017; PubMed Central PMCID: PMC6812583 at http://www.icmje.org/coi_disclosure.pdf and declare: support from the US National Institutes of Health for the submitted work; no financial relationships with any organizations that might have an interest in the submitted work in the previous three years, no other relationships or activities that could appear to have influenced the submitted work.

43. Picelli S, Faridani OR, Bjorklund AK, Winberg G, Sagasser S, Sandberg R. Full-length RNA-seq from single cells using Smart-seq2. Nat Protoc. 2014; 9(1):171–81. https://doi.org/10.1038/nprot.2014.006 PMID: 24385147.

44. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013; 29(1):15–21. Epub 2012/10/30. https://doi.org/10.1093/bioinformatics/bts635 PMID: 23104886; PubMed Central PMCID: PMC3530905.

45. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics. 2011; 12:323. Epub 2011/08/06. https://doi.org/10.1186/1471-2105-12-323 PMID: 21816040; PubMed Central PMCID: PMC3163565.

46. DeLuca DS, Levin JZ, Sivachenko A, Fennell T, Nazaire MD, Williams C, et al. RNA-SeQC: RNA-seq metrics for quality control and process optimization. Bioinformatics. 2012; 28(11):1530–2. Epub 2012/04/28. https://doi.org/10.1093/bioinformatics/bts196 PMID: 22539670; PubMed Central PMCID: PMC3356847.

47. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015; 43(7):e47. Epub 2015/01/22. https://doi.org/10.1093/nar/gkv007 PMID: 25605792; PubMed Central PMCID: PMC4402510.

48. O'Connell J, Gurdasani D, Delaneau O, Pirastu N, Ulivi S, Cocca M, et al. A general approach for haplotype phasing across the full spectrum of relatedness. PLoS Genet. 2014; 10(4):e1004234. Epub 20140417. https://doi.org/10.1371/journal.pgen.1004234 PMID: 24743097; PubMed Central PMCID: PMC3990520.

49. Howie B, Fuchsberger C, Stephens M, Marchini J, Abecasis GR. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. Nat Genet. 2012; 44(8):955–9. Epub 20120722. https://doi.org/10.1038/ng.2354 PMID: 22820512; PubMed Central PMCID: PMC3696580.

50. Delaneau O, Ongen H, Brown AA, Fort A, Panousis NI, Dermitzakis ET. A complete tool set for molecular QTL discovery and analysis. Nat Commun. 2017; 8:15452. Epub 2017/05/19. https://doi.org/10.1038/ncomms15452 PMID: 28516912; PubMed Central PMCID: PMC5454369.

51. Marderstein AR, Davenport ER, Kulm S, Van Hout CV, Elemento O, Clark AG. Leveraging phenotypic variability to identify genetic interactions in human phenotypes. Am J Hum Genet. 2021; 108(1):49–67. Epub 2020/12/17. https://doi.org/10.1016/j.ajhg.2020.11.016 PMID: 33326753; PubMed Central PMCID: PMC7820920.

52. Storey JDaB A. J. qvalue: Q-value estimation for false discovery rate control. R package version 2280. 2022;http://github.com/jdstorey/qvalue.

53. Kerimov N, Hayhurst JD, Peikova K, Manning JR, Walter P, Kolberg L, et al. A compendium of uniformly processed human gene expression and splicing quantitative trait loci. Nat Genet. 2021; 53(9):1290–9.

Epub 20210906. https://doi.org/10.1038/s41588-021-00924-w PMID: 34493866; PubMed Central PMCID: PMC8423625.

54. Pacis A, Mailhot-Leonard F, Tailleux L, Randolph HE, Yotova V, Dumaine A, et al. Gene activation precedes DNA demethylation in response to infection in human dendritic cells. Proc Natl Acad Sci U S A. 2019; 116(14):6938–43. Epub 2019/03/20. https://doi.org/10.1073/pnas.1814700116 PMID: 30886108; PubMed Central PMCID: PMC6452747.

55. Roberts AW, Popov LM, Mitchell G, Ching KL, Licht DJ, Golovkine G, et al. Cas9(+) conditionally-immortalized macrophages as a tool for bacterial pathogenesis and beyond. Elife. 2019;8. Epub 2019/06/18. https://doi.org/10.7554/eLife.45957 PMID: 31204998; PubMed Central PMCID: PMC6579556.

56. Conant D, Hsiau T, Rossi N, Oki J, Maures T, Waite K, et al. Inference of CRISPR Edits from Sanger Trace Data. CRISPR J. 2022; 5(1):123–30. Epub 20220202. https://doi.org/10.1089/crispr.2021.0113 PMID: 35119294.

57. Andreu N, Zelmer A, Fletcher T, Elkington PT, Ward TH, Ripoll J, et al. Optimisation of bioluminescent reporters for use with mycobacteria. PLoS One. 2010; 5(5):e10777. Epub 2010/06/04. https://doi.org/10.1371/journal.pone.0010777 PMID: 20520722; PubMed Central PMCID: PMC2875389.

58. Weiss AKH, Loeffler JR, Liedl KR, Gstach H, Jansen-Durr P. The fumarylacetoacetate hydrolase (FAH) superfamily of enzymes: multifunctional enzymes from microbes to mitochondria. Biochem Soc Trans. 2018; 46(2):295–309. Epub 2018/03/01. https://doi.org/10.1042/BST20170518 PMID: 29487229.

59. Wang GG, Calvo KR, Pasillas MP, Sykes DB, Hacker H, Kamps MP. Quantitative production of macrophages or neutrophils ex vivo using conditional Hoxb8. Nat Methods. 2006; 3(4):287–93. https://doi.org/10.1038/nmeth865 PMID: 16554834.

60. Thye T, Owusu-Dabo E, Vannberg FO, van Crevel R, Curtis J, Sahiratmadja E, et al. Common variants at 11p13 are associated with susceptibility to tuberculosis. Nat Genet. 2012; 44(3):257–9. Epub 2012/02/07. https://doi.org/10.1038/ng.1080 PMID: 22306650; PubMed Central PMCID: PMC3427019.

61. Blischak JD, Tailleux L, Mitrano A, Barreiro LB, Gilad Y. Mycobacterial infection induces a specific human innate immune response. Sci Rep. 2015; 5:16882. Epub 2015/11/21. https://doi.org/10.1038/srep16882 PMID: 26586179; PubMed Central PMCID: PMC4653619.

62. Seshadri C, Shenoy M, Wells RD, Hensley-McBain T, Andersen-Nissen E, McElrath MJ, et al. Human CD1a deficiency is common and genetically regulated. J Immunol. 2013; 191(4):1586–93. https://doi.org/10.4049/jimmunol.1300575 PMID: 23858036; PubMed Central PMCID: PMC3748949.

63. Kasmar AG, Van Rhijn I, Magalhaes KG, Young DC, Cheng TY, Turner MT, et al. Cutting Edge: CD1a tetramers and dextramers identify human lipopeptide-specific T cells ex vivo. J Immunol. 2013; 191(9):4499–503. https://doi.org/10.4049/jimmunol.1301660 PMID: 24089190; PubMed Central PMCID: PMC3845436.

64. Moody DB, Young DC, Cheng TY, Rosat JP, Roura-Mir C, O'Connor PB, et al. T cell activation by lipopeptide antigens. Science. 2004; 303(5657):527–31. https://doi.org/10.1126/science.1089353 PMID: 14739458.

65. Netea MG, Quintin J, van der Meer JW. Trained immunity: a memory for innate host defense. Cell Host Microbe. 2011; 9(5):355–61. https://doi.org/10.1016/j.chom.2011.04.006 PMID: 21575907.

66. Kleinnijenhuis J, Quintin J, Preijers F, Joosten LA, Ifrim DC, Saeed S, et al. Bacille Calmette-Guerin induces NOD2-dependent nonspecific protection from reinfection via epigenetic reprogramming of monocytes. Proc Natl Acad Sci U S A. 2012; 109(43):17537–42. Epub 2012/09/19. https://doi.org/10.1073/pnas.1202870109 PMID: 22988082; PubMed Central PMCID: PMC3491454.

67. DiNardo AR, Netea MG, Musher DM. Postinfectious Epigenetic Immune Modifications—A Double-Edged Sword. N Engl J Med. 2021; 384(3):261–70. https://doi.org/10.1056/NEJMra2028358 PMID: 33471978; PubMed Central PMCID: PMC8053819.

68. Lee Y, Riskedal E, Kalleberg KT, Istre M, Lind A, Lund-Johansen F, et al. EWAS of post-COVID-19 patients shows methylation differences in the immune-response associated gene, IFI44L, three months after COVID-19 infection. Sci Rep. 2022; 12(1):11478. Epub 2022/07/08. https://doi.org/10.1038/s41598-022-15467-1 PMID: 35798818; PubMed Central PMCID: PMC9261254 authors declare no competing interests.

69. Morrow G, Angileri F, Tanguay RM. Molecular Aspects of the FAH Mutations Involved in HT1 Disease. Adv Exp Med Biol. 2017; 959:25–48. Epub 2017/07/30. https://doi.org/10.1007/978-3-319-55780-9_3 PMID: 28755182.

70. Bentley AR, Callier S, Rotimi CN. Diversity and inclusion in genomic research: why the uneven progress? J Community Genet. 2017; 8(4):255–66. Epub 2017/08/05. https://doi.org/10.1007/s12687-017-0316-6 PMID: 28770442; PubMed Central PMCID: PMC5614884.

71. Swart Y, Uren C, Eckold C, Cliff JM, Malherbe ST, Ronacher K, et al. cis-eQTL mapping of TB-T2D comorbidity elucidates the involvement of African ancestry in TB susceptibility. MedRxiv pre-print:

https://wwwbiorxivorg/content/101101/20221019512814v1 2022;Accessed online on December 23rd, 2022.

72. Kerner G, Laval G, Patin E, Boisson-Dupuis S, Abel L, Casanova JL, et al. Human ancient DNA analyses reveal the high burden of tuberculosis in Europeans over the last 2,000 years. Am J Hum Genet. 2021; 108(3):517–24. Epub 2021/03/06. https://doi.org/10.1016/j.ajhg.2021.02.009 PMID: 33667394; PubMed Central PMCID: PMC8008489.

73. Chimusa ER, Zaitlen N, Daya M, Moller M, van Helden PD, Mulder NJ, et al. Genome-wide association study of ancestry-specific TB risk in the South African Coloured population. Hum Mol Genet. 2014; 23 (3):796–809. Epub 2013/09/24. https://doi.org/10.1093/hmg/ddt462 PMID: 24057671; PubMed Central PMCID: PMC3888262.

74. Daya M, van der Merwe L, van Helden PD, Moller M, Hoal EG. The role of ancestry in TB susceptibility of an admixed South African population. Tuberculosis (Edinb). 2014; 94(4):413–20. Epub 2014/05/17. https://doi.org/10.1016/j.tube.2014.03.012 PMID: 24832562.

75. Alexander TA, Machiela MJ. LDpop: an interactive online tool to calculate and visualize geographic LD patterns. BMC Bioinformatics. 2020; 21(1):14. Epub 2020/01/12. https://doi.org/10.1186/s12859-020-3340-1 PMID: 31924160; PubMed Central PMCID: PMC6954550.

76. Jani C, Solomon SL, Peters JM, Pringle SC, Hinman AE, Boucau J, et al. TLR2 is non-redundant in the population and subpopulation responses to Mycobacterium tuberculosis in macrophages and in vivo. mSystems. 2023; 8(4):e0005223. Epub 20230713. https://doi.org/10.1128/msystems.00052-23 PMID: 37439558; PubMed Central PMCID: PMC10506474.

77. Adam C, Paolini L, Gueguen N, Mabilleau G, Preisser L, Blanchard S, et al. Acetoacetate protects macrophages from lactic acidosis-induced mitochondrial dysfunction by metabolic reprograming. Nat Commun. 2021; 12(1):7115. Epub 2021/12/10. https://doi.org/10.1038/s41467-021-27426-x PMID: 34880237; PubMed Central PMCID: PMC8655019.

78. Mita-Mendoza NK, Magallon-Tejada A, Parmar P, Furtado R, Aldrich M, Saidi A, et al. Dimethyl fumarate reduces TNF and Plasmodium falciparum induced brain endothelium activation in vitro. Malar J. 2020; 19(1):376. Epub 2020/10/23. https://doi.org/10.1186/s12936-020-03447-7 PMID: 33087130; PubMed Central PMCID: PMC7579885.