

RESEARCH

Open Access



The spectrum of tuberculosis described as differential DNA methylation patterns in alveolar macrophages and alveolar T cells

Isabelle Pehrson^{1†}, Shumaila Sayyab^{1†}, Jyotirmoy Das^{1,2,3}, Nina Idh¹, Jakob Paus^{1,4}, Melissa Méndez-Aranda⁵, César Ugarte-Gil^{6,7†} and Maria Lerm^{1*†}

Abstract

Background: Host innate immune cells have been identified as key players in the early eradication of *Mycobacterium tuberculosis* and in the maintenance of an anti-mycobacterial immune memory, which we and others have shown are induced through epigenetic reprogramming. Studies on human tuberculosis immunity are dominated by those using peripheral blood as surrogate markers for immunity. We aimed to investigate DNA methylation patterns in immune cells of the lung compartment by obtaining induced sputum from *M. tuberculosis*-exposed subjects including symptom-free subjects testing positively and negatively for latent tuberculosis as well as patients diagnosed with active tuberculosis. Alveolar macrophages and alveolar T cells were isolated from the collected sputum and DNA methylome analyses performed (Illumina Infinium Human Methylation 450 k).

Results: Multidimensional scaling analysis revealed that DNA methylomes of cells from the tuberculosis-exposed subjects and controls appeared as separate clusters. The numerous genes that were differentially methylated between the groups were functionally connected and overlapped with previous findings of trained immunity and tuberculosis. In addition, analysis of the interferon-gamma release assay (IGRA) status of the subjects demonstrated that the IGRA status was reflected in the DNA methylome by a unique signature.

Conclusions: This pilot study suggests that *M. tuberculosis* induces epigenetic reprogramming in immune cells of the lung compartment, reflected as a specific DNA methylation pattern. The DNA methylation signature emerging from the comparison of IGRA-negative and IGRA-positive subjects revealed a spectrum of signature strength with the TB patients grouping together at one end of the spectrum, both in alveolar macrophages and T cells. DNA methylation-based biosignatures could be considered for further development towards a clinically useful tool for determining tuberculosis infection status and the level of tuberculosis exposure.

Keywords: DNA methylation, Tuberculosis, Biosignature, Epigenetics, Sputum induction, IGRA

[†]Isabelle Pehrson and Shumaila Sayyab contributed equally to this work

[†]César Ugarte-Gil and Maria Lerm share last authorship

*Correspondence: maria.lerm@liu.se

¹ Division of Inflammation and Infection, Department of Biomedical and Clinical Sciences, Faculty of Medicine and Health Sciences, Linköping University, Lab 1, Floor 12, 58185 Linköping, Sweden
Full list of author information is available at the end of the article

Background

Tuberculosis (TB) is a pulmonary infection and until the COVID-19 pandemic it was the leading cause of death from a single infectious agent [1]. An expansion of the current toolkit for diagnosis, prevention and treatment is critical for reaching the United Nations' Sustainable Development Goals for 2030 of ending the TB epidemic [2]. TB is caused by *Mycobacterium tuberculosis*, which transmit via aerosols and target alveolar macrophages



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

in exposed individuals [3]. In recent years, the concept of trained immunity has evolved as an epigenetically encoded immune memory that can be triggered by a variety of stimuli and is reflected in a reprogrammed immune state characterized by a higher magnitude of response to subsequent pathogen challenges. The discovery of epigenetically regulated antimicrobial defense mechanisms goes beyond the classical understanding of immune defense and opens a new field of research. Along this line, we have demonstrated that administration of the BCG vaccine to healthy subjects induced profound epigenetic alterations in immune cells, which correlated with enhanced anti-mycobacterial activity in macrophages isolated from the vaccinees [4]. The changes were reflected in the DNA methylome, with the strongest response being recorded within weeks after vaccination [4]. Our observation that BCG induces alterations of the DNA methylome of immune cells has later been confirmed by others [5, 6]. Since BCG vaccination reflects an in vivo interaction between immune cells and viable mycobacteria, we here hypothesized that natural exposure to *M. tuberculosis* would induce similar changes not only in TB patients, but also in individuals who have been exposed to TB. We performed this study to evaluate if alveolar macrophages and alveolar T cells exposed to *M. tuberculosis* undergo epigenetic reprogramming, reflected as differentiated DNA methylation (DNAm) patterns in individuals exposed to the pathogen. Alveolar macrophages and alveolar T cells were procured via induced sputum, which has been verified as an equivalent method to bronchoalveolar lavage for the seclusion of the two cell types [7], and isolated in accordance with a well-established method [8, 9]. Analyses of the DNA methylomes of the isolated alveolar macrophages and alveolar T cells allowed us to identify distinct DNAm signatures in patients with TB and TB-exposed individuals.

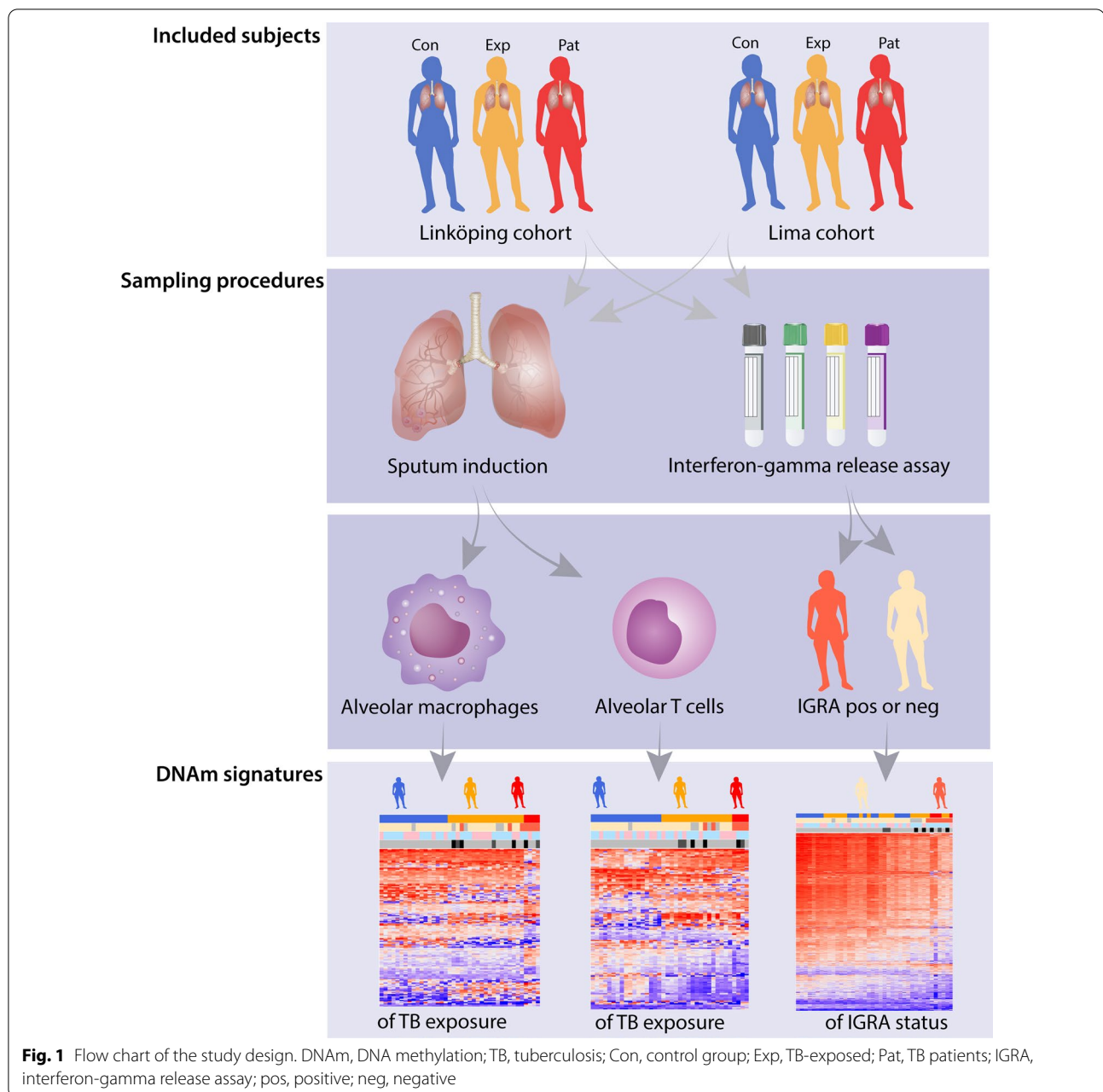
Biosignatures of TB have been proposed to lay ground for next generation of diagnostics for TB and accumulating literature have been published studying possible TB biosignatures [10, 11]. DNAm-based biosignatures exists for many other diseases e.g. inflammatory arthritis and asthma [12, 13]. Further, ethnicity and age also influence the DNA methylome [14, 15]. We included patients with tuberculosis, domestic- and occupational TB-exposed subjects and healthy controls. The first part of our pilot study was performed in Linköping, Sweden, which is a low-endemic site for TB. When we found a distinct DNAm signature of TB disease and TB-exposure we expanded the study to include TB patients and TB-exposed in Lima, Peru, a TB high-burden country. Despite the difference in TB prevalence of the two populations of Sweden and Peru it was possible to discern TB exposure based on the DNA methylomes of the subjects.

Attempts have been made to identify biomarkers of TB and TB progression, including DNAm events in peripheral immune cells of active and latent TB patients [16, 17], however no study to our knowledge has investigated DNAm in immune cells of the lung compartment of patients with TB and TB-exposed subjects, making this pilot study unique.

Results

Subject characteristics

Figure 1 represents a flow chart of the study design. Patients with TB (Pat, $n=4$), TB-exposed (Exp, $n=19$) and healthy controls (Con, $n=18$) in Sweden and in Peru (Table 1) were included. The exposed subjects from Peru all had occupational exposure to TB. The TB-exposed from Sweden were close contacts to a TB patient, enrolled via routine contact tracing at the hospital. The healthy control group of Peru and Sweden were healthy students and university employees (Universidad Peruana Cayetano Heredia and Linköping University, respectively) that weren't aware of previous exposure to TB. The two TB patients included in Lima had drug-resistant TB (MDR and XDR, respectively, (both on >3 years of treatment) and the two patients included in Linköping had drug-sensitive TB (<2 months of treatment). Two of the TB patients and all exposed subjects except one had been BCG-vaccinated (Table 1). Interferon-Gamma Release Assay (IGRA) status was determined and among the exposed individuals, three were positive and four had a borderline-positive result (T-SPOT.TB [18]). Among the controls, one subject was classified as 'borderline'-positive (QuantiFERON-TB® [19]) (Table 1). Subjects with asthma were found among the exposed and controls, but not in the patient group (Table 1). The participants donated induced sputum, from which we isolated HLA-DR-positive alveolar macrophages and CD3 positive alveolar T cells [7, 20]. DNA was prepared from the cells and DNAm data generated. Applying singular value decomposition (SVD), we found no significant differences in the DNAm data between the Pat, Exp and Con groups regarding BMI, smoking, age and sex (Additional file 2: Fig. S1). For asthma, we found significant DNAm differences (Additional file 2: Fig. S1), but we could not identify any differentially methylated CpG sites (DMCs) when comparing the subjects with asthma to those without asthma with the same significance threshold used for TB exposed versus non-exposed (Benjamini-Hochberg (BH)-corrected $p < 0.01$, data not shown). Batch effects, which co-varied with the two cohorts (= country of sampling), were corrected using a singular variance analysis algorithm (SVA) [21]. Among Exp or controls, country of sampling did not reveal any contribution to the variance.



DNA methylome data from TB-exposed individuals form a separate cluster

The 1 000 most variable positions of the DNAm data were subjected to multidimensional scaling (MDS) to investigate any inherent differences between three study groups (Fig. 2A, alveolar macrophages and 2D, alveolar T cells) using the Euclidean distances between the samples. This analysis revealed distinct differences between the study groups (Con and Exp) in the alveolar macrophages (Fig. 2A) and alveolar T cells (Fig. 2D)

on the first dimension, with the TB cases (Pat) clearly distinct from the Con and clustering closely to the Exp.

Next, we identified DMCs and differentially methylated genes (DMGs), $n = 137$ for the alveolar macrophages (Additional file 9: Table S1) and $n = 128$ for the alveolar T cells (Additional file 9: Table S2) by comparing the TB-exposed (Pat and Exp) and control groups (Con) for each cell population. To filter out the most significantly altered DMGs in the dataset, the stringency criteria of mean methylation difference (MMD) > 0.2 and p value $_{BH} < 0.05$

Table 1 Demographic characteristic data of the subjects.

Characteristics	Pat (n = 4)	Exp (n = 19)	Con (n = 18)
Country of residence			
Sweden	2 (50%)	5 (26%)	6 (33%)
Peru	2 (50%)	14 (74%)	12 (66%)
Mean age (year) ^a	27.5 ± 4.7	29.6 ± 9.5	23.5 ± 5.5
Mean height (cm) ^a	171 ± 6.2	169.2 ± 9.8	170.6 ± 7.7
Mean weight (kg) ^a	59.3 ± 3.9	70.1 ± 14.9	68.4 ± 14.4
Mean Body Mass Index (BMI) ^a	20.3 ± 0.7	24.3 ± 3.5	23.4 ± 3.8
Sex			
Male	3 (75%)	11 (58%)	10 (66%)
Female	1 (25%)	8 (42%)	8 (44%)
Smoking			
Current	1 (25%)	3 (16%)	0 (0%)
Previous	1 (25%)	2 (10%)	0 (0%)
Never	2 (50%)	14 (74%)	18 (100%)
BCG			
Vaccinated	2 (50%)	18 (95%)	15 (83%)
Non-vaccinated	2 (50%)	1 (5%)	3 (17%)
IGRA result			
Positive	2 (50%)	3 (16%)	0 (0%)
Negative	0 (0%)	12 (63%)	17 (94%)
ND	2 (50%)	0 (0%)	0 (0%)
Borderline	0 (0%)	4 (21%)	1 (6%)
Asthma			
Asthma	0 (0%)	4 (21%)	4 (22%)
No asthma	4 (100%)	15 (79%)	14 (78%)

Pat, TB patients; Exp, TB-exposed; Con, control group; BMI, body mass index; BCG, bacillus Calmette Guérin; IGRA, interferon gamma release assay; ND, no data

^a The standard deviation of the mean values is added to the age, height, weight and BMI

were applied. First, we analyzed the DMCs of the Pat vs Con groups, the Pat vs Exp groups, the Exp vs Con groups and lastly the DMCs between the Pat and Exp groups vs the Con group. For the alveolar macrophages, we found 137 DMCs and for the alveolar T cells we found 128 DMCs, when merging the results of the individual analysis (Pat-Con, Pat-Exp, Exp-Con, (Pat-Exp)-Con) (Fig. 2B, E). The results are depicted as volcano plots, which show that DNA methylomes of TB-exposed and TB patients strongly differ in the alveolar macrophages and alveolar T cells as compared to control subjects (Fig. 2B, E). The DMCs are presented in a heatmap, in which we see a clear separation between the Pat, Exp and Con groups (Fig. 2C, F). One IGRA positive subject clustered with the Con-group, identified as IGRA borderline-positive (Fig. 2A, C, D, F). For the alveolar macrophages and alveolar T cells, the patients with drug-resistant TB formed a separate group that was distinguished from

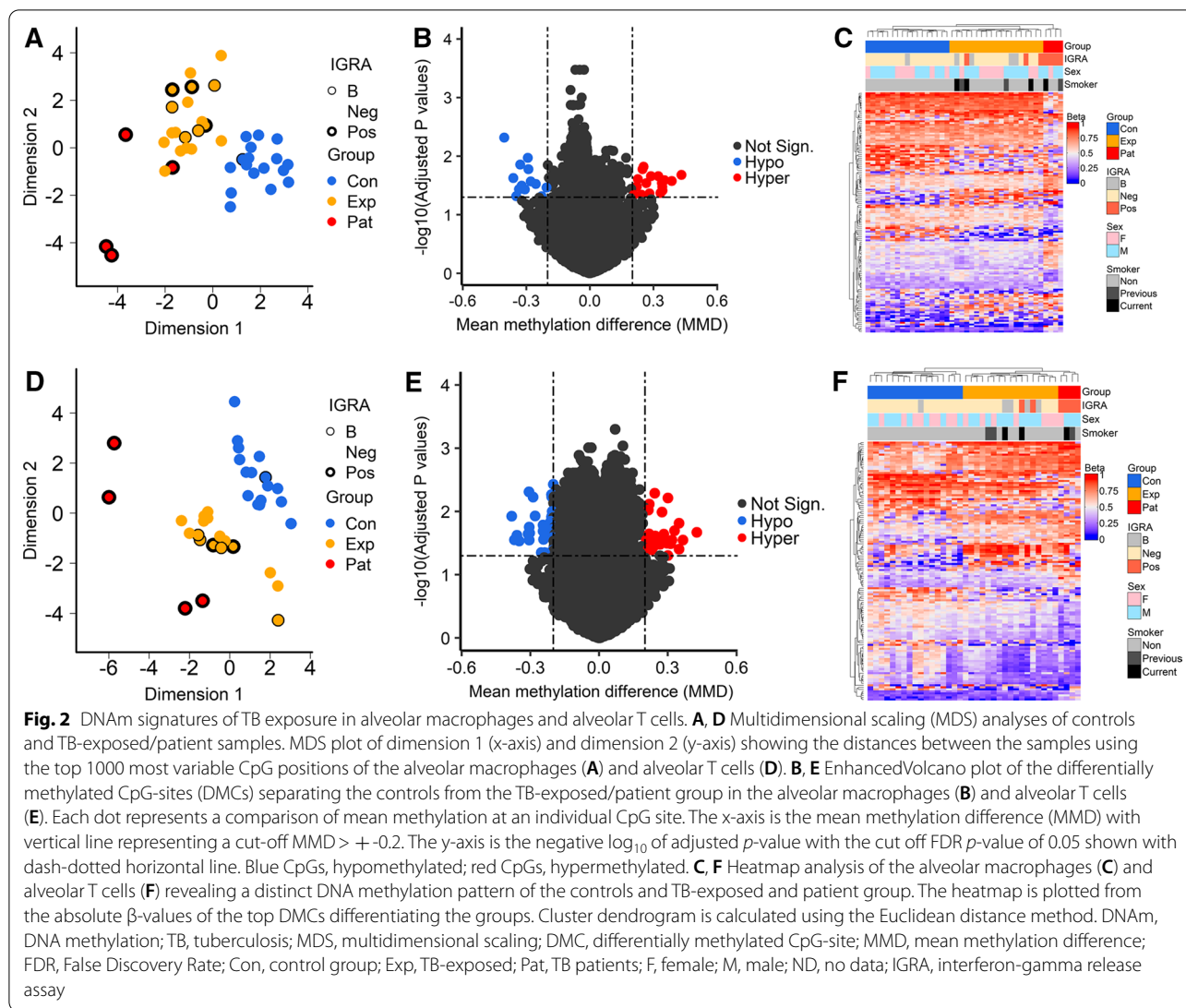
the patients with drug-sensitive TB (Additional files 3 and 4: Fig. S2A and S2B). We found 16 DMCs between the patients with drug-resistant TB and drug-sensitive TB for the alveolar T cells (Additional file 9: Table S3, $MMD > 0.2$, $p\text{-value}_{BH} < 0.1$), but no DMCs for the alveolar macrophages (data not shown).

Pathway enrichment analysis reveals common and unique interactomes in the datasets

Using the PANTHER Database, we investigated whether the identified DMGs were enriched in known pathways (Fig. 3A, B). For alveolar macrophages, the analysis revealed enrichment in pathways with relevance for TB infection, including hypoxia-inducible factor (HIF)1- α activation and the Wnt signaling pathway [22–24] (Fig. 3A). For the alveolar T cells, the overrepresentation analysis using PANTHER Database revealed B and T cell activation, PDGF signaling pathway, angiogenesis and interleukin signaling pathways as top candidate pathways (Fig. 3B). The identified DMGs from the Pat, Exp vs. Con comparison were used as seed genes in the analyses and applied to the protein–protein interaction network from STRINGdb, that resulted in module genes. We then performed a disease enrichment analysis and pathway enrichment analysis using the module genes identified by MCODE. For the enriched analysis of diseases, we found lung disease to be a top candidate (Fig. 3C). Enriched pathways based on module genes for the alveolar macrophages and alveolar T cells included PI3K-Akt, Ras and Chemokine (Fig. 3D).

Comparisons across cell populations reveals the existence of a common DNA methylome-based biosignature in mycobacteria-exposed immune cells

Given the fact that the interaction between mycobacteria and eukaryotes is evolutionary ancient, we predicted that highly conserved pathways exist that are common among the studied cell populations. Comparing the identified DMGs from the alveolar macrophages and alveolar T cells in a Venn analysis, we discovered 8 common DMGs ($MMD > 0.2$) (Additional file 5: Fig. S3) and 144 DMGs with reduced stringency ($MMD > 0.1$) (Fig. 4A). Among the 144 common DMGs between both cell types in this study, separating the Pat, Exp and Con groups we found FBRSL1 (Additional file 9: Table S4). For common DMGs between the alveolar macrophages and alveolar T cells the overrepresentation analysis revealed hypoxia response via HIF-activation, the Ras pathway and VEGF signaling pathway as important pathways (Fig. 4B). We expanded the Venn analyses to include data from our previous work on BCG vaccine-induced DMGs that correlated with enhanced mycobacterial control [4], as natural exposure to TB and BCG vaccination both

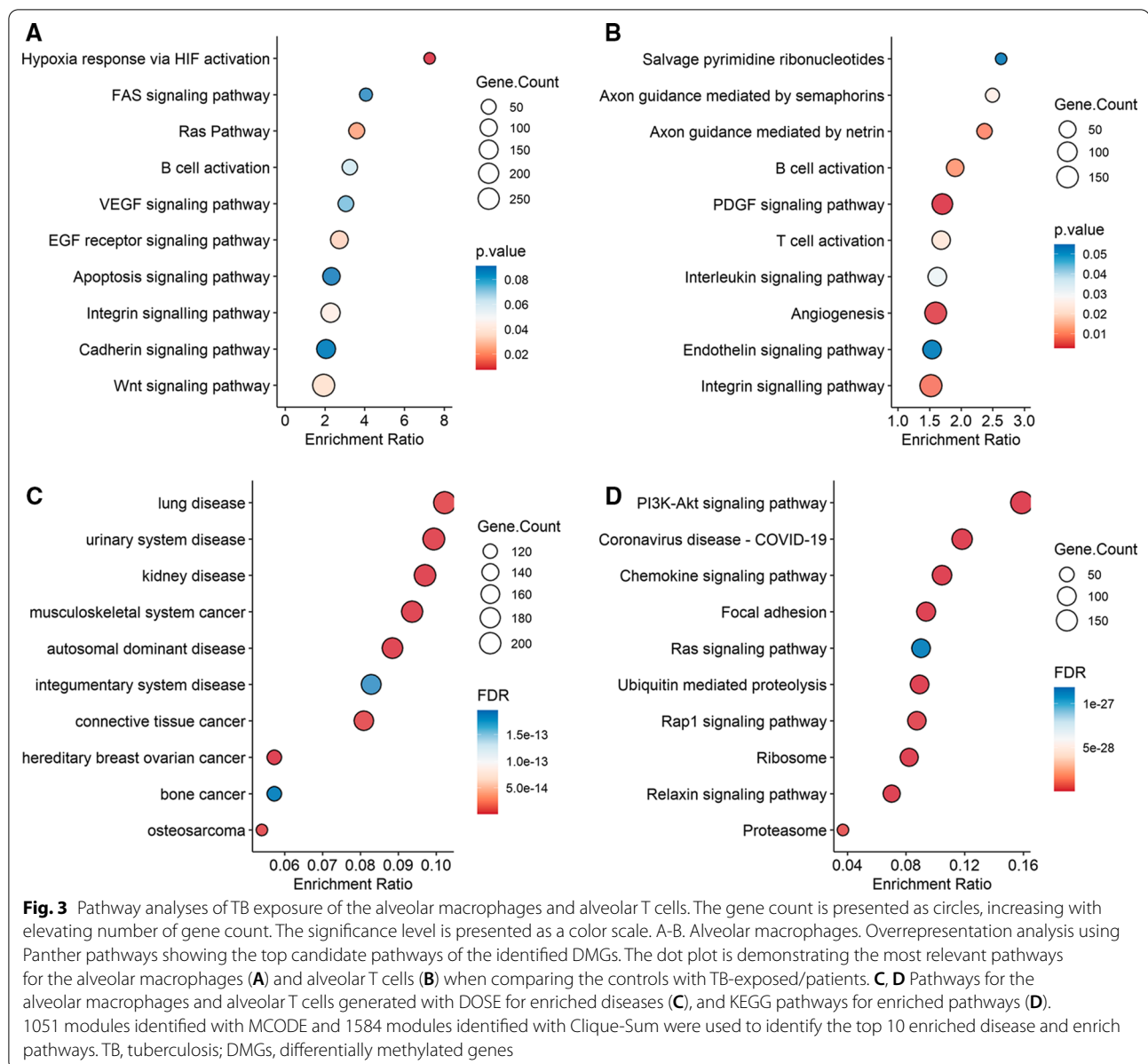


represent in vivo encounters between mycobacteria and host immune cells. Even though the routes of mycobacterial exposure differ profoundly in these settings, a set of 19 DMGs could be identified as overlapping between our previous BCG study (studying macrophages isolated from peripheral blood) and all cell populations studied here (Additional file 6: Fig. S4), suggesting that a highly conserved epigenetic response to mycobacterial challenge exists.

Differential DNA methylation analysis identified DMCs to distinguish IGRA positive from IGRA negative subjects

To determine whether IGRA status can be distinguished based on the methylation pattern, we compared the alveolar macrophage and alveolar T cell

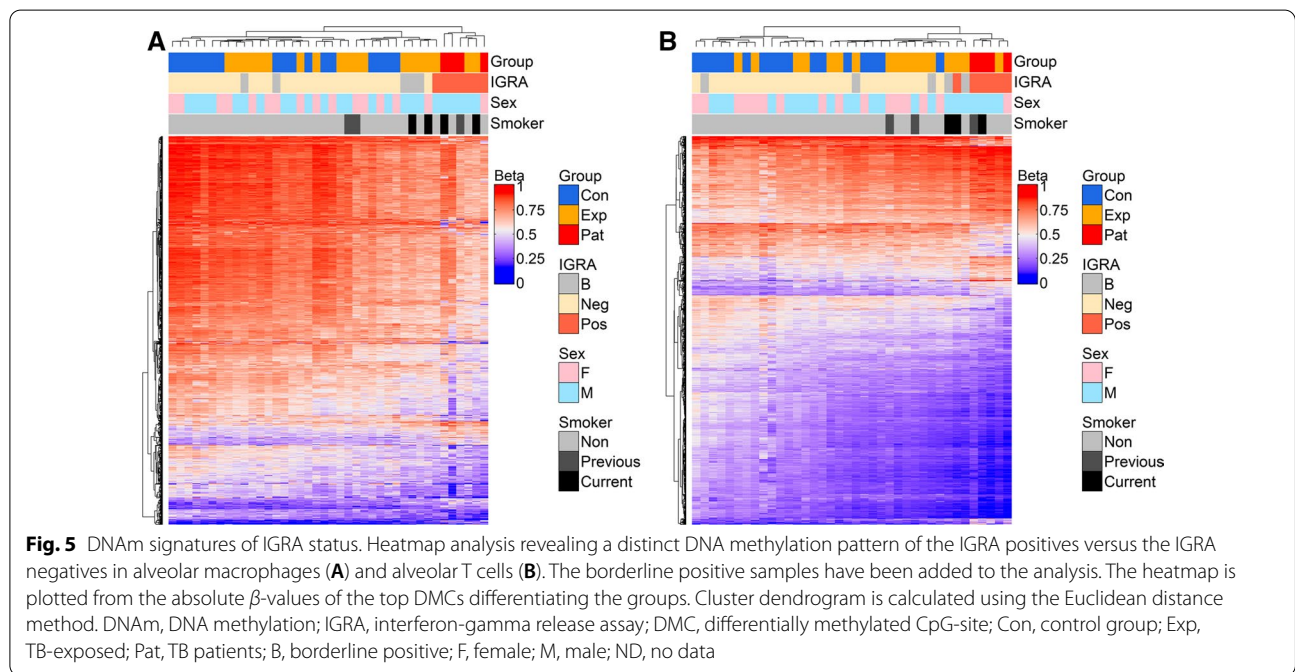
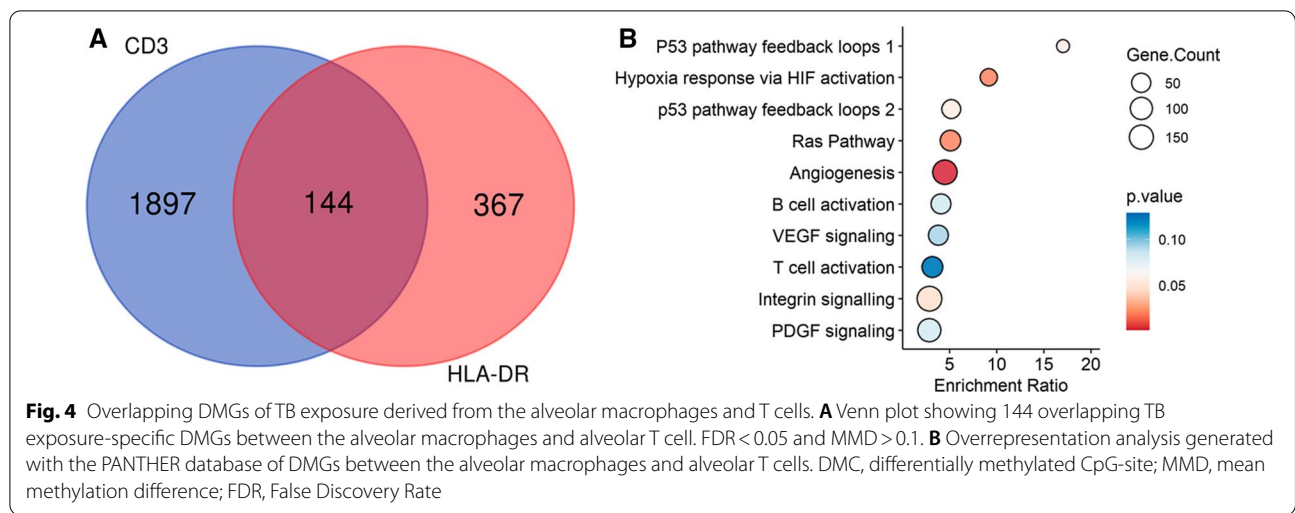
DNA methylome data obtained from the IGRA positive and negative subjects. With the p -value_{BH} < 0.05 and $MMD > 0.1$, the analysis revealed 1315 DMCs and 785 DMGs (Fig. 5A; Additional file 9: Table S5) for the alveolar macrophages and 1491 DMCs and 855 DMGs for the alveolar T cells (Fig. 5B; Additional file 9: Table S6). After demonstrating that the IGRA positive and negative groups had distinct DNA methylation traits, we performed a heatmap analysis based on the top DMCs of the DNA methylation analysis that showed that the IGRA positive group clustered into one homogenous group for both the alveolar macrophages and alveolar T cells (Fig. 5A, B). The pathway analyses of the IGRA related genes for the alveolar macrophages and alveolar T cells are shown in Additional files 7 and 8: Figure S5A and S5B.



Discussion

The epigenetic response to infection is poorly studied in humans, and the few studies that exist focus on individuals who have ongoing disease symptoms and conclude that the observed alterations in epigenetic patterns are a pathological event [25]. However, in recent years, several studies performed in fish points toward a contribution of altered DNA methylation patterns to resistance against infections [26, 27]. We previously demonstrated that a distinct DNA methylation pattern in immune cells of BCG-vaccinated subjects correlated with increased capacity of monocyte-derived macrophages to kill *M. tuberculosis* [4]. We have also described that subjects,

who IGRA converted during the course of the study, displayed a differential DNA methylation pattern before IGRA conversion [28]. In this study, we present data suggesting that exposure to TB generates a distinct DNAm signature in pulmonary immune cells. The signature was found not only in those with TB infection (TB patients or positive IGRA), but also in individuals, who are exposed to TB but IGRA negative. The finding that healthy, TB-exposed individuals also carry the signature opens the possibility that the epigenetic alterations reflect a host-beneficial reprogramming of the immune mechanisms rather than being induced by *M. tuberculosis* as a step to evade the immune defense. The way TB disease is viewed



has shifted from a binary perspective (latent or active TB) to be considered a spectrum of disease ranging from early clearance, latent infection to various presentations of active TB [29–32]. 32% of the TB-exposed subjects in this study and 11% of the controls had a positive IGRA result, corresponding to the declared exposure to TB. One theory why one healthy control with a borderline-positive result, indicating host immune responses against *M. tuberculosis*, aligned with the rest of the healthy controls, is that the subject was exposed to TB a long time ago, while the IGRA positive TB-exposed subjects had a more recent exposure. A diagnostic tool to specify who will

clear the bacteria or develop active TB upon TB exposure and to define the risk of LTBI conversion to active disease would aid in contact tracing and treatment decision making [33]. Attempts have previously been made to identify patients with LTBI with a high risk of developing active TB using RNA-sequencing, but with lacking confirmatory longitudinal data [34]. The DNAm signatures of TB and TB exposure found in this study could be used to determine who has recently been exposed to TB in a contact tracing context. Further studies need to confirm the DNAm patterns in a larger cohort and investigate longitudinally whether the signature seen in the

TB-exposed subjects reflects a beneficial response or risk of becoming ill with TB. The TB patients formed a separate group based on their DNAm pattern, but some differences could be found between the Lima and Linköping cohort (drug-resistant TB and drug-sensitive TB, respectively), suggesting that the type of TB-treatment and different treatment responses are also reflected in the DNA methylome. In line with the fact that macrophages constitute the main niche for mycobacterial replication, the strongest enrichment of DNAm changes was observed in the alveolar macrophage population. The pathways identified to be enriched in the alveolar macrophage population for both TB exposure and IGRA status have been described in the context of trained immunity, BCG exposure and TB. For example, activation of Hypoxia-Inducible Factor 1 α and glycolysis pathways are hallmarks of macrophages that have undergone the epigenetic changes reflective of trained immunity [35], which is induced in myeloid cells upon BCG-stimulation [36, 37]. VEGF-release by macrophages has been shown to recruit immune cells during granuloma formation [38]. Several studies have ascribed Wnt pathways immunomodulating functions and induction during *M. tuberculosis* infection (reviewed in [39]). The Ras pathway appeared in the pathway analysis of the alveolar macrophages, alveolar T cells and the pathway analysis of the common DMGs between the cell types. In a recent study where we found that a distinct DNA methylome signature characterized IGRA negative subjects, who later presented with a positive IGRA result [28], we showed that the Ras pathway is of importance in IGRA conversion [28]. Also, the PI3K-Akt pathways appeared for both the alveolar T cells and the common pathway analysis. The top pathway for the alveolar macrophages using MCODE was lung disease, indicating that genes of importance for lung inflammation are affected upon TB exposure. In the Venn diagram, when comparing the DMGs from this study with DMGs from our previous publication [4] we found 17 common genes for the alveolar macrophages, and 266 common genes for the alveolar T cells. FBRSL1 was found as a common DMG between alveolar macrophages and alveolar T cells. FBRSL1 has been shown to be an important gene in a recent published study on TB related DNA methylation biomarkers in peripheral blood [40]. BCG vaccination has convincingly been shown to induce heterologous immunity protecting against childhood mortality from other causes than TB [41, 42]. Based on our finding that natural TB exposure and BCG vaccination trigger similar epigenetic changes, and that the involved pathways overlap with those found in trained immunity, we propose the hypothesis that a “beneficial exposure” to TB exists, which protects against other infections through heterologous immunity. Along the same line, it

has been shown that a substantial fraction of individuals exposed to TB can be defined as ‘early clearers,’ since they remain tuberculin skin test or IGRA negative, suggesting effective eradication of the infection [43]. A limitation of the present study is that no data on bacteria clearance or antimycobacterial efficacy of the cells is available, therefore no conclusion can be drawn on the correlation between the epigenetic changes and their effect.

Conclusions

Our study suggests that DNA methylome analysis is a promising approach to study the spectrum of TB. Our results affirm previous work by us and others on mycobacteria-induced epigenetic changes in immune cells. Longitudinal studies of patients with TB during treatment are needed that cover DNA methylation changes of alveolar macrophages and alveolar T cells before the time point of IGRA conversion over the development of active TB and during TB treatment until cure. The results will allow dissection of the unique DNAm signatures for each state of TB and could be further developed as precision tools for clinical TB diagnosis, contact-tracing, decision-making in LTBI infected subjects and TB treatment monitoring.

Methods

Study design and participants

Patients with pulmonary TB, healthy participants with a history of TB-exposure and healthy controls, with an age ranging from 18 to 53 years, were enrolled at Linköping University Hospital and Linköping University, respectively. To determine epigenetic changes in the immune cells in TB-exposed individuals, we recruited subjects enrolled in a TB contact tracing at Linköping University Hospital, Sweden, according to standard contact tracing routines (household contacts and persons with >8 h of interaction with the index case). We included two index cases and five contact tracing subjects who were ethnicity-matched (Table 1). The patients were diagnosed with drug-sensitive pulmonary TB. In a secondary inclusion of subjects we directed our focus towards Peru, which has a higher prevalence of TB than Sweden. Peruvian health care workers and medical students have an increased risk to acquire a TB infection [44, 45]. Because we were interested in studying DNAm changes upon TB exposure, subjects were recruited from the Hospital Nacional Cayetano Heredia in Lima, Peru, and controls were recruited from Universidad Peruana Cayetano Heredia, also in Lima. Following oral and written consent, two patients with drug-resistant TB, 14 occupational-exposed subjects and 12 controls (18–40 years of age) were included in the study. We obtained induced sputum from each

participant and used an established protocol for the separate isolation of alveolar macrophages and alveolar T cells.

Sputum induction and pulmonary immune cell isolation

Induced sputum is a well-tolerated, non-invasive method to collect cells from the respiratory compartment after inhalation of a hypertonic saline solution. The procedure of sputum induction takes approximately 30 min and is both cost effective and safe with minimal clinical risks [46]. Sputum specimens were collected as described by Alexis et al. [47], with the following modification: pre-medication with an adrenergic β 2-agonist, salbutamol (Ventoline, 1 ml, 1 mg/ml) was administered before the inhalation of hypertonic saline, using a nebulizer (eFlow, PARI), in Sweden but not in Peru, following local protocols. The subsequent steps of sputum processing were adopted from Alexis et al. [48] and Sikkeland et al. [20]. The alveolar macrophages and alveolar T cells were isolated using superparamagnetic beads coupled with anti-human CD3 and Pan Mouse IgG antibodies and HLA-DR/human MHC class II antibodies (Invitrogen Dynabeads, ThermoFisher, cat no. 11041 and 14-9956-82, respectively). An initial positive selection was done with CD3 beads followed by a positive HLA-DR selection. Bead-coating and cell isolation were performed according to manufacturer's protocol.

DNAm sequencing and statistical analyses

DNA from the alveolar macrophages and alveolar T cells were extracted using the AllPrep DNA/RNA Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. Genome-wide DNA methylation analysis of the DNA from alveolar macrophages and alveolar T cells was performed using the Infinium Human Methylation 450 K Bead Chip array (Illumina) as per the manufacturer's instructions. The technique is based on quantitative genotyping of C/T polymorphism which is generated by DNA bisulfite conversion and allows for the assessment of over 450 000 methylation sites within the whole genome. DNA was converted and amplified and subsequently fragmented and hybridized to the Infinium HumanMethylation450 Bead Chip. The procedure was conducted at the Bioinformatics and Expression analysis Core facility at the Karolinska Institute, Sweden. The methylation profiles from the HumanMethylation450K BeadChip analysis for each cell type were analyzed from the raw IDAT files in R (v4.0.2) using the *minfi* (v1.36.0) with subset-quantile within array (SWAN) normalization [49, 50]. *ChAMP* [51] (v2.19.3). The data were pre-processed in several steps, following probes were filtered out: (1) probes below the detection p -value (>0.01), (2) non-CpG probes, (3) multi-hit probes, and (4) all probes

of X and Y chromosomes. Batch effects were corrected using *ComBat* from the *SVA* package [21] (v3.38.0) in *ChAMP* (v2.19.3). Identified sources of variation that were still present upon SVD correction were included as co-variables in the models. Differential methylation analysis was performed with the linear modeling (lmFit) using the *limma* package [52] (v3.46.0) in a contrast matrix of the TB-exposed and TB-non-exposed (Control) individuals. All differentially methylated CpGs (DMCs) were considered significant at the corrected p -value_{BH} <0.05 and mean methylation difference (MMD) >0.2 . The DMCs were mapped to their corresponding genes as differentially methylated genes (DMGs) using the hg19 annotations.

Unsupervised cluster analysis

Hierarchical clustering of the all TB-exposed and control individuals was performed with the normalized, batch corrected β values obtained after the data filtration in each cell type individually. The distance was calculated using the Euclidean distance matrix. The *dendextend* [52] (v1.14.0) and *ape* [53] (v5.4-1) packages in R were used to construct the horizontal hierarchical plots from the three different cell populations using the *hclust* and *dendrogram* functions.

Structural annotations

The *EnhancedVolcano* package [54] (v1.8.0) was used to generate the individual volcano plots from all cell populations. The heatmaps were generated from the filtered DMGs with their respective CpGs for each cell type using the *ComplexHeatmap* (v2.6.2) package [55]. The clustering dendrogram in heatmaps were plotted using the Euclidean distance matrix. The DMCs identified in all four comparisons (Exp vs Con, Pat vs Con, Pat vs Exp, Exp + Pat vs Con), were merged/combined to generate a single heatmap. A mean methylation difference (MMD) of >0.2 was used for the groupwise analyses. MMD >0.1 was used for the comparison of IGRA positives vs negatives. For the IGRA status analysis, the TB patients were regarded as IGRA positives and included in the IGRA positive group.

Pathway overrepresentation and enrichment analyses

The identified DMGs (relaxed cutoffs of MMD >0.1) in all the above comparisons were used for PANTHER pathway over-representation analyses using the WEB-based Gene SeT Analysis Toolkit (WebGestalt) webserver (v2022). The FDR in the pathway analysis is BH adjusted p -values. We used nominal p -values (significance level set to $p < 0.05$) in case FDR correction was too stringent. Top ten significantly enriched pathways were displayed in

dot plots generated in R using ggplot2 package (version 3.3.3).

Furthermore, the graph clustering algorithm MCODE and Clique Sum (Clique Susceptibility Module algorithm) in Modifier package was used to identify molecular complexes and create a large disease modules. The identified DMGs were used as seed genes and applied to the protein–protein interaction network from STRINGdb (confidence score > 0.7), that resulted in module genes. For the module genes, we used package DOSE (Disease Ontology Semantic and Enrichment analysis) to identify the enriched diseases (Fig. 3C) and using *clusterProfiler* [56] (v3.18.1), we performed KEGG pathway [57] enrichment analysis (Fig. 3D). To enhance the visualization and better understanding of the enrichment result, ggplot2 package was used to generate the dot plots.

Venn and overlap analyses

Venn analyses were performed to detect the DMGs overlapping between cell populations and a previous publication. We constructed the Venn diagrams by using *ugent* Venn diagram generator (<http://bioinformatics.psb.ugent.be/webtools/Venn/>) and in house R script to compare the DMGs.

Statistical analyses

All differences with a p -value_{BH} < 0.05 were considered significant if not otherwise stated. We calculated family-wise error rate (FWER) using the BH correction method. All analyses were performed in R (v4.0.2) with the mentioned packages.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13148-022-01390-9>.

Additional file 1. Figure legends Supplementary figures.

Additional file 2. Supplemental figure S1. Singular Value Decomposition (SVD) analysis of subject characteristics. We found no significant difference between the Pat, Exp and Con groups regarding BMI, smoking, age and sex. For asthma, IGRA and batch (also reflecting country; Peru or Sweden) we found significant differences in the DNA methylomes. Pat, TB patients; Exp, TB-exposed; Con, control group; BMI, body mass index; IGRA, interferon-gamma release assay.

Additional file 3. Supplemental figure S2. MDS plot showing the origin of the sample (Lima or Linköping). Alveolar macrophage (A) and alveolar T cell (B) samples from the Lima cohort are presented as dots and samples from the Linköping cohort are presented as triangles. The patients are shown in red. The Lima patients (n=2) have drug-resistant TB and the Linköping patients (n=2) have drug-sensitive TB. MDS, multidimensional scaling.

Additional file 4. Supplemental figure S2. MDS plot showing the origin of the sample (Lima or Linköping). Alveolar macrophage (A) and alveolar T cell (B) samples from the Lima cohort are presented as dots and samples from the Linköping cohort are presented as triangles. The patients are shown in red. The Lima patients (n=2) have drug-resistant TB and the

Linköping patients (n=2) have drug-sensitive TB. MDS, multidimensional scaling.

Additional file 5. Supplemental figure S3. Overlapping DMGs derived from the alveolar macrophages and T cells. FDR < 0.05 and MMD > 0.2. DMGs, differentially methylated genes; FDR, False Discovery Rate; MMD, mean methylation difference.

Additional file 6. Supplemental figure S4. Overlapping genes with our previous study*. We found 19 common genes for both alveolar macrophages and lymphocytes. *D. Verma et al, 2017.

Additional file 7. Supplemental figure S5. Pathway analysis of IGRA status based on DMGs found between the IGRA positive and IGRA negative participants. A-B. Pathways of IGRA status in alveolar macrophages (A), based on 785 DMGs, and in alveolar T cells (B), based on 855 DMGs. IGRA, interferon-gamma release assay; DMGs, differentially methylated genes.

Additional file 8. Supplemental figure S5. Pathway analysis of IGRA status based on DMGs found between the IGRA positive and IGRA negative participants. A-B. Pathways of IGRA status in alveolar macrophages (A), based on 785 DMGs, and in alveolar T cells (B), based on 855 DMGs. IGRA, interferon-gamma release assay; DMGs, differentially methylated genes.

Additional file 9. Supplemental table S1. Differentially methylated CpG sites (DMCs) from analyses in alveolar macrophages for TB status. Merged list of DMCs from Pat, Exp and Con contrasts using absolute MMD > 0.2 and Adj.p-val < 0.05. **Supplemental table S2.** Differentially methylated CpG sites (DMCs) from analyses in alveolar T cells for TB status. Merged list of DMCs from Pat, Exp and Con contrasts using absolute MMD > 0.2 and Adj.p-val < 0.05. **Supplemental table S3.** Differentially methylated CpG sites (DMCs) identified between Linköping and Lima patient samples (n=2+2) in Alveolar T cells with Adj.p-val < 0.1. (No DMCs identified in HLA-DR). **Supplemental table S4.** Common Differentially methylated Genes (DMGs) from analyses in alveolar macrophages and T cells for TB status. Using relaxed cutoffs MMD > 0.1 and Adj.p-val < 0.05. **Supplemental table S5.** Differentially methylated CpG sites (DMCs) identified in alveolar macrophages for IGRA status with MMD > 0.1, Adj.p-val < 0.05. **Supplemental table S6.** Differentially methylated CpG sites (DMC) identified in alveolar T cells for IGRA status with MMD > 0.1, Adj.p-val < 0.05.

Acknowledgements

We direct our gratitude to the staff at Linköping University Hospital and the Vrinnevi Hospital in Norrköping for assistance in sample collection and all the subjects for donating samples. We especially thank Hilma Hård af Segerstad, Helena Rylander, Elsa Reuterswård and Emma Marttala for establishing the study logistics and starting up inclusions. We are also grateful for the valuable help provided by Ronald Cadillo Hernandez, Rodrigo Cachay Figueroa and the medical personnel at the Instituto de Medicina Tropical Alexander von Humboldt, Universidad Peruana Cayetano Heredia. We would like to acknowledge the Bioinformatics Core Facility, Faculty of Medicine and Health Sciences and Clinical Genomics Linköping, Science for Life Laboratory, Department of Biomedical and Clinical Sciences, Linköping University, for assistance with bioinformatics analyses.

Author contributions

ML, NI, CUG, MM-A and JP designed and conceptualized the study. IP and NI performed the laboratory work and the related analyses. SS and ML designed and performed the bioinformatic analyses of the data. SS wrote the scripts for analysis. SS and IP created figures. IP, NI, JD, SS, and ML wrote the manuscript. JD and ML are co-authors of a patent application "Biomarker for detection of mycobacterial infection and exposure" filed on February 2nd with the Swedish Patent Registry (100692). All authors read and approved the final manuscript.

Funding

Open access funding provided by Linköping University. This study was funded through generous grants from Forskningsrådet Sydöstra Sverige (FORSS-932096), the Swedish Research Council (2015-02593, 2018-02961 and 2018-04246), the Swedish Heart Lung Foundation (20150709 and 20180613) and CONCYTEC—PROCIENCIA (106-2018-FONDECYT). The DNA methylome data were generated at the Bioinformatics and Expression Analysis Core Facility at the Department of Biosciences and Nutrition, which is supported by the

Board of Research at the Karolinska Institute, Stockholm. The computations were enabled by resources provided by the Swedish National Infrastructure for Computing (SNIC) at Linköping University campus partially funded by the Swedish Research Council through grant agreement (2018-05973). We further want to acknowledge the support with a travel grant from the World Infection Fund.

Availability of data and materials

The datasets generated and/or analysed during the current study are not publicly available due to intellectual property restrictions but are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

Included subjects donated peripheral blood and induced sputum samples following oral and written informed consent. Ethical approval for Sweden was obtained from the Regional Ethics Review Board in Linköping, (2016/237-31) and for Peru from the Institutional Committee on Research Ethics of the Cayetano Heredia National Hospital for the inclusion of patients (082-018) and by Universidad Peruana Cayetano Heredia Institutional Review Board for the inclusion of healthy subjects (101773). The study protocol included questionnaires on respiratory and overall health, the evaluation of IGRAs status and sputum samples for DNA extraction. The subjects' samples and questionnaires were not linked to any personal information at any stage of the study. The sputum sample collections were performed in accordance with guidelines from the Department of Respiratory Medicine at Linköping University Hospital. In Sweden, the IGRAs samples (QuantiferON-TB Gold) were collected and analysed by medical personnel according to the guidelines at the Department of Clinical Microbiology at Linköping University Hospital. In Peru, the IGRAs samples (T-SPOT.TB) was taken and analysed by medical students under supervision.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Division of Inflammation and Infection, Department of Biomedical and Clinical Sciences, Faculty of Medicine and Health Sciences, Linköping University, Lab 1, Floor 12, 58185 Linköping, Sweden. ²Bioinformatics Unit (Core Facility), Linköping University, Linköping, Sweden. ³Clinical Genomics Linköping, SciLife Laboratory, Department of Biomedical and Clinical Sciences, Linköping University, Linköping, Sweden. ⁴Division of Infectious Diseases, Department of Biomedical and Clinical Sciences, Faculty of Medicine and Health Sciences, Linköping University, Linköping, Sweden. ⁵Laboratorio de Investigación en Enfermedades Infecciosas, Facultad de Ciencias Y Filosofía, Universidad Peruana Cayetano Heredia, Lima, Peru. ⁶School of Medicine, Universidad Peruana Cayetano Heredia, Lima, Peru. ⁷Instituto de Medicina Tropical Alexander Von Humboldt, Universidad Peruana Cayetano Heredia, Lima, Peru.

Received: 15 June 2022 Accepted: 29 November 2022

Published online: 17 December 2022

References

- Global Tuberculosis Report 2021. World Heal. Organ. Geneva; 2021.
- Fitchett JR, MacPherson P, Corbett EL. Implementing the end TB strategy and the intersection with the sustainable development goals, 2016–2030. *Trans R Soc Trop Med Hyg*. 2015.
- Donald PR, Diacon AH, Lange C, Demers AM, Von Groote-Biddlingmeier F, Nardell E. Droplets, dust and Guinea pigs: an historical review of tuberculosis transmission research, 1878–1940. *Int J Tuberc Lung Dis*. 2018.
- Verma D, Parasa VR, Raffetseder J, Martis M, Mehta RB, Netea M, et al. Antimycobacterial activity correlates with altered DNA methylation pattern in immune cells from BCG-vaccinated subjects. *Sci Rep*. 2017;7:1–10.
- Hasso-Agopsowicz M, Scriba TJ, Hanekom WA, Dockrell HM, Smith SG. Differential DNA methylation of potassium channel KCa3.1 and immune signalling pathways is associated with infant immune responses following BCG vaccination. *Sci Rep*. 2018;
- Kaufmann E, Sanz J, Dunn JL, Khan N, Mendonça LE, Pacis A, et al. BCG Educates Hematopoietic Stem Cells to Generate Protective Innate Immunity against Tuberculosis. *Cell Cell Press*. 2018;172:176–190.e19.
- Das J, Idh N, Sikkeland LIB, Paues J, Lerm M. DNA methylome-based validation of induced sputum as an effective protocol to study lung immunity: construction of a classifier of pulmonary cell types. *Epigenetics*. 2021;
- Sikkeland LIB, Haug T, Stangeland AM, Flatberg G, Sørstrand P, Halvorsen B, et al. Airway inflammation in paper mill workers. *J Occup Environ Med*. 2007;49.
- Sikkeland LIB, Dahl CP, Ueland T, Andreassen AK, Gude E, Edvardsen T, et al. Increased levels of inflammatory cytokines and endothelin-1 in alveolar macrophages from patients with chronic heart failure. *PLoS One*. 2012;7.
- Lerm M, Dockrell HM. Addressing diversity in tuberculosis using multidimensional approaches. *J Intern Med*. 2018.
- Maertzdorf J, Kaufmann SHE, Weiner J. Toward a unified biosignature for tuberculosis. *Cold Spring Harb Perspect Med*. 2015;5.
- Laprise C. Cell-Specific DNA Methylation Signatures in Asthma. 2019;
- Hammaker D, Firestein GS, Jolla L. Epigenetics of inflammatory arthritis. 2019;30:188–96.
- Kader F, Ghai M. DNA methylation-based variation between human populations. *Mol Genet Genomics*. Springer Berlin Heidelberg; 2017;292:5–35.
- Bell CG, Lowe R, Adams PD, Baccarelli AA, Beck S, Bell JT. DNA methylation aging clocks : challenges and recommendations. *Genome Biology*; 2019;1–24.
- Esterhuysen MM, Weiner J, Caron E, Loxton AG, Iannaccone M, Wagman C, et al. Epigenetics and proteomics join transcriptomics in the quest for tuberculosis biomarkers. *MBio*. 2015;6.
- Sharma G, Sowpati DT, Singh P, Khan MZ, Ganji R, Upadhyay S, et al. Genome-wide non-CpG methylation of the host genome during M. tuberculosis infection. *Sci Rep*. 2016;6.
- Rego K, Pereira K, MacDougall J, Cruikshank W. Utility of the T-SPOT®.TB test's borderline category to increase test resolution for results around the cut-off point. *Tuberculosis [Internet]*. Elsevier; 2018;108:178–85. Available from: <https://doi.org/10.1016/j.tube.2017.12.005>
- Jonsson J, Westman A, Bruchfeld J, Sturegård E, Gaines H, Schön T. A borderline range for Quantiferon Gold In-Tube results. *PLoS One*. 2017;
- Sikkeland LIB, Kongerud J, Stangeland AM, Haug T, Alexis NE. Macrophage enrichment from induced sputum [3]. *Thorax*. 2007.
- Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The SVA package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics*. 2012;28:882–3.
- Li Q, Xie Y, Cui Z, Huang H, Yang C, Yuan B, et al. Activation of hypoxia-inducible factor 1 (Hif-1) enhanced bactericidal effects of macrophages to Mycobacterium tuberculosis. *Tuberculosis [Internet]*. Elsevier Ltd; 2021;126:102044. Available from: <https://doi.org/10.1016/j.tube.2020.102044>
- Villaseñor T, Madrid-paulino E, Maldonado-bravo R, Urbán-aragón A, Pérez-martínez L, Pedraza-alva G. Activation of the wnt Pathway by Mycobacterium tuberculosis : A wnt – wnt Situation. 2017;8:1–16.
- Shi L, Eugenin EA, Subbian S. Immunometabolism in Tuberculosis. 2016;7:1–15.
- Zhang D, Guo S, Schrodi SJ. Mechanisms of DNA Methylation in Virus-Host Interaction in Hepatitis B Infection: Pathogenesis and Oncogenetic Properties. *Int J Mol Sci* 2021, Vol 22, Page 9858 [Internet]. Multidisciplinary Digital Publishing Institute; 2021 [cited 2021 Dec 1];22:9858. Available from: <https://www.mdpi.com/1422-0067/22/18/9858/html>
- Marjoram L, Alvers A, Deerhake ME, Bagwell J, Mankiewicz J, Cocchiari JL, et al. Epigenetic control of intestinal barrier function and inflammation in zebrafish. *Proc Natl Acad Sci U S A [Internet]*. National Academy of Sciences; 2015 [cited 2021 Dec 1];112:2770–5. Available from: <https://www.pnas.org/content/112/9/2770>
- Xiu Y, Shao C, Zhu Y, Li Y, Gan T, Xu W, et al. Differences in DNA Methylation Between Disease-Resistant and Disease-Susceptible Chinese Tongue Sole (*Cynoglossus semilaevis*) Families. *Front Genet*. *Frontiers Media S.A.*; 2019;10:847.

28. Karlsson L, Das J, Nilsson M, Tyrén A, Pehrson I, Idh N, et al. A differential DNA methylome signature of pulmonary immune cells from individuals converting to latent tuberculosis infection. *Sci Reports* 2021 11:1 [Internet]. Nature Publishing Group; 2021 [cited 2021 Sep 30];11:1–13. Available from: <https://www.nature.com/articles/s41598-021-98542-3>
29. Migliori GB, Ong CWM, Centis R, Goletti D, Petrone L, Ambrosio LD. The definition of tuberculosis infection based on the spectrum of tuberculosis disease. 2021;17:1–12.
30. Lewinsohn DM, Meermeier EW. Early clearance versus control: What is the meaning of a negative tuberculin skin test or interferon-gamma release assay following exposure to *Mycobacterium tuberculosis*? [version 1; referees: 2 approved]. *F1000Research*. 2018.
31. Drain PK, Bajema KL, Dowdy D, Dheda K, Naidoo K, Schumacher SG, et al. Incipient and subclinical tuberculosis: A clinical review of early stages and progression of infection. *Clin. Microbiol. Rev.* 2018.
32. Lin PL, Flynn JL. The End of the Binary Era: Revisiting the Spectrum of Tuberculosis. *J Immunol.* 2018;201:2541–8.
33. Pai M. Spectrum of latent tuberculosis existing tests cannot resolve the underlying phenotypes. *Nat Rev Microbiol.* Nature Publishing Group; 2010;8:242.
34. Estévez O, Anibarro L, Garet E, Pallares Á, Barcia L, Calviño L, et al. An RNA-seq Based Machine Learning Approach Identifies Latent Tuberculosis Patients With an Active Tuberculosis Profile. *Front Immunol.* 2020;11.
35. Cheng S, Quintin J, Cramer RA, Shepardson KM, Saeed S, Kumar V, et al. mTOR- and HIF-1 α -mediated aerobic glycolysis as metabolic basis for trained immunity. 2014;345.
36. Benn CS, Fisker AB, Whittle HC, Aaby P. Revaccination with Live Attenuated Vaccines Confer Additional Beneficial Nonspecific Effects on Overall Survival: A Review. *EBioMedicine.* 2016;
37. Thysen SM, Benn CS, Gomes VF, Rudolf F, Wejse C, Roth A, et al. Neonatal BCG vaccination and child survival in TB-exposed and TB-unexposed children: A prospective cohort study. *BMJ Open.* 2020;
38. Verrall AJ, Alisjahbana B, Apriani L, Novianty N, Nurani AC, Van Laarhoven A, et al. Early clearance of mycobacterium tuberculosis: The INFECT case contact cohort study in indonesia. *J Infect Dis.* 2020;221.
39. Brandenburg J, Reiling N. The Wnt blows: On the functional role of Wnt signaling in mycobacterium tuberculosis infection and beyond. *Front. Immunol.* 2016.
40. Lyu M, Zhou J, Jiao L, Wang Y, Zhou Y, Lai H, et al. Deciphering a TB-related DNA methylation biomarker and constructing a TB diagnostic classifier. *Mol Ther - Nucleic Acids* [Internet]. The Authors; 2022;27:37–49. Available from: <https://doi.org/10.1016/j.omtn.2021.11.014>
41. Mourits VP, Koeken VACM, De Bree LCJ, Moorlag SJCFM, Chu WC, Xu X, et al. BCG-Induced Trained Immunity in Healthy Individuals: The Effect of Plasma Muramyl Dipeptide Concentrations. *J Immunol Res.* 2020;2020.
42. Kleinnijenhuis J, Quintin J, Preijers F, Joosten LAB, Iffrim DC, Saeed S, et al. Bacille Calmette-Guérin induces NOD2-dependent nonspecific protection from reinfection via epigenetic reprogramming of monocytes. [cited 2021 Oct 20]; Available from: www.pnas.org/cgi/doi/https://doi.org/10.1073/pnas.1202870109
43. Verrall AJ, Schneider M, Alisjahbana B, Apriani L, van Laarhoven A, Koeken VACM, et al. Early clearance of mycobacterium tuberculosis is associated with increased innate immune responses. *J Infect Dis.* 2020;221.
44. Pérez-Lu JE, Cárcamo CP, García PJ, Bussalleu A, Bernabé-Ortiz A. Tuberculin skin test conversion among health sciences students: A retrospective cohort study. *Tuberculosis.* 2013;93.
45. Baussano I, Nunn P, Williams B, Pivetta E, Bugiani M, Scano F. Tuberculosis among health care workers. *Emerg Infect Dis.* 2011;17.
46. Ugarte-Gil C, Elkington PT, Gotuzzo E, Friedland JS, Moore DAJ. Induced sputum is safe and well-tolerated for TB diagnosis in a resource-poor primary healthcare setting. *Am. J. Trop. Med. Hyg.* 2015.
47. Alexis N, Soukup J, Ghio A, Becker S. Sputum phagocytes from healthy individuals are functional and activated: A flow cytometric comparison with cells in bronchoalveolar lavage and peripheral blood. *Clin Immunol.* 2000;
48. Alexis NE, Lay JC, Zeman KL, Geiser M, Kapp N, Bennett WD. In vivo particle uptake by airway macrophages in healthy volunteers. *Am J Respir Cell Mol Biol.* 2006;
49. Maksimovic J, Gordon L, Oshlack A. SWAN: Subset-quantile within array normalization for illumina infinium HumanMethylation450 BeadChips. *Genome Biol.* 2012;
50. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, et al. Minfi: A flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics.* 2014;
51. Morris TJ, Butcher LM, Feber A, Teschendorff AE, Chakravarthy AR, Wojdacz TK, et al. ChAMP: 450k Chip Analysis Methylation Pipeline. *Bioinformatics.* 2014;30.
52. Ritchie M, Phipson W, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-seq and microarray studies | *Nucleic Acids Research | Oxford Academic.* limma powers Differ Expr Anal RNA-seq microarray Stud. 2015;
53. Paradis E, Schliep K. Ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics.* 2019;35.
54. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods.* 2012;
55. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics.* 2016;32.
56. Yu G, Wang LG, Han Y, He QY. ClusterProfiler: An R package for comparing biological themes among gene clusters. *Omi A J Integr Biol.* 2012;
57. Kanehisa M. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* 2000

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

