

1 **DNA methylomes derived from alveolar macrophages display distinct**  
2 **patterns in latent tuberculosis - implication for interferon gamma release**  
3 **assay status determination**

4 Isabelle Pehrson<sup>1</sup>ð, Jyotirmoy Das<sup>1</sup>ð, Nina Idh<sup>1</sup>, Lovisa Karlsson<sup>1</sup>, Helena Rylander<sup>1</sup>, Hilma  
5 Hård af Segerstad<sup>1</sup>, Elsa Reuterswård<sup>1</sup>, Emma Marttala<sup>1</sup>, Jakob Paues<sup>1,2</sup>, Melissa Méndez-  
6 Aranda<sup>3</sup>, César Ugarte-Gil<sup>4</sup>•, and Maria Lerm<sup>1</sup>\*•

7 <sup>1</sup>Division of Inflammation and Infection, <sup>2</sup>Division of Infectious Diseases, Department of  
8 Biomedical and Clinical Sciences, Faculty of Medicine and Health Sciences, Linköping  
9 University, SE-58185, Linköping, Sweden, <sup>3</sup>Facultad de Ciencias y Filosofía, Laboratorio de  
10 Investigación en Enfermedades Infecciosas, <sup>4</sup>Facultad de Medicina; Instituto de Medicina  
11 Tropical Alexander von Humboldt, Universidad Peruana Cayetano Heredia, Peru

12

13

14 ð) *Shared first authorship*

15 •) *Shared last authorship*

16 \*) *Corresponding author:*

17 Maria Lerm, Div. of Inflammation and Infection, Lab 1, floor 12

18 Dept. of Biomedical and Clinical Sciences, Faculty of Medicine and Health Sciences

19 Linköping University, SE-58185 Linköping, Sweden

20 Phone: +46-732707786, E-mail: [maria.lerm@liu.se](mailto:maria.lerm@liu.se)

21 **Keywords:** DNA methylation, tuberculosis, LTBI, IGRA, sputum induction, epigenetics

22 **Abstract**

23 Host innate immune cells, including alveolar macrophages, have been identified as key  
24 players in the early eradication of *Mycobacterium tuberculosis* and in the maintenance of an  
25 anti-mycobacterial immune memory, which is believed to be induced through epigenetic  
26 changes. The aim of the study was to elucidate whether exposure to *M. tuberculosis* induced  
27 a different DNA methylation pattern of alveolar macrophages and pulmonary T lymphocytes.

28 Alveolar macrophages and T lymphocytes were isolated from induced sputum obtained from  
29 individuals living in Lima, which is an area high endemic for tuberculosis. To determine the  
30 latent tuberculosis infection status of the subjects, an interferon- $\gamma$  release assay was  
31 performed. We evaluated the DNA methylomes of the alveolar macrophages and T  
32 lymphocytes using the Illumina Infinium Human Methylation 450K Bead Chip array,  
33 revealing a distinct DNA methylation pattern in alveolar macrophages allowing the  
34 discrimination of asymptomatic individuals with latent tuberculosis infection from non-  
35 infected individuals. Pathway analysis revealed that cell signalling of inflammation and  
36 chemokines in alveolar macrophages play a role in latent tuberculosis infection. In  
37 conclusion, we demonstrated that DNA methylation in alveolar macrophages can be used to  
38 determine the tuberculosis infection status of individuals in a high endemic setting.

## 39 **Introduction**

40 Infection with *Mycobacterium tuberculosis*, the causative agent of tuberculosis (TB), can  
41 result in active disease or latent infection. Globally, two billion individuals are currently  
42 estimated to have a latent TB infection (LTBI), and 5-15% of those will develop active TB  
43 during their lifetime<sup>1-4</sup>. In the rest of TB-exposed individuals, the innate immune system will  
44 eliminate the bacteria before the onset of an adaptive immune response, with no detectable  
45 evidence of a persistent immune response, referred to as early clearance<sup>5,6</sup>.

46 *M. tuberculosis* is primarily transmitted via airborne droplets and inhaled into the airways  
47 where it establishes an infection that could develop into pulmonary TB<sup>1</sup>. When *M.*  
48 *tuberculosis* reaches the alveoli within the lungs, the pathogen is phagocytosed by alveolar  
49 macrophages, a subgroup of lung tissue-resident macrophages, which form part of the innate  
50 immune system and are among the first host cells that the bacteria encounter<sup>7,8</sup>.  
51 Subsequently, lung T lymphocytes are recruited and begin to engage in immune responses  
52 against the pathogen<sup>8,9</sup>.

53 The only TB-vaccine available today is the *M. bovis* Bacillus Calmette-Guèrin (BCG). The  
54 vaccine is effective in protecting infants and small children against miliary TB and meningitis  
55 but is unsuccessful in protecting adults against pulmonary TB<sup>10,11</sup>. In addition to its  
56 protective effect against TB in children, the vaccine also protects them against unrelated  
57 pathogens and even reduce general mortality in children, which could be explained by  
58 epigenetic reprogramming of innate immune cells, referred to as trained immunity<sup>12-16</sup>. The  
59 trained innate immune system is in part regulated by DNA methylation, an epigenetic  
60 modifier which regulates various processes within the cell, including anti-inflammatory  
61 responses<sup>17,18</sup>.

62 For the diagnosis of LTBI, the best options today are the tuberculin skin test (TST) and  
63 interferon- $\gamma$  (IFN- $\gamma$ ) release assay (IGRA)<sup>1,19</sup>. The IGRAs detect IFN- $\gamma$  secondary to T  
64 lymphocyte activation and secretion as a response to exposure to *M. tuberculosis*-specific  
65 proteins. The IGRA T-SPOT.*TB* has shown a high reproducibility in serial testing and to have  
66 higher sensitivity than other available tests<sup>20,21</sup>.

67 LTBI is a dynamic state that can progress to active tuberculosis under immunosenescence or  
68 immunosuppression, since it reflects a balance between host immune responses and the  
69 bacilli. Thus, LTBI can be viewed as a spectrum, with variable levels of bacterial control and  
70 risk of progression to active TB infection<sup>22-24</sup>. Even though attempts have been made to  
71 identify biomarkers of LTBI and LTBI progression, including DNA methylation events in  
72 peripheral immune cells of active and latent TB patients<sup>25,26</sup>, no study to our knowledge has  
73 investigated DNA methylation in cells of the lung compartment of IGRA-positive latently  
74 infected individuals.

75 Biosignatures of TB have been proposed to lay ground for next generation of diagnostics for  
76 TB<sup>27,28</sup> and accumulating literature have been published studying possible TB biosignatures,  
77 based on e.g., global methylation analysis, transcriptional profiling and RNA sequencing  
78 enrichment<sup>29-32</sup>. However, no biomarker or signature of active or latent TB has been  
79 successful when generalized or applied in larger sample sizes and biosignature based  
80 diagnostics tools are still unavailable in the clinic.

81 In this study, we isolated alveolar macrophages from individuals residing in a Lima, Peru  
82 which has a high prevalence of tuberculosis. We performed DNA methylation analyses to  
83 identify changes in the methylome and detected 265 significant differently methylated CpG  
84 sites (DMCs) when comparing individuals with LTBI with IGRA-negative non-TB (NTB)

85 individuals. Further analysis of the DMCs revealed their importance in several pathways of  
86 the innate immune system.

87 **Methods**

88 ***Ethical approval***

89 Ethical approval was obtained from the Institutional Committee on Research Ethics “Comité  
90 Institucional de Ética en Investigación” (CIEI) of the Cayetano Heredia National Hospital, N°  
91 082-018 for the inclusion of patients and Universidad Peruana Cayetano Heredia Institutional  
92 Review Board N° 101773 for the inclusion of healthy subjects.

93 ***Subjects***

94 Peru has a high prevalence of TB and a heavy burden of multidrug-resistant (MDR) and  
95 extensively drug-resistant (XDR) TB<sup>33,34</sup>. With reference to the increased risk of Peruvian  
96 health care workers and medical students to acquire a TB infection<sup>35-38</sup>, subjects were  
97 recruited from the from the Universidad Peruana Cayetano Heredia (UPCH) and the Hospital  
98 Nacional Cayetano Heredia in Lima, Peru during April, October and November 2018. 26  
99 individuals aged 18-40 years were included in the study following oral and written consent.  
100 Based on the IGRA results, the subjects were divided into IGRA-positives (LTBI group) and  
101 IGRA-negatives (here defined as NTB group). DNA from alveolar macrophages and T  
102 lymphocytes, obtained via sputum induction, were isolated from both groups and the number  
103 of differentially methylated CpG-sites (DMCs) assessed. For the alveolar macrophages,  
104 pathway analysis and cluster analysis were performed. An auxiliary heatmap analysis was  
105 performed with the addition of two drug resistant tuberculosis (DRTB) patients, included in  
106 June 2019. Following cell isolation and DNA extraction, one HLA-DR and one CD3 sample  
107 from two different subjects (both NTB) had to be excluded due to insufficient DNA amount.

108 ***IFN- $\gamma$  release assay***

109 The latent TB infection status of all subjects except for the active TB patients was evaluated  
110 with an IGRA. The TB infection status of the active TB patients was based on current

111 medical records. Following venepuncture, the IGRA (T-SPOT.*TB*, Oxford Immunotec) was  
112 performed, measuring the IFN- $\gamma$  release of the peripheral immune cells, per the  
113 manufacturer's protocol. The T-SPOT.*TB* result is based on the number of spot-forming cells  
114 (memory T cells producing IFN- $\gamma$ , sensitized in response to the mycobacterial antigens  
115 ESAT-6 and CFP-10).

### 116 *Sputum induction and processing*

117 Sputum specimen were collected and processed as previously described <sup>39</sup>, with some  
118 modifications. The lung immune cells were obtained from each subject (here, salbutamol was  
119 avoided) after the nebulization of a hypertonic saline solution (4%) three times, provoking a  
120 profound cough via osmosis, producing pulmonic secretions. The coughing was preceded by  
121 an oral cleaning procedure. After the collection of the expectorate via sputum induction, the  
122 cell-rich samples were transferred to various cell-culture dishes and accumulations of mucus  
123 of immune cells, referred to as "plugs", were transferred to tubes using forceps. Following  
124 the plug collection, the volume of the sample was estimated and a solution containing  
125 phosphate buffer solution (PBS) and 0.1% dithiothreitol (DTT) was added to the sample  
126 equal to four times the volume of the sample, to provide a reducing agent to decrease sample  
127 viscosity. Following this, the sample was diluted with PBS, and rocked for 5 minutes before  
128 filtrated through a 50  $\mu$ m cell strainer into a new tube to remove contaminating squamous  
129 cells and centrifuged for 5 minutes.

### 130 *DNA extraction of isolated alveolar macrophages and T lymphocytes*

131 Sikkeland *et al.* have previously described the method for the isolation of HLA-DR positive  
132 alveolar macrophages <sup>39,40</sup>. The alveolar macrophages and T lymphocytes (CD3 positive  
133 cells) were isolated according to our protocol <sup>39</sup> using superparamagnetic beads coupled with  
134 anti-human CD3 and Pan Mouse IgG antibodies and HLA-DR/human MHC class II

135 antibodies (Invitrogen Dynabeads). An initial positive selection was done with CD3 beads  
136 followed by a positive HLA-DR selection. Bead coating and cell isolation was performed  
137 according to manufacturer's protocol. The DNA and RNA were extracted from the lung  
138 immune cells using the AllPrep DNA/RNA Mini Kit (Qiagen), per the manufacturer's  
139 instructions, and quantified using the spectrophotometer NanoDrop 2000/2000c (Thermo  
140 Scientific).

#### 141 *DNA methylation data analysis*

142 Genome-wide DNA methylation analysis of the DNA from the two cell types (HLA-DR and  
143 CD3 positive cells) was performed using the Infinium Human Methylation 450K Bead Chip  
144 array (Illumina) as per the manufacturer's instructions. The technique is based on quantitative  
145 genotyping of C/T polymorphism which is generated by DNA bisulfite conversion and allows  
146 for the assessment of over 450 000 methylation sites within the whole genome. DNA was  
147 converted and amplified and subsequently fragmented and hybridized to the Infinium  
148 HumanMethylation450 Bead Chip. The procedure was conducted at the Karolinska Institute,  
149 Sweden.

#### 150 *Computational analysis*

151 The IDAT files (RAW files that contain green and red signals from methylated and  
152 unmethylated CpG sites) were analyzed using the R programming language (v3.6.3) (R  
153 Foundation for Statistical Computing: Vienna, Austria) and Chip Analysis Methylation  
154 Pipeline (*ChAMP*) analysis package<sup>41</sup>. The following filters were applied: i) detection of *p*-  
155 value (<0.01), ii) remove non-CpG probes, iii) SNP removal, iv) remove multi-hit probes, v)  
156 remove all probes in X and Y chromosomes<sup>41</sup>. After filtering, quality control was performed,  
157 and normalization of the data was made using the BMIQ method. The  $\beta$ -values and M-values  
158 of the samples were calculated against each probe per sample. Differential methylation



159 analysis of CpG sites (DMCs) between the sample groups was then performed using  
160 Bonferroni-Hochberg (BH)-corrected  $p$ -value ( $p\text{-value}_{\text{BH}} < 0.05$ ). The DMCs were annotated  
161 with the Illumina manifest file using GRCh37/hg19 database.

162 The statistically significant ( $p\text{-value}_{\text{BH}} < 0.05$ ) was used to create the static volcano plot with  
163 the  $|\log_2\text{FC}| \geq 0.3$  using the *EnhancedVolcano plot* package <sup>42</sup> in R. The hierarchical  
164 cluster analyses were performed from the  $\beta$ -values of the significant DMCs from each  
165 participant using the *hclust* function of the *ade4/ape* package <sup>43</sup> in R. The R packages  
166 *ComplexHeatmap* <sup>44</sup> was used to create the heatmap from individual  $\beta$ -values of DMCs ( $p$ -  
167 value<sub>BH</sub> < 0.05 and  $\log_2$  fold change  $\geq |0.4|$ ). For the CD3 positive cells we found 0 DMCs ( $p$ -  
168 value<sub>BH</sub> < 0.05) (not shown) and thus no other analyses were performed on the CD3 positive  
169 cells. The design matrix of the DNA methylation analysis was “LTBI - NTB” meaning that a  
170 positive  $\log_2$  fold change value equals hypermethylated DMC in the LTBI group. Negative  
171  $\log_2$  fold change value means NTB DMCs are hypermethylated. Pathway enrichment analysis  
172 was performed using Signaling Pathway Impact Analysis (SPIA) <sup>45</sup> enrichment analysis and  
173 *pathfindR* <sup>46</sup>, based on information from the Kyoto Encyclopedia of Genes and Genomes  
174 (KEGG) <sup>47</sup>. Via the SPIA enrichment analysis, based on pathway information from KEGG,  
175 we used the expression rates of the LTBI vs NTB DMGs to calculate signaling pathway  
176 topology. The R package *pathfindR* is a tool for pathway enrichment analyses that uses active  
177 subnetwork to calculate the pathway enrichment. Using the enriched DMGs separating the  
178 LTBI from NTB, we identified active subnetworks and performed pathway enrichment  
179 analyses on the identified subnetworks. Via a scoring function, the overall activity of each  
180 pathway in each sample was estimated.

181 **Results**

182 *Subject characteristics*

183 In order to identify possible DMCs to reflect TB-exposure, we included healthy subjects with  
184 presumable high and low exposure to TB residing in a high endemic area in Lima, Peru and  
185 subsequently divided into two groups based on their IGRA result (Fig. 1). Among the healthy  
186 subjects that were included, six we identified as IGRA-positive (LTBI) and 19 as IGRA-  
187 negative individuals (NTB). Demographics, body metrics, Bacillus Calmette-Guèrin (BCG)-  
188 and smoking status are summarized in Table 1. The patients (DRTB) were hospitalized and  
189 were suffering from MDR-TB and XDR-TB, respectively (Table 2). One patient was smear  
190 microscopy positive and suffered from cough, dyspnea, and night sweats while the other was  
191 asymptomatic and smear microscopy negative at the time. One patient had received the  
192 MDR-antibiotic regimen since 2016. The other patient had a TB disease history of over 10  
193 years, with susceptible TB since 2008 and a later progression to MDR-TB (2016) and XDR-  
194 TB (2019). Both patients were treated with multiple antibiotics at the time of inclusion.

195 *Differential DNA methylation analysis identified DMCs to distinguish LTBI from NTB in*  
196 *alveolar macrophages*

197 To determine whether LTBI and NTB can be distinguished based on the methylation pattern,  
198 we compared the alveolar macrophage and T lymphocyte DNA methylome data obtained  
199 from LTBI and NTB subjects. The analysis revealed 265 DMCs with a  $p\text{-value}_{\text{BH}} < 0.05$  and  
200  $\log_2$  fold change value  $\geq |0.3|$  for alveolar macrophages (not shown). The DMCs  
201 differentiating LTBI from NTB are presented as DMCs with their equivalent DMGs in a  
202 volcano plot (Fig. 2).

203 In total, 36 055 DMCs was found to distinguish LTBI from NTB without any  $\log_2$  fold  
204 change value cut-off (not shown). The same DNA methylation analysis performed with the

205 lung T lymphocytes resulted in 0 DMCs ( $p\text{-value}_{\text{BH}} < 0.05$ ) (not shown). Thus, no further  
206 analysis was performed on the T lymphocytes. No DMCs were found between the subjects  
207 included as TB-exposed and non-exposed (not shown).

208 ***A heatmap analysis revealed a distinct DNA methylation pattern of LTBI and drug***  
209 ***resistant TB in alveolar macrophages***

210 After demonstrating that the LTBI and NTB groups had distinct DNA methylation traits,  
211 knowing that heterogeneity of the epigenome of LTBI subjects has previously been  
212 demonstrated<sup>48</sup> and that LTBI can be viewed as a spectrum<sup>22</sup>, we performed a heatmap  
213 analysis based on the top DMCs of the DNA methylation analysis (265 DMCs,  $p\text{-}$   
214  $\text{value}_{\text{BH}} < 0.05$ ,  $\log_2$  fold change  $\geq |0.3|$ ) that showed that the LTBI group, with the exception  
215 of one participant, clustered into one group (Fig. 3a). In order to investigate whether we could  
216 identify any TB-like subjects within the diverse LTBI group, as previously demonstrated by  
217 Estévez *et al.*<sup>48</sup>, we recruited two individuals with TB and with ongoing treatment to the  
218 study (Table 2). The cluster analysis revealed that the DRTB group clustered with the 5 LTBI  
219 subjects (Fig. 3a). A subsequent principal component analysis (PCA, Fig. 3b) revealed that  
220 the DRTB, despite the similarities with the LTBI group in the cluster analysis, formed a  
221 separate population, demonstrating that the three conditions are clearly separated by the  
222 identified DNA methylation signature.

223 ***Biological processes involved in latent tuberculosis infection***

224 Next, we investigated the biological processes involved in LTBI, based on the genes found in  
225 the differential DNA methylation analysis. We performed pathway enrichment analyses of  
226 the alveolar macrophages to identify the common pathways that distinguished the LTBI  
227 samples from the NTB and found pathways involved in the spliceosome, Wnt signaling,  
228 mTOR signaling, regulation of actin cytoskeleton, ribosome, extracellular matrix

229 organization, MAPK signaling, cellular senescence, TGF- $\beta$  signaling and the  
230 phosphatidylinositol signaling system (Fig. 4).

231 In addition, SPIA pathway enrichment analysis was performed, corroborating the enrichment  
232 of innate immune-related pathways, including chemokine signaling pathway, antigen  
233 processing and presentation, *Staphylococcus aureus* infection, the NOD-like receptor  
234 signaling pathway and intestinal immune network for IgA production (Table S1).

235 ***Counter DNA methylation status of genes persistently hypomethylated in BCG responders***

236 In our previous study, PBMCs obtained from BCG-vaccinated individuals, whose  
237 macrophages displayed an enhanced anti-mycobacterial response, displayed persistent  
238 hypomethylation in six genes, among them the gene encoding IFN- $\gamma$ <sup>49</sup>. In order to estimate  
239 whether these genes' DNA methylation status of the BCG study was coherent with the LTBI  
240 data in the present study, we compared the  $\beta$  values of these genes in the different studies.  
241 Contrary to the observations in the BCG vaccinees, the six genes were found to be  
242 hypermethylated in LTBI (Table S2).

243 **Discussion**

244 In this study, we found that five out of six individuals diagnosed with LTBI based on a  
245 positive IGRA displayed a distinct DNA methylation pattern of alveolar macrophages. The  
246 reason why one of the LTBI individual's signature aligned with the NTB group remains  
247 obscure, but it is tempting to speculate that this individual had cleared a previous latent TB  
248 infection, allowing the LTBI-specific DNA methylome signature to fade while the adaptive  
249 immune memory against TB antigens endured. However, further studies are needed to  
250 provide evidence for such a relationship. A limitation with the present study is that no  
251 information on previous IGRA test results was available, and future studies will dissect the  
252 relationship between the timing of IGRA conversion and the intensity of the DNA  
253 methylome signature.

254 In our previous study on DNA methylation alterations in response to the BCG vaccine, we  
255 identified CpG sites in six gene promoters, one of which belonging to the *IFNG* gene  
256 (encoding IFN- $\gamma$ ), as being persistently more accessible for transcription (hypomethylated) in  
257 BCG responders. In contrast, the present study showed that the very same six CpG sites were  
258 hypermethylated in the LTBI group<sup>49</sup>. We hypothesize that the transcription accessibility of  
259 these six genes has importance for the clearance of *M. tuberculosis*, reflective of a beneficial  
260 BCG response<sup>49</sup>. LTBI is a manifestation of lost mycobacterial control by innate  
261 mechanisms and possibly a consequence of hypermethylation of for example IFN- $\gamma$ , which  
262 has a non-disputable importance for the control of TB<sup>50</sup>.

263 While being a risk factor for developing active TB, LTBI has been shown to convey  
264 protection against *M. tuberculosis* reinfection, partly explained by heterologous immunity  
265 and the concomitant activation of the innate immune system<sup>51,52</sup>. The protective mechanisms  
266 of LTBI were associated with activation of alveolar macrophages and accelerated recruitment

267 of *M. tuberculosis*-specific T cells to the lung and enhanced cytokines production <sup>52</sup>,  
268 compatible with our results where we showed an overrepresentation of pathways involved in  
269 TGF- $\beta$ , NOD-like receptor (NLR) and chemokine signaling. The activation of the NLRP3  
270 inflammasome, activated via *NLRP3* promotor demethylation, has previously been identified  
271 as an important regulatory function in TB <sup>53</sup>. Interestingly, *NLRP3* was differentially  
272 methylated between the LTBI and NTB at 3 CpG sites in our data and we identified NLR  
273 signaling as an activated pathway in LTBI. Increased NLRP3 inflammasome activation  
274 enhances the anti-mycobacterial capacity of macrophages <sup>54</sup> and prevents development of  
275 active pulmonary TB <sup>55</sup>.

276 Methylation of the *VDR* promoter (regulating Vitamin D receptor expression) has previously  
277 been demonstrated to be of importance in TB <sup>56-58</sup>, especially for disease progression and risk.  
278 Vitamin D enhances macrophage capacity to control intracellular *M. tuberculosis* <sup>59,60</sup>. In our  
279 data, we found three differentially methylated *VDR* CpG sites, all hypermethylated in the  
280 LTBI group, pointing towards downregulation of *VDR* responsiveness in alveolar  
281 macrophages from subjects with LTBI, which may impair their ability to kill intracellular  
282 mycobacteria <sup>61</sup>.

283 Early clearance is observed in about 25% of individuals who are living with considerable  
284 exposure to TB <sup>62-64</sup>, and epigenetic mechanisms related to trained immunity could contribute  
285 to the elimination of the infection without any need to trigger an adaptive immune response  
286 <sup>65</sup>. We propose that the LTBI group has been unsuccessful in clearing the bacteria and that  
287 the epigenetic differences we identify between the LTBI and NTB group.

288 The heatmap analysis revealed that the DNA methylation pattern of alveolar macrophages of  
289 DRTB patients, who were on MDR treatment at the timepoint of inclusion, displayed similar  
290 pattern with the LTBI individuals, however they formed a distinct cluster the PCA analysis.

291 Since the DRTB patients had been on TB treatment for an extended period of time, our study  
292 does not reveal the DNA methylome signature of active pulmonary TB, but suggests that  
293 DNA methylome analysis is a promising approach to study the spectrum of TB. Therefore,  
294 longitudinal studies are needed that cover DNA methylation changes of alveolar  
295 macrophages before the time point of IGRA conversion over the development of active TB  
296 and the TB treatment until cure. The results will allow dissection of the unique DNA  
297 methylation signatures for each state of TB and could be further developed as tools for  
298 clinical TB diagnosis, decision-making in LTBI infected subjects and TB treatment  
299 monitoring.

300 **Data and materials availability**

301 Sequencing datasets will be made publicly available upon acceptance and prior to final  
302 publication.

303 **Acknowledgements**

304 We owe our thanks to the Bioinformatics and Expression Analysis core facility (BEA), which  
305 is supported by the board of research at the Karolinska Institute and the research committee at  
306 the Karolinska hospital. The study was funded through a grant from the Swedish Research  
307 Council N° 2018-04246 and the Consejo Nacional de Ciencia, Tecnología e Innovación  
308 Tecnológica CONCYTEC and Cienciactiva (N° 106-2018-FONDECYT). The computations  
309 and data handling were enabled by resources provided by the Swedish National Infrastructure  
310 for Computing (SNIC) at National Supercomputing Centre (NSC), Linköping University  
311 partially funded by the Swedish Research Council through grant agreement N° 2018-05973.  
312 We also owe thanks to the support by grants from the World Infection Fund. J.D is a  
313 postdoctoral fellow supported through the Medical Infection and Inflammation Center (MIIC)  
314 at Linköping University. We direct our gratitude to all the subjects for donating samples. We  
315 are also grateful for the help of Ronald Cadillo Hernandez and Rodrigo Cachay Figueroa.

316 **Authors Contribution**

317 M.L, N.I, J.P, C.U-G and M.M-A wrote the ethical application and conceived the study. I.P,  
318 E.M, H.R, E.R and H.H performed all sample collections. J.D and L.K preformed the  
319 bioinformatic analyses. I.P, M.L, J.D and L.K wrote the manuscript. M.M-A and C.U-G  
320 supervised the project in Peru.

321 **Conflicting interest**

322 All authors declare no conflict of interest.



323 **Consent for publication**

324 All authors have approved to the final version of the manuscript.

325 **References**

- 326 1. WHO | Global tuberculosis report 2020. WHO. Geneva, Switzerland. 2020
- 327 2. Vynnycky E, Fine PE. The natural history of tuberculosis: the implications of age-  
328 dependent risks of disease and the role of reinfection. *Epidemiol Infect.* 1997  
329 Oct;119(2):183-201. doi: 10.1017/s0950268897007917.
- 330 3. Comstock GW, Livesay VT, Woolpert SF. The prognosis of a positive tuberculin  
331 reaction in childhood and adolescence. *Am J Epidemiol.* 1974 Feb;99(2):131-8.  
332 doi: 10.1093/oxfordjournals.aje.a121593.
- 333 4. Sester M, van Leth F, Bruchfeld J, Bumbacea D, Cirillo DM, Dilektasli AG,  
334 Domínguez J, Duarte R, Ernst M, Eyuboglu FO, Gerogianni I, Girardi E, Goletti D,  
335 Janssens JP, Julander I, Lange B, Latorre I, Losi M, Markova R, Matteelli A,  
336 Milburn H, Ravn P, Scholman T, Soccac PM, Straub M, Wagner D, Wolf T, Yalcin  
337 A, Lange C; TBNET. Risk assessment of tuberculosis in immunocompromised  
338 patients. A TBNET study. *Am J Respir Crit Care Med.* 2014 Nov 15;190(10):1168-  
339 76. doi: 10.1164/rccm.201405-0967OC.
- 340 5. Meermeier EW, Lewinsohn DM. Early clearance versus control: what is the  
341 meaning of a negative tuberculin skin test or interferon-gamma release assay  
342 following exposure to *Mycobacterium tuberculosis*? *F1000Res.* 2018 May  
343 25;7:F1000 Faculty Rev-664. doi: 10.12688/f1000research.13224.1.
- 344 6. Koeken VACM, Verrall AJ, Netea MG, Hill PC, van Crevel R. Trained innate  
345 immunity and resistance to *Mycobacterium tuberculosis* infection. *Clin Microbiol*  
346 *Infect.* 2019 Dec;25(12):1468-1472. doi: 10.1016/j.cmi.2019.02.015. Epub 2019  
347 Feb 23.

- 348 7. Cohen SB, Gern BH, Delahaye JL, Adams KN, Plumlee CR, Winkler JK, Sherman  
349 DR, Gerner MY, Urdahl KB. Alveolar Macrophages Provide an Early  
350 Mycobacterium tuberculosis Niche and Initiate Dissemination. *Cell Host Microbe*.  
351 2018 Sep 12;24(3):439-446.e4. doi: 10.1016/j.chom.2018.08.001. Epub 2018 Aug  
352 23.
- 353 8. Mayer-Barber KD, Barber DL. Innate and Adaptive Cellular Immune Responses to  
354 Mycobacterium tuberculosis Infection. *Cold Spring Harb Perspect Med*. 2015 Jul  
355 17;5(12):a018424. doi: 10.1101/cshperspect.a018424.
- 356 9. Eghtesad M, Jackson HE, Cunningham AC. Primary human alveolar epithelial cells  
357 can elicit the transendothelial migration of CD14<sup>+</sup> monocytes and CD3<sup>+</sup>  
358 lymphocytes. *Immunology*. 2001 Feb;102(2):157-64. doi: 10.1046/j.1365-  
359 2567.2001.01172.x.
- 360 10. Roy A, Eisenhut M, Harris RJ, Rodrigues LC, Sridhar S, Habermann S, Snell L,  
361 Mangtani P, Adetifa I, Lalvani A, Abubakar I. Effect of BCG vaccination against  
362 Mycobacterium tuberculosis infection in children: systematic review and meta-  
363 analysis. *BMJ*. 2014 Aug 5;349:g4643. doi: 10.1136/bmj.g4643.
- 364 11. Mangtani P, Abubakar I, Ariti C, Beynon R, Pimpin L, Fine PE, Rodrigues LC,  
365 Smith PG, Lipman M, Whiting PF, Sterne JA. Protection by BCG vaccine against  
366 tuberculosis: a systematic review of randomized controlled trials. *Clin Infect Dis*.  
367 2014 Feb;58(4):470-80. doi: 10.1093/cid/cit790. Epub 2013 Dec 13.
- 368 12. Butkeviciute E, Jones CE, Smith SG. Heterologous effects of infant BCG  
369 vaccination: potential mechanisms of immunity. *Future Microbiol*. 2018  
370 Aug;13(10):1193-1208. doi: 10.2217/fmb-2018-0026. Epub 2018 Aug 17.

- 371 13. Ferluga J, Yasmin H, Al-Ahdal MN, Bhakta S, Kishore U. Natural and trained  
372 innate immunity against Mycobacterium tuberculosis. *Immunobiology*. 2020  
373 May;225(3):151951. doi: 10.1016/j.imbio.2020.151951. Epub 2020 Apr 27.
- 374 14. Netea MG, Joosten LA, Latz E, Mills KH, Natoli G, Stunnenberg HG, O'Neill LA,  
375 Xavier RJ. Trained immunity: A program of innate immune memory in health and  
376 disease. *Science*. 2016 Apr 22;352(6284):aaf1098. doi: 10.1126/science.aaf1098.  
377 Epub 2016 Apr 21.
- 378 15. Arts RJW, Moorlag SJCFM, Novakovic B, Li Y, Wang SY, Oosting M, Kumar V,  
379 Xavier RJ, Wijmenga C, Joosten LAB, Reusken CBEM, Benn CS, Aaby P,  
380 Koopmans MP, Stunnenberg HG, van Crevel R, Netea MG. BCG Vaccination  
381 Protects against Experimental Viral Infection in Humans through the Induction of  
382 Cytokines Associated with Trained Immunity. *Cell Host Microbe*. 2018 Jan  
383 10;23(1):89-100.e5. doi: 10.1016/j.chom.2017.12.010.
- 384 16. Covián C, Fernández-Fierro A, Retamal-Díaz A, Díaz FE, Vasquez AE, Lay MK,  
385 Riedel CA, González PA, Bueno SM, Kalergis AM. BCG-Induced Cross-  
386 Protection and Development of Trained Immunity: Implication for Vaccine Design.  
387 *Front Immunol*. 2019 Nov 29;10:2806. doi: 10.3389/fimmu.2019.02806.
- 388 17. Maruthai K, Kalaiarasan E, Joseph NM, Parija SC, Mahadevan S. Assessment of  
389 global DNA methylation in children with tuberculosis disease. *Int J Mycobacteriol*.  
390 2018 Oct-Dec;7(4):338-342. doi: 10.4103/ijmy.ijmy\_107\_18.
- 391 18. Morales-Nebreda L, McLafferty FS, Singer BD. DNA methylation as a  
392 transcriptional regulator of the immune system. *Transl Res*. 2019 Feb;204:1-18.  
393 doi: 10.1016/j.trsl.2018.08.001. Epub 2018 Aug 9.

- 394 19. Furin J, Cox H, Pai M. Tuberculosis. *Lancet*. 2019 Apr 20;393(10181):1642-1656.  
395 doi: 10.1016/S0140-6736(19)30308-3. Epub 2019 Mar 20.
- 396 20. Meier T, Enders M. High reproducibility of the interferon-gamma release assay T-  
397 SPOT.TB in serial testing. *Eur J Clin Microbiol Infect Dis*. 2021 Jan;40(1):85-93.  
398 doi: 10.1007/s10096-020-03997-3. Epub 2020 Aug 8.
- 399 21. Pai M, Zwerling A, Menzies D. Systematic review: T-cell-based assays for the  
400 diagnosis of latent tuberculosis infection: an update. *Ann Intern Med*. 2008 Aug  
401 5;149(3):177-84. doi: 10.7326/0003-4819-149-3-200808050-00241. Epub 2008 Jun  
402 30.
- 403 22. Barry CE 3rd, Boshoff HI, Dartois V, Dick T, Ehrt S, Flynn J, Schnappinger D,  
404 Wilkinson RJ, Young D. The spectrum of latent tuberculosis: rethinking the  
405 biology and intervention strategies. *Nat Rev Microbiol*. 2009 Dec;7(12):845-55.  
406 doi: 10.1038/nrmicro2236. Epub 2009 Oct 26.
- 407 23. Chee CBE, Reves R, Zhang Y, Belknap R. Latent tuberculosis infection:  
408 Opportunities and challenges. *Respirology*. 2018 Oct;23(10):893-900. doi:  
409 10.1111/resp.13346. Epub 2018 Jun 14.
- 410 24. Drain PK, Bajema KL, Dowdy D, Dheda K, Naidoo K, Schumacher SG, Ma S,  
411 Meermeier E, Lewinsohn DM, Sherman DR. Incipient and Subclinical  
412 Tuberculosis: a Clinical Review of Early Stages and Progression of Infection. *Clin*  
413 *Microbiol Rev*. 2018 Jul 18;31(4):e00021-18. doi: 10.1128/CMR.00021-18.
- 414 25. Esterhuyse MM, Weiner J 3rd, Caron E, Loxton AG, Iannaccone M, Wagman C,  
415 Saikali P, Stanley K, Wolski WE, Mollenkopf HJ, Schick M, Aebersold R, Linhart  
416 H, Walzl G, Kaufmann SH. Epigenetics and Proteomics Join Transcriptomics in the

- 417 Quest for Tuberculosis Biomarkers. *mBio*. 2015 Sep 15;6(5):e01187-15. doi:  
418 10.1128/mBio.01187-15.
- 419 26. Sharma G, Sowpati DT, Singh P, Khan MZ, Ganji R, Upadhyay S, Banerjee S,  
420 Nandicoori VK, Khosla S. Genome-wide non-CpG methylation of the host genome  
421 during *M. tuberculosis* infection. *Sci Rep*. 2016 Apr 26;6:25006. doi:  
422 10.1038/srep25006.
- 423 27. Haas CT, Roe JK, Pollara G, Mehta M, Noursadeghi M. Diagnostic 'omics' for  
424 active tuberculosis. *BMC Med*. 2016 Mar 23;14:37. doi: 10.1186/s12916-016-  
425 0583-9.
- 426 28. Lerm M, Dockrell HM. Addressing diversity in tuberculosis using  
427 multidimensional approaches. *J Intern Med*. 2018 May 27. doi:  
428 10.1111/joim.12776. Epub ahead of print.
- 429 29. Maertzdorf J, Kaufmann SH, Weiner J 3rd. Toward a unified biosignature for  
430 tuberculosis. *Cold Spring Harb Perspect Med*. 2014 Oct 23;5(1):a018531. doi:  
431 10.1101/cshperspect.a018531.
- 432 30. Petrilli JD, Araújo LE, da Silva LS, Laus AC, Müller I, Reis RM, Netto EM, Riley  
433 LW, Arruda S, Queiroz A. Whole blood mRNA expression-based targets to  
434 discriminate active tuberculosis from latent infection and other pulmonary diseases.  
435 *Sci Rep*. 2020 Dec 16;10(1):22072. doi: 10.1038/s41598-020-78793-2.
- 436 31. Pan L, Wei N, Jia H, Gao M, Chen X, Wei R, Sun Q, Gu S, Du B, Xing A, Zhang  
437 Z. Genome-wide transcriptional profiling identifies potential signatures in  
438 discriminating active tuberculosis from latent infection. *Oncotarget*. 2017 Dec  
439 4;8(68):112907-112916. doi: 10.18632/oncotarget.22889.

- 440 32. Zheng L, Leung ET, Wong HK, Lui G, Lee N, To KF, Choy KW, Chan RC, Ip M.  
441 Unraveling methylation changes of host macrophages in Mycobacterium  
442 tuberculosis infection. *Tuberculosis (Edinb)*. 2016 May;98:139-48. doi:  
443 10.1016/j.tube.2016.03.003. Epub 2016 Mar 21.
- 444 33. Alarcón V, Alarcón E, Figueroa C, Mendoza-Ticona A. Tuberculosis en el Perú:  
445 situación epidemiológica, avances y desafíos para su control [Tuberculosis in Peru:  
446 epidemiological situation, progress and challenges for its control]. *Rev Peru Med*  
447 *Exp Salud Publica*. 2017 Apr-Jun;34(2):299-310. Spanish. doi:  
448 10.17843/rpmesp.2017.342.2384.
- 449 34. Woodman M, Haeusler IL, Grandjean L. Tuberculosis Genetic Epidemiology: A  
450 Latin American Perspective. *Genes (Basel)*. 2019 Jan 16;10(1):53. doi:  
451 10.3390/genes10010053.
- 452 35. Pérez-Lu JE, Cárcamo CP, García PJ, Bussalleu A, Bernabé-Ortiz A. Tuberculin  
453 skin test conversion among health sciences students: a retrospective cohort study.  
454 *Tuberculosis (Edinb)*. 2013 Mar;93(2):257-62. doi: 10.1016/j.tube.2012.10.001.  
455 Epub 2012 Oct 30.
- 456 36. Baussano I, Nunn P, Williams B, Pivetta E, Bugiani M, Scano F. Tuberculosis  
457 among health care workers. *Emerg Infect Dis*. 2011 Mar;17(3):488-94. doi:  
458 10.3201/eid1703.100947.
- 459 37. Soto-Cabezas MG, Chávez-Pachas AM, Arrasco-Alegre JC, Yagui-Moscoso MJ.  
460 Tuberculosis en trabajadores de salud en el Perú, 2013-2015 [Tuberculosis in health  
461 workers in Peru, 2013-2015]. *Rev Peru Med Exp Salud Publica*. 2016 Oct-  
462 Dec;33(4):607-615. Spanish. doi: 10.17843/rpmesp.2016.334.2542.

- 463 38. Soto Cabezas MG, Munayco Escate CV, Chávez Herrera J, López Romero SL,  
464 Moore D. Prevalencia de infección tuberculosa latente en trabajadores de salud de  
465 establecimientos del primer nivel de atención. Lima, Perú [Prevalence of latent  
466 tuberculosis infection in health workers from primary health care centers in Lima,  
467 Peru]. *Rev Peru Med Exp Salud Publica*. 2017 Oct-Dec;34(4):649-654. Spanish.  
468 doi: 10.17843/rpmesp.2017.344.3035.
- 469 39. Das J, Idh N, Sikkeland L. I. B, Paues J, Lerm M. DNA methylome-based  
470 validation of induced sputum as an effective protocol to study lung immunity:  
471 construction of a classifier of pulmonary cell types. *bioRxiv* (2021)  
472 doi:10.1101/2021.03.12.435086
- 473 40. Sikkeland LI, Kongerud J, Stangeland AM, Haug T, Alexis NE. Macrophage  
474 enrichment from induced sputum. *Thorax*. 2007 Jun;62(6):558-9. doi:  
475 10.1136/thx.2006.073544.
- 476 41. Morris TJ, Butcher LM, Feber A, Teschendorff AE, Chakravarthy AR, Wojdacz  
477 TK, Beck S. ChAMP: 450k Chip Analysis Methylation Pipeline. *Bioinformatics*.  
478 2014 Feb 1;30(3):428-30. doi: 10.1093/bioinformatics/btt684. Epub 2013 Dec 12.
- 479 42. Blighe K, Rana S, Lewis M (2020). EnhancedVolcano: Publication-ready volcano  
480 plots with enhanced colouring and labeling. R package version 1.8.0,  
481 <https://github.com/kevinblighe/EnhancedVolcano>.
- 482 43. Paradis E, Schliep K (2019). “ape 5.0: an environment for modern phylogenetics  
483 and evolutionary analyses in R.” *Bioinformatics*, 35, 526-528.
- 484 44. Gu Z, Eils R, Schlesner M (2016). “Complex heatmaps reveal patterns and  
485 correlations in multidimensional genomic data.” *Bioinformatics*.



- 486 45. Tarca AL, Kathri P, Draghici S (2020). SPIA: Signaling Pathway Impact Analysis  
487 (SPIA) using combined evidence of pathway over-representation and unusual  
488 signaling perturbations. R package version 2.42.0,  
489 <http://bioinformatics.oxfordjournals.org/cgi/reprint/btn577v1>.
- 490 46. Ulgen E, Ozisik O, Sezerman OU. pathfindR: An R Package for Comprehensive  
491 Identification of Enriched Pathways in Omics Data Through Active Subnetworks.  
492 *Front Genet.* 2019 Sep 25;10:858. doi: 10.3389/fgene.2019.00858.
- 493 47. KEGG as a reference resource for gene and protein annotation *Nucleic Acids*  
494 *Research*, Volume 44, Issue D1, 4 January 2016, Pages D457–D462,  
495 <https://doi.org/10.1093/nar/gkv1070>
- 496 48. Estévez O, Anibarro L, Garet E, Pallares Á, Barcia L, Calviño L, Maueia C, Mussá  
497 T, Fdez-Riverola F, Glez-Peña D, Reboiro-Jato M, López-Fernández H, Fonseca  
498 NA, Reljic R, González-Fernández Á. An RNA-seq Based Machine Learning  
499 Approach Identifies Latent Tuberculosis Patients With an Active Tuberculosis  
500 Profile. *Front Immunol.* 2020 Jul 14;11:1470. doi: 10.3389/fimmu.2020.01470.
- 501 49. Verma D, Parasa VR, Raffetseder J, Martis M, Mehta RB, Netea M, Lerm M. Anti-  
502 mycobacterial activity correlates with altered DNA methylation pattern in immune  
503 cells from BCG-vaccinated subjects. *Sci Rep.* 2017 Sep 26;7(1):12305. doi:  
504 10.1038/s41598-017-12110-2.
- 505 50. Moreira-Teixeira L, Mayer-Barber K, Sher A, O'Garra A. Type I interferons in  
506 tuberculosis: Foe and occasionally friend. *J Exp Med.* 2018 May 7;215(5):1273-  
507 1285. doi: 10.1084/jem.20180325. Epub 2018 Apr 17.
- 508 51. Andrews JR, Noubary F, Walensky RP, Cerda R, Losina E, Horsburgh CR. Risk of  
509 progression to active tuberculosis following reinfection with *Mycobacterium*

- 510 tuberculosis. *Clin Infect Dis*. 2012 Mar;54(6):784-91. doi: 10.1093/cid/cir951.  
511 Epub 2012 Jan 19.
- 512 52. Nemeth J, Olson GS, Rothchild AC, Jahn AN, Mai D, Duffy FJ, Delahaye JL,  
513 Srivatsan S, Plumlee CR, Urdahl KB, Gold ES, Aderem A, Diercks AH. Contained  
514 *Mycobacterium tuberculosis* infection induces concomitant and heterologous  
515 protection. *PLoS Pathog*. 2020 Jul 16;16(7):e1008655. doi:  
516 10.1371/journal.ppat.1008655.
- 517 53. Wei M, Wang L, Wu T, Xi J, Han Y, Yang X, Zhang D, Fang Q, Tang B. NLRP3  
518 Activation Was Regulated by DNA Methylation Modification during  
519 *Mycobacterium tuberculosis* Infection. *Biomed Res Int*. 2016;2016:4323281. doi:  
520 10.1155/2016/4323281. Epub 2016 Jun 6.
- 521 54. Eklund D, Welin A, Andersson H, Verma D, Söderkvist P, Stendahl O, Särndahl E,  
522 Lerm M. Human gene variants linked to enhanced NLRP3 activity limit  
523 intramacrophage growth of *Mycobacterium tuberculosis*. *J Infect Dis*. 2014 Mar  
524 1;209(5):749-53. doi: 10.1093/infdis/jit572. Epub 2013 Oct 24.
- 525 55. Souza de Lima D, Ogusku MM, Sadahiro A, Pontillo A. Inflammasome genetics  
526 contributes to the development and control of active pulmonary tuberculosis. *Infect*  
527 *Genet Evol*. 2016 Jul;41:240-244. doi: 10.1016/j.meegid.2016.04.015. Epub 2016  
528 Apr 19.
- 529 56. Andraos C, Koorsen G, Knight JC, Bornman L. Vitamin D receptor gene  
530 methylation is associated with ethnicity, tuberculosis, and TaqI polymorphism.  
531 *Hum Immunol*. 2011 Mar;72(3):262-8. doi: 10.1016/j.humimm.2010.12.010. Epub  
532 2010 Dec 16.

- 533 57. Jiang C, Zhu J, Liu Y, Luan X, Jiang Y, Jiang G, Fan J. The methylation state of  
534 VDR gene in pulmonary tuberculosis patients. *J Thorac Dis.* 2017 Nov;9(11):4353-  
535 4357. doi: 10.21037/jtd.2017.09.107.
- 536 58. Wang M, Kong W, He B, Li Z, Song H, Shi P, Wang J. Vitamin D and the  
537 promoter methylation of its metabolic pathway genes in association with the risk  
538 and prognosis of tuberculosis. *Clin Epigenetics.* 2018 Sep 12;10(1):118. doi:  
539 10.1186/s13148-018-0552-6.
- 540 59. Chen L, Eapen MS, Zosky GR. Vitamin D both facilitates and attenuates the  
541 cellular response to lipopolysaccharide. *Sci Rep.* 2017 Mar 27;7:45172. doi:  
542 10.1038/srep45172.
- 543 60. Eklund D, Persson HL, Larsson M, Welin A, Idh J, Paues J, Fransson SG, Stendahl  
544 O, Schön T, Lerm M. Vitamin D enhances IL-1 $\beta$  secretion and restricts growth of  
545 *Mycobacterium tuberculosis* in macrophages from TB patients. *Int J*  
546 *Mycobacteriol.* 2013 Mar;2(1):18-25. doi: 10.1016/j.ijmyco.2012.11.001. Epub  
547 2012 Dec 20.
- 548 61. Liu PT, Stenger S, Li H, Wenzel L, Tan BH, Krutzik SR, Ochoa MT, Schaubert J,  
549 Wu K, Meinken C, Kamen DL, Wagner M, Bals R, Steinmeyer A, Zügel U, Gallo  
550 RL, Eisenberg D, Hewison M, Hollis BW, Adams JS, Bloom BR, Modlin RL. Toll-  
551 like receptor triggering of a vitamin D-mediated human antimicrobial response.  
552 *Science.* 2006 Mar 24;311(5768):1770-3. doi: 10.1126/science.1123933. Epub  
553 2006 Feb 23.
- 554 62. Hill PC, Brookes RH, Fox A, Jackson-Sillah D, Jeffries DJ, Lugos MD, Donkor  
555 SA, Adetifa IM, de Jong BC, Aiken AM, Adegbola RA, McAdam KP.

- 556 Longitudinal assessment of an ELISPOT test for Mycobacterium tuberculosis  
557 infection. PLoS Med. 2007 Jun;4(6):e192. doi: 10.1371/journal.pmed.0040192.
- 558 63. Stein CM, Zalwango S, Malone LL, Thiel B, Mupere E, Nsereko M, Okware B,  
559 Kisingo H, Lancioni CL, Bark CM, Whalen CC, Joloba ML, Boom WH, Mayanja-  
560 Kizza H. Resistance and Susceptibility to Mycobacterium tuberculosis Infection  
561 and Disease in Tuberculosis Households in Kampala, Uganda. Am J Epidemiol.  
562 2018 Jul 1;187(7):1477-1489. doi: 10.1093/aje/kwx380.
- 563 64. Verrall AJ, Alisjahbana B, Apriani L, Novianty N, Nurani AC, van Laarhoven A,  
564 Ussher JE, Indrati A, Ruslami R, Netea MG, Sharples K, van Crevel R, Hill PC.  
565 Early Clearance of Mycobacterium tuberculosis: The INFECT Case Contact Cohort  
566 Study in Indonesia. J Infect Dis. 2020 Mar 28;221(8):1351-1360. doi:  
567 10.1093/infdis/jiz168.
- 568 65. Verrall AJ, Schneider M, Alisjahbana B, Apriani L, van Laarhoven A, Koeken  
569 VACM, van Dorp S, Diadani E, Utama F, Hannaway RF, Indrati A, Netea MG,  
570 Sharples K, Hill PC, Ussher JE, van Crevel R. Early Clearance of Mycobacterium  
571 tuberculosis Is Associated With Increased Innate Immune Responses. J Infect Dis.  
572 2020 Mar 28;221(8):1342-1350. doi: 10.1093/infdis/jiz147.

573 **Figure legends**

574 **Figure 1.** A flow chart of the study design. TB = tuberculosis. IGRA = interferon- $\gamma$  release  
575 assay. LTBI = latent tuberculosis infection. NTB = non-tuberculosis. DRTB = drug resistant  
576 tuberculosis. DMCs = differentially methylated CpG-sites.

577 **Figure 2.** Volcano plot of the differentially methylated CpG-sited (DMCs) separating the  
578 latent tuberculosis infection (LTBI) from the non-tuberculosis (NTB) group. The plot is  
579 showing the distribution of the DMCs fold changes in LTBI relative to NBT in alveolar  
580 macrophages. Each dot represents one CpG site ( $p$ -value<sub>BH</sub> < 0.05). x-scale =  $\log_2FC$ , y-scale  
581 =  $-\log_{10} p$  adjusted; blue CpGs = hypomethylated; red CpGs = hypermethylated; black CpGs  
582 = significant but  $|\log_2FC| < 0.3$  horizontal black line in parallel to x-axis =  $p$ -value<sub>BH</sub> < 0.05,  
583 and vertical green lines parallel to y-axis =  $|\log_2FC| \geq 0.3$ .

584 **Figure 3a.** Heatmap of the LTBI, NTB and DRTB groups. The heatmap is plotted from  
585 the  $\beta$ -values of the top 265 CpGs ( $\log_2$  fold change  $\geq |0.3|$ ,  $p$ -value<sub>BH</sub> < 0.05) differentiating  
586 the LTBI and NTB groups. The drug-resistant TB patients (DR-TB) are added to the  
587 analysis. The color scale represent the  $\beta$ -value (from 0 (blue) to 1 (red)) of each CpG site.  
588 Orange: LTBI = latent tuberculosis infection, green: NTB = non-tuberculosis, red: DRTB =  
589 drug-resistant tuberculosis as shown in the bar plot of the heatmap. Cluster dendrogram is  
590 calculated using the Euclidean distance method.

591 **Figure 3b.** Principal components analysis (PCA) of the LTBI, NTB and DRTB group. The  
592 PCA plot is generated from the  $\beta$ -values of the top 265 CpGs ( $\log_2$  fold change  $\geq |0.3|$ ,  $p$ -  
593 value<sub>BH</sub> < 0.05) differentiating the LTBI and NTB groups, with the addition of the  $\beta$ -  
594 values of the DRTB patients of those DMCs.

595 **Figure 4.** Top 10 clusters of the enriched pathways from pathfindR based on the  
596 differentially methylated genes (DMGs) differentiating LTBI from NTB. The dot plot is

597 demonstrating the most relevant pathways for the alveolar macrophages when comparing the  
598 LTBI with the NTB group. The gene count is presented as circles, increasing with elevating  
599 number of gene count. The significance level is presented as a color scale ( $-\log_{10}(p)$ ).

600 **Table legends**

601 **Table 1.** Demographic and clinical data of the LTBI and NTB study population.

602 **Table 2.** Demographic and clinical data of the two drug resistant pulmonary tuberculosis  
603 patients.

604 **Supplementary table S1.** Table of 7 significantly over-represented pathways identified via  
605 the generation with the Signaling Pathway Impact Analysis (SPIA) enrichment analysis. The  
606 pathway information is based on Kyoto Encyclopedia of Genes and Genomes (KEGG)  
607 generated pathways. Global  $p$ -value =  $P_G$ . FDR = false discovery rate.

608 **Supplementary table S2.** Comparison of the DNA methylation status of 6 genes between  
609 our data and data from Verma *et al.* (2017)<sup>49</sup>.

610

611 **Table 1.** Demographic and clinical data of the LTBI and NTB study population.

	<b>LTBI</b> (N=6)	<b>NTB</b> (N=19)	<b>Total</b> (N=25)
Age (mean)	23 (19-26)	25 (18-40)	25 (18-40)
Sex (male/female)	5/1	10/9	15/10
Weight (kg)	78.2 (54-94)	63.2 (47-85)	66.8 (47-94)
Height (cm)	175.7 (168-183)	165.4 (148-178)	167.8 (148-183)
BMI (kg/m <sup>2</sup> )	25.2 (19.1-29.3)	22.9 (19.1-28.7)	23.5 (19.1-29.3)
Smoker/non-smoker	2/4	1/18	3/22
IGRA result (positive/negative)	6/0	0/19	6/19
History of TST-positivity (yes/no/ND)	1/4/1	3/6/10	4/10/11
BCG (vaccinated/non- vaccinated)	6/0	16/3	22/3
Included as TB-exposed/non- exposed	5/1	9/10	14/11
Previous TB-treatment (yes/no)	0/6	1/18	1/24
Previous dust exposure (yes/no)	0/6	0/19	0/25

612

613 LTBI = latent tuberculosis infection. NTB = non-tuberculosis. BMI = body mass index. BCG

614 = Bacillus Calmette-Guèrin. ND = no data. TB = tuberculosis. TST = tuberculin skin test.

615



616 **Table 2.** Demographic and clinical data of the two drug resistant pulmonary tuberculosis  
617 patients.

	<b>DRTB (N=2)</b>
Age (mean)	30 (28-32)
Sex (male/female)	2/0
Weight (kg)	61.5 (58-65)
Height (cm)	173 (166-180)
BMI (kg/m <sup>2</sup> )	20.6 (20.1-21.1)
Smoker/previous smoker/non-smoker	0/1/1
BCG (vaccinated/non-vaccinated/ND)	1/0/1
Previous TB-treatment (yes/no)	2/0
Previous dust exposure (yes/no)	1/1
MDR-TB/XDR-TB	1/1
Smear microscopy (positive/negative)	1/1
Years of TB-disease	7 (3-11)
Duration of MDR-TB-treatment (years)	4.5 (3-6)

618

619 DRTB = drug-resistant TB. BMI = body mass index. BCG = Bacillus Calmette-Guèrin. ND =  
620 no data. TB = tuberculosis. MDR = multi-drug resistant. XDR = extensively drug resistant.

Figure 1

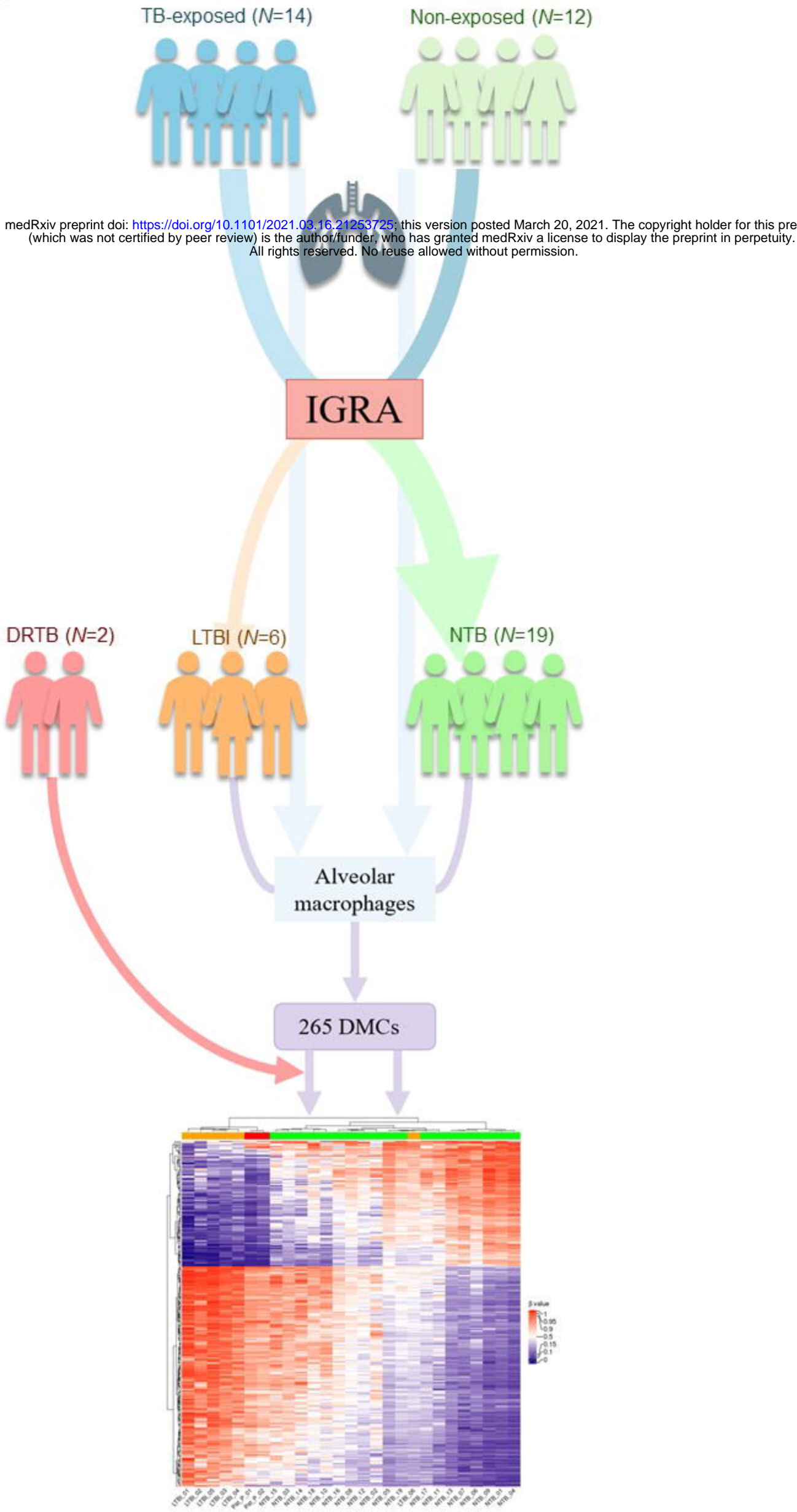


Figure 2

medRxiv preprint doi: <https://doi.org/10.1101/2021.03.16.21253725>; this version posted March 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

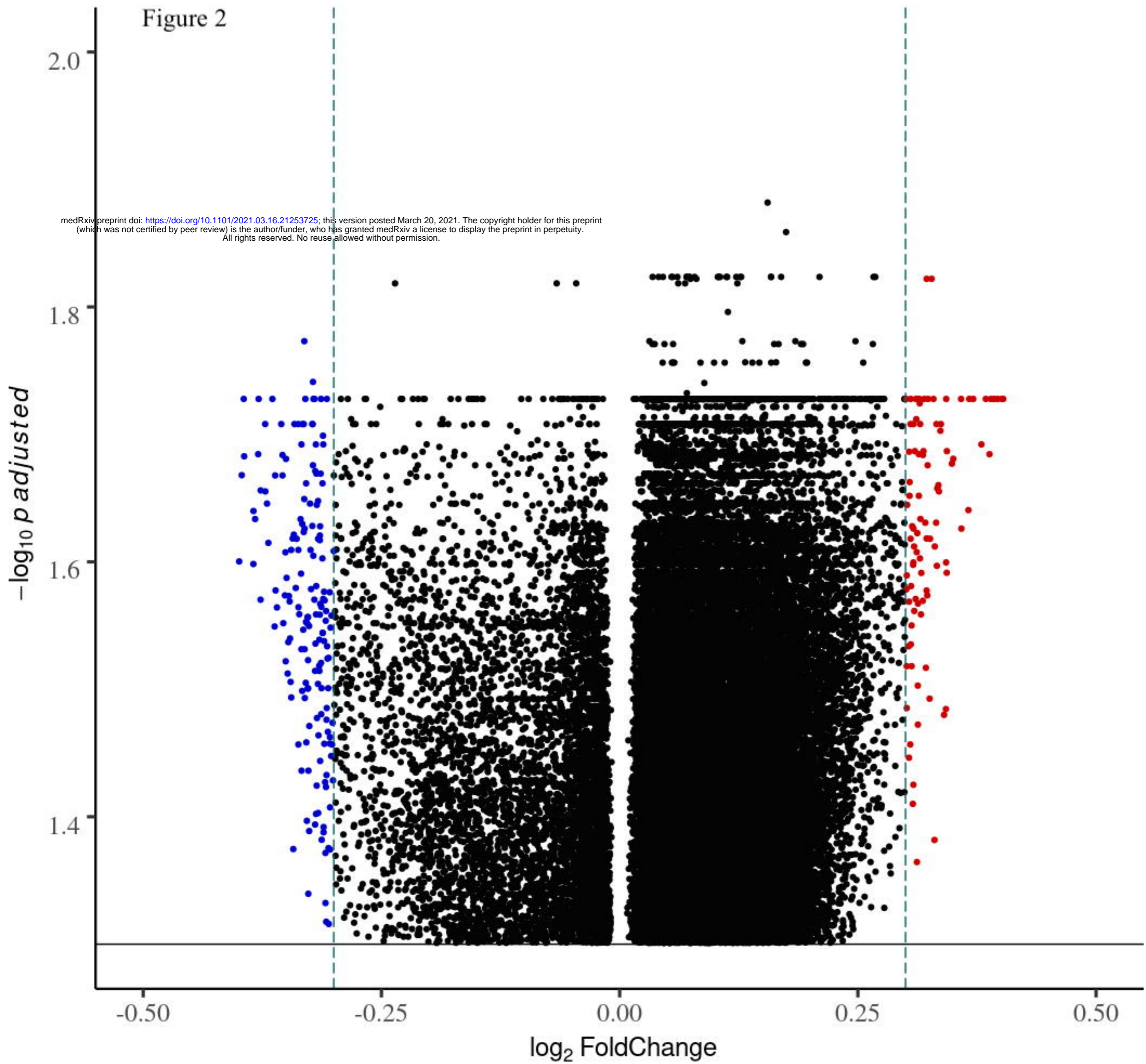




Figure 3a.

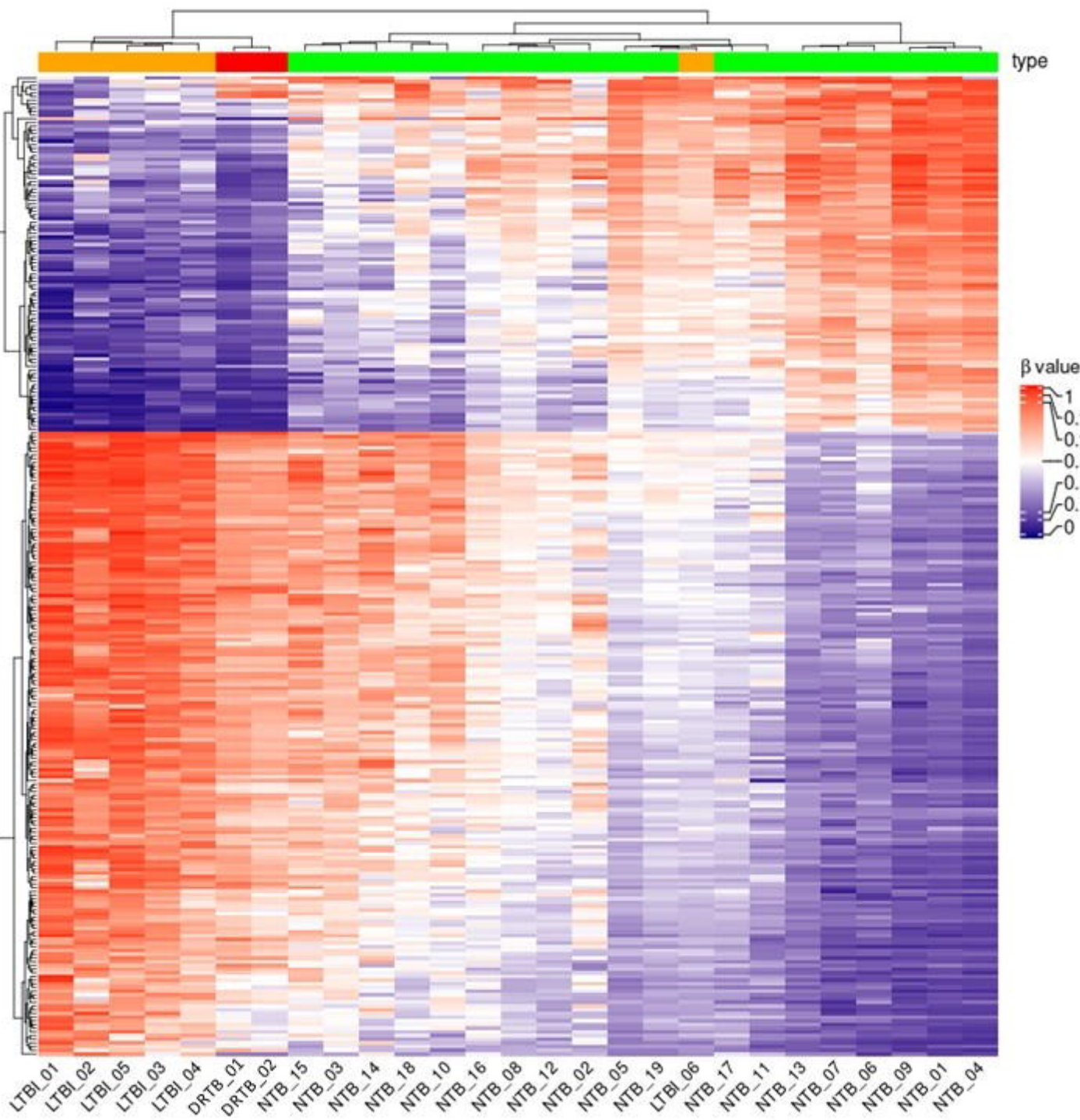


Figure 3b

