

1 **Unravelling the transcriptome of the human tuberculosis lesion and its clinical**
2 **implications**

3 Kaori L. Fonseca^{1,2}, Juan José Lozano³, Albert Despuig^{1,2}, Dominic Habgood-Coote⁴,
4 Julia Sidorova³, Lilibeth Arias^{1,2}, Álvaro Del Río-Álvarez^{5,6}, Juan Carrillo-Reixach^{5,6},
5 Aaron Goff⁷, Leticia Muraro Wildner⁷, Shota Gogishvili⁸, Keti Nikolaishvili⁸, Natalia
6 Shubladze⁸, Zaza Avaliani^{8,9}, Pere-Joan Cardona^{1,2,10}, Federico Martinón-Torres^{2,11,12},
7 Antonio Salas^{2,12,13}, Alberto Gómez-Carballa^{2,12,13}, Carolina Armengol^{5,6}, Simon J
8 Waddell⁷, Myrsini Kafourou^{4,14}, Anne O'Garra^{15,16}, Sergo Vashakidze^{8,17,&}, Cristina
9 Vilaplana^{1,2,10,18, &*}

- 10
11 1. Experimental Tuberculosis Unit (UTE). Fundació Institut Germans Trias i Pujol (IGTP).
12 Edifici Mar. Can Ruti Campus. Crtra. de Can Ruti, Camí de les Escoles, s/n. 08916, Badalona;
13 Catalonia; Spain.
14 2. Centro de Investigación Biomédica en Red de Enfermedades Respiratorias (CIBERES). Av.
15 Monforte de Lemos, 3-5. Pabellón 11. Planta 0. 28029, Madrid, Spain.
16 3. Bioinformatic Platform, Centro de Investigación Biomédica en Red de Enfermedades Hepática
17 y Digestivas (CIBERehd), Hospital Clinic, Barcelona, Spain.
18 4. Department of Infectious Diseases, Imperial College London, London, United Kingdom.
19 5. Childhood Liver Oncology Group, Germans Trias i Pujol Research Institute (IGTP),
20 Translational Program in Cancer Research (CARE), Edifici Muntanya. Can Ruti Campus. Ctra.
21 de Can Ruti, camí de les Escoles, s/n, 08916 Badalona, Catalonia, Spain.
22 6. Liver and Digestive Diseases Networking Biomedical Research Centre (CIBER), 28029
23 Madrid, Spain.
24 7. Global Health and Infection, Brighton and Sussex Medical School, University of Sussex,
25 Falmer, Brighton, BN1 9PX, UK.
26 8. National Center for Tuberculosis and Lung Diseases (NCTLD). 50, Maruashvili Str. 0101
27 Tbilisi, Georgia.
28 9. European University, Tbilisi, Georgia
29 10. Hospital Universitari Germans Trias i Pujol (HUGTIP). Microbiological dept. Crtra. Del
30 Canyet, s/n. 08916. Badalona, Catalonia; Spain.
31 11. Genetics, Vaccines and Pediatric Infectious Diseases Research Group (GENVIP), Instituto de
32 Investigación Sanitaria de Santiago (IDIS) and Universidad de Santiago de Compostela (USC),
33 15706, Galicia, Spain.
34 12. Translational Pediatrics and Infectious Diseases, Department of Pediatrics, Hospital Clínico
35 Universitario de Santiago de Compostela (SERGAS), 15706, Galicia, Spain.
36 13. Unidade de Xenética, Instituto de Ciencias Forenses (INCIFOR), Facultade de Medicina,
37 Universidade de Santiago de Compostela, and GenPoB Research Group, Instituto de
38 Investigación Sanitaria (IDIS), Hospital Clínico Universitario de Santiago (SERGAS), 15706,
39 Galicia, Spain.
40 14. Centre for Paediatrics and Child Health, Imperial College London, United Kingdom
41 15. Immunoregulation and Infection Laboratory, The Francis Crick Institute, London, UK
42 16. National Heart and Lung Institute, Imperial College, London, UK.
43 17. The University of Georgia, Tbilisi, Georgia.
44 18. Direcció Clínica Territorial de Malalties Infeccioses i Salut Internacional. Gerència territorial
45 Metropolitana Nord, Institut Català de la Salut. Can Ruti Campus. Crtra. del Canyet, s/n. 08916,
46 Badalona; Catalonia; Spain.
47
48 & Sergo Vashakidze and Cristina Vilaplana should be considered last co-authors

49
50 ***Corresponding Author:**

51 Cristina Vilaplana, MD, PhD

52 ORCID: <https://orcid.org/0009-0002-2808-7270>
53 NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Principal Investigator

54 Tuberculosis Experimental Unit (UTE)
55 Fundació Institut Germans Trias i Pujol (IGTP)
56 Campus Can Ruti. Crtra. De Can Ruti, Camí de les Escoles, s/n.
57 Badalona, Barcelona, 08916
58 Catalonia (Spain)
59 Telephone: +34 93 033 0527
60 cvilaplan@igtp.cat
61

62 ABSTRACT

63 The granuloma is a complex structure, contributing to the overall spectrum of
64 tuberculosis (TB). We characterised 44 fresh human pulmonary TB lesion samples from
65 13 patients (drug-sensitive and multi-drug resistant TB) undergoing therapeutic surgery
66 using RNA-Sequencing. We confirmed a clear separation between the granuloma and
67 adjacent non-lesional tissue, with the granuloma samples consistently displaying
68 increased inflammatory profile despite heterogeneity. Using weighted correlation
69 network analysis, we identified 17 transcriptional modules associated with granulomata
70 and demonstrated a gradient of immune-related transcript abundance according to the
71 granuloma's spatial organization. Furthermore, we associated the modular transcriptional
72 signature of the TB granuloma with clinical surrogates of treatment efficacy and TB
73 severity. We show that in patients with severe disease, the IFN/cytokine signalling and
74 neutrophil degranulation modules were overabundant, while tissue organization and
75 metabolism modules were under-represented. Our findings provide evidence of a
76 relationship between clinical parameters, treatment response and immune signatures at
77 the infection site.

78 Tuberculosis (TB) is an infectious disease caused by *Mycobacterium tuberculosis*
79 (*Mtb*), and a major cause of ill-health and mortality worldwide¹. Globally, TB
80 chemotherapy is successful in 85% of drug-sensitive (DS) TB cases². Nevertheless, there
81 is a fraction of patients who will fail treatment and are therefore prone to disease relapse
82 and death, especially in multi drug-resistant (MDR) TB cases³. The formation of
83 granulomas is a hallmark of TB and is crucial for containing and controlling the spread
84 of *Mtb* within the host⁴. The existing literature has demonstrated a high degree of
85 heterogeneity in TB granulomatous lesions⁵. Animal studies involving macaques have
86 provided valuable information on granuloma nature and evolution, showing high
87 diversity even within the same host with different grades of bacteria clearance⁶.
88 Moreover, this diversity is observed over the course of infection⁷. These preclinical
89 studies are key to understanding how TB lesions evolve, as human studies cannot provide
90 this information unless using surrogates, such as 18-F-fluorodeoxyglucose positron
91 emission tomography-computed tomography (18F-FDG-PET-CT), as demonstrated by
92 Malherbe et al. when correlating images of TB patients obtained using this method with
93 bacillary load^{8,9}.

94 The development of lung cavitary lesions from a granuloma is crucial in TB
95 pathogenesis and is linked to increased transmission and poor outcomes¹⁰. Human TB
96 lesion lung biopsies are limited and scarce¹¹, and host factors that drive cavitary lesion
97 formation or trigger or reflect poor clinical outcomes remain unknown. The resection of
98 human pulmonary lesions during therapeutic surgery or autopsies has revealed insights
99 into TB lesion architecture, and immunopathology at the site of disease, which may
100 contribute to the emergence of MDR *Mtb* populations^{12,13}. Moreover, mycobacterial
101 culture from resected granuloma tissue demonstrated that a subset of patients still
102 harboured live *Mtb* bacilli despite showing preoperative microbiological clearance in
103 sputa in both, DS- and MDR-TB patients^{12,14}.

104 Subbian et al demonstrated molecular correlation of immune responses to the
105 heterogeneity of granuloma samples from four MDR-TB patients, diversity that the
106 authors linked to lesion maturation¹⁵. Marakalala et al.¹⁶ suggested that the response to
107 *Mtb* might be shaped by the anatomical localization within the granuloma¹⁶. They
108 demonstrated a pro-inflammatory center and an anti-inflammatory surrounding tissue by
109 mass spectrometry and lipid quantification. These authors worked with different types of
110 granulomata from six MDR-TB patients and highlighted the heterogeneity of the
111 lesions¹⁶. Dheda et al. were the first authors to characterize the transcriptional response

112 at anatomically different locations within the granulomas of 14 MDR-TB patients¹¹. They
113 showed the cavity wall as the main source of pro-inflammatory activity compared to the
114 lesion centre. However, none of these human studies included DS-TB patients. Finally, a
115 recent study constructed a spatial cell atlas using 6 patients' samples (two DS-TB and one
116 MDR-TB patient undergoing surgery, and three autopsies) to map granuloma structure
117 and composition and contrast it with the peripheral immune responses¹⁷.

118 To our knowledge, no published TB studies have correlated their results with the
119 clinical characteristics of the patients. Based on the hypothesis that characterising human
120 TB lesions at the transcriptomic level could help us to understand the heterogeneity of
121 granulomas and TB pathogenesis in relation to disease presentation, we determined the
122 modular transcriptome signatures of human TB pulmonary lesions from DS- and MDR-
123 TB patients who underwent surgery. We also investigated their link to clinical and
124 microbiological surrogates of TB severity and treatment response (Fig. 1).

125

126

127

128

129 **RESULTS**

130 **The human TB granuloma signature shows a distinct and heterogeneous
131 transcriptional profile as compared with non-lesional lung tissue**

132 In our study, outlined in Fig. 1, we collected 48 samples from 14 patients and
133 analysed 44 paired samples from 13 patients (7 DS-TB and 6 MDR-TB) to evaluate the
134 human TB lung granuloma transcriptomic changes by RNA sequencing. We analysed
135 total RNA from three different sections: Central Lesion (C; $n=6$), Internal Wall (I; $n=12$)
136 and External Wall (E; $n=13$) collected from each patient's granuloma biopsy. Fewer C-
137 samples could be analysed compared to I and E, due to poorer RNA recovery.
138 Additionally, surrounding non-lesional (NL) tissue from the involved lung was collected
139 as a comparator ($n=13$) (Fig. 2a). Patients were matched according to their sex and *Mtb*
140 drug-sensitivity classification to avoid potential confounding factors (Supplementary
141 Table 1). Moreover, clinical and demographic data and, resected TB lesion characteristics
142 of each participant were assessed at the time of surgery (Supplementary Tables 1 and 2).

143 We found a total of 4,630 significantly differentially expressed genes (DEGs),
144 using DESeq2 with adjusted $p\leq 0.05$ (Extended data Fig. 1a and Supplementary Data Set
145 1). Of these, 2,496 genes were over-expressed in lung granuloma tissues, whereas 2,134
146 were under-expressed, as compared to NL lung tissue (Extended data Fig. 1a). The top
147 40 ranked DEGs clearly separated granuloma samples from NL lung samples (Fig. 2b),
148 showing distinct transcriptional profiles for the two tissues. Among them, genes involved
149 in immune system/cytokine signalling (*IRF4*, *CCL19*, *LTB*, *JAK3*, *INPP5D*, *FCER2*,
150 *MMP1*) and B cell activation and differentiation (*CD22*, *BLNK*, *CARD11*) were over
151 expressed, suggesting an inflammatory signature in the granuloma. Seven granuloma
152 lesion samples clustered with NL tissue samples pointing towards heterogeneity in the
153 granuloma transcriptional profiles (Fig. 2b).

154 Altogether, our data show a distinct segregation of the granuloma lesion when
155 compared to the NL lung tissue with respect to an inflammatory profile, as previously
156 proposed¹¹. Our findings also indicated a diverse range of molecular diversity within the
157 granuloma samples, prompting our decision to delve deeper into the heterogeneity at a
158 transcriptional level.

159

160

161

162

163 **Compartments within the granuloma reveal distinct gene expression profiles with
164 an enriched inflammatory response across the lesion**

165 To further explore the granuloma heterogeneity and investigate the contribution
166 of each granuloma compartment to the TB lesion architecture, we first performed an
167 enrichment analysis derived from single sample Gene Set Enrichment analysis (ssGSEA)
168 using the top 40 DEGs discriminating the granuloma from the NL lung tissue. The
169 expression of these genes in the different tissue compartments revealed a more
170 pronounced enrichment score of these DEGs in central and internal lesion samples,
171 suggesting that these two compartments might be the main contributors for the overall
172 granuloma transcriptional signature (Fig. 3a).

173 Next, we compared the expression profiles derived from each granuloma
174 compartment with the NL tissue. The list of DEGs (DESeq2 with adjusted $p \leq 0.05$) for
175 the C, I and E vs NL tissue comparisons respectively constituted 3,228 (1,539 genes were
176 over-expressed, whereas 1,689 were under-expressed); 5275 (2,676 over-expressed and
177 2,599 under-expressed); and 1,045 genes (552 over-expressed and 493 under-expressed)
178 (Supplementary Data Set 1 and Extended data Fig. 1b). For central and internal
179 compartments, the hierarchical clustering of the 40 most significantly DEGs showed an
180 evident separation when compared each compartment against the NL lung tissue (Fig.
181 3b). Though less noticeable, the external compartment was still distinguishable from the
182 NL tissue. Therefore, the magnitude of differential expression relative to NL decreased
183 gradually towards the edge of the granuloma structures (Fig. 3b).

184 Among the highly variable genes in central lesion, we found genes involved in the
185 immune system/cytokine signalling (*CCL19*, *CXCL10*) to be upregulated in comparison
186 to the NL tissue (Fig. 3b). On the other hand, we found extracellular matrix organization-
187 related genes to be downregulated (*LRP4*, *MUC15*), while others were upregulated
188 (*ADAM12*, *CTSK*). Moreover, collagen-encoding genes (*COL1A1*, *COL3A1*, *COL11A1*)
189 were upregulated in the central compartment, which could reflect collagen turnover or
190 fibrosis; as well as of genes associated to immunoglobulin heavy and light chains
191 (*IGHV4-61*, *IGKV1-39*, *IGKVID-39*, *IGHVI-18*, *IGHV3-74*, *IGLV1-40*,
192 *IGHV4-34*, *IGHVI-3*, *IGLV1-51*, *IGLV2-18*, *IGHV6-1*, *IGHV4-55*), related to
193 humoral immunity (Fig. 3b). Furthermore, genes involved in complement fixing (*C1QA*,
194 *C1QB*, *C1QC*) were significantly upregulated, although not among the top 40 DEGs. For
195 the internal compartment, genes involved in the immune system/cytokine signalling
196 (*LTB*, *FCMR*, *AIM2*, *CXCL10*, *IRF8*, *IRF4*) were upregulated compared to NL (Fig. 3b).

197 Furthermore, immune system/cytokine signalling genes (*LTB*, *CCL19*, *CXCL9*,
198 *TNFAIP3*, *TNFRSF13C*, *FCMR*, *AIM2*, *CXCL10*) were over-expressed in the external
199 compartment relative to NL (Fig. 3b), evidencing an inflammatory signature throughout
200 the granuloma lesion.

201 We then applied weighted gene co-expression analysis (WGCNA) to perform a
202 modular analysis of co-expressed genes in the granuloma and in the three compartments
203 separately, comparing all samples to NL control tissues. We identified 17 modules from
204 co-expression networks related to the human TB whole granuloma (Fig. 3c,
205 Supplementary Data Set 2). The identified granuloma modular signature showed that
206 neutrophil degranulation, cell signalling, humoral immunity, extracellular matrix,
207 interferon/cytokine signalling, and innate/pathogen recognition receptors (PRR) modules
208 were overabundant. These observations were consistent throughout the compartments,
209 except for the neutrophil degranulation and innate/PRR modules, which were apparent in
210 the total granuloma and the internal lesion only, but not in the central or external lesions
211 (Fig. 3c). Conversely, the Epithelial to Mesenchymal Transition (EMT), cholesterol
212 biosynthesis and metabolism modules were found to be under-abundant in the whole
213 granuloma and external and internal, but not in the central compartment. In addition to
214 cholesterol and metabolism, the organelle biogenesis module was underrepresented,
215 across all compartments, suggesting that some pathways present in the healthy lung are
216 diminished in granuloma tissue (Fig. 3c).

217 In summary, our results showed a significant enrichment of modules related to
218 inflammation, including pathways of innate immunity in the TB granuloma and the
219 internal compartment, and of humoral immunity across all compartments. Meanwhile a
220 decrease in modules related to extracellular matrix organisation and cholesterol
221 biosynthesis and metabolism was observed in granuloma tissues.
222

223 **Patients' clinical status is associated with differential modular transcriptomic 224 profiles in TB lesions**

225 The heterogeneity in the host immune response to infection, considering the
226 involvement and contribution of physically distinct compartments, together with the
227 bacteria and the inflammatory environment, defines granuloma fate and disease
228 manifestation^{16,18}. Hence, we next aimed to associate the modular signature changes in
229 the granulomata (considering the three compartments together) from TB patients with
230 surrogates of treatment response and disease severity. As surrogates of response to

231 treatment we used: achieving stable sputum culture conversion (SCC) before or later than
232 2 months after the start of anti-TB treatment (fast or slow converters); drug sensitivity of
233 the infecting *Mtb* isolate; whether the individual is a TB relapse or new TB case; and the
234 number of pulmonary lesions present (Supplementary Data Set 2). As surrogates of
235 disease severity, we used presence of symptoms assessed by the St. George's Respiratory
236 Questionnaire (SGRQ) symptoms score, stratifying individuals according to SGRQ score
237 > 20 or < 20. We quantitatively tested the association of each clinical parameter with each
238 of the significant module's eigengene (ME) expression patterns (Wilcoxon $p \leq 0.05$).

239 Regarding the SCC, the modular signature of the granuloma revealed a significant
240 association of DNA binding and interferon/cytokine signalling modules with SCC, with
241 the enrichment of these modules being significantly higher in those individuals
242 converting the sputum culture later (FDR < 0.1; Figs. 4a and b). No significant modular
243 expression was found to be associated with *Mtb* drug sensitivity of the individuals,
244 relapsed or new cases, or number of lesions (data not shown).

245 When considering severity of TB disease, we found that DNA binding, neutrophil
246 degranulation, interferon/cytokine signalling, cholesterol biosynthesis and myeloid
247 activation modules were significantly overabundant and associated with higher SGRQ
248 symptoms score (FDR < 0.1; Fig. 4a and c), pointing to higher inflammation status with
249 more severe disease manifestation. In contrast, the EMT module was significantly
250 underabundant in these individuals' granulomas (Fig. 4c).

251 Our data from the granulomas of this patient cohort suggest that those patients
252 exhibiting slow SCC, which reflects a poorer response to treatment and slower clearance
253 of *Mtb*, were associated with greater lung inflammatory responses, mainly driven through
254 the interferon/cytokine signalling axis. Additionally, the granuloma signature in patients
255 with the most severe TB cases was also associated with an increased inflammatory
256 response.

257

258 **DISCUSSION**

259 The immune response to *Mtb* constitutes a complex and dynamic interaction
260 between the host immune system, the bacteria and lung microenvironment. Throughout
261 infection, the inflammatory response leads to granuloma formation primarily for bacterial
262 containment, while causing extensive tissue remodelling and destruction, which
263 contributes to the clinical spectrum of TB¹⁹. In our study, we transcriptionally
264 characterised the host response in human pulmonary TB lesions from patients undergoing
265 therapeutic surgery. Fresh human TB lesion specimens obtained from these lung
266 resections (without formalin fixation or paraffin embedding) were transcriptionally
267 profiled using RNA-Seq. Our study provides various advances over previous
268 approaches^{15,20,21}. Firstly, we have used a more robust data set with increased patient
269 numbers, which included 44 TB granuloma samples from 13 DS- and MDR-TB patients
270 from the SH-TBL cohort. Secondly, we confirmed a distinct demarcation between
271 granuloma and adjacent non-lesional tissue from the same patient lung. Thirdly, and to
272 our knowledge for the first time, we have identified a transcriptional modular signature
273 within granulomata and linked our findings to clinical/microbiological parameters used
274 as surrogates of TB severity and response to treatment. In patients with more severe
275 disease, our results showed an increased eigengene expression of pro-inflammatory-
276 related modules and a decreased eigengene expression of tissue organization modules.
277 Strikingly, those patients with a delayed response to treatment showed an increased DNA
278 binding and interferon/cytokine modular granuloma response, which has major
279 implications for modifying treatment and subsequent clinical management of disease.

280 Granuloma heterogeneity in TB is a well-accepted concept and has been reported
281 in non-human primate models of infection and human lesions^{16,22}. Besides the
282 heterogeneity among samples, we were able to identify a clear pattern across all TB
283 patients compared to their own NL control lung tissue. Among the top 40 DEGs, we found
284 genes predominantly encoding proteins related to the inflammatory processes that
285 orchestrate the antimycobacterial response such as *CCL19*, *LTB*, *FCMR*, *MMPI* and
286 *IRF4*²³⁻²⁷. *CCL19* gene expression was found to be increased in mouse lungs post-*Mtb*-
287 infection to induce lymphoid structures²³. Regarding human studies, previously published
288 blood microarray-derived signatures found up-regulated levels of *LTB*²⁸ when comparing
289 baseline to end-of-treatment samples from TB patients. Moreover, MMP-1 was found to
290 be increased in the respiratory secretions from TB patients and to drive extracellular
291 matrix remodelling in a TB murine model²⁶. Recently, *MMPI* was also found to be

292 differentially expressed in TB lymph node biopsies compared to control samples in a
293 study by Reichmann et al²⁹. Furthermore, *FCMR* has been considered a target for host-
294 directed therapies²⁵, while the transcription factor *IRF4* was found to be required for the
295 generation of Th1 and Th17 subsets of helper T cells and follicular helper T-like cellular
296 responses²⁷. In samples from our patient cohort, we found those genes to be over-
297 expressed in the granuloma lesion of patients undergoing surgery for unresolved TB.
298 Additionally, the overexpression of immunoglobulin genes in the granuloma suggests
299 their involvement in complement fixation processes, since *C1QA*, *C1QB* and *C1QC*
300 transcripts were also found to be upregulated in our granuloma samples. Previously
301 published blood signatures found up-regulated levels of *C1QC*^{30,31}, when comparing
302 baseline to end-of-treatment samples. Moreover, the expression of this gene has been
303 proposed as a disease severity biomarker³² and linked to poor treatment response³³.
304 Collectively, these findings suggest that information on the host response within the
305 granulomata from lung resections from critically ill patients, may inform the clinical
306 management of disease.

307 We explored the granuloma heterogeneity by analysing DEGs between different
308 compartments and applied WGCNA modular analysis to the whole granuloma and
309 separated compartments. In line with Dheda's and Marakalala's studies^{11,16}, we showed
310 that both the centre and internal compartments of the TB granuloma lesion present a more
311 perturbed transcript abundance compared to the external lesion compartment, potentially
312 pointing to the trajectory of the host response to *Mtb*. Our results confirm the involvement
313 of pro-inflammatory-related genes and modules in the human TB granuloma and
314 compartments. The spatial differences within the granuloma in terms of gene expression
315 are supported by other studies^{11,15}. Dheda et al. described the pathways involved in
316 different parts of cavitary lesions from 14 failed MDR-TB subjects that underwent
317 surgery, pointing to the cavity wall as the main source of pro-inflammatory activity¹¹. In
318 line with our findings, they showed that pro-inflammatory pathways were especially over-
319 represented in the cavity wall, including nitric oxide production, reactive oxygen species,
320 IL-1, IL-6, IFN- γ and NF- $\kappa\beta$ transcriptional signatures. Furthermore, Subbian et al.
321 demonstrated using four granuloma samples, the involvement of immune cell signalling
322 and activation, interferon response and tissue remodelling processes in the complex TB
323 granuloma microenvironment¹⁵. The modular granuloma signature that we describe
324 herein, provides comparable and additional data, albeit in a larger and independent patient
325 cohort.

326 We identified 17 modules from co-expressed networks and mapped a TB
327 granuloma modular signature, consisting of increased neutrophil degranulation, cell
328 signalling, humoral immunity, extracellular matrix, interferon/cytokine signalling and
329 innate/PRR. In our cohort, patients presented advanced TB disease with cavitary TB. We
330 found the neutrophil module to be increased in whole granuloma but not in the central or
331 external lesions, possibly explained by more necrosis in this region, coupled with
332 relatively low RNA abundance in neutrophils. Moreover, the *MMP1* gene was
333 overexpressed in the granuloma as compared to NL tissue, which might suggest the
334 involvement of neutrophils through the activity of matrix metalloproteinases. In humans,
335 neutrophil accumulation in the lungs of active TB patients and has been correlated with
336 increased lung pathology and consequent disease progression^{34,35}. The role of neutrophils
337 in TB disease progression and pathology has been well documented in experimental
338 mouse models³⁶⁻³⁹. Additionally, the extracellular matrix, interferon/cytokine signalling
339 and innate/PRR modules have been reported to be upregulated in blood from TB
340 patients^{40,41}, demonstrating that whatever is happening at the site of infection can also be
341 translated to the periphery. Interestingly, we found that the humoral module was increased
342 in whole granuloma samples, corroborating with the expression of immunoglobulin
343 heavy and light chains transcripts in both central and internal compartments. Our findings
344 are in keeping with a report from Krause et al., demonstrating the presence of different B
345 cell subsets and high levels of *Mtb*-reactive antibodies in human lung tissue⁴².

346 Conversely, the EMT, cholesterol biosynthesis and metabolism modules were
347 decreased in granuloma relative to NL parenchyma. The EMT is linked to wound healing
348 but also to fibrogenesis and scarring⁴³. The downregulation of this module in granuloma,
349 along with its decreased enrichment in more severe TB cases compared to those with
350 milder symptoms, might suggest a disruption in critical processes needed for tissue repair.
351 We also observed increased cholesterol synthesis in individuals experiencing more severe
352 disease in terms of presenting more pronounced symptomatology. Kim et al. proposed
353 dysregulation of host lipid metabolism caused by *Mtb*, tracing the progression of TB
354 granulomas to caseation, cavitation, and eventual disease transmission²⁰. The authors
355 suggested that bacterial components could trigger the host's innate immune system,
356 potentially augmenting the synthesis or storage of host lipids. Consequently, in line with
357 these findings²⁰ the upsurge in cholesterol synthesis which we observe herein might
358 mirror the impact of *Mtb* on the host lipid metabolism.

Indeed, a major novelty of our work is the use of unbiased modular analysis to link the transcriptional signature generated from TB granuloma lesions with patients' clinical surrogates of TB severity and the time taken to clear *Mtb* in sputum, and thus response to treatment. Specifically, our results show an important inflammation component in granulomas from patients presenting with greater severity of disease and slower response to treatment. Inflammation has been described as key for tissue damage and linked to a blood transcriptional signature in individuals suffering from active TB disease even months before being diagnosed⁴⁴, and radiographic lung disease extension^{7,36,40}, decreasing upon treatment^{40,45}. Tabone et al. revealed differential responses in the blood transcriptional signature among various clinical TB subgroups following treatment, observing a reduction in the inflammation and IFN modules alongside B and T cell modular signatures accompanying successful treatment⁴¹. This observation raises the logical assumption that changes induced by treatment at the blood level might parallel alterations occurring at the infection site. In our study, we found an overabundance of the IFN/cytokine signalling and DNA binding modules associated with severe disease, characterized by worsened symptoms and slower bacterial clearance. Our observations also hint at a potential connection between heightened neutrophil degranulation in severe cases and the damaging mechanisms associated with neutrophil-mediated inflammation⁴⁶, suggesting a plausible role for this process in exacerbating the severity of the condition. As our study samples were collected after the end of treatment, all cases examined here could be considered difficult or inadequate responders to treatment. This inadequate response may result in a sustained pro-inflammatory profile at the granuloma site or be the consequence of it, and our findings may thus help in future management of disease treatment. Both our data and data from other sources collectively offer the opportunity to use these transcript patterns to monitor treatment response.

Our results suggest that the clinical picture mirrors the inflammation happening at the site of infection and confirms what has been previously seen by others indirectly, both in humans and in experimental animal models. Malherbe et al., showed through 18F-FDG-PET-CT lung scans that some patients still have an increased FDG uptake in the lesions when compared to surrounding healthy tissue after six months of treatment⁸, and more recently, the authors have related both a larger burden of disease and a slower rate of reduction in scan metrics with delayed sputum conversion⁴⁷. Our data showed that slow sputum converters present different modular expression profiles in TB granulomas when compared to fast sputum converters. To date, the SCC constitutes the only tool endorsed

393 by the WHO to monitor treatment response⁴⁸ and can be considered a surrogate of disease
394 severity. Therefore, achieving SCC after two months of starting treatment has been
395 associated with TB cavity persistence¹⁴ and poor prognosis^{49,50}. In our cohort, the
396 proportion of DS-TB/MDR-TB cases in the fast SCC converters was higher (5/8) than in
397 slow converters (4/6), as previously described⁴⁸. Importantly, from our findings, the
398 analysis of granulomatous tissue in resected lung samples from patients suggests that the
399 slower SCC converters exhibit the highest expression of inflammatory genes in the lesion
400 after several months, supporting the potential use of SCC at the two-month mark after
401 treatment initiation as a prognostic indicator. On the other hand, time to sputum clearance
402 could enhance treatment monitoring and refine clinical management and might assist in
403 identifying patients who should be prioritized for further therapeutic interventions.
404 Interesting, measures of microbiological treatment success and clinical severity of disease
405 have also been associated with *Mtb* transcriptional profiles in patient sputa⁵¹, suggesting
406 that the lesion immunopathology described here also impact *Mtb* lung phenotypes.

407 Our study has some limitations. These TB individuals presented advanced TB
408 disease with cavitary TB, and underwent lung resection surgery, so the results may not
409 be generalisable to the whole spectrum of TB disease. In countries with a high prevalence
410 of MDR-TB, adjunctive surgical resection is a common therapeutic tool which, despite
411 being a major invasive procedure, reduces the transmission burden of MDR-TB and
412 results in favourable outcomes for the patients⁵². However, this approach is uncommon
413 in most countries, thus our results help to understand TB host response but may have a
414 direct impact on TB treatment at short term only in high burden countries where resection
415 is practised. Furthermore, given that individuals with TB may have several lesions at
416 varying stages, which can evolve and recede (as shown in experimental animal
417 studies^{4,7,53,54}), expanding the sample size to include several lesions from the same
418 individuals would be beneficial. However, achieving this is practically unfeasible without
419 conducting a complete pulmonectomy or lung section resection. Consequently, working
420 with samples collected *post-mortem* could offer a viable solution, offering substantial
421 insights and information in this regard, although this is limited by the number of TB
422 patients from whom *post-mortem* samples would be available and the quality of the
423 samples collected.

424 In conclusion, we have defined a robust signature for human advanced TB
425 granuloma lesions, despite the inter granulomata heterogeneity. Moreover, this is the first
426 study showing different modular transcriptomic signature patterns, integrating and co-

427 analysing our findings with TB patients' clinical/microbiological measurements,
428 including severity and response to treatment. Our study provides a considerable dataset
429 on TB granulomas gene expression which will undoubtedly be of broad utility, interest
430 and significance to the scientific community, contributing to an increase in knowledge on
431 TB immunopathology. A better understanding of disease processes and host protective
432 immune responses may help in the clinical management of TB and development of
433 treatment strategies. Most importantly, our findings provide evidence of the clinical
434 picture with a relationship between clinical parameters, treatment response and immune
435 signatures at the infection site.

436

437 **METHODS**

438 **Ethics**

439 This study is part of the SH-TBL project (ClinicalTrials.gov Identifier: NCT02715271).
440 The protocol, research methodology and all associated documents (informed consent
441 sheet, informed consent form) were reviewed and approved by both ethics' committees
442 at the National Center of Tuberculosis and Lung Diseases (NCTLD) (IRB00007705
443 NCTLD Georgia #1, IORG0006411) and the Germans Trias i Pujol University Hospital
444 (EC: PI-16-171). Written informed consent was obtained from all study participants
445 before enrolment.

446

447 **Study design and patient cohort**

448 Study participants constituted the SH-TBL cohort, a cross-sectional study recruiting 40
449 adult patients was conducted in the NCTLD in Tbilisi (Georgia), from May 2016 to May
450 2018, after receiving therapeutic surgery indication for their pulmonary TB. All
451 volunteers received standard anti-TB treatment (ATT) regimen according to Georgia
452 national guidelines and achieved sputum culture conversion. Surgery was required due to
453 persistent radiological signs of cavitary lesions on chest X-Ray and computed
454 tomography scan, according to the Georgian National Guideline "Surgical Treatment of
455 Patients with Pulmonary Tuberculosis" based on official guidelines of surgery
456 recommendations for pulmonary TB. Thoracic surgery decisions were made by the
457 NCTLD Resistant Tuberculosis Treatment Committee, composed of two surgeons and 18
458 pulmonary TB specialists.

459

460 **Data and sample collection**

461 Anonymised data regarding the socio-epidemiological factors, clinical aspects, and
462 information referring to the current TB episode for the SH-TBL cohort were collected
463 using an electronic case report form. TB individuals were classified taking into
464 consideration clinical surrogates of disease severity. According to their sputum
465 conversion date, individuals were categorized as fast (SCC < 2 months) or slow
466 converters (SCC > 2 months). Drug sensitivity of the *Mtb*-infecting isolate was recorded,
467 classifying TB cases as DS or MDR-TB. Presence of symptoms was assessed with the
468 SGRQ symptom score, and a cut off defined by the median SGRQ symptoms value
469 (SGRQ cut off = 20) was used to stratify the individuals: those with an SGRQ score > 20

470 were considered to have more severe TB disease than those with SGRQ score < 20.
471 Relapsed cases and number of lesions were also considered (Supplementary Data Set 2
472 and supplementary Tables 1 and 2). During surgery, TB granuloma lesions of 3.2 cm in
473 diameter (median; 2.2-5.5 cm) were removed according to surgical criteria. In total, 48
474 SH-TBL biological samples were collected for the study. Biopsies from three different
475 sections of TB lesion were collected, namely Central Lesion (C), Internal Wall (I),
476 External Wall (E). In addition, surrounding non-lesional (NL) lung parenchyma tissue,
477 unaffected, by eye and by palpation, was collected from the same patient (Fig. 2a). Biopsy
478 fragments (~0.5 cm³) of tissue samples from 14 patients were collected for RNA-Seq.
479 Fresh tissue samples were incubated in RNAlater solution (Qiagen) at 4°C overnight,
480 before storage at -80°C. Samples were processed in BioSafety Level 3 (BSL-3)
481 laboratory.

482

483 **Total RNA extraction**

484 For an optimal RNA recovery, TB lesion biopsy samples were divided into 0.16-0.21g
485 single pieces and placed into new tubes. Samples were reduced to powder by mechanical
486 cryofracturing using a BioPulverizer device (Biospec Products) after being cooled in
487 liquid nitrogen. The powdered tissue was then transferred to 2 mL Lysing Matrix D tubes
488 together with lysis solution for homogenization by FastPrep® instrument (MP
489 Biomedicals). RNA was purified using the mirVana miRNA Isolation Kit (Thermo Fisher
490 Scientific), followed by genomic DNA digestion using the DNA-free DNA Removal Kit
491 (Thermo Fisher Scientific) according to manufacturer's instructions. Quantitative and
492 qualitative RNA integrity number equivalent (RINe) values were obtained by Agilent
493 Bioanalyzer 2100 (Agilent Technologies). In general, a standard RINe score for good
494 quality RNA is set at seven. Considering the source and status of the human TB lesion
495 samples, a minimum RINe cut-off of four was established, accepting that low RNA
496 integrity is a potential source of bias in RNA-Sequencing (RNA-Seq) experiments.

497

498 **RNA-Sequencing library preparation, sequencing, and gene alignment**

499 Purified RNA was diluted to 25 ng/μl per aliquot and then shipped on dry ice to Macrogen
500 (Seoul, South Korea), where the RNA-sequencing (RNA-Seq) was performed. Libraries
501 were constructed using the TruSeq Stranded Total RNA LT Sample Prep Kit (Human
502 Mouse Rat) (Illumina) following the TruSeq Stranded Total RNA Sample Prep guide
503 (Part #15031048 Rev. E), including prior removal of ribosomal RNA using the RNA

504 Ribo-Zero rRNA Removal Kit (Human/Mouse/Rat) (Illumina). RNA-Seq was performed
505 on an Illumina platform HiSeq 4000 (Illumina), at 50 million reads per sample, 100bp
506 stranded paired-end reads. Pre-processing of raw data included quality control through
507 FastQC (v.011.7) and MultiQC (v.1.9)⁵⁵. Before further steps in read pre-processing,
508 Illumina adapters were trimmed off with Trimmomatic (v.0.39)⁵⁶. The human genome
509 sequence GRCh38.89 and human gene annotations were downloaded from the
510 ENSEMBL web repository. Files from each sample were aligned to the human reference
511 genome using the Spliced Transcripts Alignment to a Reference (STAR) package
512 (v.2.7.5b)⁵⁷, with the built-in gene counts quantification mode for stranded RNA-Seq
513 data. BAM files were generated, and the SAMtools package applied to calculate the
514 percentage of successful read alignment against the reference human genome (v.1.10)⁵⁷

515

516 **RNA-seq data analysis**

517 The overall pipeline for data handling, plotting and statistical analysis was conducted in
518 R (v.3.6.2). After STAR mapping, a gene count data table was obtained including C, I, E,
519 and NL samples. Genes with a lower than 50 counts among all the samples were discarded
520 to avoid confounding the differential gene expression analysis, as they had low expression
521 to be reliably quantified. Paired statistical analyses were done globally and separately for
522 each compartment. The set for the RNA-Seq experiments comprised 44 paired samples
523 from 13 patients. Samples from patient SH-TBL03 weren't taken into consideration for
524 the RNA-Seq analysis as the NL tissue sample control was missing. The differential
525 expression analysis from tissue count tables was conducted using the DESeq2
526 Bioconductor package (v.1.28.1)⁵⁸. Genes were considered as significant DEGs when the
527 Benjamini-Hochberg adjusted *p*-value was equal to or less than 0.05 (*p* ≤ 0.05). The R
528 package heatmap (v.1.0.12) was used to generate heatmaps and dendograms for the
529 genes and samples by hierarchical clustering after DESeq2 depth normalization.
530 Heatmaps describe the Euclidean distances between samples.

531

532 **Enrichment score for the different tissue compartments**

533 The expression across compartments of upregulated selected genes differentiating
534 granuloma from non-lesional tissue was performed using single sample Gene Set
535 Enrichment analysis (ssGSEA). ssGSEA is a variation of the GSEA algorithm that instead
536 of calculating enrichment scores for groups of samples and sets of genes, it provides a
537 score for each sample and gene set pair⁵⁹.

538

539 **Weighted gene co-expression network analysis and functional annotation**

540 Weighted gene co-expression network analysis (WGCNA) was performed to identify
541 modules using the R package WGCNA. The TB granuloma modules were constructed
542 using the 10,000 most variable genes across all TB samples collected (log2 RNA-seq
543 expression values). To satisfy the scale-free topology criteria and the recommendations
544 for WGCNA use, we chose an optimal soft-threshold ($\beta = 22.5$) to obtain an adjacency
545 matrix from a signed weighted correlation matrix containing pairwise Pearson
546 correlations, generating the corresponding topological overlap measure (TOM). To detect
547 the modules, we applied a dynamic hybrid tree-cut algorithm to detect the computed
548 modules of co-expressed genes (minimum module size of 20, and deep split = 1). Finally,
549 21 modules were obtained. An additional “grey” module was identified in TB granuloma
550 modules, consisting of genes that were not co-expressed with any other genes. The grey
551 module was discarded from further analysis. Moreover, only modules with more than 40
552 genes were annotated. We computed their intramodular connectivity and selected the top
553 five most interconnected genes⁶⁰. Significantly enriched Gene Ontology and canonical
554 pathways from the MSigDB website⁶¹ were computed using clusterProfiler R package⁶².
555 Modules were annotated based on representative biological processes from pathways and
556 processes from all three reference databases. Fold enrichment for the WGCNA modules
557 was calculated using the quantitative set analysis for gene expression with the
558 Bioconductor package QuSAGE⁶³. To identify the modules of genes over- or under-
559 abundant in TB granuloma, compared to the respective non-lesional lung tissue using
560 log2 expression values using the three compartments. Only modules with enrichment
561 scores with FDR *p*-value < 0.1 were considered significant.
562

563

Association between modules and clinical characteristics

564 The following parameters were used as surrogates of TB severity: presence of symptoms
565 assessed by the SGRQ symptoms score; achieving the SCC before or later than 2 months
566 after the start of anti-TB treatment; DS vs MDR-TB case; being a relapse or new TB case;
567 and number of lesions present. We computed the eigengene for each module, defined as
568 the first principal component of the module representing the overall expression level of
569 the module. The relationship of the transcriptomic modules with clinical surrogates of TB
570 severity (SCC and SGRQ symptoms score) was tested using Wilcoxon-rank sum test.
571 Nominal *p*- values were adjusted using the Benjamini-Hochberg approach⁶⁴.

572 **ACKNOWLEDGEMENTS**

573 The authors would like to thank the patients who agreed to participate in the study and
574 the staff from the NCTLID who helped with the SH-TBL project. The authors thank the
575 support from Comparative Medicine and Bioimage Centre of Catalonia (CMCiB),
576 namely at BSL 3 facility, Eric Garcia for technical support in acquiring the data. The
577 authors, thank William J. Branchett, from The Francis Crick Institute, London, for his
578 input on the manuscript and providing suggestions/changes for improvement.

579 **FUNDING**

580 This work was supported by: 1) the Plan Nacional I + D + I co-financed by ISCIII-
581 Subdirección General de Evaluación and Fondo-EU de Desarrollo Regional (FEDER)
582 through PI16/01511, PI20/01424, CP13/00174, CPII18/00031 and CB06/06/0031. 2) The
583 European Union's Horizon 2020 research and innovation program under grant agreement
584 No 847762. 3) The Catalan Agency for Management of University and Research Grants
585 (AGAUR) through 2017SGR500, 2021 SGR 00920 and 2017 FI_B_00797. 4) The
586 "Spanish Society of Pneumology and Thoracic Surgery" (SEPAR) (16/023). 5) The
587 Wellcome Trust, the Medical Research Foundation grants (206508/Z/17/Z and MRF-160-
588 0008-ELP-KAFO-C0801) and the NIHR Imperial College BRC to M.K. 6) Wellcome
589 Trust (204538/Z/16/Z) and National Centre for the Replacement, Refinement and
590 Reduction of Animals in Research (NC3Rs) (NC/R001669/1) grants to S.J.W. 7) Á.D.R.-
591 Á and J.C-R were supported by AGAUR grant 2022_FI_B00528 and 2019 FI_B01024,
592 respectively. 8) CA received funding from CIBERehd (CB06/04/0033) and AGAUR
593 (2017-SGR-490 and 2021 SGR-01186) and was supported by Ramón y Cajal grant of the
594 Ministry of Science and Innovation of Spain (RYC-2010-07249). 9) AOG is supported
595 by The Francis Crick Institute which receives its core funding from Cancer Research UK
596 (FC001126), the UK Medical Research Council (FC001126), and the Wellcome Trust
597 (FC001126); before that by the UK Medical Research Council (MRC U117565642).

598 **AUTHOR CONTRIBUTIONS**

599 Conception: CV and SV. Design of the work: CV and SV, with substantial contributions
600 of MK and AOG. KLF, AD, JJL, JS, LA, DHC, Ad R, JCR, AG, LMW, PJC, SV, SG,
601 KN, NS and ZA worked on the acquisition and analysis of data for the work, and all
602 authors made substantial contributions to its interpretation. RNA-Sequencing library
603 preparation, sequencing, gene alignment and initial bioinformatics analysis was
604 performed by AD and DHC, supervised by MK; and paired statistical analysis, single
605 sample Gene Set Enrichment analysis, WGCNA analysis and the association between
606 modules and clinical characteristics were performed by KLF and JJL, supervised by
607 AOG. KLF and CV drafted the work; and CA, FMT, AS, AGC, SV, SW, MK and AOG
608 reviewed it critically for important intellectual content. All authors revised and gave their
609 final approval of the version to be published and agreed to be accountable for all aspects
610 of the work in ensuring that questions related to the accuracy or integrity of any part of
611 the work are appropriately investigated and resolved.

612

613 **COMPETING INTERESTS**

614 The authors declare the following competing interests: C.V salary is covered by ISCIII
615 (CPII18/00031) and Fundacio Institut d'Investigació en Ciencies de la Salut Germans
616 Trias I Pujol (IGTP). C.V is an unpaid board member of the following non-profit
617 organizations: the FUITB foundation and the ACTMON foundation. Neither the ISCIII,
618 the IGTP, the FUITB nor ACTMON have had any role in the conceptualization, design,
619 data collection, analysis, decision to publish, or preparation of the manuscript.

620

621 **DATA AVAILABILITY**

622 Patient's data were transferred to a data frame (SH-TBL Cohort metadata) and are freely
623 accessible (doi: 10.17632/knhvdbjv3r.1). All metadata and sequencing data from this
624 manuscript have been deposited in the National Center for Biotechnology Information
625 Gene Expression Omnibus (GEO) database and will be released after manuscript
626 acceptance in a peer-reviewed scientific journal.

627 **REFERENCES**

- 628 1. World Health Organization (WHO). *Global Tuberculosis Report*. (2022).
- 629 2. World Health Organization (WHO). *WHO Report on TB 2020*. (2020).
- 630 3. World Health Organization. *Global Tuberculosis Report*. (2021).
- 631 4. Gil, O. *et al.* Granuloma encapsulation is a key factor for containing tuberculosis
632 infection in minipigs. *PLoS One* **5**, e10030 (2010).
- 633 5. Lenaerts, A., Barry, C. E. & Dartois, V. Heterogeneity in tuberculosis pathology,
634 microenvironments and therapeutic responses. *Immunol Rev* **264**, 288–307 (2015).
- 635 6. Lin, P. L. *et al.* Sterilization of granulomas is common in active and latent
636 tuberculosis despite within-host variability in bacterial killing. *Nat Med* **20**, 75–79
637 (2014).
- 638 7. Gideon, H. P. *et al.* Multimodal profiling of lung granulomas in macaques reveals
639 cellular correlates of tuberculosis control. *Immunity* **55**, 827–846 (2022).
- 640 8. Malherbe, S. T. *et al.* Persisting PET-CT lesion activity and M. tuberculosis
641 mRNA after pulmonary tuberculosis cure. *Nat Med* **22**, 1094 (2016).
- 642 9. Penn-Nicholson, A. *et al.* RISK6, a 6-gene transcriptomic signature of TB disease
643 risk, diagnosis and treatment response. *Sci Rep* **10**, 8629 (2020).
- 644 10. Urbanowski, M. E., Ordonez, A. A., Ruiz-Bedoya, C. A., Jain, S. K. & Bishai, W.
645 R. Cavitary tuberculosis: the gateway of disease transmission. *Lancet Infect Dis*
646 **20**, e117–e128 (2020).
- 647 11. Dheda, K. *et al.* Spatial network mapping of pulmonary multidrug-resistant
648 tuberculosis cavities using RNA sequencing. *Am J Respir Crit Care Med* **200**, 370–
649 380 (2019).
- 650 12. Kempker, R. R. *et al.* Additional drug resistance in *Mycobacterium tuberculosis*
651 isolates from resected cavities among patients with multidrug-resistant or
652 extensively drug-resistant pulmonary tuberculosis. *Clinical Infectious Diseases* **54**,
653 51–54 (2012).
- 654 13. Dartois, V. & E. Barry, C. Clinical pharmacology and lesion penetrating properties
655 of second- and third-line antituberculous agents used in the management of
656 multidrug-resistant (MDR) and extensively-drug resistant (XDR) tuberculosis.
657 *Curr Clin Pharmacol* **5**, 96–114 (2010).
- 658 14. Vashakidze, S. *et al.* Retrospective study of clinical and lesion characteristics of
659 patients undergoing surgical treatment for Pulmonary Tuberculosis in Georgia. *Int
660 J Infect Dis* **56**, 200–207 (2017).

- 661 15. Subbian, S. *et al.* Lesion-specific immune response in granulomas of patients with
662 pulmonary tuberculosis: A pilot study. *PLoS One* **10**, 1–21 (2015).
- 663 16. Marakalala, M. J. *et al.* Inflammatory signaling in human tuberculosis granulomas
664 is spatially organized. *Nat Med* **22**, 531–538 (2016).
- 665 17. McCaffrey, E. F. *et al.* The immunoregulatory landscape of human tuberculosis
666 granulomas. *Nature Immunology* **23**, 318–329 (2022).
- 667 18. Lin, P. L. & Flynn, J. L. The End of the Binary Era: Revisiting the Spectrum of
668 Tuberculosis. *The Journal of Immunology* **201**, 2541–2548 (2018).
- 669 19. McCaffrey, E. F. *et al.* The immunoregulatory landscape of human tuberculosis
670 granulomas. *Nat Immunol* **23**, 318–329 (2022).
- 671 20. Kim, M. J. *et al.* Caseation of human tuberculosis granulomas correlates with
672 elevated host lipid metabolism. *EMBO Mol Med* **2**, 258–274 (2010).
- 673 21. McCaffrey, E. F. *et al.* The immunoregulatory landscape of human tuberculosis
674 granulomas. *Nature Immunology* **23**, 318–329 (2022).
- 675 22. Cadena, A. M., Fortune, S. M. & Flynn, J. L. Heterogeneity in tuberculosis. *Nature*
676 *Reviews Immunology* **17**, 691–702 (2017).
- 677 23. Kahnert, A. *et al.* Mycobacterium tuberculosis triggers formation of lymphoid
678 structure in murine lungs. *Journal of Infectious Diseases* **195**, 46–54 (2007).
- 679 24. Agyekum, S. *et al.* Expression of lymphotoxin-beta (LT-beta) in chronic
680 inflammatory conditions. *J Pathol* **199**, 115–121 (2003).
- 681 25. Wang, H., Coligan, J. E. & Morse, H. C. Emerging Functions of Natural IgM and
682 Its Fc Receptor FCMR in Immune Homeostasis. *Front Immunol* **7**, (2016).
- 683 26. Elkington, P. *et al.* MMP-1 drives immunopathology in human tuberculosis and
684 transgenic mice. *J Clin Invest* **121**, 1827 (2011).
- 685 27. Swanson, R. V. *et al.* Antigen-specific B cells direct T follicular-like helper cells
686 into lymphoid follicles to mediate Mycobacterium tuberculosis control. *Nature*
687 *Immunology* **24**, 855–868 (2023).
- 688 28. Cliff, J. M. *et al.* Distinct phases of blood gene expression pattern through
689 tuberculosis treatment reflect modulation of the humoral immune response. *J Infect*
690 *Dis* **207**, 18–29 (2013).
- 691 29. Reichmann, M. T. *et al.* Integrated transcriptomic analysis of human tuberculosis
692 granulomas and a biomimetic model identifies therapeutic targets. *Journal of*
693 *Clinical Investigation* **131**, (2021).

- 694 30. Bloom, C. I. *et al.* Detectable Changes in The Blood Transcriptome Are Present
695 after Two Weeks of Antituberculosis Therapy. *PLoS One* **7**, e46191 (2012).
- 696 31. Ottenhoff, T. H. M., Dass, R. H., Yang, N., Zhang, M. M. & Wong, H. E. E.
697 Genome-Wide Expression Profiling Identifies Type 1 Interferon Response
698 Pathways in Active Tuberculosis. *PLoS One* **7**, 45839 (2012).
- 699 32. Cai, Y. *et al.* Increased complement C1q level marks active disease in human
700 tuberculosis. *PLoS One* **9**, (2014).
- 701 33. Cliff, J. M. *et al.* Distinct phases of blood gene expression pattern through
702 tuberculosis treatment reflect modulation of the humoral immune response. *J Infect*
703 *Dis* **207**, 18–29 (2013).
- 704 34. Eum, S.-Y. *et al.* Neutrophils Are the Predominant Infected Phagocytic Cells in
705 the Airways of Patients With Active Pulmonary TB. *Chest* (2010).
- 706 35. Gopal, R. *et al.* S100A8/A9 proteins mediate neutrophilic inflammation and lung
707 pathology during tuberculosis. *Am J Respir Crit Care Med* **188**, 1137–1146 (2013).
- 708 36. Moreira-Teixeira, L. *et al.* Mouse transcriptome reveals potential signatures of
709 protection and pathogenesis in human tuberculosis. *Nature Immunology* **2020** *21*:4
710 **21**, 464–476 (2020).
- 711 37. Fonseca, K. L. *et al.* Deficiency in the glycosyltransferase Gcnt1 increases
712 susceptibility to tuberculosis through a mechanism involving neutrophils. *Mucosal*
713 *Immunology* **13**, 836–848 (2020).
- 714 38. Nandi, B. & Behar, S. M. Regulation of neutrophils by interferon- γ limits lung
715 inflammation during tuberculosis infection. *J Exp Med* **208**, 2251–2262 (2011).
- 716 39. Dorhoi, A. *et al.* The adaptor molecule CARD9 is essential for tuberculosis control.
717 *J Exp Med* **207**, 777–792 (2010).
- 718 40. Berry, M. P. R. *et al.* An interferon-inducible neutrophil-driven blood
719 transcriptional signature in human tuberculosis. *Nature* **466**, 973–977 (2010).
- 720 41. Tabone, O. *et al.* Blood transcriptomics reveal the evolution and resolution of the
721 immune response in tuberculosis. *Journal of Experimental Medicine* **218**, (2021).
- 722 42. Krause, R. *et al.* B cell heterogeneity in human tuberculosis highlights
723 compartment-specific phenotype and putative functional roles. *PREPRINT*
724 (*Version 1*) available at *Research Square* (2023).
- 725 43. Stone, R. C. *et al.* Epithelial-Mesenchymal Transition in Tissue Repair and
726 Fibrosis. *Cell Tissue Res* **365**, 495 (2016).

- 727 44. Scriba, T. J. *et al.* Sequential inflammatory processes define human progression
728 from M. tuberculosis infection to tuberculosis disease. *PLoS Pathog* **13**, 1–24
729 (2017).
- 730 45. Thompson, E. G. *et al.* Host blood RNA signatures predict the outcome of
731 tuberculosis treatment. *Tuberculosis* **107**, 48–58 (2017).
- 732 46. Tiwari, D. & Martineau, A. R. Inflammation-mediated tissue damage in pulmonary
733 tuberculosis and host-directed therapeutic strategies. *Semin Immunol* **65**, 101672
734 (2023).
- 735 47. Malherbe, S. T. *et al.* Quantitative 18F-FDG PET-CT scan characteristics correlate
736 with tuberculosis treatment response. *EJNMMI Res* **10**, 1–15 (2020).
- 737 48. Kurbatova, E. V. *et al.* Sputum culture conversion as a prognostic marker for end-
738 of-treatment outcome in patients with multidrug-resistant tuberculosis: a
739 secondary analysis of data from two observational cohort studies. *Lancet Respir
740 Med* **3**, 201–209 (2015).
- 741 49. Wallis, R. S. *et al.* Biomarkers for tuberculosis disease activity, cure, and relapse.
742 *Lancet Infect Dis* **9**, 162–172 (2009).
- 743 50. Assemie, M. A., Alene, M., Petrucca, P., Leshargie, C. T. & Ketema, D. B. Time
744 to sputum culture conversion and its associated factors among multidrug-resistant
745 tuberculosis patients in Eastern Africa: A systematic review and meta-analysis. *Int
746 J Infect Dis* **98**, 230–236 (2020).
- 747 51. Honeyborne, I. *et al.* Profiling persistent tubercle bacilli from patient sputa during
748 therapy predicts early drug efficacy. *BMC Med* **14**, (2016).
- 749 52. Kempker, R. R., Vashakidze, S., Solomonia, N., Dzidzikashvili, N. & Blumberg,
750 H. M. Surgical treatment of drug-resistant tuberculosis. *Lancet Infect Dis* **12**, 157–
751 166 (2012).
- 752 53. Hunter, L., Hingley-Wilson, S., Stewart, G. R., Sharpe, S. A. & Salguero, F. J.
753 Dynamics of Macrophage, T and B Cell Infiltration Within Pulmonary
754 Granulomas Induced by *Mycobacterium tuberculosis* in Two Non-Human Primate
755 Models of Aerosol Infection. *Front Immunol* **12**, 776913 (2022).
- 756 54. Canetti, G. *The Tubercle Bacillus in the Pulmonary Lesion of Man: Histobacteriology and Its Bearing on the Therapy of Pulmonary Tuberculosis*.
757 (Springer Publishing Company, 1955).
- 758

- 759 55. Ewels, P., Magnusson, M., Lundin, S. & Käller, M. MultiQC: Summarize analysis
760 results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047–
761 3048 (2016).
- 762 56. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer for
763 Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
- 764 57. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics*
765 **25**, (2009).
- 766 58. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and
767 dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).
- 768 59. Hänelmann, S., Castelo, R. & Guinney, J. GSVA: Gene set variation analysis for
769 microarray and RNA-Seq data. *BMC Bioinformatics* **14**, 1–15 (2013).
- 770 60. Langfelder, P. & Horvath, S. WGCNA: An R package for weighted correlation
771 network analysis. *BMC Bioinformatics* **9**, 1–13 (2008).
- 772 61. Liberzon, A. *et al.* Molecular signatures database (MSigDB) 3.0. *Bioinformatics*
773 **27**, 1739–1740 (2011).
- 774 62. Yu, G., Wang, L. G., Han, Y. & He, Q. Y. clusterProfiler: an R package for
775 comparing biological themes among gene clusters. *OMICS* **16**, 284–287 (2012).
- 776 63. Yaari, G., Bolen, C. R., Thakar, J. & Kleinstein, S. H. Quantitative set analysis for
777 gene expression: A method to quantify gene set differential expression including
778 gene-gene correlations. *Nucleic Acids Res* **41**, 1–11 (2013).
- 779 64. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical
780 and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical
781 Society: Series B (Methodological)* **57**, 289–300 (1995).
- 782

RNA-sequencing of 44 samples from 13 patients (DSTB=7/ MDR-TB=6) from the SH-TBL cohort (Figures 2a & b)

Tissue transcriptional changes in TB granuloma compared to non-lesional lung tissue



Differential expression analysis of genes in the granuloma compartments versus non-lesional tissue. (Figures 3a, b)

Tissue transcriptional response across TB granuloma

Whole granuloma and granuloma compartments modular analysis (WGCNA). (Figure 3c)



Evaluation of modular signature in whole granuloma versus non-lesional tissue, stratifying patients according to clinical surrogates of treatment response efficacy and disease severity.

Sputum culture conversio | SGRQ symptoms score

Fast (SCC < 2 months)	SGRQ score > 20
Slow (SCC > 2 months)	SGRQ score < 20

(Figures 4a, b & c)

Granuloma modular transcriptional signature according to TB patients'

Figure 1. Overall study plan. Overview of the analysis undertaken in the study. Figures associated with each objective are stated.

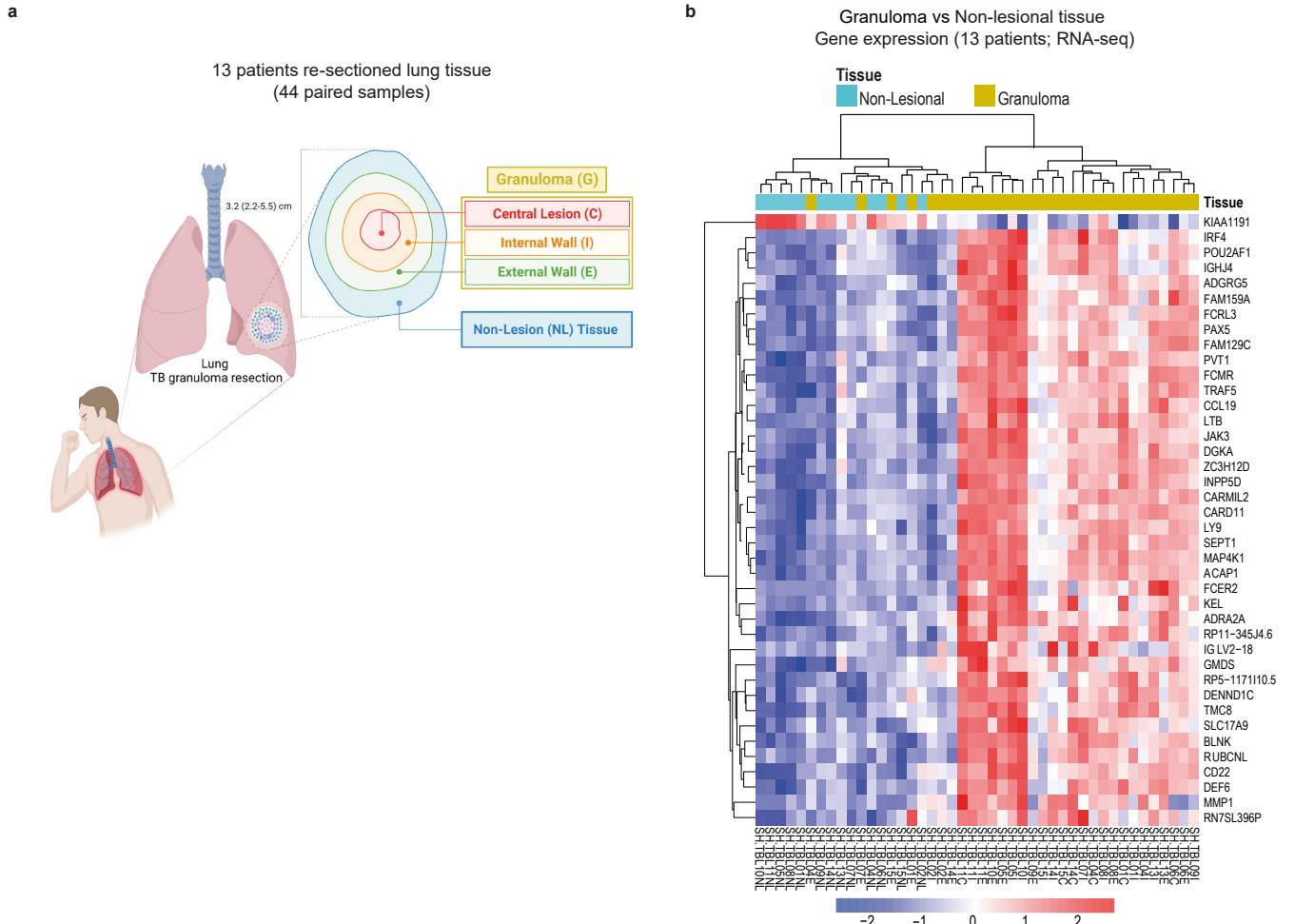
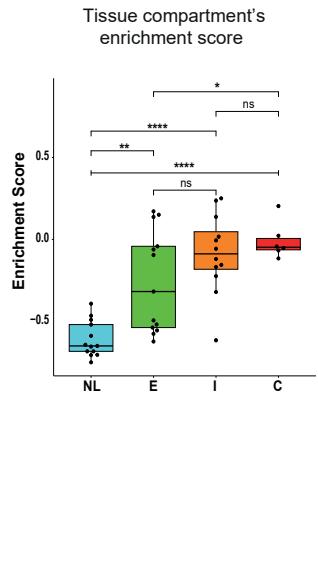
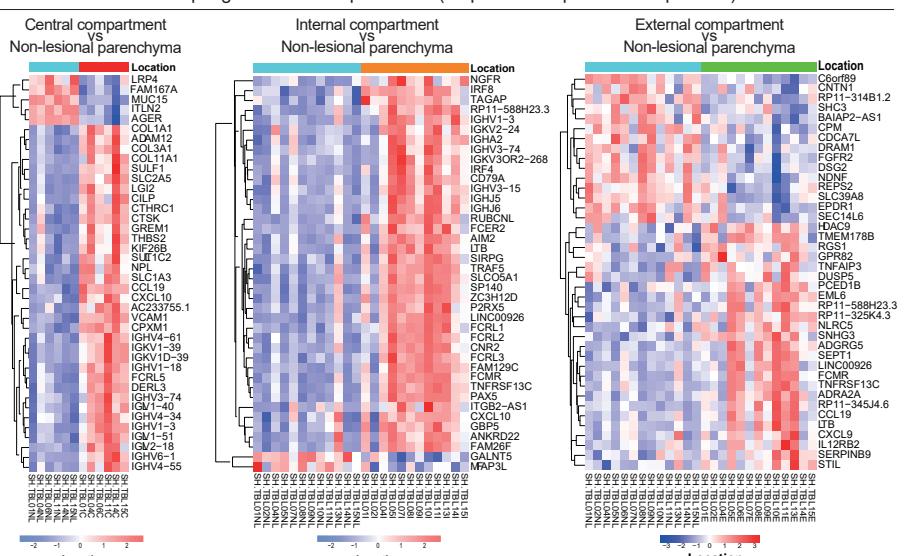


Figure 2. The human TB granuloma signature shows a distinct and heterogeneous transcriptional profile as compared with non-lesional lung tissue. Granuloma samples were collected from each patient included in the SH-TBL cohort: central lesion (C), internal wall (I) and external wall (E) and, altogether, samples from each patient represent the human TB granuloma (G). An additional sample from surrounding non-lesional lung tissue (NL) was also collected from the same patient as control (**Panel a**). 44 samples from 13 patients (7 DS-TB and 6 MDR/XDR-TB) were RNA sequenced to evaluate the human TB lung granuloma transcriptomic changes. A set of 4,630 DGES was identified after comparing the human TB granuloma counts with NL lung tissue expression, using DESeq2 with adjusted $p < 0.05$. **Panel b** heatmap depicts the top 40 DEGs ranked by the adjusted p -value comparing the human TB granuloma versus NL lung tissue expression profiles. The intensity of each colour denotes the standardized ratio between each value and the average expression of each gene across all samples. Red pixels correspond to an increased abundance of mRNA in the indicated sample, whereas blue pixels indicate decreased mRNA levels.

a**b****c**

Key modules associated to granuloma compartments

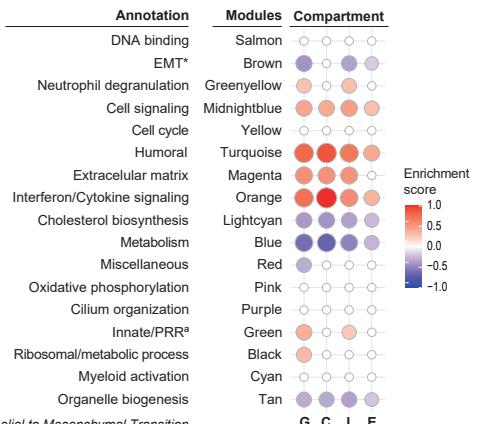


Figure 3. The human lung granuloma compartments have different gene expression profiles and are enriched for immune inflammatory response pathways. Panel a show the enrichment score derived from single sample analysis GSEA using the top 40 genes discriminating granuloma from NL lung tissue. Heatmaps showing differences in the top 40 ranked genes from DESeq2 with adjusted $p < 0.05$ by separately comparing the central (C), internal (I) and external (E) compartments with the NL lung tissue gene expression derived (Panel b). The intensity of each colour denotes the standardized ratio between each value and the average expression of each gene across all samples. Red pixels correspond to an increased abundance of mRNA in the indicated sample, whereas blue pixels indicate decreased mRNA levels. Panel c pictures modular transcriptional of the seventeen modules of co-expressed genes derived from WGCNA for our granuloma dataset separated by compartment. Fold enrichment scores derived using QuSAGE are depicted, with red and blue indicating modules over or under expressed compared to the control. Only modules with fold enrichment ($FDR < 0.1$) were considered significant.

*Epithelial to Mesenchymal Transition

^aPattern recognition receptors

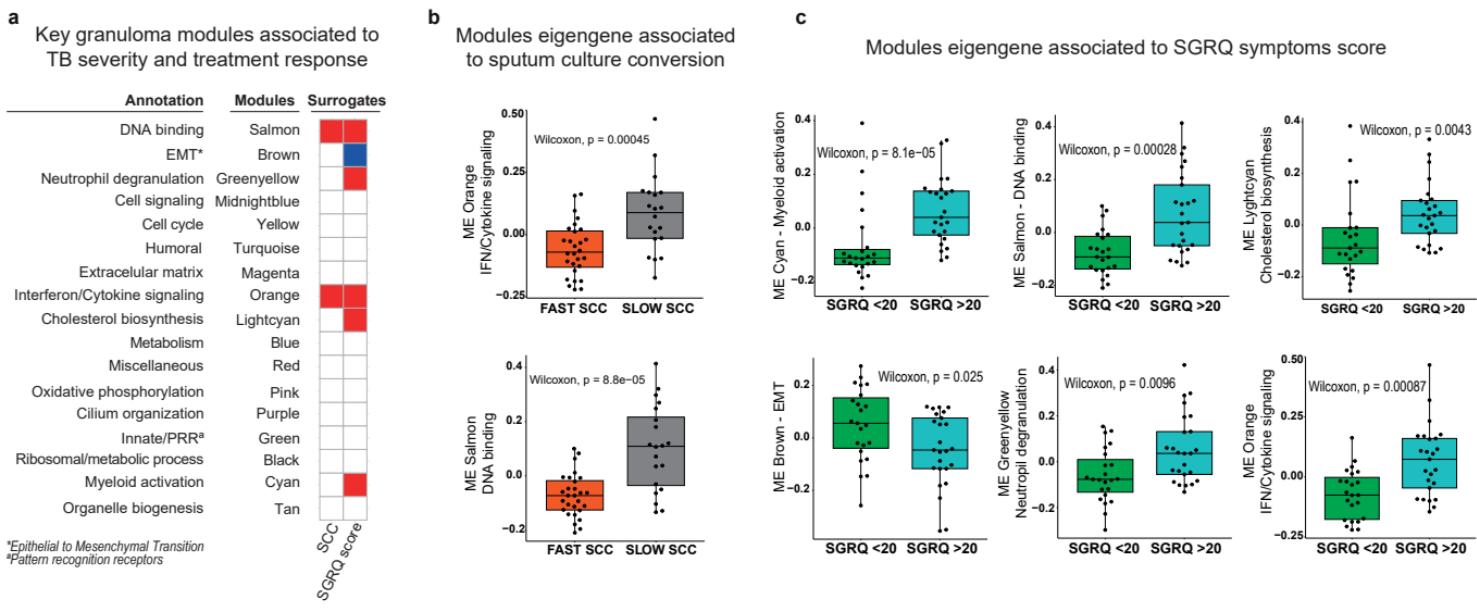
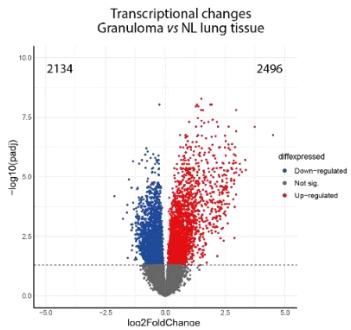


Figure 4. Granuloma modular transcriptional signature correlates with TB clinical and microbiological characteristics revealing differential responses between patient's group. Modular analysis of granuloma RNA-seq samples from 13 patients. Patients clinically defined accordingly to sputum culture conversion (SCC) and TB disease impact on lung function, measured using the Saint George's Respiratory Questionnaire (SGRQ) symptom score, as surrogates of TB severity and treatment response. Heatmap represent the key granuloma modules significantly associated to individual's clinical surrogates of TB severity and treatment response (**Panel a**). Fold enrichments were calculated for each WGCNA module using hypergeometric distribution to assess whether the number of genes associated with each clinical status is larger than expected. Fold enrichment scores derived using QuSAGE are depicted, with red and blue indicating modules over or under expressed compared to the control. The colour intensity represents degree of perturbation. Modules with fold enrichment scored FDR p-value < 0.1 are considered significant. **Panel b** and **c** show TB individuals' stratification according to SCC (fast or slow converters) and SGRQ symptom score (low impact if SGRQ < 20 or high impact if SGRQ > 20), respectively, and the significant association using their corresponding derived WGCNA significant eigengene modules (ME) (p<0.05).

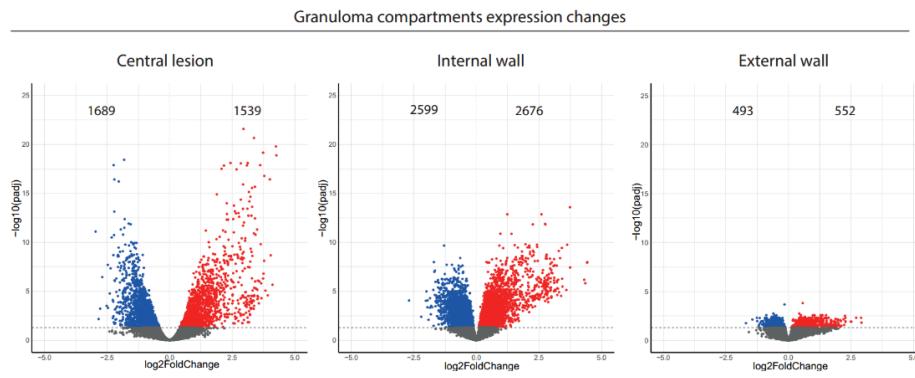
1

Extended Data Figure

a



b



2

3 **Extended data Figure 1. Differential expression analysis of human TB lung granuloma and**
4 **separated compartments relative to non-lesional lung tissue. Panel a)** Volcano plot depicts
5 differentially expressed genes (DEGs) between granuloma and non-lesional (NL) lung tissue.
6 **Panel b)** Volcano plot depicts DEGs between Central, Internal and External compartments and
7 the non-lesional (NL) lung tissue. Genes with the adjusted p -value ≤ 0.05 were considered
8 significantly down-regulated when the log2 fold change was < 0.1 (blue) and up-regulated when
9 log2fold change > 0.1 (red).

10

11

12

Supplementary Information

Supplementary Table 1. Demographic and clinical and TB-related patients' characteristics at time of surgery.

Variable	Category	Overall N= 14	Male n= 7	Female n= 7	p-value
N					
Clinical & epidemiological characteristics					
Age [mean (SD)]		34.36 (12.02)	37 (15.56)	31 (7.39)	0.4387 [a]
BMI [median (range)]		23.55 [19.1-32]	23.4 [21.5-31-7]	23.7 [19.1-32]	0.6841 [b]
Smoker (%)	Yes	4 (28.57)	4 (57.14)	0 (0)	
	No	10 (71.43)	3 (42.86)	7 (100)	0.0699 [c]
Alcohol (%)	Yes	5 (35.71)	5 (71.43)	0 (0)	
	No	9 (64.29)	2 (28.57)	7 (100)	0.0210* [c]
Comorbidities (HCV, HBV, Diabetes) (%)	Yes	6 (42.86)	3 (42.86)	3 (42.86)	
	No	8 (57.14)	4 (57.14)	4 (57.14)	1 [c]
Characteristics of current TB episode					
Patient history (%)	Relapse	5 (37.71)	2 (28.57)	3 (42.86)	
	New patient	9 (64.29)	5 (71.43)	4 (57.14)	1 [c]
Drug Sensitivity (%)	DS-TB	7 (50)	3 (42.86)	3 (42.86)	
	MDR/XDR-TB	7 (50)	4 (57.14)	4 (57.14)	1 [c]
Fast (\leq 2 months) and slow ($>$ 2 months) time to sputum culture conversion	Fast converters	8 (57.14)	3 (42.86)	5 (71.43)	
	Slow converters	6 (42.86)	4 (57.14)	2 (28.57)	0.5921 [c]
Multiple lesions in the CXR	Yes (\geq 2)	4 (28.57)	2 (28.57)	2 (28.57)	
	No (<2)	10 (71.43)	5 (71.43)	5 (71.43)	1 [c]
Macroscopic and microbiological characteristics of the TB lesions					
Size of resected lesions (mm) [median (range)]		30 [29-42]	30 [30-35]	30 [29-42]	0.4371 [b]
Necrosis (%)	Fresh	13 (92.86)	6 (85.71)	7 (100)	
	Not fresh necrosis	1 (7.14)	1 (14.29)	0 (0)	1 [c]
Granuloma biopsy culture conversion by patient (%)	Positive	1 (17.14)	1 (14.29)	0 (0)	
	Negative	13 (92.86)	6 (85.71)	7 (100)	1 [c]

15

Statistics: [a] t-test, [b] Mann-Whitney U-test, [c] Fisher's exact test. *Statistically significant differences found in alcohol use only).

18

19

20 **Supplementary Table 2** - Characteristics of resected lesions from TB patients according to treatment
21 response.

22

Variable	Category	Overall	Fast Responders	Slow Responders	p-value
N		N = 14	n = 8	n = 6	
Size of resected lesion (mm) [median (range)]		30 [29-42]	30 [29-42]	32.5 [30-40]	0.5035 [a]
Necrosis (%)	Fresh necrosis	13 (92.86)	8 (100)	5 (83.33)	0.4286 [b]
	Not fresh necrosis	1 (7.14)	0 (0)	1 (16.67)	
Cavitation (%)	Cavitation	11 (78.57)	5 (62.5)	6 (100)	0.2088 [b]
	Tuberculoma	3 (21.43)	3 (37.5)	0 (0.00)	
Biopsy culture positivity					
Biopsy culture conversion by patient (%)	Positive	1 (7.14)	0 (0.00)	1 (16.67)	0.4286 [b]
	Negative	13 (92.86)	8 (100)	5 (83.33)	

S24 statistics: [a] Mann-Whitney U-test, [b] Fisher's exact test.

C25 comparisons refer to patients presenting microbiological conversion at two months or less against those
26 required more than two months to respond to standard anti-TB treatment.

27

28