Whole blood transcriptional profiles and the pathogenesis of tuberculous meningitis

Hoang Thanh Hai[1]*, Le Thanh Hoang Nhat[1]*, Trinh Thi Bich Tram[1], Artika P Nath[6], Joseph Donovan[1,5], Nguyen Thi Anh Thu[1], Dang Van Thanh[1], Nguyen Duc Bang[2], Dang Thi Minh Ha[2], Nguyen Hoan Phu[1,3], Ho Dang Trung Nghia[1,3,4], Le Hong Van[1], Michael Inouye[6,7], Guy E Thwaites[1,5], Nguyen Thuy Thuong Thuong[1,5#]

[1] Oxford University Clinical Research Unit, Ho Chi Minh City, Viet Nam

[2] Pham Ngoc Thanh Hospital, Ho Chi Minh City, Viet Nam

[3] Hospital for Tropical Diseases, Ho Chi Minh City, Viet Nam

[4] Pham Ngoc Thach University of Medicine, Ho Chi Minh City, Viet Nam

[5] Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine, University of Oxford, United Kingdom

[6] Cambridge Baker Systems Genomics Initiative, Baker Heart and Diabetes Institute, Melbourne, Victoria, Australia

[7] Cambridge Baker Systems Genomics Initiative, Department of Public Health and Primary Care, University of Cambridge, Cambridge, United Kingdom

*equal contribution, [#]corresponding author

Keywords: tuberculosis meningitis, mortality, prognostic, RNA sequencing, gene expression, weighted gene co-expression network analysis

eLife

No maximum length for Research Articles, but try not to exceed 5,000 words, **word count: 4073 / 5000** (excluding the Materials and methods, References, and Figure legends)

No limits on the number of display items, **9 figures + 3 tables**

Abstract

Background: Mortality and morbidity from tuberculous meningitis (TBM) are frequent and strongly associated with the inflammatory response to *Mycobacterium tuberculosis* infection. However, the mechanisms driving the associations are uncertain. We sought to identify the gene modules, hubs and pathways associated with the pathogenesis and mortality from TBM, and to identify which best-predicted death.

Methods: We used whole blood RNA sequencing to obtain transcriptional profiles from 281 Vietnamese adults with TBM (207 HIV-negative; 74 HIV-positive), 295 with pulmonary TB (PTB), and 30 healthy controls. The TBM cohort was divided randomly into a discovery cohort (n=142) and a validation cohort (n=139). Weighted gene co-expression network analysis identified clusters of genes (or 'modules') and hub genes associated with death or disease severity. An overrepresentation analysis identified pathways associated with TBM mortality, with a consensus analysis identifying consensual patterns between HIV-positive and HIV-negative individuals. A multivariate elastic-net Cox regression model selected the candidate predictors of TBM mortality, then model prediction performance using logistic regression and internal bootstrap validation to choose best predictors.

Results: Overall, TBM mortality was associated with increased neutrophil activation and decreased T and B cell activation pathways. Death from TBM was associated with increased angiogenesis in HIV-positive adults, and with activated TNF signaling and down-regulated extracellular matrix organization in HIV-negative adults. PTB and TBM have similar transcriptional profiles compared to healthy controls, although inflammatory genes were more activated in HIV-positive than HIV-negative TBM. The expression of four hub genes – *MCEMP1*, *NELL2*, *ZNF354C* and *CD4* – were strongly predictive of death from TBM (AUC 0.80 and 0.86 for HIV-negative and HIV-positive, respectively).

Conclusions: Whole blood transcriptional profiling revealed that TBM is associated with a characteristic systemic inflammatory response, similar to that invoked by pulmonary tuberculosis, but with key gene modules, hubs and pathways strongly associated with death. Our analysis suggests a novel 4-gene biomarker for predicting death from TBM, but also opens a new window into TBM pathogenesis that may reveal novel therapeutic targets for this lethal disease.

53 **Introduction**

54 Of 7.1 million new tuberculosis (TB) cases in 2019, tuberculous meningitis (TBM) is estimated to have
55 developed in 164,000 adults, around 25% of whom were living with HIV [1]. TBM is the most severe form
56 of TB, causing death or neurological disability in half of all cases. Overall, TBM mortality is around 25%,
57 but rises to around 50% in those with HIV, with most deaths occurring in the first 3 months of treatment
58 [1,2].

59

60 The poor outcomes from TBM are strongly associated with the inflammatory response [3,4], with both a
61 paucity and an excess of inflammation linked to death from TBM [5-11]. However, the mechanisms behind
62 these observations remain uncertain. Immune responses in TBM are thought to be compartmentalized
63 within the central nervous system. Studies have shown that immune cell counts, cytokine
64 concentrations, metabolites and transcriptional responses differ between the peripheral blood and the
65 cerebrospinal fluid (CSF) [10,12,13]. In adults with TBM, leukocyte activation is higher in the CSF than in
66 peripheral blood, although a marked myeloid response in peripheral blood has been reported [10]. Blood
67 transcriptomic analysis has found increased neutrophil-associated transcripts and inflammasome
68 signaling in those with HIV-associated TBM and immune reconstitution inflammatory syndrome [5]. In
69 children with TBM, whole blood transcriptional profiles showed increased inflammasome activation and
70 decreased T-cell activation [12]. Taken together, these studies suggest the inflammatory response
71 associated with TBM may have a greater systemic component than originally thought and its
72 characterization may help define the immune mechanisms leading to death and disability.

73

74 The accurate and early detection of patients at highest risk of complications and death from TBM may
75 help to target treatment for those most in need. The British Medical Research Council (MRC) grades
76 have been used to categorize TBM severity for almost 80 years [14], and the system strongly predicts TBM

77 mortality[15]. Previously, we and others developed new prognostic models from studies of 1699 adults
78 with TBM using clinical and laboratory parameters, including MRC grade[4,15]. These models predicted
79 outcome more accurately than MRC grade alone. However, they might be improved by measures of host
80 inflammatory response. Host-based peripheral blood gene expression analysis has been used to identify
81 active or progressive pulmonary TB and in pulmonary TB treatment monitoring [16,17], but has yet to be
82 applied to TBM.
83
84 In the current study, we investigated whole blood transcriptional profiles in 281 Vietnamese adults with
85 TBM, 295 with pulmonary TB (PTB), and 30 healthy controls. Our objective was to use weighted gene co-
86 expression network analysis, an unbiased and well-evaluated approach, to identify the biological
87 pathways and hub genes associated with TBM pathogenesis and assess the predictive value of gene
88 expression for early mortality from TBM.
89
90 **Results**
91 **Characteristics and outcomes of the cohorts**
92 Four cohorts (all ≥18 years) were used in the study, representing a total of 606 participants. The
93 characteristics of these cohorts are provided in **Table 1**. There were 281 adults with TBM; 207 HIV-
94 negative and 74 HIV-positive. In the HIV-negative TBM adults, the median age was 46 years (IQR 34, 58),
95 127 (61%) were male, and the median Body Mass Index (BMI) was 20.2 (IQR 18.2, 22.3). HIV-positive
96 adults were more likely than HIV-negative adults to be male, younger, have lower BMI, have previously
97 received TB treatment, and to have microbiologically confirmed TBM. Total white cell counts in blood
98 and CSF in HIV-positive adults were lower than in HIV-negative adults. The PTB cohort consisted of 295
99 adults with the median age of 44 years (IQR 31, 52), 228 (77%) were male, the median of BMI was 19.4

100 (IQR 17.7, 21.6) and 129 (48%) had pulmonary cavities on chest X-ray. Of the 30 healthy controls, 11
101 (37%) were male, and the median age was 33 (IQR 29, 37) (**Table 1**).
102
103 The clinical variables associated with three-month mortality of the 281 adults with TBM are given in
104 **Table 2**. The discovery (n=142) and validation (n=139) cohorts had similar characteristics (**Table 2**).
105 47.5% (133/280) had definite TBM [18], with microbiologically confirmed disease, accounting for 45.9%
106 (101/220) of survivors and 52.4% (32/61) of those who died. The overall three-month mortality rate was
107 21.7% (61/281): 16.4% (34/207) in HIV-negative and 36.5% (27/74) in HIV-positive (p<0.001). We did not
108 observe differences in mortality by sex, age and diagnostic category. Greater disease severity, MRC
109 grades 2 and 3 at enrolment, was associated with increased mortality compared to grade 1 (p<0.001). In
110 those who died, CSF and peripheral blood neutrophil counts were higher and peripheral blood
111 lymphocyte count lower, compared to those who survived.
112
113 **Whole blood transcriptional profiles of the four cohorts**
114 We analyzed the whole blood transcriptomics, using bulk RNA sequencing from 606 participants in the 4
115 cohorts. On average 35.1 million reads/sample and 57% reads mapping accuracy to human reference
116 genome (GRCh.38 release 99) was obtained. The study objectives and cohorts flow are presented in
117 **Figure 1.** Principal component analysis on the transcriptomic profile of the top 20,000 genes with most
118 variation across 4 studies showed distinct separation between healthy controls from both PTB and TBM
119 cases (**Figure 2**). The PTB profile substantively overlapped with TBM, with some separation between
120 HIV-negative and HIV-positive TBM (**Figure 2**).
121
122 Enrichment scores from single sample gene set enrichment analysis (ssGSEA), which based on
123 expression rank of genes relevant to pathways, were measured in some pathways already known to be

important mediators of TB or TBM pathogenesis (**Figure 2**). As anticipated, inflammatory response, cytokine signaling, interferon signaling, TNF signaling, and inflammasome activation pathways, were enriched in PTB and TBM cohorts as compared to healthy controls. In TBM, enrichment of genes in these pathways were generally higher in HIV-positive than in HIV-negative individuals (**Table S1, Table S5**).

Transcriptional gene modules associated with TBM severity and mortality

Transcriptional profiles associated with TBM mortality were generated by identifying differentially expressed genes. Of the top 20,000 genes with most variation, we observed 724 (3.6%) genes that were differentially expressed (FDR <0.05, absolute fold change (FC) >1.5) in all those with TBM (**Figure 3, Table S2**). Next, we applied weighted gene co-expression network analysis (WGCNA) to 5000 most variable genes from 281 TBM samples (n= 207, HIV-negative; n=74, HIV-positive) to define clusters of highly correlated genes (modules) associated with TBM severity and mortality. Gene modules are clusters of genes that have highly interconnected expression, usually related to their biological functions. Hub genes are genes with high connectivity to other genes within a respective module. First, we used WGCNA to construct a network of gene modules in the discovery cohort. Then we validated the presence of these transcriptional modules in the validation cohort, labelling the modules with different colors. In the discovery cohort (n=142), 15 modules were identified overall, consisting of 44 to 1350 genes per module (**Figure S1, Table S6**). All 15 modules were preserved in the validation cohort (n=139) through the preservation analysis (**Figure S2**).

The associations between the 15 modules and TBM severity and mortality are presented in **Figure 4** for both the discovery and validation cohorts. Modules were linked to each other in a hierarchical structure, with major biological processes annotated. Associations of the modules with TBM disease severity (MRC grade) at baseline were estimated by Spearman correlations between MRC grade and the first principle

148 component (PC1) of each module. Similarly, associations of the modules with mortality were measured
149 by hazard ratio (HR) per increase 1 unit of PC1 using Cox regression model adjusted for age, HIV status
150 and dexamethasone treatment (**Figure 4**) in both discovery and validation cohorts.
151
152 Of the 15 preserved modules, five modules were significantly associated with mortality in the discovery
153 and validation cohorts, with false discovery rate (FDR) < 0.05 (**Figure 4**). These five modules were
154 separated into two big module clusters. The first cluster contained the blue module (n=799 genes),
155 involved in inflammatory and innate immune responses, and the cyan module (n=44 genes) with
156 unknown biological function. These modules were upregulated in those who died, as shown in the heat-
157 map in **Figure 4** (HR: 3.0 and 2.2 for the blue module, and 2.2 and 1.7 for the cyan module, FDR < 0.05
158 for all comparisons). The black (n= 207 genes), brown (n=698 genes) and red (n=229 genes) modules
159 were in the second cluster and were generally involved in adaptive immunity including T and B cell
160 signaling pathways. These three modules were down-regulated in those who died (HR: 0.43 and 0.36 for
161 brown, 0.46 and 0.39 for red, 0.59 and 0.49 for black) (**Figure 4**).
162
163 It is known that TBM MRC grade before treatment initiation strongly predicts outcome from TBM. Here,
164 we investigated correlations between each module and MRC grade and their association with mortality.
165 The pink module, involved in hemostasis and platelet activation, was positively correlated with MRC
166 grade, but not mortality. Of the five modules associated with death from TBM, all were correlated with
167 MRC grade (**Figure 4**). Four of these five modules, were enriched for immune responses.
168
169 **Transcriptional hub genes associated with TBM severity and mortality**
170 We next identified hub genes within the four biologically functional modules associated with TBM
171 mortality in both the discovery and validation cohorts. Hub genes showed higher connectivity within the

172 modules, and stronger association with TBM mortality as compared to less connected genes within a
173 module (**Figure S3**). Seven hub genes (*ETS2, PGD, UPP1, CYSTM1, FCAR, KIF1B* and *MCEMP1*) were
174 upregulated in death, all from the acute inflammation (blue) module. Hub genes from the brown, red
175 and black modules were downregulated in death and were involved in adaptive immune response. Ten
176 hub genes associated with mortality (*CD96, TNFRSF25, TBC1D4, CD28, ABLIM1, RASGRP1, NELL2, TRAF5,*
177 *TESPA1, TRABD2A*) were from the brown module, three (*EVL, PLCG1, NLRC3*) from the red, and six (*CD2,*
178 *CD247, TGFBR3, ARL4C, KLRK1, MATK*) from the black module (**Figure 5, Table S7**).
179
180 We examined patterns of shared and distinct gene expression of some hub across the four cohorts
181 (healthy controls, PTB, and TBM with and without HIV-infection) to reveal disease progression and
182 pathogenesis of different TB forms (**Figure 6)**. There were two upregulated genes from the acute
183 inflammation module (*FCAR* and *MCEMP1*) and six downregulated genes (*NELL2, TRABD2A, PLCG1,*
184 *NLRC3, CD247* and *MATK*) from the other three adaptive immunity modules. The patterns of up and
185 downregulation, relative to healthy controls, were similar in PTB and TBM, although tended to be more
186 pronounced in those with TBM as well as those with HIV (**Figure 6**).
187
188 **Transcriptional immune pathways associated with TBM mortality**
189 To better understand the biological functions of the five modules associated with TBM mortality, a gene
190 set associated with mortality in each module was functionally annotated using known pathway
191 databases, such as gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG)[19].
192 Pathway enrichment analysis was performed using fold enrichment to determine if the prevalence of
193 genes in a pathway were different from that expected by chance. We also used ssGSEA enrichment
194 scores based on gene expression ranking of genes in a particular pathway within a single individual, to

195 show the activity of this pathway between death and survival TBM as well as across 4 cohorts. These
196 analyses helped to identify the pathways linked to mortality within the gene modules.
197
198 In the blue module, we found TBM mortality was associated with upregulated inflammatory and innate
199 immune response transcripts, particularly in pathways involved in the acute inflammatory response and
200 the regulation of inflammatory responses, including responses to bacteria and neutrophil activation.
201 KEGG pathway analysis suggested this signal was associated with transduction and immune system
202 pathways, such as toll-like receptor signaling, TNF signaling, NF-kappa B signaling and neutrophil
203 extracellular trap formation (**Figure 5, Table S8**).
204
205 We did not find any pathway significantly associated with mortality in the upregulated cyan gene
206 module, although in the hierarchical clustering the cyan module was highly correlated with the blue -
207 inflammatory response module (**Figure 4**). In contrast, analyses of the down-regulated brown, red and
208 black modules highlighted pathways involved in adaptive immunity, predominantly those mediated by
209 lymphocytes. These included downregulations of lymphocyte proliferation, T cell activation, B cell
210 activation, natural killer cell mediated cytotoxicity and their signaling pathways such as antigen
211 receptor-mediated, T and B cell receptor (**Figure 5**, **Table S8**).
212
213 We also investigated the association between pathways known to be important to TB pathogenesis and
214 mortality (**Figure 7**). Looking at those who died vs. survived amongst HIV-positive TBM, there was little
215 difference between the groups with respect to inflammatory response, cytokine signaling, and icosanoid
216 and inflammasome activation. This was also the case for HIV-negative individuals. However, interferon
217 and TNF pathways were more active in those with HIV, and TNF signaling expression was significantly
218 higher in HIV-negative adults who died rather than survived.

219

220 **HIV influence on modules, hub genes, and pathways associated with TBM mortality**

221 Transcriptional profiles associated with TBM mortality stratified by HIV status are shown in **Figure 3**. In
222 HIV-negative individuals, 786 (3.9%) genes were differentially expressed, whereas in HIV-positive the
223 number was 1620 (8.1%) genes (**Figure 3, Table S2 & S3)**. We hypothesized that host transcriptional
224 signatures associated with TBM mortality differ according to HIV status. To test this hypothesis, we
225 constructed gene co-expression networks from all genes in HIV-negative (n=207) and HIV-positive (n=74)
226 individuals separately, then performed consensus gene co-expression network analysis (**Figure S4**).
227 Modules were identified with colors and to discriminate them from the modules linked to mortality
228 alone we added an annotated function on the module name. We focused on modules that failed to form
229 consensus associations with mortality due to opposite associations in the two cohorts.

230

231 Sixteen gene co-expression modules (**Figure 8**, **Table S9**), ranging from 60 to 958 genes, were obtained
232 from the HIV-negative and positive cohorts. Of these, 12 modules formed consensus association with
233 mortality (**Figure 8**). Of the 4 modules which failed to form consensus association, two modules were
234 significantly associated with mortality (FDR<0.05). The blue-angiogenesis module was up-regulated in
235 death in HIV-positive adults (HR: 1.7) and the yellow-extracellular matrix organization (EMO) module
236 was down-regulated in death in HIV-negative adults (HR: 0.31) (**Figure 8**).

237

238 Hub genes associated with mortality, which are highly correlated with other genes in the module, were
239 identified in the two modules. Five hub genes, with three downregulated (*ZNF354C, ZNF721* and *ETAA1*)
240 and two upregulated (*CEP295NL* and *RNF24*), were identified in the blue-angiogenesis module. Other
241 five hub genes, with four downregulated (*IL11RA*, *CD4*, *ZNF395* and *BCL9L*) and one upregulated
242 (*PPP4R2*), were identified in the yellow-EMO module (**Figure 8, Table S10**). Expression of some hub

243 genes across the four cohorts showed the patterns of up (*RNF24*) and downregulation (*ZNF721, BCL9L*
244 and *IL11RA*), and relative to healthy controls they were similar in PTB and TBM. Expression of *RNF24* and
245 *ZNF721* genes were significantly associated with death in HIV-positive adults, whereas expression of
246 *BCL9L* and *IL11RA* genes were significantly associated with death in HIV-negative adults (**Figure S5**).
247
248 Pathway enrichment analysis showed genes in the blue-angiogenesis module were significantly enriched
249 for angiogenesis or blood vessel development, leukocyte and neutrophil activation, and signal
250 transduction pathways (**Figure 8**, **Table S11**). Gene expression in the yellow- module was strongly
251 enriched in extracellular matrix organization, cell-substrate adhesion and protein localization pathways
252 (**Figure 8**, **Table S11**). Hierarchical clustering of modules indicated that these two modules were highly
253 correlated (**Figure 8**) suggesting they share similar functions, which appeared in some pathways such as
254 angiogenesis and extracellular matrix organization.
255
256 In addition, almost all of the pathways significantly associated with TBM mortality were similar in HIV-
257 negative and positive cohorts, confirming that these important pathways were common to all TBM
258 (**Figure S6**). However, pathway enrichment scores were higher in HIV-positive than in HIV-negative,
259 especially for the TNF signaling pathway and TNF transcripts (**Figure 9**). Death in HIV-negative TBM was
260 associated with increased enrichment of these pathways, but with less TNF transcript compared to HIV-
261 positive individuals. The enrichment scores indicated mortality was strongly associated with
262 downregulation of adaptive immune responses, including T cell activation, T and B cell receptor signaling
263 pathways, with little impact of HIV status on these pathways (**Figure 9**). These data suggest an
264 inadequate adaptive immune response contributes to disease pathogenesis and mortality in all those
265 with TBM, regardless of HIV status.
266

267 **Predictors of TBM mortality**

268 We aimed to identify baseline gene signatures that might help predict death or survival from TBM. We
269 selected potential gene predictors from 26 common and 10 HIV-specific hub genes; these hub genes
270 represented the dominant modules associated with mortality. Three predefined clinical factors known
271 to be associated with outcome (age, MRC grade and CSF lymphocyte count) were also used for
272 candidate predictor selection. Using a multivariate Cox elastic-net regression model for 39 predictors, six
273 predictors were selected for HIV-negative TBM (MRC grade, age, CSF lymphocyte count, gene set 1:
274 *MCEMP1*, *TRABD2A* and *CD4*), and 6 predictors were selected for HIV-positive TBM (MRC grade, age,
275 CSF lymphocyte count, gene set 2: *NELL2*, *MCEMP1* and *ZNF354C*) (**Table S12**). Gene expression and
276 association with mortality of five distinct genes in gene sets 1 and 2 were presented in **Figure S7**. By
277 combining the two gene sets above and reducing genes in the same module, we generated gene set 3
278 (*MCEMP1*, *TRABD2A*, *CD4* and *ZNF354C*) and gene set 4 (*MCEMP1*, *NELL2*, *ZNF354C* and *CD4*). These two
279 gene sets, with and without clinical factors (MRC grade and age), were tested for their performance in
280 HIV-negative and HIV-positive cohorts together with a reference model [15]. Model 7 from gene set 4 and
281 clinical factors outperformed other gene sets and reference model, with the best overall model
282 performance (lowest optimism-corrected Brier score, 0.11 and 0.14 for HIV-negative and HIV-positive,
283 respectively) and the best discriminant performance (highest optimism-corrected AUC 0.80 and 0.86 for
284 HIV-negative and positive) (**Table 3, Figure S8**).

285

286 **Discussion**

287 The biological pathways involved in pathogenesis of TBM are unclear. In general, previous studies
288 investigating TBM pathogenesis have been small, testing for relatively small numbers of selected genes
289 or molecules, and have been unable to take an unbiased and broader view of the inflammatory
290 response. This study investigated the pathways associated with death from TBM at a whole-genome

291 transcriptome level, characterizing a global dysregulation in immune responses, including inflammation,
292 and revealing specific functional pathways and hub genes involved in TBM and the mechanisms leading
293 to death.
294
295 We sought to understand better the pathogenesis of TBM by identifying pretreatment blood
296 transcriptional gene modules associated TBM disease severity and three-month mortality. Four out of
297 five identified modules were involved in immunological functions, indicating multiple functional
298 pathways of systemic immunity are involved in the pathogenesis of TBM. In particular, mortality was
299 strongly associated with increased acute inflammation and neutrophil activation, and decreased
300 adaptive immunity and T and B cell activation. Whilst there appeared to be many common pathways
301 involved in TBM mortality in HIV-positive and negative individuals, there were differences: death was
302 associated with increased expression of angiogenesis genes in HIV-positive adults, and with TNF
303 signaling and down regulated extracellular matrix organisation in HIV-negative adults. We also identified
304 a four-gene signature as a potential host response biomarker for mortality, regardless of HIV status.
305
306 TBM mortality was associated with increased acute inflammatory responses, the regulation of
307 inflammatory responses, and neutrophil activation. Previous blood transcriptome studies have shown
308 that IFN-inducible neutrophil-driven transcripts were over-expressed in blood neutrophils from active TB
309 patients compared to those with latent TB[20,21]. Another study described that TBM adults who developed
310 IRIS during treatment, also had significantly more abundant neutrophil-associated transcripts that
311 preceded the development of TBM-IRIS [5]. Our own earlier studies have also suggest a role for
312 neutrophils in TBM pathogenesis, with high pretreatment CSF bacterial loads being correlated with high
313 neutrophil numbers in both CSF and blood and more frequent new neurological events or paradoxical
314 inflammatory complications [8].

315

316 Taken together with previous studies, our findings support an important role for over-activation of
317 neutrophil-mediated inflammatory responses in TBM pathogenesis and its lethal complications. Looking
318 further into specific pathways, TBM mortality was neither associated with inflammatory response and
319 cytokine signaling pathways in general, nor with interferon or inflammasome activation. However,
320 increased transcripts in some specific immunity pathways, including TNF signaling, Toll-like receptor, NF-
321 kappa B and neutrophil extracellular trap formation, were associated with TBM mortality. These
322 pathways are involved as activators or effectors in the process of neutrophil activation and regulation,
323 leading to an exacerbated inflammatory response [22-24].

324

325 We found blood transcriptional responses of T cells and B cells were under-expressed in those who died
326 from TBM, indicating an impairment in adaptive immunity in fatal disease. A reduction of activities in
327 both T cell and B cell receptor signaling pathways in death were identified, independent of over-
328 expression of neutrophil-mediated immune responses, indicating that multiple functional pathways
329 influence TBM mortality. TBM pathogenesis is known to be associated with T cell impairment. Previous
330 studies have shown lower numbers of T cells, reduced ability to respond to *Mycobacterium tuberculosis*
331 antigens or reduced expression of activation markers and cytokine production in TBM compared to PTB
332 and healthy individuals [10,25,26]. This impaired T cell function has correlated with disease severity and poor
333 clinical outcomes in participants with PTB and TBM [10,27,28].

334

335 B cells and antibodies also influence humoral immunity against *Mycobacterium tuberculosis*. Studies
336 have shown a decreased memory B cell proportion, and lower levels of IgG, IgM antibodies and Fcγ
337 receptors binding capacity in plasma in those with active lung TB compared to those with latent TB or
338 healthy volunteers [29,30]. But little is known about the role of B cells in TBM. The observed association

339 between TBM mortality and decreased transcriptional responses in B cell activation and B cell receptor
340 signaling pathways suggest an unanticipated role for B cells and humoral immunity in TBM pathogenesis
341 that needs further investigation.
342
343 Transcriptomic profiles from four cohorts, including healthy controls and PTB, provide a broad view of
344 host responses in different TB clinical forms. Our data showed common transcriptional pathways and
345 genes between PTB and TBM. A range of immune responses, involving in inflammation, cytokines,
346 interferon, inflammasome and neutrophil signaling pathways, were activated in both PTB and TBM, but
347 with significantly greater activation in HIV-positive TBM than HIV-negative TBM. This finding aligns with
348 our previous data showing a dyregulated hyper-inflammation in HIV-associated TBM, with significantly
349 higher CSF cytokine concentrations than in those without HIV infection [7]. These data suggest that
350 different forms of TB are associated with similar inflammatory responses, but with different degrees of
351 host responses, exemplified by hyper-inflammation in those with HIV.
352
353 The HIV-driven differences in TBM-associated inflammation appear sufficient to influence response to
354 adjunctive anti-inflammatory therapy with corticosteroids. An earlier randomized controlled trial of
355 corticosteroids in 545 predominantly HIV-negative Vietnamese adults showed they significantly
356 improved survival [31]. Our group recently completed a similar sized trial exclusively in HIV-positive adults
357 without any clear benefit upon mortality (NEJM 22-16218.R2, in press). Therefore, the differences in
358 gene expression associated with TBM mortality in the current study – with angiogenesis activation
359 linked to HIV-positive TBM mortality only, for example – may provide an explanation for the poor
360 response and offer alternative therapeutic strategies. Recent studies have shown that angiogenesis is
361 induced by *Mycobacterium tuberculosis* infection, which then contributes to inflammation, tissue
362 damage and is correlated with disease severity [32].

363

364 Developing prognostic models for TBM is important for guiding clinical decision making and improving
365 outcomes. Several studies have developed and validated prognostic models for TBM using clinical,
366 laboratory and radiological variables. These models have demonstrated moderate to high accuracy in
367 predicting mortality and functional outcomes from TBM patients [15,33-35]. In this study, we used blood
368 transcriptional signatures and co-expression network analysis to identify module-representative hub
369 genes. We identified a four-gene set at the start of treatment (*MCEMP1*, *NELL2*, *ZNF354C* and *CD4*)
370 whose expression strongly predicted three-month TBM mortality. Our prognostic models combining this
371 four-gene host response in blood and two routine clinical predictors achieved very good performance,
372 with AUC 0.80 and 0.86 for HIV-negative and HIV-positive individuals with TBM, respectively. This is
373 proof-of-concept that whole blood RNA host response might be a useful pre-treatment biomarker to
374 predict early TBM mortality. Although further investigation and validation is needed, we have identified
375 potential gene candidates for future development as prognostic biomarkers.

376

377 In summary, we present a comprehensive and unbiased analysis of the gene transcripts associated with
378 TBM severity and mortality. Our data open a new window on TBM pathogenesis, with dysregulation in
379 both innate and adaptive immune responses strongly associated with death from TBM. Furthermore, we
380 have identified similarities and differences in the inflammatory responses associated with TBM in HIV-
381 positive and negative adults, which may explain the different therapeutic effects of adjunctive
382 corticosteroid treatment. We also revealed a 4-gene host response signature in blood that might
383 represent a novel biomarker for defining those at highest risk of death, regardless of their HIV status.

384

385

386 **Materials and Methods**

387 **Participants**

388 We collected data and whole peripheral blood samples for transcriptomic profiling from adults (≥18
389 years) with TBM enrolled into two randomized controlled trials conducted in Vietnam. The two trials
390 investigated whether adjunctive dexamethasone improves outcome from TBM and the protocols for
391 both trials have been published[36,37] The ACT-HIV trial (NCT03092817) completed enrolment and follow-
392 up of 520 HIV-positive adults with TBM in April 2023, with the results accepted for publication (NEJM
393 22-16218.R2, in press). The LAST-ACT trial (NCT03100786) completed enrollment of 720 HIV-negative
394 adults with TBM in March 2023, with the last participants due to complete follow-up in March 2024.

395

396 Peripheral blood samples were taken from 281 trial participants: the first 207 consecutively enrolled
397 HIV-negative participants from the LAST-ACT trial, and 74 randomly selected HIV-positive participants
398 from the ACT-HIV trial. The samples were taken at enrollment, when patients could not have received
399 more than 6 days of anti-tuberculosis drugs. All trial participants then received standard 4-drug anti-
400 tuberculosis treatment for 2 months, followed by 3 drugs for 10 months, and were randomly allocated
401 to dexamethasone or identical placebo for the first 6-8 weeks[36,37]. The investigators remain blind to the
402 treatment allocation until the last participant completes follow-up and the database has been locked.
403 Therefore, an analysis of the direct influence of corticosteroids on inflammatory response and outcome
404 is not included in the current study.

405

406 Peripheral blood samples were taken for transcriptional profiling from two other cohorts: 295 adults
407 with PTB and 30 healthy controls. The PTB cohort was enrolled in Pham Ngoc Thach hospital and District
408 TB Units in Ho Chi Minh city, from a prospective observational study of the host and bacterial
409 determinants of outcome from PTB. Participants had culture-confirmed PTB, either drug susceptible TB

410 or a new diagnosis of multidrug-resistant TB. All participants in this cohort had <7 days of anti-TB drugs
411 at enrolment and did not have clinical evidence of extra-pulmonary TB. Healthy controls were enrolled
412 in Hospital for Tropical Diseases in Ho Chi Minh city, from a prospective study for epidemiological
413 characteristics of human resistance to *Mycobacterium tuberculosis* infection. Participants in this cohort
414 were adults without signs or symptoms of TB nor history of TB contact within the last two years.
415
416 Ethics approval was obtained from the institutional review board at the Hospital for Tropical Diseases,
417 Pham Ngoc Thach Hospital and the ethics committee of the Ministry of Health in Vietnam, and the
418 Oxford Tropical Research Ethics Committee, UK (OxTREC 52-16, 36-16, 24-17, 33-17 and 532-22). All
419 participants provided their written informed consent to take part in the study, or from their relatives if
420 they were incapacitated.
421
422 **Study design**
423 The objectives and cohorts used in this study are presented in **Figure 1** and a workflow of data analysis is
424 shown in **Figure S9**. Briefly, blood transcriptional profiling was generated from four cohorts of 207 TBM
425 HIV-negative, 74 TBM HIV-positive, 279 PTB and 30 healthy controls. To define transcriptional signatures
426 associated with TBM mortality, data from all 281 TBM participants were used. To ensure reproducibility,
427 in our main analysis we randomly split our TBM data into two datasets, a discovery cohort (n=142) and a
428 validation cohort (n=139). After identifying gene modules associated with mortality, related functional
429 pathways and hub genes were determined with more details of the data analysis provided in **Figure S9**.
430 In a broader view, gene enrichment and expression from significant pathways and hub genes were
431 illustrated across all four cohorts. Finally, outcome prediction models were developed for TBM.
432
433

434 **Sample processing and RNA-seq**

435 Whole blood samples, collected from participants at enrollment, were stored in PAXgene collection
436 tubes at -80^{0}C. RNA samples were isolated at the same time using the PAXgene Blood RNA kits (Qiagen,
437 Valencia, CA, USA) following the manufacturer's instructions, except for an additional washing step
438 before RNA elution. DNA was digested on columns using the RNase-free DNase Set (Qiagen, Valencia,
439 CA, USA). Quality control of the RNA extraction was performed using the Epoch spec for quantity and
440 quality, and Tapestation Eukaryotic RNA Screentape for integrity. Samples with RNA integrity number
441 below 4 were exclude for further steps. RNA-seq was performed by the Ramaciotti Centre for Genomics
442 (Sydney, Australia). One microgram of total RNA was used as input for each sample, using the TruSeq
443 Stranded Total RNA Ribo-zero Globin kit (Illumina). Libraries were generated on the Sciclone G3 NGS
444 (Perkin Elmer, Utah, USA) and the cDNA was amplified using 11 PCR cycles. Libraries were pooled 75
445 samples per pool and sequenced using NovaSeq 6000 S4 reagents at 2x100bp to generate about 30
446 million reads per sample.

447

448 **RNA-seq data quality control and pre-processing**

449 Quality control and alignment were performed using an in-house pipeline modified from previously
450 published practices for RNA-seq analysis [38,39] in linux command line. Briefly, the quality of the
451 sequencing fastq files was analyzed using FastQC (v0.11.5) and poor quality samples were excluded from
452 further analysis. Sequence reads were adapter and quality trimmed using Trimmomatic (v0.36), followed
453 by duplicated optical read removal using BBMap (v38.79) tool. STAR aligner (v2.5.2a) was used to align
454 the reads to the human reference genome (GRCh38 build 99) downloaded from Ensembl, allowing for
455 maximum 2 mismatches in each 25 bp segment and a maximum of 20 alignment hits per read [40,41]. The
456 alignment results were sorted and indexed for downstream analyses as BAM format files. The aligned
457 reads were further utilized to generate gene expression counts using FeatureCounts (v2.0.0) against the

458 human reference annotation (GRCh38 build 99) [42]. Next, 60,067 genes in the expression matrix were
459 first normalized by variance stabilizing transformation method built in DESeq2 package in R [43].
460 Subsequently, the batch effect was removed for 20,000 most variable genes using combat function in
461 the SVA package [44].
462
463 **Statistical analysis**
464 The primary outcome examined in this study was TBM three-month mortality. In the descriptive
465 analysis, we summarized and tested association of patient characteristics with three-month mortality
466 using univariate Cox regression analysis. We presented the proportion for binary variables and median
467 (1st and 3rd interquartile range) for continuous variables.
468
469 To explore transcriptional profiles associated with mortality in TBM, differential expression analysis was
470 performed on the 20,000 normalized gene expression matrix to find differentially expressed genes
471 (DEGs). We constructed the contradicted matrix with survival and death status adjusted for covariates
472 including: age, HIV status, dexamethasone treatment (LAST-ACT trial investigators remained blind), and
473 sequencing batch. Linear models were used to assess DEGs using limma R package. Empirical Bayes
474 moderated-t p-values were computed for each genes and Benjamini-Hochberg were used for correcting
475 multiple testing (FDR). We defined the DEGs with the parameters (fold change > 1.5 and FDR <0.05). To
476 visualize samples clustering in the 4 cohorts, a unsupervised principal component analysis was
477 performed on 20,000 normalized genes, and the first principle component was plotted against the
478 second principle component. The 95% confidence ellipse was drawn for each group using multivariate t-
479 distribution ellipse. To compare enrichment activity of interested pathways between death vs survival,
480 TBM vs PTB or healthy controls, enrichment scores of these pathways were calculated for single patient

481 using single sample Gene Set Enrichment Analysis (GSEA) algorithm (ssGSEA). Pairwise Wilcoxon rank
482 sum test was used for comparison between any two cohorts.
483
484 In the primary analysis, we initially conducted *weighted gene co-expression network analysis* (WGCNA)
485 on the whole-blood transcriptomic profiling of the discovery cohort to identify clusters of genes (or
486 'modules') in the gene co-expression network among host transcriptomic genes. We then performed a
487 *network modules preservation analysis* to assess the reproducibility of these modules and the network
488 in the validation cohort. Subsequently, we conducted a *module-trait association analysis* to identify
489 modules associated with the clinical traits of mortality and TBM severity. For the association analysis
490 with mortality, we used a *Cox regression model* with the PC1 of the module as an independent variable.
491 We adjusted the model for age, HIV status, and corticosteroid treatment (dexamethasone vs. placebo).
492
493 To analyze the association with TBM severity, we calculated the *Spearman correlation* between TBM
494 severity (MRC grade) and the module's PC1. We corrected for false discovery rate (FDR) for the
495 dependent hypotheses by using the Benjamini-Yekutieli procedure. Target modules were defined as
496 modules associated with mortality in both the discovery and validation cohorts at FDR < 0.05. Hub genes
497 within each targeted module were identified based on 3 criteria: protein coding gene, module
498 membership cut off 0.85 (MM) and high rank of gene significance rank (GS). MM defined as correlation
499 between individual gene and the module's PC1. GS defined as $-\log_{10}$ p value association of gene with
500 mortality using Cox regression model adjusted for age, HIV status and treatment. Genes were first
501 filtered for protein coding function and MM above 0.85. Genes then was ranked based on its GS and top
502 3% were selected if number of genes in module above 500, otherwise top 5% were selected. Overlap
503 hub genes between discovery and validation cohort were consider validated.
504

505 In the subsequent analysis, to determine the biological funcions or processes potentially related to each
506 module we conducted *overrepresentation analysis* (ORA) using ShinyGO v0.77 [19]. Genes were first filters
507 for association with three-month mortality at $p < 0.05$ and then were input into ShinyGO. We used both
508 *GO and KEGG databases* for this analysis. We pre-specified significant pathways with overlap gene in
509 pathway > 5 and FDR < 0.05.
510
511 Furthermore, to compare *enrichment activity* of resulting pathways between death vs survival, TBM vs
512 PTB or healthy controls, enrichment scores of these pathways were calculated for single patient using
513 ssGSEA algorithm [45]. Pairwise Wilcoxon rank rum test were applied for comparison between 2 cohorts.
514 To compare or visualize the difference of pathway effect on mortality between HIV-negative and HIV-
515 positive TBM, *fold change enrichment* of interested pathways were calculated using Quantitative Set
516 Analysis for Gene Expression (QuSAGE) method [46].
517
518 In the secondary analysis, we performed a *consensus analysis* to identify consensus patterns of co-
519 expression networks between HIV-positive and HIV-negative conditions in our TBM cohort. We
520 conducted similar association analyses for TBM severity and mortality as in the preceding WGCNA
521 analysis for both the HIV-negative and HIV-positive cohorts. We visualized the results using a heatmap.
522 We defined strongly associated modules in HIV-positive or HIV-negative cohort as HIV-positive-specific
523 signals and HIV-negative-specific signals. Additionally, we functional enrichmet ORA was performed for
524 these consensus modules and the HIV-specific modules
525
526 We then performed variable selection using a *multivariate elastic-net Cox regression model* to select the
527 important predictors for HIV-negative and HIV-positive TBM prognosis separately. Candidate predictors
528 in HIV-negative TBM consisted of the combination of common hub genes, HIV-negative specific hub

529 genes and clinical biomarkers: age, TBM severity and CSF lymphocyte count. Similarly, for HIV-positive
530 TBM, candidate predictors consisted of common hub genes, HIV-positive specific hub genes and clinical
531 biomarkers. We performed the analysis with 1,000 bootstrap sampling with replacement. The chosen
532 variables were 1) among the top 75% of selected variables and 2) one representative hub gene per
533 module. Subsequently, we developed a prediction model based on a *logistic regression model*
534 incorporating only the chosen variables. We also assessed the overall model performance (by *optimism-*
535 *corrected Brier score*), the discrimination (by *optimism-corrected AUC*), and calibration (by *optimism-*
536 *corrected calibration-slope*) of the developed model [47].
537
538 More details on performing ORA, SSGSEA, QuSAGE and WGCNA preservation and consensus analyses
539 can be found in the *Supplementary method document.*
540

541 **Acknowledgements**

542 We would like to thank and the clinical staff who recruited patients into our study from Hospital for
543 Tropical Diseases and Pham Ngoc Thach hospital, Ho Chi Minh city, Vietnam, and all participants.

544

545 **Funding**

546 This work was supported by Wellcome Trust Fellowship in Public Health and Tropical Medicine to NTTT
547 (206724/Z/17/Z), Wellcome Trust Investigator Award to GT and Wellcome Trust to Vietnam Africa Asia
548 Programme.

549

550 **Conflicts of Interest**

551 No conflicts of interest

552

553 **References**

554 1. Dodd, P.J. *et al.* The global burden of tuberculous meningitis in adults: A modelling study. *PLOS*
555 *Glob Public Health* **1**, e0000069 (2021).
556 2. Stadelman, A.M. *et al.* Treatment Outcomes in Adult Tuberculous Meningitis: A Systematic
557 Review and Meta-analysis. *Open Forum Infect Dis* **7**, ofaa257 (2020).
558 3. Wilkinson, R.J. *et al.* Tuberculous meningitis. *Nat Rev Neurol* **13**, 581-598 (2017).
559 4. Huynh, J. *et al.* Tuberculous meningitis: progress and remaining questions. *Lancet Neurol* **21**,
560 450-464 (2022).
561 5. Marais, S. *et al.* Inflammasome Activation Underlying Central Nervous System Deterioration in
562 HIV-Associated Tuberculosis. *J Infect Dis* **215**, 677-686 (2017).
563 6. Rohlwink, U.K. *et al.* Biomarkers of Cerebral Injury and Inflammation in Pediatric Tuberculous
564 Meningitis. *Clin Infect Dis* **65**, 1298-1307 (2017).
565 7. Thuong, N.T.T. *et al.* Leukotriene A4 Hydrolase Genotype and HIV Infection Influence
566 Intracerebral Inflammation and Survival From Tuberculous Meningitis. *J Infect Dis* **215**, 1020-
567 1028 (2017).
568 8. Thuong, N.T.T. *et al.* Pretreatment Cerebrospinal Fluid Bacterial Load Correlates With
569 Inflammatory Response and Predicts Neurological Events During Tuberculous Meningitis
570 Treatment. *J Infect Dis* **219**, 986-995 (2019).
571 9. van Laarhoven, A. *et al.* Clinical Parameters, Routine Inflammatory Markers, and LTA4H
572 Genotype as Predictors of Mortality Among 608 Patients With Tuberculous Meningitis in
573 Indonesia. *J Infect Dis* **215**, 1029-1039 (2017).
574 10. van Laarhoven, A. *et al.* Immune cell characteristics and cytokine responses in adult HIV-
575 negative tuberculous meningitis: an observational cohort study. *Sci Rep* **9**, 884 (2019).

576 11. Cresswell, F.V. *et al.* Recent Developments in Tuberculous Meningitis Pathogenesis and
577 Diagnostics. *Wellcome Open Res* **4**, 164 (2019).
578 12. Rohlwink, U.K. *et al.* Tuberculous meningitis in children is characterized by compartmentalized
579 immune responses and neural excitotoxicity. *Nat Commun* **10**, 3767 (2019).
580 13. Ardiansyah, E. *et al.* Tryptophan metabolism determines outcome in tuberculous meningitis: a
581 targeted metabolomic analysis. *Elife* **12**(2023).
582 14. STREPTOMYCIN treatment of tuberculous meningitis. *Lancet* **1**, 582-96 (1948).
583 15. Thao, L.T.P. *et al.* Prognostic Models for 9-Month Mortality in Tuberculous Meningitis. *Clin Infect*
584 *Dis* **66**, 523-532 (2018).
585 16. Sweeney, T.E., Braviak, L., Tato, C.M. & Khatri, P. Genome-wide expression for diagnosis of
586 pulmonary tuberculosis: a multicohort analysis. *Lancet Respir Med* **4**, 213-24 (2016).
587 17. Penn-Nicholson, A. *et al.* RISK6, a 6-gene transcriptomic signature of TB disease risk, diagnosis
588 and treatment response. *Sci Rep* **10**, 8629 (2020).
589 18. Marais, S. *et al.* Tuberculous meningitis: a uniform case definition for use in clinical research.
590 *Lancet Infect Dis* **10**, 803-12 (2010).
591 19. Ge, S.X., Jung, D. & Yao, R. ShinyGO: a graphical gene-set enrichment tool for animals and
592 plants. *Bioinformatics* **36**, 2628-2629 (2020).
593 20. Berry, M.P. *et al.* An interferon-inducible neutrophil-driven blood transcriptional signature in
594 human tuberculosis. *Nature* **466**, 973-7 (2010).
595 21. Singhania, A. *et al.* A modular transcriptional signature identifies phenotypic heterogeneity of
596 human tuberculosis infection. *Nat Commun* **9**, 2308 (2018).
597 22. Geng, X. *et al.* Whole transcriptome sequencing reveals neutrophils' transcriptional landscape
598 associated with active tuberculosis. *Front Immunol* **13**, 954221 (2022).
599 23. Castro-Alcaraz, S., Miskolci, V., Kalasapudi, B., Davidson, D. & Vancurova, I. NF-kappa B
600 regulation in human neutrophils by nuclear I kappa B alpha: correlation to apoptosis. *J Immunol*
601 **169**, 3947-53 (2002).
602 24. Sabroe, I., Dower, S.K. & Whyte, M.K. The role of Toll-like receptors in the regulation of
603 neutrophil migration, activation, and apoptosis. *Clin Infect Dis* **41 Suppl 7**, S421-6 (2005).
604 25. Davoudi, S. *et al.* CD4+ cell counts in patients with different clinical manifestations of
605 tuberculosis. *Braz J Infect Dis* **12**, 483-6 (2008).
606 26. Shridhar, A. *et al.* Prevalence of primary immunodeficiency syndromes in tuberculous
607 meningitis: A case-control study. *J Infect Public Health* **15**, 29-35 (2022).
608 27. An, H.R. *et al.* The relationship between absolute counts of lymphocyte subsets and clinical
609 features in patients with pulmonary tuberculosis. *Clin Respir J* **16**, 369-379 (2022).
610 28. Hemingway, C. *et al.* Childhood tuberculosis is associated with decreased abundance of T cell
611 gene transcripts and impaired T cell function. *PLoS One* **12**, e0185973 (2017).
612 29. Lu, L.L. *et al.* A Functional Role for Antibodies in Tuberculosis. *Cell* **167**, 433-443 e14 (2016).
613 30. La Manna, M.P. *et al.* Impact of Mycobacterium tuberculosis Infection on Human B Cell
614 Compartment and Antibody Responses. *Cells* **11**(2022).
615 31. Thwaites, G.E. *et al.* Dexamethasone for the treatment of tuberculous meningitis in adolescents
616 and adults. *N Engl J Med* **351**, 1741-51 (2004).
617 32. Kumar, N.P., Banurekha, V.V., Nair, D. & Babu, S. Circulating Angiogenic Factors as Biomarkers of
618 Disease Severity and Bacterial Burden in Pulmonary Tuberculosis. *PLoS One* **11**, e0146318
619 (2016).
620 33. Thao, L.T.P. *et al.* Dynamic Prediction of Death in Patients With Tuberculous Meningitis Using
621 Time-updated Glasgow Coma Scale and Plasma Sodium Measurements. *Clin Infect Dis* **70**, 827-
622 834 (2020).

623 34. Feng, B. *et al.* Prognostic factors of adult tuberculous meningitis in intensive care unit: a single-
624 center retrospective study in East China. *BMC Neurol* **21**, 308 (2021).
625 35. Sharawat, I.K. *et al.* Development and Validation of a Prognostic Model and Bedside Score for
626 the Neurological Outcome in Children with Tuberculous Meningitis. *Am J Trop Med Hyg* **107**,
627 1288-1294 (2022).
628 36. Donovan, J. *et al.* Adjunctive dexamethasone for the treatment of HIV-uninfected adults with
629 tuberculous meningitis stratified by Leukotriene A4 hydrolase genotype (LAST ACT): Study
630 protocol for a randomised double blind placebo controlled non-inferiority trial. *Wellcome Open*
631 *Res* **3**, 32 (2018).
632 37. Donovan, J. *et al.* Adjunctive dexamethasone for the treatment of HIV-infected adults with
633 tuberculous meningitis (ACT HIV): Study protocol for a randomised controlled trial. *Wellcome*
634 *Open Res* **3**, 31 (2018).
635 38. Yalamanchili, H.K., Wan, Y.W. & Liu, Z. Data Analysis Pipeline for RNA-seq Experiments: From
636 Differential Expression to Cryptic Splicing. *Curr Protoc Bioinformatics* **59**, 11 15 1-11 15 21
637 (2017).
638 39. Conesa, A. *et al.* A survey of best practices for RNA-seq data analysis. *Genome Biol* **17**, 13 (2016).
639 40. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2013).
640 41. Bolger, A.M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence
641 data. *Bioinformatics* **30**, 2114-20 (2014).
642 42. Liao, Y., Smyth, G.K. & Shi, W. featureCounts: an efficient general purpose program for assigning
643 sequence reads to genomic features. *Bioinformatics* **30**, 923-30 (2014).
644 43. Love, M.I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-
645 seq data with DESeq2. *Genome Biol* **15**, 550 (2014).
646 44. Leek, J.T., Johnson, W.E., Parker, H.S., Jaffe, A.E. & Storey, J.D. The sva package for removing
647 batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* **28**,
648 882-3 (2012).
649 45. Barbie, D.A. *et al.* Systematic RNA interference reveals that oncogenic KRAS-driven cancers
650 require TBK1. *Nature* **462**, 108-12 (2009).
651 46. Yaari, G., Bolen, C.R., Thakar, J. & Kleinstein, S.H. Quantitative set analysis for gene expression: a
652 method to quantify gene set differential expression including gene-gene correlations. *Nucleic*
653 *Acids Res* **41**, e170 (2013).
654 47. Steyerberg, E.W. *et al.* Internal validation of predictive models: efficiency of some procedures
655 for logistic regression analysis. *J Clin Epidemiol* **54**, 774-81 (2001).

656

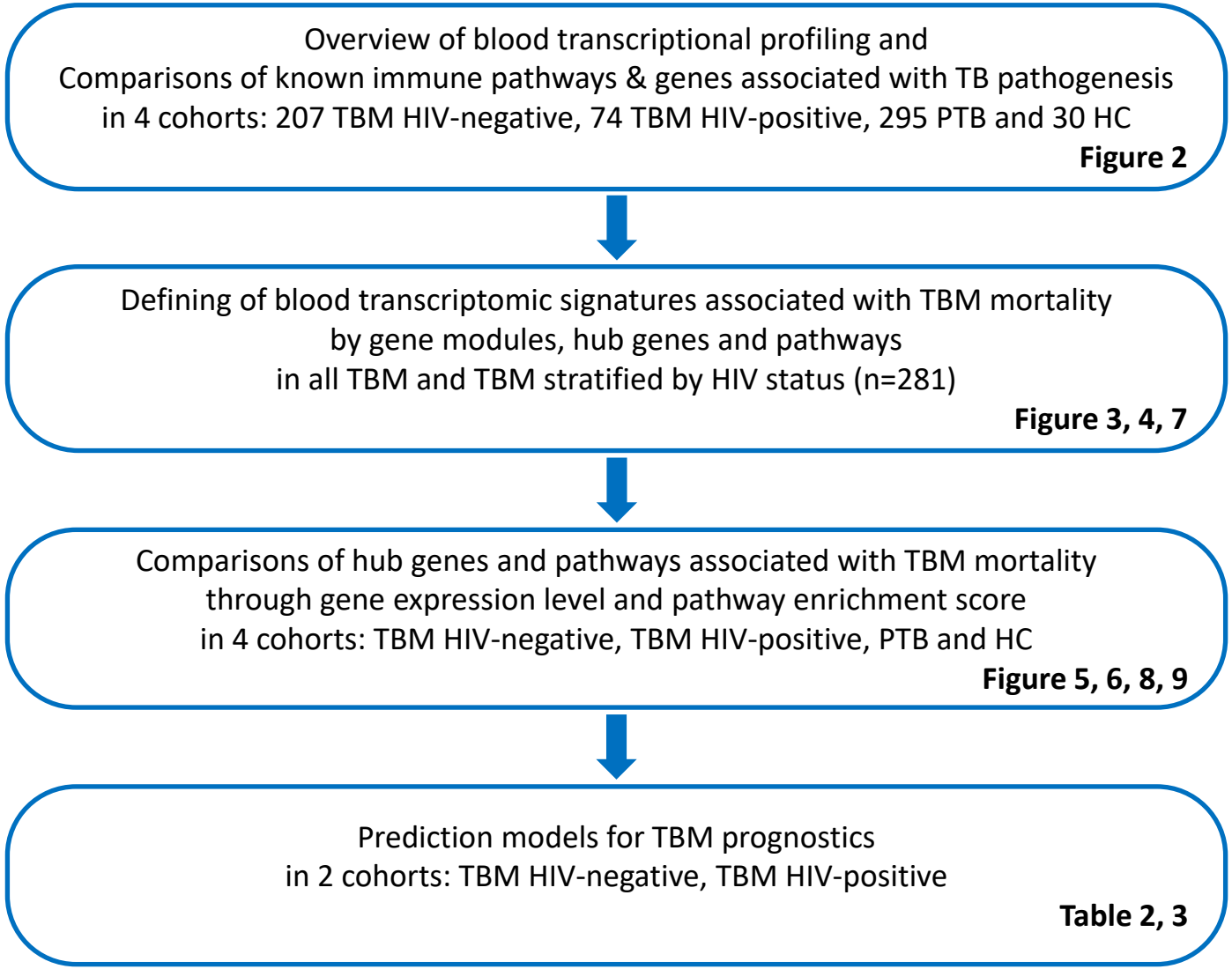


Figure 1. Objectives and cohorts flow

TBM: TB meningitis, HIV: human immunodeficiency virus, PTB: pulmonary TB, HC; healthy controls

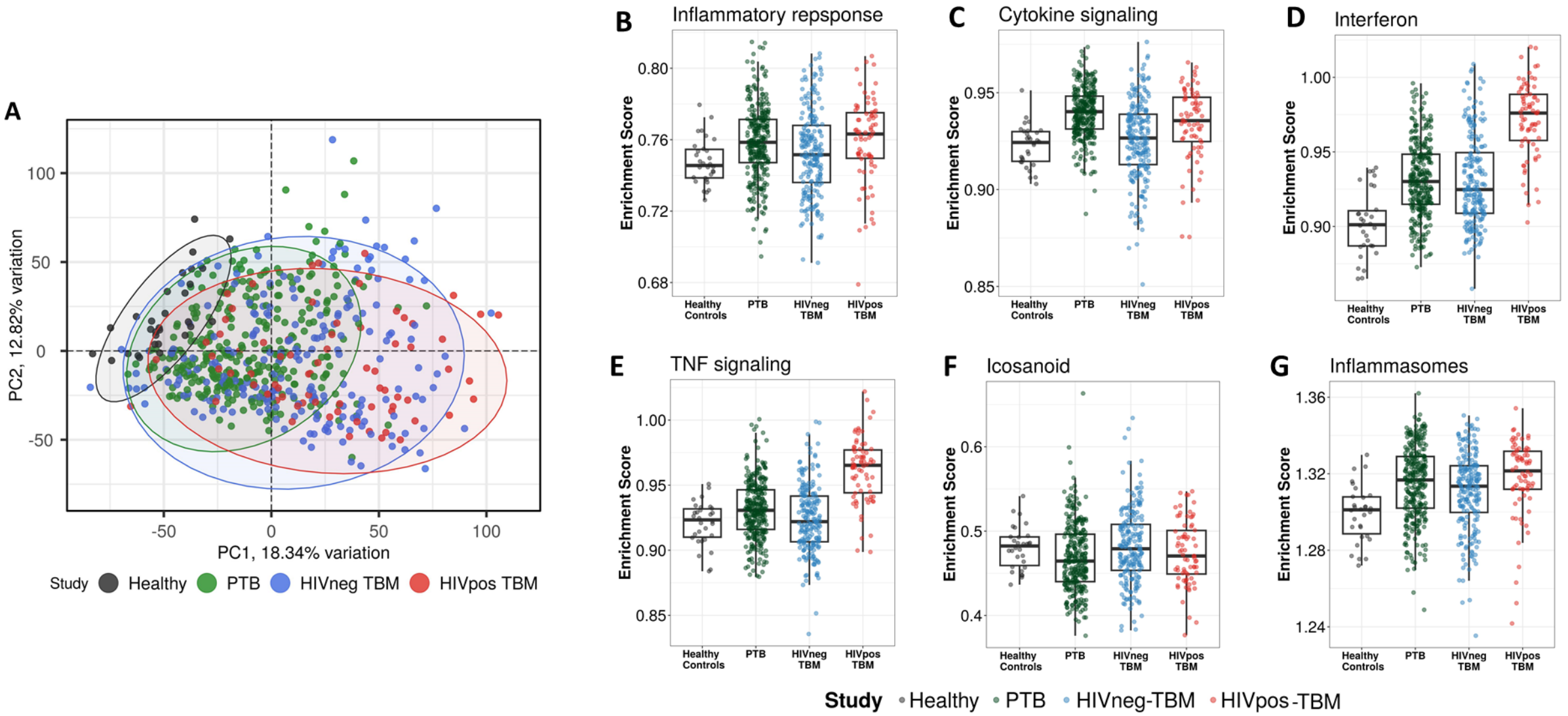


Figure 2: Blood transcriptomic profiles of four cohorts: healthy controls (n=30), PTB (n=295), HIV-negative TBM (n=207) and HIV-positive TBM (n=74)
(**A**) Principle component analysis (PCA) of whole transcriptomic profile of HC, PTB PTB and TBM with and without HIV. Each symbol represents one individual with color coding different cohorts. The x-axis represents principle component (PC) 1, while y-axis represents PC2. **(B-G)** Enrichment scores of known innate immunity pathways associated with TBM pathogenesis. Gene list of these pathway were depicted in additional file 1, table S5. Pathway enrichment scores were calculated using single sample GSEA algorithm (ssGSEA) (Barbie DA, et al 2009). Each dot represents one participant. The box presents median, 25th to 75th percentile and the whiskers present the minimum to the maximum points in the data.

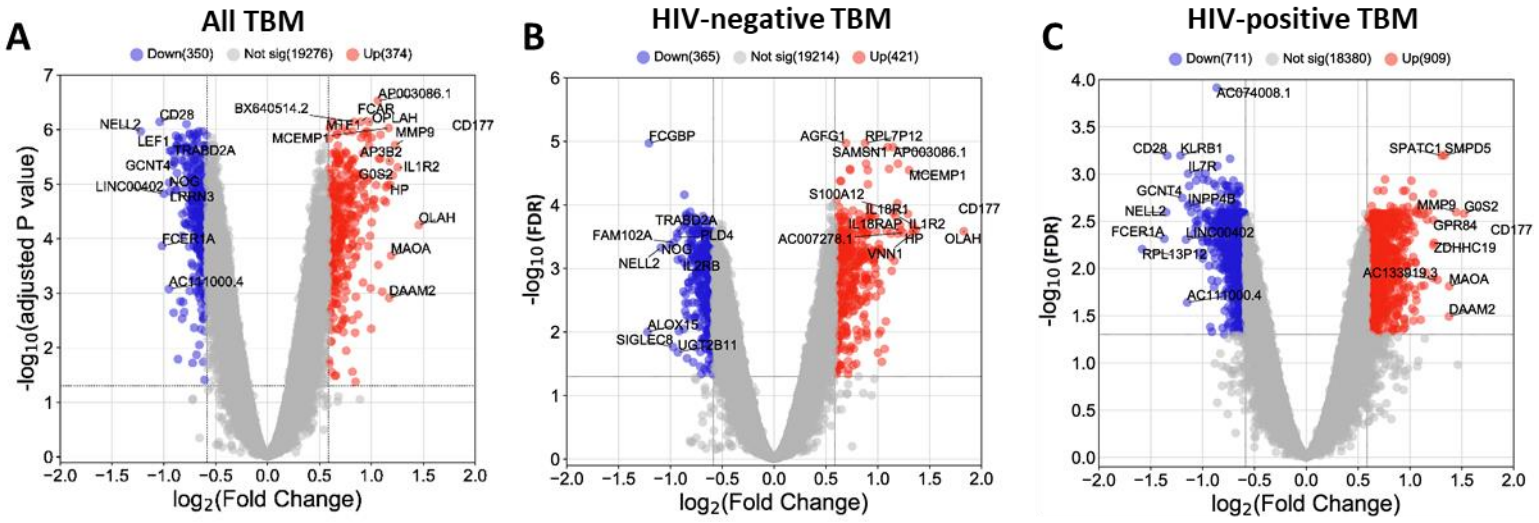


Figure 3: Blood transcriptomic profiles of three-month mortality at baseline in all TBM and TBM stratified by HIV status
Volcano plot showed differential expression (DE) genes by fold change (FC) between death and survival in all TBM **(A)**, HIV-negative **(B)** and HIV-positive TBM **(C)**. Each dot represents one gene. The x-axis represents $\log_2$ FC. The y-axis showed $-\log_{10}$ FDR of genes. DE genes were colored with red indicating up-regulated, blue indicating down-regulated genes which having fold discovery rate (FDR) <0.05 and absolute FC > 1.5.

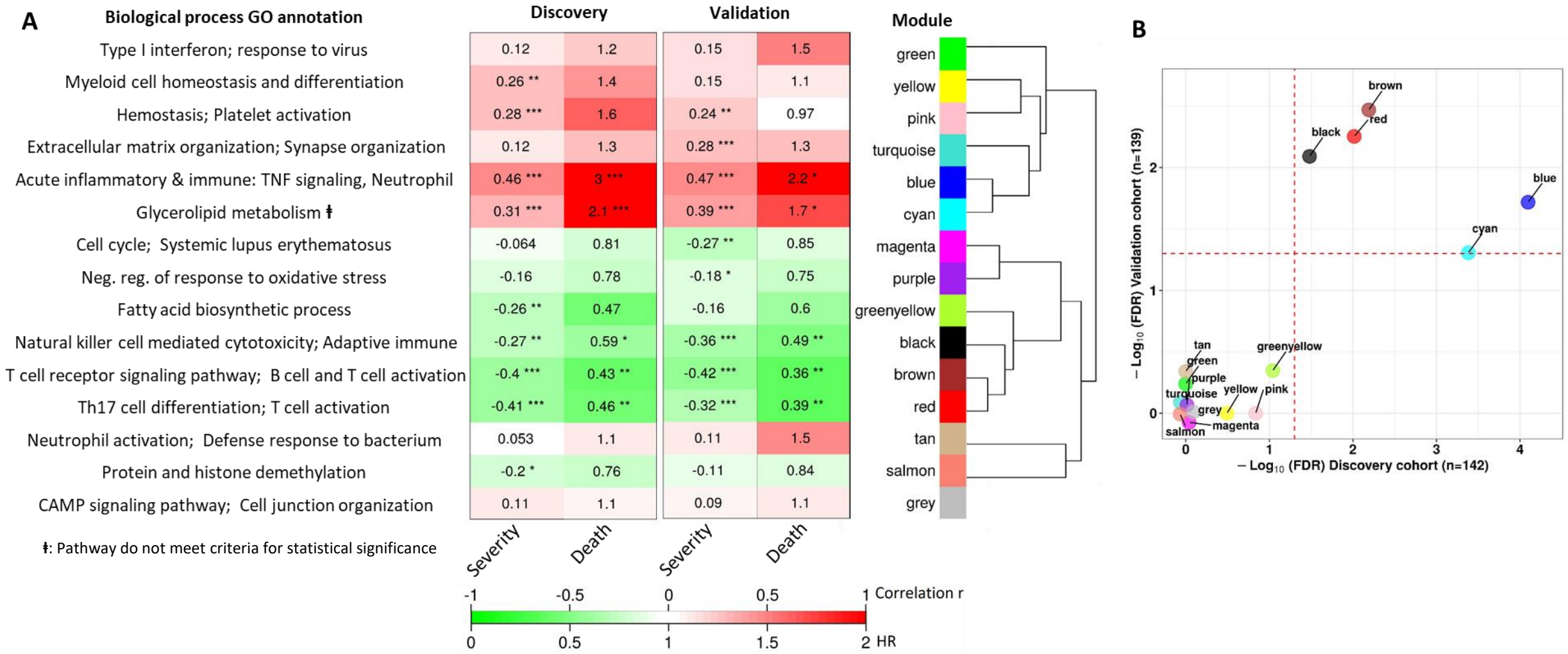


Figure 4. Blood transcriptional modules associated with mortality in TBM

(A) Associations between WGCNA modules with two clinical phenotypes TBM disease severity (MRC grade) and three-month mortality in discovery and validation cohorts, and their associated biological processes. The heatmap showed the association between principle component 1 (PC1) of each module and the phenotypes, particularly Spearman correlation r for MRC grade and hazard ratio per increase 1/10 unit of PC1 (HR) for mortality. The HRs were estimated using a Cox regression model adjusted for age, HIV status and dexamethasone treatment. False discovery rate (FDR) corrected based on Benjamini & Yekutieli procedure, with significant level denoted as * < .05, ** < .01 and *** <.001. Gradient colors were used to fill the cell with green indicating negative r or HR < 1, red color indicating positive r or HR > 1. The order of modules was based on hierarchical clustering using Pearson correlation distance for module eigengene. On the left, biological processes, corresponding to modules, were identified using Gene Ontology and KEGG database. **(B)** Validation of the association between WGCNA modules and mortality in discovery and validation cohorts. X-axis represents $-\log_{10}$ FDR in discovery cohort and Y-axis represents $-\log_{10}$ FDR in validation cohort. Red dash lines indicate FDR = 0.05 as the threshold for statistically significant in both cohorts. Five modules (blue, brown, red, black and cyan) with FDR < 0.05 were validated.

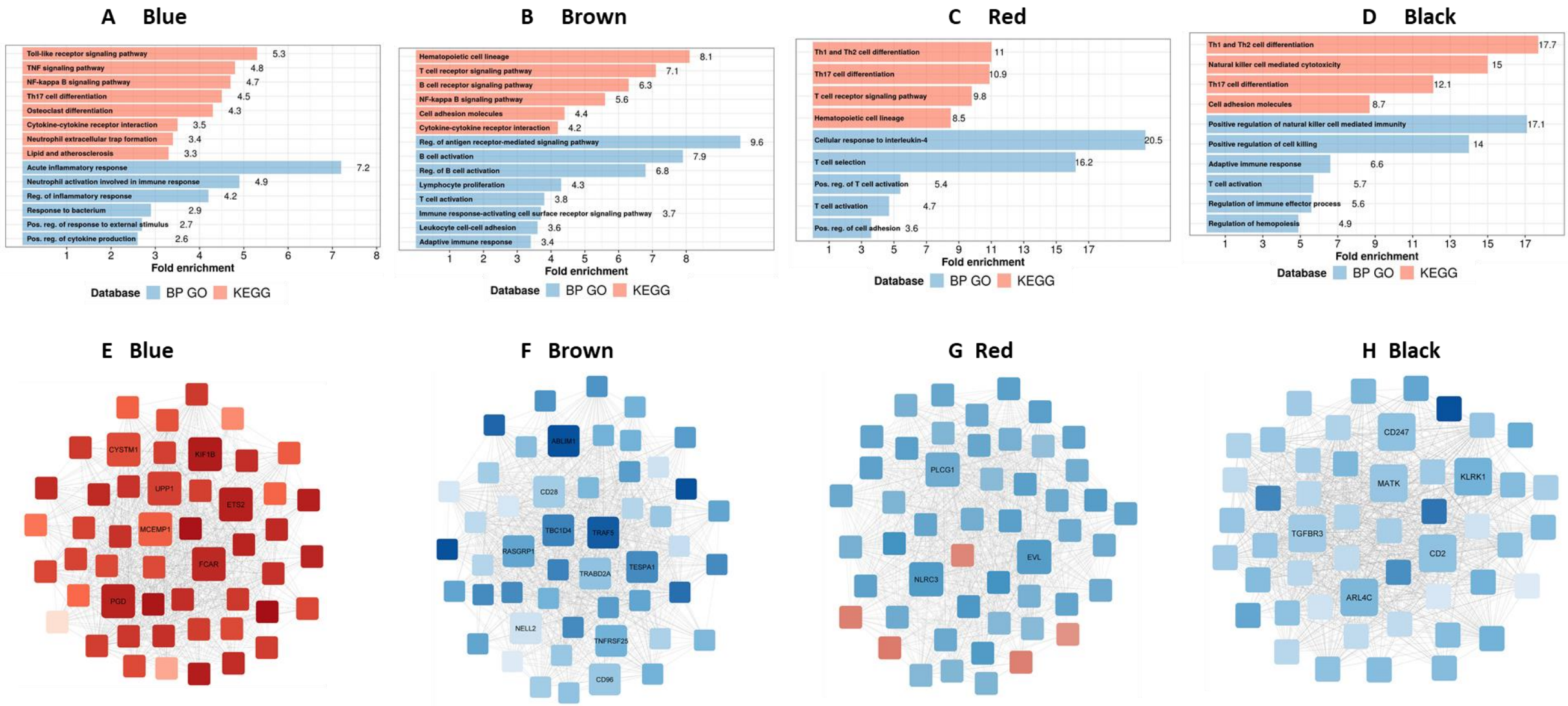


Figure 5. Biological processes, pathways and hub genes of validated modules associated with mortality

(A-D) showed biological processes and pathways identified in four mortality associated modules: blue, brown, red and black module, by over representation analysis (ORA). Bar plots show the top representative GO biological processes or KEGG pathways. The bars indicates biological processes or pathways having ORA FDR < 0.05 and size corresponding to fold enrichment calculated as the ratio of gene number of pathway in the input list divided by the ratio of gene number of the pathway in reference. **(E-H)** showed gene co-expression networks and hub genes of blue, brown, red and black module, respectively. Each node represents one gene. Each edge represents the link between two genes. Hub genes were shown by bigger nodes and bold text. The gradient color of node corresponds to its HR to mortality, with red indicating HR > 1, and blue HR < 1.

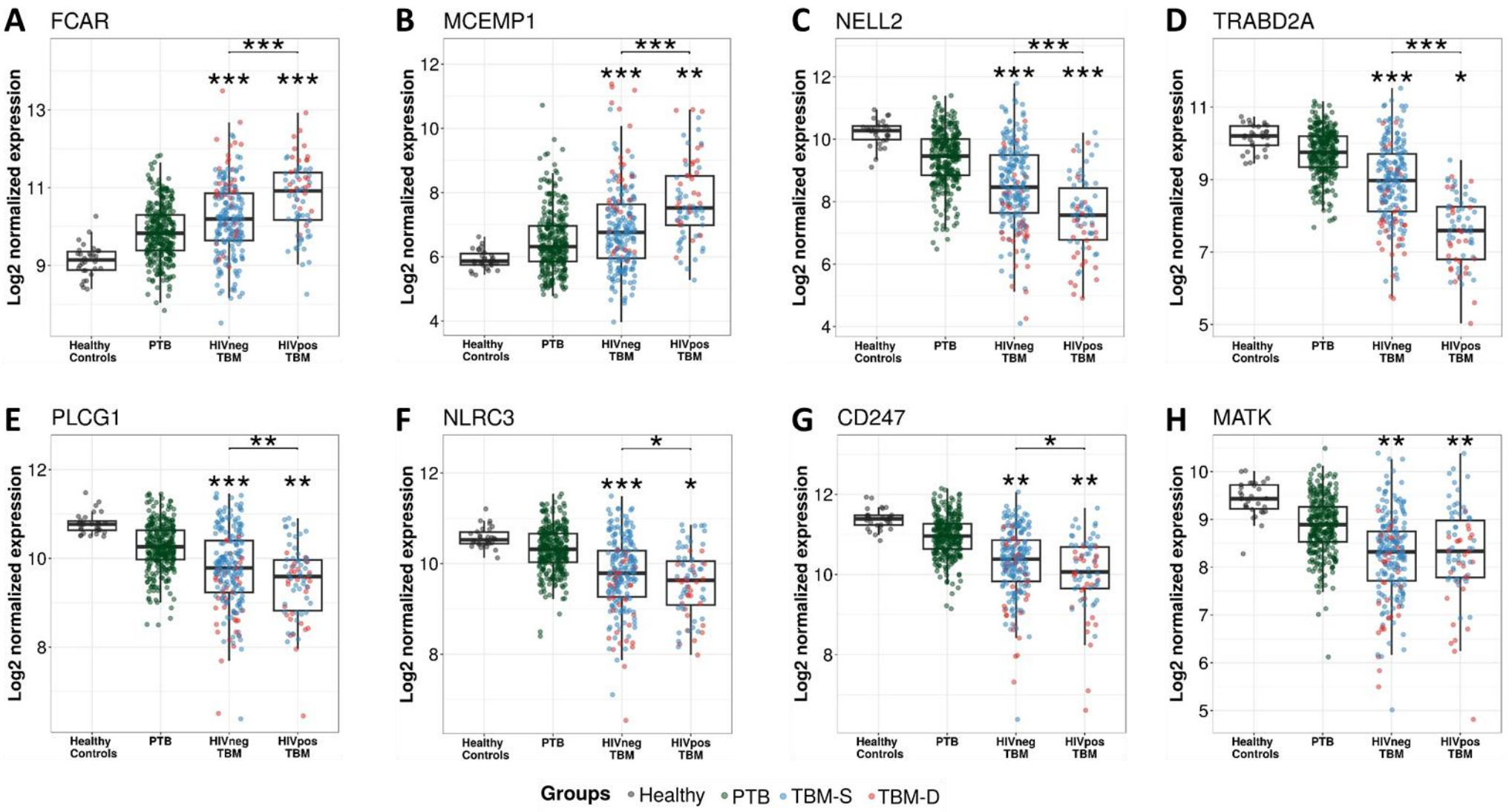
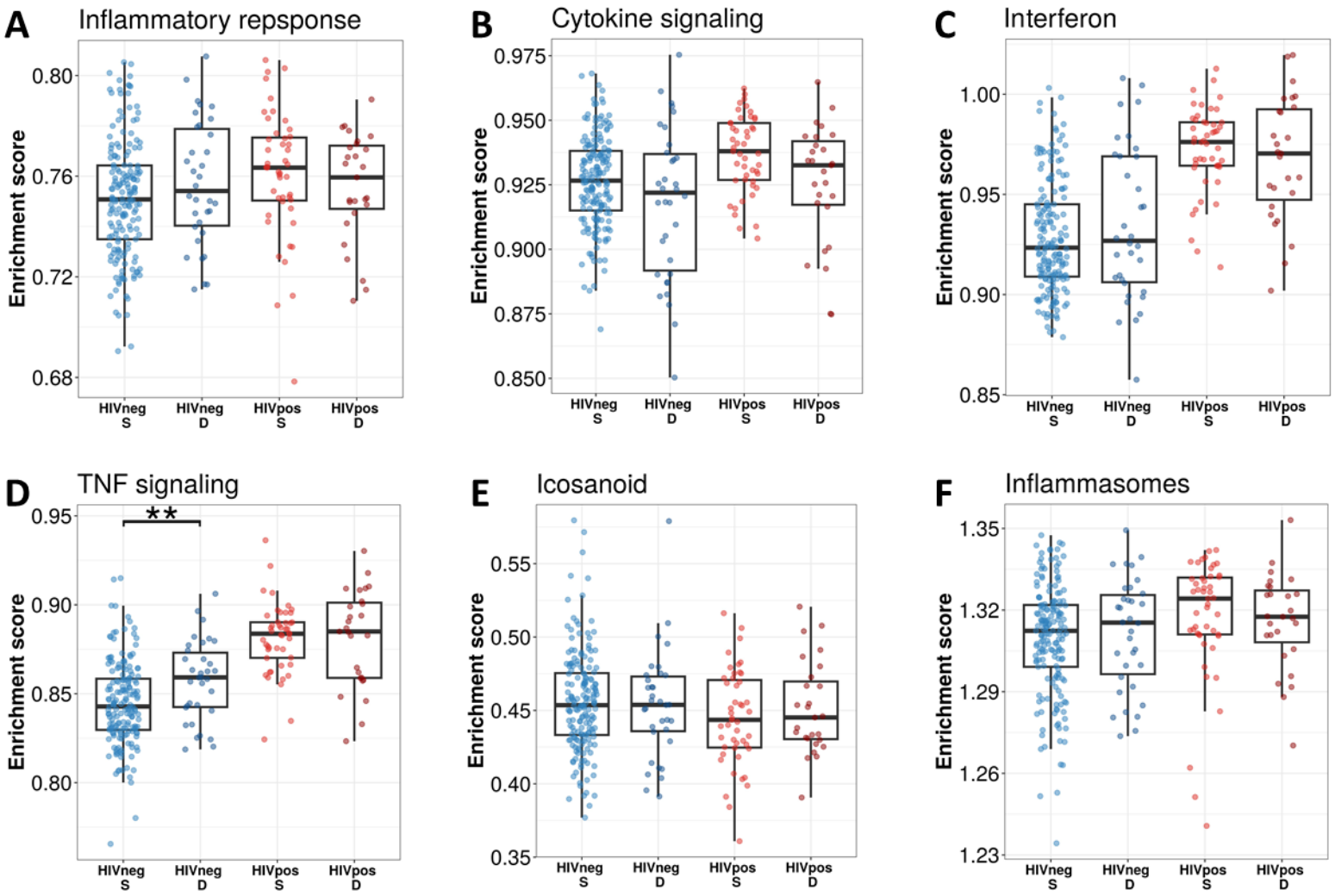


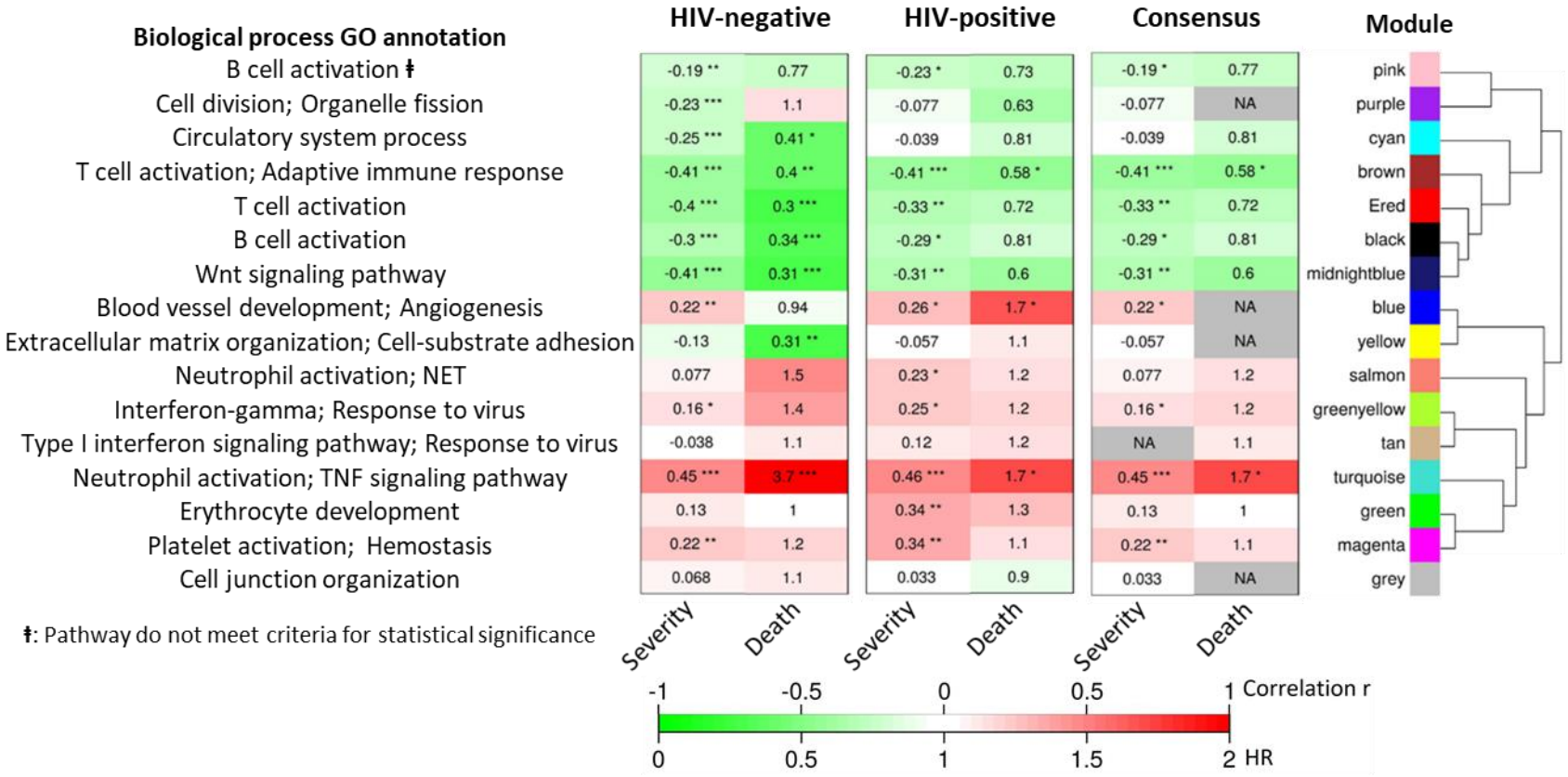
Figure 6. Gene expression of representative hub genes in healthy controls (n=30), PTB (n=295), HIV-negative TBM (n=207) and HIV-positive TBM (n=74)
Each dot represents gene expression from one participant. **(A, B)** expression of FCAR and MCEMP1 hub genes from the blue module. **(C, D)** expression of NELL2 and TRABD2A hub genes from the brown modules. **(E, F)** expression of PLCG1 and NLRC3 hub genes from the red module. **(G, H)** expression of CD247 and MATK hub genes from the black module. The box presents median, 25th to 75th percentile and the whiskers present the minimum to the maximum points in the data. Comparisons were made between dead (red) with survival (blue) or between HIV-negative and HIV-positive TBM by Wilcoxon rank sum test with p-values displayed as significance level above the boxes and the horizontal bars (* < .05, ** < .01, *** <.001).

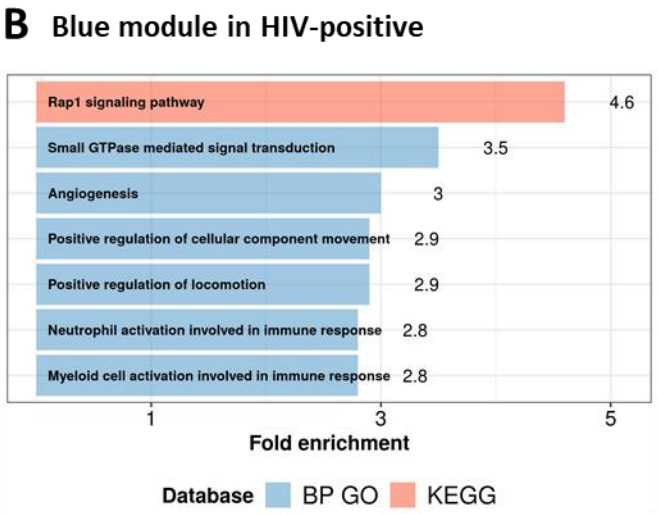


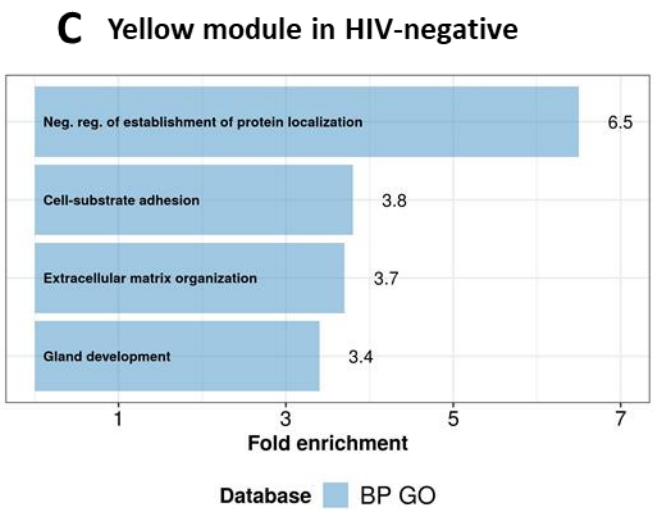
* HIVneg S: HIV-negative TBM survival, HIVneg D: HIV-negative TBM death, HIVpos S: HIV-positive TBM survival, HIVpos D: HIV-positive TBM death

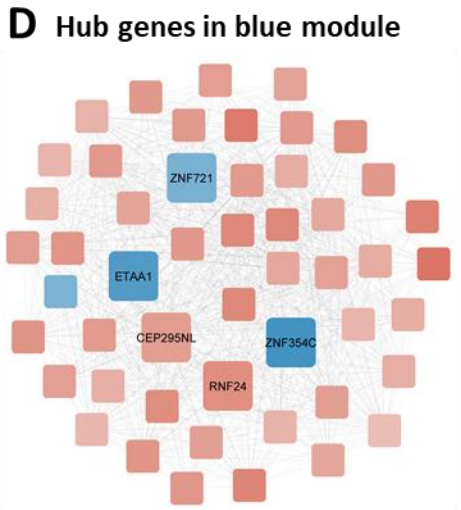
Figure 7: Relationship between known pathways associated with TBM pathogenesis and mortality

Enrichment scores of known immune pathways associated with TBM pathogenesis. Gene list of these pathway were depicted in additional file 1, table S5. Pathway enrichment scores were calculated using single sample GSEA algorithm (ssGSEA) (Barbie DA, et al 2009). Each dot represents one participant. The box presents median, 25th to 75th percentile and the whiskers present the minimum to the maximum points in the data. The comparisons were made between survival and death using Wilcoxon rank sum test. Only significant results are presented with * < .05, ** < .01, *** <.001.


A
Biological process GO annotation
HIV-negative
HIV-positive
Consensus
Module
B cell activation ǂ
Cell division; Organelle fission
Circulatory system process
T cell activation; Adaptive immune response
T cell activation
B cell activation
Wnt signaling pathway
Blood vessel development; Angiogenesis
Extracellular matrix organization; Cell-substrate adhesion
Neutrophil activation; NET
Interferon-gamma; Response to virus
Type I interferon signaling pathway; Response to virus
Neutrophil activation; TNF signaling pathway
Erythrocyte development
Platelet activation; Hemostasis
Cell junction organization
ǂ: Pathway do not meet criteria for statistical significance
Severity
Death
pink
purple
cyan
brown
Ered
black
midnightblue
blue
yellow
salmon
greenyellow
tan
turquoise
green
magenta
grey
Correlation r
HR


B Blue module in HIV-positive
Rap1 signaling pathway 4.6
Small GTPase mediated signal transduction 3.5
Angiogenesis 3
Positive regulation of cellular component movement 2.9
Positive regulation of locomotion 2.9
Neutrophil activation involved in immune response 2.8
Myeloid cell activation involved in immune response 2.8
Fold enrichment
Database BP GO KEGG


C Yellow module in HIV-negative
Neg. reg. of establishment of protein localization 6.5
Cell-substrate adhesion 3.8
Extracellular matrix organization 3.7
Gland development 3.4
Fold enrichment
Database BP GO


D Hub genes in blue module
ZNF721
ETAA1
CEP295NL
ZNF354C
RNF24

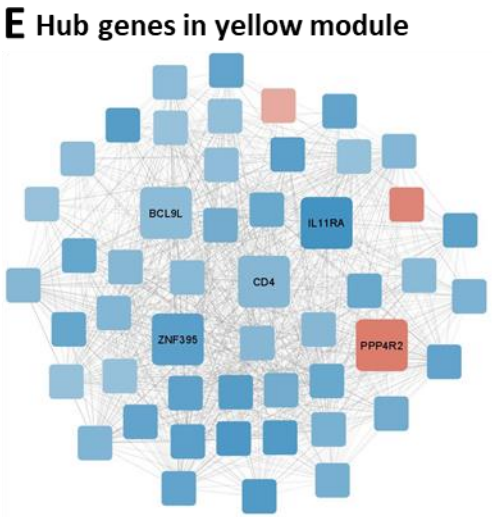

E Hub genes in yellow module
BCL9L
IL11RA
CD4
ZNF395
PPP4R2

Figure 8. Consensus transcriptional modules associated with TBM mortality stratified by HIV-infection

(A) Associations between 16 consensus WGCNA modules with two clinical phenotypes TBM severity (MRC grade) and mortality in HIV-negative (n=207) and HIV-positive (n=74) TBM participants, and their associated BP Gene ontology or KEEG database. The heatmap showed the association between modules and the phenotypes, with Spearman correlation r for MRC grade and hazard ratio per increase 1 unit of PC1 of module (HR) for mortality in HIV-positive and HIV-negative cohorts. The consensus sub-panel presented associations of the consensus modules and clinical phenotypes with same trend detected in both HIV cohorts, otherwise were annotated with missing (NA) values. False discovery rate (FDR) corrected using Benjamini & Yekutieli procedure, with significant level denoted as * < .05, ** < .01 and *** <.001. Gradient colors were used to fill the cell with green indicating negative r or HR < 1, red color indicating positive r or HR > 1. The order of modules was based on hierarchical clustering using Pearson correlation distance for module eigengene. It is noted that these consensus modules were not identical to the identified modules in the primary analysis in Figure 1. **(B-C)** Functional enrichment analysis of HIV-positive pathway (blue module) and HIV-negative pathway (yellow module), respectively. **(D-E)** Gene co-expression network of blue and yellow modules. Each node represents one gene. Each edge represents the link between two genes. Hub genes were shown by bigger nodes with bold text. The gradient color of node corresponds to its HR to mortality, with red indicating HR>1, and blue HR<1.

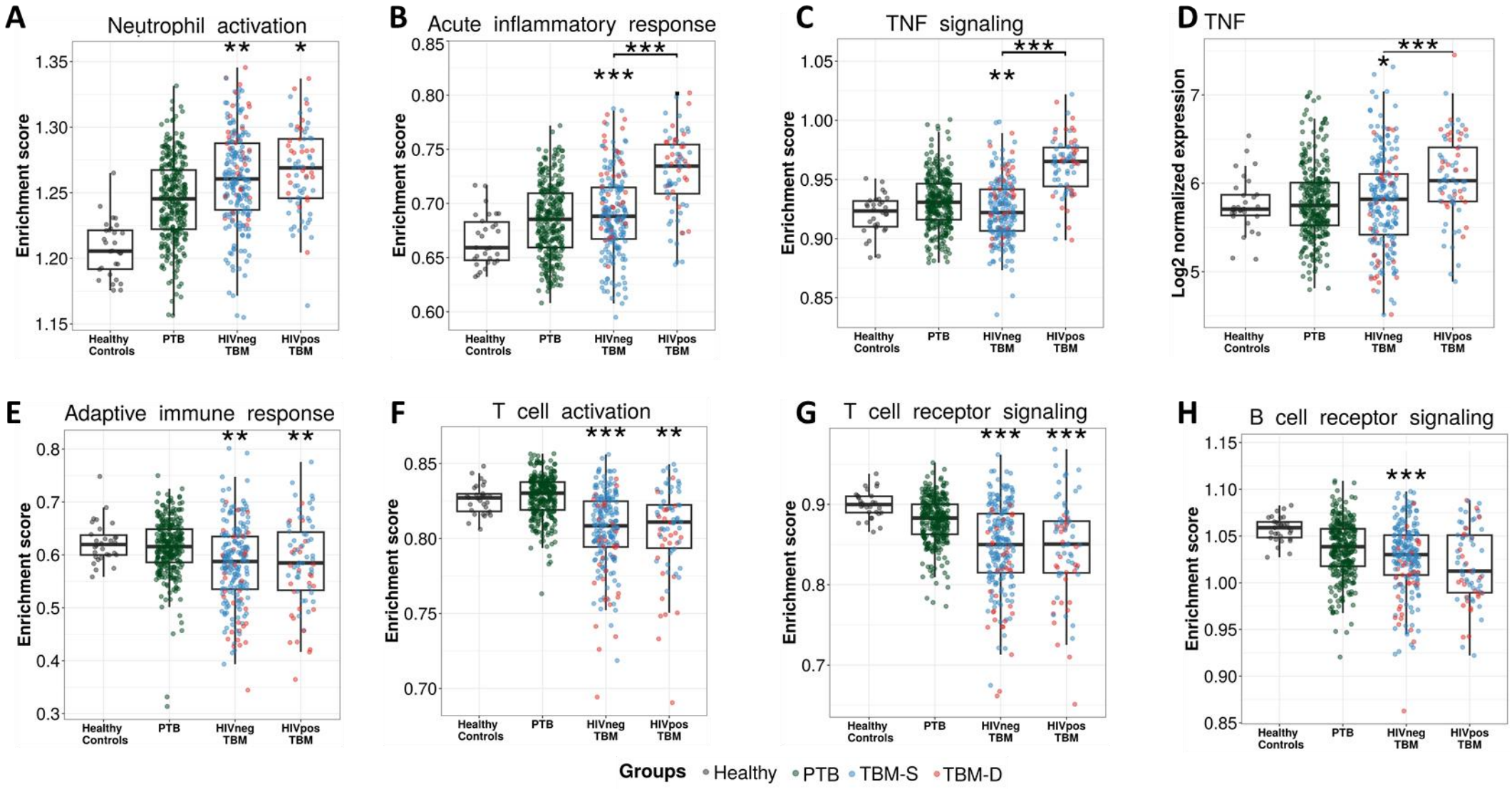


Figure 9. Enrichment score of immunity pathways in healthy controls (n=30), PTB (n=295), HIV-negative TBM (n=207) and HIV-positive TBM (n=74)

Pathway enrichment scores were calculated using single sample GSEA algorithm (ssGSEA) (Barbie DA, et al 2009). Each dot represents one participant. **(A-C)** showed box-plots depicting enrichment scores of the innate immunity pathways from the blue module. **(E-H)** enrichment scores of the adaptive immunity pathways from the red and brown modules. **(D)** normalized expression of TNF. Gene list of these pathway were depicted in additional file 1, table S5. The box presents median, 25th to 75th percentile and the whiskers present the minimum to the maximum points in the data. Comparisons were made between dead (red) with survival (blue) or between HIV-negative and HIV-positive TBM by Wilcoxon rank sum test with p-values displayed as significance level above the boxes and the horizontal bars, respectively (* < .05, ** < .01, *** <.001).

Table 1. Baseline characteristics of TBM, PTB, and healthy controls

Characteristics	HIV-negative TBM n=207		HIV-positive TBM n=74		PTB n=295		Healthy controls n=30	
	n	Summary	n	Summary	n	Summary	n	Summary
Age (years)	207	46 (34, 58)	74	34 (29, 40)	295	44 (31, 52)	30	33 (29, 37)
Male sex	207	127 (61)	74	56 (76)	295	228 (77)	30	11 (37)
BMI (kg/m2)	205	20.0 (18.2, 22.3)	72	19.3 (17.2, 20.4)	295	19.4 (17.7, 21.6)		
Symptom duration (days)	207	14 (11, 20)	73	16 (10, 30)	294	20 (10, 30)		
History of TB treatment	204	5 (2.5)	74	15 (20)	295	100 (34)		
Glasgow coma score	207	14 (12, 15)	72	14 (13, 15)				
Cavity chest X-ray					270	129 (48)		
TB microbiological tests								
MGIT culture positive	199	50 (25)	69	41 (59)	295	279 (95)		
Xpert/Ultra positive	198	42 (21)	70	39 (56)	295	287 (97)		
Microscopy positive	205	48 (23)	67	35 (52)	203	169 (83)		
Blood (10^6 cells/ml)								
Leucocyte count	204	9.4 (7.0, 11.9)	74	6.4 (5.0, 9.2)	242	9.2 (7.4, 11.4)	26	6.4 (5.6, 7.2)
Neutrophil count	204	7.1 (4.8, 9.1)	74	5.0 (3.3, 6.9)	241	6.1 (4.7, 8.2)	26	3.4 (3.1, 4.1)
Lymphocytes count	204	1.2 (0.9, 2.0)	74	0.7 (0.4, 1.2)	242	1.9 (1.4, 2.3)	26	2.2 (1.9, 2.6)
CSF (10^3 cells/ml)								

Leucocyte count	207	142 (19, 323)	73	124 (10, 453)				
Neutrophil count	207	0 (0, 39)	73	17 (0, 144)				
Lymphocyte count	207	106 (18, 223)	73	58 (10, 216)				

Values were displayed as median (1st and 3rd interquartile) for continuous variables and frequency (%) for categorical variables.

TBM = tuberculous meningitis, PTB= pulmonary tuberculosis, CSF= cerebrospinal fluid

Table 2. Association between baseline clinical characteristics with TBM mortality

Characteristics	All TBM						Discovery cohort		Validation cohort	
	n	Survival n=220	Death n=61	HR*	95% CI	p value*	Survival n=111	Death n=31	Survival n=109	Death n=30
Male sex	281	144 (65)	39 (64)	1.01	0.99, 1.02	0.3	70 (63)	19 (61)	74 (68)	20 (67)
Age (years)	281	41 (32, 53)	39 (30, 60)	0.92	0.55, 1.56	0.8	41 (32, 53)	47 (33, 64)	41 (32, 53)	34 (29, 58)
HIV infection	281	47 (21%)	27 (44%)	2.34	1.41, 3.89	**< 0.001**	24 (22%)	14 (45%)	23 (21%)	13 (43%)
****Diagnostic category**	280									
definite TBM		101 (46%)	32 (52%)				55 (50%)	17 (55%)	46 (42%)	15 (50%)
possible TBM		43 (20%)	6 (9.8%)	1.01	0.59, 1.72	0.9	20 (18%)	2 (6.5%)	23 (21%)	4 (13%)
probable TBM		75 (34%)	23 (38%)	0.50	0.21, 1.19	0.12	35 (32%)	12 (39%)	40 (37%)	11 (37%)
***MRC grade**	281									
grade 1		114 (52%)	7 (11%)				58 (52%)	3 (9.7%)	56 (51%)	4 (13%)
grade 2		100 (45%)	35 (57%)	4.96	2.20, 11.2	**< 0.001**	51 (46%)	18 (58%)	49 (45%)	17 (57%)
grade 3		6 (2.7%)	19 (31%)	26.4	11.0, 63.2	**< 0.001**	2 (1.8%)	10 (32%)	4 (3.7%)	9 (30%)
Blood (10^6 cells/ml)										
Leukocyte count	278	8.1 (6.2, 10.7)	10.1 (6.7, 12.2)	4.49	1.13, 17.9	**0.033**	7.8 (6.2, 10.6)	8.6 (6.4, 12.0)	8.4 (6.4, 10.7)	10.8 (7.6, 12.4)
Neutrophil count	278	6.0 (4.0, 8.1)	8.0 (5.0, 10.3)	7.44	2.34, 23.6	**< 0.001**	5.9 (4.0, 8.1)	6.7 (4.9, 9.6)	6.4 (4.0, 8.2)	9.0 (5.8, 10.3)
Lymphocyte count	278	1.2 (0.8, 1.9)	0.9 (0.5, 1.2)	0.31	0.15, 0.66	**0.002**	1.2 (0.7, 2.0)	1.0 (0.6, 1.3)	1.3 (0.9, 1.9)	0.8 (0.5, 1.2)
CSF (10^3 cells/ml)										

Leukocyte count	280	129 (19, 340)	148 (15, 360)	0.98	0.72, 1.33	0.9	130 (17, 348)	148 (15, 444)	124 (27, 335)	146 (25, 283)
Neutrophil count	117	0 (0, 49)	0 (0, 97)	1.97	1.03, 3.76	**0.040**	3 (0, 53)	0 (0, 106)	0 (0, 36)	25 (0, 94)
Lymphocyte count	280	101 (18, 229)	91 (14, 176)	0.83	0.60, 1.16	0.3	104 (16, 208)	123 (10, 214)	97 (21, 248)	86 (14, 147)

Values were displayed as median (1st and 3rd interquartile) for continuous variables and frequency (%) for categorical variables.

*Association of the corresponding variables with three-month mortality using a Cox regression model. Hazard ratio (HR) and p-value from the Cox regression model were presented in the table.

** Diagnostic categories were assigned according to the consensus case definition [18].

*** MRC grade denotes modified British Medical Research Council criteria [14].

TBM = tuberculous meningitis, CSF= cerebrospinal fluid

Table 3. Comparison of gene signatures in distinguishing survival and death in TBM prognostic models

No.	Predictor set	All TBM n=281		HIV-negative cohort n=207		HIV-positive cohort n=74	
		AUC	Brier score	AUC	Brier score	AUC	Brier score
1	**Reference model** (PMID: 29029055)	0.78	0.14	0.77	0.12	0.82	0.18
2	**Gene set 1** (MCEMP1, TRABD2A, and CD4)	0.78	0.14	0.78	0.11	0.65	0.23
3	**Gene set 2** (MCEMP1 ,NELL2 and ZNF354C)	0.78	0.14	0.77	0.11	0.75	0.20
4	**Gene set 3** (MCEMP1, TRABD2A, CD4 and ZNF354C)	0.78	0.14	0.77	0.11	0.73	0.21
5	**Gene set 4** (MCEMP1 ,NELL2, ZNF354C and CD4)	0.79	0.14	0.77	0.11	0.75	0.20
6	**Gene set 3** and clinical risk factors	0.82	0.13	0.80	0.11	0.84	0.15
7	**Gene set 4** and clinical risk factors	0.82	0.13	0.80	0.11	0.86	0.14

Clinical risk factors were age and MRC grade [14]. The prediction models for three-month mortality were based on multivariable logistic regression models with top-hit genes and clinical risk factors. Area under the curve (AUC) and Brier score were corrected for optimism using internal bootstrap resampling over 1,000 iterations to evaluate the model performance. The Brier score is an overall performance measure,

calculated as the mean squared difference between the predicted probability and the actual outcome, with smaller values indicating superior model performance.