Whole blood transcriptional profiles and the pathogenesis of tuberculous meningitis

Hoang Thanh Hai[1]*, Le Thanh Hoang Nhat[1]*, Trinh Thi Bich Tram[1], Do Dinh Vinh[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: 4429 / 5000** (excluding the Materials and methods, References, and Figure legends)

No limits on the number of display items, **9 figures + 4 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 RNA sequencing (RNA-seq) transcriptional profiles in
85 281 Vietnamese adults with TBM, 295 with pulmonary TB (PTB), and 30 healthy controls. Our objective
86 was to use weighted gene co-expression network analysis, an unbiased and well-evaluated approach, to
87 identify the biological pathways and hub genes associated with TBM pathogenesis and assess the
88 predictive value of gene expression for early mortality from TBM.
89
90 **Results**
91 **Characteristics and outcomes of the cohorts**
92 Four RNA-seq cohorts (all ≥18 years) were used in the study, representing a total of 606 participants.
93 The 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 20.0 (IQR 18.2, 22.3). HIV-positive
96 TBM were more likely than HIV-negative TBM 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 TBM were lower than in HIV-negative TBM. Median CD4 cell counts in HIV-
99 positive TBM were 67 cells/mm^3 (IQR 19, 124) and 28 (39%) were receiving antiretroviral therapy. The
100 PTB cohort consisted of 295 HIV-negative adults with the median age of 44 years (IQR 31, 52), 228 (77%)

101 were male, the median of BMI was 19.4 (IQR 17.7, 21.6) and 129 (48%) had pulmonary cavities on chest
102 X-ray. Of the 30 healthy controls, 11 (37%) were male, and the median age was 33 (IQR 29, 37). In real-
103 time quantitative polymerase chain reaction (qPCR) validation cohort, 132 HIV-negative TBM adults
104 have similar characteristics as HIV-negative TBM RNA-seq cohort (**Table 1**).
105
106 The clinical variables associated with three-month mortality of the 281 adults with TBM in RNA-seq
107 cohort are given in **Table 2**. The discovery (n=142) and validation (n=139) cohorts had similar
108 characteristics (**Table 2**). 47.5% (133/280) had definite TBM [18], with microbiologically confirmed disease,
109 accounting for 45.9% (101/220) of survivors and 52.4% (32/61) of those who died. The overall three-
110 month mortality rate was 21.7% (61/281) for TBM regardless of HIV status: 16.4% (34/207) in HIV-
111 negative and 36.5% (27/74) in HIV-positive (p<0.001). We did not observe differences in mortality by
112 sex, age and diagnostic category. Greater disease severity, MRC grades 2 and 3 at enrolment, was
113 associated with increased mortality compared to grade 1 (p<0.001). In those who died, CSF and
114 peripheral blood neutrophil counts were higher and peripheral blood lymphocyte count lower,
115 compared to those who survived.
116 In the qPCR validation cohort (n=132), three-month mortality rate was 28.8% (38/132) and those who
117 died were of older age, had more severe disease, and had lower number of CSF leukocytes and
118 lymphocytes (Table S1). Death was not associated with CSF or peripheral blood neutrophil counts.
119
120 **Whole blood transcriptional profiles of the four RNA-seq cohorts**
121 We analyzed the whole blood transcriptomics, using bulk RNA sequencing from 606 participants in the 4
122 cohorts. On average 35.1 million reads/sample were obtained with 89.4% reads mapping accuracy to
123 human reference genome (GRCh.38 release 99) and 65.4% of reads were uniquely mapped. The study
124 objectives and cohorts flow are presented in **Figure 1.** Principal component analysis on the

125 transcriptomic data of 20,000 genes across 4 studies showed different profiles between healthy controls
126 from both PTB and TBM cases (**Figure 2**). The PTB profile substantively overlapped with TBM, with some
127 separation between HIV-negative and HIV-positive TBM (**Figure 2**).
128
129 Enrichment scores from single sample gene set enrichment analysis (ssGSEA), which based on
130 expression rank of genes relevant to pathways, were measured in some pathways already known to be
131 important mediators of TB or TBM pathogenesis (**Figure 2**). As anticipated, inflammatory response,
132 cytokine signaling, interferon signaling, TNF signaling, and inflammasome activation pathways, were
133 enriched in PTB and TBM cohorts as compared to healthy controls. In TBM, enrichment of genes in these
134 pathways were generally higher in HIV-positive than in HIV-negative individuals (**Tables S2, S6**).
135
136 **Transcriptional gene modules associated with TBM severity and mortality**
137 Transcriptional profiles associated with TBM mortality were generated by identifying differentially
138 expressed genes. Of the top 20,000 genes with most variation, we observed 724 (3.6%) genes that were
139 differentially expressed (FDR <0.05, absolute fold change (FC) >1.5) in all those with TBM (**Figure 3,**
140 **Table S3**). Next, we applied weighted gene co-expression network analysis (WGCNA) to 5000 most
141 variable genes from 281 TBM samples (n= 207, HIV-negative; n=74, HIV-positive) to define clusters of
142 highly correlated genes (modules) associated with TBM severity and mortality. Gene modules are
143 clusters of genes that have highly interconnected expression, usually related to their biological
144 functions. Hub genes are genes with high connectivity to other genes within a respective module. First,
145 we used WGCNA to construct a network of gene modules in the discovery cohort. Then we validated the
146 presence of these transcriptional modules in the validation cohort, labelling the modules with different
147 colors. In the discovery cohort (n=142), 15 modules were identified overall, consisting of 44 to 1350

genes per module (**Figure S1, Table S7**). 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 component (PC1) of each module. Similarly, associations of the modules with mortality were measured by hazard ratio (HR) per increase 1 unit of PC1 using Cox regression model adjusted for age, HIV status and dexamethasone treatment (**Figure 4**) in both discovery and validation cohorts.

Of the 15 preserved modules, five modules were significantly associated with mortality in the discovery and validation cohorts, with false discovery rate (FDR) < 0.05 (**Figure 4**). These five modules were separated into two big module clusters. The first cluster contained the blue module (n=799 genes), involved in inflammatory and innate immune responses, and the cyan module (n=44 genes) with unknown biological function. These modules were upregulated in those who died, as shown in the heat-map in **Figure 4** (HR: 3.0 and 2.2 for the blue module, and 2.1 and 1.7 for the cyan module, FDR < 0.05 for all comparisons). The black (n= 207 genes), brown (n=698 genes) and red (n=229 genes) modules were in the second cluster and were generally involved in adaptive immunity including T and B cell signaling pathways. These three modules were down-regulated in those who died (HR: 0.43 and 0.36 for brown, 0.46 and 0.39 for red, 0.59 and 0.49 for black) (**Figure 4**).

It is known that TBM MRC grade before treatment initiation strongly predicts outcome from TBM. Here, we investigated correlations between each module and MRC grade and their association with mortality.

172 The pink module, involved in hemostasis and platelet activation, was positively correlated with MRC
173 grade, but not mortality. Of the five modules associated with death from TBM, all were correlated with
174 MRC grade (**Figure 4**). Four of these five modules, were enriched for immune responses.
175
176 **Transcriptional hub genes associated with TBM severity and mortality**
177 We next identified hub genes within the four biologically functional modules associated with TBM
178 mortality in both the discovery and validation cohorts. Hub genes showed higher connectivity within the
179 modules, and stronger association with TBM mortality as compared to less connected genes within a
180 module (**Figure S3**). Seven hub genes (*ETS2, PGD, UPP1, CYSTM1, FCAR, KIF1B* and *MCEMP1*) were
181 upregulated in death, all from the acute inflammation (blue) module. Hub genes from the brown, red
182 and black modules were downregulated in death and were involved in adaptive immune response. Ten
183 hub genes associated with mortality (*CD96, TNFRSF25, TBC1D4, CD28, ABLIM1, RASGRP1, NELL2, TRAF5,*
184 *TESPA1, TRABD2A*) were from the brown module, three (*EVL, PLCG1, NLRC3*) from the red, and six (*CD2,*
185 *CD247, TGFBR3, ARL4C, KLRK1, MATK*) from the black module (**Figure 5, Table S8**). For qPCR validation,
186 available samples from HIV-negative TBM patients (n=132) were used to evaluate 11 hub genes selected
187 from the blue and brown modules. 8/11 hub genes were found to be significantly associated with
188 mortality, as determined by univariate Cox regression (**Table 3**). Three genes (FCAR, PGD, ETS2) were
189 also found to be associated with mortality, but in the reverse direction.
190
191 We next examined patterns of shared and distinct gene expression of some hub across the four cohorts
192 (healthy controls, PTB, and TBM with and without HIV-infection) to reveal disease progression and
193 pathogenesis of different TB forms (**Figure 6)**. There were two upregulated genes from the acute
194 inflammation module (*FCAR* and *MCEMP1*) and six downregulated genes (*NELL2, TRABD2A, PLCG1,*
195 *NLRC3, CD247* and *MATK*) from the other three adaptive immunity modules. The patterns of up and

196 downregulation, relative to healthy controls, were similar in PTB and TBM, although tended to be more
197 pronounced in those with TBM as well as those with HIV (**Figure 6**).
198
199 **Transcriptional immune pathways associated with TBM mortality**
200 To better understand the biological functions of the five modules associated with TBM mortality, a gene
201 set associated with mortality in each module was functionally annotated using known pathway
202 databases, such as gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) [19].
203 Pathway enrichment analysis was performed using fold enrichment to determine if the prevalence of
204 genes in a pathway were different from that expected by chance. We also used ssGSEA enrichment
205 scores based on gene expression ranking of genes in a particular pathway within a single individual, to
206 show the activity of this pathway between death and survival TBM as well as across 4 cohorts. These
207 analyses helped to identify the pathways linked to mortality within the gene modules.
208
209 In the blue module, we found TBM mortality was associated with upregulated inflammatory and innate
210 immune response transcripts, particularly in pathways involved in the acute inflammatory response and
211 the regulation of inflammatory responses, including responses to bacteria and neutrophil activation.
212 KEGG pathway analysis suggested this signal was associated with transduction and immune system
213 pathways, such as toll-like receptor signaling, TNF signaling, NF-kappa B signaling and neutrophil
214 extracellular trap formation (**Figure 5, Table S9**).
215
216 We did not find any pathway significantly associated with mortality in the upregulated cyan gene
217 module, although in the hierarchical clustering the cyan module was highly correlated with the blue -
218 inflammatory response module (**Figure 4**). In contrast, analyses of the down-regulated brown, red and
219 black modules highlighted pathways involved in adaptive immunity, predominantly those mediated by

220 lymphocytes. These included downregulations of lymphocyte proliferation, T cell activation, B cell
221 activation, natural killer cell mediated cytotoxicity and their signaling pathways such as antigen
222 receptor-mediated, T and B cell receptor (**Figure 5**, **Table S9**).
223
224 We also investigated the association between pathways known to be important to TB pathogenesis and
225 mortality (**Figure 7**). Looking at those who died vs. survived amongst HIV-positive TBM, there was little
226 difference between the groups with respect to inflammatory response, cytokine signaling, and icosanoid
227 and inflammasome activation. This was also the case for HIV-negative individuals. However, interferon
228 and TNF pathways were more active in those with HIV, and TNF signaling expression was significantly
229 higher in HIV-negative adults who died rather than survived.
230
231 **HIV influence on modules, hub genes, and pathways associated with TBM mortality**
232 Transcriptional profiles associated with TBM mortality stratified by HIV status are shown in **Figure 3**. In
233 HIV-negative individuals, 786 (3.9%) genes were differentially expressed, whereas in HIV-positive the
234 number was 1620 (8.1%) genes (**Figure 3, Table S3 & S4)**. We hypothesized that host transcriptional
235 signatures associated with TBM mortality differ according to HIV status. To test this hypothesis, we
236 constructed gene co-expression networks from all genes in HIV-negative (n=207) and HIV-positive (n=74)
237 individuals separately, then performed consensus gene co-expression network analysis (**Figure S4**).
238 Modules were identified with colors and to discriminate them from the modules linked to mortality
239 alone we added an annotated function on the module name. We focused on modules that failed to form
240 consensus associations with mortality due to opposite associations in the two cohorts.
241
242 Sixteen gene co-expression modules (**Figure 8**, **Table S10**), ranging from 60 to 958 genes, were obtained
243 from the HIV-negative and positive cohorts. Of these, 12 modules formed consensus association with

244 mortality (**Figure 8**). Of the 4 modules which failed to form consensus association, two modules were
245 significantly associated with mortality (FDR<0.05). The blue-angiogenesis module was up-regulated in
246 death in HIV-positive adults (HR: 1.7) and the yellow-extracellular matrix organization (EMO) module
247 was down-regulated in death in HIV-negative adults (HR: 0.31) (**Figure 8**).
248
249 Hub genes associated with mortality, which are highly correlated with other genes in the module, were
250 identified in the two modules. Five hub genes, with three downregulated (*ZNF354C, ZNF721* and *ETAA1*)
251 and two upregulated (*CEP295NL* and *RNF24*), were identified in the blue-angiogenesis module. Other
252 five hub genes, with four downregulated (*IL11RA*, *CD4*, *ZNF395* and *BCL9L*) and one upregulated
253 (*PPP4R2*), were identified in the yellow-EMO module (**Figure 8, Table S11**). Expression of some hub
254 genes across the four cohorts showed the patterns of up (*RNF24*) and downregulation (*ZNF721, BCL9L*
255 and *IL11RA*), and relative to healthy controls they were similar in PTB and TBM. Expression of *RNF24* and
256 *ZNF721* genes were significantly associated with death in HIV-positive adults, whereas expression of
257 *BCL9L* and *IL11RA* genes were significantly associated with death in HIV-negative adults (**Figure S5**).
258
259 Pathway enrichment analysis showed genes in the blue-angiogenesis module were significantly enriched
260 for angiogenesis or blood vessel development, leukocyte and neutrophil activation, and signal
261 transduction pathways (**Figure 8**, **Table S12**). Gene expression in the yellow- module was strongly
262 enriched in extracellular matrix organization, cell-substrate adhesion and protein localization pathways
263 (**Figure 8**, **Table S12**). Hierarchical clustering of modules indicated that these two modules were highly
264 correlated (**Figure 8**) suggesting they share similar functions, which appeared in some pathways such as
265 angiogenesis and extracellular matrix organization.
266

267 In addition, almost all of the pathways significantly associated with TBM mortality were similar in HIV-
268 negative and positive cohorts, confirming that these important pathways were common to all TBM
269 (**Figure S6**). However, pathway enrichment scores were higher in HIV-positive than in HIV-negative,
270 especially for the TNF signaling pathway and TNF transcripts (**Figure 9**). Death in HIV-negative TBM was
271 associated with increased enrichment of these pathways, but with less TNF transcript compared to HIV-
272 positive individuals. The enrichment scores indicated mortality was strongly associated with
273 downregulation of adaptive immune responses, including T cell activation, T and B cell receptor signaling
274 pathways, with little impact of HIV status on these pathways (**Figure 9**). These data suggest an
275 inadequate adaptive immune response contributes to disease pathogenesis and mortality in all those
276 with TBM, regardless of HIV status.
277
278 **Predictors of TBM mortality**
279 We aimed to identify baseline gene signatures that might help predict death or survival from TBM. We
280 selected potential gene predictors from 26 common and 10 HIV-specific hub genes; these hub genes
281 represented the dominant modules associated with mortality. Three predefined clinical factors known
282 to be associated with outcome (age, MRC grade and CSF lymphocytes) were also used for candidate
283 predictor selection. Using a multivariate Cox elastic-net regression model for 39 predictors, six
284 predictors were selected for HIV-negative TBM (MRC grade, age, CSF lymphocyte count, gene set 1:
285 *MCEMP1*, *TRABD2A* and *CD4*), and six predictors were selected for HIV-positive TBM (MRC grade, age,
286 CSF lymphocyte count, gene set 2: *NELL2*, *MCEMP1* and *ZNF354C*) (**Table S13**).
287 Gene expression and association with mortality of five distinct genes in gene sets 1 and 2 were
288 presented in **Figure S7**. By combining the two gene sets above and reducing genes in the same module,
289 we generated gene set 3 (*MCEMP1*, *TRABD2A*, *CD4* and *ZNF354C*) and gene set 4 (*MCEMP1*, *NELL2*,
290 *ZNF354C* and *CD4*). These two gene sets, with and without clinical factors (age, MRC grade and CSF

291 lymphocytes), were tested for their predictive performance in HIV-negative and HIV-positive RNA-seq
292 cohorts together with a reference model [15]. Model 7 using gene set 4 and clinical factors outperformed
293 other gene sets and reference model, with the best overall model performance (lowest optimism-
294 corrected Brier score, 0.11 and 0.14 for HIV-negative and HIV-positive, respectively) and the best
295 discriminant performance (highest optimism-corrected AUC 0.80 and 0.86 for HIV-negative and positive)
296 (**Table 4, Figure S8**).
297 Given that gene expression could reflect cellular composition changes of the peripheral blood, with
298 neutrophils being the most abundant sub-population, we performed a sensitivity analysis including
299 blood neutrophil count as a potential gene predictor in a multivariate Cox elastic-net regression model.
300 In the sensitivity analysis, all six predictors for either HIV-negative TBM or HIV-positive TBM were
301 repeatedly selected (**Table S14**). We evaluated the predictive values of blood neutrophil alone and in
302 combination with identified gene sets, but adding neutrophils did not improve the overall performance
303 of the predictive models.
304 Next, we evaluated the prognostic signature of model 7 in another sample set of 132 HIV-negative TBM
305 patients with gene expressions measured by qPCR. Only three genes (*MCEMP1, NELL2* and *ZNF354C*)
306 from gene set 4 were used because qPCR data of *CD4* was not available. The analysis demonstrated
307 good predictive performance in model 8, with an optimism-corrected AUC 0.92 and optimism-corrected
308 Brier score 0.12 (**Table 4**, **Figure S8**). Despite using three-gene signature in the validation, the results still
309 confirmed the ability to predict TBM mortality in HIV-negative patients.
310

311 **Discussion**

312 The biological pathways involved in pathogenesis of TBM are unclear. In general, previous studies
313 investigating TBM pathogenesis have been small, testing for relatively small numbers of selected genes
314 or molecules, and have been unable to take an unbiased and broader view of the inflammatory

315 response. This study investigated the pathways associated with death from TBM at a whole-genome
316 transcriptome level in whole blood, characterizing a global dysregulation in immune responses,
317 including inflammation, and revealing specific functional pathways and hub genes involved in TBM and
318 the mechanisms leading to death.
319
320 We sought to understand better the pathogenesis of TBM by identifying pretreatment blood
321 transcriptional gene modules associated TBM disease severity and three-month mortality. Four out of
322 five identified modules were involved in immunological functions, indicating multiple functional
323 pathways of systemic immunity are involved in the pathogenesis of TBM. In particular, mortality was
324 strongly associated with increased acute inflammation and neutrophil activation, and decreased
325 adaptive immunity and T and B cell activation. Whilst there appeared to be many common pathways
326 involved in TBM mortality in HIV-positive and negative individuals, there were differences: death was
327 associated with increased expression of angiogenesis genes in HIV-positive adults, and with TNF
328 signaling and down regulated extracellular matrix organization in HIV-negative adults. We also identified
329 a four-gene signature as a potential host response biomarker for mortality, regardless of HIV status.
330
331 TBM mortality was associated with increased acute inflammatory responses, the regulation of
332 inflammatory responses, and neutrophil activation. Previous blood transcriptome studies have shown
333 that IFN-inducible neutrophil-driven transcripts were over-expressed in blood neutrophils from active TB
334 patients compared to those with latent TB [20,21]. Another study described that TBM adults who developed
335 IRIS during treatment, also had significantly more abundant neutrophil-associated transcripts that
336 preceded the development of TBM-IRIS [5]. Our own earlier studies have also suggested a role for
337 neutrophils in TBM pathogenesis, with high pretreatment CSF bacterial loads being correlated with high

338 neutrophil numbers in both CSF and blood and more frequent new neurological events or paradoxical
339 inflammatory complications [8].
340
341 Taken together with previous studies, our findings support an important role for over-activation of
342 neutrophil-mediated inflammatory responses in TBM pathogenesis and its lethal complications. Looking
343 further into specific pathways, TBM mortality was neither associated with inflammatory response and
344 cytokine signaling pathways in general, nor with interferon or inflammasome activation. However,
345 increased transcripts in some specific immunity pathways, including TNF signaling, Toll-like receptor, NF-
346 kappa B and neutrophil extracellular trap formation, were associated with TBM mortality. These
347 pathways are involved as activators or effectors in the process of neutrophil activation and regulation,
348 leading to an exacerbated inflammatory response [22-24].
349
350 We found blood transcriptional responses of T cells and B cells were under-expressed in those who died
351 from TBM, indicating an impairment in adaptive immunity in fatal disease. A reduction of activities in
352 both T cell and B cell receptor signaling pathways in death were identified, independent of over-
353 expression of neutrophil-mediated immune responses, indicating that multiple functional pathways
354 influence TBM mortality. TBM pathogenesis is known to be associated with T cell impairment. Previous
355 studies have shown lower numbers of T cells, reduced ability to respond to *Mycobacterium tuberculosis*
356 antigens or reduced expression of activation markers and cytokine production in TBM compared to PTB
357 and healthy individuals [10,25,26]. This impaired T cell function has correlated with disease severity and poor
358 clinical outcomes in participants with PTB and TBM [10,27,28].
359
360 B cells and antibodies also influence humoral immunity against *Mycobacterium tuberculosis*. Studies
361 have shown a decreased memory B cell proportion, and lower levels of IgG, IgM antibodies and Fcγ

362 receptors binding capacity in plasma in those with active lung TB compared to those with latent TB or
363 healthy volunteers [29,30]. But little is known about the role of B cells in TBM. The observed association
364 between TBM mortality and decreased transcriptional responses in B cell activation and B cell receptor
365 signaling pathways suggest an unanticipated role for B cells and humoral immunity in TBM pathogenesis
366 that needs further investigation.
367
368 Transcriptomic profiles from four cohorts, including healthy controls and PTB, provide a broad view of
369 host responses in different TB clinical forms. Our data showed common transcriptional pathways and
370 genes between PTB and TBM. A range of immune responses, involving in inflammation, cytokines,
371 interferon, inflammasome and neutrophil signaling pathways, were activated in both PTB and TBM, but
372 with significantly greater activation in HIV-positive TBM than HIV-negative TBM. This finding aligns with
373 our previous data showing a dyregulated hyper-inflammation in HIV-associated TBM, with significantly
374 higher CSF cytokine concentrations than in those without HIV infection [7]. These data suggest that
375 different forms of TB are associated with similar inflammatory responses, but with different degrees of
376 host responses, exemplified by hyper-inflammation in those with HIV.
377
378 The HIV-driven differences in TBM-associated inflammation appear sufficient to influence response to
379 adjunctive anti-inflammatory therapy with corticosteroids. An earlier randomized controlled trial of
380 corticosteroids in 545 predominantly HIV-negative Vietnamese adults showed they significantly
381 improved survival [31]. Our group recently completed a similar sized trial exclusively in HIV-positive adults
382 without any clear benefit upon mortality [32]. Therefore, the differences in gene expression associated
383 with TBM mortality in the current study – with angiogenesis activation linked to HIV-positive TBM
384 mortality only, for example – may provide an explanation for the poor response and offer alternative
385 therapeutic strategies. Recent studies have shown that angiogenesis is induced by *Mycobacterium*

386 *tuberculosis* infection, which then contributes to inflammation, tissue damage and is correlated with
387 disease severity [33].
388
389 Developing prognostic models for TBM is important for guiding clinical decision making and improving
390 outcomes. Several studies have developed and validated prognostic models for TBM using clinical,
391 laboratory and radiological variables. These models have demonstrated moderate to high accuracy in
392 predicting mortality and functional outcomes from TBM patients [15,34-36]. In this study, we used blood
393 transcriptional signatures and co-expression network analysis to identify module-representative hub
394 genes. We identified a four-gene set at the start of treatment (*MCEMP1*, *NELL2*, *ZNF354C* and *CD4*)
395 whose expression strongly predicted three-month TBM mortality. Our prognostic models combining this
396 four-gene host response in blood and three routine clinical predictors achieved very good performance,
397 with AUC 0.80 and 0.86 for HIV-negative and HIV-positive individuals with TBM, respectively. This is
398 proof-of-concept that whole blood RNA host response might be a useful pre-treatment biomarker to
399 predict early TBM mortality. Although further investigation and validation is needed, we have identified
400 potential gene candidates for future development as prognostic biomarkers.
401
402 In summary, we present a comprehensive and unbiased analysis of the gene transcripts in blood
403 associated with TBM severity and mortality. Our data open a new window on TBM pathogenesis, with
404 dysregulation in both innate and adaptive immune responses strongly associated with death from TBM.
405 Furthermore, we have identified similarities and differences in the inflammatory responses associated
406 with TBM in HIV-positive and negative adults, which may explain the different therapeutic effects of
407 adjunctive corticosteroid treatment. We also revealed a 4-gene host response signature in blood that
408 might represent a novel biomarker for defining those at highest risk of death, regardless of their HIV
409 status. The signature needs validation in future cohorts.

410

411 **Materials and Methods**

412 **Participants**

413 We collected data and whole peripheral blood samples for transcriptomic profiling from adults (≥18
414 years) with TBM enrolled into two randomized controlled trials conducted in Vietnam. The two trials
415 investigated whether adjunctive dexamethasone improves outcome from TBM and the protocols for
416 both trials have been published [37,38]. The ACT-HIV trial (NCT03092817) completed enrolment and follow-
417 up of 520 HIV-positive adults with TBM in April 2023, with the results accepted for publication [32]. The
418 LAST-ACT trial (NCT03100786) completed enrollment of 720 HIV-negative adults with TBM in March
419 2023, with the last participants due to complete follow-up in March 2024.

420

421 In TBM RNA-seq cohorts, peripheral blood samples were taken from 281 trial participants: the first 207
422 consecutively enrolled HIV-negative participants from the LAST-ACT trial, and 74 randomly selected HIV-
423 positive participants from the ACT-HIV trial. In qPCR TBM cohorts, 132 HIV-negative participants were
424 randomly selected from remaining participants of LAST-ACT trial with an enrichment for non-survival
425 cases. The samples were taken at enrollment, when patients could not have received more than 6
426 consecutive days of two or more drugs active against *Mycobacterium tuberculosis*. All trial participants
427 then received standard 4-drug anti-tuberculosis treatment for 2 months, followed by 3 drugs for 10
428 months, and were randomly allocated to dexamethasone or identical placebo for the first 6-8 weeks
429 [37,38]. The investigators remain blind to the treatment allocation until the last participant completes
430 follow-up and the database has been locked. Therefore, an analysis of the direct influence of
431 corticosteroids on inflammatory response and outcome is not included in the current study. To adjust
432 for corticosteroid effect, treatment outcome and effect were blinded in analyses by extracting only the

433 fold change difference between survival and death in the linear regression model, in which gene
434 expression was outcome and survival and treatment were covariates.
435
436 Peripheral blood samples were taken for transcriptional profiling from two other cohorts: 295 adults
437 with PTB and 30 healthy controls. The 295 PTB was enrolled in Pham Ngoc Thach hospital and District TB
438 Units in Ho Chi Minh city, randomly selected from HIV-negative participants from a prospective
439 observational study of the host and bacterial determinants of outcome from PTB (n=900) due to the
440 very low prevalence of HIV in the study population. Participants had culture-confirmed PTB, either drug
441 susceptible TB or a new diagnosis of multidrug-resistant TB. All participants in this cohort had <7 days of
442 anti-TB drugs at enrolment and did not have clinical evidence of extra-pulmonary TB. Healthy controls
443 were enrolled in Hospital for Tropical Diseases in Ho Chi Minh city, from a prospective study for
444 epidemiological characteristics of human resistance to *Mycobacterium tuberculosis* infection.
445 Participants in this cohort were adults without signs or symptoms of TB nor history of TB contact within
446 the last two years.
447
448 Ethics approval was obtained from the institutional review board at the Hospital for Tropical Diseases,
449 Pham Ngoc Thach Hospital and the ethics committee of the Ministry of Health in Vietnam, and the
450 Oxford Tropical Research Ethics Committee, UK (OxTREC 52-16, 36-16, 24-17, 33-17 and 532-22). All
451 participants provided their written informed consent to take part in the study, or from their relatives if
452 they were incapacitated.
453
454 **Study design**
455 The objectives and cohorts used in this study are presented in **Figure 1** and a workflow of data analysis is
456 shown in **Figure S10**. Briefly, blood transcriptional profiling was generated from four cohorts of 207 TBM

457 HIV-negative, 74 TBM HIV-positive, 295 PTB and 30 healthy controls. To define transcriptional signatures
458 associated with TBM mortality, data from all 281 TBM participants were used. To ensure reproducibility,
459 in our main analysis we randomly split our TBM data into two datasets, a discovery cohort (n=142) and a
460 validation cohort (n=139). After identifying gene modules associated with mortality, related functional
461 pathways and hub genes were determined with more details of the data analysis provided in **Figure S10**.
462 In a broader view, gene enrichment and expression from significant pathways and hub genes were
463 illustrated across all four cohorts. Next, outcome prediction models were developed for TBM. Finally,
464 the association of hub genes and outcome prediction then was validated in qPCR HIV-negative TBM
465 cohort.
466
467 **Sample processing and RNA-seq**
468 Whole blood samples, collected from participants at enrollment, were stored in PAXgene collection
469 tubes at -80^{0}C. RNA extraction and RNA-seq were done in 2 batches. Batch 1 was done in 2020 including
470 207 HIV-negative TBM, 31 HIV-positive TBM and 295 PTB. Batch 2 was done in 2022 including 43 HIV-
471 positive TBM and 30 Healthy control. RNA samples were isolated using the PAXgene Blood RNA kits
472 (Qiagen, Valencia, CA, USA) following the manufacturer's instructions, except for an additional washing
473 step before RNA elution. DNA was digested on columns using the RNase-free DNase Set (Qiagen,
474 Valencia, CA, USA). Quality control of the RNA extraction was performed using the Epoch spec for
475 quantity and quality, and Tapestation Eukaryotic RNA Screentape for integrity. Samples with RNA
476 integrity number below 4 were exclude for further steps. RNA-seq was performed by the Ramaciotti
477 Centre for Genomics (Sydney, Australia). One microgram of total RNA was used as input for each
478 sample, using the TruSeq Stranded Total RNA Ribo-zero Globin kit (Illumina). Libraries were generated
479 on the Sciclone G3 NGS (Perkin Elmer, Utah, USA) and the cDNA was amplified using 11 PCR cycles.

480 Libraries were pooled 75 samples per pool and sequenced using NovaSeq 6000 S4 reagents at 2x100bp
481 to generate about 30 million reads per sample.
482
483 **RNA-seq data quality control and pre-processing**
484 Quality control and alignment were performed using an in-house pipeline modified from previously
485 published practices for RNA-seq analysis [39,40] in linux command line. Briefly, the quality of the
486 sequencing fastq files was analyzed using FastQC (v0.11.5) and poor quality samples were excluded from
487 further analysis. Sequence reads were adapter and quality trimmed using Trimmomatic (v0.36), followed
488 by duplicated optical read removal using BBMap (v38.79) tool. STAR aligner (v2.5.2a) was used to align
489 the reads to the human reference genome (GRCh38 build 99) downloaded from Ensembl, allowing for
490 maximum 2 mismatches in each 25 bp segment and a maximum of 20 alignment hits per read [41,42]. The
491 alignment results were sorted and indexed for downstream analyses as BAM format files. The aligned
492 reads were further utilized to generate gene expression counts using FeatureCounts (v2.0.0) against the
493 human reference annotation (GRCh38 build 99) [43]. Next, 60,067 genes in the expression matrix were
494 first normalized by variance stabilizing transformation method built in DESeq2 package in R [44].
495 Subsequently, the batch effect was removed for 20,000 most variable genes using combat function in
496 the SVA package with correcting for 2 RNA-seq batches [45]. The results of batch effect removal were
497 visualized by principle component analysis. The first component was plotted against the second
498 component. The variation explained by RNA-seq batches was removed after using combat (Figure S9).
499 All later analysis and data visualization were done used batch corrected data.
500 **Validation of hub genes by Microfluidic multiplex RT-qPCR**
501 Whole blood samples were collected, stored and RNA extraction was performed as described in the
502 previous section. The expression of housekeeping genes (*GAPDH* and *TMBIM6*) and other hub genes
503 were evaluated by microfluidic RT-qPCR using Biomark™ 48.48 Complete Bundle with Delta Gene™

504 Assays and BioMark™ HD system (Fluidigm Corporation, South San Francisco, CA, USA) following
505 manufacturer's instructions with some optimized modifications. Briefly, 2 µL of total RNA with
506 concentration of 50 ng/µL was reverse transcribed to cDNA. The specific target amplification (STA) of
507 cDNA was used for 14 cycles of preamplification, then the STA products were treated with Exonuclease I
508 (New England Biolabs, Ipswich, MA, USA) before 20-fold dilution. SsoFast™ EvaGreen® Supermix with
509 Low ROX (Bio-Rad Laboratories, Hercules, CA, USA) was used in RT-qPCR before applying to IFC
510 Controllers MX and BioMark™ HD system (Fluidigm Corporation, South San Francisco, CA, USA). The
511 cycle threshold (CT) values of target hub genes were normalized by subtracting for CT of *GAPDH* before
512 analysis.
513
514 **Statistical analysis**
515 The primary outcome examined in this study was TBM three-month mortality. In the descriptive
516 analysis, we summarized and tested association of patient characteristics with three-month mortality
517 using univariate Cox regression analysis. We presented the proportion for binary variables and median
518 (1st and 3rd interquartile range) for continuous variables.
519
520 To explore transcriptional profiles associated with mortality in TBM, differential expression analysis was
521 performed on the 20,000 normalized gene expression matrix to find differentially expressed genes
522 (DEGs). We constructed the contradicted matrix with survival and death status adjusted for covariates
523 including: age, HIV status, dexamethasone treatment (LAST-ACT trial investigators remained blind), and
524 sequencing batch. Linear models were used to assess DEGs using limma R package. Empirical Bayes
525 moderated-t p-values were computed for each genes and Benjamini-Hochberg were used for correcting
526 multiple testing (FDR). We defined the DEGs with the parameters (fold change > 1.5 and FDR < 0.05). To
527 visualize samples clustering in the 4 cohorts, an unsupervised principal component analysis was

528 performed on 20,000 normalized genes, and the first principle component was plotted against the
529 second principle component. The 95% confidence ellipse was drawn for each group using multivariate t-
530 distribution ellipse. To compare enrichment activity of interested pathways between death vs survival,
531 TBM vs PTB or healthy controls, enrichment scores of these pathways were calculated for single patient
532 using single sample Gene Set Enrichment Analysis algorithm (ssGSEA). Pairwise Wilcoxon rank sum test
533 was used for comparison between any two cohorts.
534
535 In the primary analysis, we initially conducted *weighted gene co-expression network analysis* (WGCNA)
536 on the whole-blood transcriptomic profiling of the discovery cohort to identify clusters of genes (or
537 'modules') in the gene co-expression network among host transcriptomic genes. We then performed a
538 *network modules preservation analysis* to assess the reproducibility of these modules and the network
539 in the validation cohort. Subsequently, we conducted a *module-trait association analysis* to identify
540 modules associated with the clinical traits of mortality and TBM severity. For the association analysis
541 with mortality, we used a *Cox regression model* with the PC1 of the module as an independent variable.
542 We adjusted the model for age, HIV status, and corticosteroid treatment (dexamethasone vs. placebo).
543
544 To analyze the association with TBM severity, we calculated the *Spearman correlation* between TBM
545 severity (MRC grade) and the module's PC1. We corrected for false discovery rate (FDR) for the
546 dependent hypotheses by using the Benjamini-Yekutieli procedure. Target modules were defined as
547 modules associated with mortality in both the discovery and validation cohorts at FDR < 0.05. Hub genes
548 within each targeted module were identified based on 3 criteria: being protein coding gene, module
549 membership cut off 0.85 (MM) and high rank of gene significance rank (GS). MM defined as correlation
550 between individual gene and the module's PC1. GS defined as $-\log_{10}$ p value association of gene with
551 mortality using Cox regression model adjusted for age, HIV status and corticosteroid treatment. Genes

552 were first filtered for protein coding function and MM above 0.85. Genes then was ranked based on its
553 GS and top 20 hub genes were selected if number of genes in module above 500, otherwise top 10
554 genes were selected. Overlap hub genes between discovery and validation cohort were consider
555 validated.
556
557 In the subsequent analysis, to determine the biological functions or processes potentially related to each
558 module we conducted *overrepresentation analysis* (ORA) using ShinyGO v0.77 [19]. Genes were first filters
559 for association with three-month mortality at $p < 0.05$ and then were input into ShinyGO. We used both
560 *GO and KEGG databases* for this analysis. We pre-specified significant pathways with overlap gene in
561 pathway > 5 and FDR < 0.05.
562
563 Furthermore, to compare *enrichment activity* of resulting pathways between death vs survival, TBM vs
564 PTB or healthy controls, enrichment scores of these pathways were calculated for single patient using
565 ssGSEA algorithm [46]. Pairwise Wilcoxon rank rum test were applied for comparison between 2 cohorts.
566 To compare or visualize the difference of pathway effect on mortality between HIV-negative and HIV-
567 positive TBM, *fold change enrichment* of interested pathways were calculated using Quantitative Set
568 Analysis for Gene Expression (QuSAGE) method [47].
569
570 In the secondary analysis, we performed a *consensus analysis* to identify consensus patterns of co-
571 expression networks between HIV-positive and HIV-negative conditions in our TBM cohort. We
572 conducted similar association analyses for TBM severity and mortality as in the preceding WGCNA
573 analysis for both the HIV-negative and HIV-positive cohorts. We visualized the results using a heatmap.
574 We defined strongly associated modules in HIV-positive or HIV-negative cohort as HIV-positive-specific
575 signals and HIV-negative-specific signals. Additionally, we functional enrichment ORA was performed for

576 these consensus modules and the HIV-specific modules. Top 5 hub genes from HIV-specific modules
577 were selected based on criteria described in the primary analysis, which included protein-coding genes
578 and having high MM and high rank GS. GS in this analysis was calculated as $-\log_{10}$ p value association of
579 genes with mortality using Cox regression model adjusted for age and corticosteroid treatment.
580
581 We then performed variable selection using a *multivariate elastic-net Cox regression model* to select the
582 important predictors for HIV-negative and HIV-positive TBM prognosis separately. Candidate predictors
583 in HIV-negative TBM consisted of the combination of common hub genes, HIV-negative specific hub
584 genes and clinical biomarkers: age, TBM severity and CSF lymphocyte count. Similarly, for HIV-positive
585 TBM, candidate predictors consisted of common hub genes, HIV-positive specific hub genes and clinical
586 biomarkers. We performed the analysis with 1,000 bootstrap sampling with replacement. The chosen
587 variables were 1) among the top 75% of selected variables and 2) one representative hub gene per
588 module. Subsequently, we developed a prediction model based on a *logistic regression model*
589 incorporating only the chosen variables for all TBM cohort or stratified by HIV-negative and HIV-positive
590 cohort. We also assessed predictive performance of chosen variables in each cohort by 1000 times
591 bootstrapping sampling approach and reported the overall model performance (*optimism-corrected*
592 *Brier score*), the discrimination (*optimism-corrected AUC*), and calibration (*optimism-corrected*
593 *calibration-slope*) of the developed model [48].
594
595 More details on performing ORA, SSGSEA, QuSAGE and WGCNA preservation and consensus analyses
596 were provided in in the *Supplementary method document.*
597

598 **Acknowledgements**

599 We would like to thank and the clinical staff who recruited patients into our study from Hospital for

600 Tropical Diseases and Pham Ngoc Thach hospital, Ho Chi Minh city, Vietnam, and all participants.

601

602 **Funding**

603 This work was supported by Wellcome Trust Fellowship in Public Health and Tropical Medicine to NTTT

604 (206724/Z/17/Z), Wellcome Trust Investigator Award to GT and Wellcome Trust to Vietnam Africa Asia

605 Programme.

606

607 **Conflicts of Interest**

608 No conflicts of interest

609

610 **References**

611 1. Dodd, P.J. *et al.* The global burden of tuberculous meningitis in adults: A modelling study. *PLOS*
612 *Glob Public Health* **1**, e0000069 (2021).
613 2. Stadelman, A.M. *et al.* Treatment Outcomes in Adult Tuberculous Meningitis: A Systematic
614 Review and Meta-analysis. *Open Forum Infect Dis* **7**, ofaa257 (2020).
615 3. Wilkinson, R.J. *et al.* Tuberculous meningitis. *Nat Rev Neurol* **13**, 581-598 (2017).
616 4. Huynh, J. *et al.* Tuberculous meningitis: progress and remaining questions. *Lancet Neurol* **21**,
617 450-464 (2022).
618 5. Marais, S. *et al.* Inflammasome Activation Underlying Central Nervous System Deterioration in
619 HIV-Associated Tuberculosis. *J Infect Dis* **215**, 677-686 (2017).
620 6. Rohlwink, U.K. *et al.* Biomarkers of Cerebral Injury and Inflammation in Pediatric Tuberculous
621 Meningitis. *Clin Infect Dis* **65**, 1298-1307 (2017).
622 7. Thuong, N.T.T. *et al.* Leukotriene A4 Hydrolase Genotype and HIV Infection Influence
623 Intracerebral Inflammation and Survival From Tuberculous Meningitis. *J Infect Dis* **215**, 1020-
624 1028 (2017).
625 8. Thuong, N.T.T. *et al.* Pretreatment Cerebrospinal Fluid Bacterial Load Correlates With
626 Inflammatory Response and Predicts Neurological Events During Tuberculous Meningitis
627 Treatment. *J Infect Dis* **219**, 986-995 (2019).
628 9. van Laarhoven, A. *et al.* Clinical Parameters, Routine Inflammatory Markers, and LTA4H
629 Genotype as Predictors of Mortality Among 608 Patients With Tuberculous Meningitis in
630 Indonesia. *J Infect Dis* **215**, 1029-1039 (2017).
631 10. van Laarhoven, A. *et al.* Immune cell characteristics and cytokine responses in adult HIV-
632 negative tuberculous meningitis: an observational cohort study. *Sci Rep* **9**, 884 (2019).

633 11. Cresswell, F.V. *et al.* Recent Developments in Tuberculous Meningitis Pathogenesis and
634 Diagnostics. *Wellcome Open Res* **4**, 164 (2019).
635 12. Rohlwink, U.K. *et al.* Tuberculous meningitis in children is characterized by compartmentalized
636 immune responses and neural excitotoxicity. *Nat Commun* **10**, 3767 (2019).
637 13. Ardiansyah, E. *et al.* Tryptophan metabolism determines outcome in tuberculous meningitis: a
638 targeted metabolomic analysis. *Elife* **12**(2023).
639 14. STREPTOMYCIN treatment of tuberculous meningitis. *Lancet* **1**, 582-96 (1948).
640 15. Thao, L.T.P. *et al.* Prognostic Models for 9-Month Mortality in Tuberculous Meningitis. *Clin Infect*
641 *Dis* **66**, 523-532 (2018).
642 16. Sweeney, T.E., Braviak, L., Tato, C.M. & Khatri, P. Genome-wide expression for diagnosis of
643 pulmonary tuberculosis: a multicohort analysis. *Lancet Respir Med* **4**, 213-24 (2016).
644 17. Penn-Nicholson, A. *et al.* RISK6, a 6-gene transcriptomic signature of TB disease risk, diagnosis
645 and treatment response. *Sci Rep* **10**, 8629 (2020).
646 18. Marais, S. *et al.* Tuberculous meningitis: a uniform case definition for use in clinical research.
647 *Lancet Infect Dis* **10**, 803-12 (2010).
648 19. Ge, S.X., Jung, D. & Yao, R. ShinyGO: a graphical gene-set enrichment tool for animals and
649 plants. *Bioinformatics* **36**, 2628-2629 (2020).
650 20. Berry, M.P. *et al.* An interferon-inducible neutrophil-driven blood transcriptional signature in
651 human tuberculosis. *Nature* **466**, 973-7 (2010).
652 21. Singhania, A. *et al.* A modular transcriptional signature identifies phenotypic heterogeneity of
653 human tuberculosis infection. *Nat Commun* **9**, 2308 (2018).
654 22. Geng, X. *et al.* Whole transcriptome sequencing reveals neutrophils' transcriptional landscape
655 associated with active tuberculosis. *Front Immunol* **13**, 954221 (2022).
656 23. Castro-Alcaraz, S., Miskolci, V., Kalasapudi, B., Davidson, D. & Vancurova, I. NF-kappa B
657 regulation in human neutrophils by nuclear I kappa B alpha: correlation to apoptosis. *J Immunol*
658 **169**, 3947-53 (2002).
659 24. Sabroe, I., Dower, S.K. & Whyte, M.K. The role of Toll-like receptors in the regulation of
660 neutrophil migration, activation, and apoptosis. *Clin Infect Dis* **41 Suppl 7**, S421-6 (2005).
661 25. Davoudi, S. *et al.* CD4+ cell counts in patients with different clinical manifestations of
662 tuberculosis. *Braz J Infect Dis* **12**, 483-6 (2008).
663 26. Shridhar, A. *et al.* Prevalence of primary immunodeficiency syndromes in tuberculous
664 meningitis: A case-control study. *J Infect Public Health* **15**, 29-35 (2022).
665 27. An, H.R. *et al.* The relationship between absolute counts of lymphocyte subsets and clinical
666 features in patients with pulmonary tuberculosis. *Clin Respir J* **16**, 369-379 (2022).
667 28. Hemingway, C. *et al.* Childhood tuberculosis is associated with decreased abundance of T cell
668 gene transcripts and impaired T cell function. *PLoS One* **12**, e0185973 (2017).
669 29. Lu, L.L. *et al.* A Functional Role for Antibodies in Tuberculosis. *Cell* **167**, 433-443 e14 (2016).
670 30. La Manna, M.P. *et al.* Impact of Mycobacterium tuberculosis Infection on Human B Cell
671 Compartment and Antibody Responses. *Cells* **11**(2022).
672 31. Thwaites, G.E. *et al.* Dexamethasone for the treatment of tuberculous meningitis in adolescents
673 and adults. *N Engl J Med* **351**, 1741-51 (2004).
674 32. Donovan, J. *et al.* Adjunctive Dexamethasone for Tuberculous Meningitis in HIV-Positive Adults.
675 *N Engl J Med* **389**, 1357-1367 (2023).
676 33. Kumar, N.P., Banurekha, V.V., Nair, D. & Babu, S. Circulating Angiogenic Factors as Biomarkers of
677 Disease Severity and Bacterial Burden in Pulmonary Tuberculosis. *PLoS One* **11**, e0146318
678 (2016).

679 34. Thao, L.T.P. *et al.* Dynamic Prediction of Death in Patients With Tuberculous Meningitis Using
680 Time-updated Glasgow Coma Scale and Plasma Sodium Measurements. *Clin Infect Dis* **70**, 827-
681 834 (2020).
682 35. Feng, B. *et al.* Prognostic factors of adult tuberculous meningitis in intensive care unit: a single-
683 center retrospective study in East China. *BMC Neurol* **21**, 308 (2021).
684 36. Sharawat, I.K. *et al.* Development and Validation of a Prognostic Model and Bedside Score for
685 the Neurological Outcome in Children with Tuberculous Meningitis. *Am J Trop Med Hyg* **107**,
686 1288-1294 (2022).
687 37. Donovan, J. *et al.* Adjunctive dexamethasone for the treatment of HIV-uninfected adults with
688 tuberculous meningitis stratified by Leukotriene A4 hydrolase genotype (LAST ACT): Study
689 protocol for a randomised double blind placebo controlled non-inferiority trial. *Wellcome Open*
690 *Res* **3**, 32 (2018).
691 38. Donovan, J. *et al.* Adjunctive dexamethasone for the treatment of HIV-infected adults with
692 tuberculous meningitis (ACT HIV): Study protocol for a randomised controlled trial. *Wellcome*
693 *Open Res* **3**, 31 (2018).
694 39. Yalamanchili, H.K., Wan, Y.W. & Liu, Z. Data Analysis Pipeline for RNA-seq Experiments: From
695 Differential Expression to Cryptic Splicing. *Curr Protoc Bioinformatics* **59**, 11 15 1-11 15 21
696 (2017).
697 40. Conesa, A. *et al.* A survey of best practices for RNA-seq data analysis. *Genome Biol* **17**, 13 (2016).
698 41. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2013).
699 42. Bolger, A.M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence
700 data. *Bioinformatics* **30**, 2114-20 (2014).
701 43. Liao, Y., Smyth, G.K. & Shi, W. featureCounts: an efficient general purpose program for assigning
702 sequence reads to genomic features. *Bioinformatics* **30**, 923-30 (2014).
703 44. Love, M.I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-
704 seq data with DESeq2. *Genome Biol* **15**, 550 (2014).
705 45. Leek, J.T., Johnson, W.E., Parker, H.S., Jaffe, A.E. & Storey, J.D. The sva package for removing
706 batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* **28**,
707 882-3 (2012).
708 46. Barbie, D.A. *et al.* Systematic RNA interference reveals that oncogenic KRAS-driven cancers
709 require TBK1. *Nature* **462**, 108-12 (2009).
710 47. Yaari, G., Bolen, C.R., Thakar, J. & Kleinstein, S.H. Quantitative set analysis for gene expression: a
711 method to quantify gene set differential expression including gene-gene correlations. *Nucleic*
712 *Acids Res* **41**, e170 (2013).
713 48. Steyerberg, E.W. *et al.* Internal validation of predictive models: efficiency of some procedures
714 for logistic regression analysis. *J Clin Epidemiol* **54**, 774-81 (2001).

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

Characteristics	RNA-seq cohorts								qPCR validation cohort	
	HIV-negative TBM n=207		HIV-positive TBM n=74		PTB n=295		Healthy controls n=30		HIV-negative TBM n=132	
	n	Summary	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)	132	48 (35, 60)
Male sex	207	127 (61)	74	56 (76)	295	228 (77)	30	11 (37)	132	84 (65)
BMI (kg/m2)	205	20.0 (18.2, 22.3)	72	19.3 (17.2, 20.4)	295	19.4 (17.7, 21.6)			132	20.0 (18.2, 22.3)
Symptom duration (days)	207	14 (11, 20)	73	16 (10, 30)	294	20 (10, 30)			132	16 (13, 24)
History of TB treatment	204	5 (2.5)	74	15 (20)	295	100 (34)			130	12 (9.2)
Glasgow coma score	207	14 (12, 15)	72	14 (13, 15)					132	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)			105	18 (17)
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)	129	10.0 (7.7, 12.4)
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)	129	7.8 (6.9, 8.5)
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)	129	1.2 (0.7, 1.8)

CSF (10^3 cells/ml)										
Leucocyte count	207	142 (19, 323)	73	124 (10, 453)					106	122 (38, 328)
Neutrophil count	207	0 (0, 39)	73	17 (0, 144)					60	20 (3, 73)
Lymphocyte count	207	106 (18, 223)	73	58 (10, 216)					106	95 (79, 100)
CD4 cell count (cells/ml)			73	67 (19, 124)						
Antiretroviral therapy			70	28 (39%)						
HIV load (10^3 cells/ml)			70	77.3 (8.28, 672)						

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 in RNA-seq cohorts

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]. Definite TBM refers to cases that are microbiological confirmed TB by microscopy, culture and Xpert using CSF samples.

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

TBM = tuberculous meningitis, CSF= cerebrospinal fluid

Table 3. Validation of hub genes in PCR validation cohort

No.	Genes	RNA-seq cohort HIV-negative TBM (n=207)			qPCR validation cohort HIV-negative TBM (n=132)		
		HR	95%CI	p value	HR	95%CI	p value
1	MCEMP1	1.93	2.37, 1.57	4.07E-10	1.74	2.37, 1.27	4.99E-04
2	FCAR	2.37	3.30, 1.70	3.52E-07	0.59	0.87, 0.40	7.87E-03
3	ETS2	2.65	3.81, 1.85	1.24E-07	0.27	0.44, 0.16	3.27E-07
4	PGD	2.65	3.81, 1.84	1.47E-07	0.18	0.34, 0.09	2.69E-07
5	NELL2	0.56	0.71, 0.45	6.08E-07	0.80	0.96, 0.66	1.53E-02
6	TRABD2A	0.43	0.58, 0.31	4.94E-08	0.64	0.8, 0.51	1.05E-04
7	TRAF5	0.31	0.49, 0.2	4.52E-07	0.47	0.61, 0.36	1.66E-08
8	CD28	0.57	0.76, 0.42	1.42E-04	0.60	0.76, 0.47	2.83E-05
9	TESPA1	0.36	0.52, 0.25	2.74E-08	0.65	0.89, 0.48	7.35E-03
10	ABLIM1	0.31	0.49, 0.2	2.21E-07	0.47	0.63, 0.36	1.55E-07
11	RASGRP1	0.45	0.67, 0.31	6.10E-05	0.59	0.77, 0.45	9.77E-05

Association of the hub genes with three-month mortality using a univariate Cox regression model. Hazard ratio (HR), 95% CI of HR and p-value from the Cox regression model were presented in the table. HR or 95% mean per increase 1 unit of log2 normalized expression of gene in RNA-seq cohort or decrease 1 unit of cycle threshold in qPCR validation cohort.

Table 4. Comparison of gene signatures in distinguishing survival and death in TBM prognostic models

A. RNA-seq cohorts							
No.	**Predictor set**	**All TBM n=281**		**HIV-negative TBM n=207**		**HIV-positive TBM 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, CD4 and ZNF354C)	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

B. qPCR validation cohort			
No.	**Predictor set**	**HIV-negative TBM (n=132)**	
		AUC	**Brier score**
8*	MCEMP1, NELL2, ZNF354C and clinical predictors	0.92	0.12

Clinical risk factors were age, MRC grade and CSF lymphocytes [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.

* CD4 data is unavailable for the analysis.

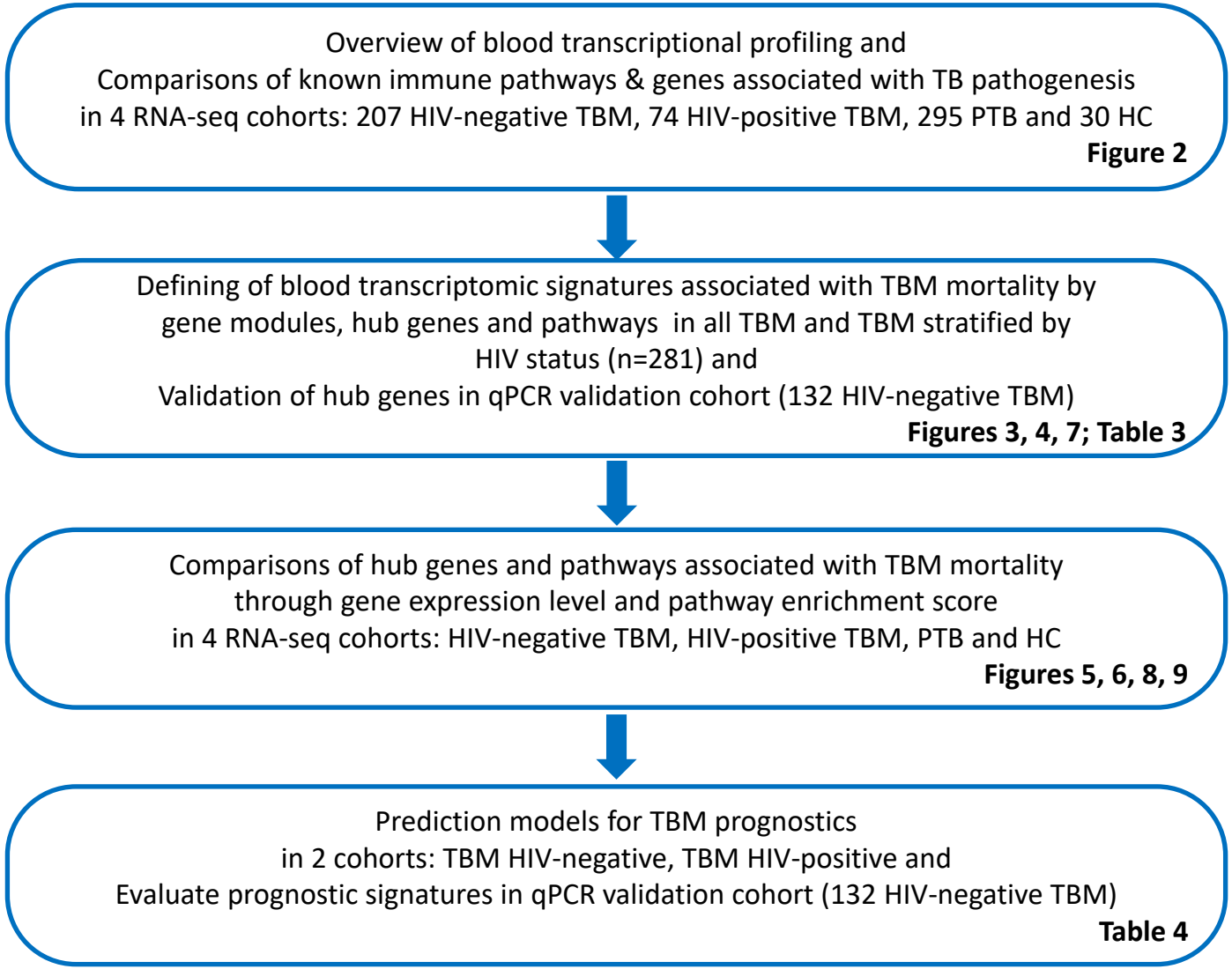
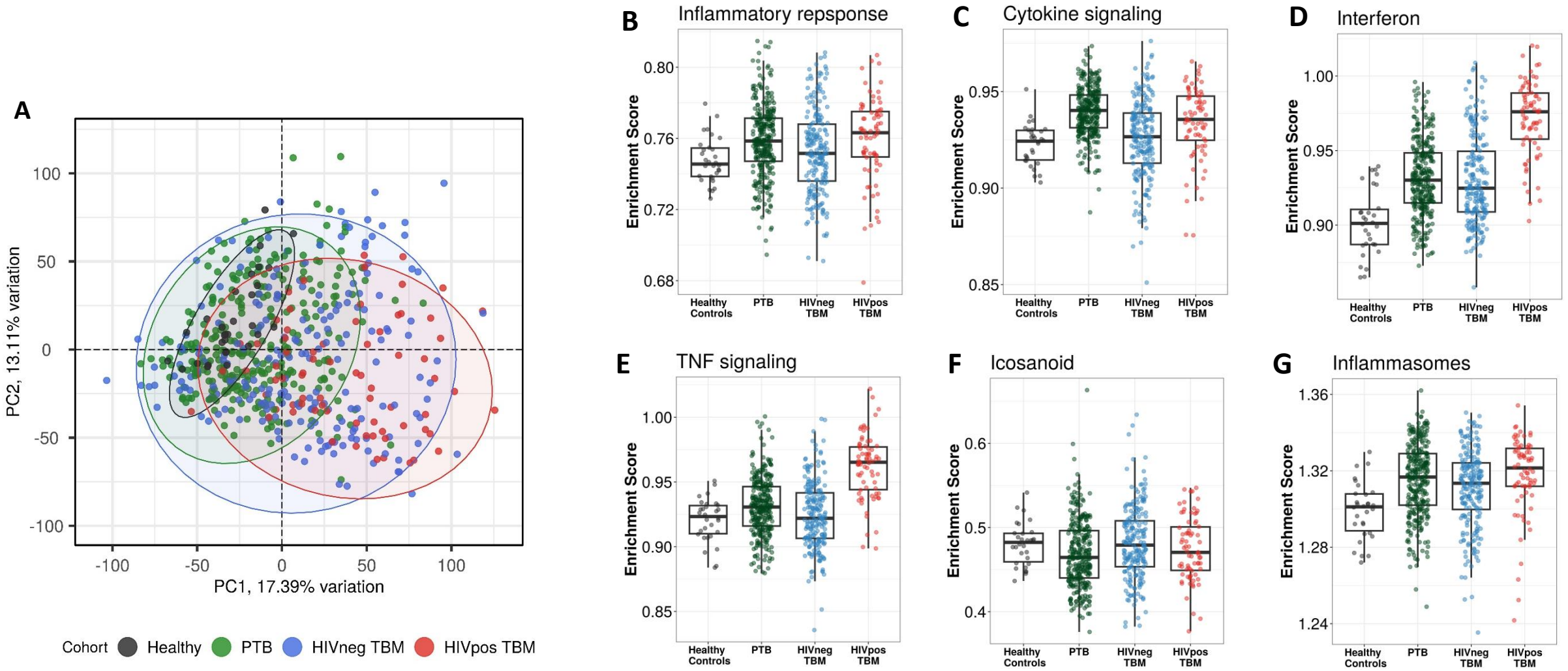


Figure 1. Objectives and cohorts flow

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



*HIVneg TBM: HIV-negative TBM, HIVpos TBM: HIV-positive TBM

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.

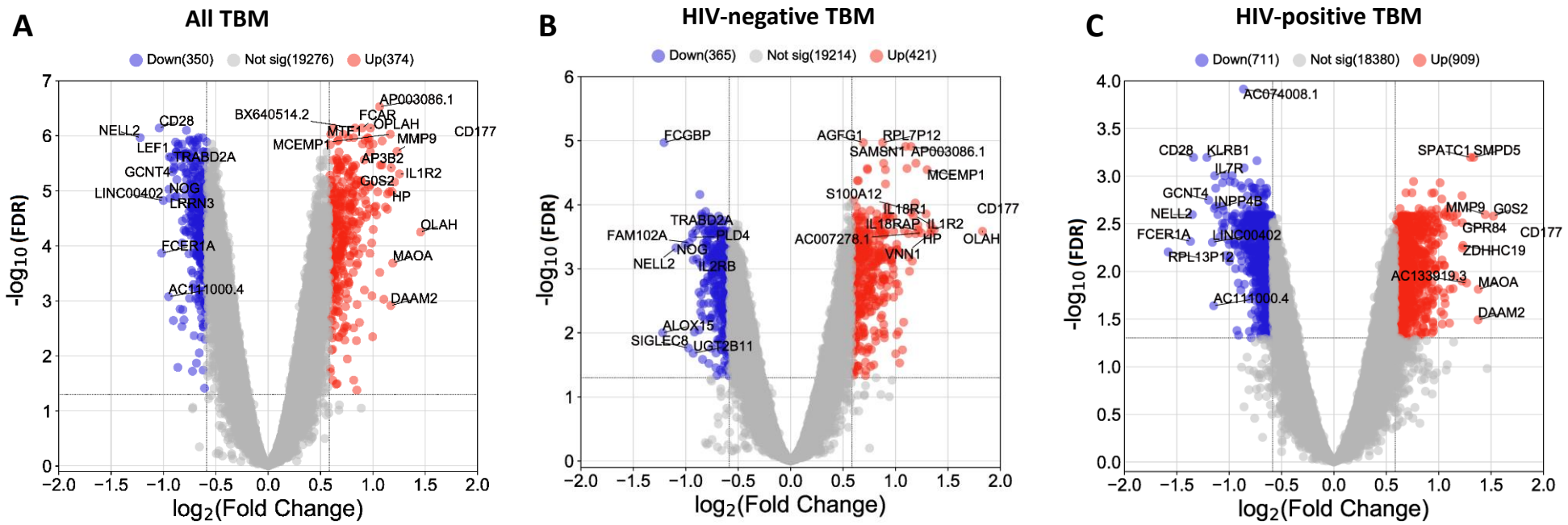


Figure 3: Blood transcriptomic profiles of three-month mortality at baseline in all TBM and TBM stratified by HIV status

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.

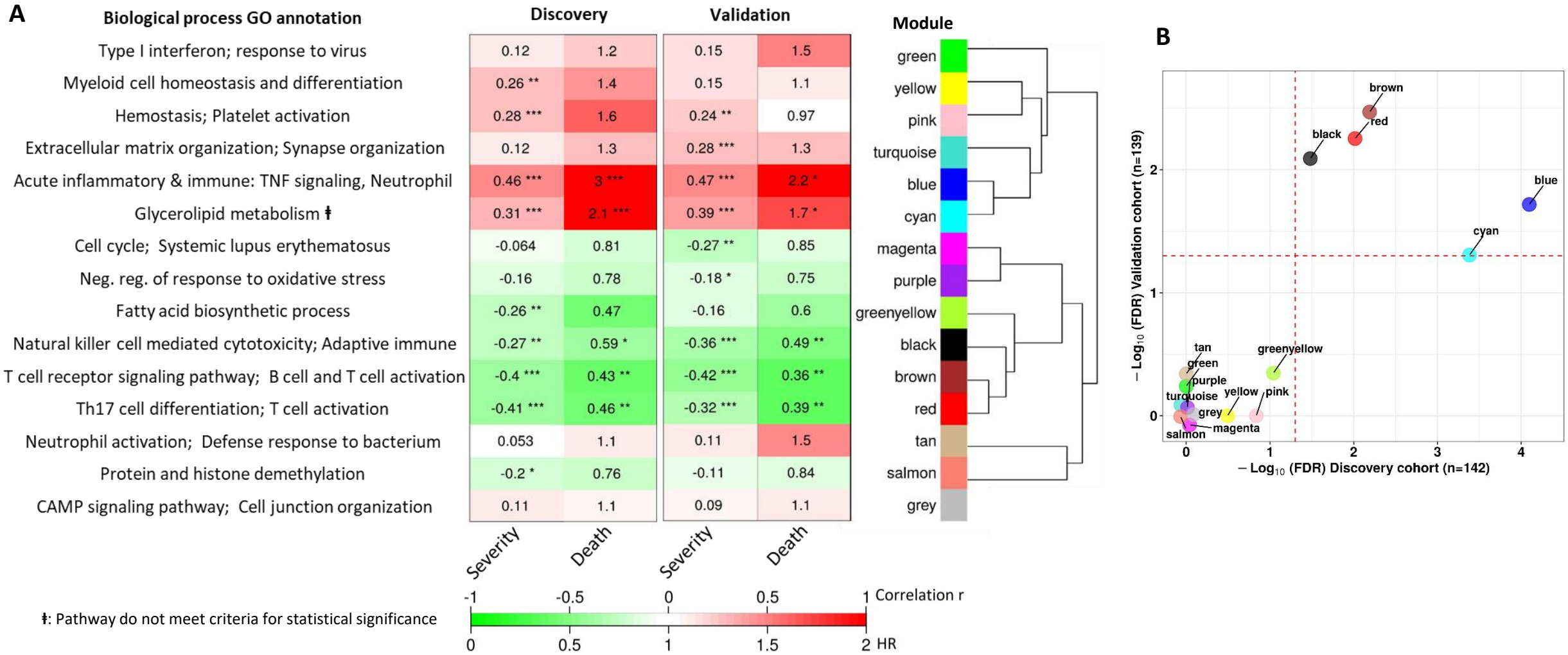


Figure 4. Blood transcriptional modules associated with mortality in TBM

(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.

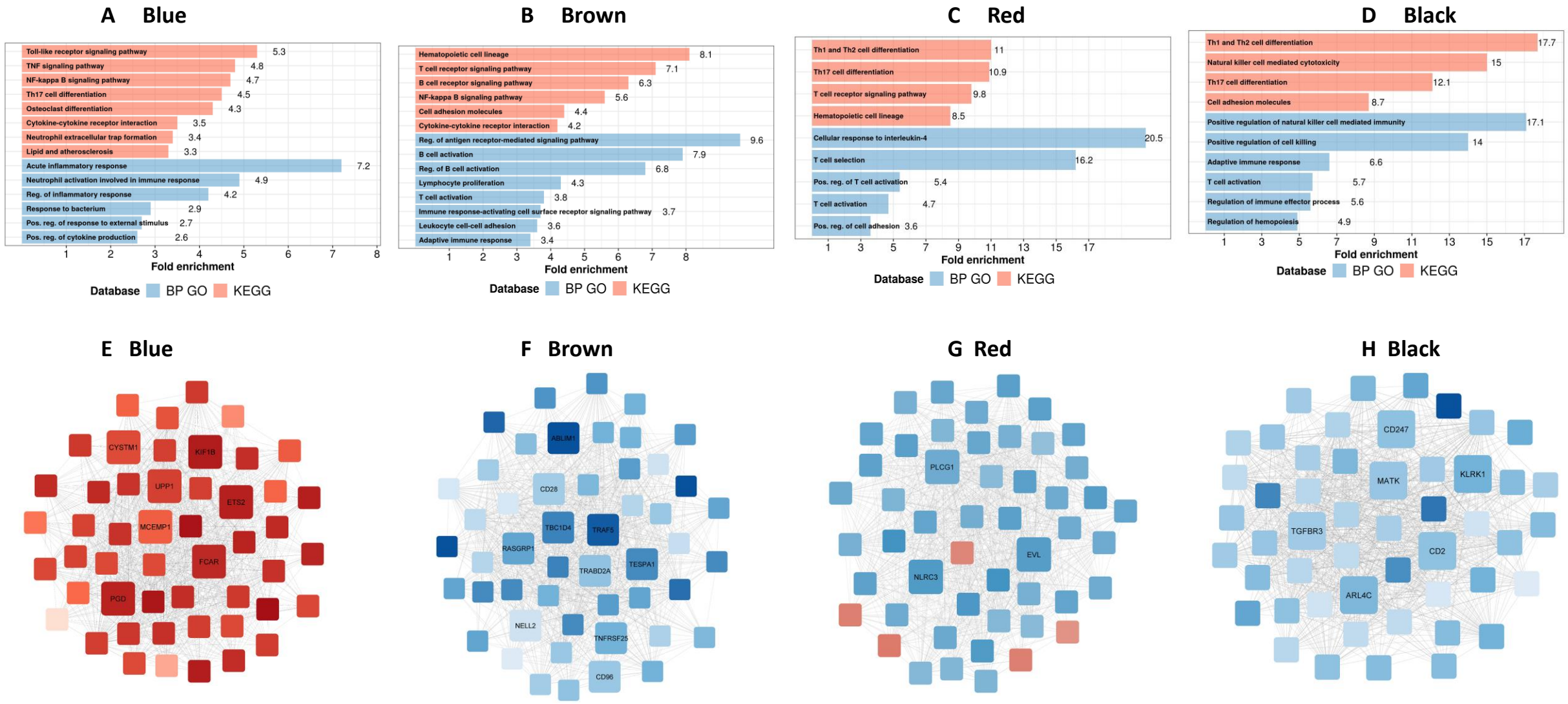


Figure 5. Biological processes, pathways and hub genes of validated modules associated with mortality

(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.

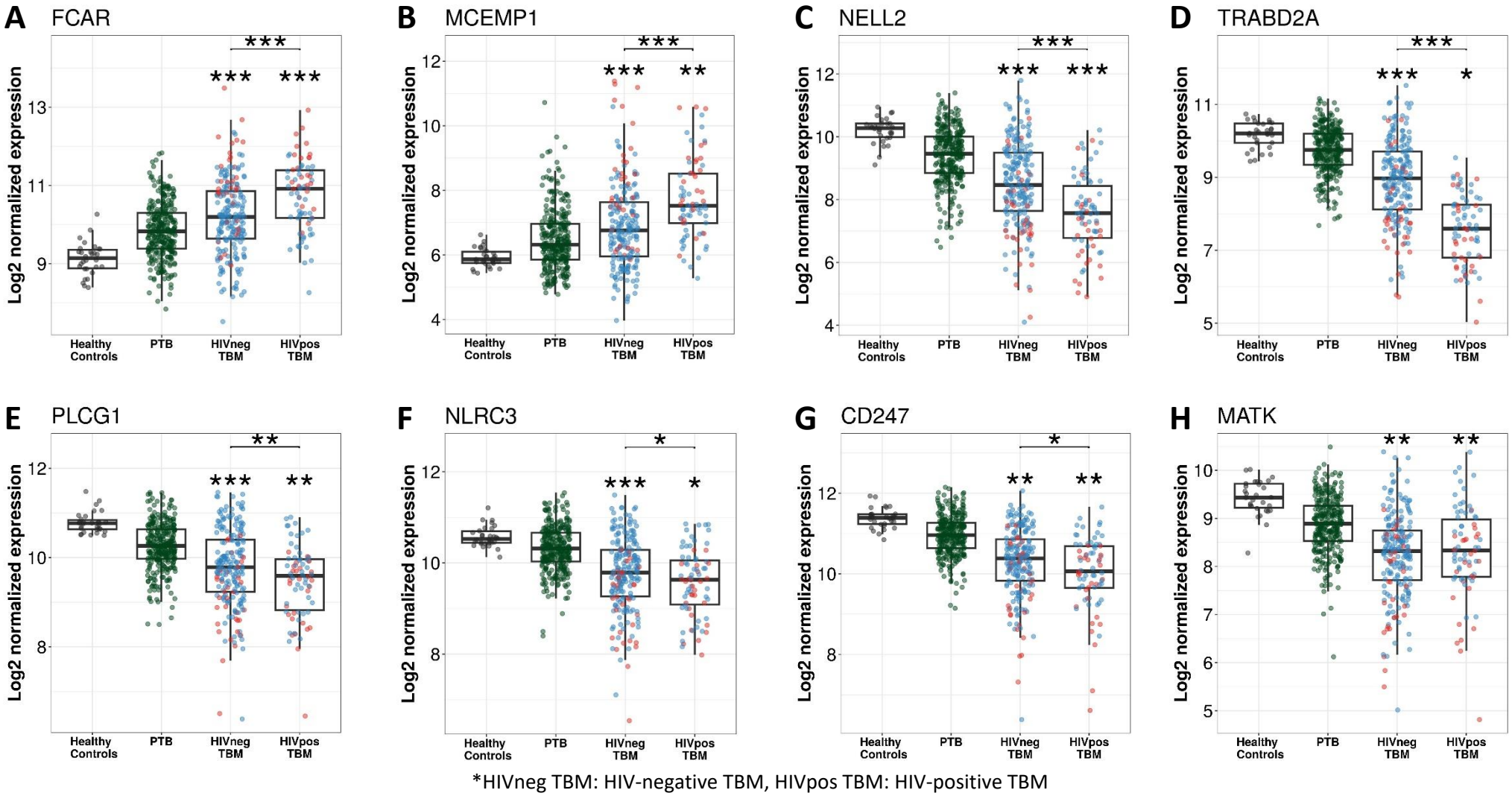
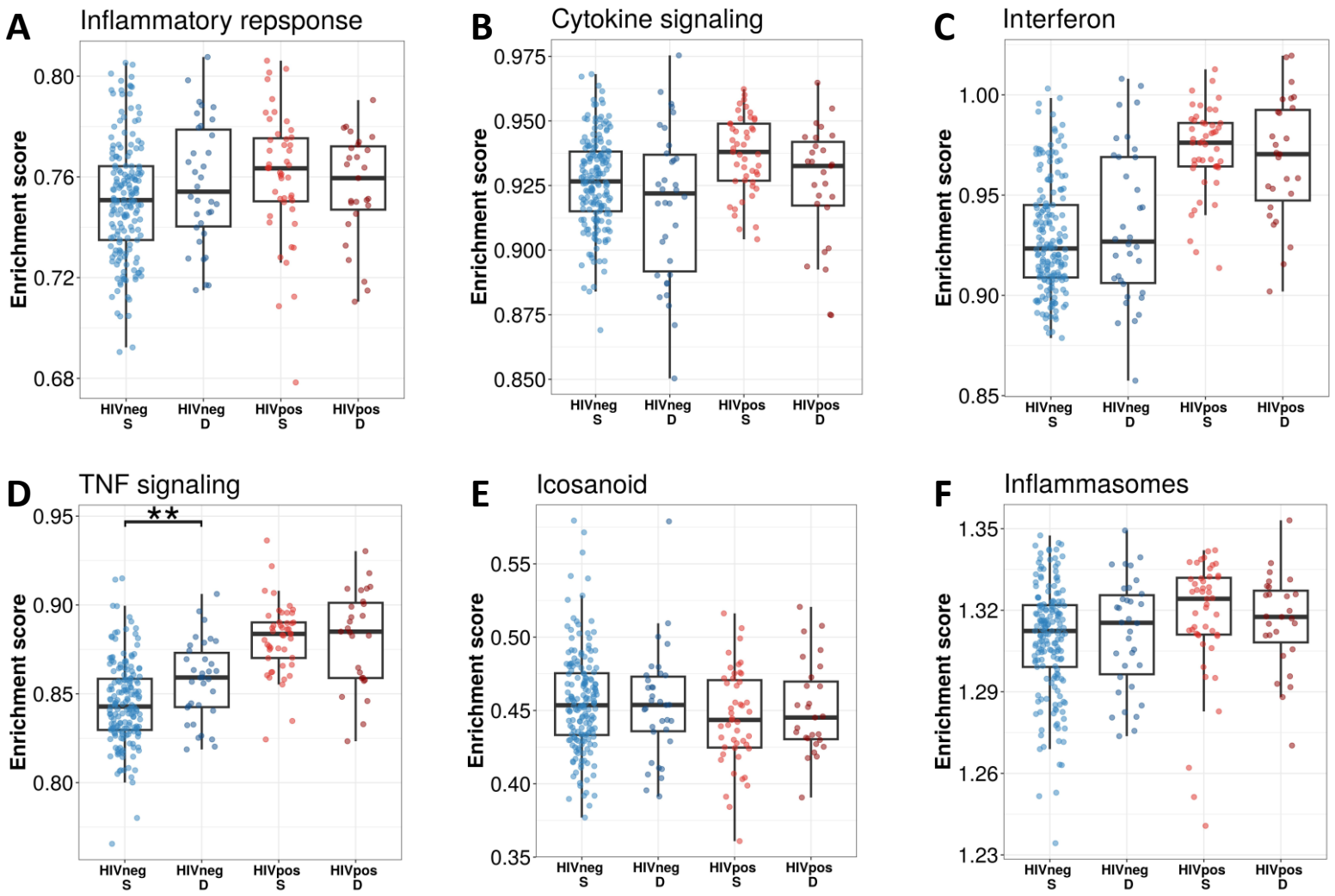


Figure 6. Gene expression of representative hub genes in healthy controls (n=30), PTB (n=295), HIV-negative TBM (n=207) and HIV-positive TBM (n=74)
Each dot represents gene expression from one participant. **(A, B)** expression of FCAR and MCEMP1 hub genes from the blue module. **(C, D)** expression 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

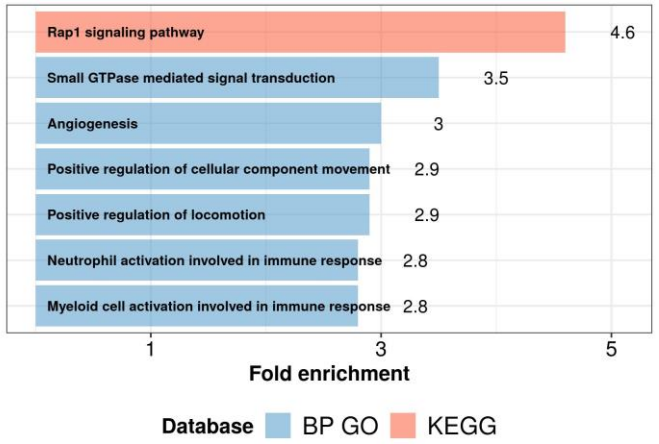
Biological process GO annotation	HIV-negative Severity	HIV-negative Death	HIV-positive Severity	HIV-positive Death	Consensus Severity	Consensus Death	Module
B cell activation ǂ	-0.19 **	0.77	-0.23 *	0.73	-0.19 *	0.77	pink
Cell division; Organelle fission	-0.23 ***	1.1	-0.077	0.63	-0.077	NA	purple
Circulatory system process	-0.25 ***	0.41 *	-0.039	0.81	-0.039	0.81	cyan
T cell activation; Adaptive immune response	-0.41 ***	0.4 **	-0.41 ***	0.58 *	-0.41 ***	0.58 *	brown
T cell activation	-0.4 ***	0.3 ***	-0.33 **	0.72	-0.33 **	0.72	Ered
B cell activation	-0.3 ***	0.34 ***	-0.29 *	0.81	-0.29 *	0.81	black
Wnt signaling pathway	-0.41 ***	0.31 ***	-0.31 **	0.6	-0.31 **	0.6	midnightblue
Blood vessel development; Angiogenesis	0.22 **	0.94	0.26 *	1.7 *	0.22 *	NA	blue
Extracellular matrix organization; Cell-substrate adhesion	-0.13	0.31 **	-0.057	1.1	-0.057	NA	yellow
Neutrophil activation; NET	0.077	1.5	0.23 *	1.2	0.077	1.2	salmon
Interferon-gamma; Response to virus	0.16 *	1.4	0.25 *	1.2	0.16 *	1.2	greenyellow
Type I interferon signaling pathway; Response to virus	-0.038	1.1	0.12	1.2	NA	1.1	tan
Neutrophil activation; TNF signaling pathway	0.45 ***	3.7 ***	0.46 ***	1.7 *	0.45 ***	1.7 *	turquoise
Erythrocyte development	0.13	1	0.34 **	1.3	0.13	1	green
Platelet activation; Hemostasis	0.22 **	1.2	0.34 **	1.1	0.22 **	1.1	magenta
Cell junction organization	0.068	1.1	0.033	0.9	0.033	NA	grey

ǂ: Pathway do not meet criteria for statistical significance

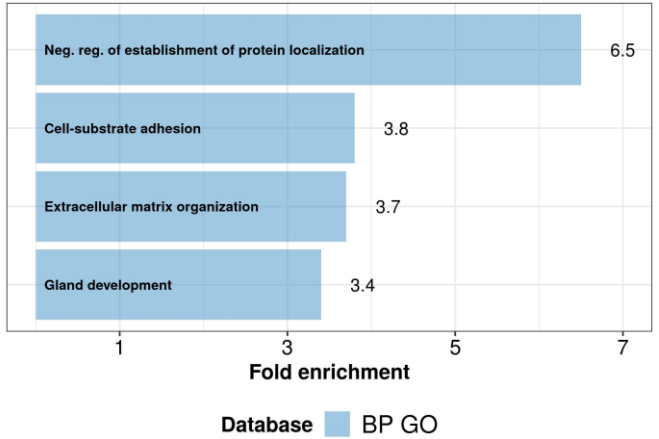
Correlation r: -1, -0.5, 0, 0.5, 1

HR: 0, 0.5, 1, 1.5, 2

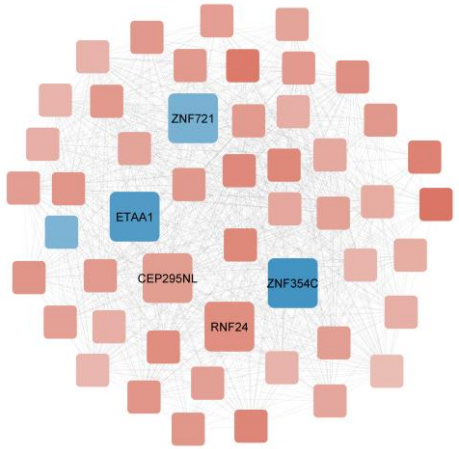
B Blue module in HIV-positive



C Yellow module in HIV-negative



D Hub genes in blue module



E Hub genes in yellow module

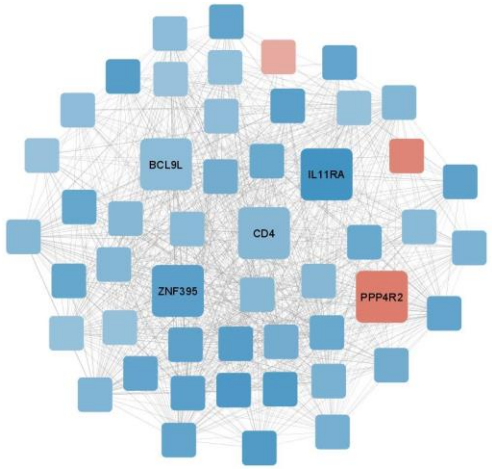


Figure 8. Consensus transcriptional modules associated with TBM mortality stratified by HIV-infection

(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.

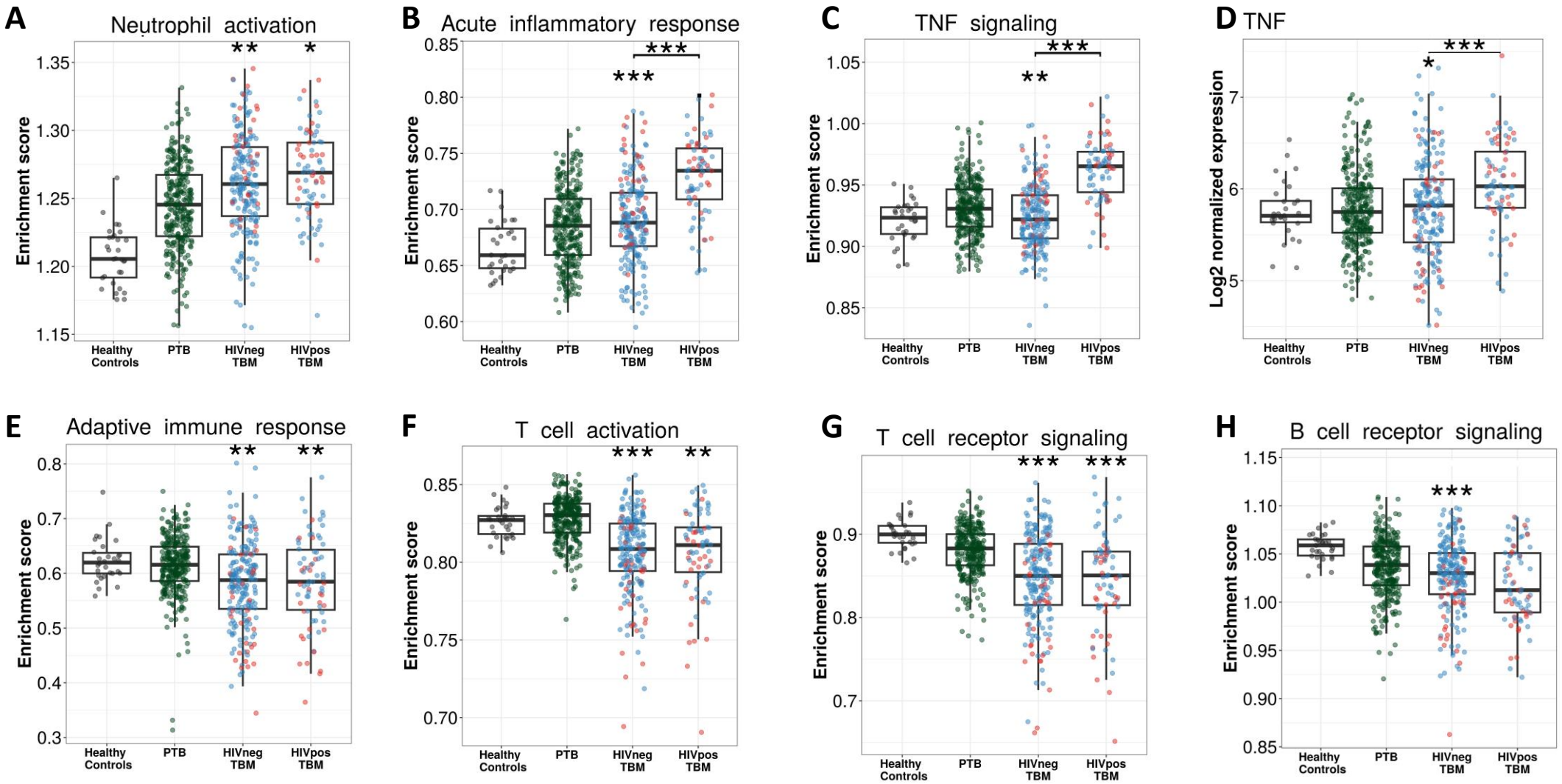


Figure 9. Enrichment score of immunity pathways in healthy controls (n=30), PTB (n=295), HIV-negative TBM (n=207) and HIV-positive TBM (n=74)
Pathway enrichment scores were calculated using single sample GSEA algorithm (ssGSEA) (Barbie DA, et al 2009). Each dot represents one participant. **(A-C)** showed 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).

Supplementary Figures (n=10)

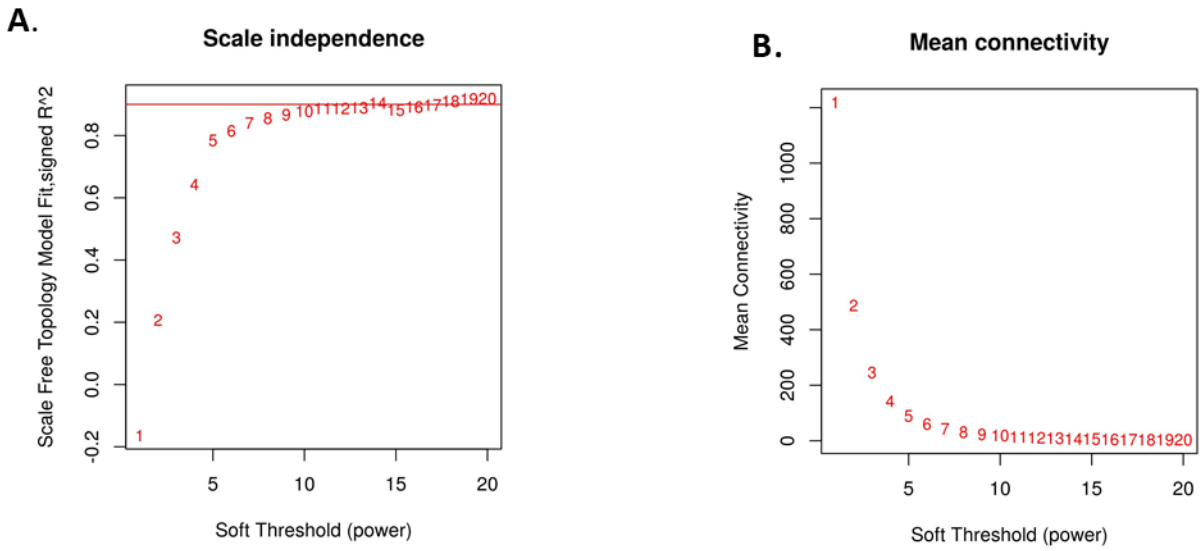
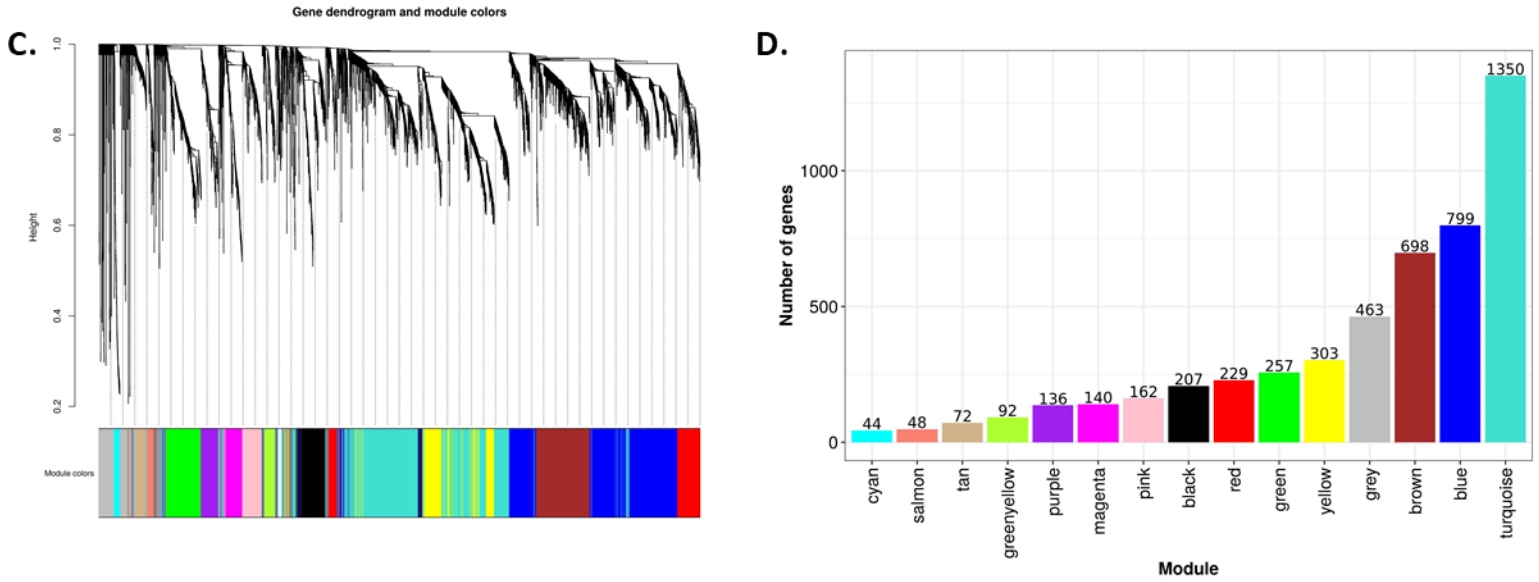


Figure S1. Construction of WGCNA in discovery cohort

Analysis of network topology for various soft-thresholding powers of the top 5,000 most variant genes based on the scale-free network model. **(A)** showed the scale-free fit index (y-axis) as a function of the soft-thresholding power (x-axis). The red horizontal line corresponds to $R^2 = 0.85$ and soft-thresholding power β=8, which was chosen for the construction gene-expression network. **(B)** displayed the mean connectivity on the y-axis as a function of the soft-thresholding power on the x-axis. The adjacency matrix of the scale-free network between genes was determined as $A = (a_{ij})$, where $a_{ij} = |cor(gene_i, gene_j)|^{\beta}$. **(C)** displays the dendrogram resulted from the hierarchical clustering analysis using topological overlap of the adjacency matrix A served as a dissimilarity metric. Each cluster was referred as a module and assigned with a color. **(D)** showed bar-plots indicate the number of genes contained in each module.



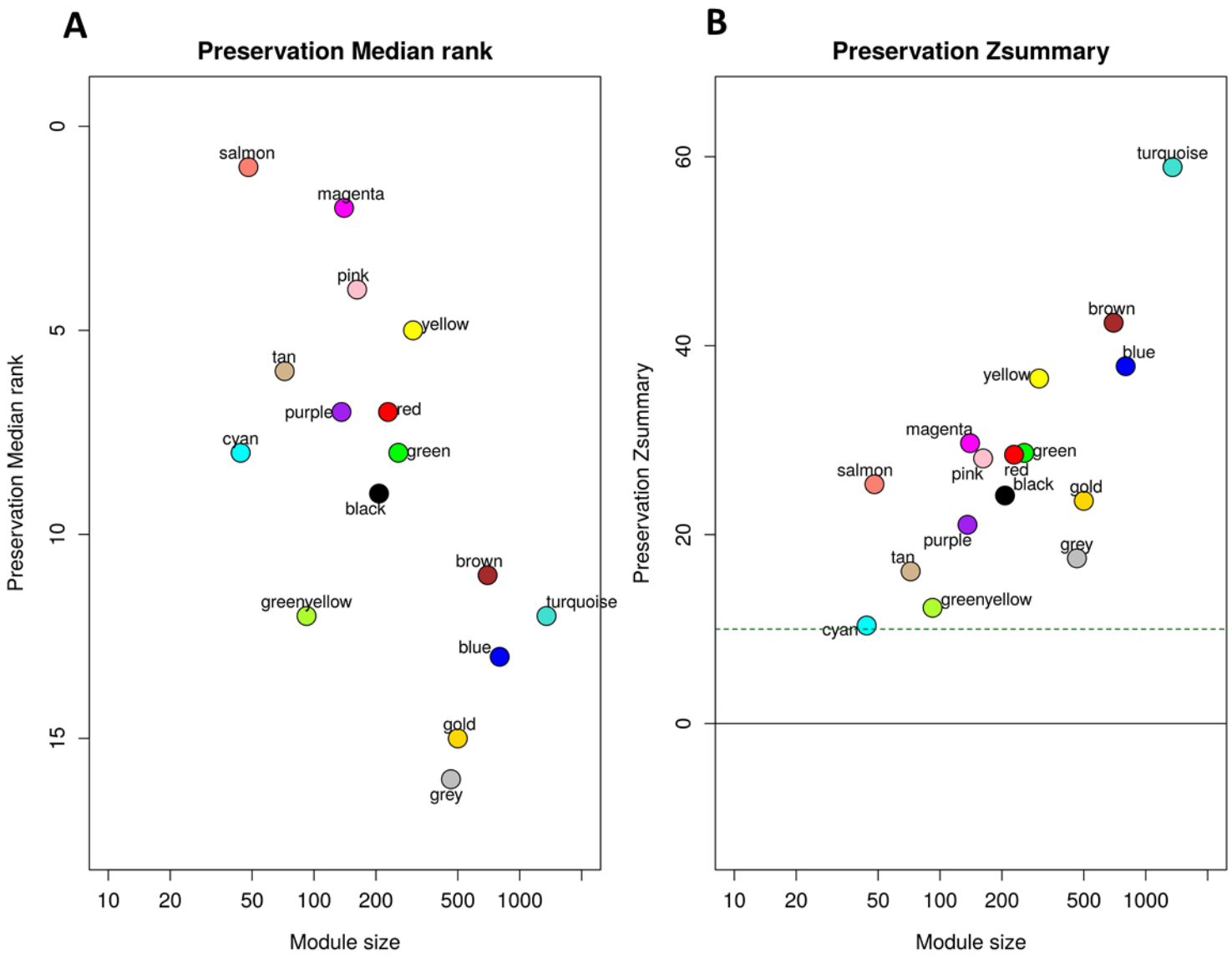


Figure S2. Preservation of discovery modules in validation cohort
(A) The median rank preservation statistics of the modules. Each module was represented by a point, labeled with the corresponding color and name. The Y-axis represents the median rank of observed preservation statistics per module, while the X-axis indicates the number of genes within each module. A low preservation median rank value indicates a high level of preservation. The gold module was an artificial module comprised of 500 randomly selected genes. The grey module consisted of non-connected genes identified in the WGCNA analysis of the discovery cohort. These two modules exhibited the highest median rank, indicating low preservation, which ensured our preservation analysis controlling well the background noise signal. **(B)** The Zsummary preservation statistics of the modules. Each module was represented by a point, labeled with the corresponding color and name. The Y-axis represents the Zsummary statistic of each module based on 1,000 permutations of module labels, while the X-axis indicates the number of genes within each module. The horizontal green dash line corresponds to the threshold of Zsummary>10 indicating strongly persevered modules. The presence of all modules in discovery were validated in validation cohort with Zsummary>10.

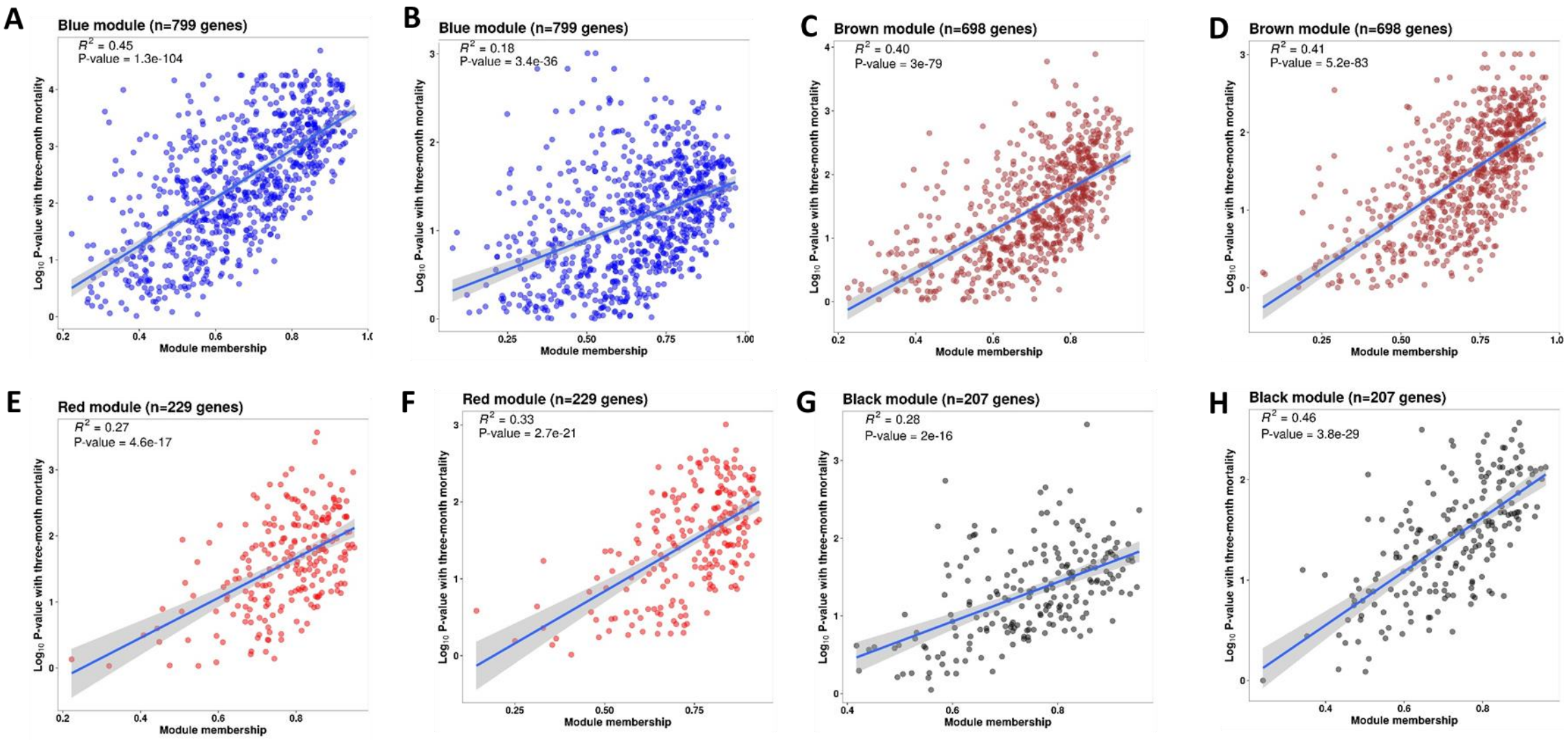


Figure S3. Correlations between gene module membership and gene significance with mortality in 4 associated modules in discovery and validation cohorts
(A, C, E and G) are scatter plots for blue, brown, red and black modules in discovery cohort. (**B, D, F and H)** are scatter plots for blue, brown, red and black modules in validation cohort. For each module, each dot represents one gene in the module. The X-axis represents module membership calculated by Pearson correlation between gene expression level and its corresponding PC1 of that module. The Y-axis represents $-\log_{10}$(p-value) of the association between gene expression and three-month mortality based on the Cox regression model adjusted for age, HIV, and dexamethasone treatment. Hub-genes were those on the right corner of the plot. The displayed R^2 and P value were taken from linear regression models.

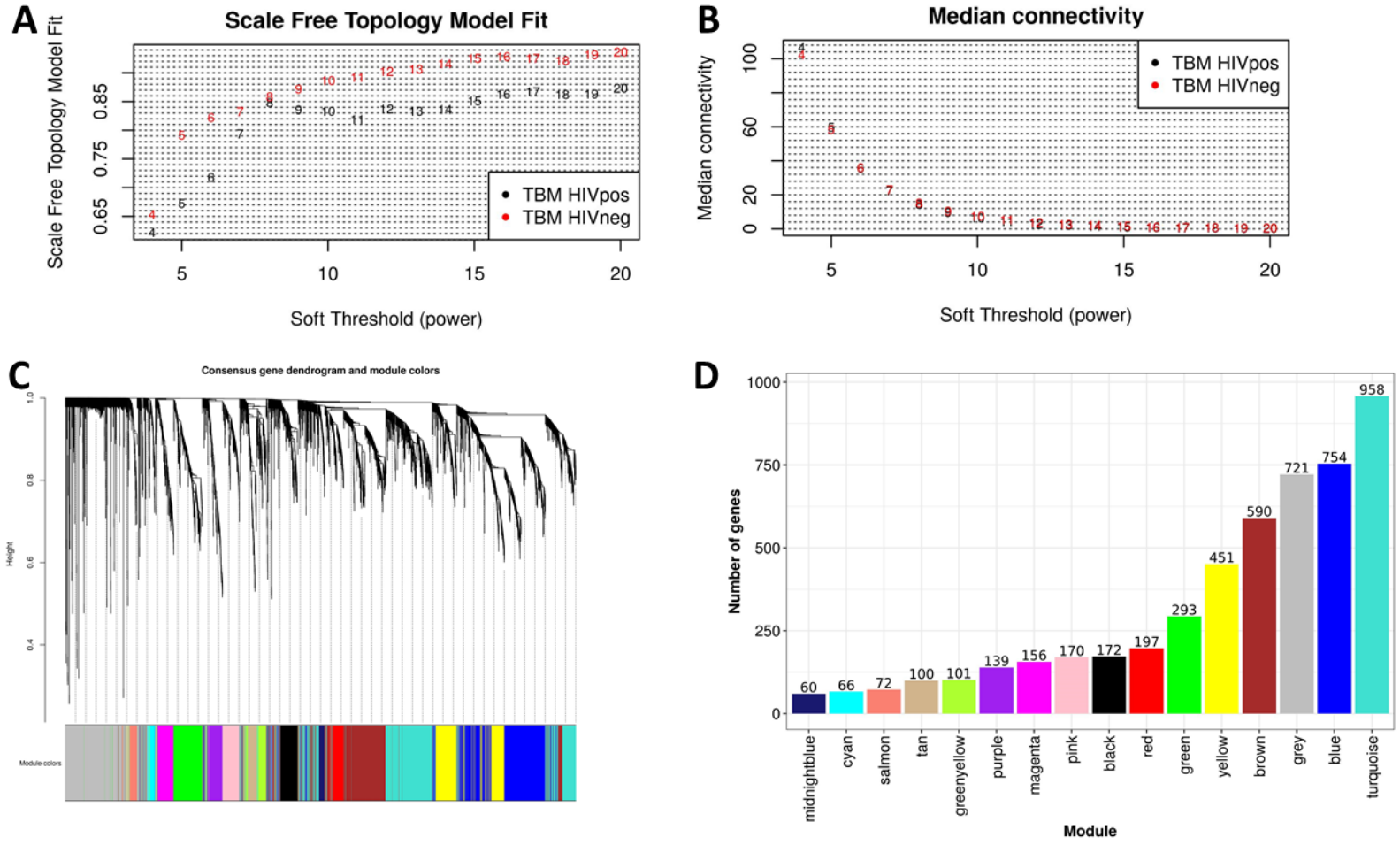
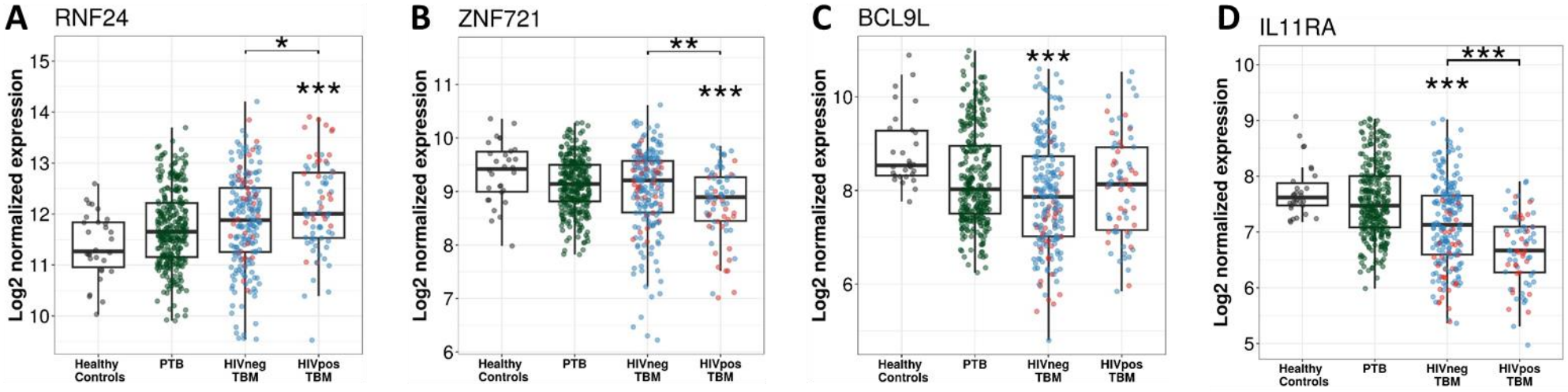


Figure S4. Construction of consensus WGCNA in HIV-negative (n=207) and HIV-positive (n=74)

Consensus analysis of network topology for various soft-thresholding powers of the top 5,000 most variant genes based on the scale-free network model. **(A, B)** showed the summary network indices (y-axes) as functions of the soft thresholding power (x-axes). Numbers in the plots indicate the corresponding soft thresholding powers. The plots indicated that approximate scale-free topology is attained around the soft-thresholding power of 8 for both sets, which is the lowest power that satisfies the approximate scale-free topology criterion for both cohorts, HIV-positive (black dots) and HIV-negative (red dots). The adjacency matrix of the scale-free network between genes of each cohort was determined and scaled. The consensus topological overlap of two adjacency matrix were identified and served as dissimilarity metric for the hierarchical clustering. **(C)** displayed the dendrogram of the consensus gene co-expression network. Each cluster was referred as a module and assigned with a color. **(D)** the bar-plot indicates the number of genes contained in each module.



*HIVneg TBM: HIV-negative TBM, HIVpos TBM: HIV-positive TBM

Figure S5. 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 RNF24 and ZNF721 hub genes from the specific blue module. **(C, D)** expression of BCL9L and IL11RA hub genes from the specific yellow modules. 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).

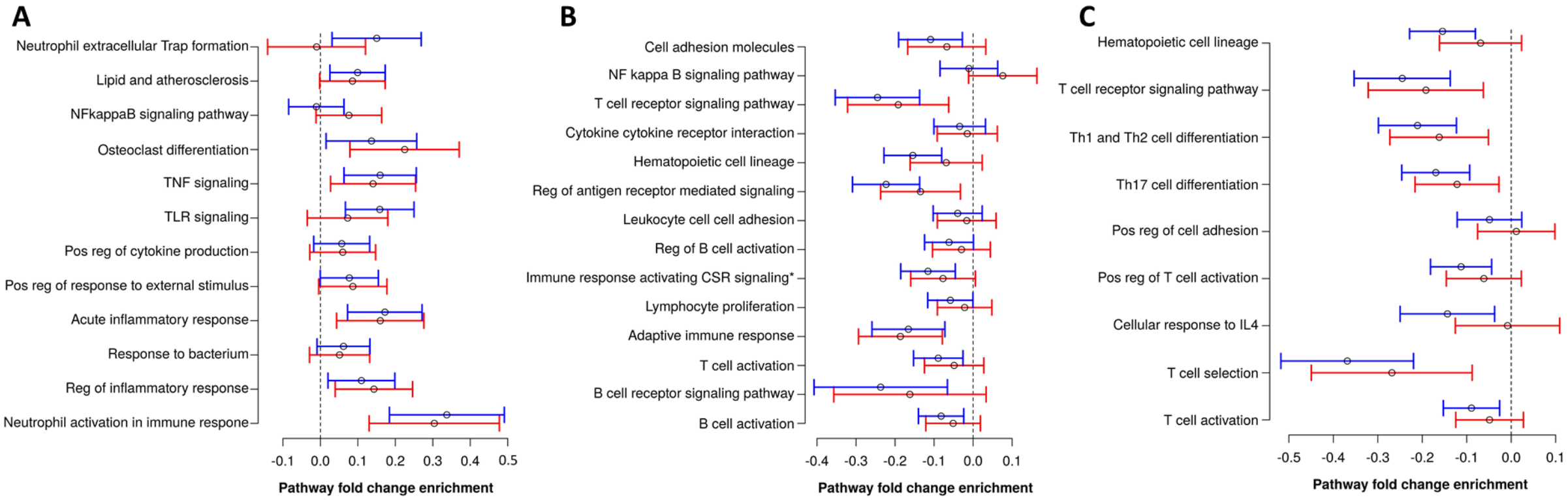
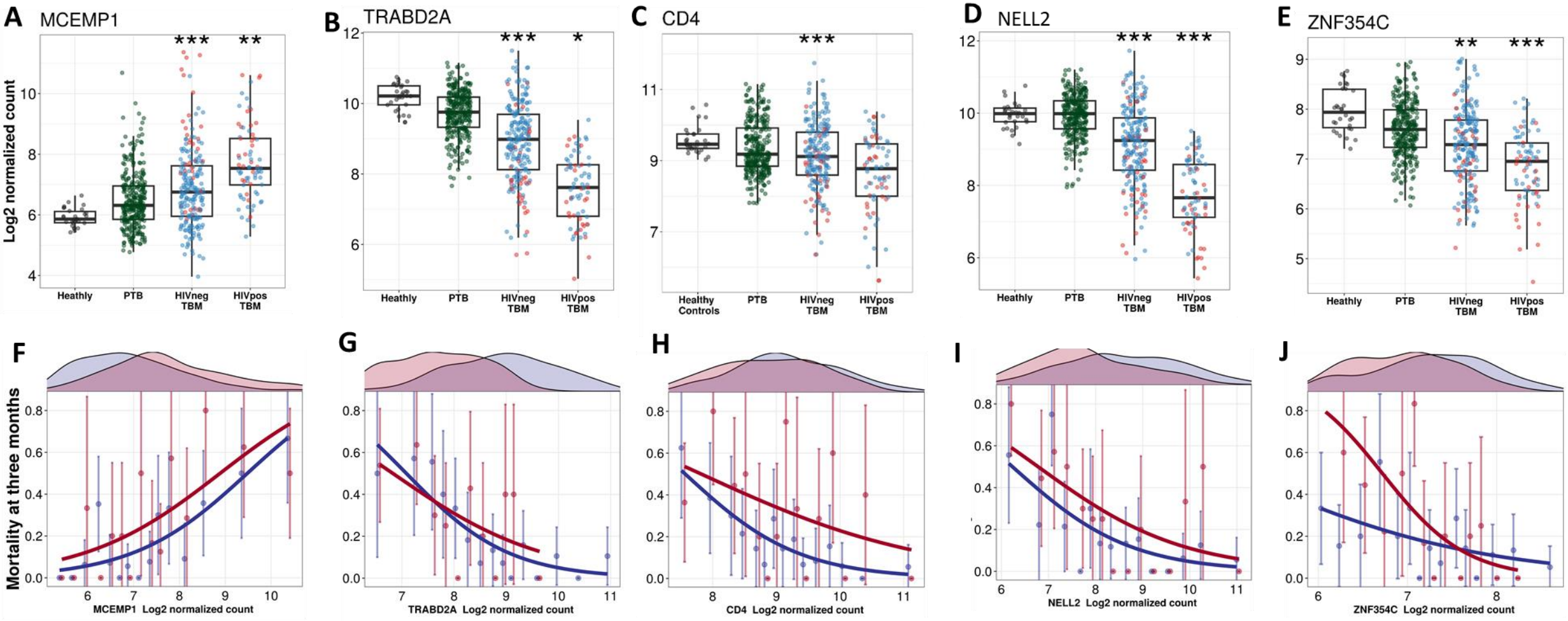


Figure S6. Pathway fold change enrichment to mortality of top hit pathways in blue, brown and red modules from the primary analysis

Pathway fold change enrichment to mortality of each biological pathway defined as the distribution of the total of mean differential expression between dead and survival TBM of all genes involved in the pathway based on Qusage method (Yaari G at el, 2013). Median and 95% confidence interval of pathway fold change of each top-hit biological pathway in the blue **(A)**, brown **(B)**, and red module **(C)** in the primary analysis. Blue lines are from HIV-negative cohort and red lines from HIV-positive cohort. Pathway fold change enrichment above 0 indicates up-regulation while below 0 indicates down-regulation in dead. Vertical dashed lines indicate no difference in dead compared to survival TBM.



*HIVneg TBM: HIV-negative TBM, HIVpos TBM: HIV-positive TBM

Figure S7. Gene expression and mortality of selected hub genes in TBM, PTB and healthy participants

(A-E) Boxplots visualized distribution of gene expression of MCEMP1, TRABD2A, CD4, NELL2 and ZNF354C in healthy (n=30), PTB (n=295), TBM without HIV (n=207) and with HIV (n=74). Boxes indicate median and inter-quantile range. Dots indicate data in individuals. The comparisons were made between dead (red) and survival (blue) by Wilcoxon rank sum test with p-value displayed as * < .05, ** < .01 and *** <.001. **(F-J)** Associations between mortality and gene expression of MCEMP1, TRABD2A, PKD1, NELL2 and ZNF354C. In each figure, the upper panel corresponds to gene distribution in HIV-negative (blue) and HIV-positive (red), the lower panel presents the approximation of the association between that gene and mortality. Their gene expression were divided into 15 groups using equal-distant quantiles (1/15,…, 14/15) intervals and the proportion of mortality within those groups of patients were computed. Each point (error bar) presented the proportion with its confidence interval of mortality per group. The line represents the logistic curve that illustrates the mortality trend corresponding to a two-fold increase in gene expression.

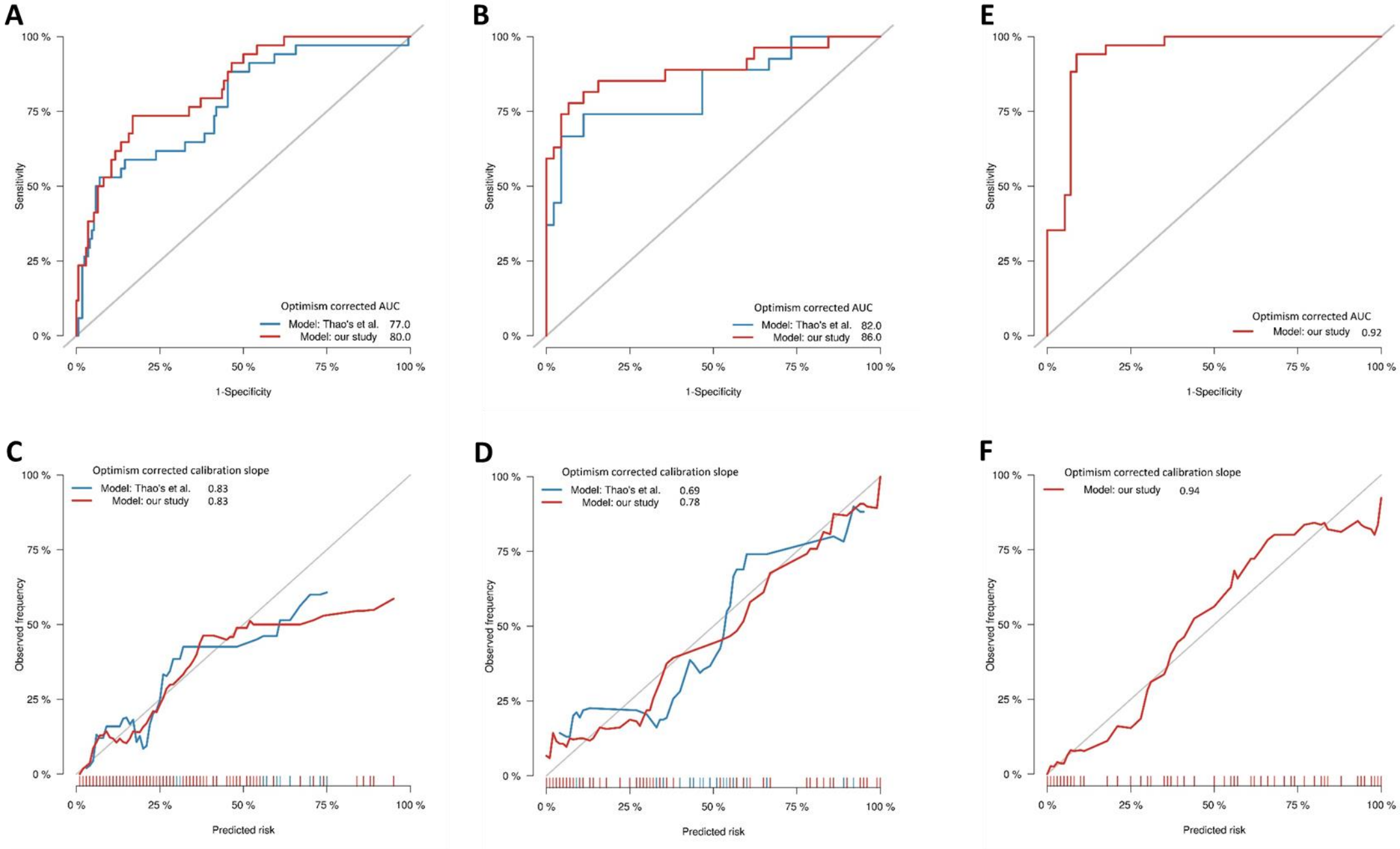

A
Sensitivity
1-Specificity
Optimism corrected AUC
Model: Thao's et al. 77.0
Model: our study 80.0
B
Sensitivity
1-Specificity
Optimism corrected AUC
Model: Thao's et al. 82.0
Model: our study 86.0
E
Sensitivity
1-Specificity
Optimism corrected AUC
Model: our study 0.92
C
Optimism corrected calibration slope
Model: Thao's et al. 0.83
Model: our study 0.83
Observed frequency
Predicted risk
D
Optimism corrected calibration slope
Model: Thao's et al. 0.69
Model: our study 0.78
Observed frequency
Predicted risk
F
Optimism corrected calibration slope
Model: our study 0.94
Observed frequency
Predicted risk
0 %
25 %
50 %
75 %
100 %

Figure S8. Performance of optimal gene set for TBM mortality prediction

Receiver operating characteristic (ROC) curves for three-month mortality for our developed model using expression level of 4 genes (MCEMP1, NELL2, CD4 and ZNF354C) with 2 clinical predictors (age, MRC grade) (red line) and the model develop by Thao et al. (PMID: 29029055) (blue line) for HIV-negative cohort **(A)**, for HIV-positive cohort **(B) and for HIV-negative PCR validation cohort (E)**. The optimism corrected area under the curve (AUC) was calculated based on 1,000 times bootstrap subsampling as described in the methods. Calibration plots were presented for the corresponding prediction models for HIV-negative **(C)**, HIV-positive cohort **(D) and HIV-negative PCR validation cohort (E)**.

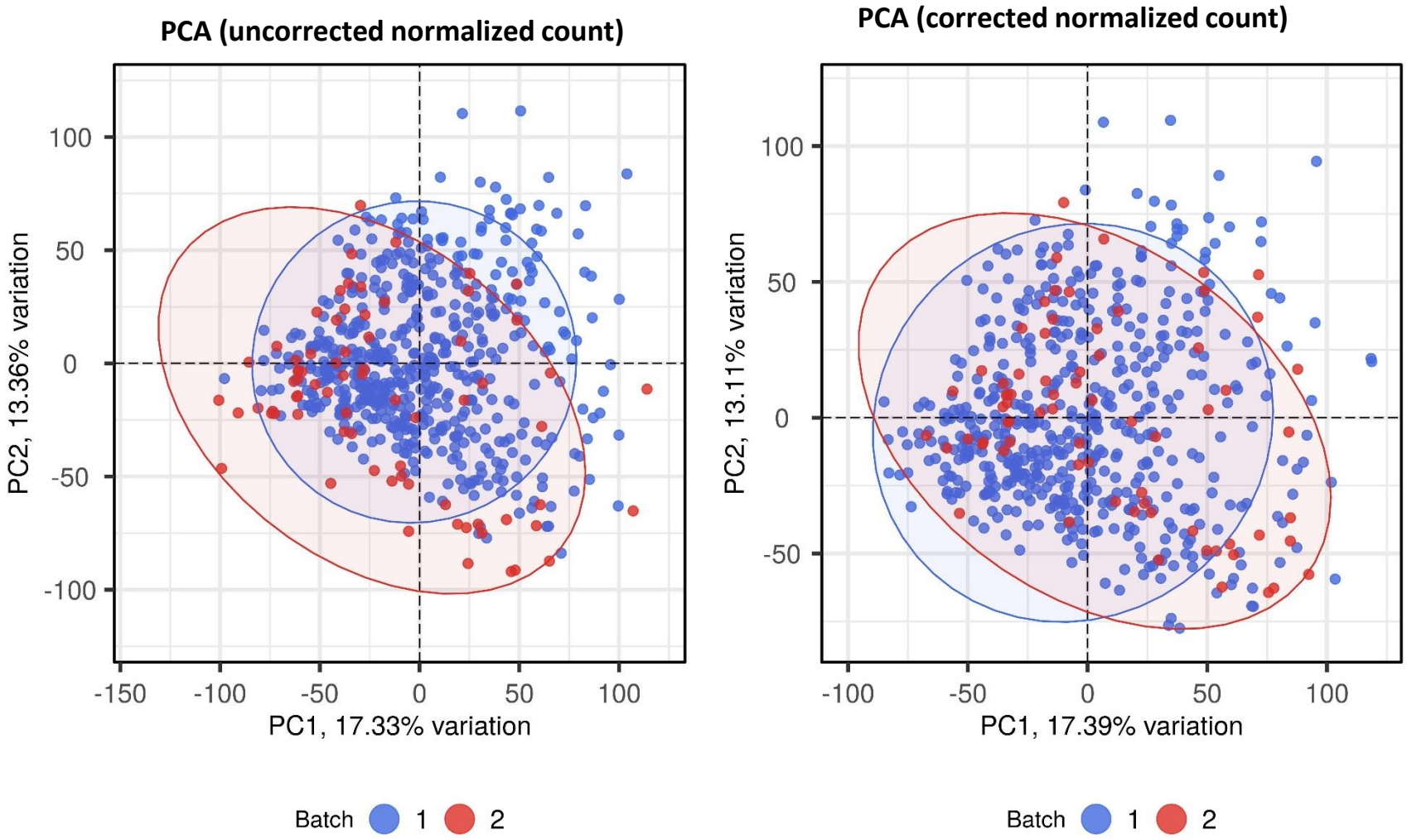


Figure S9. Batch correction for RNA-seq normalized count data

Principle component analysis (PCA) of transcriptomic data before (A) and after batch correction by combat function in SVA r package. Each symbol represents one individual with color coding different cohorts. The x-axis represents principle component (PC) 1, while y-axis represents PC2

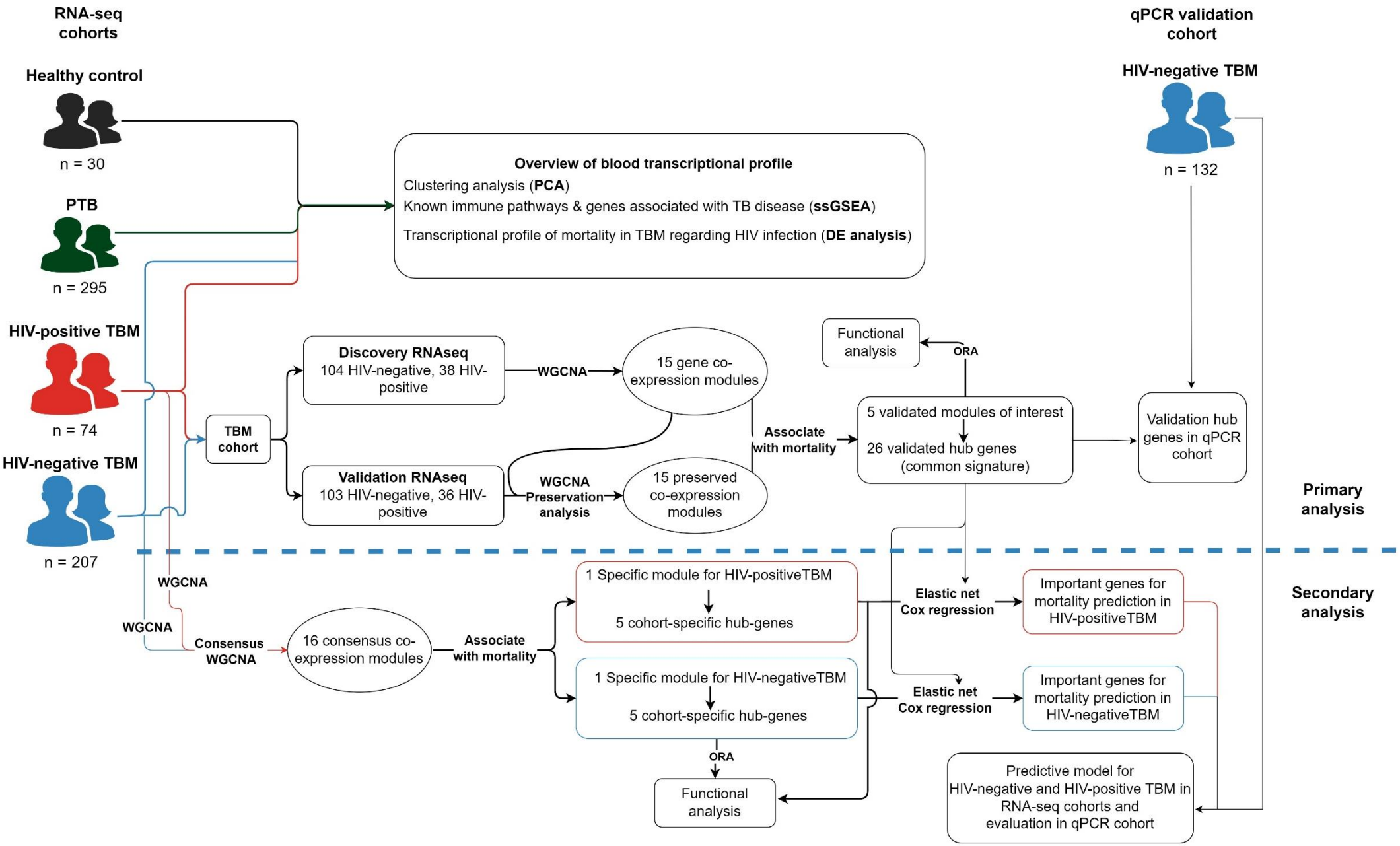

RNA-seq cohorts
Healthy control
n = 30
PTB
n = 295
HIV-positive TBM
n = 74
HIV-negative TBM
n = 207
qPCR validation cohort
HIV-negative TBM
n = 132
Overview of blood transcriptional profile
Clustering analysis (PCA)
Known immune pathways & genes associated with TB disease (ssGSEA)
Transcriptional profile of mortality in TBM regarding HIV infection (DE analysis)
TBM cohort
Discovery RNAseq
104 HIV-negative, 38 HIV-positive
WGCNA
15 gene co-expression modules
Validation RNAseq
103 HIV-negative, 36 HIV-positive
WGCNA Preservation analysis
15 preserved co-expression modules
Associate with mortality
Functional analysis
ORA
5 validated modules of interest
26 validated hub genes (common signature)
Validation hub genes in qPCR cohort
Primary analysis
Secondary analysis
WGCNA
WGCNA
Consensus WGCNA
16 consensus co-expression modules
Associate with mortality
1 Specific module for HIV-positiveTBM
5 cohort-specific hub-genes
1 Specific module for HIV-negativeTBM
5 cohort-specific hub-genes
ORA
Functional analysis
Elastic net Cox regression
Important genes for mortality prediction in HIV-positiveTBM
Elastic net Cox regression
Important genes for mortality prediction in HIV-negativeTBM
Predictive model for HIV-negative and HIV-positive TBM in RNA-seq cohorts and evaluation in qPCR cohort

Figure S10. Analysis workflow diagram

TBM participants were recruited from two randomized control trials for adjunctive dexamethasone treatment of HIV-negative and HIV-positive adults with TBM (*LAST ACT: NCT03100786; ACT HIV: NCT03092817). Samples from 207 HIV-negative and 74 HIV-positive patients were used for RNA sequencing and randomly divided into discovery and validation cohorts to identify biological pathways and hub genes associated with three-month mortality.