Redefining hematopoietic progenitor cells and reaffirming lineage commitment by transcriptional atlas

Yongjian Yue[1]

[1]Shenzhen Institute of Respiratory Disease, the First Affiliated Hospital of Southern University of Science and Technology, Shenzhen, Guangdong, China

*** Correspondence:** Yongjian Yue
Address: Shenzhen People's Hospital, The First Affiliated Hospital of Southern University of Science and Technology, No. 1017 Dongmen North Road, Luohu District, Shenzhen 518020, China.
Email: yueyongj@163.com
Tel: +86 15919909356

Abstract

Tracing the differentiation of the hematopoietic progenitor cells(HPCs) in human blood, particularly with the scarcity in adults, has been challenging owing to unable capture the entire differentiation stages and subgroups of HPCs. Deciphering the underlying hematopoiesis mechanism, including progenitor cells definition and fate decisions identification, is critical to understanding fundamental questions of hematopoietic lineage commitment. Often, the use of classic and traditional surface markers has been shown to be inadequate, speciosity, and ineffective in accurately defining hematopoietic progenitors. The lineage definition and commitment of hematopoietic progenitors in adult peripheral blood remains to be elucidated by a reliable transcriptional atlas. Here, single cell RNA sequencing (scRNA seq) was employed to define the lineage of HPCs and reaffirm the hierarchy of hematopoiesis. A total of 27,566 adult peripheral blood CD34+ cells were enriched and sequenced, capturing most differentiation stages and subgroups of HPCs. HPCs lineages were redefined by unique makers and key factor genes. We are the first to clearly depict the transcriptional atlas and define each subgroup of adult peripheral blood progenitors. We recovered the key transcript factors involved in lineage fate decision, and restructured the lineage commitment model of HPCs that can reconcile previous models and explain conflicts among functional studies. Several fate decision key factors of progenitors were identified as major switches that controlled hematopoietic lineage formation. Moreover, we are also the first to identify the cell fate decision transcriptional factors in each branch among hematopoietic hierarchy formation. Our reaffirmed lineage commitment also revealed that hematopoiesis is a continuum process controlled by the activation and suppression of transcription factors with identified boundary. With these information, an elaborate and comprehensive transcriptional atlas of all types of hematopoietic progenitors was then constructed, which enabled the whole map of the transcriptional signal and pathway among lineage commitment. Importantly, this newly created atlas can enable accurate definition, transplantation, and biological insights of hematopoietic progenitors for facilitating new research development in cell engineering, regenerative medicine, and disease

treatment, which we call Artificial Operation of HPCs. Furthermore, this opens up a wide variety of possibilities for artificial precise control progenitor cell differentiation and targeted treatment of immune disease, providing hope for untreatable diseases.

Graphical abstract

The reaffirmed lineage commitment model of HPCs

Highlights

- Classic and traditional surface markers are often inadequate, speciosity, and ineffective in accurately defining hematopoietic progenitor cells.
- Multiple hematopoietic lineages progenitor cells of adult peripheral blood were clearly depicted and redefined by unique makers and key transcript factors.
- Key fate decision transcript factors and genes in each branch among hematopoietic hierarchy formation were recovered by transcriptional atlas.
- Several key factors, including NFE2, LYL1, MYB, CDK4, ETV6, TESPA1, ZBTB16, and GATA2 of myeloid progenitors and RUNX2, HOXA9, JUN/FOS, TCF4, DDIT4, HOPX, KLF10, HOXA3, and TSC22D1 of lymphoid progenitors, were identified as major regulators that controlled hematopoietic lineage formation.
- Fate decision transcription factors were identified in progenitor cells of erythrocyte (GATA1, KLF1), monocyte (MAFB, IRF8), B cells (PAX5, PRDM1), NK cells (DDIT4, HOPX), and T cells (TCF7, GATA3).
- Plasma progenitor cells were first identified and defined, and the rearrangements of IGHV, IGKV, and IGLV genes were already existed and occurred during the progenitor status of plasma cells.
- Hematopoiesis is a continuum procedure controlled by the activation and

suppression of fate decision factors, with a relative identified boundary.

- The erythroid lineage commitment process is a continuous process but the lymphoid lineage commitment process is more likely a discrete stepwise process that showed distinct lineage-committed populations during the hematopoiesis.
- The lineage commitment model of HPCs was reaffirmed that can reconcile previous models and explain the conflict among functional studies.
- Transcriptional atlas of HPCs helps to provide a well explanation of the immune reconstitution in adults and on its cytokines secretion and regulation.
- Application of this atlas can enable artificial precise control progenitor cell differentiation, in vitro accurate culture of each lineage cells, and targeted treatment of immune cell disease to be made possible for future development.

Introduction

Deciphering the hematopoiesis mechanism that underlies progenitor cells definition and fate decision are critical to understanding fundamental questions of hematopoietic lineage commitment. Toward the diversities of peripheral blood cells in the human body, it is like the multifarious stars existing in our galaxy. Although we know their existence, the exact characterization of these cell subsets remains a challenge.

Among the various cell types in peripheral blood (PB), hematopoietic stem cells (HSCs) are one of such rare cell subsets. These cells possess self-renewal capacity and can differentiate to all blood lineage cells. Such that, HSCs reside at the top of a hierarchy and differentiate to all blood lineage cells with tree-like pattern in classical hematopoietic hierarchical models [1]. Single-cell sequencing technique has been successfully implemented to explore molecular heterogeneity and hierarchy of HSCs[2-4]. The lineage commitment is affected by transcription factors (TFs) with a continuum fate decision procedure[5, 6]. Though *in vitro* and *in vivo* studies have reaffirmed their critical role of transcription factors in lineage commitment, the function and regulatory mechanism of HSCs is still unclear [7]. In addition, the transcriptional atlas and the differentiation among subtypes of progenitors in adult peripheral blood remain to be elucidated.

The reason for the molecular mechanisms that controlled lineage commitment and progenitor behaviors remaining an enigma is due to various reasons. Firstly, the definition of hematopoietic progenitor cells (HPCs) is not standardized. Based on the classical and traditional definition, the characterization of HPCs is based on the presence of surface markers, such as CD34, CD38, FLT3 and KIT[5]. Similarly, the combination of surface markers has been proposed for the identification toward common lymphoid progenitors (CLP) and common myeloid progenitors (CMP). These markers include CD10 (MME) and CD38. Yet, owing to the rare and transient nature, the subtypes of CLP and CMP from HSCs have never been stringently identified and accurately captured. Secondly, the functional studies were also used surface markers for the characterization of HPCs. These include the selection of CD34

positive cell from bone marrow or cord blood. However, this approach can be affected by the type of culture methods and surface markers being used for the identification. Simply, the use of classical and traditional markers may be not insufficient, such as KIT, to differentiate mature subset cells. Thus, the use of combined surface markers for HPCs may cause biased and arbitrary.

In recent years, many studies have identified the lineage by using single cell sequencing. However, there are still several limitations of the transcriptional atlas being defined. The scRNA expression data of CD34 positive cells showed two well-defined subtypes of progenitor CLP and CMP being scattered and mixed distribution in T-SNE plots. This result implies that the transcriptional atlas for these types of cells are not distinguishable by unsupervised dimensionality reduction and clustering. When compared with others mature cells, HPCs should exist as a relatively homogeneous progenitor population. The idea classification and defined subset of lineage cells should be separately distinguishable by unsupervised dimensionality reduction which carried the PCA key markers of genes as factors.

In addition, the hematopoietic hierarchical models are also inconsistent among the literature studies [1, 8, 9]. Most of the established lineage fate decision models were based on *in-vitro* / *in-vivo* functional studies or cell fate induces experiments [10, 11]. The use of cellular conditions can often lead to inaccurate interpretations of HPCs cell fate decision because cell lineages were induced by diverse experimental conditions and cannot be permissive to all possible branches of lineage commitment [11]. Similarly, single cell studies would have modified some aspects of classical hematopoiesis hierarchical model, as the hematopoietic progenitors are changing continuously among the progenitors differentiation process [5, 12]. In addition, progenitors cells were defined by classical markers and selected by surface markers [5, 6]. As each type of progenitors is very rare in whole blood, the captured cell numbers are often very low with less than one hundred cells[5]. This implies that some key progenitor subgroups may be missed. Most of the bone marrow HPCs was quiescent status progenitors. scRNA Sequencing of bone marrow progenitors can't clearly depict the whole lineage commitment of hematopoietic progenitors. Thus, this

147 would affect the reliability of results. As the classic definition of hematopoietic
148 progenitors remains inefficient and inaccurate, this can cause the disorganization
149 among the subgroups and without a reliable identification of the boundary among
150 them [2]. Moreover, unsupervised dimensionality reduction and clustering of lineage
151 cells demonstrated farraginous model [4]. The measurement of lineage definition, fate
152 decision, and establishing the reliable lineage commitment between them are still a
153 challenge [8, 13].

154 Herein, we redefined the adult hematopoietic progenitors, and recovered the key
155 transcript factors in lineages fate decision, and reaffirmed the lineage commitment
156 model of HPCs by single-cell sequencing. Our HPCs atlas defined the cellular biology
157 of differentia stage progenitors among hematopoietic lineages generation. Our study
158 can help to open new perspectives for the characterization, definition, and
159 classification of HPCs.

160

161

162 **Results**

163 ***Single-cell identified two subtypes of progenitor cells in adults peripheral blood.***

164 In order to enrich most of the subset of HPCs and later progenitors, such as CLP
165 and CMP, lymphocyte and granulocyte were removed by surface marker panel
166 (Lineage−) and progenitor cells were enriched by negative deletion (Fig. 1A). 10X
167 genomics scRNA was conducted and totally 79,655 cells were captured after quality
168 filters of mitochondrial RNA, low gene expression cells and erythrocyte cell.
169 Unsupervised dimensionality reduction and clustering analysis showed 17 clusters by
170 Seurat and UMAP (Fig. 1B). Total 3180 CD34 positive cells were identified in the
171 dataset. Besides, samples from all donors were highly reproducible for cell
172 composition. The HPCs marker of CD34 mainly expressed in two nearby clusters 6
173 and 14 with the co-expression of FLT3 and KIT (Fig. 1C).

174 To confirm whether the two clusters are subset of progenitor cells, the different
175 expression of lymphocyte and granulocyte surface markers and known HSC related
176 genes were compared (Fig. 1C). Both of the two clusters were found to express KIT

177 and FLT3, but do not contained the expression of lymphocyte and granulocyte surface
178 markers. Based on this, the two clusters can then be well distinguishable by three
179 markers, defined with CD34+CD38+MME- and CD34+CD38-MME+ (Fig. 1C).
180 However, a portion of the cells can positively expressed CD38 or MME in the two
181 subgroups, respectively. KIT (CD117) and FLT3 were mainly expressed in one of the
182 CD34+ clusters, with low positive percentages but also high expressed in the other
183 clusters. This indicates that the progenitors of CD34+KIT and CD34+FLT3 belonging
184 to two different clusters which may control the two differentiation directions.

185 We further analyzed the significantly identified genes among the two clusters
186 with the other clusters. As found, we presented 16 progenitors related genes which
187 expressed in more than 80% cells in the two clusters (Fig. 1C). Several HSCs key
188 regulate genes were identified, such as high expression of *CDK6, SERPINB1, SPINT2,*
189 *IMPDH2, TXN, SOX4, MYADM, HMGA1, CDK2AP1, CD82, TSC22D1, LYL1*, and
190 middle expression of SPI1[6], IFITM3, HMGB1, and PPIA. Based on the surface
191 markers and the above hematopoietic stem cells functional key gens, we can confirm
192 that the two identified clusters are hematopoietic progenitor cells. We also identified
193 three new progenitor key genes of *PPIB, HMGA1,* and *EGFL7*. It is very obviously
194 that the gene of EGFL7 are very high and specific expressed in the two progenitors
195 clusters, so that EGFL7 is a key functional gene in hematopoietic progenitor cells
196 lineage differentiation and mature (Fig. 1D). The single cell sequencing identified two
197 subgroups which are progenitor cells with category and functional heterogeneity.

198

199 ***The differentiation of transcript markers among progenitors***

200 Our single cell expression atlas level showed the two clusters of myeloid
201 progenitors (CD38+) and lymphoid progenitors (MME+) presented heterogeneity in
202 these progenitors, suggesting that functional of the two-cluster subset are diversity. To
203 determine the diversity, re-clustering of analysis was conducted and 11 clusters was
204 generated (Fig 2A). Based on our finding, the subgroup of cells aggregated in two
205 main directions. One is the cluster 0 called lymphoid progenitors, which showed the
206 up-regulated expression of MME. The other is myeloid progenitors comprised with

207 cluster 1, cluster 2, and cluster 5, with the gradual up-regulation of CD38 and KIT
208 (Fig 2B). The three genes expression demonstrated contrary among the direction of
209 myeloid progenitors and lymphoid progenitors. But the expression is not whole
210 positive of the two direction cell types. This implies that CD38 and MME can be used
211 as the markers distinguished myeloid progenitors and lymphoid progenitor
212 differentiation directions, but not suitable to use in subtypes classification.

213 The efficient and accurately of classical hematopoietic progenitors markers which
214 be used to distinguished CMP, GMP (Granulocyte-monocyte progenitors) and CLPs,
215 were also re-analyzed. Our data demonstrated the expression of THY1 (CD90) is very
216 rare (only 17 cells) in the two progenitor clusters which may cause as the progenitors
217 were isolated from adult PBMC. To confirm whether the traditional classification
218 markers THY1 and CD49f (ITGA6) expressed in progenitor, we analyzed four bone
219 marrow single cell projects data of GSE120221, GSE117498, GSE145531 and
220 GSE181989 (Fig. S1). Total 8296 CD34 positive cells were extracted and merged of
221 the four projects data. The results showed positive expression cell numbers of THY1
222 is only 84 and CD49f is 824, and the distribution is inconsistent. This is similar with
223 the other identified markers of FLT3, CD38 and KIT (Fig. S1). The different
224 progenitor types of cells (CMP or CLP) were not distinguishable by unsupervised
225 dimensionality reduction and clustering in bone marrow CD34+ cells. In addition, our
226 data showed that the use of CD49f, THY1, FLT3 and KIT to classify or define the
227 progenitor cells groups were very artificial and subjective. With the progenitors of
228 CD34+FLT3+ cells defined a mixed population of progenitors, this finding may help
229 to provide a well explanation that previously functional study of selected progenitors
230 (such as CD34+FLT3+) presented multiple differential potential.

231

232 ***The essential regulate genes of HPCs lineages differentiation***

233 To confirm the essential genes in HPCs status maintain and differentiation. We
234 compared the expressions among three single cells data sheets, including bone
235 marrow, PBMC and our CD34- cells subsets. The genes of SERPINB1, CDK6,
236 HMGA1, IMPDH2, SPINT2, TXN, and EGFL7 were highly expressed in bone

237 marrow and in our identified progenitor cells when compared with mature PBMC
238 cells (Fig S2A). Those genes also high expressed in most of the subgroups of our
239 CD34+ cells. This indicates those genes are essential to regulating genes in HPCs
240 status for maintenance and differentiation.
241
242 ***Classifications of the lineages differentiation***
243 The transcriptomes among membrane marker-defined and transcriptomes-defined
244 cells were thoroughly compared. Traditionally, membrane markers of progenitor CMP
245 were defined with CD38+MME-FLT3+, while megakaryocytic–erythroid progenitors
246 (MEP) were defined with CD38+MME-FLT3-. However, single cell clusters based on
247 these markers were not distinguishable among the subgroups. As the expressions of
248 CD38 and FLT3 were contradictory, this demonstrated FLT3 positive cells were
249 negative of CD38 expression in most clusters (Fig 2B). With the CD38-FLT3+
250 subgroups progenitors were more original than the FLT3 negative cells, they are more
251 likely to be CMPs. The expression is changed gradually and presented a continuum
252 procedure. Although the expression tendency of FLT3+ was consistent with
253 previously definition, it remained inefficient and inaccurate to classify the
254 differentiation stages among the CMPs lineages. On the other hand, our result
255 demonstrated clearly that CMPs lineages clusters can be distinguishable with the
256 expression of KLF1, GATA1 and CSF3R (Fig 2B).
257 Our results showed KLF1 and GATA1 were uniquely expressed in the three
258 clusters of 3, 8 and 9. This indicated that three clusters are likely to be either
259 megakaryocyte progenitors (MkPs) or erythrocyte progenitors (ErP) of MEPs
260 progenitors. The expression of KLF1 and GATA1 only presented in the late stage of
261 CMPs directions, which indicated those two genes are the keys regulated genes in the
262 differentiation from CMPs to MkPs or ErP. The other gene of GATA2 was widely
263 expressed in most of the others clusters but mainly highly expressed in CMPs clusters.
264 Interestingly, the expression of CSF3R disappeared in the MEPs or MkPs clusters
265 (Fig 2B). During the lineage differentiation of CMPs, GATA2 play a very important
266 role in all types of myeloid cell lines. The procedure of MEPs or MkPs differentiation

267 was managed by up-regulated KLF1 and GATA1, and down-regulated with CSF3R.
268 Hence, the cluster 5 was MEPs would act as the key progenitor to control the
269 differentiation to MkPs or ErP progenitors.

270 As AVP was known to be uniquely expressed in HSC, our data showed both of
271 AVP and CSF3R were mainly expressed in cluster 1 and 2 (Fig 2C and 2D). The high
272 expression of CSF3R was consistent with G-CSF mobilization in peripheral blood
273 HSC transplantation. Those clusters were the original multipotent clusters named as
274 HPCs, which can prime to lineages of CLPs and CMPs. MPO was uniquely expressed
275 in neutrophils in the border of cluster 1 (Fig S2B and S2C). This indicated that the
276 granulocyte-monocyte progenitor branches were also present in the CMPs clusters.
277 Our results showed the expression of IGHM, MME, and HOXA9[14] was presented
278 in cluster 0, which confirmed the cluster was lymphoid progenitor cells (Fig 2C). The
279 expression level is increasing gradually of the transcription factors HOXA9 from
280 cluster 1 to cluster 0 (Fig 2C).

281 Based on traditional surface markers, our data showed the classification of the
282 progenitor subgroups or stages among lineage differentiation were inefficient and
283 inaccurate. Importantly, this demonstrated that the expression and activation of
284 transcription factors combined with unique marker genes would be more efficient for
285 lineage definition of HPCs.

286

287 ***Differentiation and enrichment analysis***

288 Differentiation and enrichment analysis was conducted among the clusters. As the
289 expression changed gradually and presented in a continuum procedure, simple
290 analysis on the different genes among clusters, especially when compared with nearby
291 clusters, may not be efficient. To overcome this, we conducted the differentiation
292 analysis among cluster 0 (CLPs) and cluster 3 (MEPs). Both of the two clusters are
293 original clusters of CLPs or MEPs. The biology procedure analysis of Gene Ontology
294 showed the highly expression genes of CLPs were enriched in the regulation of
295 leukocyte and T cell activation (Fig. 2E). While for CMPs, the genes expressions were
296 high enriched in mononuclear cell differentiation, erythrocyte differentiation and

homeostasis (Fig. 2F). This was consistent with our previous results on CLPs and CMPs progenitors differentiation to two directions with strictly lineage affiliated by unsupervised dimensionality reduction and clustering.

Dynamics of progenitors cells subpopulations

It is clearly visible that CD38 and MME subgroups predominantly occupied the two directions of the progenitor differentiation, respectively. To confirm the trajectory change of the cells, trajectory analysis was conducted and three directions of progenitor cells differentiation were presented. Along the trajectory of right direction (State 1) which was specific expressed MME, only the cluster 0 and 4 were distributed (Fig. 3A and 3B). In contrast, the left direction (State 3) was composed by the CD38 positive expression clusters, such as cluster 1, 2, 3 and 5 (Fig. 3A and 3B) (Fig. S3). The expression tendency of MME and FLT3 was consistent with CLP differentiation, with contrary of KIT and CD38 which expressed with CMP lineages.

Heatmap analysis showed the expression tendency of selected marker genes in CMPs and CLPs. The CLPs branch highly expressed lymphoid marker genes of MME, IGHM and FLT3 (Fig. 3C). The CMPs branch was highly expressed with myeloid marker genes of GATA1, KLF1 and CD38. The expression of MME and CD38 were clearly inverse during the CLP and CMP lineages priming procedures (Fig. 3D). The last branch was highly expressed with NKG7, CD68 and FCGR3A (CD16), which are the markers of NK cells, monocyte and macrophages progenitor cells. This indicated lineage priming of these three types of cells may be distinct with CLPs and CMPs. Our study showed CD16 ~~and CD14~~ were widely positive expression in monocyte progenitor cells. In addition, the expression of transcription factors was distinguished among CLPs and CMPs directions. The expression of KLF1 and GATA1 only appeared in the later stage of myeloid progenitors named MEPs. This indicated that the HPCs lineages priming is likely controlled by key transcription factors. This may also indicated that the unperturbed adult hematopoiesis depends on replenishment from multiple potential progenitors.

327

328 ***Transcription factors promote the progenitor lineage priming with key receptors***

329 To identify and characterize the committed cellular status, the critical
330 co-expression of transcription factors for lineage fates were re-analyzed. Our study
331 showed the CMP multipotent progenitors, which co-express with transcription factors
332 GATA2, can differentiate into the lineage of myeloid and megakaryocyte cells (Fig.
333 2B). The expression of KLF1 was only present in the later stage of MEPs status.
334 Furthermore, the expression of CSF3R was mainly showed in the original progenitors
335 and almost disappeared in the later stage of MEPs progenitor. The cytokine receptors
336 will co-express or co-silence with the relevant lineage key regulators, which indicated
337 the progenitor lineages priming was multiple genes regulate procedure.

338 Likewise, the transcription factors regulation by SCENIC analysis showed the
339 expression of transcription factors was completely distributed differently among CLP
340 and CMP progenitors priming directions (Fig. 4A). The activation and expression of
341 key regulators was much different among the priming stage of progenitors, such as
342 GATA1 (Fig. 4B). As shown in our analysis, the transcription factors of BCL11A,
343 YY1 and REL were the key regulators in CLP priming. And the KLF1, GATA1, and
344 GATA2 were found to play a critical role during the lineage procedure of CMPs (Fig.
345 4C, Fig. S4).

346 To confirm which cluster was a megakaryocyte progenitor, the analysis on the
347 expression of marker genes was performed. When compared with the key different
348 genes and transcription factors among cluster 8 and cluster 9, the expression of KLF1
349 and GATA1 were gradually up-regulated in cluster 9 (Fig. 4A). The expression of
350 KLF1 and GATA1 were up-regulated when the progenitor cells differentiate to
351 megakaryocyte and erythrocyte lineages (Fig. S5). Our results also showed the
352 expression of KIT, MS4A2 and TPSB2 were up-regulated or unique expression in
353 cluster 8, indicating that the cluster was consisted of mast cell and basophil
354 progenitors (Fig. S5).

355

356 ***G-CSF mobilization various type of progenitors in peripheral blood***

357 Various types of progenitors were shown to exist in the PBMC of adults in the
358 quiescent status. To examine whether any different among the mobilization and
359 quiescent (unperturbed) progenitors, single cell sequencing was conducted in the
360 G-CSF mobilization peripheral blood progenitors. A total of 159603 cells were
361 captured and 27566 CD34+ cells were enriched and sequenced, among which 17
362 clusters subgroups were showed after re-clustering (Fig. 5A, B). Our results indicated
363 that the total number of all types of progenitors were remarkable increased. All kinds
364 of subtype progenitors were shown to be existed in the quiescent status when
365 compared before and after mobilization (Fig. 5C). This signifies that all types of
366 progenitors were existed in the quiescent status of peripheral blood.

367 Since the progenitor was shown to gradually differentiate to mature stages in the
368 peripheral blood, it would be expected fold changes of the more primitive progenitors
369 to be higher than mature progenitors. Based on our analysis, the highest cell numbers
370 and folder changed among the clusters were found to be in clusters 0 and 3 (Fig. 5D).
371 The distribution and characterization of the featured markers were consistent with
372 previously CD34 subgroups (Fig. 5E, F). Additionally, the AVP was expressed in
373 cluster 0 which was consistent with the analyzed results in the quiescent status of the
374 progenitor cells. Hence, cluster 0 was the most original progenitor in the peripheral
375 blood that can be identified as HPCs. In contrast, the expression of THY1 (CD49f)
376 were only identified in very few cells. Gene expression analysis also demonstrated
377 five genes, including *AVP*, *SOX4*, CDK6, *C1QTNF4*, and *JUN*, were high expressed
378 in cluster 0 (Fig. 5G).

379 The progenitors were clustering into several directions which also consistent with
380 the cluster distributions of quiescent status (Fig. 5B and Fig. 5E). One new direction
381 of cluster 5 was presented in the all cells merged populations (Fig. 5F). The markers
382 genes of that cluster were MPO and CSF3R. As myeloperoxidase (MPO) is the most
383 abundant protein in human neutrophil and plays an important role in neutrophil
384 antimicrobial responses[15, 16]. The other marker gene PRSS57 of neutrophils was
385 also presented and up-regulated expression in the cluster[17]. So the cluster 5 was the

386 progenitor of neutrophils, this was shown to be in a completely different direction
387 with megakaryocyte or erythrocyte progenitors. Based on this analysis, it further
388 showed NK cells progenitors and monocyte progenitor cells were not clustering
389 nearby the other lineages (Fig. 5B, G). Most of the top five specific marker genes for
390 each lineage progenitor were known marker genes of classically defined hemopoietic
391 cells, indicating the reliable characterization of subgroups (Fig. 5H). Genes set
392 enrichment analysis of clusters by GSVA revealed the pathways and signaling that
393 affect the differentiation of each branch and progenitor subgroup (Fig5 I).
394 In summary, there are three clusters which nearby the original progenitor of HPCs.
395 It is clearly shown that during the HPCs priming and differentiation procedures, the
396 HPCs priming into three progenitor cell types coinstantaneous and later maturing to
397 diversity hematopoietic lineages. The lymphoid progenitors, neutrophils progenitors,
398 and megakaryocytic–erythroid progenitors were presented at the same time point and
399 stage during the HPCs first step of priming process.
400
401
402 ***Re-definition of hematopoietic progenitor cells (HPCs)***
403 As previously shown, cluster 0 was the most original progenitor in the peripheral
404 blood that can be identified and defined as HPCs. All of the later mature progenitors
405 are found to be developed from those original cells. Based on this rationale, we
406 re-defined the three directions with common lymphoid progenitors (CLPs), GAPs
407 (GATA genes controlled progenitors), and neutrophilic and monocyte progenitors
408 (NMPs), which cannot be characterized by CD38, FLT3, and KIT (Fig. 6A, Table 1).
409 Most of the unique and specific marker genes and transcription factors were identified
410 in re-defined subgroups (Table 1). The transcriptional atlas between cluster 3 and
411 cluster 0 is very similar. The expression of unique genes gradually accumulated
412 during the differentiation procedure, such as GATA2 and SLC40A1. That provided
413 evidence of hematopoiesis is a continuum procedure controlled by the activation and
414 suppression of fate decision genes. The differentiation of GAPs lineages was
415 independently with the other two branches. The shape of each branch was controlled

416 by the expression of key regulated genes which was shown to function like a
417 gradually activate switch. Through the transcription factors expression switches
418 gradually on-or-off and further gradually accumulated, generally synergistic with key
419 receptor genes expression and signals activation, the original progenitor cells will be
420 primed and differentiated to various progenitors.

421 During the myeloid progenitor differentiation procedure, the expression of
422 GATA2 and CSF3R genes is the first branch being formed and presented (Fig. 6).
423 This was shown to be followed by the down-regulation of CSF3R genes and
424 up-regulation of GATA2, which implies the progenitors were priming to oligo-potent
425 progenitor GAPs. The expression of GATA1 and KLF1 were presented and gradually
426 up-regulated at later stage (Fig 5E). This subgroup was defined as GAPs mature stage
427 named with megakaryocytic–erythroid progenitors (MEPs). With the differentiation
428 of MEPs enhancement, the second branch was presented in the mature stage of MEPs.
429 Toward the final stages of the MEPs, the expression of KIT, CDK4, GATA1, , KLF1
430 or LMO4, controlled the differentiation of progenitor to megakaryocyte progenitors,
431 erythrocyte progenitors (Pro-ME), or mast cell and basophil progenitors (MBPs). In
432 contrast to the other branch of myeloid progenitor differentiation process, the
433 expression of GATA2 and KLF1 were restrained during differentiation trajectories
434 when the progenitors were priming to neutrophilic and monocyte progenitors (NMPs,
435 Fig 6A, B). In the lymphoid progenitor branch priming, critical transcription factor of
436 HOXA9 and HOPX were identified and was shown to gradually up-regulated during
437 the process of CLPs lineage differentiation (Fig. 6B). Similarly, the marker genes
438 IGHM and MME of lymphocyte were also found to be up-regulated. This result
439 provided strong evidence that the whole peripheral blood progenitors include a
440 gradient of heterogeneity hematopoietic progenitors. The progenitor cells fate
441 decisions and lineage commitment were mainly controlled by a combination of
442 critical TFs and key receptor genes (Fig. 6). Also, the hierarchical organization of
443 hematopoiesis was a continuum procedure, with key gene expressions presented and
444 gradually being regulated. After the switch genes were activated and up-regulated, the
445 progenitor cells would then differentiated into the next stages of priming. During this

446 process, the decision of cell fate will be controlled and divided into different lineage
447 branches. The fate deciding boundaries were then identified among the differentiation
448 stages. The expression of GATA2 and CSF3R genes was shown to be the boundary
449 for GAPs and NMPs. This finding is the first to validate that hematopoiesis was
450 controlled by a switch-like model with identified relatively boundary, opening a new
451 perspective on hematopoiesis (Fig. 6C). Interestingly, the erythroid lineage process is
452 a continuous process but the lymphoid lineage process is more likely a discrete
453 stepwise process that showed distinct lineage-committed populations during the
454 hematopoiesis.
455
456 ***Cell Atlas and characterization of all re-defined subtype progenitors cells***
457 To clearly define the subgroups of progenitors hematopoietic, the characterization
458 of all re-defined subtype progenitors cells were re-analyzed. Each of subgroups
459 presents specific highly expression genes or negative express genes. These activated
460 genes may enrich in some signal pathways which play a critical role in the lineages
461 differentiation.
462
463 ***Pan markers of HPCs and unique makers of each subgroups***
464 Our study showed the classical markers of CD49f and THY1 were not found in
465 the multipotent HPCs progenitors. In contrast, ITGA4 (CD49d) was highly expressed
466 in most of the progenitors, which is key molecular regulation of HSC homing[18].
467 Some reported HSC functional related marker genes was broadly expressed in most of
468 the progenitor cells, such as PTPRC, HMGB1, CD74, CD37, PFN1, TXN, ZFP36L2,
469 SERPINB1, TMSB4X, SMIM24, and HSP90AA1 (Table 1 and Fig. 1B)[19-25].
470 These genes were shown to mediate the self-renewal, differentiation, mobilization,
471 and maintenance of hematopoietic progenitor cells. In addition, two new genes of
472 PPIA and EIF1 were identified. These genes were similarly to previous reported genes
473 that pan-expressed in progenitors and highly expressed in matured subgroups of
474 peripheral blood cells. A few progenitors' specific genes of CDK6, SOX4, CD82,
475 EGFL7, IMPDH2, TSC22D1, CYTL1, FOXP1, SPI1, and ETV6 were also widely

476 identified in most progenitors[26-32] (Table 1). And ERG, BCL11A, LMO2, and
477 RUNX1 were also widely expressed in half percentages of cells in each progenitor
478 subgroup (Fig. S6). Interestingly, some transcription factors of them showed different
479 expression tendency in specific progenitor subgroups, such as SPI1 for monocyte
480 progenitors and ETV6 for eosinophil progenitors. Among these specific genes,
481 EGFL7 and CYTL1 were identified as being pan-expressed in HPCs, CLPs and MEPs
482 (Table 1). As most of these genes were confirmed by functional validation, our data
483 demonstrated that those high expression genes of CD34 progenitor subsets were
484 essential genes in progenitor status and differentiation.

485 Several key factors were identified, including NFE2, LYL1, MYB, CDK4, ETV6,
486 TESPA1, ZBTB16, and GATA2 of myeloid progenitors and RUNX2, HOXA9,
487 JUN/FOS, TCF4, DDIT4, HOPX, KLF10, HOXA3, and TSC22D1 of lymphoid
488 progenitors (Fig. 6B). Our data revealed RUNX1 monitored the differentiation of
489 most progenitors (Fig. S6), while RUNX2 mainly controlled the differentiation of
490 lymphoid progenitors. Importantly, these key factors were shown to play reversely
491 functional role during the differentiation of the HPCs. Through the up-regulated
492 combination of NFE2, LYL1, MYB, CDK4, ETV6, TESPA1, GATA2 and in
493 restraining the expression of RUNX2, HOXA9, JUN/FOS,TCF4, DDIT4, HOPX,
494 KLF10, HOXA3, and TSC22D1, the HPCs will diverse to differentiate to myeloid
495 progenitors. Together with the activation of other receptors and unique transcript
496 factors in each lineages progenitors, such as CSF3R for neutrophils and GATA2 for
497 GAPs, the whole hematopoietic lineage commitment was precisely controlled and
498 gradually formed (Table 1, Fig. 6C).

499 Our result also showed the selected specific marker genes and key transcript
500 factors of each lineage progenitors (Table 1, Fig 5. G, H, Fig6. B). Additionally, the
501 negatively expressed genes were listed to distinguish each subgroup and clearly
502 define the lineages progenitors (Table 1). It was identified that SLC40A1 and
503 CNRIP1 were unique new markers and key regulator of the whole lineages during
504 GAPs differentiation. The CNRIP1 gene is highly potential a key neurotransmitter
505 related receptor, which activate related signaling pathways and plays role during

506 HPCs differentiation. Our studies showed TAL1 only expressed in MEPs progenitors,
507 which is the later stages of GAPs. LYL1 was up-regulated in the lineages of GAPs but
508 decreased in the lymphoid lineages. Our studies further confirmed that those two
509 genes were critical regulators during myeloid lineages differentiation.
510
511 ***Fate decision genes of each redefined subgroups***
512 Advanced analysis was conducted to clearly re-define the progenitor subgroups
513 and identify fate decision genes, particularly transcription factors. Unique and specific
514 marker genes were used to define each progenitor subgroup (Fig 7.A), while negative
515 expression genes helped distinguish similar or identical original lineages subgroups
516 (Table 1). Six genes of SPINK2, HOPX, C1QTNF4, CSF3R, NPDC1, and SELL were
517 highly expressed in the original progenitors of HPCs, but gradually restrain during
518 differentiation from GAPs to MEPs (Fig 7.B). CSF3R, S100A10, and C1QTNF4 were
519 also gradually restrained during differentiation of lymphoid early progenitors (C4
520 cluster, Fig 7.B). Interestingly, we identified a few genes which are completely
521 contrary and specifically expressed among erythrocyte lineages (C2,C6,13) and other
522 progenitors, such as SPINK2, HOPX, C1QTNF4, and FCER1A, GATA1,
523 SLC40A1(Fig 7.A, B). This indicated the fate decision among erythrocyte lineages
524 (C2, C6, 13) and other progenitors may be controlled by those genes during the
525 lineage commitment. Two class A basic helixloop-helix protein genes of TAL1 and
526 LYL1 were highly expressed among erythrocyte lineages. HBD was the early protein
527 in erythrocyte cell development. mini-chromosome maintenance proteins gene of
528 MCM2, MCM6, MCM5, MCM4, MCM3, and MCM7 which acts as component of
529 the MCM2-7 complex, were highly expressed in erythrocyte progenitors(Pro-ME) and
530 NMPs (Table 1). MCM complex is the replicative helicase that may play an
531 essential role in erythrocyte and neutrophilic progenitors maturation. LMO4 is the
532 most specific TF gene of MBPs.
533 Three key TFs of GATA3, TCF7, and DDIT4 were identified in T progenitor cells
534 (Fig 7.C), especially the two novel expression genes of GATA3 and TCF7. These
535 three TFs play a critical role in the fate decision of T cells lineages differentiation

536 (Table 1, Fig 6B). Previous studies have reported TCF7 and GATA3 to be important
537 TFs in T cell development and differentiation[33, 34], which is highly consistent with
538 our redefined characterization of T progenitor cells. LTB was widely expressed in
539 Pro-TBN, Pro-T, and Pro-B progenitor cells, with an up-regulation in the later stages
540 of lymphoid progenitor differentiation. Two distinct types of B progenitor cells were
541 identified and redefined as having different fate decision TFs and marker genes in C9
542 and C12 clusters. This indicates the development and differentiation of B cells occur
543 early on in the progenitor stages. Cluster 9 is the progenitor of plasma cells, whereas
544 cluster 12 represents classical B progenitor cells (Table 1). Interestingly,
545 immunoglobulin variable genes were mainly expressed in the progenitors of Cluster 9
546 cells, and rearrangements of IGHV, IGKV and IGLV genes had already occurred
547 during the progenitor status of plasma cells. It's widely accepted that plasma cell are
548 the final stage of B cell proliferation and are unable to proliferate. However, our
549 single cell data confirmed the existence of plasma progenitor cells in whole blood.
550 The two factors PRDM1 and IRF4 were specifically expressed in cluster 9, while
551 PAX5 was specifically expressed in cluster 12. XBP1 was widely expressed in both
552 clusters and up-graduated in cluster 9. These four TFs factors play a critical role in the
553 fate decision of B lineage cells. HOPX was found to be very highly expressed in the
554 progenitor and mature NK cells but negative in the erythrocyte lineages (C2, C6, 13),
555 which was one of the most important TF for NK cells (Fig 6B, Fig 7A, Table 1).

556 Two types of monocyte progenitor cell in C7 and C14 clusters were identified and
557 redefined. One of the most specially expressed genes in monocyte lineages is IFI30,
558 which is related to interferon gamma signaling pathway. SPI1 and LST1, were widely
559 expressed in most progenitors but were highly expressed in the monocyte
560 subgroups(C7), which were also highly expressed in mature states of monocyte
561 original cells, such as macrophages and dendritic cells(DCs, Data not showed). This
562 suggests that SPI1 and LST1 are critical genes for monocyte progenitors. The
563 monocyte progenitors of C7 and C14 were controlled by various TFs including KLF4,
564 EGR1, MAFB, IRF8, and DDIT4 (Table 1, Fig 7B). The two clusters of C7 and C14
565 were CD86 and CD68 positive. C7 highly expressed EGR1, CEBPB, and MAFB, but

566 were negative for IRF8 and DDIT4. C7 was FCGR3A positive, which indicates it is
567 more likely to be macrophage progenitors, and CSF1R was novel and highly
568 expressed in that group. The cluster C7 preferred differentiation to macrophages
569 which expressed the secreted cytokine of TNF. The monocyte progenitor of C14 was
570 positively expressed CLEC9A, ANPEP, and THBD (Table 1), which indicates it is
571 more likely to differentiate into DCs. These TFs and receptors were the main fate
572 decision factors that affect the differentiation of monocyte progenitor lineages.
573 Although we have identified the fate decision factors of B cell and monocyte
574 progenitor, the association between these cell types is still unknown in the HPCs
575 differentiation procedure. Future investigation is required to determine whether the
576 two types of progenitor cells derive from the same or independent original cells. It's
577 widely accepted that neutrophilic and monocyte progenitors originate from the same
578 myeloid progenitors. However, our results showed that fate decision TFs and
579 characterization genes among the two types of progenitors were highly different,
580 which reminds us that the original and association of these two types of lineages
581 should be reconsider.

582 *PECAM1* (CD31) was widely expressed in more than 50% of progenitors and
583 CST3 was widely expressed in more than 60% of progenitors, both of them are
584 mainly high expression in monocyte progenitors. Previous studies defined endothelial
585 progenitor cell by the markers of CD31, CD34, and THBD (CD141). Our study
586 showed one cluster (C14) cells positively expressed those markers which were
587 re-defined and confirmed as monocyte progenitor (DCs) (Table 1, Fig. 7A). That
588 reminds us the classical definition of endothelial progenitor cell should be
589 reconsideration.

590

591

592 ***Pathways of each subgroups***

593 Diversity pathways affect and control the differentiation of HPCs. Our current
594 analyses results showed the JUN/FOS (AP-1) and JUND were enriched and highly
595 expressed in most of the progenitors (Fig. 5 and 5F). This indicated that the JUN

596 related signaling pathway may play a critical role in differentiation, mobilization, and
597 maintenance of hematopoietic progenitor cells. The DEPTOR gene was found to be
598 highly expressed in the clusters of MEPs (Fig. S6). As DEPTOR was a negative
599 regulator of the mTORC1 and mTORC2 signaling pathways [35]. That meaning the
600 mTOR signaling pathway should be suppressed in the MEPs mature and
601 differentiation. NFKBIA was widely expressed in more than 80% of the progenitor
602 cells. The expression of NFKBIZ was very high expressed in the progenitors of
603 monocyte and macrophages (Fig. S6), which is a member of the NF-kappa-B inhibitor
604 family [36]. This meanings the NFKB signaling pathway was suppressed in
605 progenitors, especially in the monocyte and macrophages mature and differentiation.

606 All enriched gene sets were shown in figure 5I by GSVA analysis. TGF-b
607 signaling was enriched in the original progenitor cells (C0, HPCs), which plays a key
608 role in maintaining the quiescent state of HPCs[37]. Our study showed that the Notch,
609 KRAS, and TNF-α signaling pathways were enriched and may play a crucial role in
610 maintaining monocyte progenitors (Fig. 5I). Other studies have indicated the novel
611 role of Notch in regulating monocyte survival during hematopoiesis[38]. Heme
612 metabolism and hedgehog pathway gene sets were enriched in erythroid progenitors.
613 The activation of TGF-β and hedgehog pathway gradually decreased during the
614 mature stage of erythroid progenitors. Hallmark cell proliferation-related gene sets
615 (G2M Checkpoint, E2F Targets, MYC Targets V1 and V2, p53 pathway) defined by
616 MSigDB were highly enriched among erythroid and neutrophilic progenitors, but the
617 TNF signaling gene set were restrained (Fig. 5I). The Wnt beta catenin signaling was
618 enriched in NK and T cell progenitors, which was reported to regulate aspects of
619 peripheral T-cell activation and migration[39]. Interferon response gene sets were
620 enriched in B cell and monocyte progenitors. It is clear that the enriched gene sets
621 among myeloid progenitors (C2, C5, C6, C13) and lymphoid progenitors were
622 different.

623

624

625 ***Functional validation and primary culture***

626 Previously studies already conducted many functional validations of HSCs key
627 regulated genes. We hypothesis that our HPCs definitions and hematopoiesis model
628 were right, that meanings the functional studies phenotype results would be consistent
629 or can be explanation by our reaffirmed model. And furthermore, those conflicting
630 and confusing results can be clearly interpretation by the reaffirmed hematopoiesis
631 hierarchical organization model.

632 Our analysis results were found to be consistent with other functional studies
633 results. KLF1 and GATA1 were well-known play a key role in the erythroid and
634 megakaryocytic differentiation[40, 41]. The expression of GATA1 progenitors
635 generated mast cells, eosinophils, megakaryocytes and erythroid cells, and lacking
636 expression can drive generated monocytes, neutrophils, and lymphocytes[40]. Our
637 results revealed that GATA1 was unique expressed in the later stages of MEPs and
638 negatively expressed in the monocytes, neutrophils and lymphocytes progenitors (Fig.
639 5E, F). The result implies that the key transcript factor of GATA1 was very critical for
640 the MEPs progenitor lineages differentiation. Other evidence is transcript factor NFE2.
641 It was reported that NFE2 acts as a critical regulator in hematopoietic cells. The
642 knockout of NFE2 in mouse model displayed mild erythroid abnormalities and major
643 defect in megakaryocyte biogenesis[42, 43]. Likewise, our results revealed NFE2 was
644 up-regulated in the NMPs and GAPs lineage directions (Fig. 6B).

645 Based on findings from primary culture, M-CSF was shown to be critical for the
646 survival of macrophage, which is consistent with our results that the receptor CSF1R
647 of M-CSF only expression in monocyte progenitors and matured macrophages (Fig 7.
648 D). In addition, SCGF, also known as CLEC11A, is a secreted glycoprotein that
649 exhibits erythroid burst (BFU-E) promoting activity in the absence of SCF/c-kit
650 Ligand[44, 45]. Our results showed stem cell growth factor (SCGF) was highly
651 expression in NMP and GAP progenitors, especially in the whole lineages of MEPs
652 that can explain SCGF burst promote erythroid activity. Furthermore, SCGF
653 cooperates with GM-CSF, or G-CSF to enhance the growth of
654 granulocyte-macrophage colonies (CFU-GM)[44, 45]. Because the CSF3R receptor of
655 G-CSF was highly expression in progenitors of myeloid, SCGF cooperates with

656 G-CSF would activate and induced the progenitor differentiation to neutrophils and
657 macrophages but reduce MEPs lineages generation. The IL15 receptor of IL2RB was
658 novel expressed in NK progenitors (Fig 7. D). The common receptor beta chain
659 CSF2RB of IL-3, IL-5, and GM-CSF, was novel expressed in erythroid lineage
660 progenitors. IL7R was novel expressed in T cell progenitors which is an essential
661 driver of T cell development and homeostasis. IL2RG and IL1B were expressed
662 across most progenitor subgroups (Fig 7. D). Our study indicated that cytokine IL1B
663 plays a critical role in hematopoiesis. The transcription of IL1B was regulated by TFs
664 of SPI1 and CEBPB[46], which exhibit high expression in monocytes progenitors
665 (Table 1). ALDH1A1 was highly expressed in erythroid lineage progenitors, and its
666 product participates in retinoic acid synthesis. It's widely reported that retinoic acid is
667 required for progenitor cell maintenance [47]. The expression characterization of
668 these receptors and genes was highly consistent with previous studies, providing
669 strong evidence that our re-definition of these progenitors was reliable and reasonable.
670
671
672 **Discussion**
673 Taken together, we demonstrated the transcription atlas data of original status
674 HPCs in adult peripheral blood with large number of HPCs. We redefined the
675 characterization of progenitor lineages and identified specific TFs and marker genes
676 that are critical for their development during hematopoiesis. Our transcription atlas
677 data provided novel insights into the fate decision factors and characterization genes
678 of hematopoietic progenitor cells. Our study further reaffirmed and established a
679 comprehensive model of hematopoiesis that can reconcile previous models and
680 explain conflicts among functional studies. This helps to provide a detailed insight
681 into lineage commitment from HPCs into almost branches of human peripheral blood
682 progenitors.
683
684 Currently, there is without a standard consensus for the definition of all HPCs
685 lineages. This is owing to the inconsistent classification with conflict functional

686 outcomes among the current literatures [9, 13]. Our study succeed enriched the large
687 numbers and almost whole subgroups of HPCs. This will be particularly attractive for
688 large single cells studies to establish the comprehensive model of hematopoiesis with
689 clearly HPCs definition and reliable lineage commitment. All of the hematopoietic
690 progenitors were reaffirmed by unique makers and key factor genes, which can be
691 well distinguishable. We are the first clearly definition all lineages of HPCs by
692 unsupervised and inartificial clustering, which provide the unambiguous transcription
693 atlas of each type of HPCs lineages. The unique markers genes of progenitor
694 subgroups were highly consistent with previously characterization descriptions, such
695 as KLF1 of MEPs, HPGDS of mast cell, and MPO of neutrophils[15, 48, 49].
696
697 It is widely known that the HPCs would differentiate into common myeloid
698 progenitors (CMPs) and common lymphoid progenitors (CLPs). Our data showed
699 none of the MME positive cells co-expressed CD38, suggesting that classification of
700 CLPs using solely CD38+ and MME+ was not sufficient. CMPs are defined with
701 CD34+CD38+MME- which is consistent with classic definition, but our single cell
702 expression atlas showed the CD34+CD38-MME+ progenitor cells is inconsistent with
703 the traditional define of CLPs[4, 5]. Based on our study, we identified two subgroups
704 of progenitors by single cell sequencing which are not consistent with classical
705 definition, and showed specific classify surface or characterization markers genes,
706 such as GATA1, MPO and KLF1. Likewise, we provided a clearly evidence that the
707 classical surface markers of CD38, KIT, THY1, CD49f and FLT3 were non-specific
708 and inefficient to identify HSCs. This remind us the traditional lineage definition of
709 HSC by above markers was imperfect and need to be improved. Evidently, the
710 combination of transcriptional factors with unique marker genes would be more
711 efficient for lineage definition of HPCs.
712
713 Our classification and definition of HPCs also help to explain the conflict and
714 mechanism of *in vitro* and *in vivo* progenitor cell studies. It is well known that KLF1
715 and GATA1 are the keys regulate genes of erythrocyte or megakaryocyte

716 differentiation[50, 51]. Functional study have been shown that the upregulated
717 transcriptional factor of GATA-1 can affect mast cells, eosinophils, megakaryocytes
718 and erythroid cells generation, and downregulation to affect monocytes, neutrophils
719 and lymphocytes generation[40]. Our finding reaffirmed the definition of GAPs
720 lineages to show GATA1 to be solely expressed in the MEPs and up-regulated during
721 maturation. Previously studies reported TAL1 and LYL1 are essential for maintenance
722 of adult HSC function [52-54]. Since the induced pluripotent stem cells (iPSCs) was
723 established through activation of four transcript factors, reprogram of pluripotent stem
724 cells was widely investigated in various somatic cells[55]. GATA1 and TAL1 have
725 been validated in the reprogram of megakaryocytes and erythroid cells[50, 56].
726 Although they finally induced mature erythroid cells based on those key transcript
727 factors, but there are still many functional capacity limited during differentiation.
728
729 The CSF3R was mainly expressed in the original clusters (HPCs), which was the
730 receptor of G-CSF [57]. The underlying molecular mechanisms of G-CSF induces
731 enhancement of hematopoietic remain unclear. Our results provide a new model of
732 hematopoietic through scRNA of unperturbed and G-CSF mobilized progenitors
733 lineages in peripheral blood. All types of the progenitors were remarkable increased,
734 especially the HPCs. G-CSF directly activates the downstream activation signals of
735 HPCs through bind with CSF3R, which converted the quiescent status HPCs to
736 activation progenitors. After the mobilization procedures, HPCs was priming and
737 differentiation to all kinds of progenitors which can differentiate to mature myeloid
738 and lymphocyte cells. As key transcriptional factors and regulators controlled the
739 whole hematopoietic lineage commitment formation, the use of primary cultures also
740 showed factors of G-CSF and SCGF to induce the lineages formation. Both of these
741 mechanisms would cooperate to play key roles during the lineage commitment
742 formation. That reminds us two questions: First, how did the key transcript factors
743 was activated during the lineage commitment procedure? Secondly, the receptors
744 activated the signaling cascade and transcript factors, or the transcript factors were
745 activated independently and factors enhance the commitment formation?

746

747 In brief, the unique markers were highly consistent with characterization of most
748 progenitor subgroups. Secondly, the definition can very well explanation the conflict
749 of functional studies. Thirdly, the expression of distinguished receptor of cytokines
750 can perfectly explain the primary culture phenomenon of progenitor subgroups, such
751 as CSF3R for neutrophilic and CSF1R for monocyte culture[58]. These three aspects
752 strongly support the definition and classification of HPCs to be distinctly verifiable,
753 reliable and reasonable.

754

755 Moreover, our transcriptional atlas revealed the major role of each progenitor
756 subgroups in hematopoiesis and the fate decision of downstream various
757 hematopoietic cells. Through the expression changes of the three key transcription
758 factors of KIT, KLF1 and GATA1, the later stage of MEPs can differentiate into two
759 directions. The expression of transcriptional factors, both in original and mature stage
760 of progenitors, revealed a basic split between myeloid-lymphocyte development and
761 megakaryocyte-erythroid development. This provides a strong evidence that the
762 transcription factors were the key regulators in each stages of the progenitor's
763 differentiation. Generally, progenitor cells can be priming by key transcription factor
764 activation, which derived progenitor differentiated to downstream specific mature
765 lineages[59]. But previously study showed that transcriptional programs were not
766 strictly lineage affiliated, and expression changes between HSCs and early
767 progenitors occur independently of lineage choice[60]. That may cause by not reliable
768 classification and definition of the subgroups and different stage of progenitor among
769 lineage commitment, which based on the traditional cell surface markers in most of
770 those investigations. We demonstrated the unique markers or transcriptional factors
771 which control the progenitors differentiate to all kind of blood lineages in a
772 continuum differentiation procedure. Our results strongly supported that in an ideally
773 cell fate decision model, the activation of transcription factors and expression of
774 progenitor lineage unique markers demonstrated a consistent lineage affiliated with
775 relatively obvious subgroups boundary.

776

777 Our analysis on the transcriptional atlas revealed that the transcriptional factors
778 activation and suppression play a critical role in the differentiation among lineage
779 commitment. The bias during HPCs differentiation was gradually accumulated by
780 activation of multiple lineage-specifying TFs, such as NFE2, LYL1, MYB, and
781 GATA2 of myeloid and RUNX2, HOXA9, and HOPX of lymphoid progenitors.
782 These TFs play critical role during the progenitor priming with dynamic entail
783 quantitative expression and abundances change. By the multiple TFs activation and
784 suppression, the downstream signal pathway would be activated. The stepwise
785 controlled the differentiation of each branch during HPCs lineage commitment.
786 During the process, individual progenitors would gradually acquire lineage biases
787 along multiple directions during homeostasis. Our atlas results showed that key
788 transcriptional factors priming HPCs for distinct downstream lineages, which finally
789 generated and formatted the whole lineage commitment. The expression of key TFs
790 could be considered as the relative reliable boundary among stages during progenitor
791 mature. Our findings have highlighted the importance of TFs and characterization
792 genes in hematopoietic lineage fate decision, providing additional insights into the
793 molecular mechanisms underlying the hematopoietic lineage commitment. These
794 transcription factors hold great potential for reprograming progenitor cell. Through
795 the control of TFs genes, the artificial precise committed progenitor cell
796 differentiation can be made possible.

797

798 It was reported the HSC lineage commitment is a continuous process in recently
799 single cell sequencing studies [5, 12]. But HSC heterogeneity and unclearly definition
800 of HSC subgroups caused the mechanism of hierarchy formation still an enigma. The
801 established hematopoietic lineage models were inconsistent and incompatible. This
802 was due to the limitation of captured cell numbers and inefficient definition by
803 identification using surface markers. Based on our clear definition and atlas of HPCs,
804 it revealed that a gradient of hematopoietic progenitors existed during the lineage
805 commitment and can be distinguished by the expression of unique TFs. Those

806 progenitors gradually acquired lineage biases, divergence and prime to each branch.
807 Our elaborate and comprehensive transcriptional atlas of HPCs can provide accurate
808 definition, hierarchy, and biological insights of hematopoietic progenitors. As our
809 study further explained how the diversified progenitor cells were primed, induced,
810 and differentiated to the whole lineages of immune system cells, this enable the
811 transcriptional atlas of HPCs to provide new perspective and well explanation about
812 how the immune was reconstituted in adult. Recently, cytokines storm was reported in
813 severe infection and coronavirus[61, 62]. Our atlas can clearly demonstrate the
814 expression cytokines receptor in specific progenitor. By controlling the original stage
815 and generation of cytokines, especially suppress the initial cytokines activation of the
816 cascade activation, this could inhibit or reduce the effect of the follow-up cascade
817 activation. Technically, this would be helpful in controlling the excessive activation
818 by cytokines and preventing the occurrence of cytokines storm as a new therapeutic
819 strategy. Our reaffirmed taxonomy will enhance more accurate prevention and
820 treatment as well as immune monitoring in health and disease.
821
822 There were some limitations in our current study. Firstly, our study did not define
823 the original progenitors with HSCs. Even though the characterization of these
824 progenitors was most likely originated with stem cells, the markers of CD49f and
825 THY1 were found to be negatively expressed. Our study showed two types of
826 progenitor cells that are not clustering nearby the other lineages. One is the NK cells
827 progenitors and the other is monocyte progenitors. Monocyte progenitors have more
828 than two types of status progenitors, which located far away from the main cell
829 populations. We speculated two possibilities to cause this phenomenon. One may be
830 from the limitation of our enrichment method. This can cause the loss of major status
831 progenitors from these two types of cells. This will have an effect to cause only the
832 capture from the later stages. Secondly, the original source of NK cells progenitors
833 and the monocyte progenitors were found to be independent and distinct with all of
834 other progenitor cell types. As the whole atlas of our definition was established by
835 unsupervised dimensionality reduction and clustering from statistical approach, the

836 accuracy and efficiency should be considered. Artificial intelligence and deep learning
837 would provide a more accurate, reliable, and distinguishable definition for each status
838 and all progenitor subgroups during the whole procedure of hematopoiesis.
839
840 Importantly, our study helped to re-defined hematopoietic progenitor cells and
841 reaffirming lineage commitment by transcriptional atlas. We are the first to clearly
842 depict the transcriptional atlas and defined each subgroups of adult peripheral blood
843 progenitor. Our results will enable the whole map of the transcriptional signal and
844 pathway among lineage commitment. Moreover, we are also the first to identify the
845 cell fate decision transcriptional factors in each branch among hematopoietic
846 hierarchy formation. The precise control on the hematopoietic progenitor cells
847 priming and differentiation can then be interfere by the key transcriptional factors and
848 receptors signals, which we call Artificial Operation of HPCs.
849
850 In summary, our transcriptional atlas of HPCs established a comprehensive
851 continuum and key regulators controlled model of the human adult peripheral blood
852 hematopoiesis. Toward this transcriptional atlas, it provides accurate definition,
853 hierarchy, and biological insights of hematopoiesis and lineage commitment of HPCs.
854 Importantly, this can enable new research development in cell engineering,
855 regenerative medicine, and disease treatment. Also, this can have critical insight and
856 implications for the etiology and treatment of hematologic disorders. Moreover, this
857 opens up a wide variety of possibilities for artificial precise control progenitor cells
858 differentiation and targeted treatment of immune disease, providing hope for
859 untreatable diseases.
860

Acknowledgments

We acknowledge Professor Chen Qiu, Rongchang Chen, and Yingyun Fu for supporting this research. We also thank Lipeng Chen for bioinformatics and statistical analysis of sequencing data, Qijun Huang and Guoqiang Li for sample collection, and Sujin Chen for draw the schematic diagram. We also thank ChiWai Man of CUHK and Qing Sun of Shenzhen Zhongshan Urology Hospital for manuscript revise. Moreover, we are grateful to the subjects for their participation in this study.

Author Contributions

Y.Y conceived the original idea, designed the study, and performed experiments on HPC's enrichment.Y.Y conducted the bioinformatics and statistical analysis of sequencing data.Y.Y wrote the manuscript and revised the manuscript. All authors approved the final draft of the manuscript.

Declaration of Interests

The authors have declared no competing financial interest.

Declaration

The correspondence author YJ Yue would like to clarify that we already identified most marker genes, receptors, and transcript factor genes of each redefined subgroups based on the transcriptional atlas, which we didn't and can't show all of them in that version. Based on the redefined marker genes, all of the other receptors and fate decision factors can be easily identify by our transcriptional atlas of progenitors. We will present these results and reference data in the future.

Funding

The correspondence author Dr. Yongjian Yue did not even get one coin supported from any funding or organizations in the study.

892 **Data Availability Statement**

893 The datasets used or analyzed during the current study are available from the
894 corresponding author on reasonable request.

895

896 **Methods**

897 **Sample collection and HPCs enrichment**

898 Totally, 8mL peripheral blood was draw from each donor. Prior to blood collection,
899 mobilizated donors will received subcutaneous recombinant human granulocyte
900 colony-stimulating factor with 300ug in each 12 hours for five days. After five days of
901 continuous treatment, a total of 10mL of peripheral blood was collected in heparinized
902 tubes from the stem cell donor, and kept at room temperature. PBMC and total white
903 blood cells were obtained according to standard protocols. The protocol of peripheral
904 blood mononuclear cell (PBMC) isolation was based on separation by density
905 gradient centrifugation with Ficoll® (Cytiva, GE), a highly branched polysaccharide.
906 Lineages (Lin−, CD3-CD19-CD56-) negative cells were removed by STEMCELL
907 reagents (15272). Self-design method was applied to enrich the HPCs. Ethical
908 approval was obtained from the ethics committee of the Shenzhen People's Hospital,
909 as per the ethical standards recommended in the 1964 Declaration of Helsinki.
910 Informed consent was obtained from the subjects in the study.

911

912

913 **Single cell RNA preparation and sequencing**

914 RNA-sequencing was also performed on human enriched HPCs (Lin−) purified from
915 10 donors, including one mobilization PBMC progenitor cells by G-CSF. The single
916 cells suspension was loaded and captured by Chromium microfluidic chips and
917 barcoded with a 10× Chromium Controller (10X Genomics). Barcoded single cells
918 RNA was subsequently reverse-transcribed and sequencing libraries constructed with
919 reagents from a Chromium Next GEM Single Cell 5' reagent kits v2 according to the
920 manufacturer's instructions (10X Genomics). Sequencing was conducted by Illumina
921 Nova 6000 (PE150).

922

Single-cell RNA-seq data processing

The raw sequencing reads of all samples were demultiplexed and aligned to the reference genome hg38 by Cellranger pipeline. All downstream single-cell unsupervised dimensionality reduction and clustering was conducted by Seurat[63]. Gene Ontology (GO) enrichment analysis of marker genes was implemented by the cluster Profiler R package. Trajectory analysis was conducted by Monocle2 [64]. Transcription factors regulation inference and gene regulatory networks analyses were performed using SCENIC [65]. Genes set variation enrichment analysis was conducted by a non-parametric, unsupervised method of GSVA[66].

Integrated analysis with online single cell dataset

The online scRNA data of GSE120221[67], GSE117498[4], GSE145531[68], and GSE181989[69] was downloaded from GSE database. Integrated analysis was performed by Seurat.

References

1. Cheng H, Zheng Z, Cheng T: **New paradigms on hematopoietic stem cell differentiation**. *Protein & cell* 2020, **11**(1):34-44.
2. Xie X, Liu M, Zhang Y, Wang B, Zhu C, Wang C, Li Q, Huo Y, Guo J, Xu C *et al*: **Single-cell transcriptomic landscape of human blood cells**. *Natl Sci Rev* 2021, **8**(3):nwaa180.
3. Ranzoni AM, Tangherloni A, Berest I, Riva SG, Myers B, Strzelecka PM, Xu J, Panada E, Mohorianu I, Zaugg JB *et al*: **Integrative Single-Cell RNA-Seq and ATAC-Seq Analysis of Human Developmental Hematopoiesis**. *Cell stem cell* 2021, **28**(3):472-487 e477.
4. Pellin D, Loperfido M, Baricordi C, Wolock SL, Montepeloso A, Weinberg OK, Biffi A, Klein AM, Biasco L: **A comprehensive single cell transcriptional landscape of human hematopoietic progenitors**. *Nature communications* 2019, **10**(1):2395.
5. Velten L, Haas SF, Raffel S, Blaszkiewicz S, Islam S, Hennig BP, Hirche C, Lutz C, Buss EC, Nowak D *et al*: **Human haematopoietic stem cell lineage commitment is a continuous process**. *Nat Cell Biol* 2017, **19**(4):271-281.
6. Sugimura R, Jha DK, Han A, Soria-Valles C, da Rocha EL, Lu YF, Goettel JA, Serrao E, Rowe RG, Malleshaiah M *et al*: **Haematopoietic stem and progenitor cells from human pluripotent stem cells**. *Nature* 2017, **545**(7655):432-438.
7. Zhu J, Emerson SG: **Hematopoietic cytokines, transcription factors and lineage commitment**. *Oncogene* 2002, **21**(21):3295-3313.
8. Jacobsen SEW, Nerlov C: **Haematopoiesis in the era of advanced single-cell technologies**. *Nat Cell Biol* 2019, **21**(1):2-8.

959 9. Haas S, Trumpp A, Milsom MD: **Causes and Consequences of Hematopoietic Stem Cell**
960 **Heterogeneity**. *Cell stem cell* 2018, **22**(5):627-638.
961 10. Manz MG, Miyamoto T, Akashi K, Weissman IL: **Prospective isolation of human clonogenic**
962 **common myeloid progenitors**. *Proceedings of the National Academy of Sciences of the United*
963 *States of America* 2002, **99**(18):11872-11877.
964 11. Akashi K, Traver D, Miyamoto T, Weissman IL: **A clonogenic common myeloid progenitor that**
965 **gives rise to all myeloid lineages**. *Nature* 2000, **404**(6774):193-197.
966 12. Buenrostro JD, Corces MR, Lareau CA, Wu B, Schep AN, Aryee MJ, Majeti R, Chang HY,
967 Greenleaf WJ: **Integrated Single-Cell Analysis Maps the Continuous Regulatory Landscape of**
968 **Human Hematopoietic Differentiation**. *Cell* 2018, **173**(6):1535-1548 e1516.
969 13. Laurenti E, Gottgens B: **From haematopoietic stem cells to complex differentiation**
970 **landscapes**. *Nature* 2018, **553**(7689):418-426.
971 14. Huang Y, Sitwala K, Bronstein J, Sanders D, Dandekar M, Collins C, Robertson G, MacDonald J,
972 Cezard T, Bilenky M *et al*: **Identification and characterization of Hoxa9 binding sites in**
973 **hematopoietic cells**. *Blood* 2012, **119**(2):388-398.
974 15. Metzler KD, Fuchs TA, Nauseef WM, Reumaux D, Roesler J, Schulze I, Wahn V,
975 Papayannopoulos V, Zychlinsky A: **Myeloperoxidase is required for neutrophil extracellular**
976 **trap formation: implications for innate immunity**. *Blood* 2011, **117**(3):953-959.
977 16. Schultz J, Kaminker K: **Myeloperoxidase of the leucocyte of normal human blood. I. Content**
978 **and localization**. *Archives of Biochemistry and Biophysics* 1962, **96**(3):465-467.
979 17. Perera NC, Wiesmuller KH, Larsen MT, Schacher B, Eickholz P, Borregaard N, Jenne DE: **NSP4**
980 **is stored in azurophil granules and released by activated neutrophils as active**
981 **endoprotease with restricted specificity**. *Journal of immunology* 2013, **191**(5):2700-2707.
982 18. Li D, Xue W, Li M, Dong M, Wang J, Wang X, Li X, Chen K, Zhang W, Wu S *et al*: **VCAM-1(+)**
983 **macrophages guide the homing of HSPCs to a vascular niche**. *Nature* 2018,
984 **564**(7734):119-124.
985 19. Becker-Herman S, Rozenberg M, Hillel-Karniel C, Gil-Yarom N, Kramer MP, Barak A, Sever L,
986 David K, Radomir L, Lewinsky H *et al*: **CD74 is a regulator of hematopoietic stem cell**
987 **maintenance**. *PLoS biology* 2021, **19**(3):e3001121.
988 20. Sundaramoorthy P, Wang Q, Zheng Z, Jiao Y, Chen BJ, Doan PL, Chao NJ, Kang Y: **Thioredoxin**
989 **mitigates radiation-induced hematopoietic stem cell injury in mice**. *Stem cell research &*
990 *therapy* 2017, **8**(1):263.
991 21. Zheng J, Lu Z, Kocabas F, Bottcher RT, Costell M, Kang X, Liu X, Deberardinis RJ, Wang Q, Chen
992 GQ *et al*: **Profilin 1 is essential for retention and metabolism of mouse hematopoietic stem**
993 **cells in bone marrow**. *Blood* 2014, **123**(7):992-1001.
994 22. Zhang L, Prak L, Rayon-Estrada V, Thiru P, Flygare J, Lim B, Lodish HF: **ZFP36L2 is required for**
995 **self-renewal of early burst-forming unit erythroid progenitors**. *Nature* 2013,
996 **499**(7456):92-96.
997 23. Benarafa C, LeCuyer TE, Baumann M, Stolley JM, Cremona TP, Remold-O'Donnell E: **SerpinB1**
998 **protects the mature neutrophil reserve in the bone marrow**. *Journal of leukocyte biology*
999 2011, **90**(1):21-29.
1000 24. Altmann S, Lange S, Pommerencke J, Murua Escobar H, Bullerdiek J, Nolte I, Freund M,
1001 Junghanss C: **High Mobility Group Box 1-Protein expression in canine haematopoietic cells**
1002 **and influence on canine peripheral blood mononuclear cell proliferative activity**. *Veterinary*

1003 *immunology and immunopathology* 2008, **126**(3-4):367-372.
1004 25. Kruta M, Sunshine MJ, Chua BA, Fu Y, Chawla A, Dillingham CH, Hidalgo San Jose L, De Jong B,
1005 Zhou FJ, Signer RAJ: **Hsf1 promotes hematopoietic stem cell fitness and proteostasis in**
1006 **response to ex vivo culture stress and aging**. *Cell stem cell* 2021, **28**(11):1950-1965 e1956.
1007 26. Saito-Reis CA, Balise VD, Pascetti EM, Jiminez M, Gillette JM: **Tetraspanin CD82 regulates**
1008 **S1PR1-mediated hematopoietic stem and progenitor cell mobilization**. *Stem cell reports*
1009 2021, **16**(10):2422-2431.
1010 27. Karigane D, Kobayashi H, Morikawa T, Ootomo Y, Sakai M, Nagamatsu G, Kubota Y, Goda N,
1011 Matsumoto M, Nishimura EK *et al*: **p38alpha Activates Purine Metabolism to Initiate**
1012 **Hematopoietic Stem/Progenitor Cell Cycling in Response to Stress**. *Cell stem cell* 2016,
1013 **19**(2):192-204.
1014 28. Scheicher R, Hoelbl-Kovacic A, Bellutti F, Tigan AS, Prchal-Murphy M, Heller G,
1015 Schneckenleithner C, Salazar-Roa M, Zochbauer-Muller S, Zuber J *et al*: **CDK6 as a key**
1016 **regulator of hematopoietic and leukemic stem cell activation**. *Blood* 2015, **125**(1):90-101.
1017 29. Doulatov S, Vo LT, Chou SS, Kim PG, Arora N, Li H, Hadland BK, Bernstein ID, Collins JJ, Zon LI
1018 *et al*: **Induction of multipotential hematopoietic progenitors from human pluripotent stem**
1019 **cells via respecification of lineage-restricted precursors**. *Cell stem cell* 2013, **13**(4):459-470.
1020 30. Calvanese V, Capellera-Garcia S, Ma F, Fares I, Liebscher S, Ng ES, Ekstrand S,
1021 Aguade-Gorgorio J, Vavilina A, Lefaudeux D *et al*: **Mapping human haematopoietic stem cells**
1022 **from haemogenic endothelium to birth**. *Nature* 2022, **604**(7906):534-540.
1023 31. Naudin C, Hattabi A, Michelet F, Miri-Nezhad A, Benyoucef A, Pflumio F, Guillonneau F,
1024 Fichelson S, Vigon I, Dusanter-Fourt I *et al*: **PUMILIO/FOXP1 signaling drives expansion of**
1025 **hematopoietic stem/progenitor and leukemia cells**. *Blood* 2017, **129**(18):2493-2506.
1026 32. Yu J, Ershler M, Yu L, Wei M, Hackanson B, Yokohama A, Mitsui T, Liu C, Mao H, Liu S *et al*:
1027 **TSC-22 contributes to hematopoietic precursor cell proliferation and repopulation and is**
1028 **epigenetically silenced in large granular lymphocyte leukemia**. *Blood* 2009,
1029 **113**(22):5558-5567.
1030 33. Zhou W, Gao F, Romero-Wolf M, Jo S, Rothenberg EV: **Single-cell deletion analyses show**
1031 **control of pro–T cell developmental speed and pathways by Tcf7, Spi1, Gata3, Bcl11a, Erg,**
1032 **and Bcl11b**. *Science immunology* 2022, **7**(71).
1033 34. van de Wetering M, Oosterwegel M, Dooijes D, Clevers H: **Identification and cloning of TCF-1,**
1034 **a T lymphocyte-specific transcription factor containing a sequence-specific HMG box**. *The*
1035 *EMBO journal* 1991, **10**(1):123-132.
1036 35. Peterson TR, Laplante M, Thoreen CC, Sancak Y, Kang SA, Kuehl WM, Gray NS, Sabatini DM:
1037 **DEPTOR is an mTOR inhibitor frequently overexpressed in multiple myeloma cells and**
1038 **required for their survival**. *Cell* 2009, **137**(5):873-886.
1039 36. Totzke G, Essmann F, Pohlmann S, Lindenblatt C, Janicke RU, Schulze-Osthoff K: **A novel**
1040 **member of the IkappaB family, human IkappaB-zeta, inhibits transactivation of p65 and its**
1041 **DNA binding**. *The Journal of biological chemistry* 2006, **281**(18):12645-12654.
1042 37. Blank U, Karlsson S: **TGF-β signaling in the control of hematopoietic stem cells**. *Blood* 2015,
1043 **125**(23):3542-3550.
1044 38. Ohishi K, Varnum-Finney B, Flowers D, Anasetti C, Myerson D, Bernstein ID: **Monocytes**
1045 **express high amounts of Notch and undergo cytokine specific apoptosis following**
1046 **interaction with the Notch ligand, Delta-1**. *Blood* 2000, **95**(9):2847-2854.

1047 39. Staal FJT, Luis TC, Tiemessen MM: **WNT signalling in the immune system: WNT is spreading**
1048 **its wings**. *Nature Reviews Immunology* 2008, **8**(8):581-593.
1049 40. Drissen R, Buza-Vidas N, Woll P, Thongjuea S, Gambardella A, Giustacchini A, Mancini E, Zriwil
1050 A, Lutteropp M, Grover A *et al*: **Distinct myeloid progenitor-differentiation pathways**
1051 **identified through single-cell RNA sequencing**. *Nature immunology* 2016, **17**(6):666-676.
1052 41. Siatecka M, Bieker JJ: **The multifunctional role of EKLF/KLF1 during erythropoiesis**. *Blood*
1053 2011, **118**(8):2044-2054.
1054 42. erythroid abnormalitiesGasiorek JJ, Blank V: **Regulation and function of the NFE2**
1055 **transcription factor in hematopoietic and non-hematopoietic cells**. *Cellular and molecular*
1056 *life sciences : CMLS* 2015, **72**(12):2323-2335.
1057 43. Lecine P, Villeval JL, Vyas P, Swencki B, Xu Y, Shivdasani RA: **Mice lacking transcription factor**
1058 **NF-E2 provide in vivo validation of the proplatelet model of thrombocytopoiesis and show**
1059 **a platelet production defect that is intrinsic to megakaryocytes**. *Blood* 1998,
1060 **92**(5):1608-1616.
1061 44. Hiraoka A, Yano Ki K, Kagami N, Takeshige K, Mio H, Anazawa H, Sugimoto S: **Stem cell growth**
1062 **factor: in situ hybridization analysis on the gene expression, molecular characterization and**
1063 **in vitro proliferative activity of a recombinant preparation on primitive hematopoietic**
1064 **progenitor cells**. *The hematology journal : the official journal of the European Haematology*
1065 *Association* 2001, **2**(5):307-315.
1066 45. Hiraoka A, Ohkubo T, Fukuda M: **Production of human hematopoietic survival and growth**
1067 **factor by a myeloid leukemia cell line (KPB-M15) and placenta as detected by a monoclonal**
1068 **antibody**. *Cancer research* 1987, **47**(19):5025-5030.
1069 46. Pulugulla SH, Workman R, Rutter NW, Yang Z, Adamik J, Lupish B, Macar DA, el Abdouni S,
1070 Esposito EX, Galson DL *et al*: **A combined computational and experimental approach reveals**
1071 **the structure of a C/EBPβ–Spi1 interaction required for IL1B gene transcription**. *Journal of*
1072 *Biological Chemistry* 2018, **293**(52):19942-19956.
1073 47. Chute JP, Muramoto GG, Whitesides J, Colvin M, Safi R, Chao NJ, McDonnell DP: **Inhibition of**
1074 **aldehyde dehydrogenase and retinoid signaling induces the expansion of human**
1075 **hematopoietic stem cells**. *Proceedings of the National Academy of Sciences of the United*
1076 *States of America* 2006, **103**(31):11707-11712.
1077 48. Jiang J, Faiz A, Berg M, Carpaij OA, Vermeulen CJ, Brouwer S, Hesse L, Teichmann SA, Ten
1078 Hacken NHT, Timens W *et al*: **Gene signatures from scRNA-seq accurately quantify mast cells**
1079 **in biopsies in asthma**. *Clinical and experimental allergy : journal of the British Society for*
1080 *Allergy and Clinical Immunology* 2020, **50**(12):1428-1431.
1081 49. Klimchenko O, Mori M, Distefano A, Langlois T, Larbret F, Lecluse Y, Feraud O, Vainchenker W,
1082 Norol F, Debili N: **A common bipotent progenitor generates the erythroid and**
1083 **megakaryocyte lineages in embryonic stem cell-derived primitive hematopoiesis**. *Blood*
1084 2009, **114**(8):1506-1517.
1085 50. Capellera-Garcia S, Pulecio J, Dhulipala K, Siva K, Rayon-Estrada V, Singbrant S, Sommarin
1086 Mikael NE, Walkley Carl R, Soneji S, Karlsson G *et al*: **Defining the Minimal Factors Required**
1087 **for Erythropoiesis through Direct Lineage Conversion**. *Cell reports* 2016, **15**(11):2550-2562.
1088 51. Wontakal SN, Guo X, Smith C, MacCarthy T, Bresnick EH, Bergman A, Snyder MP, Weissman
1089 SM, Zheng D, Skoultchi AI: **A core erythroid transcriptional network is repressed by a master**
1090 **regulator of myelo-lymphoid differentiation**. *Proceedings of the National Academy of*

1091 *Sciences* 2012, **109**(10):3832-3837.
1092 52. Bergiers I, Andrews T, Vargel Bolukbasi O, Buness A, Janosz E, Lopez-Anguita N, Ganter K,
1093 Kosim K, Celen C, Itir Percin G *et al*: **Single-cell transcriptomics reveals a new dynamical**
1094 **function of transcription factors during embryonic hematopoiesis**. *eLife* 2018, **7**.
1095 53. Souroullas GP, Salmon JM, Sablitzky F, Curtis DJ, Goodell MA: **Adult hematopoietic stem and**
1096 **progenitor cells require either Lyl1 or Scl for survival**. *Cell stem cell* 2009, **4**(2):180-186.
1097 54. Chan WY, Follows GA, Lacaud G, Pimanda JE, Landry JR, Kinston S, Knezevic K, Piltz S,
1098 Donaldson IJ, Gambardella L *et al*: **The paralogous hematopoietic regulators Lyl1 and Scl are**
1099 **coregulated by Ets and GATA factors, but Lyl1 cannot rescue the early Scl-/- phenotype**.
1100 *Blood* 2007, **109**(5):1908-1916.
1101 55. Takahashi K, Yamanaka S: **Induction of Pluripotent Stem Cells from Mouse Embryonic and**
1102 **Adult Fibroblast Cultures by Defined Factors**. *Cell* 2006, **126**(4):663-676.
1103 56. Moreau T, Evans AL, Vasquez L, Tijssen MR, Yan Y, Trotter MW, Howard D, Colzani M,
1104 Arumugam M, Wu WH *et al*: **Large-scale production of megakaryocytes from human**
1105 **pluripotent stem cells by chemically defined forward programming**. *Nature communications*
1106 2016, **7**(1).
1107 57. Beekman R, Touw IP: **G-CSF and its receptor in myeloid malignancy**. *Blood* 2010,
1108 **115**(25):5131-5136.
1109 58. Hume DA, MacDonald KP: **Therapeutic applications of macrophage colony-stimulating**
1110 **factor-1 (CSF-1) and antagonists of CSF-1 receptor (CSF-1R) signaling**. *Blood* 2012,
1111 **119**(8):1810-1820.
1112 59. Hoppe PS, Schwarzfischer M, Loeffler D, Kokkaliaris KD, Hilsenbeck O, Moritz N, Endele M,
1113 Filipczyk A, Gambardella A, Ahmed N *et al*: **Early myeloid lineage choice is not initiated by**
1114 **random PU.1 to GATA1 protein ratios**. *Nature* 2016, **535**(7611):299-302.
1115 60. Laurenti E, Doulatov S, Zandi S, Plumb I, Chen J, April C, Fan JB, Dick JE: **The transcriptional**
1116 **architecture of early human hematopoiesis identifies multilevel control of lymphoid**
1117 **commitment**. *Nature immunology* 2013, **14**(7):756-763.
1118 61. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, Zhang L, Fan G, Xu J, Gu X *et al*: **Clinical features of**
1119 **patients infected with 2019 novel coronavirus in Wuhan, China**. *Lancet* 2020,
1120 **395**(10223):497-506.
1121 62. Fajgenbaum DC, Longo DL, June CH: **Cytokine Storm**. *New England Journal of Medicine* 2020,
1122 **383**(23):2255-2273.
1123 63. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, 3rd, Hao Y, Stoeckius M,
1124 Smibert P, Satija R: **Comprehensive Integration of Single-Cell Data**. *Cell* 2019,
1125 **177**(7):1888-1902 e1821.
1126 64. Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, Lennon NJ, Livak KJ, Mikkelsen
1127 TS, Rinn JL: **The dynamics and regulators of cell fate decisions are revealed by**
1128 **pseudotemporal ordering of single cells**. *Nature biotechnology* 2014, **32**(4):381-386.
1129 65. Aibar S, González-Blas CB, Moerman T, Huynh-Thu VA, Imrichova H, Hulselmans G, Rambow F,
1130 Marine J-C, Geurts P, Aerts J *et al*: **SCENIC: single-cell regulatory network inference and**
1131 **clustering**. *Nature Methods* 2017, **14**(11):1083-1086.
1132 66. Hänzelmann S, Castelo R, Guinney J: **GSVA: gene set variation analysis for microarray and**
1133 **RNA-Seq data**. *BMC Bioinformatics* 2013, **14**(1).
1134 67. Oetjen KA, Lindblad KE, Goswami M, Gui G, Dagur PK, Lai C, Dillon LW, McCoy JP, Hourigan CS:

1135 **Human bone marrow assessment by single-cell RNA sequencing, mass cytometry, and flow**
1136 **cytometry**. *JCI insight* 2018, **3**(23).
1137 68. Zhu C, Lian Y, Wang C, Wu P, Li X, Gao Y, Fan S, Ai L, Fang L, Pan H *et al*: **Single-cell**
1138 **transcriptomics dissects hematopoietic cell destruction and T-cell engagement in aplastic**
1139 **anemia**. *Blood* 2021, **138**(1):23-33.
1140 69. Tonglin H, Yanna Z, Xiaoling Y, Ruilan G, Liming Y: **Single-Cell RNA-Seq of Bone Marrow Cells**
1141 **in Aplastic Anemia**. *Front Genet* 2021, **12**:745483.
1142
1143

1144 **List of Table and Legends**

1145

1146 **Table 1. The selected specific marker genes and fate decision genes (partly) of all**

1147 **re-defined subgroups**

Progenitor Cell Types	Marker genes and fate decision genes (partly)	Transcription factors	Negative
Pan-Progenitors	EIF1,PPIA, HMGB1, CD74, PFN1, TXN, ZFP36L2, CD37, SERPINB1,HSP90AA1, TMSB4X, SMIM24		
Progenitors(Specific)	CD34, CDK6, SOX4, CD49d, CD82, EGFL7, CYTL1, IMPDH2, TSC22D1, SPI1(PU.1), FOXP1, EVT6, XBP1	CDK6, SOX4, IMPDH2, TSC22D1, SPI1(PU.1), EVT6, XBP1	
CLP (C4)	SPINK2, HOPX, HOXA9, RUNX2, LTB, IGHM, DNTT, PRSS2, SLC2A5, MME, CCR7, NKG7, LST1, BASP1, CD79A	HOPX,DDIT4, HOXA9, RUNX2	CNRIP1,FCER1A,GATA1, S100A10
Pro-NK (C8)	GNLY,NKG7, CD247, CCL5, FCGR3A, PRF1, GZMA, GZMB, KLRD1, KLRB1 ,CD3E,CD7. HOPX	DDIT4, HOPX	IL7R,GATA3
Pro-B2 (C9, Pro-Plasma cell)	CD27, IGKC, IGHA1, SLAMF7, CD79A, CD79B, PRDM1, IRF4, CD38, JCHAIN	PRDM1, IRF4	MS4A1, FCER2
Pro-T (C11)	TCF7, IL7R, GATA3, KLRB1, CD3E, CD3D, CD7, CD247, LTB, DDIT4	TCF7,GATA3,DDIT4	GNLY, FCGR3A, GZMA
Pro-B1 (C12)	CD19, MS4A1,FCER2,CD79A, CD79B, IGHM, LTB, IGKC, PAX5, VPREB3	PAX5	CD27
NMPs(C5)	CSF3R, MPO, MGST1, IGLL1, S100A10, C1QTNF4, NPDC1, MYB, CDK4, CDCA7	MYB, CDK4	GATA2, SLC40A1, CNRIP1,LTB
Lineages(C2,C6,C13)	CNRIP1, CPA3 ,FCER1A, GATA1, KLF1, HPGDS, SLC40A1, GATA2, ZBTB16 ,TAL1,	GATA1, ZBTB16,TAL1, CDK4, TESPA1,KLF1	SPINK2,HOPX, C1QTNF4, CSF3R

	TESPA1, MYB, CDK4		
GAPs (C1)	GATA2, NFE2, LYL1, MYB, SLC40A1, TESPA1, CSF3R	GATA2, NFE2, LYL1, MYB	GATA1, KLF1
Pro-ME(C6)	HBD,CDT1, MCM2, MCM6, MCM5, MCM4, MCM3, MCM7, CDCA7, CDK4, TYMS	CDK4	MS4A2, MS4A3
MBPs (C13)	TPSAB1,LMO4, HDC, MS4A2, TPSB2, MS4A3, KIT, HPGDS, PRG2, CLC, MCM2-MCM7	LMO4, CDK4	HBD
Eosinophil (C10)	CLC, HDC, RFLNB, ETV6	ETV6	MS4A2,TPSB2,MS4A3
Pro-Monocytes1 (C7, Macrophages)	EGR1,SPI1,KLF4,CEBPB,FCGR3A, CSF1R,CD68, CD86, ITGAX, FCGR2A , LYZ , LST1, EGR2, CEBPA, MAFB, TNF	SPI1,KLF4,CEBPB, EGR1, EGR2, CEBPA, MAFB	CLEC9A, THBD, IRF8
Pro-Monocytes2 (C14, DCs)	CLEC9A, ANPEP, THBD, IRF8, KLF4, CD68, CD86, ITGAX, LYZ, SPI1 , LST1, DDIT4, SLAMF7	SPI1,KLF4,IRF8,DDIT4	FCGR3A, CSF1R, MAFB

Graphical abstract

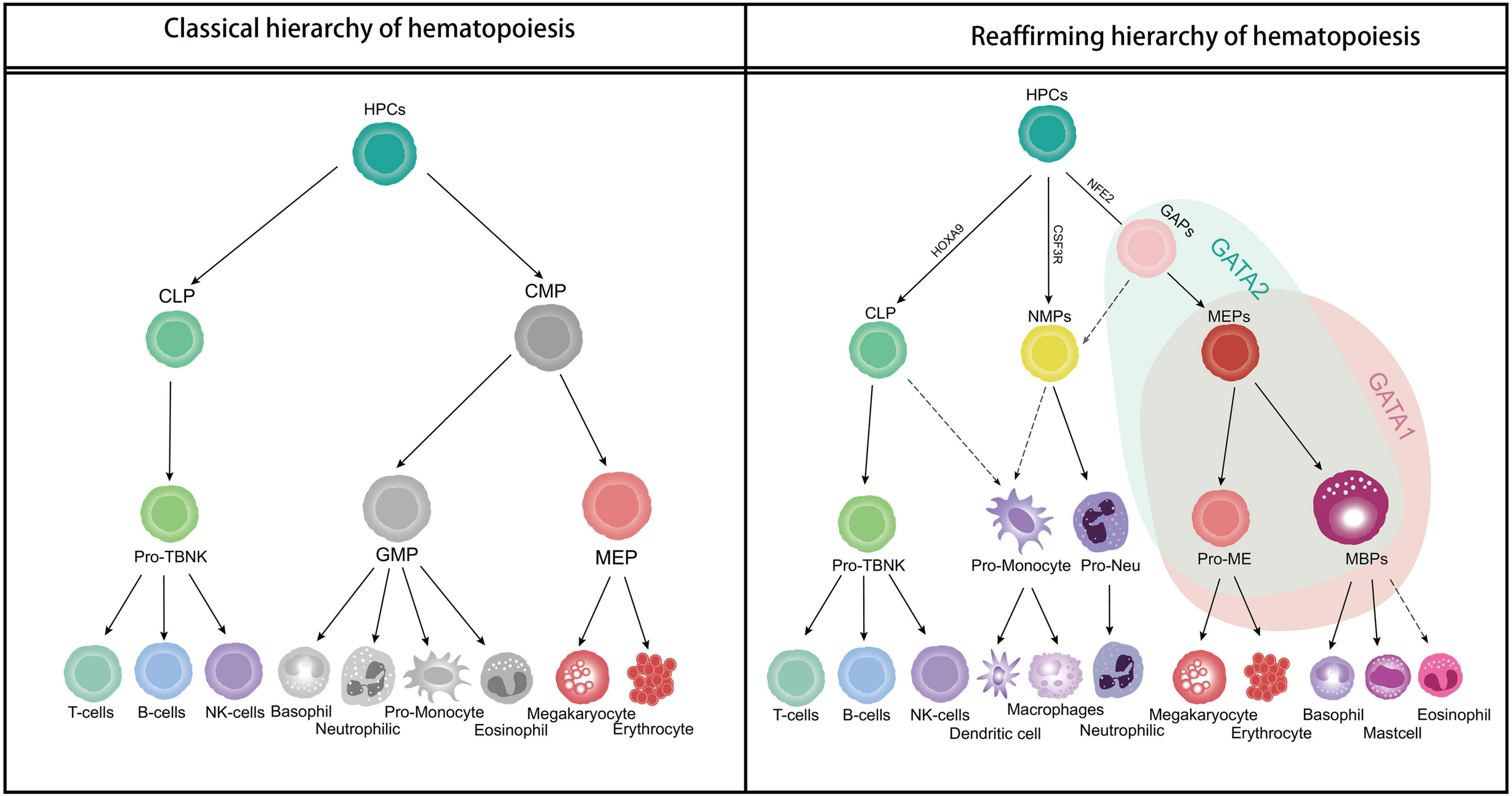
The reaffirmed lineage commitment model of HPCs

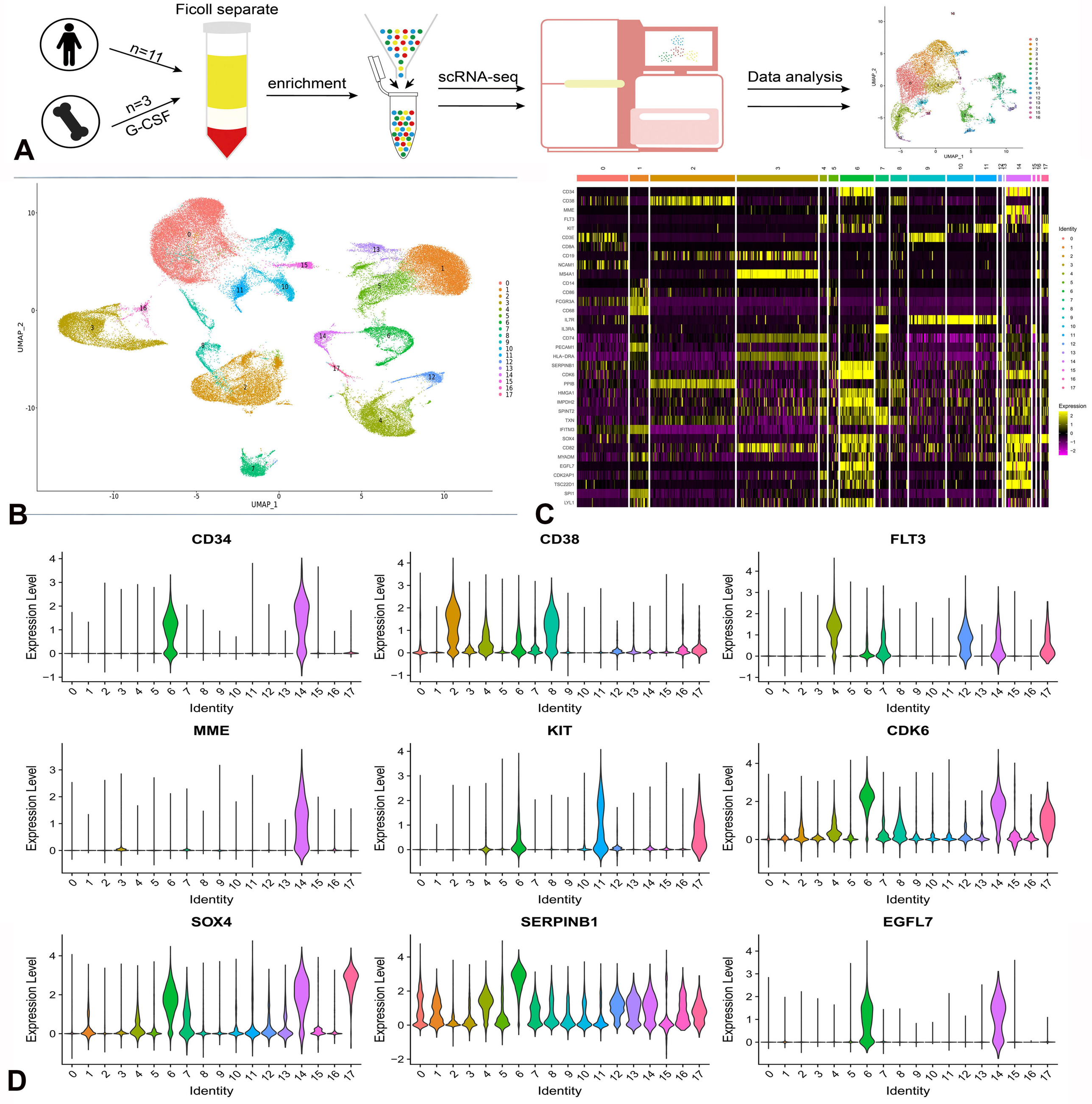
Figure 1. Two distinct subtypes of progenitors that were identified by unsupervised dimensionality reduction and clustering analysis of single cell data. A, Schematic showing the progenitor cell enrichment and sequencing workflow. B,17 clusters were identified by UMAP analysis. The progenitors marker of CD34 is mainly expressed in two adjacent clusters 6 and 14. C, The heatmap results illustrated the expression of different surface markers and known progenitors cell related markers in all clusters. Progenitor cells key regulate genes were identified, such as high expression of CDK6, SERPINB1, and SOX4. D, Violin plot presented the

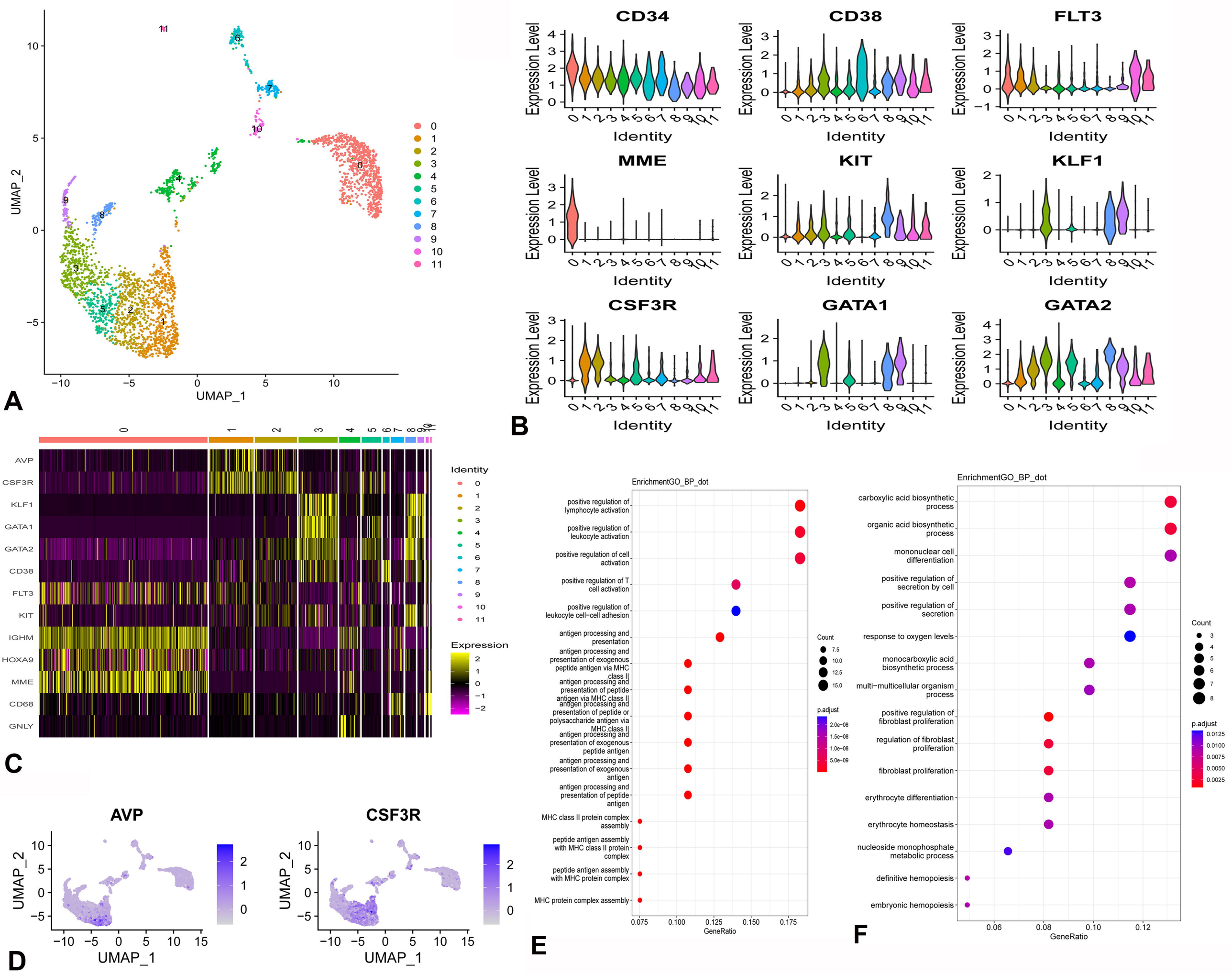
1162 expression of HSCs traditional classification marker genes and known progenitor cell
1163 marker genes.
1164
1165 **Figure 2. The transcriptome differentiation and characterization among**
1166 **re-clustered CD34 positive subgroups of progenitor cells.** A, UMAP analysis
1167 reveals the aggregation of CD34 positive subgroup cells in two mainly distinguished
1168 directions. B, Violin plots presented the expression of unique marker genes, which
1169 demonstrated GATA2 expression of common myeloid progenitors (CMPs) and MME
1170 expression of common lymphoid progenitors (CLPs) characterization. C, The
1171 heatmap results displayed the different expression marker genes, which confirmed
1172 that the two distinguished directions belong to different lineage cells of CLPs and
1173 CMPs. D, Feature plot of AVP and CSF3R expression indicating cluster 1 was the
1174 most original progenitors in all of CD34 positive cells. E, The biological process
1175 analysis of Gene Ontology, indicating that the highly expressed genes of CLPs were
1176 enriched in the regulation of leukocyte and T cell activation. F, CMPs mainly high
1177 expression genes were enriched in mononuclear cell differentiation, erythrocyte
1178 differentiation and homeostasis.
1179
1180 **Figure 3. Dynamics of myeloid progenitor and lymphoid progenitor cell**
1181 **transcriptomes.** A, Trajectory analysis illustrated all cell clusters distributed in three
1182 branches. B, The expression tendency of unique markers genes was completely
1183 different among the two lineages of myeloid and lymphoid progenitor. MME and
1184 FLT3 were up-regulated with CLP differentiation branch, contrary to KIT and GATA2
1185 up-regulated in CMP lineages. C, Heatmap analysis displayed the expression
1186 tendency of selected marker genes in CMP and CLP. D, The expression of MME and
1187 CD38 showed an inverse correlation during the CLP and CMP lineages priming
1188 procedures.
1189
1190 **Figure 4. Different transcription factors promote the progenitor lineage priming.**
1191 A, Heatmap results depicted the different expression of transcription factors among all

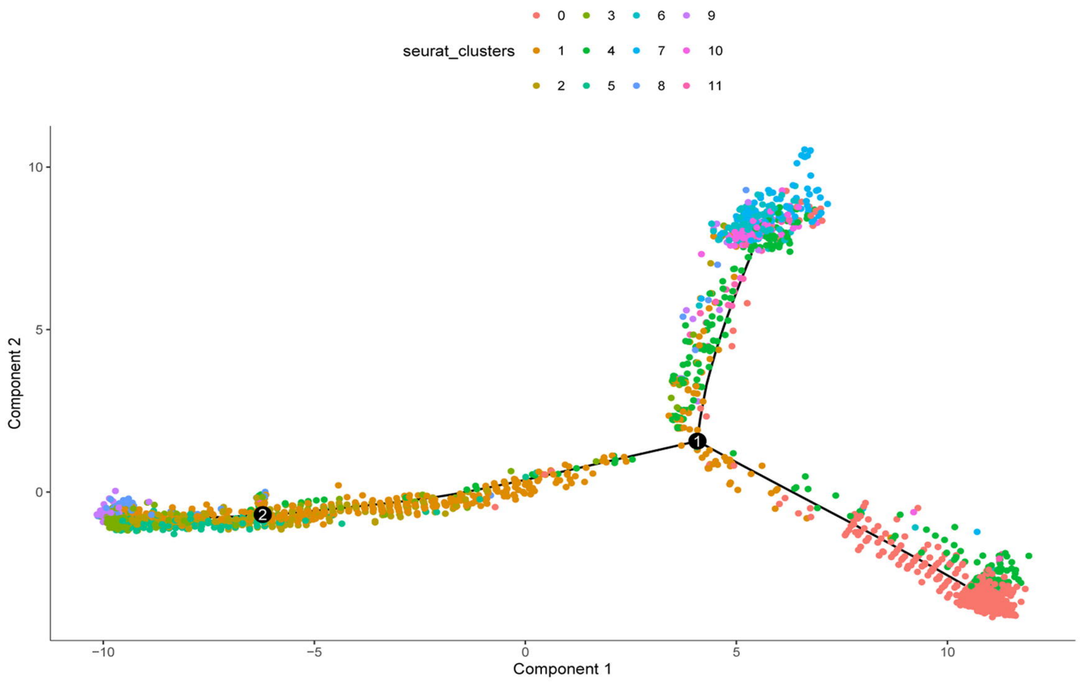
1192 progenitor subgroups. B, The activation and expression of the key regulators GATA1
1193 were obviously different during the priming stages of progenitors. C, Feature plots of
1194 key transcription factors. BCL11A, YY1, and REL were identified as key regulators
1195 in CLP priming. KLF1, GATA1, and GATA2 played a critical role in the lineage
1196 priming of CMPs.
1197
1198 **Figure 5. Characterization of progenitor subtypes in quiescent and activated**
1199 **status.** A, All integrated cells were presented by UMAP analysis. B, UMAP analysis
1200 illustrated 17 clusters of integrated CD34 positive progenitor cells . C, The cells
1201 distribution among the quiescent (Control) and activate (G_CSF) status, which
1202 showed all kinds of progenitors were persisting existence in the quiescent status of
1203 adult peripheral blood. D, Column plot displayed percentage changes of CLP and
1204 CMP clusters. The primitive cluster 0 showed the most obviously changes in cell
1205 numbers and folder. E, Violin plot showed the distribution and characterization of
1206 feature markers in activated status were consistent with quiescent status. F, Feature
1207 plots of unique marker genes. G, Dot plots presented expression of unique marker
1208 genes and lineages specific genes. H, Volcano plot showed the top five fold change
1209 genes in progenitor clusters. I, Heatmap depicted gene set enrichment analysis results
1210 of clusters by GSVA.
1211
1212 **Figure 6. Newly reaffirmed hematopoietic hierarchy of hematopoietic progenitor**
1213 **cells differentiation.** A, T-SNE plots showed the re-defined progenitor clusters. B,
1214 Dot plots presented the expression of fate decision transcription factor genes of each
1215 subgroup progenitors. C, Reaffirmed lineage commitment model through scRNA
1216 transcriptional atlas. In the new hematopoietic hierarchy, the HPCs priming into three
1217 progenitor cell types simultaneously and later maturing to diverse hematopoietic
1218 lineages. The first progenitor cell type is lymphoid progenitors (CLPs), which
1219 stepwise branch out and differentiate to T cell, B cell and NK cells. The second
1220 progenitor cell type is GAPs, which continuously differentiate into the middle stage of
1221 megakaryocytic–erythroid progenitors (MEPs) first, controlled by GATA2 and

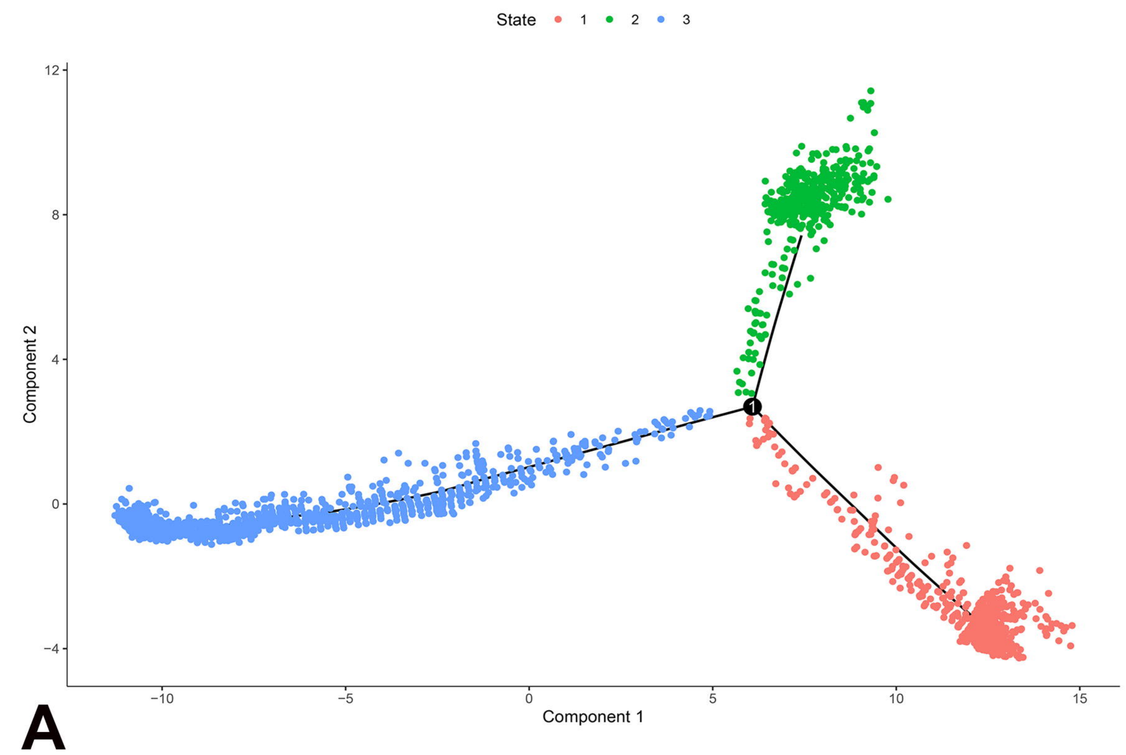
1222 GATA1. In the later stages, MEPs differentiate into two directions controlled by KIT,
1223 LMO4, and KLF1. The KIT and LMO4 high expression branches will be derived to
1224 mast cell or basophil, eosinophil progenitors (MBPs). The other branch differentiates
1225 into megakaryocytic and erythroid pro-progenitors (Pro-ME). The third progenitor
1226 cell type is neutrophilic and monocyte progenitors (NMPs), which differentiate into
1227 neutrophils and monocytes. Dashed lines meanings partial lineage connections that
1228 lack strong direct evidence of differentiation relationship. The shaded area indicates
1229 the GATA2 and GATA1 expression domains.
1230
1231
1232 **Figure 7. Expression of markers genes and fate decision genes of redefinition**
1233 **hematopoietic progenitor cells.** A, Heatmap showed the expression of markers genes
1234 and fate decision genes of re-defined progenitor clusters. B, Violin plot illustrated the
1235 expression of feature markers genes of SPINK2, HOPX, C1QTNF4, CSF3R, NPDC1,
1236 and SELL, which were highly expressed in the original progenitors of HPCs but
1237 restrained during the differentiation from GAPs to MEPs. C, The expression feature
1238 plot of LTB, TCF7, and MAFB. D, Violin plot demonstrated the expression of
1239 receptor genes, including CSF1R, IL7R, IL2RB, and others, which were specifically
1240 expressed in the redefined progenitor subgroups.


Classical hierarchy of hematopoiesis
HPCs
CLP
CMP
Pro-TBNK
GMP
MEP
T-cells
B-cells
NK-cells
Basophil
Neutrophilic
Pro-Monocyte
Eosinophil
Megakaryocyte
Erythrocyte
Reaffirming hierarchy of hematopoiesis
HPCs
HOXA9
CSF3R
NFE2
GAPs
GATA2
GATA1
CLP
NMPs
MEPs
Pro-TBNK
Pro-Monocyte
Pro-Neu
Pro-ME
MBPs
T-cells
B-cells
NK-cells
Dendritic cell
Macrophages
Neutrophilic
Megakaryocyte
Erythrocyte
Basophil
Mastcell
Eosinophil

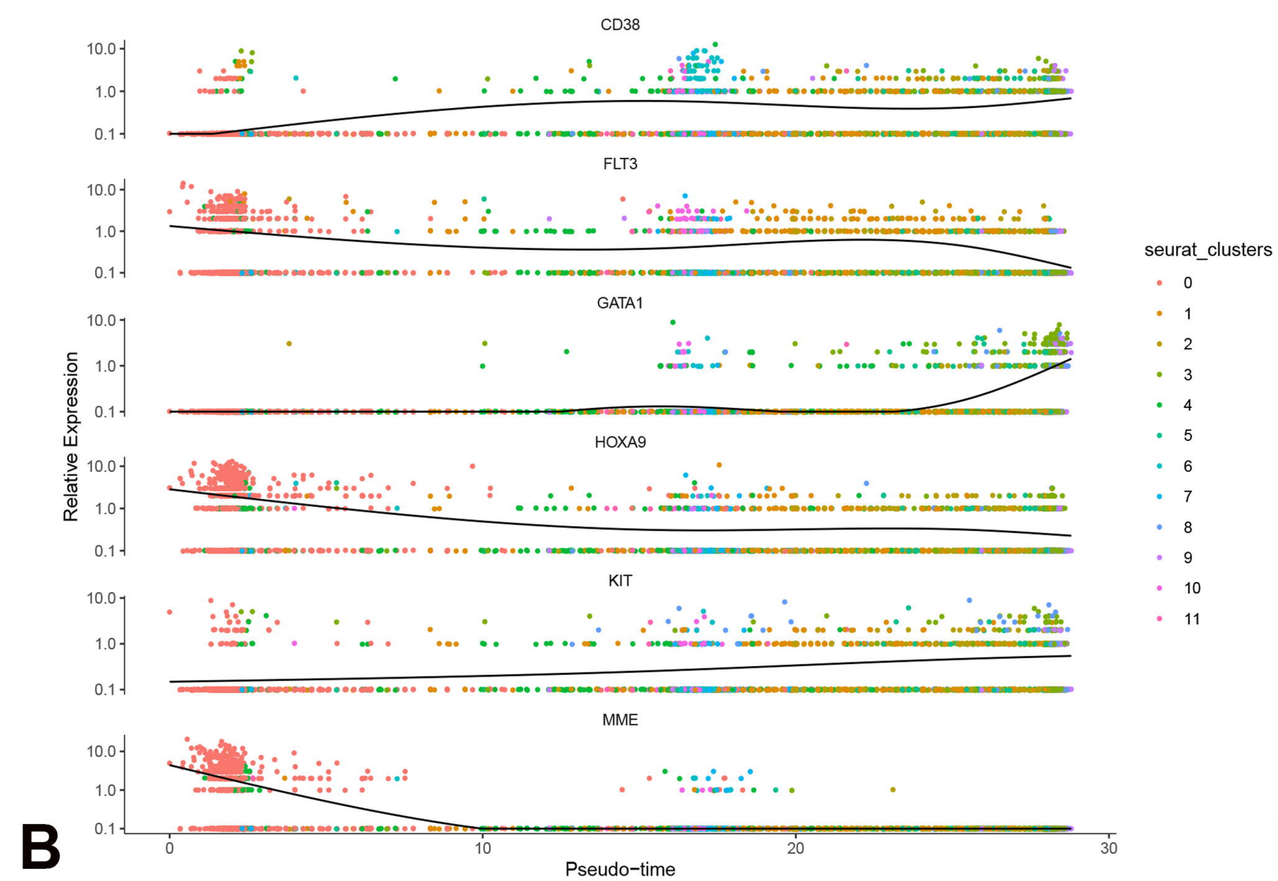

A
Ficoll separate
n=11
n=3
G-CSF
enrichment
scRNA-seq
Data analysis
UMAP_1
UMAP_2
B
C
Identity
Expression
D
CD34
CD38
FLT3
MME
KIT
CDK6
SOX4
SERPINB1
EGFL7
Expression Level
Identity


A
UMAP_1
UMAP_2
0
1
2
3
4
5
6
7
8
9
10
11
B
CD34
CD38
FLT3
MME
KIT
KLF1
CSF3R
GATA1
GATA2
Expression Level
Identity
C
AVP
CSF3R
KLF1
GATA1
GATA2
CD38
FLT3
KIT
IGHM
HOXA9
MME
CD68
GNLY
Identity
Expression
D
AVP
CSF3R
E
EnrichmentGO_BP_dot
positive regulation of lymphocyte activation
positive regulation of leukocyte activation
positive regulation of cell activation
positive regulation of T cell activation
positive regulation of leukocyte cell−cell adhesion
antigen processing and presentation
antigen processing and presentation of exogenous peptide antigen via MHC class II
antigen processing and presentation of peptide antigen via MHC class II
antigen processing and presentation of peptide or polysaccharide antigen via MHC class II
antigen processing and presentation of exogenous peptide antigen
antigen processing and presentation of exogenous antigen
antigen processing and presentation of peptide antigen
MHC class II protein complex assembly
peptide antigen assembly with MHC class II protein complex
peptide antigen assembly with MHC protein complex
MHC protein complex assembly
GeneRatio
Count
p.adjust
F
EnrichmentGO_BP_dot
carboxylic acid biosynthetic process
organic acid biosynthetic process
mononuclear cell differentiation
positive regulation of secretion by cell
positive regulation of secretion
response to oxygen levels
monocarboxylic acid biosynthetic process
multi−multicellular organism process
positive regulation of fibroblast proliferation
regulation of fibroblast proliferation
fibroblast proliferation
erythrocyte differentiation
erythrocyte homeostasis
nucleoside monophosphate metabolic process
definitive hemopoiesis
embryonic hemopoiesis
GeneRatio
Count
p.adjust

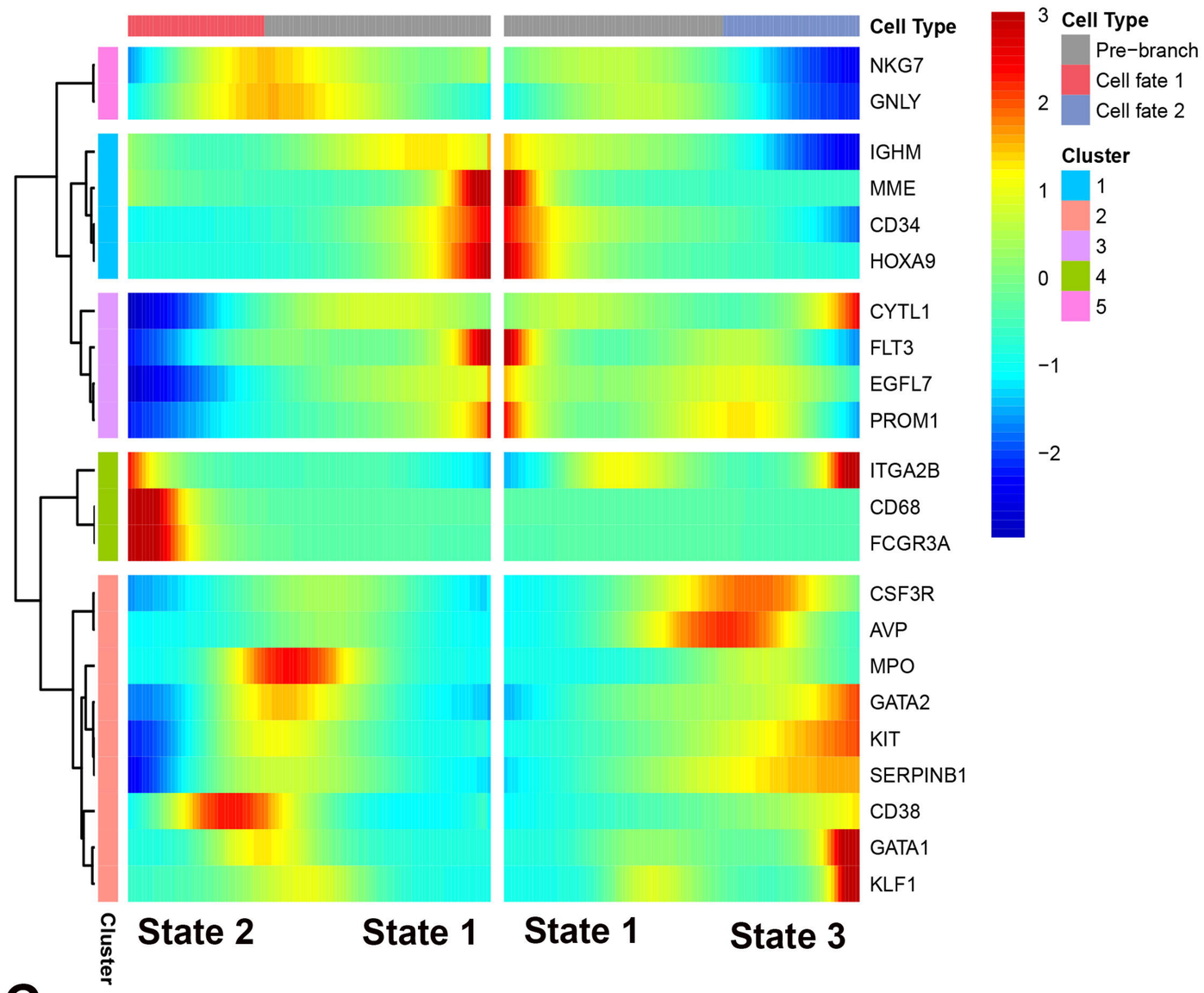

seurat_clusters
0 3 6 9
1 4 7 10
2 5 8 11
Component 2
Component 1


State 1 2 3
Component 2
Component 1

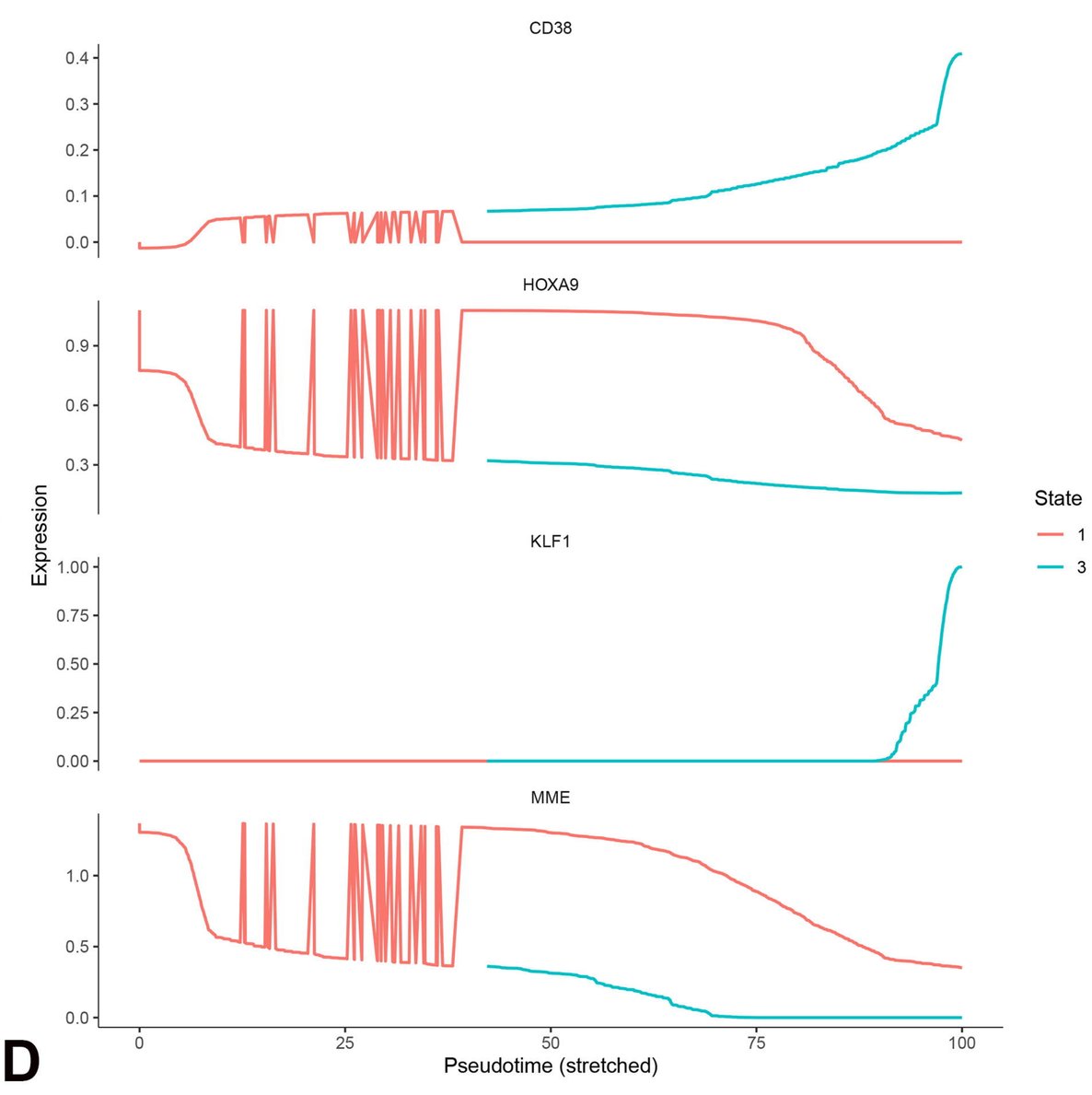
A


CD38
FLT3
GATA1
HOXA9
KIT
MME
Relative Expression
Pseudo−time
seurat_clusters
0
1
2
3
4
5
6
7
8
9
10
11

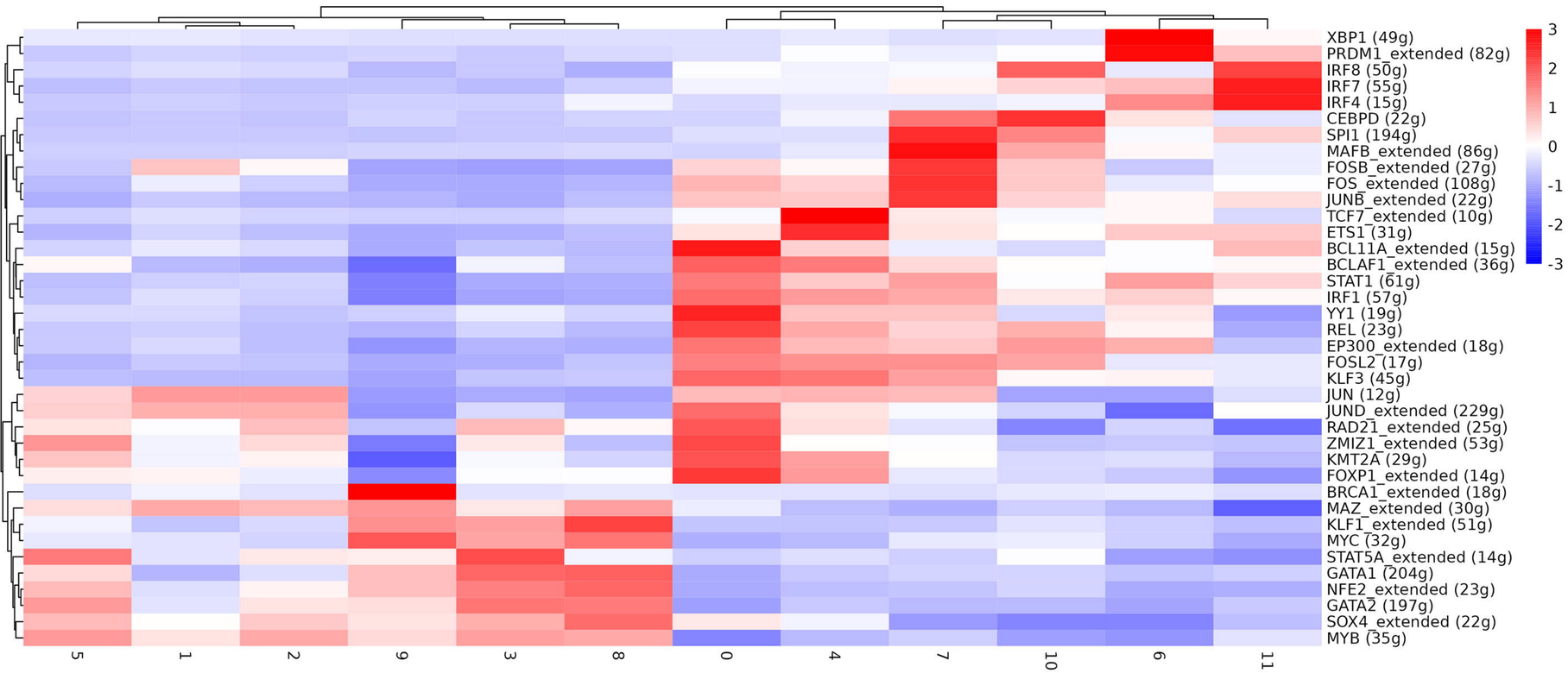
B

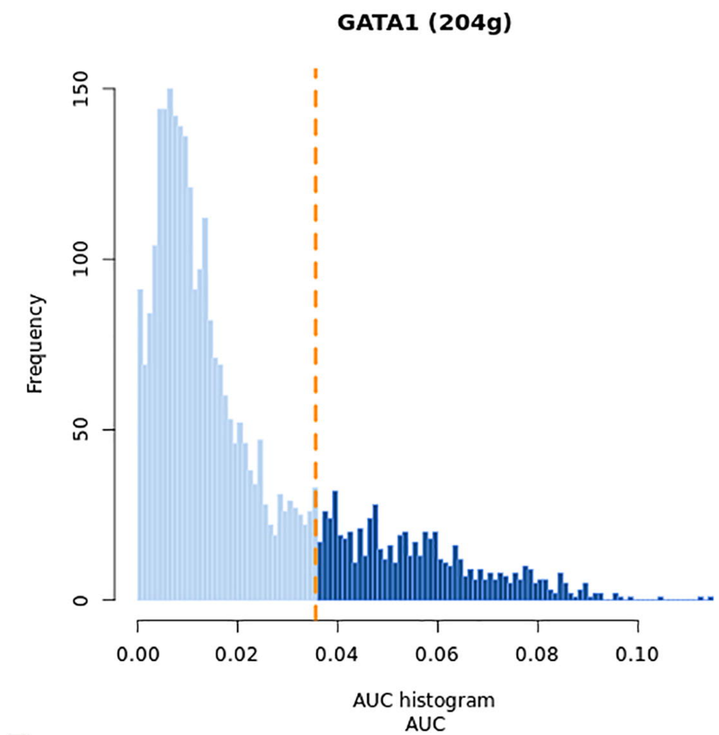

Cell Type
NKG7
GNLY
IGHM
MME
CD34
HOXA9
CYTL1
FLT3
EGFL7
PROM1
ITGA2B
CD68
FCGR3A
CSF3R
AVP
MPO
GATA2
KIT
SERPINB1
CD38
GATA1
KLF1
Cluster
State 2
State 1
State 1
State 3
Cell Type
Pre−branch
Cell fate 1
Cell fate 2
Cluster
1
2
3
4
5

C


CD38
HOXA9
KLF1
MME
Expression
Pseudotime (stretched)
State
1
3

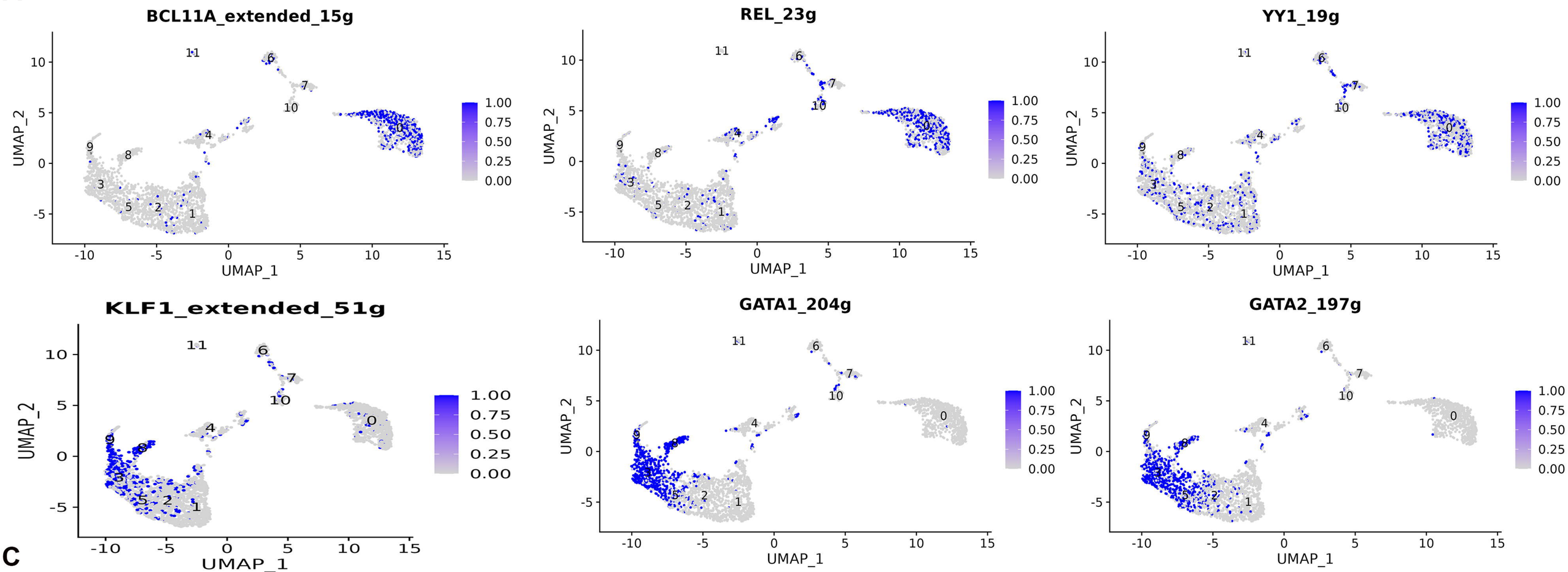
D


XBP1 (49g)
PRDM1_extended (82g)
IRF8 (50g)
IRF7 (55g)
IRF4 (15g)
CEBPD (22g)
SPI1 (194g)
MAFB_extended (86g)
FOSB_extended (27g)
FOS_extended (108g)
JUNB_extended (22g)
TCF7_extended (10g)
ETS1 (31g)
BCL11A_extended (15g)
BCLAF1_extended (36g)
STAT1 (61g)
IRF1 (57g)
YY1 (19g)
REL (23g)
EP300_extended (18g)
FOSL2 (17g)
KLF3 (45g)
JUN (12g)
JUND_extended (229g)
RAD21_extended (25g)
ZMIZ1_extended (53g)
KMT2A (29g)
FOXP1_extended (14g)
BRCA1_extended (18g)
MAZ_extended (30g)
KLF1_extended (51g)
MYC (32g)
STAT5A_extended (14g)
GATA1 (204g)
NFE2_extended (23g)
GATA2 (197g)
SOX4_extended (22g)
MYB (35g)
5 1 2 9 3 8 0 4 7 10 6 11
3 2 1 0 -1 -2 -3

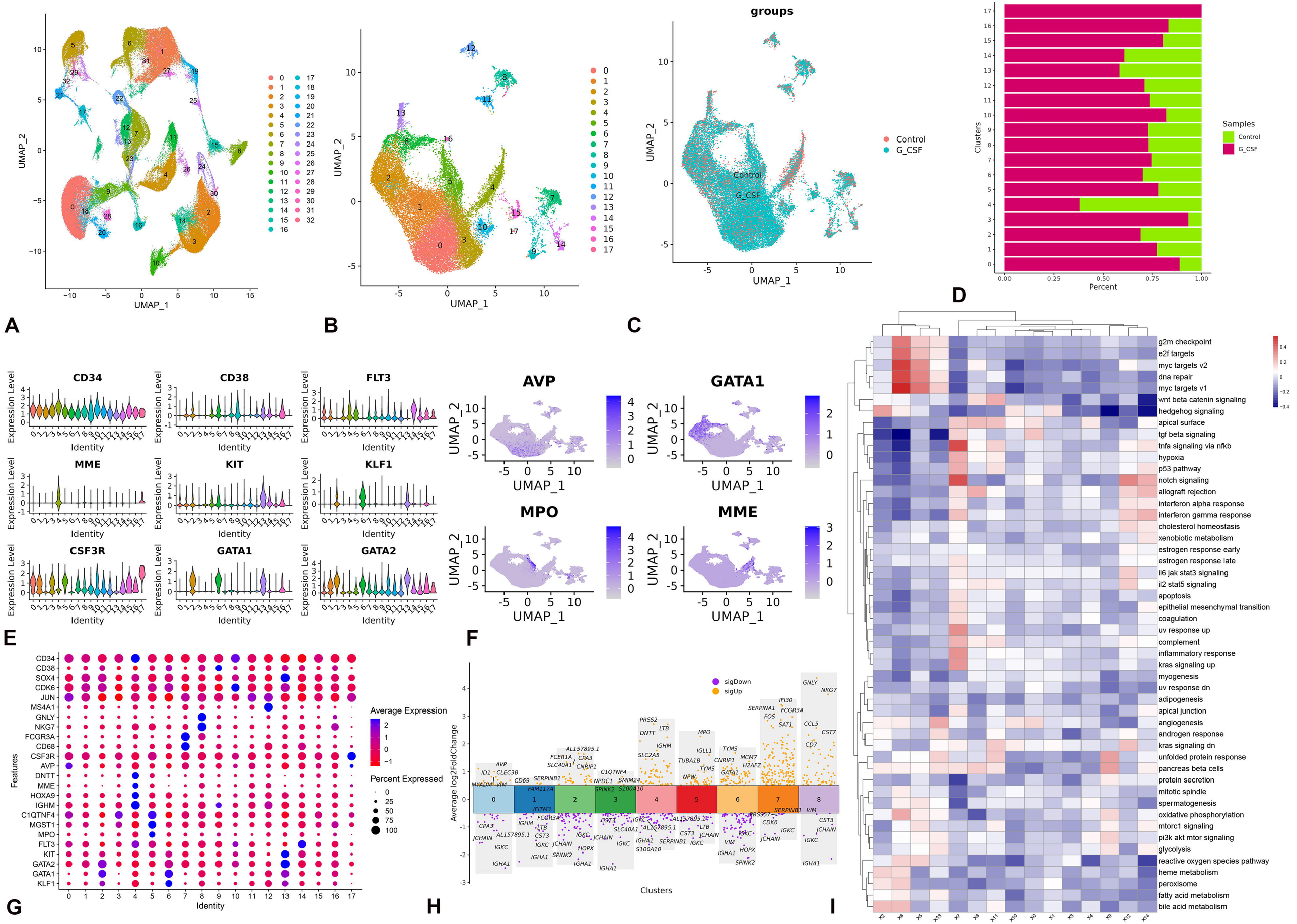

GATA1 (204g)
Frequency
0 50 100 150
0.00 0.02 0.04 0.06 0.08 0.10
AUC histogram
AUC

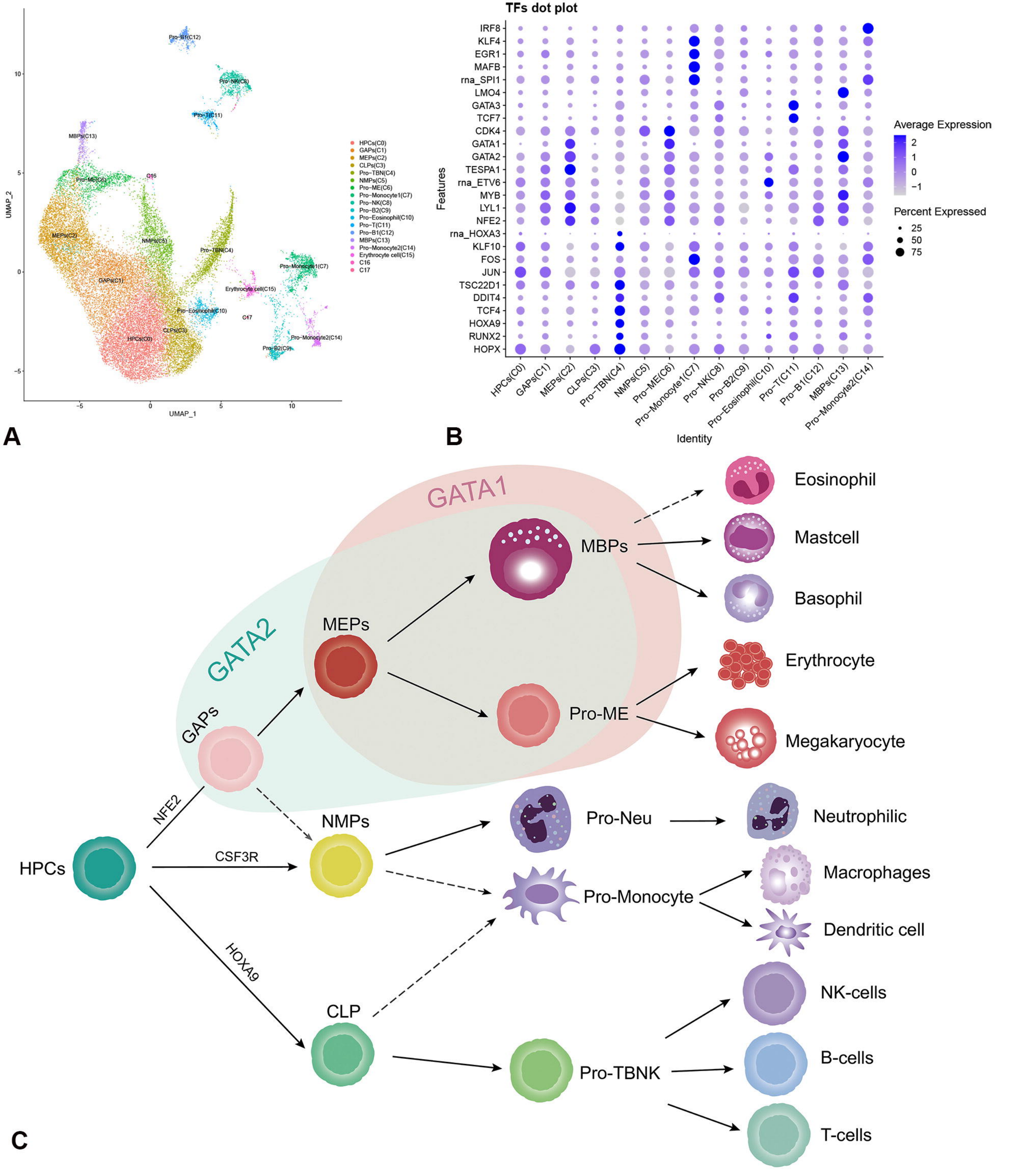
A

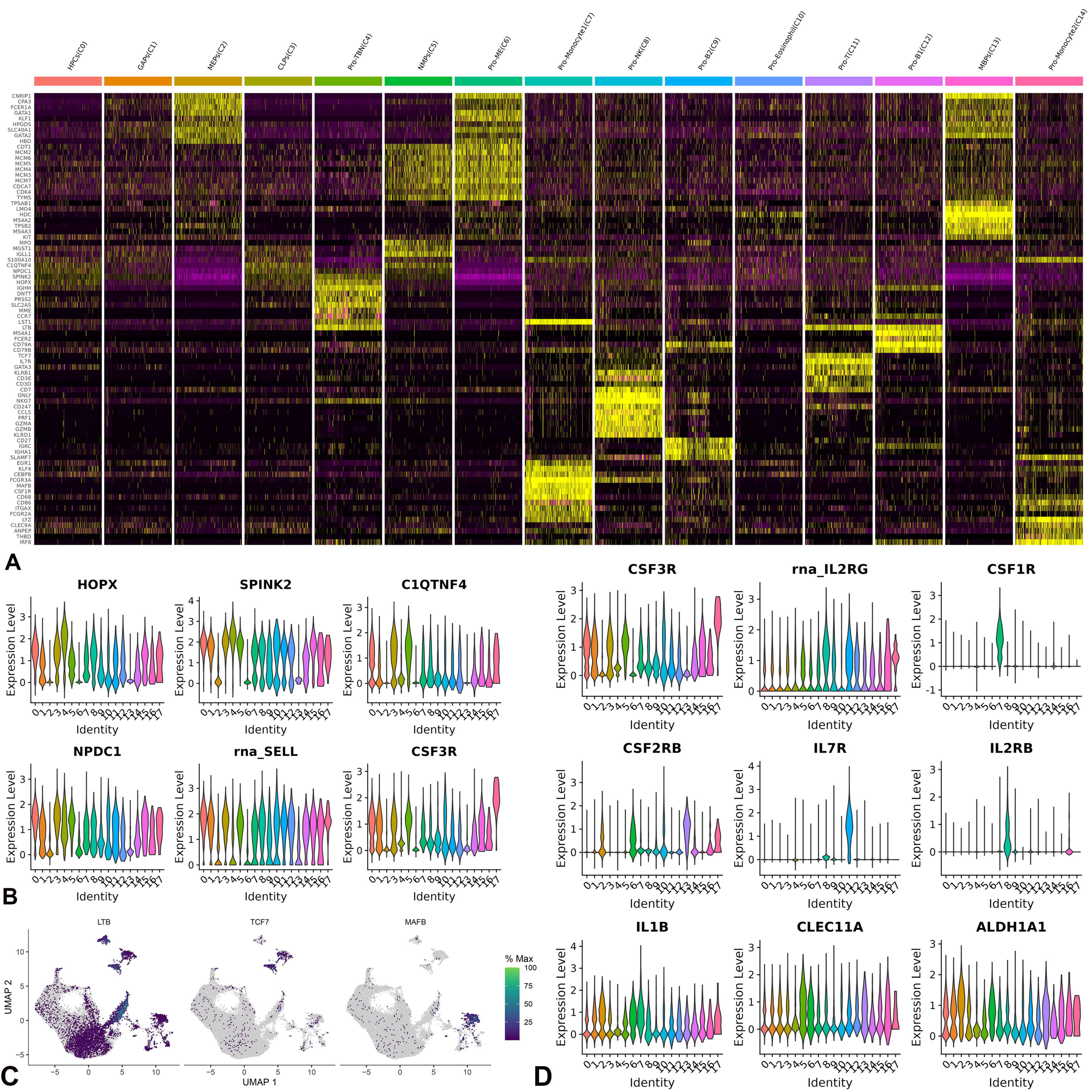
B


BCL11A_extended_15g
REL_23g
YY1_19g
KLF1_extended_51g
GATA1_204g
GATA2_197g
UMAP_1
UMAP_2
1.00 0.75 0.50 0.25 0.00

C


A
UMAP_1
UMAP_2
B
C
groups
Control
G_CSF
D
Clusters
Percent
Samples
E
CD34
CD38
FLT3
MME
KIT
KLF1
CSF3R
GATA1
GATA2
Expression Level
Identity
F
AVP
GATA1
MPO
MME
G
Features
Average Expression
Percent Expressed
H
sigDown
sigUp
Average log2FoldChange
Clusters
I
g2m checkpoint
e2f targets
myc targets v2
dna repair
myc targets v1
wnt beta catenin signaling
hedgehog signaling
apical surface
tgf beta signaling
tnfa signaling via nfkb
hypoxia
p53 pathway
notch signaling
allograft rejection
interferon alpha response
interferon gamma response
cholesterol homeostasis
xenobiotic metabolism
estrogen response early
estrogen response late
il6 jak stat3 signaling
il2 stat5 signaling
apoptosis
epithelial mesenchymal transition
coagulation
uv response up
complement
inflammatory response
kras signaling up
myogenesis
uv response dn
adipogenesis
apical junction
angiogenesis
androgen response
kras signaling dn
unfolded protein response
pancreas beta cells
protein secretion
mitotic spindle
spermatogenesis
oxidative phosphorylation
mtorc1 signaling
pi3k akt mtor signaling
glycolysis
reactive oxygen species pathway
heme metabolism
peroxisome
fatty acid metabolism
bile acid metabolism


A
UMAP_1
UMAP_2
Pro-B1(C12)
Pro-NK(C8)
Pro-T(C11)
MBPs(C13)
Pro-ME(C6)
C16
MEPs(C2)
NMPs(C5)
Pro-TBN(C4)
Pro-Monocyte1(C7)
GAPs(C1)
Erythrocyte cell(C15)
Pro-Eosinophil(C10)
C17
CLPs(C3)
HPCs(C0)
Pro-Monocyte2(C14)
Pro-B2(C9)
HPCs(C0)
GAPs(C1)
MEPs(C2)
CLPs(C3)
Pro-TBN(C4)
NMPs(C5)
Pro-ME(C6)
Pro-Monocyte1(C7)
Pro-NK(C8)
Pro-B2(C9)
Pro-Eosinophil(C10)
Pro-T(C11)
Pro-B1(C12)
MBPs(C13)
Pro-Monocyte2(C14)
Erythrocyte cell(C15)
C16
C17
TFs dot plot
Features
IRF8
KLF4
EGR1
MAFB
rna_SPI1
LMO4
GATA3
TCF7
CDK4
GATA1
GATA2
TESPA1
rna_ETV6
MYB
LYL1
NFE2
rna_HOXA3
KLF10
FOS
JUN
TSC22D1
DDIT4
TCF4
HOXA9
RUNX2
HOPX
Identity
Average Expression
2
1
0
−1
Percent Expressed
25
50
75
B
GATA1
GATA2
MBPs
MEPs
Pro-ME
GAPs
NFE2
NMPs
CSF3R
HPCs
HOXA9
CLP
Pro-Neu
Pro-Monocyte
Pro-TBNK
Eosinophil
Mastcell
Basophil
Erythrocyte
Megakaryocyte
Neutrophilic
Macrophages
Dendritic cell
NK-cells
B-cells
T-cells
C


HPCs(C0)
GAPs(C1)
MEPs(C2)
CLPs(C3)
Pro-TBN(C4)
NMPs(C5)
Pro-ME(C6)
Pro-Monocyte1(C7)
Pro-NK(C8)
Pro-B2(C9)
Pro-Eosinophil(C10)
Pro-T(C11)
Pro-B1(C12)
MBPs(C13)
Pro-Monocyte2(C14)
CNRIP1
CPA3
FCER1A
GATA1
KLF1
HPGDS
SLC40A1
GATA2
HBD
CDT1
MCM2
MCM6
MCM5
MCM4
MCM3
MCM7
CDCA7
CDK4
TYMS
TPSAB1
LMO4
HDC
MS4A2
TPSB2
MS4A3
KIT
MPO
MGST1
IGLL1
S100A10
C1QTNF4
NPDC1
SPINK2
HOPX
IGHM
DNTT
PRSS2
SLC2A5
MME
CCR7
LST1
LTB
MS4A1
FCER2
CD79A
CD79B
TCF7
IL7R
GATA3
KLRB1
CD3E
CD3D
CD7
GNLY
NKG7
CD247
CCL5
PRF1
GZMA
GZMB
KLRD1
CD27
IGKC
IGHA1
SLAMF7
EGR1
KLF4
CEBPB
FCGR3A
MAFB
CSF1R
CD68
CD86
ITGAX
FCGR2A
LYZ
CLEC9A
ANPEP
THBD
IRF8
A
HOPX
SPINK2
C1QTNF4
NPDC1
rna_SELL
CSF3R
Expression Level
Identity
B
LTB
TCF7
MAFB
% Max
100
75
50
25
UMAP 2
UMAP 1
C
CSF3R
rna_IL2RG
CSF1R
CSF2RB
IL7R
IL2RB
IL1B
CLEC11A
ALDH1A1
D