Redefining hematopoietic progenitor cells and reforming the hierarchy of hematopoiesis

Lipeng Chen[1,4], Qing Sun[2,4], Guoqiang Li[3], Qijun Huang[1], Sujin Chen[1], Yingyun Fu[1], Yongjian Yue[1,*]

[1]Shenzhen Institute of Respiratory Disease, the First Affiliated Hospital of Southern University of Science and Technology, Shenzhen People's Hospital, Shenzhen, Guangdong, China
[2]Shenzhen Key Laboratory of Reproductive Immunology for Peri-implantation, Fertility Center, Shenzhen Zhongshan Obstetrics & Gynecology Hospital (formerly Shenzhen Zhongshan Urology Hospital), Shenzhen, China
[3]Department of Hematology, The Second Clinical Medical College of Jinan University, Shenzhen People's Hospital, Shenzhen, Guangdong, China
[4]Lipeng Chen and Qing Sun contributed equally to this work.

***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 0755-2553 3018

Abstract

Deciphering the mechanisms underlying progenitor cell differentiation and cell-fate decisions is critical for answering fundamental questions regarding hematopoietic lineage commitment. Here, we redefine the entire spectrum of original hematopoietic progenitor cells (HPCs) using a comprehensive transcriptional atlas that effectively delineates the transitional progenitors. This is the first study to fully distinguish the transitional state along hematopoietic progenitor cell differentiation, reconciling previous controversial definitions of common myeloid progenitors (CMPs), granulocyte–monocyte progenitors (GMPs), and lymphoid-primed multipotent progenitors (LMPPs). Moreover, plasma progenitor cells are identified and defined. Transcription factors associated with key cell-fate decisions are identified at each level of the hematopoietic hierarchy, providing novel insights into the underlying molecular mechanisms. The hematopoietic hierarchy roadmap was reformed that reconciles previous models concerning pathways and branches of hematopoiesis commitment. Initial hematopoietic progenitors are simultaneously primed into megakaryocytic–erythroid, lymphoid, and neutrophilic progenitors during the first differentiation stage of hematopoiesis. During initial progenitor commitment, *GATA2*, *HOPX*, and *CSF3R* determine the co-segregation of the three transitional lineage branches. Two types of lineage-commitment processes occur during hematopoiesis: the megakaryocytic–erythroid lineage commitment process is continuous, while the lymphoid-lineage commitment is stepwise. Collectively, these results raise numerous possibilities for precisely controlling progenitor cell differentiation, facilitating advancements in regenerative medicine and disease treatment.

Keywords: hematopoietic progenitor cells, redefinition, lineage commitment, fate decision, hematopoietic hierarchy

Highlights

- Hematopoietic progenitors are redefined using a comprehensive transcriptional

54 atlas.

55 • Cell fate decision-related transcription factors are revealed in the hematopoietic
56 hierarchy.

57 • Progenitor lineage commitment includes continuous and stepwise processes.

58 • The initial hematopoietic hierarchy is simultaneously primed into three lineages.

59

Introduction

Hematopoietic stem cells (HSCs) comprise a rare cell population in the peripheral blood, capable of self-renewal and differentiation into full blood cell lineages with tree-like differentiation patterns in classical hematopoietic hierarchical models [1]. While the diversity of human peripheral blood progenitor cells is well-known, their conflicting definitions, lineage heterogeneity, and dynamic differentiation impede the characterization of the molecular mechanisms underlying lineage commitment. Moreover, the definition of hematopoietic progenitor cells (HPCs) has not been standardized due to the limitations in sorting methods and traditional definitions based on surface markers, including CD34, CD38, FLT3, and KIT [2]. Due to progenitors' rare and transient nature, surface markers fail to efficiently capture their entire spectrum and internal state or reflect the dynamic differentiation mechanisms underlying hematopoiesis [3, 4]. Moreover, using original quiescent progenitors derived from bone marrow, which are widely used for functional studies and subgroup identification, may cause bias as they do not cover the entire spectrum of progenitor stages [5, 6]. HPCs have been characterized in functional studies using selected $CD34^+$ cells; however, diverse cellular experimental conditions can lead to the inaccurate interpretation of HPC fate decisions as the conditions do not apply to all branches of lineage commitment [7]. Therefore, determining cell lineage definitions and fate decisions, as well as establishing reliable lineage commitment processes during hematopoiesis, remains challenging [8, 9].

Most established hierarchies of hematopoiesis are based on *in vitro* and *in vivo* functional experimental results [7, 10]. Single-cell RNA (scRNA) sequencing technology has been successfully implemented to explore the molecular heterogeneity and hierarchy of HPCs [11, 12]. However, the scRNA transcription atlas tends to modify aspects of the classical hematopoietic hierarchy, in which progenitors continuously differentiate during hematopoiesis [4, 13]. Lineage commitment is affected by transcription factors (TFs) within a continuum of cell fate-decision processes [2, 14]. However, $CD34^+$ progenitor clusters exhibit a scattered and mixed distribution based on typical markers, which contradicts the unsupervised

90 dimensionality reduction and clustering using the transcription atlas [12]. Additionally,
91 the unsupervised dimensionality reduction and clustering of lineage commitment can
92 result in a farraginous hierarchy model [12]. Furthermore, the inefficient and
93 inaccurate classical definitions of hematopoietic progenitors can result in disorganized
94 subgrouping without reliable boundaries [11, 15]. Indeed, several previously proposed
95 hematopoietic hierarchical models contradict each other [8, 16]. Meanwhile, the
96 critical functional roles of TFs in lineage commitment and the regulatory mechanisms
97 underlying the hierarchy of hematopoiesis remain unclear [17, 18].

98 In this study, we aimed to redefine adult HPCs, identify key TFs involved in
99 lineage-fate decisions, and reform the hematopoietic hierarchy. Our HPC atlas
100 redefines the cellular properties of progenitors at different stages throughout
101 hematopoietic cell-lineage commitment. Thus, our results provide new perspectives
102 for characterizing, defining, and classifying HPCs. This work represents a significant
103 step in clarifying the initial primed transitional progenitor and reforming the
104 hematopoietic hierarchy, resolving complexities in hematopoiesis commitment
105 pathways and models.

106

107 **Results**

108 ***Determining the characteristics of progenitor cells in adult peripheral blood via***
109 ***scRNA sequencing***

110 To enrich the entire HPC spectrum and intermediate-state progenitors, mature
111 lymphocytes and granulocytes were removed using a surface marker panel (Lineage^{-})
112 via negative deletion (Figure 1A). The mean percentage of CD34^{+} cells reached
113 10.3% and 46.1% in control peripheral blood mononuclear cells (PBMC) and
114 granulocyte colony-stimulating factor (G-CSF) mobilized peripheral blood groups,
115 respectively (Figure 1B). We performed 10× Genomics scRNA sequencing to
116 characterize the progenitor cells by unsupervised dimensionality reduction, and
117 clustering analysis using the Seurat toolkit and uniform manifold approximation and
118 projection (UMAP) analysis. Overall, 16,2579 qualified cells were captured and
119 sequenced after filtering the doublets and low-quality cells, identifying 29 clusters

120 (Figure 1C). Notably, the single-cell expression atlas showed that several $CD34^+$
121 progenitor clusters are gathered together and present heterogeneity.
122
123 Furthermore, we analyzed significantly differential genes identified among
124 progenitor cells and other cell clusters. *CD34, SXO4*, *CDK6*, *CYTL1*, and *EGFL7*
125 were highly expressed across the progenitor clusters (Figure 1D, E). However, the
126 classical markers of progenitors, namely *HLF*, *THY1* (CD90), and *ITGA6* (CD49f),
127 were negatively expressed in most progenitor clusters, excluding CRHBP (Figure 1F
128 and S1A, B). *EGFL7* and *CYTL1* were highly and specifically expressed in the
129 progenitor clusters accompanying expression differences among lineages, indicating
130 that they likely represent unique marker genes for the HPC lineage. The 30 most
131 highly expressed genes in progenitor cells are presented in Figure 1F and may have
132 essential roles in maintaining progenitor functions and stemness. Six highly specific
133 HPC genes were identified, namely *EGFL7, CYTL1, LAPTM4B, PRSS57, NPR3*, and
134 *SMIM24*. Combining these marker genes achieved high efficiency in identifying and
135 characterizing HPCs. Gene Ontology (GO) and pathway analysis further revealed the
136 enrichment of immune activity and regulation of hematopoietic cell lineages in HPC
137 clusters (Figure S1C, D). Gene set enrichment analysis of differentially expressed
138 genes showed that the hematopoietic cell lineage, MAPK, and PI3K–Akt pathways
139 were enriched in progenitor cells, but not significantly (Figure S1E). Therefore, based
140 on the identified conventional and novel markers and the aforementioned key
141 functional genes associated with HPCs, we successfully characterized the identified
142 $CD34^+$ HPC clusters.
143
144 ***Conventional surface markers are ineffective at accurately defining HPCs***
145 To reveal the heterogeneity of HPCs, we first extracted and re-clustered the
146 aggregated HPC subsets. Results showed that *CYTL1* and *CD34* negative cells were
147 clustered in the same subgroup, indicating that they were not progenitor cells (Figure
148 S2A). Notably, the expression characterization of *EGFL7, CYTL1, LAPTM4B*, and
149 *NPR3* was highly aggregated in CD34 positive clusters and presented as more

150 efficient and sensitive than CRHBP or HLF in progenitor cell identification (Figure
151 S2B). Given that the expression of CD34 was more uniform and pronounced than that
152 of other markers, further unsupervised dimensionality reduction re-clustering analysis
153 was performed on 25,450 $CD34^+$ cells to evaluate diversity, resulting in 14 clusters,
154 including 1,9953 mobilized and 5,497 quiescent progenitors (Figure 2A). A sufficient
155 number of progenitor cells were present in each cluster to generate a reliable
156 transcriptional atlas (Table S1). These progenitor cell subgroups exhibited categorical
157 and functional heterogeneity, clustering in three main differentiation directions
158 (Figure 2A). Conventionally, membrane markers of common myeloid progenitors
159 (CMPs) and granulocyte–monocyte progenitors (GMPs) are defined based on the
160 $CD38^+MME^-FLT3^+CD45RA$ phenotype, whereas megakaryocytic–erythroid
161 progenitors (MEPs) are defined according to the $CD38^+MME^-FLT3^-$ phenotype [2].
162 However, our results showed that CD38, FLT3, IL3RA, and KIT are pan-expressed in
163 most progenitor clusters and were indistinguishable among the subgroups (Figure 2B,
164 C). The $FLT3^+$ progenitor cells were a diverse population, which may explain the
165 varying differentiation potentials observed in previous functional studies of select
166 $FLT3^+$ progenitors [19, 20]. Moreover, *GATA1* and *FLT3* showed opposing expression
167 patterns in the clusters (Figure 2B), indicating that $CD34^+GATA1^+$ and $CD34^+FLT3^+$
168 progenitors represented two distinct differentiation directions. Canonical markers,
169 including *CD49f* (*ITGA6*) and *THY1*, were also expressed at very low levels in HPCs
170 (Figure 2C). However, the TFs *GATA1* and *GATA2* showed high effectiveness in
171 distinguishing HPC intermediate stages (Figure 2B).

172 We subsequently merged a bone marrow single-cell cohort of 7,037 $CD34^+$ cells
173 from seven projects to validate the identified markers. The bone marrow progenitor
174 cells exhibited scattered expression distributions, potentially caused by the inefficient
175 capture of the intermediate state, and an insufficient number of cells in each cluster
176 (Figure S2C). Expression of the canonical markers *THY1, CD49f*, and *HLF* also
177 exhibited low efficiency. Similarly, clusters in bone marrow $CD34^+$ cells could not be
178 clearly distinguished according to the traditional classification criteria of *FLT3,*
179 *IL3RA*, and *KIT*. In contrast, the newly identified markers *EGFL7*, *CYTL1*, and *SOX4*

displayed high efficiency in classifying clusters (Figure S2D), supporting the reliability of the marker identification (Figure 1E, F). Overall, our data show that using the classical and canonical markers *CD49f*, *THY1*, *FLT3*, *IL3RA*, and *KIT* to classify or define progenitor cell groups is relatively arbitrary and subjective. This leads to a heterogeneous mixture of various progenitors that are ineffectively distinguished as CMPs, GMPs, and CLPs [18, 21].

Redefinition of adult peripheral blood progenitor cells

To accurately and efficiently redefine the progenitor clusters, new marker genes, especially TFs for distinguishing progenitor clusters, were identified in the reliable transcriptional atlas. The subset distribution reproducibility across samples was exemplary, indicating the reliable repeatability of each atlas (Figure S3A). The clustered progenitors were redefined by identifying the unique markers of each subset (Figure 2C). Expression of the canonical markers *CD49f* (*ITGA6*), *SLAMF1,* and *THY1* was absent or very low in all HPC clusters. *AVP* and *CRHBP* were expressed in HSCs and, along with *CSF3R*, were primarily expressed in cluster 0 (C0). Hence, the cluster likely comprised the original HSC progenitors designated multipotent progenitor cells (MPCs; Figure 2B and C). Myeloid progenitors were detected in C1, C3, C6, C2, and C5 with gradual upregulation of *CD38, GATA2*, *NFE2*, and *KIT* (Figure 2B). The unidirectional lineages C1, C3, and C6 could be distinguished based on the expression levels of TFs. For example, expression of the erythrocyte fate decision TFs *KLF1*, *GATA1*, and *GATA2* (Figure 2B) indicated the presence of megakaryocytic–erythroid (MkE) lineage progenitors. In contrast, C2 and C5 expressed myeloperoxidase (*MPO*), *IGLL1*, and *CSF3R* marker genes (Figure 2C). MPO is the most abundant and uniquely expressed protein in human neutrophils, with a vital role in neutrophil antimicrobial responses [22, 23]. Hence, the progenitor branch was likely neutrophil–monocyte progenitors (NMPs), with an entirely unique orientation from megakaryocyte-erythrocyte and lymphoid progenitors. In contrast, *IGHM*, *MME*, and *HOXA9* [24] were expressed in C4 and C7, confirming that this cluster comprised common lymphoid progenitor cells (Figure 2C). Expression of the

210 TF *HOXA9* increased sequentially from C4 to C7. The characterization of CSF3R and
211 GATA2 expression clearly demonstrated the co-segregation among NMPs,
212 megakaryocyte-erythrocyte, and lymphoid progenitors (Figure 2B). Accordingly, TF
213 expression fully distinguished the intermediate state and reflected the mechanisms
214 underlying hematopoietic lineage commitment. Assessing the expression and
215 activation of TFs with unique marker genes was efficient for defining HPC lineages.

216 To assess differences between mobilized and unperturbed progenitor cells,
217 cell-distribution analysis was performed. The total number of progenitor cells was
218 remarkably increased after G-CSF mobilization. A comparison of the progenitor cells
219 before and after mobilization revealed that all progenitor cell subtypes existed in the
220 unperturbed state (Figure 2D). *CSF3R* expression is consistent with G-CSF
221 mobilization in the peripheral blood during HSC transplantation [25]. This signifies
222 that the entire spectrum of progenitor cells exists in an unperturbed state within the
223 peripheral blood, consistent with the cluster distribution in the mobilized state (Figure
224 2D). This might also indicate that unperturbed adult hematopoiesis depends on
225 replenishment from multiple progenitor cell types. As the progenitor cells gradually
226 differentiate into mature cells in the peripheral blood, we posited that the number and
227 proportion of more primitive progenitors would be higher than those of mature
228 progenitors. The largest relative abundance and increase in absolute number were
229 found in C0 (Figure 2E), the original multipotent cluster, which can differentiate into
230 multiple lineages.

231 To identify the CMPs, all early priming-stage clusters were re-clustered. A
232 common intermediate progenitor stage was not detected among the priming-stage
233 progenitors, and the clusters could not be distinguished based on FLT3, CD38, and
234 IL3RA expression (Figure S3B). Hence, defining the CMPs or GMPs by surface
235 markers would create a heterogeneous mixture of various progenitors (Figure S3C).
236 The original MPCs were simultaneously primed to three transitional lineages, which
237 committed to lymphoid, neutrophil/monocyte, and megakaryocytic–erythroid
238 progenitor cell types; they subsequently matured into different hematopoietic lineages
239 (Figure 2A). The subsets of the three transitional lineages were adjacent, lacking clear

240 boundaries, indicating the continuous differentiation trajectories of the early
241 commitment process. Moreover, the three transitional lineages were co-segregated
242 simultaneously, contradicting the conclusion that lymphoid-primed multipotent
243 progenitors (LMPPs) are co-segregated between neutrophils/monocytes and lymphoid
244 cells [4, 15]. Although the surface markers (CD38-CD10+) of CLPs are consistent
245 with LMPPs, they should not be defined as lymphoid-primed multipotent progenitors
246 as they only present lymphoid-primed potential, not neutrophil/monocyte-primed
247 potential. Rather, LMPPs (CD38-CD10+CSF3R+) likely represent a heterogeneous
248 mixture of neutrophil and lymphoid transition progenitors [15]. Based on the
249 successful capture of the entire spectrum of adult peripheral blood progenitor cells,
250 including those in mobilized and quiescent progenitor states, we generated a
251 comprehensive transcriptional atlas for adult peripheral blood progenitor cells.

252 To accurately redefine the HPC subgroups, we analyzed the characteristics of
253 various progenitor cell subtypes, identifying unique and specific marker genes and
254 key TFs of progenitor cells in each lineage (Table 1 and Figure 2C, F). Additionally,
255 negatively expressed genes were identified to distinguish each similar or identically
256 redefined subgroup (Table 1). The transcriptional atlases of the two lineages of C2
257 and C5 (NMPs) or C4 and C7 (CLPs) exhibited remarkable comparability, with C2
258 and C4 presenting an early tendency toward committed NMPs or lymphoid progenitor
259 cells. We redefined the three MPC-primed intermediate directions of CLPs, *GATA*
260 gene-controlled progenitors (GAPs), and NMPs that could not be characterized based
261 on CD38, FLT3, or KIT expression (Figure 4A and Table 1). Some of the specific
262 marker genes for each progenitor lineage overlapped with canonical marker genes of
263 classically defined hematopoietic cells, providing complementary evidence of the
264 reliable redefinition (Figure 2F).

265

266

267 ***Trajectories and dynamics of progenitor cell differentiation***

268 As commitment trajectories can be readily revealed based on the aforementioned
269 T-SNE distribution and reliable definition, MPCs were found to be gradually primed

270 into three transitional progenitor cell branches and subsequently transitioned into
271 diverse lineages (Figure 3A). To confirm the differentiation trajectories of the
272 progenitor cells, complementary pseudo-time analysis was performed using Monocle2,
273 which showed a consistent tendency of differentiation trajectory commitment,
274 although the NMP lineages were a farraginous prediction (Figure S4A). The
275 pseudo-time results showed that the redefined clusters of CLPs and MkEs were
276 differentiated along the distinct pseudo-time trajectory states, respectively (Figure 3B,
277 C). The dynamic changes in *GATA2* and *HOPX* expression were co-segregated and
278 separated from each other along the lineage differentiation trajectory (Figure 3D and
279 S4B, C). Additionally, lineage signature genes, such as *KLF1*, *GATA1*, and *GATA2* of
280 MkEs and *HOXA9*, *HOPX*, *MME*, and *RUNX2* of CLPs, exhibited unique expression
281 patterns along the differentiation trajectory (Figure 3E and S4E, F). The results of
282 in-silico prediction provided complementary evidence of the progenitor differentiation
283 trajectory. The expression trajectories of the signature genes effectively validated the
284 reliability of the differentiation trajectories of lineages (Figure 3A).

285 Similarly, the inference of the single-cell regulatory network and clustering
286 assessment based on TF and regulon activities showed the activation of key TFs in
287 cell-fate decisions during progenitor commitment (Figure 3F). The original progenitor
288 of MPCs exhibited RFX2 activity, whereas the TFs *MAZ*, *HOXA9*, *KLF3*, *STAT1*,
289 *JUND*, *RUNX2*, and *IRF1* were identified as key regulators of CLP priming (Figure
290 3G). Several TFs, including *SF1*, *RUNX1*, *ZMIZ1*, and *BCLAF1*, were identified in
291 eosinophilic progenitors. Moreover, *GATA1*, *MYC*, *SOX4*, and *HMGB1* played critical
292 roles in the progression of MkE lineages (Figure 3F). XBP1 was activated in plasma
293 progenitor cells, and the TFs *IRF8*, *CEBPD*, *JUNB/FOS*, *SPI1*, *EGR1*, and *CEBPB*
294 determined cell-fate decisions during monocytic lineage differentiation. TF and
295 regulon activation showed a high overlap with the unique expression pattern in
296 clusters that support cell-fate decisions in progenitor commitment. The characteristics
297 of TF activation were similar in the same origin lineages of MkE lineages (Figure 3F).
298 The two adjacent progenitor CLPs and the eosinophil clusters had similar
299 TF-activation characteristics, indicating that they may have been primed by similar

300 origin cells (C4) but still required confirmation. Therefore, integrating multi-omics
301 and epigenetic analyses using the scRNA assay for transposase-accessible chromatin
302 may improve our understanding of the HPC-fate-decision process. Our findings
303 revealed highly consistent TF expression and activation patterns throughout the
304 lineage-committing process, reflecting the synergistic effects of the suppression and
305 activation of key regulators.
306
307 ***Reforming the hematopoietic hierarchy***
308 Based on the reliable redefinition, precise characteristics of the entire progenitor
309 spectrum, and clear differentiation trajectory, the hematopoietic hierarchy roadmap
310 was reformed (Figure 4A). The shapes of the three transitional branches were
311 controlled by the expression of key regulatory genes, which functioned similarly to a
312 gradual activation switch. During initial progenitor-differentiation, GATA2, HOPX,
313 and CSF3R regulated the formation of the first branches of the three transitional
314 lineages (Figure 4A, B). Downregulation of *CSF3R* and upregulation of *GATA2*
315 implied that the progenitors were primed to generate oligo-potent progenitor GAPs
316 (Figure 3B). Inversely, the downregulation of *CSF3R* and *GATA2* would prime the
317 MPC committed to CLPs with the upregulation of HOPX and HOXA9. Hence, the
318 separation between NMPs, CLPs, and megakaryocyte/erythroid lineages was
319 determined by GATA2, HOPX, and CSF3R, not GATA1 [4, 26]. That is, *GATA2,*
320 *HOPX,* and *CSF3R* expression marked the relative boundary between GAPs, CLPs,
321 and NMPs (Figure 4B). The MkE lineages differentiated independently from the other
322 two branches. The expression of *GATA1* and *KLF1* was gradually increased at later
323 mature stages, at which point cells were determined and defined as MEPs (Figure 2B).
324 CSF3R was upregulated in the original MPCs and NMPs but was nearly absent during
325 the later stages of MEPs (C1) and CLPs (C7) (Figure 2B and 4B). GATA2, HOPX,
326 and HOXA9 expression was restrained when the progenitors were primed to generate
327 the NMP branch (Figure 3B, C). Cytokine receptors were co-expressed or co-silenced
328 with relevant key lineage regulators, indicating that multiple genes regulated the
329 progenitor lineage-priming process. During the differentiation of the MkE lineages,

330 the expression levels of unique genes, such as *GATA2*, *CPA3*, *CNRIP1*, and *SLC40A1*,
331 gradually increased. We also identified genes with the opposite expression pattern that
332 were specific among MkE lineages (C1, C6, and C11) and other progenitors, such as
333 *SPINK2*, *C1QTNF4*, *NPDC1*, *S100A10*, and *SELL*. Enhanced MEP differentiation
334 resulted in the formation of the second branch of mature MEPs. Toward the final
335 maturation stages of MEPs, KIT, GATA2, KLF1, or LMO4 expression guided the
336 differentiation of progenitors to megakaryocyte–erythrocyte (Pro-ME) cells or mast
337 cell and basophil progenitors (MBPs) (Figure 4A). The receptor of TPO, MPL, was
338 absent in Pro-ME, indicating that megakaryocyte progenitors were more likely
339 originated from MEPs (C1) but not co-segregated with erythrocytes in the Pro-ME
340 cluster. Additionally, *LMO4*, *MS4A2*, and *TPSB2* were uniquely upregulated in MBPs.
341 The hemoglobin subunit delta was identified as a protein involved in early
342 erythrocytic development. Mini-chromosome maintenance protein genes, encoding
343 components of the MCM2–7 complex, were highly expressed in Pro-ME cells and
344 NMPs (Table 1). The MCM complex is the replicative helicase that plays an essential
345 role in the eukaryotic DNA replication-licensing system [27].Our results indicate that
346 the status of the two lineage progenitors (NMPs and Pro-MEs) may switch from
347 differentiation (commitment) to proliferation. Finally, the continuous differentiation
348 process of neutrophils was directly primed from MPCs and underwent two
349 commitment stages before transfer to proliferation (Figure 4A).
350
351 Our results revealed that the original progenitors of basophils, eosinophils, and
352 neutrophils differed (Figure 4A), consistent with the results of a previous study, which
353 reported that the origin of neutrophils is distinct from that of basophils and
354 eosinophils during lineage commitment [4]. However, the identity of the original
355 eosinophil progenitor was not confirmed. The original progenitor cells were primed
356 and differentiated into various progenitors through gradual activation/deactivation
357 followed by TF accumulation, synergism with key receptor genes, and signal
358 activation. MEP-representative genes gradually accumulated during differentiation,
359 revealing their distinct continuous differentiation characteristics (Figure 2B and 3C).

360 These findings suggest that the hierarchical organization of MkE lineages during
361 hematopoiesis is a continuous process controlled by the activation and suppression of
362 cell-fate decision-associated genes.

363 During priming of the lymphoid progenitor branch, critical TFs (including
364 *HOXA9* and *HOPX*) were upregulated (Figure 3C, D). Similarly, marker genes of
365 lymphocytic lineages (such as *MS4A1*, *GNLY*, and *GATA3*) were uniquely expressed
366 during specific stages and upregulated stepwise. Certain genes were prominently and
367 specifically co-expressed in lineage-restricted progenitors and CLPs: CLPs expressed
368 *IGHM*, *FCMR*, *BLNK*, and *LTB*; B progenitors expressed *CD79A* and *JCHAIN*;
369 monocyte progenitors expressed *SPI1*, *FOS*, *PECAM1*, and *LST1*; and natural killer
370 (NK) progenitors expressed *NKG7* and *GZMA*. The expression profiles of these genes
371 also indicated the underlying mechanisms of lymphoid lineage commitment. The
372 lymphoid lineage-commitment process is likely a discrete stepwise process, with
373 distinct lineage-committed populations formed during hematopoiesis. This was also
374 observed for monocyte-lineage differentiation into macrophages and dendritic cell
375 (DC) progenitors. We also identified the co-expression of *ETV6, LYL1,* and *NFE2* in
376 NMPs and monocyte lineage-restricted progenitors, resembling the co-expression of
377 *SPI1, FOS, PECAM1*, and *LST1* in monocytes and CLPs (Figure 2C and 3B).
378 However, we were still unable to confirm the original monocyte lineage during
379 hematopoiesis. The commitment of each progenitor lymphoid lineage was controlled
380 by the expression of key regulatory genes, which collectively functioned as a stepwise
381 activation switch. By switching the sequential expression of TFs, which generally
382 showed a synergistic relationship with key receptor gene expression and signal
383 activation, the original progenitor cells were primed and assigned to the subsequent
384 stages of lymphoid-lineage progenitors (Figure 4). This study is the first to clear
385 define the initial primed transitional progenitor and reformed the hierarchy, effectively
386 reconciled the puzzle concerning pathways and models of hematopoiesis
387 commitment.

388

389 ***Identification of cell fate decision-associated genes in HPC maintenance and***
390 ***lineage commitment***

391 Cell-fate decision-associated factors were identified in different lineages,
392 including *NFE2, LYL1, MYB, CDK4, ETV6, TESPA1, ZBTB16*, and *GATA2* in
393 myeloid progenitors and *RUNX2, HOXA9, JUN/FOS, TCF4, DDIT4, HOPX, KLF10,*
394 *HOXA3*, and *TSC22D1* in lymphoid progenitors (Figure 4C). Notably, these key
395 factors were determined to play contrasting functional roles during HPC
396 differentiation. HPCs differentiated into diverse myeloid progenitors by enhancing the
397 expression and functions of *NFE2, LYL1, MYB, CDK4, ETV6, TESPA1*, and GATA2
398 while suppressing the expression and functions of *RUNX2, HOXA9, TCF4, DDIT4,*
399 *HOPX, KLF10, HOXA3*, and *TSC22D1*. Together with the activation of other
400 receptors and unique TFs (such as CSF3R in neutrophils and CSF2RB in MkEs) in
401 each progenitor lineage, the entire hematopoietic lineage-commitment process was
402 precisely controlled and occurred gradually (Table 1 and Figure 4A, C).

403 To identify the essential genes involved in maintaining HPC states, we compared
404 the expression levels in three cohorts of bone marrow cells, control PBMCs, and
405 CD34+ single cell subsets. We classified the progenitor-maintaining genes using the
406 following criteria: first, the gene must be pan-expressed across most PBMC and bone
407 marrow HPC clusters but should be excluded once it serves as a specific lineage fate
408 decision gene (such as GATA1 for MkE lineages); second, the expression of this gene
409 should be limited to progenitor cell types. Such genes, pan-expressed in both
410 progenitor and mature cells, would be considered HPC functionally-related genes but
411 would not be associated with progenitor maintenance. In this regard, our data revealed
412 that *RUNX1* regulated the differentiation of most progenitors (Figure 3C), whereas
413 *RUNX2* primarily regulated the differentiation of lymphoid progenitors. We also
414 found that ITGA4 (CD49d) was highly expressed in most progenitors, which is
415 important for molecular regulation during HSC homing [28]. Functional HSC-marker
416 genes, including *PTPRC*, *HMGB1*, *FOXP1, CD74*, *CD37*, *PFN1*, *TXN*, *ZFP36L2*,
417 *SERPINB1*, *TMSB4X*, and *HSP90AA1,* were broadly expressed in most progenitor
418 cells but were not specific (Table 1 and Figure 1F). These genes may mediate the

419 self-renewal and mobilization, but not differentiation and maintenance, of HPCs.
420 Additionally, two new genes (*PPIA* and *EIF1*) were identified, which were similar to
421 previously reported genes as they exhibit pan-expression in progenitors and were
422 highly expressed in mature subgroups of peripheral blood cells.

423 *XBP1, SOX4, CD82, SPINT2, IMPDH2, TSC22D1, SPI1* [14], and *ETV6* were
424 also widely expressed in most progenitors and progenitor-differentiated cells; however,
425 their expression was relatively absent in all other cell types and control cohorts
426 (Figure 1F and Table 1). Moreover, *ERG, BCL11A, LMO2*, and *RUNX1* were widely
427 expressed in approximately 50% of the cells in each progenitor subgroup. These
428 genes, similar to EGFL7 and CYTL1 (Figure 1F), were more highly or specifically
429 expressed in the bone marrow and progenitor cells than in mature PBMCs (Figure S5),
430 which highlights their significance in regulating the HPC state for maintenance and
431 differentiation.

432

433 ***Identification and validation of fate decision-associated genes in each redefined***
434 ***subgroup***

435 To identify cell-fate decision-associated genes, particularly those encoding TFs,
436 advanced analysis was performed on enriched CD34+ subsets. The unique cell-fate
437 decision-associated TFs were identified in each cluster of lineages (Figure 4C). Four
438 key TFs (GATA3, TCF7, BCL11B, and DDIT4) were identified in T progenitor cells
439 (Figure 4C). These four TFs were critical for cell fate decisions during T-cell lineage
440 differentiation (Table 1 and Figure 4C). Previous findings suggest that *TCF7* and
441 *GATA3* are critical TFs for T-cell development and differentiation [29, 30], which is
442 highly consistent with our redefined characterization of T progenitor cells. We
443 observed that LTB was widely expressed by CLPs, Pro-T, and Pro-B progenitor cells
444 and that it was upregulated during the later stages of lymphoid progenitor
445 differentiation. Two distinct types of B progenitor cells were identified and redefined
446 as expressing different cell-fate decision-associated TFs and marker genes in C13
447 and C14. Two TFs (*PRDM1* and *IRF4*) were specifically expressed in C14, whereas
448 *PAX5*, *IRF8*, and *BACH2* were specifically expressed in C13 (Figure 4C). *XBP1* was

449 widely expressed in both clusters and upregulated in C14. C14 comprised the
450 progenitors of plasma cells, whereas C13 represented classical B progenitor cells
451 (Table 1, Figure 4C). These data indicate that the development and differentiation of
452 B cells occur during the early progenitor stages. Plasma cells are thought to represent
453 the final stage of B cell proliferation and are incapable of further proliferation [31].
454 However, our scRNA data confirmed the existence of plasma progenitor cells in
455 whole blood samples. The immunoglobulin variable (IGV) genes were expressed
456 primarily by the progenitors of C14 cells, with rearrangements in the *IGHV*, *IGKV*,
457 and *IGLV* genes having already occurred during the progenitor state of plasma cells.
458 These six TFs, *PRDM1*, *IRF4 PAX5*, *IRF8*, *BACH2,* and *XBP1,* play critical roles in
459 the cell fate decisions of B lineage cells. HOPX was significantly highly expressed
460 in progenitor and mature NK cells but was not expressed in MkE lineages,
461 suggesting that HOPX was a crucial TF for NK cell development (Figure 4C).

462 Monocyte progenitor cells in C10 were identified and redefined. One of the most
463 highly expressed genes in monocyte lineages was *IFI30* (Figure 2F), which is related
464 to the interferon-gamma signaling pathway [32]. *SPI1*, *PECAM1*, and *LST1* were
465 widely expressed in most monocyte progenitors, suggesting that *SPI1* and *LST1* are
466 critical genes for monocyte progenitors. The monocyte progenitors were controlled by
467 various cell-fate decision-associated TFs, including *KLF4, EGR1, MAFB, CEBPB,*
468 *BCL6, IRF8*, and *DDIT4* (Table 1 and Figure 4C). To identify and validate the fate
469 decision genes among the two monocyte cell types (macrophages and DCs), TF
470 expression patterns were reanalyzed in the control cohorts. The macrophage cells
471 expressed high levels of *EGR1*, *CEBPB*, and *MAFB* but did not express *IRF8* or
472 *DDIT4* (Figure 4D and S5). Furthermore, the macrophages included FCGR3A^{+} cells
473 and showed high levels of CSF1R and TNF expression. The DC cells positively
474 expressed *CLEC9A, ANPEP*, and *THBD* (Table 1) and highly expressed *IRF8, BLC6*,
475 and *DDIT4* (Figure S5). The fate decision of macrophages and DCs were determined
476 by these TFs.

477 Validation was conducted in three cohorts of bone marrow cells, control PBMCs,
478 and enriched lineages (Figure 4D and S5). We analyzed the expression

479 characteristics of marker genes and identified TFs of progenitors in control cohorts.
480 MAFB was uniquely expressed in mature macrophages, TBX21 was uniquely
481 expressed in mature NK cells, and PRDM1 was uniquely expressed in mature
482 plasma cells (Figure S5). These findings were highly consistent with the
483 fate-decision identification in redefined progenitor clusters. Results showed that the
484 cell-fate decision-associated TFs were uniquely expressed in their respective
485 differentiated mature cell types. Furthermore, these cell-fate decision-associated TFs
486 showed more accuracy and efficiency in mature immune cell classification (Figure
487 S5). Overall, the characteristics of cell-fate decision-associated TFs in lymphoid,
488 monocyte, and MEP redefined progenitor lineages were highly consistent with the
489 mature cell types. Cell-fate decision-associated TFs in each branch of the
490 hematopoietic hierarchy were clearly recovered.
491
492 ***Signaling pathways associated with commitment in each lineage***
493 Diverse pathways affect and regulate HPC differentiation. Our analysis showed
494 that most progenitors expressed high levels of *JUN/FOS* (*AP-1*), *JUNB*, and *JUND*
495 (Figure 2C and 3F). These data indicate that the JUN-related signaling pathway is
496 critical for the differentiation, mobilization, and maintenance of HPCs. MkE clusters
497 expressed high levels of *DEPTOR* (Figure S6), which negatively regulates the
498 mTORC1 and mTORC2 signaling pathways [33], indicating that the mTOR signaling
499 pathway should be suppressed during MkE maturation and differentiation. *NFKBIA*
500 was widely expressed in >80% of the progenitor cells, whereas *NFKBIZ* was highly
501 expressed in the progenitors of monocytes and macrophages (Figure S6). *NFKBIZ* is a
502 member of the NF-κB inhibitor family [34], suggesting that the NF-κB signaling
503 pathway may be suppressed during monocyte and macrophage maturation and
504 differentiation.
505 Gene set enrichment analysis of clusters using gene set variation analysis
506 revealed the pathways and signals that affect the differentiation of each branch and
507 progenitor subgroup (Figure 5A). The transforming growth factor-beta (TGF-β)
508 signaling pathway was enriched in the original progenitor cells (C0, MPCs),

509 demonstrating its importance in maintaining the quiescent state of HPCs [35].
510 Moreover, the Notch, interferon alpha and gamma response, TGF-β, and TNF-α
511 signaling pathways were enriched, indicating their potential importance in
512 maintaining monocyte progenitors (Figure 5A). However, expression of the
513 TNF-signaling gene set was suppressed in MEP progenitors (Figure 5A). A previous
514 study has revealed the novel role of Notch in regulating monocyte survival during
515 hematopoiesis [36]. Heme-metabolism and hedgehog-pathway gene sets were
516 enriched in erythroid progenitors (GAPs). Activation of TGF-β and the hedgehog
517 pathway gradually decreased during the committed stage of erythroid progenitors
518 (MEPs), whereas oxidative phosphorylation increased. The NOTCH signaling
519 pathway was highly enriched in MBPs. Hallmark cell proliferation-related gene sets
520 related to the G2M checkpoint, E2F targets, MYC targets (V1 and V2), and the DNA
521 repair pathway (defined by the molecular-signature database) were highly enriched
522 among erythroid and neutrophilic progenitors (Figure 5A). The presence of cell
523 proliferation signaling in erythroid (C6) and neutrophil (C5) progenitors indicates that
524 these cells completed differentiation and tended to proliferate. The IL6/JAK/ STAT3
525 signaling pathway was enriched in CLPs. The Wnt/β-catenin signaling pathway was
526 enriched in T cell progenitors (C12), which was previously reported to regulate
527 peripheral T cell activation and migration [37]. Moreover, interferon-response gene
528 sets were enriched in B cells and monocyte progenitors.
529
530
531 ***Functional validation and primary culture confirmation of the reliability of the***
532 ***redefinition of HPCs***
533 The functions of HSC-fate decision-associated genes have been validated
534 previously [14, 38]. We hypothesized that the phenotypic results of functional studies
535 would be consistent with or explained by our HPC definitions. Furthermore, the
536 reformed model of the hierarchical organization of hematopoiesis can be used to
537 interpret conflicting or unclear results. For example, KLF1 and GATA1 are key
538 cell-fate decision-associated TFs involved in erythroid and megakaryocytic

539 differentiation [26, 39]. Functional studies have confirmed that GATA1 progenitors
540 generate mast cells, basophils, megakaryocytes, and erythroid cells [26, 40], while
541 their absence results in the production of monocytes, neutrophils, and lymphocytes
542 [26]. Our transcriptional atlas revealed that GATA1 was uniquely expressed in the
543 MEP but not in monocyte, neutrophil, or lymphocyte progenitors (Figure 2B). These
544 expression characteristics explain the functional study results, indicating that the
545 absence of GATA1 halts MkE lineage differentiation. Conversely, the absence of
546 GATA1 causes progenitor cells to differentiate into NMPs and monocytes. In primary
547 cultures, HPCs differentiated into erythroid and mast cells under the same *in vitro*
548 conditions, utilizing IL3 and SCF; this indicates that both cell types originated from
549 the same lineage. These results confirmed the reliability of the reformed
550 hematopoietic hierarchy of the MkE lineage.

551 Macrophage colony-stimulating factor (M-CSF) is critical for macrophage
552 survival during *in vitro* primary cultures [41]. This observation supports our
553 transcriptional atlas findings, in which the M-CSF receptor, CSF1R, was only
554 expressed in monocyte progenitors and mature macrophages (Figure 5B). Stem cell
555 growth factor (SCGF), also known as CLEC11A, is a secreted glycoprotein that
556 exhibits erythroid burst-promoting activity in the absence of the SCF/c-Kit ligand [42,
557 43]. Our results showed that SCGF was highly expressed in NMPs and MEPs,
558 whereas CSF3R, a G-CSF receptor, was highly expressed in NMPs (Figure S6).
559 SCGF cooperates with GM-CSF or G-CSF and enhances the growth of
560 granulocyte-macrophage colonies; however, if only SCGF is present in the culture,
561 erythroid-burst activity is promoted [42, 43]. The expression characteristics of CSF3R
562 and G-CSF strongly suggest that SCGF, together with G-CSF, activates and induces
563 the differentiation of neutrophil and macrophage progenitors, suppressing MkE
564 lineage generation. The expression of the IL15 receptor IL2RB in NK progenitors was
565 unique (Figure 5B). The common beta chain receptor CSF2RB of IL3, IL5, and
566 GM-CSF was uniquely expressed in erythroid lineage progenitors. Results based on
567 the primary HPC cultures showed that progenitors could differentiate into a higher
568 percentage of $CD235A^{+}$ cells after treatment with IL-5, erythropoietin (EPO), and

569 SCF than with IL3 (Figure 5C). The expression patterns of IL3RA and CSF2RB in
570 our atlas revealed that IL3 can non-specifically bind to both receptors and induce
571 progenitor differentiation into other types of lineages (Figure 5B). Moreover, flow
572 cytometric results showed that the expression of the identified marker *SLC40A1*
573 gradually decreased and that its expression became more unique throughout erythroid
574 differentiation (Figure 5C). These results validated the expression patterns of markers
575 unique to MkE lineages. In addition, IL7R, an essential driver of T-cell development
576 and homeostasis, was solely expressed in T-cell progenitors, whereas *IL2RG* and *IL1B*
577 were expressed across most progenitor subgroups (Figure 5B). IL18 was identified as
578 one of the most essential cytokines in early progenitors (data not shown), playing a
579 critical role in hematopoiesis. *IL1B* transcription was regulated by the TFs *SPI1* and
580 *CEBPB* [44], which were highly expressed in monocyte progenitors (Table 1).
581 Moreover, *ALDH1A1* was highly expressed in erythroid-lineage progenitors (Figure
582 S6), the product of which has previously been shown to participate in retinoic acid
583 synthesis. Retinoic acid is required for progenitor cell maintenance [45]. The
584 expression characterization of these cytokine receptors and genes explains the
585 importance of cytokine ligands in cell survival and proliferation within *in vitro*
586 primary cultures. In summary, the expression characteristics of cell-fate
587 decision-associated TFs and cytokine receptors support that the redefinition of
588 progenitor lineages is verifiable, reliable, and reasonable (Figure 3A).
589
590 ***Are the definition and hierarchy reliable?***
591 *In vivo* or *in vitro* functional studies can be affected by various conditions, which
592 can't accurately reflect the true physiological state of original progenitors. And colony
593 forming assay also exhibits obviously defects due to variations in cytokines, samples,
594 and selection methods (see introduction). Most scRNA studies that enrich bone
595 marrow progenitors using flow cytometry enrichment demonstrate inefficient
596 reproducibility, false positives, and subjectivity definitions [2, 12]. Some
597 CD34-positive cells are transcriptionally negative, and subsets are either missing or
598 inconsistently distribution among samples. Fate mapping is nearly impossible for

599 adult original progenitors. Here, we present three aspects of evidence supporting the
600 reliability of our definition and hierarchy.
601
602 **Robustly Validating the Definition by a Reliable Atlas**
603 Here, we utilized original progenitors from adult peripheral blood, which faithfully
604 reflect the physiological state of human progenitors. We comprehensively captured
605 the entire differentiation stages of these physiological progenitors. The substantial
606 numbers of cells in each progenitor subset provided a reliable transcriptional atlas
607 (Table S1). By utilizing unsupervised data analysis, we avoided bias and subjective
608 effects on results. The subset distribution reproducibility across sample is exemplary
609 (Figure 2D). Furthermore, the expression patterns of some unique genes in each
610 redefined subset are highly consistent with conventional markers (Table 1). Based on
611 truly physiological original progenitors and a reliable atlas, the scRNA data robustly
612 validates the definition, and is more reliable than definition based on *in vivo* or *in*
613 *vitro* experiments.
614
615 **Straightforward Results Confirm the Reliability of the Hierarchy**
616 Our T-SNE plot clearly and perfectly displayed the three differentiation directions of
617 progenitors after the initial priming step of MPC (Figure 3A). The different stages of
618 the megakaryocytic–erythroid lineages are distinctly distributed in the same directions,
619 expressing key well-known fate decision genes (GATA1,KLF1,TAL1) (Figure
620 2B).The unique expression characteristics of these well-known genes
621 (GATA1,KLF1,TAL1) confirmed the definition of MEP lineages, and we are the first
622 to identify and define the intermediate stage of GAP(Figure 3A). The subsets of the
623 megakaryocytic–erythroid lineages are adjacent and lack clear boundaries, indicating
624 continuous differentiation trajectories. The gradual and continuous expression
625 characteristics of other well-known genes (GATA2,SLC40A1) also provide robust
626 evidence of the continuous lineage commitment. Furthermore, the unique expression
627 patterns of well-known receptor genes (EPO, MPL, CSF2RB) confirm the definition
628 of MEP lineages, which is consistent with *in vitro* culture. The termination of MEP

629 lineages differentiation yields two type of cells (mast and erythliod), which already
630 validated by numerous studies that mast and erythroid cell are generated under the
631 same culture conditions, and further validation is needn't.
632
633 **Physiological Evidence**
634 The megakaryocytic–erythroid lineage commitment process is continuous, aligning
635 with the constant production of erythroid cells every second in the blood. While the
636 commitment process in the lymphoid lineage is discrete and stepwise, reflecting the
637 obviously increase in numbers and heightened activity of lymphoid cells specifically
638 in response to infections or pathological conditions (adaptive immunity). The
639 continuous differentiation process of neutrophils is the shortest in the hematopoietic
640 hierarchy (only two stages) (Figure 3A, C), which also matching with physiological
641 phenomenon. Neutrophils are the primary and rapid responders during infections of
642 innate immunity and can quickly generate an enormous number of mature
643 neutrophils for immune defense [46]. The shortest differentiation trajectories of
644 neutrophils enable a rapid and robust response against infection. In summary, the
645 reformed hierarchy of hematopoiesis perfectly explains the physiological
646 phenomenon of immunity. Often, the truth is simple, yet we have been confused so
647 long time.
648
649
650 **Discussion**
651 Overall, we redefined the characteristics of the original progenitor lineages and
652 identified specific cell-fate decision-associated TFs and marker genes critical for
653 hematopoiesis. This study is the first to define the initial primed transitional
654 progenitor using reliable TFs and markers, effectively solving the century-old puzzle
655 concerning whether CMPs, GMPs, or LMPPs exist. In addition, we reformed the
656 comprehensive hierarchy of hematopoiesis that reconciled previous models and
657 explained the conflicting results among functional studies [4, 16, 47].
658 During the first commitment stage of hematopoiesis, the initial hematopoietic

659 progenitors are simultaneously primed into three transitional progenitor branches. Our
660 findings provide detailed insights into HPC lineages, revealing the molecular
661 mechanism underlying their commitment and multipotent differentiation. Our scRNA
662 atlas provides an integrated landscape of the heterogeneity, development, and
663 hierarchy of adult peripheral-blood progenitor cells. Collectively, the results of this
664 study represent significant breakthroughs in the fundamental discipline of
665 hematopoiesis, overcoming disciplinary challenges that have plagued the field for
666 more than half a century.

667 Currently, no consensus has been reached regarding the definition of all HPC
668 lineages, owing to inconsistent classifications and conflicting functional outcomes [9,
669 16]. Here, the comprehensive enrichment of the original HPCs in adult peripheral
670 blood was particularly helpful for establishing clear and reliable HPC definitions. We
671 demonstrated that classical surface markers, including CD38, KIT, THY1, CD49f, and
672 FLT3, were non-specific and inefficient for identifying HPC lineages [2, 12]. Hence,
673 the conventional lineage definition of HSCs using these markers requires
674 improvement. Accordingly, we redefined the lineages using reliable TFs and markers,
675 including initial primed transitional progenitors. Certain unique marker genes of the
676 redefined progenitor subgroups were highly consistent with previous characterizations,
677 such as *KLF1*, *HPGDS*, and *MPO* for MEPs, mast cells, and neutrophils, respectively
678 [22, 48, 49]. We redefined most HPC lineages using unsupervised and inartificial
679 clustering for the first time and provided an unambiguous transcription atlas for each
680 type.

681 Through our transcriptional atlas, we identified the key fate decision-associated
682 factors at each hematopoietic hierarchy level. The activation of key TFs primes
683 progenitor cells, initiating them to differentiate into downstream-specific committed
684 lineages [50]. The expression of TFs during the original and mature stages of
685 progenitors revealed a basic co-segregation between myeloid–lymphocyte and
686 mast–erythroid development. Additionally, during the later stage, MEPs differentiated
687 into two directions owing to changes in the expression of key TFs (*KLF1*, *LMO4*, and
688 *GATA2*). Our findings highlight the importance of TFs and the characterization of

689 genes involved in hematopoietic-lineage fate decisions, providing novel insights into
690 the molecular mechanisms underlying the hematopoietic hierarchy and addressing the
691 inconsistencies among the *in vitro* and *in vivo* studies of progenitor cells [26]. The
692 expression characteristics of known cell-fate decision-associated TFs, such as *NFE2*,
693 *TAL1*, and *LYL1* [51-53], provide complementary evidence that the redefinition of
694 progenitor lineages is verifiable, reliable, and reasonable. Furthermore, we found that
695 changes in the expression of unique TFs were independent and strictly
696 lineage-affiliated, contradicting previous conclusions [54]. Previously, induced
697 pluripotent stem cells were established by activating four TFs [55], and *GATA1* and
698 *TAL1* have been validated in reprogramming megakaryocytes and erythroid cells [56,
699 57]. To our knowledge, this is the first study to clearly identify the cell-fate
700 decision-associated TFs in each branch of the adult peripheral-blood hematopoietic
701 hierarchy. In turn, precise control of the priming and differentiation of HPCs can be
702 achieved by artificially and precisely controlling these key TFs and receptor signals,
703 which we refer to as the Artificial Operation of HPCs.

704 CSF3R, the G-CSF receptor [58], was primarily expressed in the original clusters
705 (MPCs). However, the molecular mechanisms underlying G-CSF
706 induced-enhancement of hematopoiesis remain unclear. Our study reformed the
707 hematopoiesis hierarchy based on the scRNA analysis of unperturbed and
708 G-CSF-mobilized progenitor lineages in peripheral blood. G-CSF directly stimulated
709 downstream activation signals in HPCs by binding to CSF3R, resulting in the
710 conversion of quiescent-state HPCs into activated progenitors. These HPCs were then
711 primed and committed into various progenitors, after which they further differentiated
712 into mature myeloid and lymphocyte cells. The key TFs and cytokine regulators
713 synergistically controlled hematopoietic-lineage commitment, which was validated
714 using primary culture, lineage formation, and reprogramming studies [14]. However,
715 the mechanisms underlying lineage commitment raise certain questions regarding the
716 following concepts: how key TFs are activated during the lineage-commitment
717 process; whether the TFs are activated independently (internally) or via
718 cytokine-signaling cascades (externally); and which factors regulate the activation of

719 lineage-fate decision-associated factors (TFs and cytokine-signaling cascades) that
720 determine subsequent differentiation after the initial priming of HPC with G-CSF.
721 Generally, environmental factors and epigenetic modifications also play key roles.

722 The following three aspects strongly support the definition and classification of
723 HPCs as distinctly verifiable, reliable, and reasonable. First, the unique markers were
724 highly consistent with the characteristics of most progenitor subgroups. Second, the
725 redefinition addresses the conflicting results reported in previous functional studies
726 [26]. Third, the expression profiles of distinguished cytokine receptors precisely
727 explain the phenomena observed in primary cultures of progenitor subgroups, such as
728 CSF3R and CSF1R in neutrophil and monocyte cultures, respectively [59].

729 Recent findings have suggested that HSC lineage commitment is a continuous
730 process [2, 13], contradicting the classical stepwise hierarchy of hematopoiesis [60].
731 However, owing to HSC heterogeneity and the conflicting definitions of HSC
732 subgroups, the mechanism underlying this hierarchy is enigmatic, resulting in
733 inconsistent and incompatible established hematopoietic hierarchies [8]. Our elaborate
734 and comprehensive transcriptional atlas of HPCs provides accurate definitions and a
735 new biological perspective on hematopoiesis. The clear definition and characteristics
736 of HPCs showed that both stepwise and continuous processes occurred during lineage
737 commitment, which can be distinguished by the expression of unique TFs and marker
738 genes. Bias during HPC differentiation gradually accumulated through the activation
739 of multiple lineage-specific TFs, such as *KLF1* and *GATA2*, in MEP cells. The MkE
740 lineage commitment process is continuous, corresponding with the constant
741 production of erythroid cells in the blood [61]. In contrast, the lymphoid
742 lineage-commitment process is discrete and stepwise, reflecting the obvious increase
743 in numbers and heightened activity of lymphoid cells in response to infections or
744 pathological conditions (adaptive immunity). Neutrophils have the shortest
745 differentiation trajectories, which enable a rapid and robust response against infection.
746 These findings facilitate mapping the entire set of transcriptional signaling and
747 pathways involved in lineage commitment and provide strong evidence of a
748 heterogeneity gradient among lymphoid and MkE lineage progenitors.

749 Previous studies have proposed updated models of the human hematopoietic
750 hierarchy [4, 15]. The major difference between these models and our hierarchy is that
751 we clearly defined the first priming (three directions) and transitional stages (GAP) of
752 the commitment process. We confirmed that initial hematopoietic progenitors were
753 simultaneously primed into three transitional branches of megakaryocyte/erythroid,
754 lymphoid, and neutrophilic progenitors during the first differentiation stage of
755 hematopoiesis. LMPPs, GMPs, and CMPs likely represented a heterogeneous mixture
756 of transitional progenitors. CSF3R was positively expressed in the early stages of
757 erythroids (GAPs) but not in the lymphoid lineages (CLPs). The reformed
758 hematopoietic hierarchy further explains how diversified progenitor cells are primed,
759 induced, and differentiate into all lineages of immune-system cells. The HPC
760 transcriptional atlas provides a new perspective and an adequate explanation of how
761 the immune system is reconstituted in adults. This reliable redefinition and hierarchy
762 will clarify clonal hematopoiesis dynamics. Controlling the initial stage and
763 generation of cytokines, particularly by suppressing the initial cytokine-mediated
764 activation of cascade processes, could inhibit or reduce the effect of subsequent
765 cascade activation. This approach may be a novel therapeutic strategy to control the
766 excessive activation of cytokines and prevent cytokine storms [62, 63]. Thus, our
767 reformed taxonomy and hierarchy will improve immune monitoring in healthy and
768 diseased states and precise prevention and treatment.

769 This study has certain limitations. First, granulocyte, neutrophil, and monocyte
770 progenitors were identified as originating from the same myeloid progenitors [9].
771 However, our results showed that the cell-fate decision-associated TFs and
772 characteristic genes significantly differed between these cells, implying that the
773 original association between these lineages should be reconsidered. Our results further
774 showed that monocyte lineage-restricted progenitors and CLPs had similar expression
775 patterns. We speculate that monocyte progenitors could also originate from CLPs or
776 have an independent origin. Although we identified cell-fate decision-associated TFs
777 in two types of B progenitor cells, further investigations are required to determine
778 whether both types of progenitor cells were derived from the same or independent

original cells. Second, our enrichment method might have caused the loss of some intermediate-state progenitors, such as the committed megakaryocyte progenitor. Given that our entire atlas was established based on unsupervised dimensionality reduction and clustering from a statistical approach, its accuracy and efficiency remain imperfect. Artificial intelligence and deep learning could provide an accurate, reliable, and distinguishable definition for each state and all progenitor subgroups during hematopoiesis. Moreover, DNA barcoding and the clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated protein 9 (Cas9)-mediated genome-editing system could help evaluate the clonal dynamics of progenitor cells during hematopoiesis [64].

In conclusion, our results enabled us to reform the comprehensive hierarchy of hematopoiesis and identify key regulators that control hematopoiesis in human adult peripheral blood. Our new model accurately defines a hierarchy that provides biological insights into the hematopoiesis and lineage commitment of HPCs, aiding future research in cell engineering, regenerative medicine, and disease treatment. The reformed lineage-commitment results provide critical insights into the etiology and treatment of hematologic disorders. Moreover, this hematopoietic hierarchy has the potential to achieve artificial, precise, and controlled progenitor cell differentiation and targeted treatment of immune diseases.

Acknowledgments

We acknowledge Professor Chen Qiu, Rongchang Chen, and Yingyun Fu for supporting this research. Moreover, we are grateful to the participants for their involvement in this study. We would like to express our appreciation for the laboratory support provided by the Key Laboratory of Shenzhen Respiratory Diseases (ZDSYS201504301616234). We also thank Chi-Wai Man of the Chinese University of Hong Kong for manuscript revision. The study was supported by the National Natural Science Foundation of China (grant number 21807072).

Authors' contributions

Y.Y. conceived the original idea, designed the study, and performed HPC-enrichment experiments. Y.Y. and L.C. conducted the bioinformatics and statistical analyses of the sequencing data. Y.Y. and Q.S. wrote and revised the manuscript. Q.H., Y.F., and G.L. conducted the sample collection, and S.C generated the schematic diagram. Y.Y. and S.C. conducted the primary culture and laboratory work. All authors approved the final draft of the manuscript.

Declaration of interests

The authors declare no competing interests.

Declaration of Generative AI and AI-assisted technologies in the writing process

Not applicable

823 **Figure Legends**
824
825 **Figure 1. Progenitor subtypes identified in adult peripheral blood using**
826 **unsupervised dimensionality reduction and clustering analysis.**
827 A, Schematic showing the progenitor cell enrichment and sequencing workflow. B,
828 Mean percentage of CD34-positive cells in control (PBMC) and G-CSF mobilized
829 peripheral blood groups. C, Twenty-nine integrated clusters were identified via
830 UMAP analysis. D, Violin plot of the identified progenitor cell marker genes *CDK6*,
831 *EGFL7*, *CYTL1*, and *SOX4*. E, Expression-feature plots of the progenitor markers
832 *CD34* and *CYTL1*. F, Dot plot illustrating the expression of progenitor cell-related
833 marker genes in all clusters.
834
835 **Figure 2. Characterization of progenitor subtypes in** unperturbed **and activated**
836 **states.** A, UMAP illustrating 14 clusters of integrated $CD34^+$ progenitor cells. B,
837 Violin plots showing the expression of conventional surface markers and unique
838 marker genes, indicating that *GATA2*, *CSF3R*, and *MME* were more efficient and
839 accurate markers than the conventional surface markers, *KIT* and *FLT3*, for
840 classifying myeloid, granulocyte, and lymphoid progenitors. C, Dot plots showing the
841 expression of representative markers and lineage-specific genes. D, UMAP showing
842 the distribution of cells in the unperturbed (Control) and activated (G-CSF) states,
843 displaying the persistence of all types of progenitors in the quiescent state of adult
844 peripheral blood. E, Bar plot displaying the cell number and proportion changes in all
845 progenitor clusters. The primitive clusters (C0) exhibited the most pronounced
846 changes in terms of cell numbers and fold changes. F, Volcano plot showing the top
847 five genes with the highest fold-changes in expression among the progenitor clusters.
848 C0, cluster 0; UMAP, uniform manifold approximation and projection.
849
850
851 **Figure 3. Trajectories and dynamics of progenitor cell differentiation and the**
852 **activities of lineage-priming transcription factor regulons.** A, UMAP plots

853 demonstrating the redefined progenitor clusters, with commitment trajectories directly
854 characterized without requiring in-silico prediction; red arrows: three simultaneously
855 primed lineages. B, UMAP visualizing the pseudo-time trajectory scores of progenitor
856 cells, which indicates three developmental paths from MPCs. C, UMAP feature plots
857 of the states of pseudo-time trajectories analysis, showing that the CLP lineages were
858 in state one (red), whereas the MkE lineages were in state three (blue). D, Plots
859 showing the relative expression of GATA2 and HOPX across the pseudo-time
860 trajectory. E, Heatmap showing the three distinct lineage marker gene expression
861 patterns along the pseudo-time trajectory. F, Heatmap showing the differential
862 activation of transcription factor regulons for all progenitor subgroups. G,
863 Transcriptional activity of lineage signature genes related to different differentiation
864 activation. UMAP, uniform manifold approximation and projection.
865
866
867 **Figure 4. Reformed hematopoietic hierarchy of progenitor cell differentiation.** A,
868 Reformation of the **hematopoietic hierarchy** based on an scRNA transcriptional atlas.
869 In the new hematopoietic hierarchy, the MPCs are simultaneously primed into three
870 transitional progenitor cell types, which subsequently mature into diverse lineages.
871 The first progenitor cell type is the CLPs, which differentiate stepwise into T, B, and
872 NK progenitor cells. The second progenitor cell type, GAPs, initially undergoes
873 continuous differentiation into intermediate-stage MEPs through processes controlled
874 by GATA1 and GATA2. During the later stages, MEPs differentiate into two
875 directions, controlled by KIT, KLF1, and LMO4. The high-expression branch
876 differentiates into MBPs. The other branch differentiates into Pro-MEs. The third type
877 of progenitor cells are NMPs, which differentiate into neutrophils and monocytes. The
878 dashed lines represent partial lineage connections that lack strong direct evidence of
879 differentiation relationships. The shaded area indicates the GATA2 and GATA1
880 expression domains. B, Feature plots illustrating the expression patterns of GATA2,
881 HOPX, and CSF3R in co-segregated progenitors. C, Dot plot illustrating the
882 expression of cell-fate decision-associated TF genes in enriched CD34 progenitors. D,

883 Dot plot illustrating the expression of mature cell marker genes and cell-fate
884 decision-associated TF genes in enriched lineage cohorts. CLP, common lymphoid
885 progenitors; GAP, gene-controlled progenitors; MBP, mast cell, basophil, or
886 eosinophil progenitors; MEP, megakaryocytic–erythroid progenitors; MPC,
887 multipotent progenitor cells; NMP, neutrophilic and monocytic progenitors; Pro-ME,
888 megakaryocytic and erythroid progenitors; scRNA, single-cell RNA; UMAP, uniform
889 manifold approximation and projection.
890
891
892 **Figure 5.** Gene set enrichment analysis **and flow cytometry validation of the**
893 **redefined hematopoietic progenitor cells.** A, Heatmap plot showing gene set
894 enrichment analysis results for GSVA clusters. B, Violin plot illustrating the
895 expression of receptor genes. The receptors encoded by *CSF1R*, *CSF2RB*, and *MPL*
896 were specifically expressed in the redefined progenitor subgroups. C, D, Flow
897 cytometric analysis demonstrating the frequency of $CD235A^{+}$ cells and the
898 distribution of $SLC40A1^{+}$ cells during 1 (C) or 2 weeks (D) of primary culture.
899 Higher percentages of $CD235A^{+}$ cells were detected in the IL-5 groups than in the IL3
900 groups. The plots illustrate the dynamic change in the distribution of $SLC40A1^{+}$ cells,
901 which was more efficient than that of CD71 during distinct erythrocyte cell stages in
902 *in vitro* culture. GAPs, gene-controlled progenitors; GSVA, gene set variation analysis;
903 HPCs, hematopoietic progenitor cells; MEPs, megakaryocytic–erythroid progenitors.
904
905
906 **Table 1. Specific marker genes and genes associated with cell-fate decisions in all**
907 **redefined subgroups.**
908
909

910 **Methods**

911 ***Sample collection and HPC enrichment***

912 We obtained original peripheral blood HPCs from quiescent donors who had not
913 received any treatment (11 individuals) and mobilized (G-CSF) donors (three
914 individuals). PBMCs and total white blood cells were obtained from 8 mL of
915 peripheral blood according to standard protocols via density-gradient centrifugation
916 with Ficoll® (Cytiva, GE Healthcare). Since lineage commitment is a continuous
917 process, samples collected at any time point can be used to capture all progenitor
918 stages in peripheral blood. Prior to blood collection, mobilized donors received 300
919 μg of subcutaneous recombinant human G-CSF at 12 h intervals for 5 days.
920 Subsequently, 10 mL of peripheral blood was collected in heparinized tubes from each
921 stem cell donor. Lineage (Lin−, CD3−CD19−CD56−) negative cells were removed
922 using the RosetteSep Enrichment Kit (STEMCELL Technologies, 15272) to enrich
923 the progenitors. The self-designed negative deletion method was applied to enrich the
924 HPCs, enabling the successful capture of the entire spectrum of original HPCs with a
925 sufficient number of cells. Ethical approval was obtained from the ethics committee
926 of the Shenzhen People's Hospital in adherence with the ethical standards
927 recommended in the 1964 Declaration of Helsinki. Informed consent was obtained
928 from each study participant.

929

930 ***ScRNA preparation and sequencing***

931 RNA sequencing was performed with human-enriched HPCs (Lin−) purified from 14
932 donors, including progenitor cells mobilized with G-CSF. Each single-cell suspension
933 was loaded, captured using chromium microfluidic chips, and barcoded with a 10×
934 Chromium Controller (10x Genomics). The barcoded scRNA was reverse-transcribed,
935 and sequencing libraries were constructed using reagents from the Chromium Next
936 GEM Single Cell 5′ Reagent Kit v2, according to the manufacturer's instructions (10x
937 Genomics). Paired-end sequencing (150 base pairs) was performed using an Illumina
938 NovaSeq 6000 System.

939

940 ***ScRNA sequencing data processing***

941 The raw sequencing reads of all samples were demultiplexed and aligned to the
942 reference genome hg38 via the Cellranger pipeline. All downstream single-cell
943 unsupervised dimensionality-reduction and clustering functions were performed using
944 Seurat [65]. Progenitor definition and annotation were conducted based on our custom
945 standards without the use of software. GO enrichment analysis of marker genes was
946 performed using the cluster profiler package in R software. The lineage priming
947 trajectories were directly confirmed by progenitor characterization without the need
948 for in-silico prediction. Complementary trajectory analysis was performed using
949 Monocle 2 and Monocle 3 [66]. TF regulation inference and gene regulatory network
950 analyses were performed using Single-Cell rEgulatory Network Inference and
951 Clustering (SCENIC) [67]. Gene set variation enrichment analysis was performed
952 using a non-parametric and unsupervised method of gene set variation analysis [68].

953

954 ***Integrated analysis with an online single-cell dataset***

955 The online scRNA datasets GSE120221 [69], GSE117498 [12], GSE145668 [70], ,
956 and GSE181989 [71] were downloaded from the Genomic Spatial Event (GSE)
957 database. Integrated analysis was performed using Seurat. The bone marrow scRNA
958 datasets of 136,489 cells from GSE120221, GSE117498, GSE145668, GSE137864,
959 GSE193138, GSE116256, and GSE181989 were merged. Similarly, the PBMC
960 scRNA datasets of 33,719 cells from GSE226488, GSE168732, and CNP0001102
961 were merged.

962

963 ***Cytokine-based primary culture of HPCs***

964 $CD34^+$ mononuclear cells were separated from human cord blood samples using
965 Ficoll® (Cytiva, GE Healthcare) and enriched using RosetteSep Human
966 Hematopoietic Progenitor Cell Enrichment Cocktail (STEMCELL Technologies,
967 15026). Purified $CD34^+$ cells were induced to generate MkE-lineage cells in a
968 cytokine-based primary culture. Purified $CD34^+$ cells were cultured in StemSpan
969 SFEM (STEMCELL Technologies, 09605) for 21 days at 37 °C with 5% fetal bovine

970 serum under the following two conditions: (1) 10 ng/mL human SCF (PeproTech), 20
971 ng/mL EPO (MCE), and 10 ng/mL IL3 (MCE); (2) 10 ng/mL human SCF, 20 ng/mL
972 EPO, and 10 ng/mL IL5 (MCE). Cultures were supplemented with 1×
973 insulin–transferrin–selenium–ethanolamine (ProCell) and
974 penicillin–streptomycin–glutamine (Life Technologies). After 2 weeks, human IL-3
975 and SCF were withdrawn from the culture medium, and an enucleation culture was
976 performed for each group.
977
978 ***Flow cytometric analysis of erythrocytes***
979 Cultured cells were counted using a cytometer (Count Star) before and after culture.
980 Flow cytometric analysis of surface markers was conducted weekly. Cells were
981 stained with anti-human CD34-phycoerythrin (1:100), anti-human
982 CD235a-fluorescein isothiocyanate, anti-human SLC40A1–Alexa Fluor 647, and
983 anti-human CD71-adenomatous polyposis coli (BioLegend, 1:100) antibodies,
984 followed by analysis using a flow cytometer (Beckman). The antibodies used are
985 presented in Table S2. The results were analyzed using the CytExpert software.
986
987 ***Glucose-6-phosphate dehydrogenase (G6PD)-activity assay***
988 G6PD enzyme activity in primary erythrocyte cells was detected using the G6PD
989 Activity Assay Kit (GD9650, Leadman, Beijing, China). Suspensions of differentiated
990 $CD34^+$ cells were centrifuged at room temperature (25 °C) and counted using a
991 cytometer. Next, the cells were washed with phosphate-buffered saline, resuspended
992 in 200 μL of distilled water, and incubated for 30 min at room temperature to lyse the
993 differentiated erythrocytes and release G6PD. Thereafter, G6PD activity was detected
994 using a biochemical analyzer (Leadman) according to the manufacturer's instructions
995 (cobas c 702, Roche).
996
997 **Supplemental information: titles and legends**
998 Figure S1. Specific expression of genes by HPCs and marker gene enrichment
999 analysis. A, Expression-feature plots of the progenitor markers CD34, HLF and

1000 EGFL7, showing that HLF expression was minimal but EGFL7 expression was
1001 highly consistent with that of CD34. B, Expression-feature plots of the progenitor
1002 markers CD34 and CRHBP, displaying CRHBP as mainly expressed in the early
1003 stages of progenitors. C, Gene Ontology enrichment results for progenitor cells. D,
1004 Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment results of signaling
1005 pathways in progenitor clusters. E, Gene set enrichment analysis of differentially
1006 expressed genes in hematopoietic cell lineages.
1007
1008
1009 Figure S2. Efficiency evaluation and comparison of canonical and newly identified
1010 marker genes across datasets. A, UMAP showing the T-SNE plot for the integrated
1011 progenitor cells of the aggregated HPC subsets. B, Violin plot showing
1012 gene-expression characterization of newly identified genes in HPCs, and displaying
1013 CD34-negative cells clustered in the same subgroup, which indicates that these cells
1014 are not progenitor cells. The expression patterns of EGFL7, CYTL1, LAPTM4B, and
1015 NPR3 are highly consistent with those of CD34 and are more efficient and sensitive
1016 than CRHBP in progenitor cell identification. C, UMAP showing the T-SNE plot of
1017 the integrated progenitor cells of bone marrow HPC scRNA cohorts. D, Violin plots
1018 displaying gene-expression levels in the bone marrow scRNA datasets. HPC,
1019 hematopoietic progenitor cell; scRNA, single-cell RNA sequencing. T-SNE:
1020 t-Distributed Stochastic Neighbor Embedding; UMAP, uniform manifold
1021 approximation and projection.
1022
1023
1024 Figure S3. Re-clustering of early priming-stage clusters of adult peripheral blood
1025 HPCs. A, UMAP showing the distribution of each individual atlas, demonstrating
1026 exemplary reproducibility across samples. B, UMAP showing the integrated
1027 progenitor cells (C0,C2,C3,C4, and C5 of Figure 2A); a common intermediate
1028 progenitor stage could not be distinguished among the myeloid lineages. C, Violin
1029 plots for early priming-stage progenitors showing no distinct differentiation

1030 characteristics among clusters. CMPs could not be distinguished by FLT3, CD38, and
1031 IL3RA expression, indicating that CMPs, GMPs, and LMPP were the heterogeneous
1032 mixture of various progenitor types. UMAP, uniform manifold approximation and
1033 projection.
1034
1035
1036 Figure S4. Pseudo-time analysis among MPC nearby clusters using Monocle2,
1037 excluding far away cell clusters. A, Differentiation trajectory of HPCs constructed via
1038 pseudo-time analysis. B, Expression trajectory of GATA2, HOPX, and CSF3R along
1039 the in-silico pseudo-time trajectories. C, Location of colored clusters in the
1040 reconstructed in-silico pseudo-time trajectory tree. D, Pseudo-time analysis showing
1041 the three states of the trajectory. E-F, Expression of MEP (E) and CLP (F) marker
1042 genes in clusters and states along pseudotime trajectories.
1043
1044 Figure S5. Comparison of gene expression levels across control datasets. Bone
1045 marrow and PBMC scRNA datasets were merged. A, T-SNE plot showing the
1046 integrated merged bone marrow scRNA datasets. B, Dot plot illustrating gene
1047 expression levels of cell marker genes and cell-fate decision-associated TF genes in
1048 bone marrow cohorts. C, T-SNE plot showing the integrated merged control PBMC
1049 scRNA datasets. D, Dot plot illustrating cell marker gene and cell-fate
1050 decision-associated TF gene expression levels in bone marrow progenitors. PBMC,
1051 peripheral blood mononuclear cell; scRNA, single-cell RNA.
1052
1053 Figure S6. Violin plot illustrating the expression levels of genes in enriched CD34
1054 positive clusters.
1055
1056 Table S1. Cell numbers in each cluster among the hematopoietic progenitor cell (HPC)
1057 datasets
1058
1059 **References**

1060 1 **Cheng, H., Zheng, Z. and Cheng, T.,** New paradigms on hematopoietic stem cell
1061 differentiation. *Protein Cell* 2020. **11**: 34-44.
1062 2 **Velten, L., Haas, S. F., Raffel, S., Blaszkiewicz, S., Islam, S., Hennig, B. P., Hirche, C., Lutz, C.,**
1063 **Buss, E. C., Nowak, D., Boch, T., Hofmann, W. K., Ho, A. D., Huber, W., Trumpp, A., Essers, M.**
1064 **A. and Steinmetz, L. M.,** Human haematopoietic stem cell lineage commitment is a
1065 continuous process. *Nat Cell Biol* 2017. **19**: 271-281.
1066 3 **Paul, F., Arkin, Y., Giladi, A., Jaitin, D. A., Kenigsberg, E., Keren-Shaul, H., Winter, D.,**
1067 **Lara-Astiaso, D., Gury, M., Weiner, A., David, E., Cohen, N., Lauridsen, F. K., Haas, S.,**
1068 **Schlitzer, A., Mildner, A., Ginhoux, F., Jung, S., Trumpp, A., Porse, B. T., Tanay, A. and Amit, I.,**
1069 Transcriptional Heterogeneity and Lineage Commitment in Myeloid Progenitors. *Cell* 2015.
1070 **163**: 1663-1677.
1071 4 **Drissen, R., Thongjuea, S., Theilgaard-Mönch, K. and Nerlov, C.,** Identification of two distinct
1072 pathways of human myelopoiesis. *Science Immunology* 2019. **4**.
1073 5 **CD-REFRanzoni, A. M., Tangherloni, A., Berest, I., Riva, S. G., Myers, B., Strzelecka, P. M., Xu,**
1074 **J., Panada, E., Mohorianu, I., Zaugg, J. B. and Cvejic, A.,** Integrative Single-Cell RNA-Seq and
1075 ATAC-Seq Analysis of Human Developmental Hematopoiesis. *Cell Stem Cell* 2021. **28**: 472-487
1076 e477.
1077 6 **de Morree, A. and Rando, T. A.,** Regulation of adult stem cell quiescence and its functions in
1078 the maintenance of tissue integrity. *Nature Reviews Molecular Cell Biology* 2023. **24**: 334-354.
1079 7 **Akashi, K., Traver, D., Miyamoto, T. and Weissman, I. L.,** A clonogenic common myeloid
1080 progenitor that gives rise to all myeloid lineages. *Nature* 2000. **404**: 193-197.
1081 8 **Jacobsen, S. E. W. and Nerlov, C.,** Haematopoiesis in the era of advanced single-cell
1082 technologies. *Nat Cell Biol* 2019. **21**: 2-8.
1083 9 **Laurenti, E. and Gottgens, B.,** From haematopoietic stem cells to complex differentiation
1084 landscapes. *Nature* 2018. **553**: 418-426.
1085 10 **Manz, M. G., Miyamoto, T., Akashi, K. and Weissman, I. L.,** Prospective isolation of human
1086 clonogenic common myeloid progenitors. *Proc Natl Acad Sci U S A* 2002. **99**: 11872-11877.
1087 11 **Xie, X., Liu, M., Zhang, Y., Wang, B., Zhu, C., Wang, C., Li, Q., Huo, Y., Guo, J., Xu, C., Hu, L.,**
1088 **Pang, A., Ma, S., Wang, L., Cao, W., Chen, S., Li, Q., Zhang, S., Zhao, X., Zhou, W., Luo, H.,**
1089 **Zheng, G., Jiang, E., Feng, S., Chen, L., Shi, L., Cheng, H., Hao, S., Zhu, P. and Cheng, T.,**
1090 Single-cell transcriptomic landscape of human blood cells. *Natl Sci Rev* 2021. **8**: nwaa180.
1091 12 **Pellin, D., Loperfido, M., Baricordi, C., Wolock, S. L., Montepeloso, A., Weinberg, O. K., Biffi,**
1092 **A., Klein, A. M. and Biasco, L.,** A comprehensive single cell transcriptional landscape of
1093 human hematopoietic progenitors. *Nat Commun* 2019. **10**: 2395.
1094 13 **Buenrostro, J. D., Corces, M. R., Lareau, C. A., Wu, B., Schep, A. N., Aryee, M. J., Majeti, R.,**
1095 **Chang, H. Y. and Greenleaf, W. J.,** Integrated Single-Cell Analysis Maps the Continuous
1096 Regulatory Landscape of Human Hematopoietic Differentiation. *Cell* 2018. **173**: 1535-1548
1097 e1516.
1098 14 **Sugimura, R., Jha, D. K., Han, A., Soria-Valles, C., da Rocha, E. L., Lu, Y. F., Goettel, J. A.,**
1099 **Serrao, E., Rowe, R. G., Malleshaiah, M., Wong, I., Sousa, P., Zhu, T. N., Ditadi, A., Keller, G.,**
1100 **Engelman, A. N., Snapper, S. B., Doulatov, S. and Daley, G. Q.,** Haematopoietic stem and
1101 progenitor cells from human pluripotent stem cells. *Nature* 2017. **545**: 432-438.
1102 15 **Zheng, S., Papalexi, E., Butler, A., Stephenson, W. and Satija, R.,** Molecular transitions in
1103 early progenitors during human cord blood hematopoiesis. *Mol Syst Biol* 2018. **14**: e8041.

1104 16 **Haas, S., Trumpp, A. and Milsom, M. D.,** Causes and Consequences of Hematopoietic Stem
1105 Cell Heterogeneity. *Cell Stem Cell* 2018. **22**: 627-638.
1106 17 **Zhu, J. and Emerson, S. G.,** Hematopoietic cytokines, transcription factors and lineage
1107 commitment. *Oncogene* 2002. **21**: 3295-3313.
1108 18 **Subramanian, S., Thoms, J. A. I., Huang, Y., Cornejo-Páramo, P., Koch, F. C., Jacquelin, S.,**
1109 **Shen, S., Song, E., Joshi, S., Brownlee, C., Woll, P. S., Chacon-Fajardo, D., Beck, D., Curtis, D.**
1110 **J., Yehson, K., Antonenas, V., O'Brien, T., Trickett, A., Powell, J. A., Lewis, I. D., Pitson, S. M.,**
1111 **Gandhi, M. K., Lane, S. W., Vafaee, F., Wong, E. S., Göttgens, B., Alinejad-Rokny, H., Wong, J.**
1112 **W. H. and Pimanda, J. E.,** Genome-wide transcription factor–binding maps reveal cell-specific
1113 changes in the regulatory architecture of human HSPCs. *Blood* 2023. **142**: 1448-1462.
1114 19 **Karsunky, H., Inlay, M. A., Serwold, T., Bhattacharya, D. and Weissman, I. L.,** Flk2+ common
1115 lymphoid progenitors possess equivalent differentiation potential for the B and T lineages.
1116 *Blood* 2008. **111**: 5562-5570.
1117 20 **Beaudin, A. E., Boyer, S. W. and Forsberg, E. C.,** Flk2/Flt3 promotes both myeloid and
1118 lymphoid development by expanding non–self-renewing multipotent hematopoietic
1119 progenitor cells. *Experimental Hematology* 2014. **42**: 218-229.e214.
1120 21 **Karamitros, D., Stoilova, B., Aboukhalil, Z., Hamey, F., Reinisch, A., Samitsch, M., Quek, L.,**
1121 **Otto, G., Repapi, E., Doondeea, J., Usukhbayar, B., Calvo, J., Taylor, S., Goardon, N., Six, E.,**
1122 **Pflumio, F., Porcher, C., Majeti, R., Göttgens, B. and Vyas, P.,** Single-cell analysis reveals the
1123 continuum of human lympho-myeloid progenitor cells. *Nature Immunology* 2017. **19**: 85-97.
1124 22 **Metzler, K. D., Fuchs, T. A., Nauseef, W. M., Reumaux, D., Roesler, J., Schulze, I., Wahn, V.,**
1125 **Papayannopoulos, V. and Zychlinsky, A.,** Myeloperoxidase is required for neutrophil
1126 extracellular trap formation: implications for innate immunity. *Blood* 2011. **117**: 953-959.
1127 23 **Schultz, J. and Kaminker, K.,** Myeloperoxidase of the leucocyte of normal human blood. I.
1128 Content and localization. *Archives of Biochemistry and Biophysics* 1962. **96**: 465-467.
1129 24 **Huang, Y., Sitwala, K., Bronstein, J., Sanders, D., Dandekar, M., Collins, C., Robertson, G.,**
1130 **MacDonald, J., Cezard, T., Bilenky, M., Thiessen, N., Zhao, Y., Zeng, T., Hirst, M., Hero, A.,**
1131 **Jones, S. and Hess, J. L.,** Identification and characterization of Hoxa9 binding sites in
1132 hematopoietic cells. *Blood* 2012. **119**: 388-398.
1133 25 **Hölig, K., Kramer, M., Kroschinsky, F., Bornhäuser, M., Mengling, T., Schmidt, A. H., Rutt, C.**
1134 **and Ehninger, G.,** Safety and efficacy of hematopoietic stem cell collection from mobilized
1135 peripheral blood in unrelated volunteers: 12 years of single-center experience in 3928 donors.
1136 *Blood* 2009. **114**: 3757-3763.
1137 26 **Drissen, R., Buza-Vidas, N., Woll, P., Thongjuea, S., Gambardella, A., Giustacchini, A.,**
1138 **Mancini, E., Zriwil, A., Lutteropp, M., Grover, A., Mead, A., Sitnicka, E., Jacobsen, S. E. W.**
1139 **and Nerlov, C.,** Distinct myeloid progenitor-differentiation pathways identified through
1140 single-cell RNA sequencing. *Nat Immunol* 2016. **17**: 666-676.
1141 27 **Cook, J. G., Chasse, D. A. D. and Nevins, J. R.,** The Regulated Association of Cdt1 with
1142 Minichromosome Maintenance Proteins and Cdc6 in Mammalian Cells. *Journal of Biological*
1143 *Chemistry* 2004. **279**: 9625-9633.
1144 28 **Li, D., Xue, W., Li, M., Dong, M., Wang, J., Wang, X., Li, X., Chen, K., Zhang, W., Wu, S.,**
1145 **Zhang, Y., Gao, L., Chen, Y., Chen, J., Zhou, B. O., Zhou, Y., Yao, X., Li, L., Wu, D. and Pan, W.,**
1146 VCAM-1(+) macrophages guide the homing of HSPCs to a vascular niche. *Nature* 2018. **564**:
1147 119-124.

29 **Zhou, W., Gao, F., Romero-Wolf, M., Jo, S. and Rothenberg, E. V.,** Single-cell deletion analyses show control of pro–T cell developmental speed and pathways by Tcf7, Spi1, Gata3, Bcl11a, Erg, and Bcl11b. *Science Immunology* 2022. **7**.

30 **van de Wetering, M., Oosterwegel, M., Dooijes, D. and Clevers, H.,** Identification and cloning of TCF-1, a T lymphocyte-specific transcription factor containing a sequence-specific HMG box. *The EMBO Journal* 1991. **10**: 123-132.

31 **Shapiro-Shelef, M. and Calame, K.,** Regulation of plasma-cell development. *Nature Reviews Immunology* 2005. **5**: 230-242.

32 **Phan, U. T., Arunachalam, B. and Cresswell, P.,** Gamma-Interferon-inducibleLysosomal Thiol Reductase (GILT). *Journal of Biological Chemistry* 2000. **275**: 25907-25914.

33 **Peterson, T. R., Laplante, M., Thoreen, C. C., Sancak, Y., Kang, S. A., Kuehl, W. M., Gray, N. S. and Sabatini, D. M.,** DEPTOR is an mTOR inhibitor frequently overexpressed in multiple myeloma cells and required for their survival. *Cell* 2009. **137**: 873-886.

34 **Totzke, G., Essmann, F., Pohlmann, S., Lindenblatt, C., Janicke, R. U. and Schulze-Osthoff, K.,** A novel member of the IkappaB family, human IkappaB-zeta, inhibits transactivation of p65 and its DNA binding. *J Biol Chem* 2006. **281**: 12645-12654.

35 **Blank, U. and Karlsson, S.,** TGF-β signaling in the control of hematopoietic stem cells. *Blood* 2015. **125**: 3542-3550.

36 **Ohishi, K., Varnum-Finney, B., Flowers, D., Anasetti, C., Myerson, D. and Bernstein, I. D.,** Monocytes express high amounts of Notch and undergo cytokine specific apoptosis following interaction with the Notch ligand, Delta-1. *Blood* 2000. **95**: 2847-2854.

37 **Staal, F. J. T., Luis, T. C. and Tiemessen, M. M.,** WNT signalling in the immune system: WNT is spreading its wings. *Nature Reviews Immunology* 2008. **8**: 581-593.

38 **Kucinski, I., Wilson, N. K., Hannah, R., Kinston, S. J., Cauchy, P., Lenaerts, A., Grosschedl, R. and Göttgens, B.,** Interactions between lineage‐associated transcription factors govern haematopoietic progenitor states. *The EMBO Journal* 2020. **39**.

39 **Siatecka, M. and Bieker, J. J.,** The multifunctional role of EKLF/KLF1 during erythropoiesis. *Blood* 2011. **118**: 2044-2054.

40 **Nei, Y., Obata-Ninomiya, K., Tsutsui, H., Ishiwata, K., Miyasaka, M., Matsumoto, K., Nakae, S., Kanuka, H., Inase, N. and Karasuyama, H.,** GATA-1 regulates the generation and function of basophils. *Proceedings of the National Academy of Sciences* 2013. **110**: 18620-18625.

41 **Brugger, W., Kreutz, M. and Andreesen, R.,** Macrophage Colony-Stimulating Factor Is Required for Human Monocyte Survival and Acts as a Cofactor for Their Terminal Differentiation to Macrophages In Vitro. *Journal of Leukocyte Biology* 1991. **49**: 483-488.

42 **Hiraoka, A., Yano Ki, K., Kagami, N., Takeshige, K., Mio, H., Anazawa, H. and Sugimoto, S.,** Stem cell growth factor: in situ hybridization analysis on the gene expression, molecular characterization and in vitro proliferative activity of a recombinant preparation on primitive hematopoietic progenitor cells. *Hematol J* 2001. **2**: 307-315.

43 **Hiraoka, A., Ohkubo, T. and Fukuda, M.,** Production of human hematopoietic survival and growth factor by a myeloid leukemia cell line (KPB-M15) and placenta as detected by a monoclonal antibody. *Cancer Res* 1987. **47**: 5025-5030.

44 **Pulugulla, S. H., Workman, R., Rutter, N. W., Yang, Z., Adamik, J., Lupish, B., Macar, D. A., el Abdouni, S., Esposito, E. X., Galson, D. L., Camacho, C. J., Madura, J. D. and Auron, P. E.,** A combined computational and experimental approach reveals the structure of a C/EBPβ–Spi1

1192 interaction required for IL1B gene transcription. *Journal of Biological Chemistry* 2018. **293**:
1193 19942-19956.
1194 45 **Chute, J. P., Muramoto, G. G., Whitesides, J., Colvin, M., Safi, R., Chao, N. J. and McDonnell,**
1195 **D. P.,** Inhibition of aldehyde dehydrogenase and retinoid signaling induces the expansion of
1196 human hematopoietic stem cells. *Proc Natl Acad Sci U S A* 2006. **103**: 11707-11712.
1197 46 **Ley, K., Hoffman, H. M., Kubes, P., Cassatella, M. A., Zychlinsky, A., Hedrick, C. C. and Catz, S.**
1198 **D.,** Neutrophils: New insights and open questions. *Science Immunology* 2018. **3**.
1199 47 **Zhang, Y., Gao, S., Xia, J. and Liu, F.,** Hematopoietic Hierarchy - An Updated Roadmap. *Trends*
1200 *Cell Biol* 2018. **28**: 976-986.
1201 48 **Jiang, J., Faiz, A., Berg, M., Carpaij, O. A., Vermeulen, C. J., Brouwer, S., Hesse, L., Teichmann,**
1202 **S. A., Ten Hacken, N. H. T., Timens, W., van den Berge, M. and Nawijn, M. C.,** Gene
1203 signatures from scRNA-seq accurately quantify mast cells in biopsies in asthma. *Clin Exp*
1204 *Allergy* 2020. **50**: 1428-1431.
1205 49 **Klimchenko, O., Mori, M., Distefano, A., Langlois, T., Larbret, F., Lecluse, Y., Feraud, O.,**
1206 **Vainchenker, W., Norol, F. and Debili, N.,** A common bipotent progenitor generates the
1207 erythroid and megakaryocyte lineages in embryonic stem cell-derived primitive
1208 hematopoiesis. *Blood* 2009. **114**: 1506-1517.
1209 50 **Hoppe, P. S., Schwarzfischer, M., Loeffler, D., Kokkaliaris, K. D., Hilsenbeck, O., Moritz, N.,**
1210 **Endele, M., Filipczyk, A., Gambardella, A., Ahmed, N., Etzrodt, M., Coutu, D. L., Rieger, M.**
1211 **A., Marr, C., Strasser, M. K., Schauberger, B., Burtscher, I., Ermakova, O., Burger, A., Lickert,**
1212 **H., Nerlov, C., Theis, F. J. and Schroeder, T.,** Early myeloid lineage choice is not initiated by
1213 random PU.1 to GATA1 protein ratios. *Nature* 2016. **535**: 299-302.
1214 51 **Bergiers, I., Andrews, T., Vargel Bolukbasi, O., Buness, A., Janosz, E., Lopez-Anguita, N.,**
1215 **Ganter, K., Kosim, K., Celen, C., Itir Percin, G., Collier, P., Baying, B., Benes, V., Hemberg, M.**
1216 **and Lancrin, C.,** Single-cell transcriptomics reveals a new dynamical function of transcription
1217 factors during embryonic hematopoiesis. *Elife* 2018. **7**.
1218 52 **erythroid abnormalitiesGasiorek, J. J. and Blank, V.,** Regulation and function of the NFE2
1219 transcription factor in hematopoietic and non-hematopoietic cells. *Cell Mol Life Sci* 2015. **72**:
1220 2323-2335.
1221 53 **Chan, W. Y., Follows, G. A., Lacaud, G., Pimanda, J. E., Landry, J. R., Kinston, S., Knezevic, K.,**
1222 **Piltz, S., Donaldson, I. J., Gambardella, L., Sablitzky, F., Green, A. R., Kouskoff, V. and**
1223 **Gottgens, B.,** The paralogous hematopoietic regulators Lyl1 and Scl are coregulated by Ets
1224 and GATA factors, but Lyl1 cannot rescue the early Scl-/- phenotype. *Blood* 2007. **109**:
1225 1908-1916.
1226 54 **Laurenti, E., Doulatov, S., Zandi, S., Plumb, I., Chen, J., April, C., Fan, J. B. and Dick, J. E.,** The
1227 transcriptional architecture of early human hematopoiesis identifies multilevel control of
1228 lymphoid commitment. *Nat Immunol* 2013. **14**: 756-763.
1229 55 **Takahashi, K. and Yamanaka, S.,** Induction of Pluripotent Stem Cells from Mouse Embryonic
1230 and Adult Fibroblast Cultures by Defined Factors. *Cell* 2006. **126**: 663-676.
1231 56 **Capellera-Garcia, S., Pulecio, J., Dhulipala, K., Siva, K., Rayon-Estrada, V., Singbrant, S.,**
1232 **Sommarin, Mikael N. E., Walkley, Carl R., Soneji, S., Karlsson, G., Raya, Á., Sankaran, Vijay G.**
1233 **and Flygare, J.,** Defining the Minimal Factors Required for Erythropoiesis through Direct
1234 Lineage Conversion. *Cell Reports* 2016. **15**: 2550-2562.
1235 57 **Moreau, T., Evans, A. L., Vasquez, L., Tijssen, M. R., Yan, Y., Trotter, M. W., Howard, D.,**

Colzani, M., Arumugam, M., Wu, W. H., Dalby, A., Lampela, R., Bouet, G., Hobbs, C. M., Pask, D. C., Payne, H., Ponomaryov, T., Brill, A., Soranzo, N., Ouwehand, W. H., Pedersen, R. A. and Ghevaert, C., Large-scale production of megakaryocytes from human pluripotent stem cells by chemically defined forward programming. *Nature Communications* 2016. **7**.

58 **Beekman, R. and Touw, I. P.,** G-CSF and its receptor in myeloid malignancy. *Blood* 2010. **115**: 5131-5136.

59 **Hume, D. A. and MacDonald, K. P.,** Therapeutic applications of macrophage colony-stimulating factor-1 (CSF-1) and antagonists of CSF-1 receptor (CSF-1R) signaling. *Blood* 2012. **119**: 1810-1820.

60 **Akashi, K., Traver, D., Miyamoto, T. and Weissman, I. L.,** A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. *Nature* 2000. **404**: 193-197.

61 **Schippel, N. and Sharma, S.,** Dynamics of human hematopoietic stem and progenitor cell differentiation to the erythroid lineage. *Experimental Hematology* 2023. **123**: 1-17.

62 **Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., Zhang, L., Fan, G., Xu, J., Gu, X., Cheng, Z., Yu, T., Xia, J., Wei, Y., Wu, W., Xie, X., Yin, W., Li, H., Liu, M., Xiao, Y., Gao, H., Guo, L., Xie, J., Wang, G., Jiang, R., Gao, Z., Jin, Q., Wang, J. and Cao, B.,** Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 2020. **395**: 497-506.

63 **Fajgenbaum, D. C., Longo, D. L. and June, C. H.,** Cytokine Storm. *New England Journal of Medicine* 2020. **383**: 2255-2273.

64 **VanHorn, S. and Morris, S. A.,** Next-Generation Lineage Tracing and Fate Mapping to Interrogate Development. *Developmental Cell* 2021. **56**: 7-21.

65 **Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W. M., 3rd, Hao, Y., Stoeckius, M., Smibert, P. and Satija, R.,** Comprehensive Integration of Single-Cell Data. *Cell* 2019. **177**: 1888-1902 e1821.

66 **Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., Lennon, N. J., Livak, K. J., Mikkelsen, T. S. and Rinn, J. L.,** The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nature Biotechnology* 2014. **32**: 381-386.

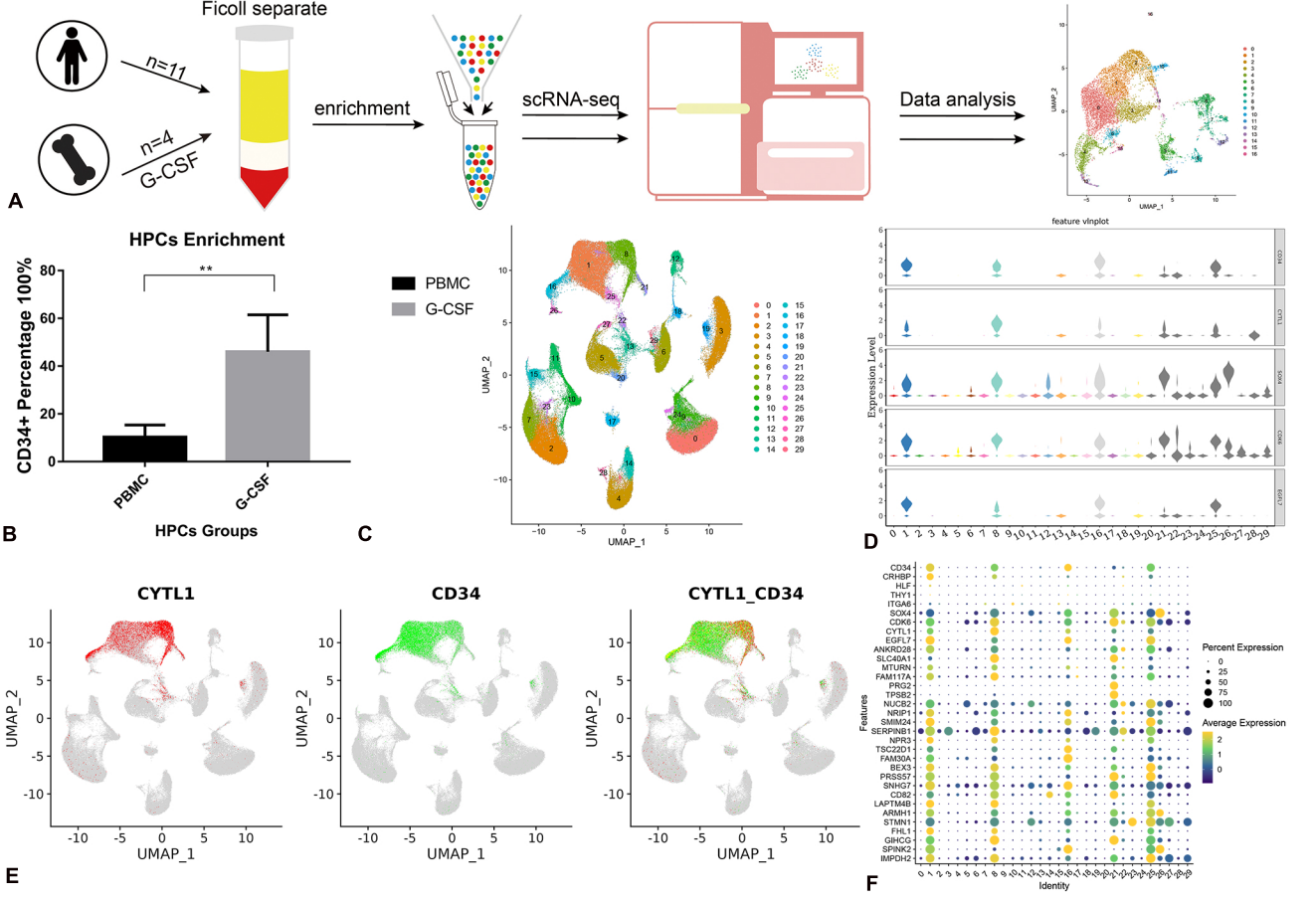
67 **Aibar, S., González-Blas, C. B., Moerman, T., Huynh-Thu, V. A., Imrichova, H., Hulselmans, G., Rambow, F., Marine, J.-C., Geurts, P., Aerts, J., van den Oord, J., Atak, Z. K., Wouters, J. and Aerts, S.,** SCENIC: single-cell regulatory network inference and clustering. *Nature Methods* 2017. **14**: 1083-1086.

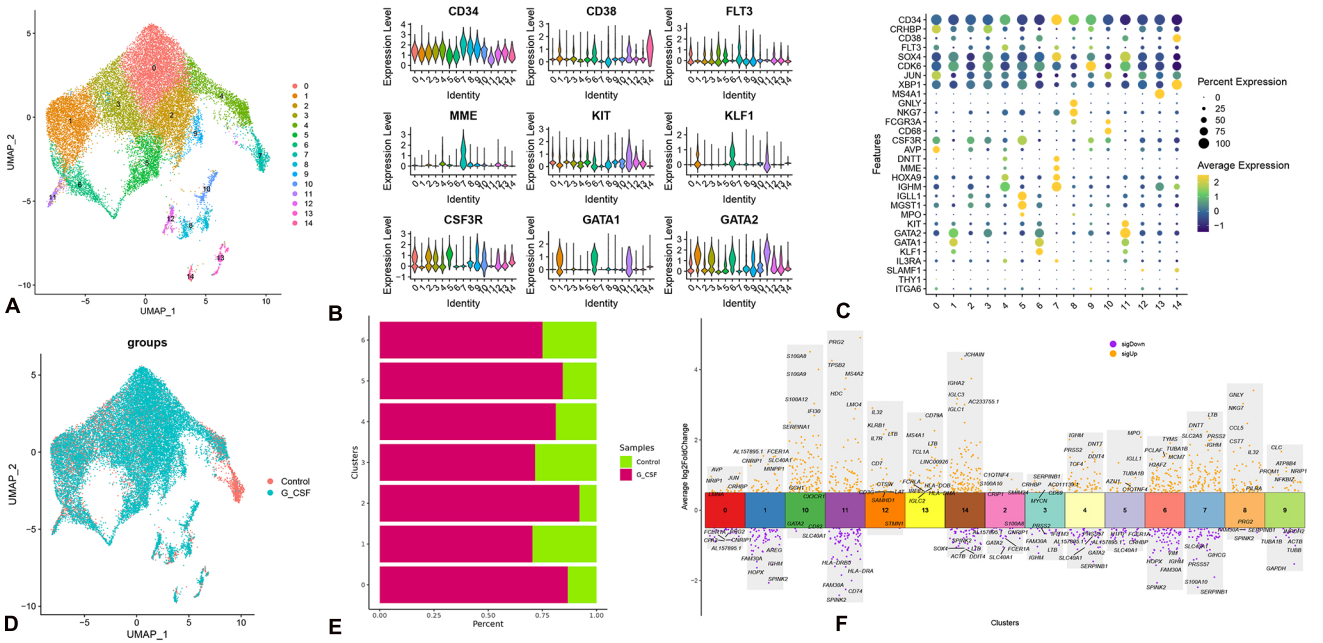
68 **Hänzelmann, S., Castelo, R. and Guinney, J.,** GSVA: gene set variation analysis for microarray and RNA-Seq data. *BMC Bioinformatics* 2013. **14**.

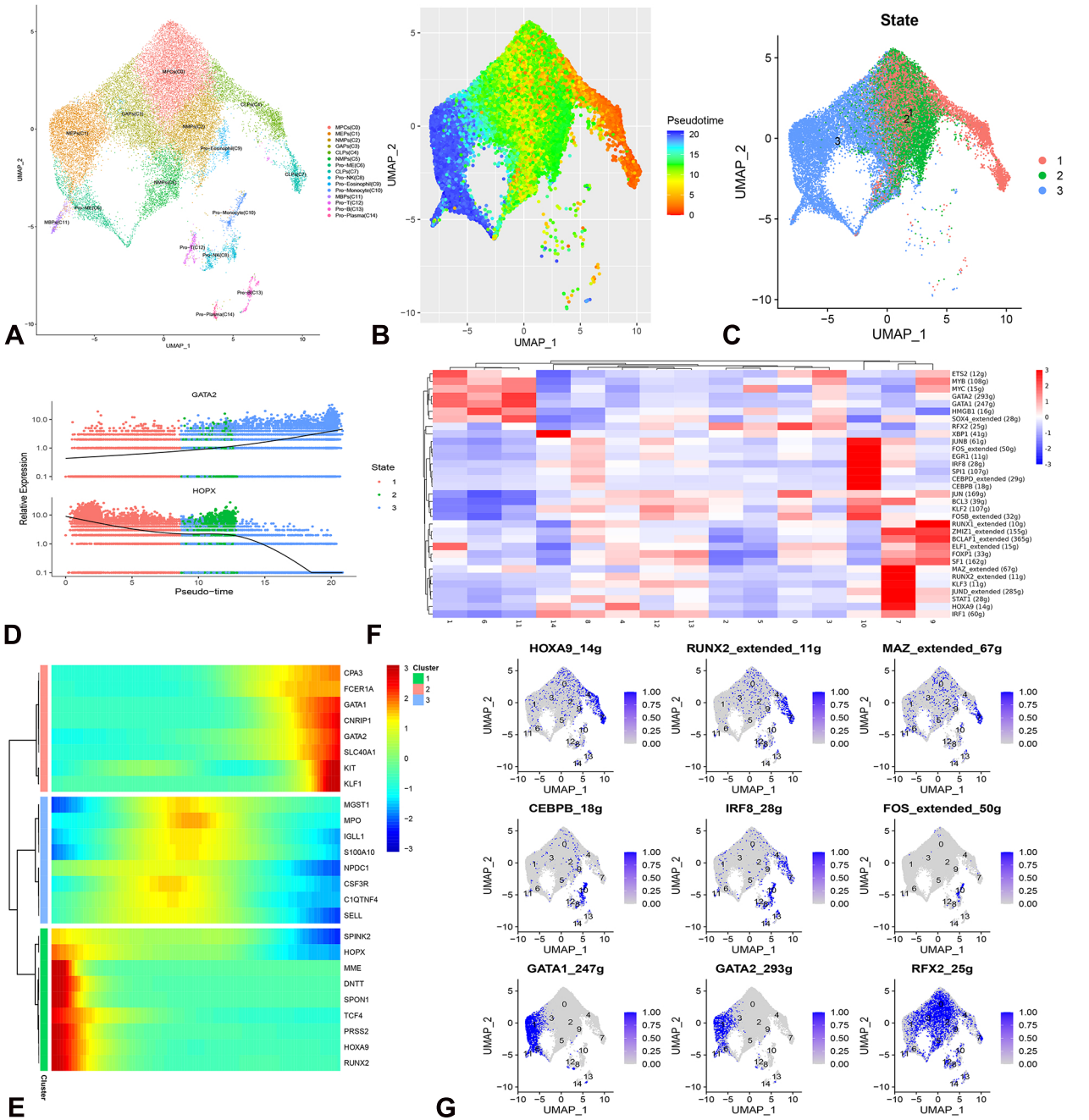
69 **Oetjen, K. A., Lindblad, K. E., Goswami, M., Gui, G., Dagur, P. K., Lai, C., Dillon, L. W., McCoy, J. P. and Hourigan, C. S.,** Human bone marrow assessment by single-cell RNA sequencing, mass cytometry, and flow cytometry. *JCI Insight* 2018. **3**.

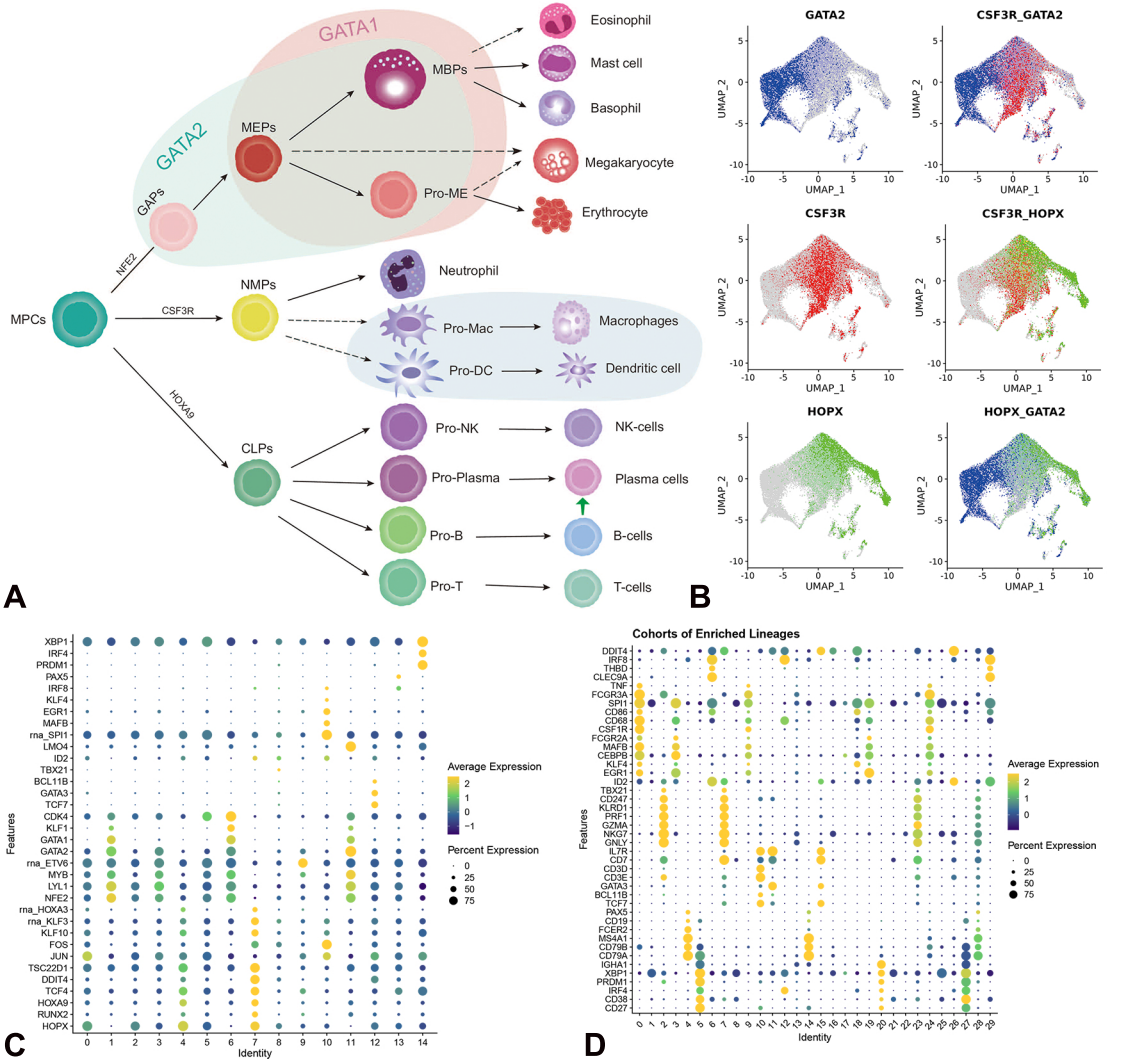
70 **Zhu, C., Lian, Y., Wang, C., Wu, P., Li, X., Gao, Y., Fan, S., Ai, L., Fang, L., Pan, H., Cheng, T., Shi, J. and Zhu, P.,** Single-cell transcriptomics dissects hematopoietic cell destruction and T-cell engagement in aplastic anemia. *Blood* 2021. **138**: 23-33.

71 **Tonglin, H., Yanna, Z., Xiaoling, Y., Ruilan, G. and Liming, Y.,** Single-Cell RNA-Seq of Bone Marrow Cells in Aplastic Anemia. *Front Genet* 2021. **12**: 745483.


Ficoll separate
n=11
n=4
G-CSF
enrichment
scRNA-seq
Data analysis
A
HPCs Enrichment
CD34+ Percentage 100%
PBMC
G-CSF
HPCs Groups
B
C
UMAP_1
UMAP_2
feature vlnplot
Expression Level
D
CYTL1
CD34
CYTL1_CD34
E
Features
Identity
Percent Expression
Average Expression
F


A
UMAP_1
UMAP_2
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
B
CD34
CD38
FLT3
MME
KIT
KLF1
CSF3R
GATA1
GATA2
Expression Level
Identity
C
Features
Percent Expression
Average Expression
D
groups
Control
G_CSF
E
Clusters
Percent
Samples
Control
G_CSF
F
Average log2Fold2hange
Clusters
sigDown
sigUp


A
B
C
D
E
F
G
UMAP_1
UMAP_2
Pseudotime
State
Pseudo−time
Relative Expression
GATA2
HOPX
Cluster
HOXA9_14g
RUNX2_extended_11g
MAZ_extended_67g
CEBPB_18g
IRF8_28g
FOS_extended_50g
GATA1_247g
GATA2_293g
RFX2_25g


GATA1
GATA2
MPCs
GAPs
MEPs
MBPs
Pro-ME
Eosinophil
Mast cell
Basophil
Megakaryocyte
Erythrocyte
NFE2
CSF3R
HOXA9
NMPs
Neutrophil
Pro-Mac
Macrophages
Pro-DC
Dendritic cell
CLPs
Pro-NK
NK-cells
Pro-Plasma
Plasma cells
Pro-B
B-cells
Pro-T
T-cells
A
GATA2
CSF3R_GATA2
CSF3R
CSF3R_HOPX
HOPX
HOPX_GATA2
UMAP_1
UMAP_2
B
Cohorts of Enriched Lineages
Features
Identity
Average Expression
Percent Expression
C
D

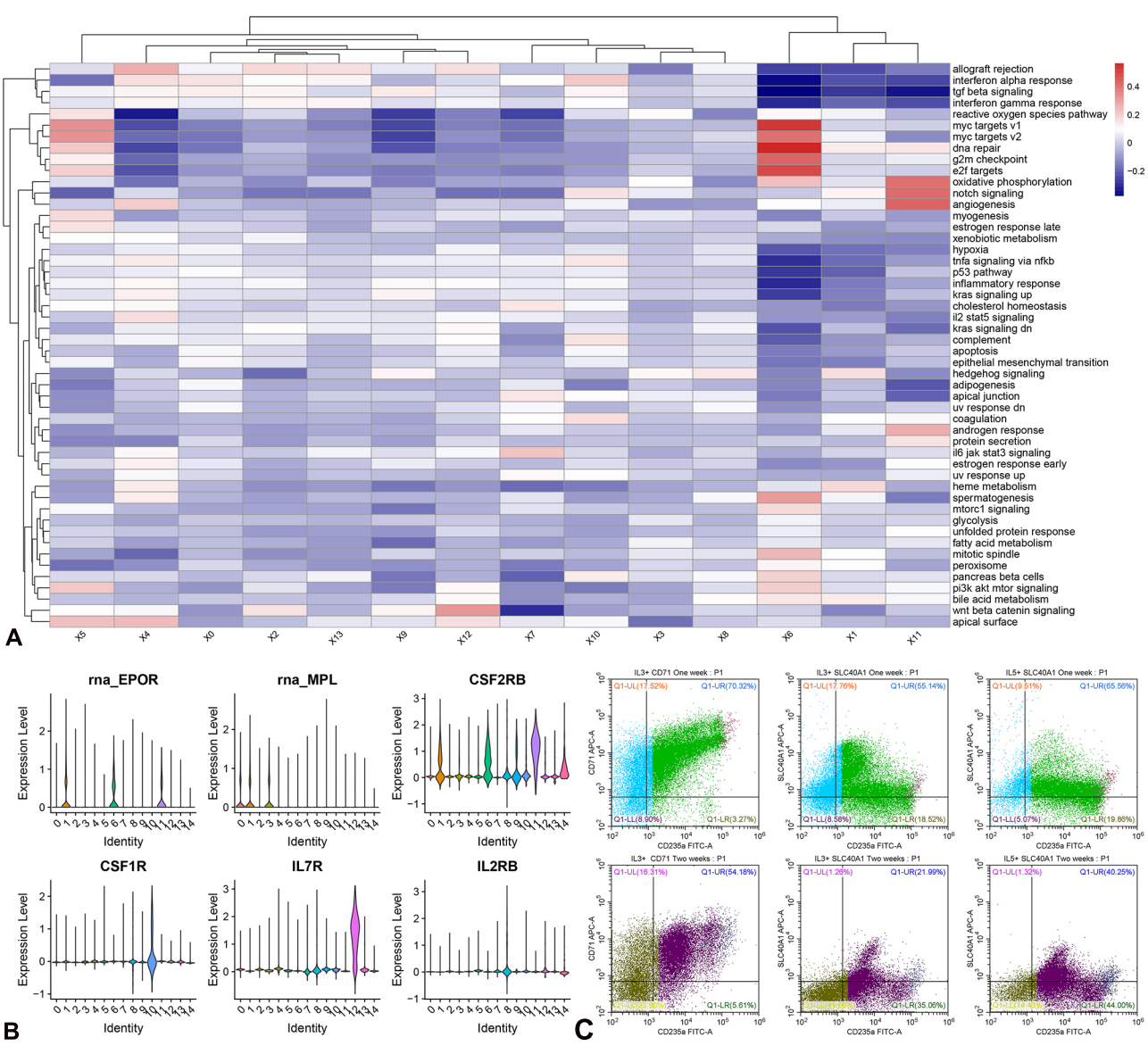

A
allograft rejection
interferon alpha response
tgf beta signaling
interferon gamma response
reactive oxygen species pathway
myc targets v1
myc targets v2
dna repair
g2m checkpoint
e2f targets
oxidative phosphorylation
notch signaling
angiogenesis
myogenesis
estrogen response late
xenobiotic metabolism
hypoxia
tnfa signaling via nfkb
p53 pathway
inflammatory response
kras signaling up
cholesterol homeostasis
il2 stat5 signaling
kras signaling dn
complement
apoptosis
epithelial mesenchymal transition
hedgehog signaling
adipogenesis
apical junction
uv response dn
coagulation
androgen response
protein secretion
il6 jak stat3 signaling
estrogen response early
uv response up
heme metabolism
spermatogenesis
mtorc1 signaling
glycolysis
unfolded protein response
fatty acid metabolism
mitotic spindle
peroxisome
pancreas beta cells
pi3k akt mtor signaling
bile acid metabolism
wnt beta catenin signaling
apical surface
X5
X4
X0
X2
X13
X9
X12
X7
X10
X3
X8
X6
X1
X11
0.4
0.2
0
−0.2
B
rna_EPOR
rna_MPL
CSF2RB
CSF1R
IL7R
IL2RB
Expression Level
Identity
C
IL3+ CD71 One week : P1
Q1-UL(17.52%)
Q1-UR(70.32%)
Q1-LL(8.90%)
Q1-LR(3.27%)
IL3+ SLC40A1 One week : P1
Q1-UL(17.76%)
Q1-UR(55.14%)
Q1-LL(8.56%)
Q1-LR(18.52%)
IL5+ SLC40A1 One week : P1
Q1-UL(9.51%)
Q1-UR(65.56%)
Q1-LL(5.07%)
Q1-LR(19.86%)
IL3+ CD71 Two weeks : P1
Q1-UL(16.31%)
Q1-UR(54.18%)
Q1-LR(5.61%)
IL3+ SLC40A1 Two weeks : P1
Q1-UL(1.26%)
Q1-UR(21.99%)
Q1-LR(35.06%)
IL5+ SLC40A1 Two weeks : P1
Q1-UL(1.32%)
Q1-UR(40.25%)
Q1-LR(44.00%)
CD71 APC-A
SLC40A1 APC-A
CD235a FITC-A

Table 1. Selected specific marker genes and genes associated with fate decisions in all redefined subgroups

Progenitor Cell Types	Marker genes and fate decision genes (partly)	Transcription factors	Negative
CLP (C4, C7)	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,TBX21	IL7R,GATA3
Pro-B2 (C14, Pro-Plasma cell)	CD27, IGKC, IGHA1, SLAMF7, CD79A, CD79B, PRDM1, IRF4, CD38, JCHAIN	PRDM1, IRF4	MS4A1, FCER2
Pro-T (C12)	TCF7, IL7R, GATA3, KLRB1, CD3E, CD3D, CD7, CD247, LTB, DDIT4	TCF7,GATA3,DDIT4,BCL11B	GNLY, FCGR3A, GZMA
Pro-B1 (C13)	CD19, MS4A1,FCER2,CD79A, CD79B, IGHM, LTB, IGKC, PAX5, VPREB3,CD22, CD24	PAX5, BACH2,IRF8	CD27
NMPs(C2, C5)	CSF3R, MPO, MGST1, IGLL1, S100A10, C1QTNF4, NPDC1, MYB, CDK4, CDCA7	MYB, CEBPA	GATA2, SLC40A1, CNRIP1,LTB
MkE Lineages(C1,C6,C11)	CNRIP1, CPA3 ,FCER1A, GATA1, KLF1, HPGDS, SLC40A1, GATA2, ZBTB16 ,TAL1, TESPA1, MYB, CDK4, ITGA2B, CSF2RB	GATA1, ZBTB16,TAL1, CDK4, TESPA1,KLF1	SPINK2,HOPX, C1QTNF4, CSF3R
GAPs (C3)	GATA2, NFE2, LYL1, MYB, SLC40A1, TESPA1, CSF3R	GATA2, NFE2, LYL1, MYB	GATA1, KLF1
MEPs (C1)	GATA2,NFE2,LYL1,MYB,GATA1,KLF1,CSF2RB,SLC40A1, ZBTB16,TAL1,CDK4,TESPA1	GATA1, KLF1, ZBTB16, TAL1	CSF3R
Pro-ME(C6)	HBD,CDT1, MCM2, MCM6, MCM5, MCM4, MCM3, MCM7, CDCA7, CDK4, TYMS	CDK4	MS4A2, MS4A3

MBPs (C11)	TPSAB1,LMO4, HDC, MS4A2, TPSB2, MS4A3, KIT, HPGDS, CLC, PRG2, MCM2-MCM7,APOC1	LMO4, CDK4,GATA2	HBD
Eosinophil (C9)	CLC, HDC, RFLNB, ETV6	ETV6	MS4A2,TPSB2,MS4A3
Pro-Monocytes1 (C0, 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 (C10, DCs)	CLEC9A, ANPEP, THBD, IRF8, KLF4, CD68, CD86, ITGAX, LYZ, SPI1 , LST1, DDIT4, SLAMF7	SPI1,KLF4,IRF8,DDIT4, BLC6	FCGR3A, CSF1R, MAFB