



**HAL**  
open science

## Single-cell analysis of megakaryopoiesis in peripheral CD34+ cells: insights into ETV6-related thrombocytopenia

Timothée Bigot, Elisa Gabinaud, Laurent Hannouche, Véronique Sbarra, Elisa Andersen, Delphine Bastelica, Céline Falaise, Manal Ibrahim-Kosta, Marie Loosveld, Paul Saultier, et al.

### ► To cite this version:

Timothée Bigot, Elisa Gabinaud, Laurent Hannouche, Véronique Sbarra, Elisa Andersen, et al.. Single-cell analysis of megakaryopoiesis in peripheral CD34+ cells: insights into ETV6-related thrombocytopenia. 2022. hal-03865977

**HAL Id: hal-03865977**

**<https://hal.science/hal-03865977v1>**

Preprint submitted on 22 Nov 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



22 **Key points**

23 - scRNAseq gain insight into *in vitro* megakaryopoiesis, identify MK-primed CMP, and a  
24 differentiation trajectory that bypasses the CMP.

25  
26 - *ETV6* variants led to the development of aberrant MEP and MK cell populations.

27

28 **Abstract**

29 Expansion of human megakaryoblasts from peripheral blood-derived CD34<sup>+</sup> cells is commonly  
30 used to characterize inherited or acquired thrombocytopenia and evaluate defects in megakaryocyte  
31 (MK) differentiation, MK maturation and proplatelet formation. We applied single-cell RNA  
32 sequencing to understand local gene expression changes during megakaryopoiesis (days 6 and 11  
33 of differentiation) in peripheral CD34<sup>+</sup> cells from healthy controls and patients with *ETV6*-related  
34 thrombocytopenia.

35 Analysis of gene expression and regulon activity revealed distinct clusters partitioned into seven  
36 major cell stages: hematopoietic stem/progenitor cells (HSPC), common-myeloid progenitors  
37 (CMP), MK-primed CMP, granulocyte-monocyte progenitors, megakaryocyte-erythroid  
38 progenitors (MEP), MK progenitor /mature MK (MKP/MK) and platelets. We observed a  
39 subpopulation of MEP that arose directly from HSPC, deviating from the canonical MK  
40 differentiation pathway.

41 *ETV6* deficiency was characterized by an increase in HSPC, a decrease in MKP/MK, and a lack  
42 of platelets. *ETV6* deficiency also led to the development of aberrant MEP and MKP/MK cell  
43 populations. Genes involved in “mitochondrial” and “DNA repair” pathways were downregulated,  
44 while genes involved in “translation” pathways were upregulated. Analysis of patient samples and  
45 hematopoietic cell lines transduced with an *ETV6* variant revealed increased translation in MK.  
46 Ribosomal protein small 6 (RPS6) levels in MK, platelets and peripheral blood mononuclear cells  
47 was consistent with the translation findings.

48 Our results provide a framework to understand peripheral CD34<sup>+</sup> cell-derived megakaryocytic  
49 cultures. Our observations also shed light on *ETV6*-variant pathology and reveal potential targets  
50 for diagnostic and therapeutic purposes.

51

52

53

54

## 55 **Introduction**

56 Megakaryocytes (MK) are fragile and only represent a small fraction of normal bone marrow cells  
57 (approximately 0.05% of mononuclear cells), which has hindered the study of megakaryopoiesis  
58 and the hierarchical structure of these cells. Most studies have analyzed flow-sorted cell  
59 populations, which limits assessments to a predefined cell subset.

60 *In vitro* culture systems for MK progenitors has enabled the analysis of megakaryopoiesis and  
61 regulation of MK differentiation. *In vitro* platelet production may represent an alternative to at least  
62 partially compensate for the increasing demand for platelet concentrates. Many investigators have  
63 attempted to increase platelet production *in vitro* by modifying culture media components<sup>1-4</sup>. An  
64 *ex-vivo* serum-free liquid culture system has also been used to expand normal human  
65 megakaryoblasts from purified peripheral-derived CD34<sup>+</sup> cells to characterize acquired or inherited  
66 thrombocytopenia (IT) and evaluate defects in MK differentiation, maturation and proplatelet  
67 formation<sup>5,6</sup>. Furthermore, these models are largely used to evaluate the effect of infectious diseases  
68 or therapeutic agents on megakaryoblast differentiation<sup>7-10</sup> and investigate novel mechanisms in  
69 MK differentiation and platelet function<sup>11-16</sup>

70 In the current study, we aimed to characterize the developmental stages of normal *in vitro*  
71 megakaryopoiesis. Primary CD34<sup>+</sup>-hematopoietic progenitor cells were induced to differentiate  
72 along the megakaryocytic lineage in liquid suspension cultures continuously exposed to  
73 thrombopoietin. We compared our results to those obtained in patients with ETV6-related  
74 thrombocytopenia (ETV6-RT), a highly penetrant form of IT with autosomal dominant  
75 inheritance<sup>17</sup>. The common phenotype observed in ETV6-RT includes moderate thrombocytopenia  
76 sometimes associated with bleeding and predisposition to acute T or B-cell lymphoblastic  
77 leukemia. ETV6-RT is also associated with B-cell lymphoma, acute myeloid leukemia and  
78 myelodysplasia to a lesser extent<sup>18-20</sup>.

79 The precise role that *ETV6* plays in megakaryocyte differentiation remains poorly understood.  
80 Studying *Etv6* function in murine models is challenging because complete loss of the gene is  
81 lethal<sup>21,22</sup> and heterozygous *Etv6* mice have obvious unperturbed hematopoiesis<sup>23</sup>. Human induced  
82 pluripotent stem cells (iPSC) harboring a pathogenic heterozygous *ETV6* mutation do not give rise  
83 to an increase in hematopoietic progenitor cells and MK. However, iPSC carrying the homozygous

84 *ETV6* mutation give rise to an increase in hematopoietic progenitor cells and immature MK, as  
85 observed in heterozygous patients using an *in vitro* model of CD34<sup>+</sup>-derived MK<sup>24,25</sup>. Therefore,  
86 further investigation using patient cells is required to better understand defective *ETV6*-  
87 megakaryopoiesis.

88 Using cells derived from controls and patients harboring the *ETV6*-variants, we applied single-cell  
89 RNA sequencing to examine the transcriptome of each cell type during differentiation from  
90 hematopoietic stem/progenitor cells (HSPC) to MK. We observed classical hematopoietic cell  
91 populations (HSPC, common-myeloid progenitors (CMP), granulocyte-monocyte progenitors  
92 (GMP), megakaryocyte-erythroid progenitors (MEP), MK progenitors (MKP), mature MK and  
93 platelets (Plt) and an unusual MK-primed CMP subpopulation. We observed a subpopulation of  
94 MEP that arose directly from HSPC, bypassing CMP. *ETV6* deficiency led to the development of  
95 aberrant MEP and MK populations, with an enrichment in “mitochondrial”, “DNA repair” and  
96 “translation” pathways. Our findings provide insight into peripheral CD34<sup>+</sup>-megakaryopoiesis,  
97 *ETV6*-variant pathology and potential targets for diagnostic and therapeutic purposes.

98

99 **Methods**

100 ***In vitro* megakaryocyte differentiation**

101 ETV6-variant carriers and healthy volunteers were recruited at the Center for the Investigation of  
102 Hemorrhagic and Thrombotic Pathologies at Marseille University Hospital (authorization number  
103 20200T2-02). Peripheral CD34<sup>+</sup>-cells were purified using magnetic cell sorting (Miltenyi-Biotec)  
104 and then cultured in StemSpan Serum-Free Expansion Medium combined with Megakaryocyte  
105 Expansion Supplement (Stemcell Technologies)<sup>26</sup>.

106 **Single-cell RNA sequencing**

107 Control (n=2) and ETV6-defective (n=2) cells were harvested from culture at days 6 and 11 (Figure  
108 1A). The cell samples were labeled with a distinct hashtag oligo (TotalSeq, Biolegend) and pooled.  
109 Single-cell isolation was then carried out with the 10X Genomics Technology using the Chromium  
110 Next GEM Single Cell 5' Kit v2. Single-cell cDNA synthesis and sequencing libraries were  
111 prepared with single-cell 5' Library and Gel Bead kit.

112 **Additional methods**

113 See the supplemental methods for additional details regarding data preprocessing and analysis, site-  
114 directed mutagenesis, western-blot, microscopy, flow-cytometry, transduction and statistical  
115 analyses.

116

## 117 **Results**

### 118 **Characterization of differentiation in control CD34<sup>+</sup>-cells**

119 We performed UMAP (Uniform Manifold Approximation and Projection) non-linear dimensional  
120 reduction to analyze cell transcriptome heterogeneity. The observed variation between the  
121 transcription profiles of cells after 6 and 11 days in culture indicated that distinct gene sets were  
122 involved in each stage of differentiation (Figure 1B).

123 Using unsupervised clustering, we identified a total of 15 clusters (Figure 1C), which were present  
124 in both controls (Supplemental figure 1A-C). We performed a detailed characterization of each cell  
125 cluster based on known cell gene sets (signature)<sup>27-30</sup> (Figure 1D) and the top differentially  
126 expressed genes (DEGs) (Figure 1E; Supplemental Figure 2, Supplemental Table 1).

127 Cluster 10 was designated HSPC based on the expression profiles of *PROM1*, *CRHBP*, *FLT3*,  
128 *HOPX* and *AVP*. Similarly, clusters 11 and 7 were designated CMP based on the expression profiles  
129 of *CPA3*, *PRG2*, *CLC* and *TPSAB1*. Cluster 11 differed from cluster 7 as it shared marker genes  
130 with the MKP/MK clusters (*VWA5A*, *LMNA*, *CAVIN2*, *LAT*, *CD9* and *ITGA2B*) (Figures 1D-F)  
131 and expressed the genes *KIT*, *KRT1*, *HPGDS* and *TPSB2*. We thus defined this cluster 11 as MK-  
132 primed CMP. Cluster 13 was designated GMP because it expressed specific markers such as *MPO*,  
133 *ELANE* and *PRTN3*. Clusters 9, 5, 4, 0 and 2 were designated MEP. Although these clusters shared  
134 a strong common MEP signature (*CD38*, *TFRC*, *DEPTOR*, *KLF1*, *TFR2*), we further characterized  
135 these five subpopulations. Cluster 9 was the most immature population and was designated “early  
136 MEP” based on the remaining expression levels of HSPC and CMP genes. Cluster 5 was designated  
137 a MEP-ERP subpopulation based on the genes associated with erythroid progenitors (ERP) such  
138 as *HBB*, *TFR2* and *ANK1*. Clusters 4, 0 and 2 were designated a MEP-MK subpopulation based on  
139 the expression levels of MK genes, such as *GP9*, *GP6*, *GP1BA* and *MPIGB*, and the gradual  
140 increase in *MYH9* expression levels between clusters 4, 0 and 2 (Figures 1E-F, Supplemental  
141 figures 1D and 2). Clusters 1, 6, 3, 8 and 12 were primarily observed at day 11 and were designated  
142 MK based on the expression of genes associated with classical MK. Specific MK markers (*GP6*,  
143 *GP9*, *PF4*, *P2RY1*) gradually increased from cluster 1 to 12 (Figure 1D, Supplemental figures 1D  
144 and 2). Of these five clusters, cluster 1 was the most immature stage (MKP) and cluster 12 was the  
145 most mature. Cluster 14 was only observed at day 11 was designated platelets based on the low

146 number of genes expressed as compared with the MKP/MK populations (mean±SD: 1,418±347 vs.  
147 4,208 ±1,399) and overall RNA count (mean±SD: 2,243±751 vs. 23,855±14,826) (Figures 2A-B,  
148 Supplemental figure 1D). Furthermore, cluster 14 expressed the same genes as cluster 12, with  
149 higher expression levels of *GP6* and *ITGA2B* (Figures 1D-E, Supplemental figures 1D and 2).

150 For each cell type, we proposed enriched gene signatures that included known cell-type specific  
151 genes, top-ranked DEGs and selective genes primarily expressed in cell stage-related clusters  
152 (Supplemental table 2). Lineage signature scores were then computed and assigned to each cell  
153 (Figure 2C). Cell types were then defined accordingly (Figure 2D).

#### 154 **Inference of megakaryopoiesis regulon activity in control cells**

155 The cell state transitions in megakaryopoiesis are tightly controlled by transcription factors (TF)<sup>17</sup>.  
156 Regulons are inferred groups of genes controlled as a unit by the same repressor or activator TF<sup>31</sup>.  
157 For each regulon, the activity score was calculated based on the cellular expression values for all  
158 genes. Cell type-specific regulons provide an opportunity to identify key regulators of cell fate  
159 decisions and establish cell signatures. Analysis of our control dataset using the single-cell  
160 regulatory network inference and clustering (SCENIC) workflow provided insight into cell type-  
161 specific regulons that drive cellular heterogeneity. Comparing each cell-types, we analyzed the top  
162 regulons to isolate key regulons at each stage of differentiation (Figure 2E, Supplemental figure 4).  
163 All 312 identified regulons are available in Supplemental table 3.

#### 164 **Characterization of the *ETV6* variants**

165 We performed a functional study of the novel *ETV6* variant pF417LTer4 compared with the  
166 previously described p.P214L<sup>25</sup> (Supplemental figure 5A). The nonsense mutation p.F417LTer4 is  
167 located in the ETS domain, while the missense mutation p.P214L is localized in the linker domain  
168 (Supplemental figure 5B). The clinical and laboratory characteristics of the patients are shown in  
169 Supplemental table 4.

170 We analyzed the repressive activity of the two variants using a dual-luciferase reporter assay. Co-  
171 transfection of the reporter plasmid containing the ETS-binding site along with expression of a  
172 plasmid encoding wild type (WT) *ETV6* resulted in almost 90% inhibition of luciferase activity.  
173 Substitution of WT *ETV6* with any of the *ETV6* variants led to a significant reduction in repressive

174 activity (85% to 100%) (Supplemental figure 5C). Western-blot analysis showed that mutant ETV6  
175 protein was expressed in the GripTite293 macrophage scavenger receptor (MSR) cell line  
176 (Supplemental figure 5D).

177 Subcellular fractionation of GripTite293 MSR cells showed increased ETV6 protein levels in the  
178 cytoplasmic fraction and decreased levels or absence of ETV6 in the nuclear fraction in cells  
179 expressing the p.P1214L or p.F417Lter4 variants compared with cells expressing the WT protein  
180 (Supplemental figure 5E). Microscopy confirmed that WT ETV6 concentrated primarily in cell  
181 nuclei, whereas both ETV6 variants were predominantly localized in the cytoplasm (Supplemental  
182 figure 5F).

### 183 **Single-cell transcriptional profiling of *ETV6*-variant CD34<sup>+</sup>-cells during megakaryopoiesis**

184 We performed UMAP non-linear dimensional reduction to analyze cell transcriptome  
185 heterogeneity and clustering. Compared with control cells, the day 6 and day 11 transcriptome  
186 profiles of patient cells overlapped much more, thereby suggesting that differentiation was delayed  
187 with accumulation of early-stage cells in both *ETV6*-variant carriers (Figure 3A, Supplemental  
188 figure 6A).

189 Using unsupervised clustering, a total of 11 clusters were identified (Figure 3B). The cell types  
190 found in controls were also observed in patients (Figures 3C-D, Supplemental figures 6 to 8,  
191 Supplemental table 5). Clusters 3 and 7 corresponded to HSPC; cluster 4 corresponded to CMP;  
192 cluster 10 corresponded to MK-primed CMP; cluster 9 corresponded to GMP; clusters 5, 2, 6, 8  
193 and 0 corresponded to MEP; and cluster 1 corresponded to MKP/MK. No platelet clusters were  
194 observed. The proposed lineage signatures (Supplemental table 2) were computed and displayed  
195 via feature plot, and cell types were assigned (Figures 3E-F).

### 196 **Developmental trajectory of megakaryocyte differentiation**

197 We then assessed the single-cell transcriptome for pseudotemporal ordering of differentiation states  
198 during megakaryopoiesis in controls and patients. Using Slingshot, we inferred differentiation  
199 trajectories in cells from controls and *ETV6*-variant patients. For each condition, we assessed the  
200 overall trajectory structure of each lineage (rooted tree) by generating the transcriptomic distance  
201 matrix using the manually set root-cluster (HSPC; cluster 10 for controls and cluster 3 for patients).

202 Several lineages were identified in each condition (5 and 4, for control and patients, respectively)  
203 (Supplemental figure 9). We observed a lineage of MK differentiation (Figures 4A-B), which  
204 bypassed the CMP cell type and differentiated directly from HSPC into MEP. In this MK trajectory,  
205 stem-cell markers such as *CD34*, *CD38* and *HLA-DRA* decreased, while the MK markers *GP9* and  
206 *PF4* displayed increased expression (Figure 4C).

207 During differentiation from HSPC to MK, the TF *GATA1*, *FLII* and *TALI* displayed increased  
208 expression in control cells, while *GATA2* displayed high expression levels in immature cells and  
209 reduced expression in later stages. During differentiation, *ETV6* and *RUNX1* expression levels were  
210 low and stable (Figure 4D, Supplementary figures 3 and 7).

211 Compared with controls, *ETV6*-deficient cells displayed stable expression of *CD34*, *HLA* (*HLA-*  
212 *DRA* (Figure 4C) and *HLA-DPB1*, *HLA-DPA1*, *HLA-DRB5*, *HLA-DRB1*, *HLA-A*, *HLA-E*, *HLA-*  
213 *DMA*, *HLA-C*, *HLA-B*, *HLA-DQA2*, *HLA-DQB1* and *HLA-DQA1* (data not shown)) and *GATA2*  
214 during MK differentiation; these genes remained highly expressed at the MKP/MK stage (fold  
215 change (FC) *ETV6* vs. WT) = 2.5 (*CD34*), 1.5 to 8.8 (*HLA*), 1.6 (*GATA2*), adjusted p-value <0.05).  
216 The expression levels of *GATA1*, *FLII* and *TALI* were lower than that observed in controls. We  
217 observed a delayed response time for *CD38*, *GP9* and *PF4* (Figures 4C-D). These results suggest  
218 that *ETV6*-variant carriers exhibit a delay in differentiation.

### 219 **Aberrant populations in *ETV6*-variant cells identified via single-cell RNA sequencing**

220 The full data set (control and patients together) was analyzed to compare controls against patients.  
221 The cell types identified in control and patient data sets were transferred to the full data set using  
222 the cell identity barcode (Figure 5A). The day 6-11 results and the differences between control and  
223 patient for each cell type are shown in Figure 5B-C. For the early stages (HSPC, CMP and GMP),  
224 the transcriptome profiles of controls and patients were markedly similar. By contrast, distinct gene  
225 expression patterns were observed for the MEP and MKP/MK populations. Cell type distribution  
226 differed between controls and patients (Figures 5D and E). Patients harboring *ETV6* variants  
227 displayed an increased proportion of HSPC (17±6.9 vs 3.0±0.7%) and a decreased proportion of  
228 MKP/MK (16.9±4.1 vs 38.5±1%) compared with controls. Furthermore, the MKP/MK populations  
229 in patients were characterized by a reduced number of detected genes (mean±SD: 2,945±927 vs  
230 4,884±1,112, p<0.0001) and RNA count (mean±SD: 12,444±6,051 vs. 29,960±14,522, p<0.0001).

231 Overall, these results indicate that *ETV6* mutations perturb megakaryopoiesis, resulting in a delay  
232 in cell differentiation and the development of aberrant MEP and MK populations.

### 233 **Patient samples displayed highly modified regulon activity and functional aberrant cell** 234 **populations**

235 We applied the SCENIC workflow to full dataset to better characterize the aberrant MEP and  
236 MKP/MK populations observed in the *ETV6* patients. The UMAP reduction, with the 312  
237 combined regulon activity scores in each cell according to transcriptome signature designation (cell  
238 type) is shown in Supplemental figure 10A. We observed differences restricted to the MEP and  
239 MKP/MK stages. To further analyze this defect, the activity of individual regulon was also  
240 evaluated in all cells (Supplemental figure 11). Compared to the initial stages (HSPC, GMP, CMP  
241 and MK-primed CMP), we observed marked differences at the MEP and MKP/MK stages between  
242 patients and controls. The top 20 differentially active regulons are shown in Supplemental Figure  
243 12. Common regulons were observed at the two cell stages (e.g., MAX, SPI1, GATA2, IRF5).  
244 Differences in regulon activities were more pronounced at the MKP/MK stage. The activity of 12  
245 remarkable regulons in each cell type is presented in Supplemental figure 10B. For the latter,  
246 hyperactivity was observed in patient cells. Overall, these results complement the previous  
247 observations and indicate that *ETV6*-variants play a functional role in MEP and MKP/MK  
248 differentiation. Hyper or hypoactivity of every regulon is available in Supplemental Table 3.

### 249 **Deregulated pathways in *ETV6*-variant cells identified via single-cell RNA sequencing**

250 When comparing controls and patients, the number of DEGs (adjusted p-value < 0.05) increased  
251 over the course of cell differentiation, especially at the MEP stage (HSPC, DEGs=38, 23  
252 upregulated genes, 15 downregulated genes; CMP, DEGs=56, 29 upregulated genes, 27  
253 downregulated genes; MK-primed CMP, DEGs=30, 15 upregulated genes, 15 downregulated  
254 genes; GMP, DEGs=32, 21 upregulated genes, 11 downregulated genes; MEP, DEGs=239, 100  
255 upregulated genes, 139 downregulated genes; and MKP/MK, DEGs=942, 339 upregulated genes,  
256 603 downregulated genes) (Figure 6A). In MEP and MKP/MK populations, DEGs were enriched  
257 in various biological pathways in multiple gene set databases (GO Biological process (Figures 6B-  
258 C), KEGG (Supplemental figure 13) and Reactome (Supplemental figure 14). Top 10 GO  
259 deregulated pathways are available in Supplemental table 6 and 7. We observed downregulation of

260 “mitochondrial”, “mRNA processing, splicing and localization to nucleus” and “DNA repair and  
261 cellular responses to DNA damage” pathways. The “translation” pathway was among the most  
262 upregulated pathways (Supplemental table 8).

### 263 **Mitochondrial pathway in patients harboring an *ETV6* variant**

264 As several mitochondria-specific processes were downregulated in *ETV6*-variant cells (i.e.,  
265 mitochondrial translation, oxidative phosphorylation, respiratory electron transport chain, ATP  
266 synthesis, assembling and biogenesis complex), we further assessed the mitochondrial defects  
267 using the MitoXplorer2.0 pipeline to evaluate 36 mitochondrial functions/pathways among the  
268 DEGs<sup>32</sup>. The greatest FC were observed for oxidative phosphorylation, ROS defense, glycolysis,  
269 import and sorting, and mitochondrial translation (Figure 7A, Supplemental figures 15 to 17).  
270 Sixty-five of 430 downregulated mitochondrial genes contained sequences that correspond to the  
271 canonical *ETV6* binding site (C/AGGAAG/A) (Normalized Enrichment Score NES 4.85, rank  
272 389e/562, Supplemental figures 15B, Supplemental table 9).

273 The number of DEGs progressively increased over the course of cell differentiation (Supplemental  
274 figures 15C and 16). *DNML1*, which encodes a key mediator of mitochondrial division DRP1, and  
275 *CYCS*, which encodes for cytochrome C, were downregulated in *ETV6*-variant MK (Supplemental  
276 figure 18).

### 277 **Increased translation in patients harboring an *ETV6* variant**

278 Translation was the dominant upregulated pathway. Translation levels were evaluated in HEL cells  
279 transduced with P214L or WT *ETV6* before and after PMA (phorbol 12-myristate 13-acetate)-  
280 induced differentiation. As expected, PMA stimulation led to reduced translation in all samples  
281 (Figure 7B). Translation levels were higher in *ETV6*-deficient cells at the basal state and after PMA  
282 stimulation (Figure 7B). Compared with control cells, higher translation levels were observed in  
283 CD34<sup>+</sup>-derived MK isolated from three *ETV6*-carriers at day 11 (F1 III-3, III-8, IV-1) and one at  
284 day 14 (F2 II-2) (Figure 7C). *RPS* and *RPL* were the most upregulated genes, which code for  
285 ribosomal protein small subunit and ribosomal protein large subunit, respectively (Supplemental  
286 figure 19). Among these 163 upregulated ribosomal genes, 35 genes contained sequences that  
287 correspond to the canonical *ETV6* binding site (NES 4.85, rank 40e/512, Supplemental table 9),  
288 including *RPS6* (Figure 7D). The *RPS6* kinase genes (*RPS6KAI*, *RPS6KA3*, *RPS6KCI*) also

289 contained ETV6 binding sites (Supplemental table 9). Compared with controls, RPS6 was  
290 overexpressed in CD34<sup>+</sup>-derived MK (day 14) from one patient and platelets from two patients  
291 (Figures 7E-F). RPS6 was also upregulated in peripheral blood mononuclear cells (PBMC) from  
292 five patients with an *ETV6* variant compared with controls (Figure 7G).

### 293 **Downregulated DNA repair pathways in *ETV6*-variant carriers**

294 Several pathways associated with DNA repair and cellular response to DNA damage were  
295 downregulated in patient MEP and MK (Supplemental table 8). Some genes involved in major  
296 DNA repair pathways displayed reduced expression levels: *FEN1* (base excision repair), *MGMT*  
297 (direct reversal of DNA damage), *RAD23A* and *RAD23B* (nucleotide excision repair), and *XRCC6*  
298 and *PRKDC* (non-homologous end joining)<sup>33</sup>. Caretaker genes indirectly involved in maintaining  
299 genomic stability (e.g., *TTK*, *NUDT1*, *DUT*, *UBE2V2*) were also downregulated in MK harboring  
300 *ETV6* variants (Figure 7H).

## 301 Discussion

302 Using single-cell transcriptome profiling of MK cell cultures derived from peripheral CD34<sup>+</sup>-cells,  
303 we established a signature for each cell stages of megakaryopoiesis. Both the gene expression  
304 profiles and the regulon profiles confirmed distinct cell-type signatures for each stage of  
305 differentiation. We also observed a differentiation trajectory in which MEP developed directly  
306 from HSPC and bypassed the CMP stage. A small MK-primed CMP population was also  
307 evidenced. Cells harboring *ETV6* variants displayed significantly distinct gene expression profiles  
308 starting at the MEP stage. Our results also reveal concomitant differences in the activity of specific  
309 regulons. Furthermore, the observed dysregulation of several pathways indicates that *ETV6*  
310 deficiency affects key cellular processes associated with mitochondria function, translation and  
311 DNA repair, which may represent promising mechanistic targets.

312 In this study, we highlight a biased megakaryopoiesis pathway that may be due to stress-driven  
313 hematopoiesis as a result of the culture conditions<sup>34</sup>. HSPC directly gave rise to MEP without  
314 passing through the CMP stage, as previously suggested<sup>35</sup>. Furthermore, it has been hypothesized  
315 that MK-biased hematopoietic stem cells (HSCs) represent the hierarchical apex with MEP  
316 developing directly from HSCs<sup>36</sup>, although this view has been challenged. We did not observe a  
317 MK-biased HSPC population<sup>37,38</sup> either because (1) the analysis time points (days 6 and 11) were  
318 too late to detect these cell types, (2) clustering was insufficient to visualize this small population,  
319 or (3) mRNA did not enable detection of MK-biased HSPC, as previously reported<sup>39</sup>. However, we  
320 identified a MK-primed CMP population that co-expressed megakaryocytic genes and displayed  
321 high *KIT* expression levels (27% of the CMP population). High *KIT* expression levels represented  
322 a good marker of MK-biased CMP. Previous murine *in vitro* and *in vivo* functional studies have  
323 demonstrated that HSCs with higher levels of c-Kit signaling preferentially differentiate into MK<sup>40</sup>.  
324 In human bone marrow, MK-primed CMP likely represented the major megakaryopoiesis pathway  
325 independent of the canonical MEP lineage<sup>41</sup>. This *in vitro* model may be a valuable tool to  
326 investigate these biased pathways.

327 We characterized each cell type by analyzing the activity of TF specific to each cell stage, which  
328 confirmed the gene expression designations. HSPC were characterized by the expression of the  
329 regulons TCFL2, TCF and HOXA9, which are involved in stem cell maintenance<sup>42-44</sup>. C/EBP  
330 family members are critical for myelopoiesis<sup>45,46</sup>. Accordingly, we observed noticeable C/EBP

331 activity at the CMP and GMP stages. Interestingly, we observed activated AhR in the MK-primed  
332 CMP population, thus indicating that AhR modulation plays a role in MK maturation, as previously  
333 described<sup>47-49</sup>. We also observed KLF1 activity in MEP, as KLF1 is involved in MEP lineage  
334 decisions and commitment<sup>50</sup>. Finally, NFE2 and GATA1 were expected markers of MK<sup>51,52</sup>.  
335 Overall, these findings demonstrate that characterizing TF activity at the single-cell level can be  
336 useful to phenotype cell types.

337 Applying the lineage signature used in control cells, we found that *ETV6*-deficient cells displayed  
338 the same cell type with a higher proportion of HSPC, a lower proportion of MK and an absence of  
339 platelets, which correlates with the thrombocytopenia phenotype observed in patients harboring an  
340 *ETV6* variant. We detected marked differences in the TF regulatory network; the most-affected  
341 regulons displayed higher activity levels in MEP and MKP isolated from patients. This finding is  
342 in accordance with a loss of *ETV6* repressor activity, which plays a role in the development of the  
343 pathology observed in patients. Also, some regulons hypoactivity was detected, challenging the  
344 unique repressor role of *ETV6*. It may be both an activator and a repressor depending of the  
345 genomic context and its co-factors, making it a potential pioneer TF like SPI1<sup>53</sup>.

346 Analysis of several gene enrichment databases displayed significant differences. Assessing the ten  
347 highest scores and after grouping the pathways with common genes, we observed deregulation of  
348 pathways linked with mitochondria function, translation and DNA repair in MEP and MKP/MK.

349 Our results indicate that patients harboring an *ETV6*-variant display decreased expression of genes  
350 involved in mitochondrial metabolism. We observed a marked decrease in the expression levels of  
351 *DNM1L*, which encodes DRP1 and has been shown to enhance the production of proplatelet  
352 forming MK<sup>54,55</sup>. These findings suggest that the mitochondria plays a role in *ETV6*-RT  
353 development, which merits further functional analysis.

354 Translation was the most upregulated pathway observed in MEP and MK patient cells. Altered  
355 expression levels of numerous ribosomal proteins (RP) genes were observed, such as  
356 downregulation of *RPS26* and upregulation of more than 30 RP with a FC>1.4. Very recent data  
357 has emphasized the unexpected contribution of ribosomal biogenesis in hematopoiesis and  
358 megakaryopoiesis. Treatment of mice or humans with the ribosomal biogenesis inhibitor CX-5461

359 results in an increase in circulating platelets, which is associated with the platelet/MK-biased  
360 hematopoietic pathway<sup>56</sup>. Additionally, ribosome biogenesis is involved in mediating the transition  
361 between proliferation and differentiation of erythroid progenitors<sup>57</sup>. GTPase-Dynamin-2 deletion  
362 in platelets and MKs induce a severe thrombocytopenia and bleeding diathesis in mice and result  
363 in upregulation of genes involved in ribosome biogenesis in erythroblast<sup>58</sup>. Several ribosomal  
364 proteins have been found to play a role in extra-ribosomal functions, including induction of  
365 apoptosis, tumor suppression, regulation of development, and DNA repair<sup>59</sup>. Our results highlight  
366 a significant decrease in DNA repair pathways in patients. Overall, these findings indicate that  
367 *ETV6* mutations are responsible for defects in translation and DNA repair pathways, which may  
368 contribute to leukemia predisposition.

369 Increased translation was confirmed at the functional level. Protein synthesis was increased in  
370 CD34<sup>+</sup> cell-derived MK in patients and hematopoietic cell lines transduced with an *ETV6* variant.  
371 As *RPS6* contains putative *ETV6* binding sites, which was confirmed via ChIP-sequencing (UCSC  
372 Genome Browser on GSM2574795, GSM2534228, GSM2574796), and *Rps6*-deficient mice  
373 display features of megakaryocytic dysplasia with thrombocytosis, this gene likely plays a role in  
374 thrombopoiesis<sup>60,61</sup>. Remarkably, total RPS6 antigen levels were increased two-fold in patient MK,  
375 platelets and PBMC compared with controls. Indeed, further investigation is required to elucidate  
376 the specific role that RPS6 plays in *ETV6*-RT.

377 In summary, our study demonstrates the heterogeneity of human CD34<sup>+</sup> cell-induced MK  
378 differentiation *in vitro*, a novel MK-primed CMP population, and a major differentiation trajectory  
379 in which HSPC develop directly into MEP and bypass the CMP stage. We found that *ETV6*  
380 variations cause defects in early hematopoietic stages and result in an aberrant MEP and MK  
381 populations with deregulated translation and DNA repair pathways. These findings provide novel  
382 insight into megakaryopoiesis and *ETV6* function that may be applied to develop targeted  
383 therapeutic strategies to alleviate platelet defects.

384

385 **Acknowledgments**

386 This work was supported by Aix-Marseille University (AMIDEX “Emergence et innovation”  
387 ngSUMMIT), the Agence Nationale de la Recherche (JCJC MOST) and the Institut National de la  
388 Santé et de la Recherche Médicale (PIA Biofit). The authors acknowledge the members of the  
389 French Reference Center for Inherited Hereditary Platelet Disorders for their contribution  
390 regarding clinical analyses and the GBiM platform for the sequencing and discussion.

391 **Authorship contributions**

392 TB, LH and DP performed bioinformatic analyses. EG, EA, VS and DB performed the culture and  
393 functional experiments. CL, MIK, ML and PS performed the clinical and biological  
394 characterization of patients. DPB and MCA conceived and supervised the project. DP and MP  
395 directed the project, designed the study, analyzed the data and wrote the manuscript.

396

397 **Disclosure of conflicts of interest**

398 The authors have declared that no conflict of interest exists.

399

400 **References**

- 401 1. Yang J, Luan J, Shen Y, Chen B. Developments in the production of platelets from stem cells (Review).  
402 *Mol Med Rep.* 2021;23(1):7.
- 403 2. Baigger A, Blasczyk R, Figueiredo C. Towards the Manufacture of Megakaryocytes and Platelets for  
404 Clinical Application. *Transfus Med Hemother.* 2017;44(3):165–173.
- 405 3. Borst S, Sim X, Poncz M, French DL, Gadue P. Induced Pluripotent Stem Cell-Derived Megakaryocytes  
406 and Platelets for Disease Modeling and Future Clinical Applications. *Arterioscler Thromb Vasc Biol.*  
407 2017;37(11):2007–2013.
- 408 4. Lambert MP, Sullivan SK, Fuentes R, French DL, Poncz M. Challenges and promises for the  
409 development of donor-independent platelet transfusions. *Blood.* 2013;121(17):3319–3324.
- 410 5. Pecci A, Balduini CL. Inherited thrombocytopenias: an updated guide for clinicians. *Blood Rev.*  
411 2021;48:100784.
- 412 6. Grodzielski M, Goette NP, Glembotsky AC, et al. Multiple concomitant mechanisms contribute to low  
413 platelet count in patients with immune thrombocytopenia. *Sci Rep.* 2019;9(1):2208.
- 414 7. Chen P-K, Chang H-H, Lin G-L, et al. Suppressive effects of anthrax lethal toxin on megakaryopoiesis.  
415 *PLoS One.* 2013;8(3):e59512.
- 416 8. Monzen S, Yoshino H, Kasai-Eguchi K, Kashiwakura I. Characteristics of myeloid differentiation and  
417 maturation pathway derived from human hematopoietic stem cells exposed to different linear  
418 energy transfer radiation types. *PLoS One.* 2013;8(3):e59385.
- 419 9. Lin G-L, Chang H-H, Lien T-S, et al. Suppressive effect of dengue virus envelope protein domain III on  
420 megakaryopoiesis. *Virulence.* 2017;8(8):1719–1731.
- 421 10. Hu L, Yin X, Zhang Y, et al. Radiation-induced bystander effects impair transplanted human  
422 hematopoietic stem cells via oxidative DNA damage. *Blood.* 2021;137(24):3339–3350.
- 423 11. Kollmann K, Warsch W, Gonzalez-Arias C, et al. A novel signalling screen demonstrates that CALR  
424 mutations activate essential MAPK signalling and facilitate megakaryocyte differentiation. *Leukemia.*  
425 2017;31(4):934–944.
- 426 12. Zeddies S, Jansen SBG, Summa F di, et al. MEIS1 regulates early erythroid and megakaryocytic cell  
427 fate. *Haematologica.* 2014;99(10):1555–1564.
- 428 13. Avanzi MP, Izak M, Oluwadara OE, Mitchell WB. Actin Inhibition Increases Megakaryocyte  
429 Proplatelet Formation through an Apoptosis-Dependent Mechanism. *PLOS ONE.*  
430 2015;10(4):e0125057.
- 431 14. Basak I, Bhatlekar S, Manne BK, et al. miR-15a-5p regulates expression of multiple proteins in the  
432 megakaryocyte GPVI signaling pathway. *J Thromb Haemost.* 2019;17(3):511–524.
- 433 15. Ramanathan G, Mannhalter C. Increased expression of transient receptor potential canonical 6  
434 (TRPC6) in differentiating human megakaryocytes. *Cell Biol Int.* 2016;40(2):223–231.
- 435 16. Song B, Miao W, Cui Q, et al. Inhibition of ferroptosis promotes megakaryocyte differentiation and  
436 platelet production. *Journal of Cellular and Molecular Medicine.* 2022;26(12):3582–3585.
- 437 17. Noetzli L, Lo RW, Lee-Sherick AB, et al. Germline mutations in ETV6 are associated with  
438 thrombocytopenia, red cell macrocytosis and predisposition to lymphoblastic leukemia. *Nat. Genet.*  
439 2015;47(5):535–538.

- 440 18. Di Paola J, Porter CC. ETV6-related thrombocytopenia and leukemia predisposition. *Blood*.  
441 2019;134(8):663–667.
- 442 19. Feurstein S, Godley LA. Germline ETV6 mutations and predisposition to hematological malignancies.  
443 *Int. J. Hematol.* 2017;106(2):189–195.
- 444 20. Porter CC, Di Paola J, Pencheva B. ETV6 Thrombocytopenia and Predisposition to Leukemia.  
445 *GeneReviews*®. 1993;
- 446 21. Wang LC, Swat W, Fujiwara Y, et al. The TEL/ETV6 gene is required specifically for hematopoiesis in  
447 the bone marrow. *Genes Dev.* 1998;12(15):2392–2402.
- 448 22. Wang LC, Kuo F, Fujiwara Y, et al. Yolk sac angiogenic defect and intra-embryonic apoptosis in mice  
449 lacking the Ets-related factor TEL. *EMBO J.* 1997;16(14):4374–4383.
- 450 23. Zhou C, Uluisik R, Rowley JW, et al. Germline ETV6 mutation promotes inflammation and disrupts  
451 lymphoid development of early hematopoietic progenitors. *Experimental Hematology*. 2022;112–  
452 113:24–34.
- 453 24. Borst S, Nations CC, Klein JG, et al. Study of inherited thrombocytopenia resulting from mutations in  
454 ETV6 or RUNX1 using a human pluripotent stem cell model. *Stem Cell Reports*. 2021;16(6):1458–  
455 1467.
- 456 25. Poggi M, Canault M, Favier M, et al. Germline variants in ETV6 underlie reduced platelet formation,  
457 platelet dysfunction and increased levels of circulating CD34+ progenitors. *Haematologica*.  
458 2017;102(2):282–294.
- 459 26. Saultier P, Vidal L, Canault M, et al. Macrothrombocytopenia and dense granule deficiency  
460 associated with FLI1 variants: ultrastructural and pathogenic features. *Haematologica*.  
461 2017;102(6):1006–1016.
- 462 27. Psaila B, Wang G, Rodriguez-Meira A, et al. Single-Cell Analyses Reveal Megakaryocyte-Biased  
463 Hematopoiesis in Myelofibrosis and Identify Mutant Clone-Specific Targets. *Mol Cell*.  
464 2020;78(3):477-492.e8.
- 465 28. Chen L, Kostadima M, Martens JHA, et al. Transcriptional diversity during lineage commitment of  
466 human blood progenitors. *Science*. 2014;345(6204):1251033.
- 467 29. Estevez B, Borst S, Jarocha DJ, et al. RUNX1 haploinsufficiency causes a marked deficiency of  
468 megakaryocyte-biased hematopoietic progenitor cells. *Blood*. 2021;
- 469 30. Pellin D, Loperfido M, Baricordi C, et al. A comprehensive single cell transcriptional landscape of  
470 human hematopoietic progenitors. *Nat Commun*. 2019;10:.
- 471 31. Aibar S, González-Blas CB, Moerman T, et al. SCENIC: single-cell regulatory network inference and  
472 clustering. *Nat Methods*. 2017;14(11):1083–1086.
- 473 32. Yim A, Koti P, Bonnard A, et al. mitoXplorer, a visual data mining platform to systematically analyze  
474 and visualize mitochondrial expression dynamics and mutations. *Nucleic Acids Res*. 2020;48(2):605–  
475 632.
- 476 33. Chatterjee N, Walker GC. Mechanisms of DNA damage, repair and mutagenesis. *Environ Mol*  
477 *Mutagen*. 2017;58(5):235–263.
- 478 34. Psaila B, Mead AJ. Single-cell approaches reveal novel cellular pathways for megakaryocyte and  
479 erythroid differentiation. *Blood*. 2019;133(13):1427–1435.
- 480 35. Adolfsson J, Månsson R, Buza-Vidas N, et al. Identification of Flt3+ Lympho-Myeloid Stem Cells  
481 Lacking Erythro-Megakaryocytic Potential: A Revised Road Map for Adult Blood Lineage  
482 Commitment. *Cell*. 2005;121(2):295–306.
- 483 36. Sanjuan-Pla A, Macaulay IC, Jensen CT, et al. Platelet-biased stem cells reside at the apex of the  
484 haematopoietic stem-cell hierarchy. *Nature*. 2013;502(7470):232–236.
- 485 37. Noetzli LJ, French SL, Machlus KR. New Insights into the Differentiation of Megakaryocytes from  
486 Hematopoietic Progenitors. *Arterioscler Thromb Vasc Biol*. 2019;39(7):1288–1300.

- 487 38. Xavier-Ferrucio J, Krause DS. Concise Review: Bipotent Megakaryocytic-Erythroid Progenitors:  
488 Concepts and Controversies. *Stem Cells*. 2018;36(8):1138–1145.
- 489 39. Haas S, Hansson J, Klimmeck D, et al. Inflammation-Induced Emergency Megakaryopoiesis Driven by  
490 Hematopoietic Stem Cell-like Megakaryocyte Progenitors. *Cell Stem Cell*. 2015;17(4):422–434.
- 491 40. Shin JY, Hu W, Naramura M, Park CY. High c-Kit expression identifies hematopoietic stem cells with  
492 impaired self-renewal and megakaryocytic bias. *J Exp Med*. 2014;211(2):217–231.
- 493 41. Miyawaki K, Iwasaki H, Jiromaru T, et al. Identification of unipotent megakaryocyte progenitors in  
494 human hematopoiesis. *Blood*. 2017;129(25):3332–3343.
- 495 42. Moreira S, Polena E, Gordon V, et al. A Single TCF Transcription Factor, Regardless of Its Activation  
496 Capacity, Is Sufficient for Effective Trilineage Differentiation of ESCs. *Cell Reports*. 2017;20(10):2424–  
497 2438.
- 498 43. Wu JQ, Seay M, Schulz VP, et al. Tcf7 is an important regulator of the switch of self-renewal and  
499 differentiation in a multipotential hematopoietic cell line. *PLoS Genet*. 2012;8(3):e1002565.
- 500 44. Thorsteinsdottir U, Mamo A, Kroon E, et al. Overexpression of the myeloid leukemia–  
501 associated Hoxa9 gene in bone marrow cells induces stem cell expansion. *Blood*. 2002;99(1):121–  
502 129.
- 503 45. Radomska HS, Huettner CS, Zhang P, et al. CCAAT/enhancer binding protein alpha is a regulatory  
504 switch sufficient for induction of granulocytic development from bipotential myeloid progenitors.  
505 *Mol Cell Biol*. 1998;18(7):4301–4314.
- 506 46. Shyamsunder P, Shanmugasundaram M, Mayakonda A, et al. Identification of a novel enhancer of  
507 CEBPE essential for granulocytic differentiation. *Blood*. 2019;133(23):2507–2517.
- 508 47. Lindsey S, Papoutsakis ET. The aryl hydrocarbon receptor (AHR) transcription factor regulates  
509 megakaryocytic polyploidization. *Br J Haematol*. 2011;152(4):469–484.
- 510 48. Strassel C, Brouard N, Mallo L, et al. Aryl hydrocarbon receptor-dependent enrichment of a  
511 megakaryocytic precursor with a high potential to produce proplatelets. *Blood*. 2016;127(18):2231–  
512 2240.
- 513 49. Smith BW, Rozelle SS, Leung A, et al. The aryl hydrocarbon receptor directs hematopoietic  
514 progenitor cell expansion and differentiation. *Blood*. 2013;122(3):376–385.
- 515 50. Tallack MR, Perkins AC. Megakaryocyte-erythroid lineage promiscuity in EKLF null mouse blood.  
516 *Haematologica*. 2010;95(1):144–147.
- 517 51. Tsang AP, Fujiwara Y, Hom DB, Orkin SH. Failure of megakaryopoiesis and arrested erythropoiesis in  
518 mice lacking the GATA-1 transcriptional cofactor FOG. *Genes Dev*. 1998;12(8):1176–1188.
- 519 52. Levin J, Peng J-P, Baker GR, et al. Pathophysiology of Thrombocytopenia and Anemia in Mice Lacking  
520 Transcription Factor NF-E2. *Blood*. 1999;94(9):3037–3047.
- 521 53. Ungerback J, Hosokawa H, Wang X, et al. Pioneering, chromatin remodeling, and epigenetic  
522 constraint in early T-cell gene regulation by SPI1 (PU.1). *Genome Res*. 2018;28(10):1508–1519.
- 523 54. Poirault-Chassac S, Nivet-Antoine V, Houvert A, et al. Mitochondrial dynamics and reactive oxygen  
524 species initiate thrombopoiesis from mature megakaryocytes. *Blood Advances*. 2021;5(6):1706–  
525 1718.
- 526 55. Tugolukova EA, Campbell RA, Hoerger KB, et al. Mitochondrial Fission Protein Drp1 Regulates  
527 Megakaryocyte and Platelet Mitochondrial Morphology, Platelet Numbers, and Platelet Function.  
528 *Blood*. 2017;130:455.
- 529 56. Bhoopalan V, Kaur A, Hein N, et al. Ribosomal biogenesis inhibition facilitates  
530 megakaryocyte/platelet biased haematopoiesis. *ISTH Congress Abstracts*. 2022;
- 531 57. Le Goff S, Boussaid I, Floquet C, et al. p53 activation during ribosome biogenesis regulates normal  
532 erythroid differentiation. *Blood*. 2021;137(1):89–102.

- 533 58. Eaton N, Boyd EK, Biswas R, et al. Endocytosis of the thrombopoietin receptor Mpl regulates  
534 megakaryocyte and erythroid maturation in mice. *Frontiers in Oncology*. 2022;12:.  
535 59. Kachaev ZM, Ivashchenko SD, Kozlov EN, Lebedeva LA, Shidlovskii YV. Localization and Functional  
536 Roles of Components of the Translation Apparatus in the Eukaryotic Cell Nucleus. *Cells*.  
537 2021;10(11):3239.  
538 60. McGowan KA, Pang WW, Bhardwaj R, et al. Reduced ribosomal protein gene dosage and p53  
539 activation in low-risk myelodysplastic syndrome. *Blood*. 2011;118(13):3622–3633.  
540 61. Keel SB, Phelps S, Sabo KM, et al. Establishing Rps6 hemizygous mice as a model for studying how  
541 ribosomal protein haploinsufficiency impairs erythropoiesis. *Exp Hematol*. 2012;40(4):290–294.

## 542 **Figure legends**

### 543 **Figure 1. Single-cell RNA sequencing analysis in megakaryopoiesis cell stages in CD34<sup>+</sup> cells** 544 **isolated from healthy volunteers.**

- 545 **A-** Schematic diagram of the experimental design. Fresh CD34<sup>+</sup> cells from human peripheral  
546 blood were isolated via density gradient and magnetic cell sorting (n=2 healthy volunteers).  
547 CD34<sup>+</sup> cells were differentiated into megakaryocytes using serum-free StemSpan SFEM II  
548 medium with megakaryocyte expansion supplement (SCF, IL6, TPO and IL9) (STEMCELL  
549 Technologies). Cells from different samples are incubated with DNA-barcoded antibodies  
550 recognizing ubiquitous cell surface proteins. Distinct barcodes (referred to as hashtag-  
551 oligos, HTO) on the antibodies allow pooling of multiple samples into one single-cell RNA  
552 (scRNA) sequencing experiment. The cells were analyzed at differentiation days 6 and 11  
553 using 10X genomics..
- 554 **B-** UMAP plot of the control cells on days 6 and 11 after merging the two data sets (D6 and  
555 D11, respectively). The cells analyzed on day 6 are shown in orange, and the cells analyzed  
556 on day 11 are shown in blue.
- 557 **C-** UMAP plot of the control cells on days 6 and 11 (D6 and D11, respectively), with cell  
558 clustering performed using a resolution of 1.2. Each cluster is indicated with a distinct color.
- 559 **D-** Seurat dot plot showing the expression levels of known specific hematopoietic cell marker  
560 genes in each cluster.
- 561 **E-** Seurat dot plot of the top five differentially expressed genes in each cluster. Dot size  
562 represents the percentage of cells expressing the gene of interest (percent expressed), while  
563 the color gradient represents the scaled average expression of the genes in each cluster (a  
564 negative value corresponds to expression levels below the mean).

565 F- Feature plot showing the expression of genes representative of each cell type (*PROM1*,  
566 *PRG2*, *MPO*, *KIT*, *LAT*, *CAVIN2*, *ITGA2B*, *HBB*, *PF4* and *CD34*).

567 **Figure 2. CD34<sup>+</sup> cell-derived cultures contain cells of various hematopoietic states.**

568 A- UMAP plot showing RNA count for hematopoietic cells derived from normal CD34<sup>+</sup> cells.

569 B- UMAP plot showing the number of genes expressed for hematopoietic cells derived from  
570 normal CD34<sup>+</sup> cells.

571 C- Feature plots showing the expression of lineage signature gene set score for hematopoietic  
572 cells derived from normal CD34<sup>+</sup> cells.

573 D- UMAP plot of hematopoietic cells derived from normal CD34<sup>+</sup> cells color-coded according  
574 to cell type. Numbers correspond to clusters.

575 E- UMAP plots displaying Top 3 regulon activity enriched in each lineage: HSPC (*TCF7L2*,  
576 *TCF4* and *HOXA9*), GMP (*CEBPD*, *ZNF502* and *CEBPE*), CMP (*GATA2*, *CEBPD* and  
577 *FOXJ1*), MK-primed CMP (*ZMAT4*, *AHR* and *BATF*), MEP (*KLF1*, *MECOM* and *SPI1*)  
578 and MKP/MK (*GATA1*, *MAFG* and *NFE2*).

579 **Figure 3. Single-cell RNA sequencing of CD34<sup>+</sup> cell-derived MK-induced cells from patients  
580 with a *ETV6* variant.**

581 A- UMAP plot of *ETV6*-variant cells at days 6 and 11 after merging the two data sets. Cells  
582 analyzed at day 6 are displayed in orange, and cells analyzed at day 11 are displayed in  
583 blue.

584 B- UMAP plot of cell clusters using a resolution of 0.8. Each of the 10 clusters is displayed  
585 with a distinct color.

586 C- Seurat dot plot showing the average relative expression of several hematopoietic cell  
587 marker genes in each cluster.

588 D- Seurat dot plot of the top five differentially expressed genes by cluster. Dot size represents  
589 the percentage of cells expressing the gene of interest, while dot color represents the scaled  
590 average gene expression (negative average expression indicates expression levels below the  
591 mean).

592 E- Feature plot of the expression profile of lineage signature gene sets.

593 F- UMAP plot of hematopoietic cells derived from *ETV6*-variant CD34<sup>+</sup> cells color-coded by  
594 cell type.

595 **Figure 4. Megakaryocyte differentiation in healthy controls and *ETV6* patients.**

- 596 A. Pseudotime ordering was performed using Slingshot to reconstruct the hierarchical structure of  
597 hematopoietic cells derived from healthy controls. Cells are color-coded by type (HSPC, CMP,  
598 MK-primed CMP, GMP, MEP, MKP/MK and platelets (Plt)).
- 599 B. Pseudotime ordering was performed using Slingshot to reconstruct the hierarchical structure of  
600 hematopoietic cells derived from *ETV6* patients. Cells are color-coded by type (HSPC, CMP,  
601 MK-primed CMP, GMP, MEP and MKP/MK).
- 602 C. Dynamic expression of immature and differentiation markers along the differentiation  
603 trajectory from HSPC to MK in healthy controls (upper panel) and *ETV6* patients (lower panel).
- 604 D. Dynamic expression of transcription factor genes along the differentiation trajectory from  
605 HSPC to MK. The x-axis shows pseudotime estimated by fitGAM, while the y-axis shows  
606 normalized gene expression.

607 **Figure 5. Single-cell RNA sequencing of cells derived from *ETV6*-variant carriers and healthy  
608 controls.**

- 609 A- UMAP plot of hematopoietic cells derived from control and *ETV6*-variant CD34<sup>+</sup> cells at  
610 days 6 and 11. Cells are color-coded by cell type (HSPC, CMP, MK-primed CMP, GMP,  
611 MEP, MKP/MK and platelets (Plt)).
- 612 B- UMAP plot of hematopoietic cells derived from control and *ETV6*-variant CD34<sup>+</sup> cells at  
613 days 6 and 11. Cells analyzed at day 6 are indicated in orange, and cells analyzed at day  
614 11 are indicated in blue.
- 615 C- UMAP plot showing one cell type per plot. *ETV6* patient cells are indicated in blue, and  
616 control cells are indicated in orange.
- 617 D- Radar plot showing the percentage of cell types for each genotype (*ETV6* patients =  
618 yellow line; control volunteers = blue line).
- 619 E- The same radar plot shown in Figure 5D with a zoom on maximal values. The maximal  
620 value is indicated for each cell type.

621 **Figure 6. Differentially expressed genes in each cell type.**

- 622 A- Volcano plots of upregulated (red) and downregulated (blue) genes in *ETV6* patients vs.  
623 controls for each cell type (HSPC, CMP, MK-primed CMP, GMP, MEP and MKP/MK),  
624 FC > 1.3 [i.e., log<sub>2</sub> FC > 0.37], adjusted p-value < 0.05).

625 B- Bubble plot of the top enriched GO biological processes based on differential gene  
626 expression by cell type (classified by p values). The upper panel shows the upregulated  
627 pathways, while the lower panel shows the downregulated pathways. Cells are color-coded  
628 by cell type (HSPC, CMP, MK-primed CMP, GMP, MEP, MKP/MK)

629 **Figure 7. Deregulated pathways in *ETV6*-variant MKP/MK.**

630 A- Ingenuity pathway schematic diagram of *ETV6* variant-mediated gene modulation in the  
631 mitochondrial electron transport chain. The color gradient displays the gene expression fold  
632 change (green = downregulated, white = unaltered, and red = upregulated).

633 B- Translation levels in HEL cells in basal conditions and after PMA (phorbol 12-myristate  
634 13-acetate) stimulation (NT = non-transduced cells, *ETV6* WT = cells transduced with wild  
635 type *ETV6*, and *ETV6* P214L = cells transduced with the *ETV6* P214L variant). Translation  
636 levels were evaluated via flow cytometry by assessing protein synthesis after incorporation  
637 of puromycin and staining with monoclonal anti-puromycin-APC antibody. \* $P < 0.05$ ,  
638 \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , one-sample t-test.

639 C- Translation levels in CD34<sup>+</sup> cell-derived MK at day 11 of culture in controls and P214L  
640 variant cells (F1 III-3 and IV1, F1III-8) and day 14 for F417Lter4 variant cells (F2 II-2).  
641 Translation levels were evaluated via flow cytometry by assessing protein synthesis after  
642 incorporation of puromycin and staining with monoclonal anti-puromycin-FITC antibody.

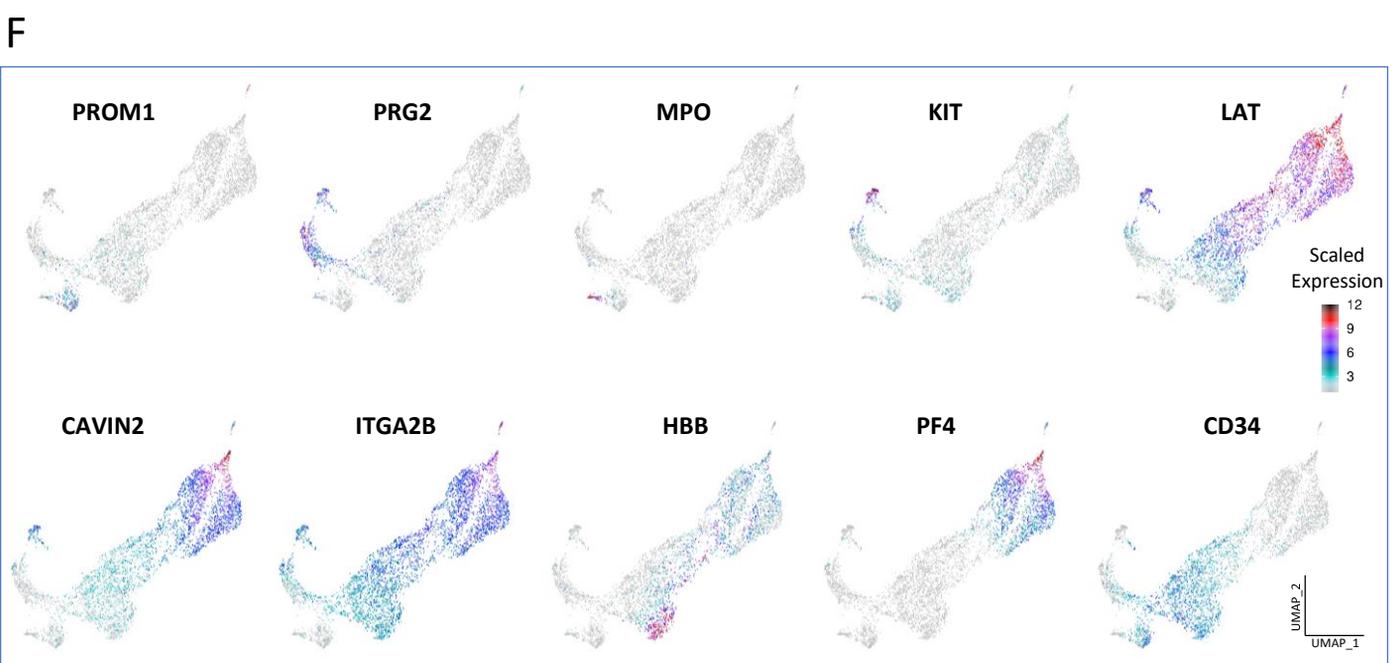
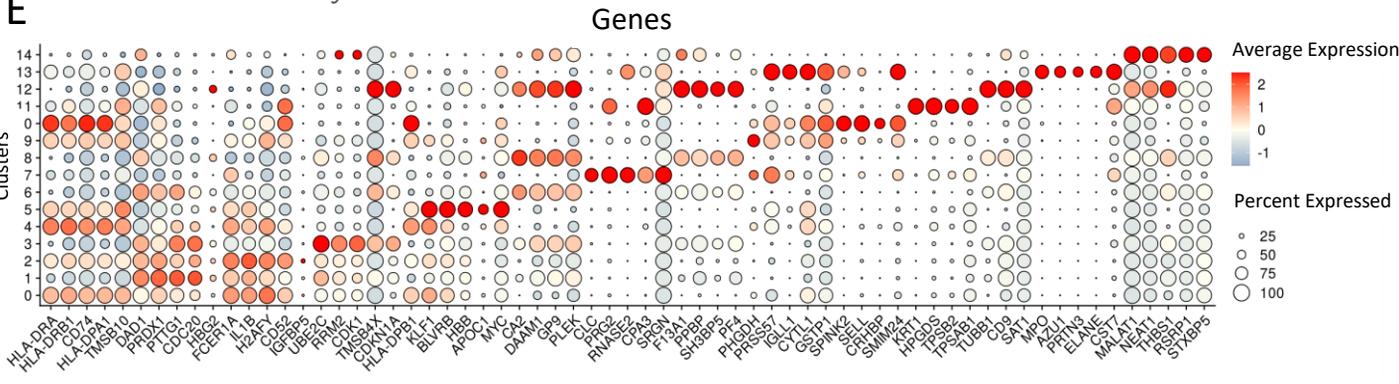
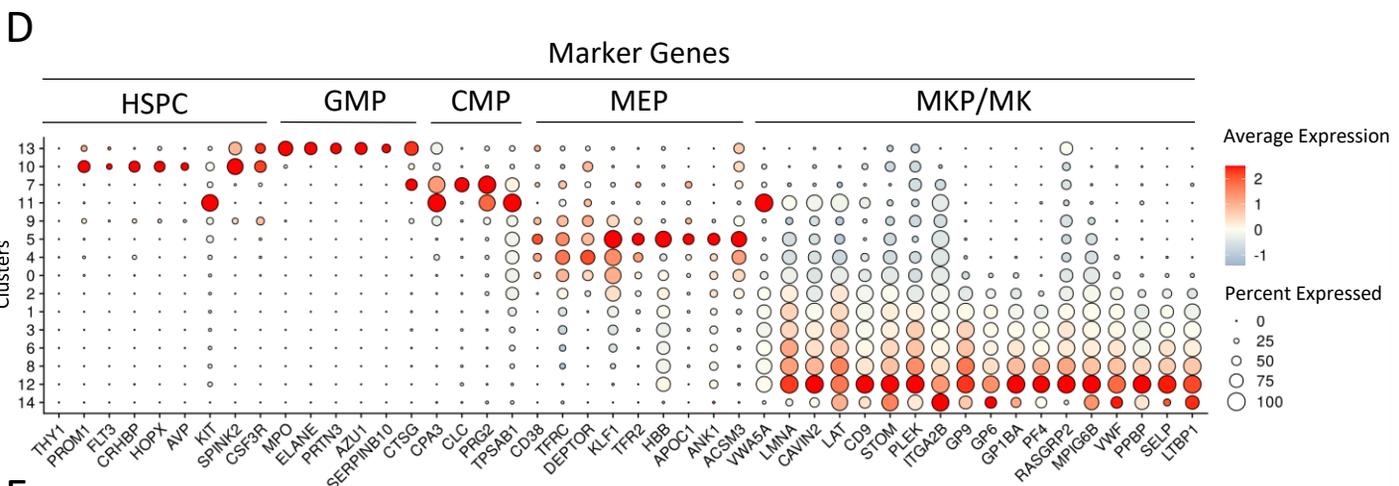
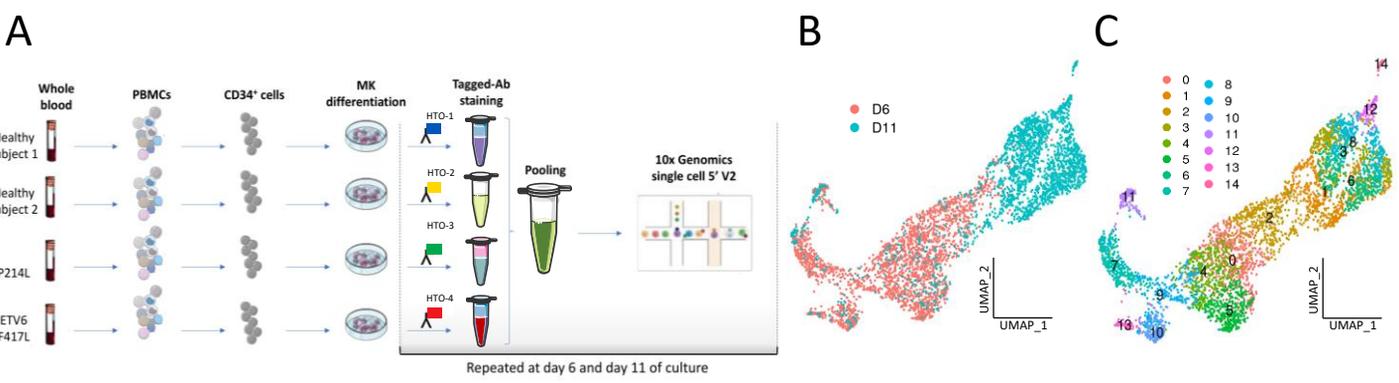
643 D- Venn diagram of upregulated genes with an *ETV6* binding site (blue circle) and ribosomal  
644 genes (yellow circle) in MKP/MK isolated from *ETV6*-patients vs controls. Among all  
645 upregulated genes carrying an *ETV6* binding site, 35 genes were associated with translation  
646 pathways.

647 E- Flow cytometry assessment of RPS6 and phosphoRPS6 in CD34<sup>+</sup> cell-derived MK from  
648 F1III-8 and four healthy controls (analyzed at the same time, at day 14) and RPS6 CD34<sup>+</sup>  
649 cell-derived MK from F1III-3 and three healthy controls (analyzed at the same time, at day  
650 13). RPS6 was analyzed by gating on CD42a<sup>+</sup> cells.

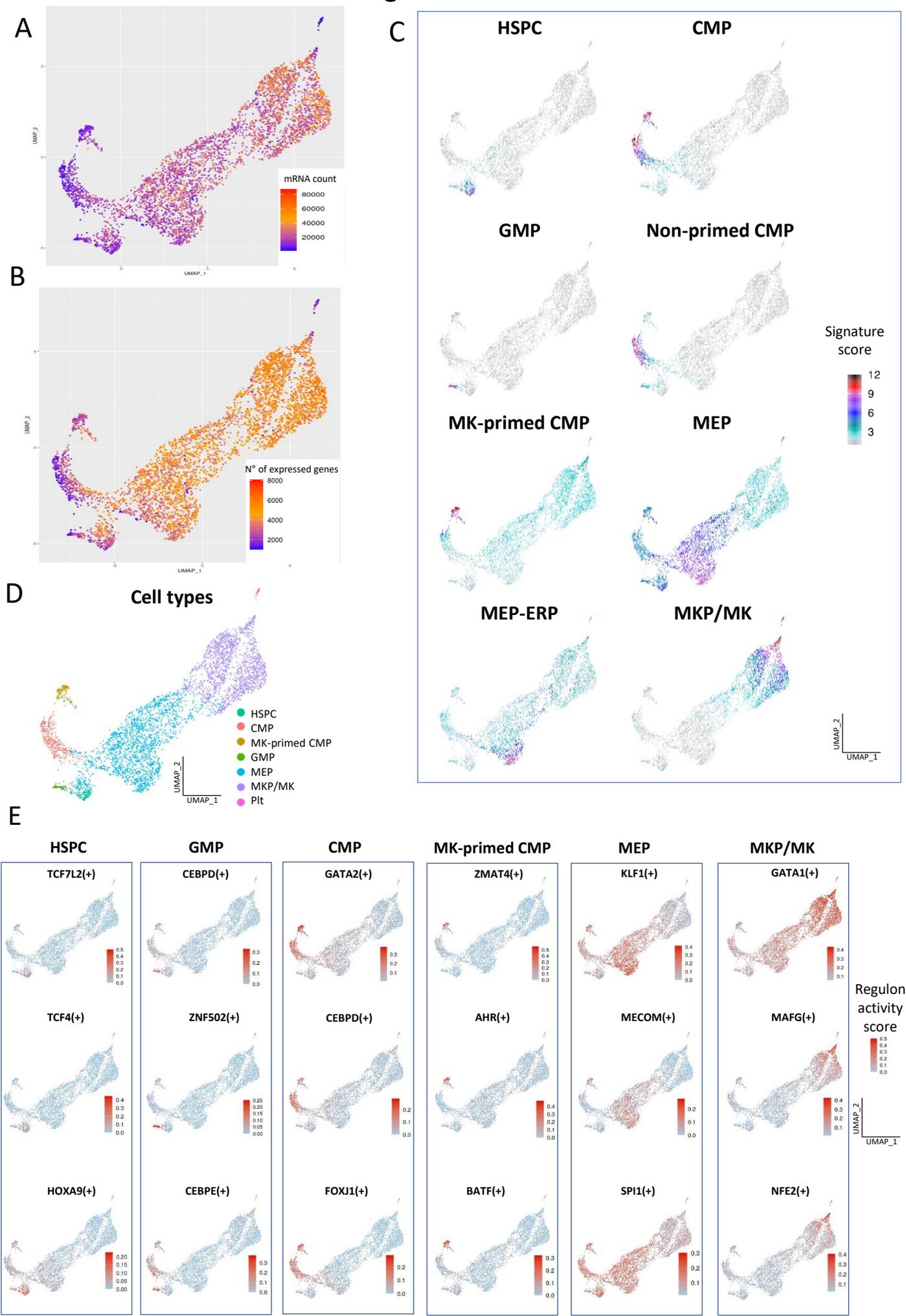
651 F- Western blot analysis of RPS6 expression in washed platelets from healthy controls and  
652 patients harboring *ETV6* variants. GAPDH was used as a protein loading control.  
653 Quantification of band intensity is shown on the right.

654 G- Flow cytometry assessment of RPS6 in peripheral blood mononuclear cells (PBMC) from  
655 five controls and five *ETV6* patients (F1-II4, III3, III8, VI1 and F2-II2). PBMC were frozen  
656 on the day of patient consultation and then thawed and cultured for four days before RPS6  
657 analysis (all samples were analyzed at the same time). \*  $P < 0.05$ , Mann-Whitney t-test.  
658 H- Seurat dot plot of the DEG in DNA repair pathways in MEP and MKP/MK cell populations  
659 (patient and control samples). Dot size represents the percentage of cells expressing the  
660 gene of interest, while dot color represents the scaled average gene expression (negative  
661 average expression indicates expression levels below the mean).

# Figure 1

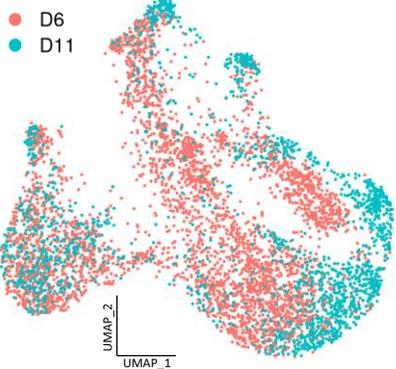


# Figure 2

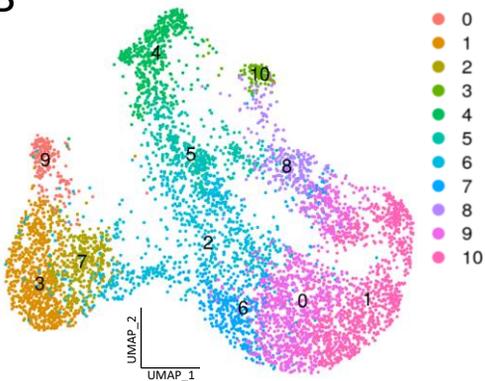


# Figure 3

**A**

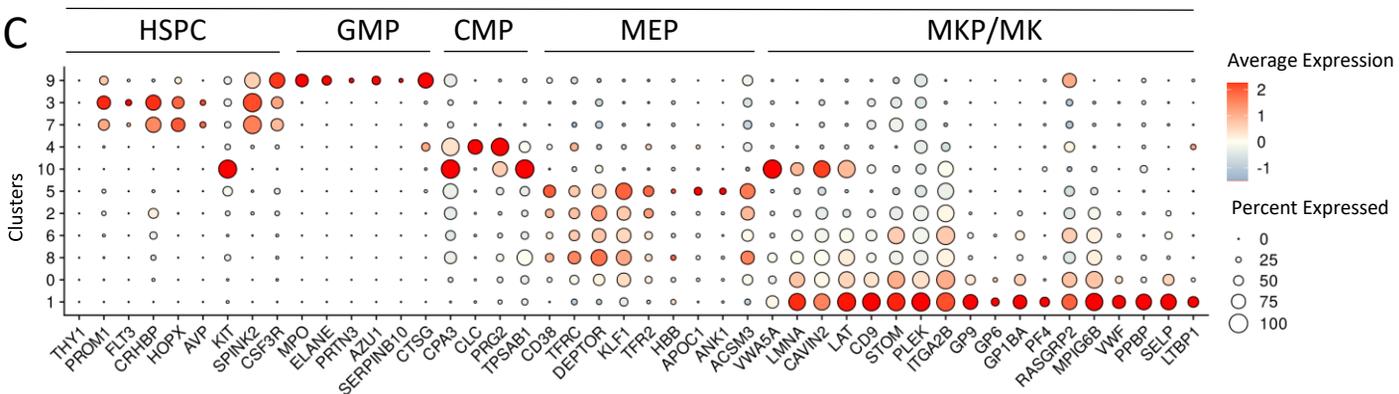


**B**

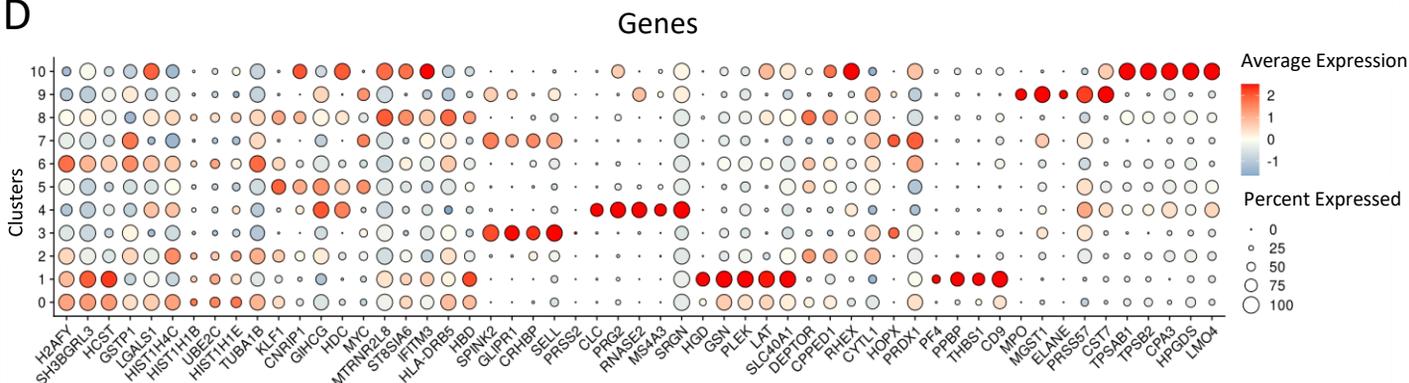


Marker Genes

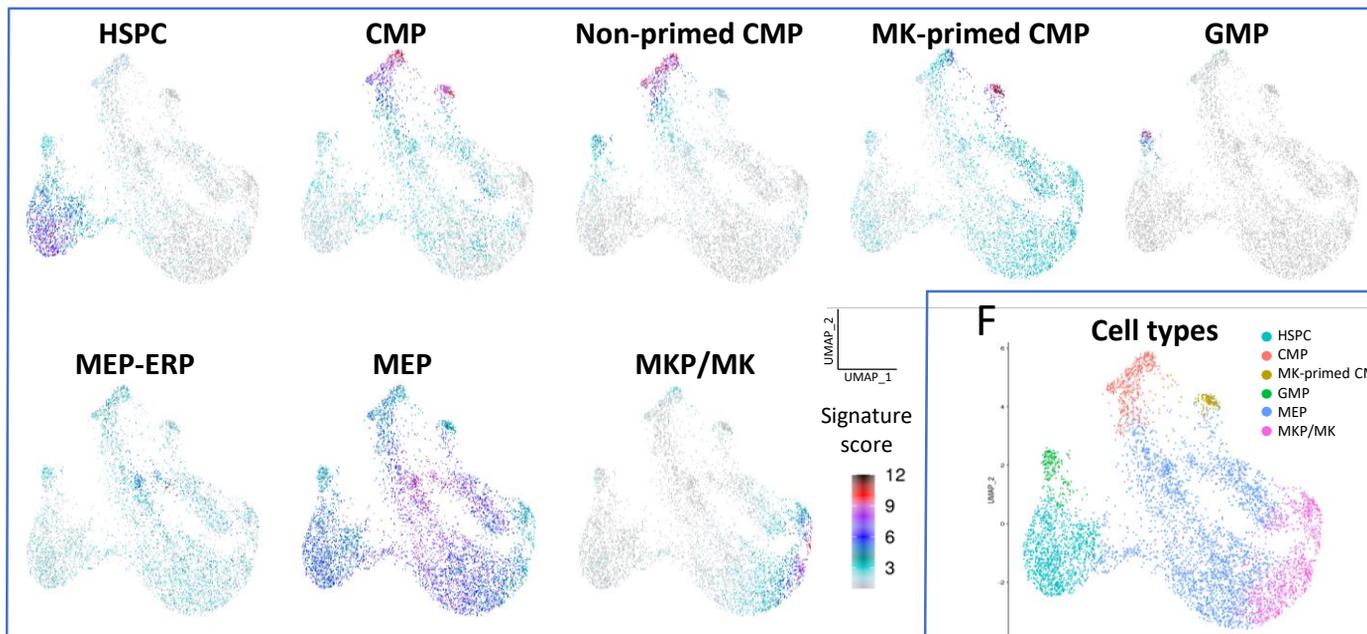
**C**



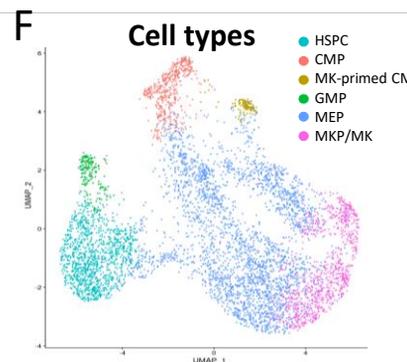
**D**



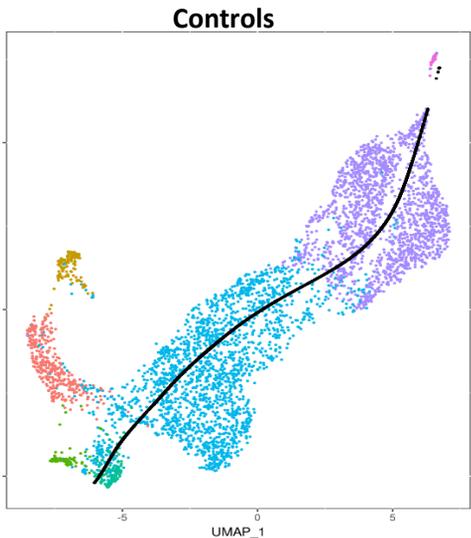
**E**



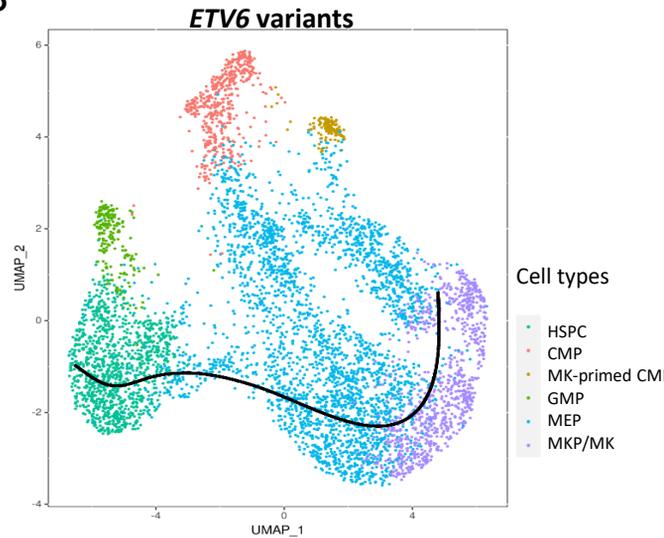
**F**



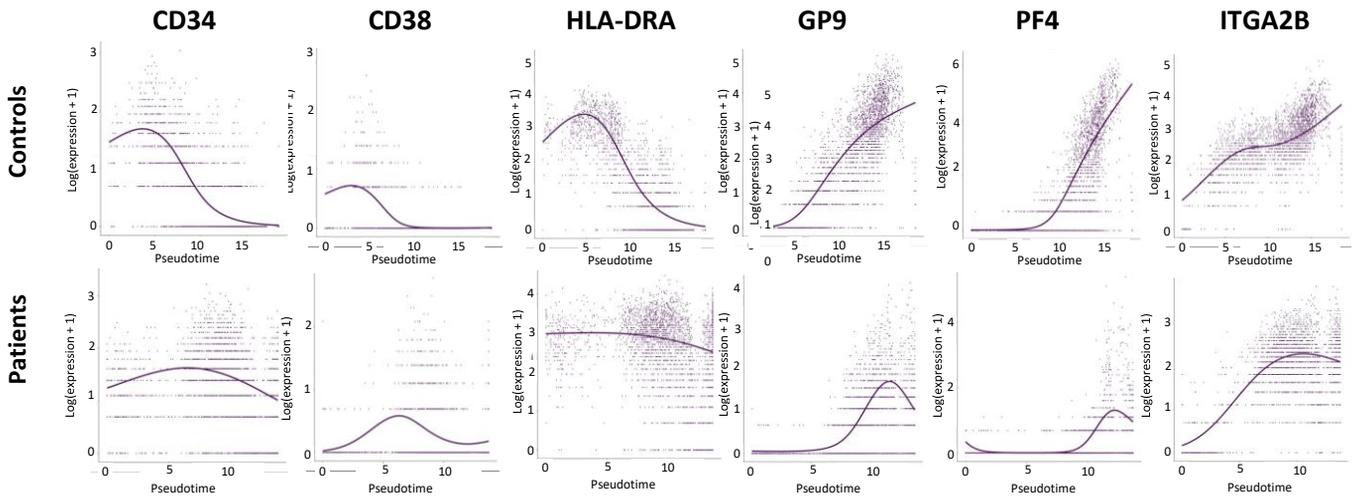
A



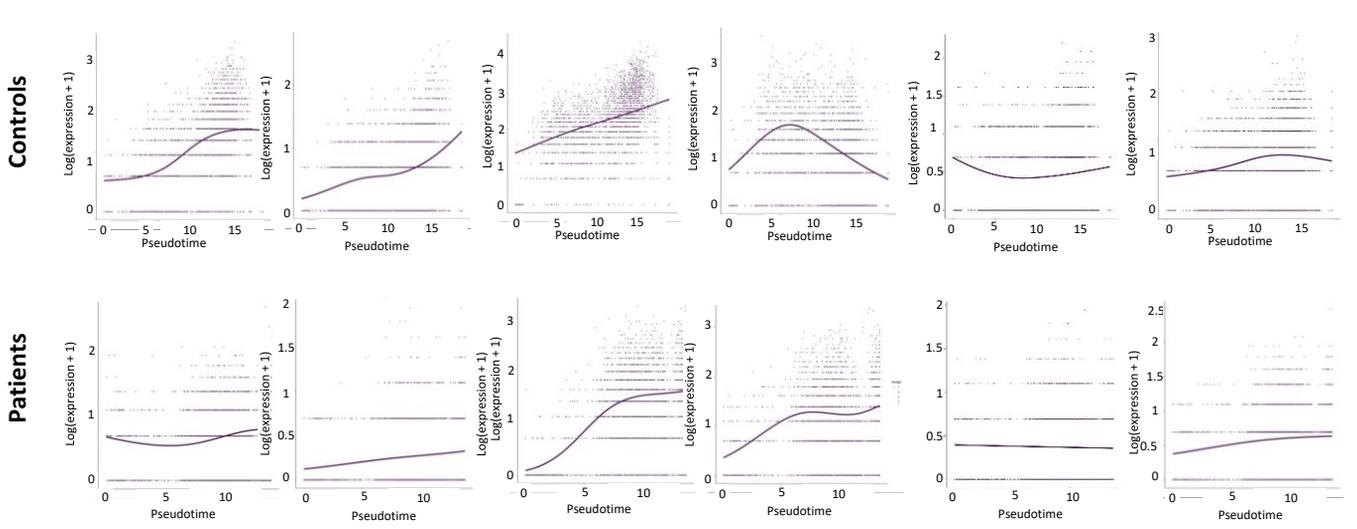
B



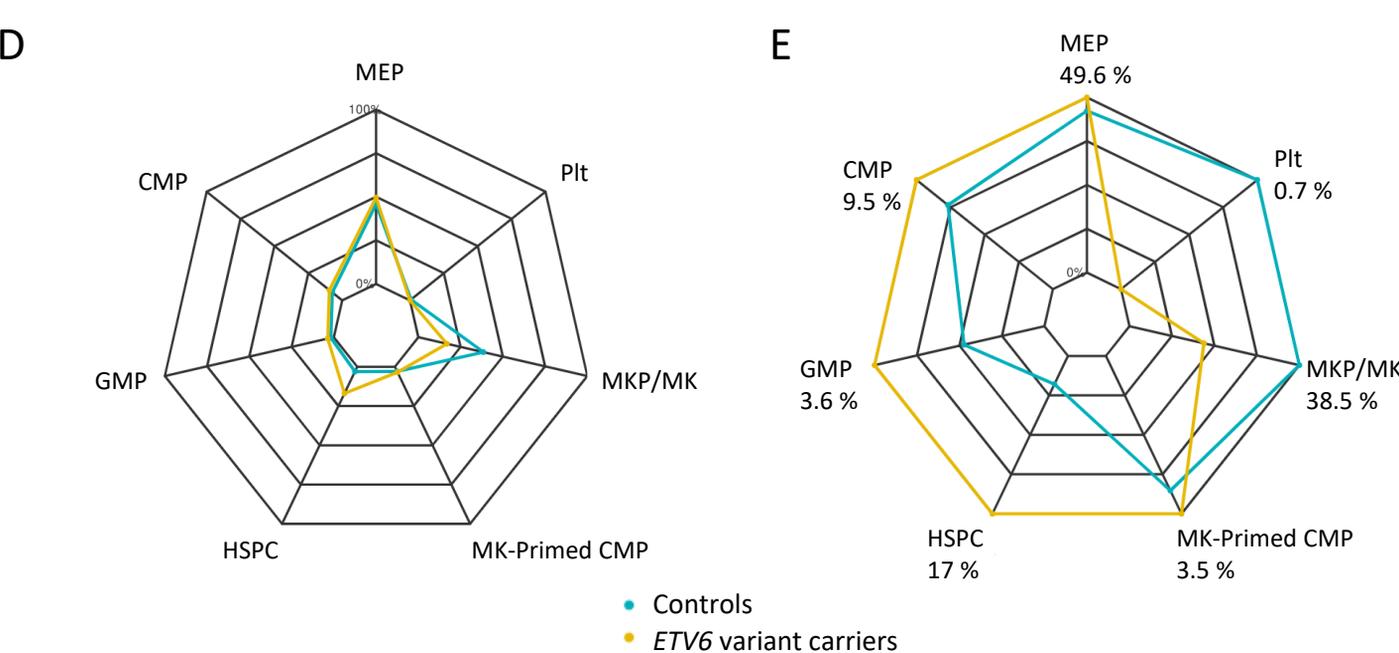
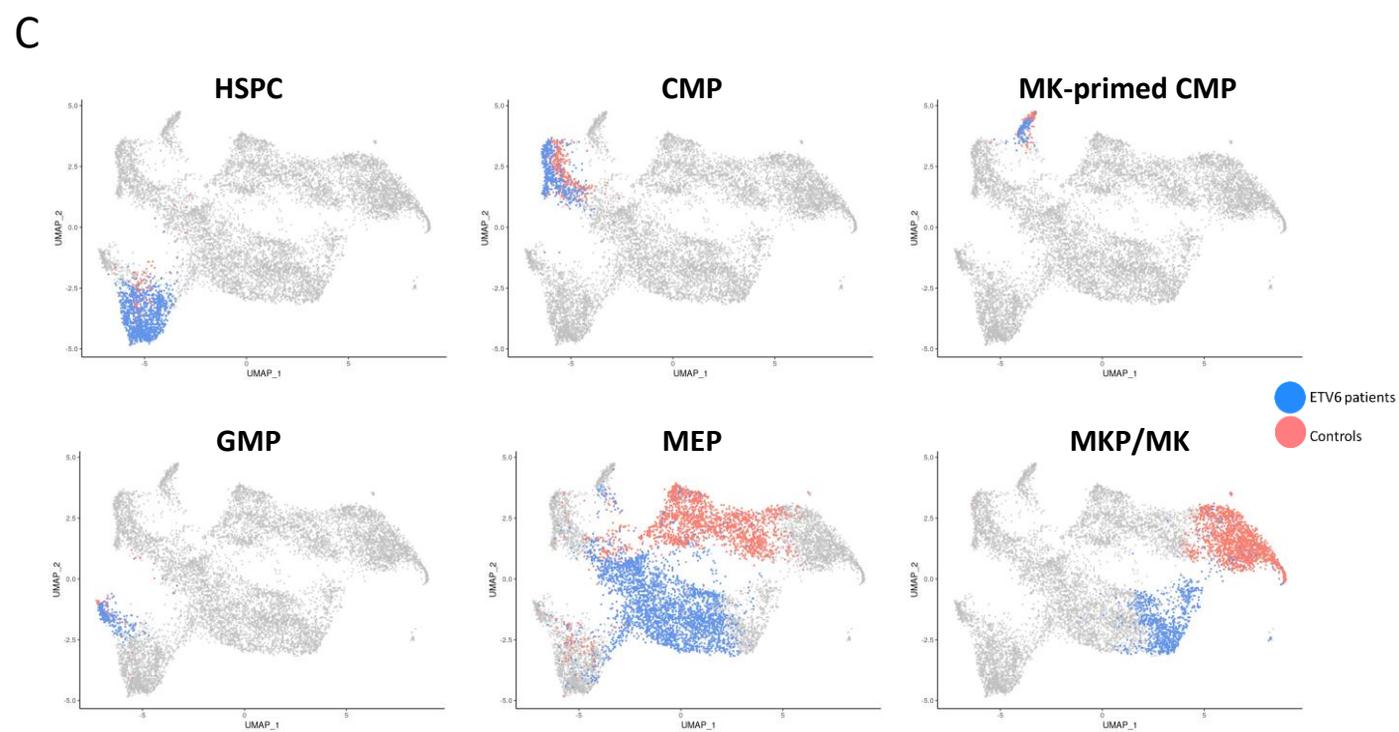
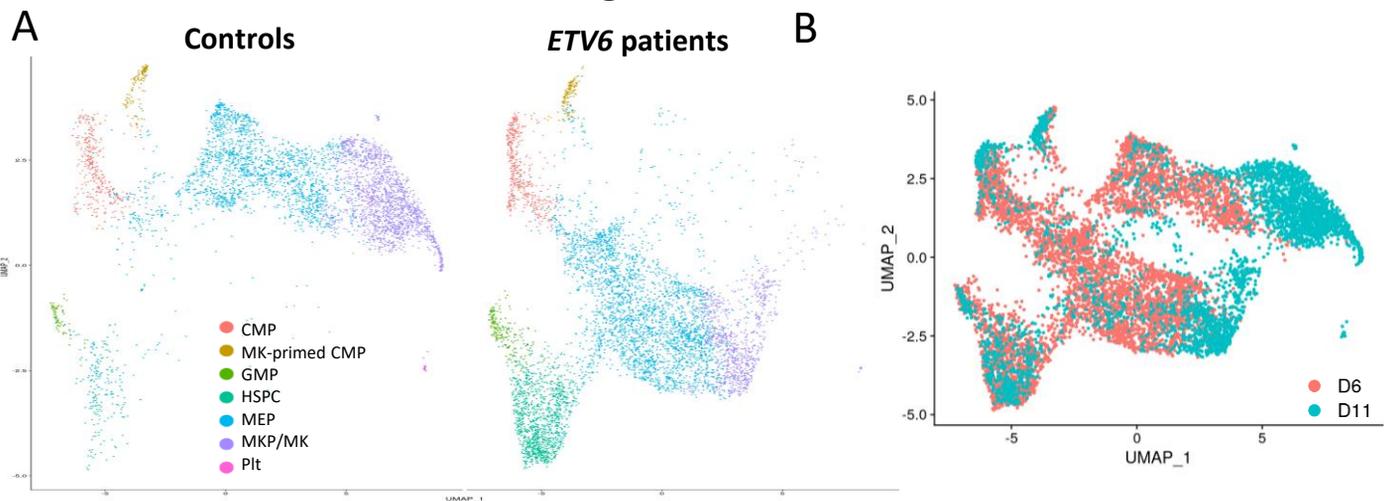
C



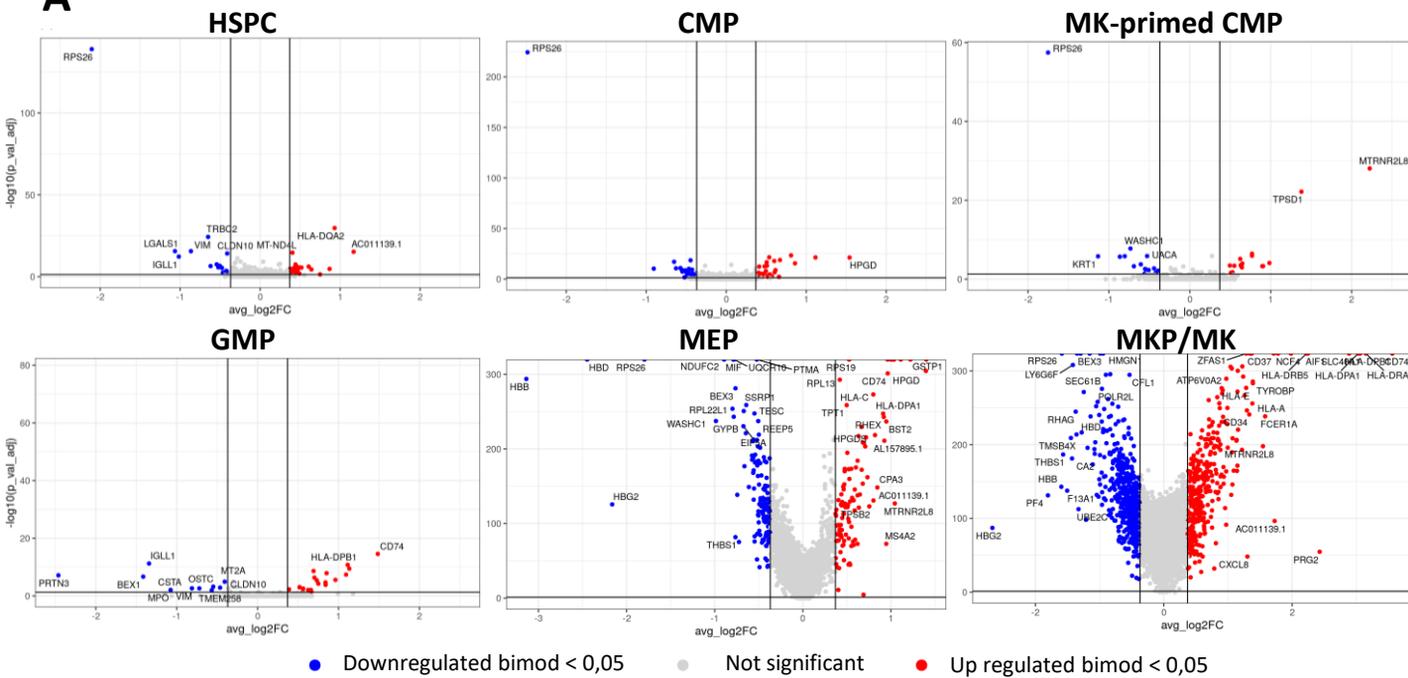
D



# Figure 5



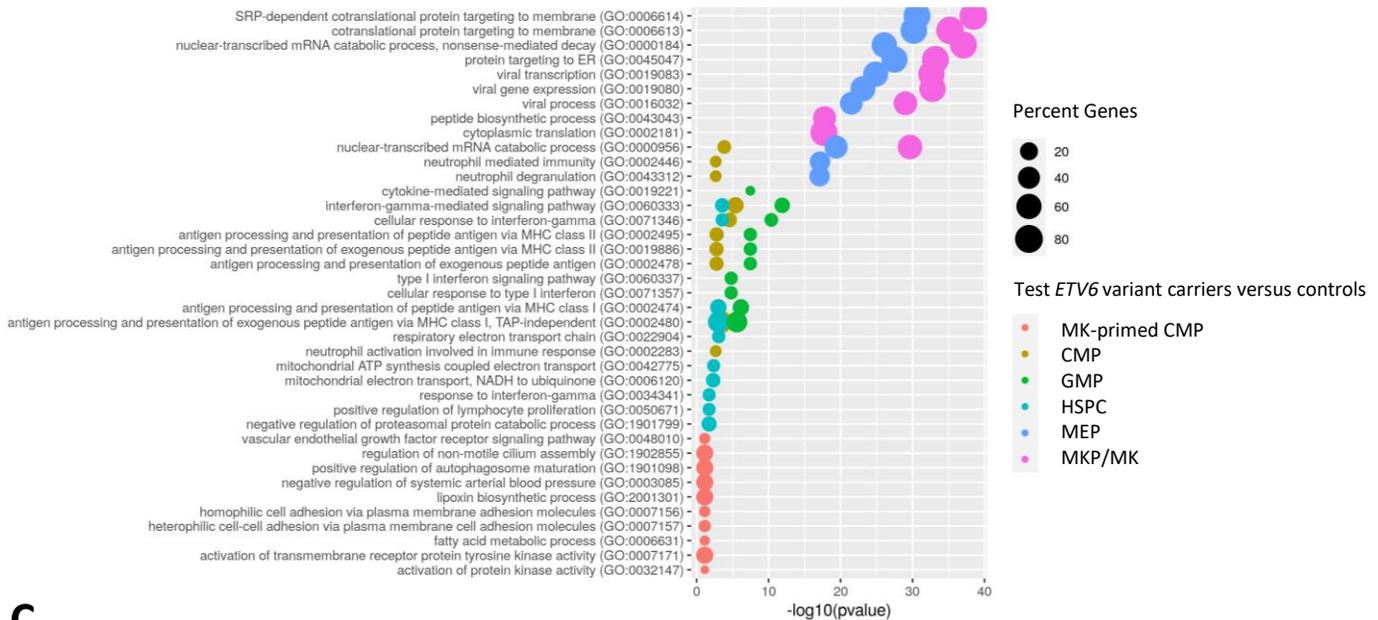
**A**



● Downregulated bimod <math>< 0,05</math>    ● Not significant    ● Up regulated bimod <math>< 0,05</math>

**B**

**Enrichment analysis of GO biological process upregulated in patient cells**



**C**

**Enrichment analysis of GO biological process downregulated in patient cells**

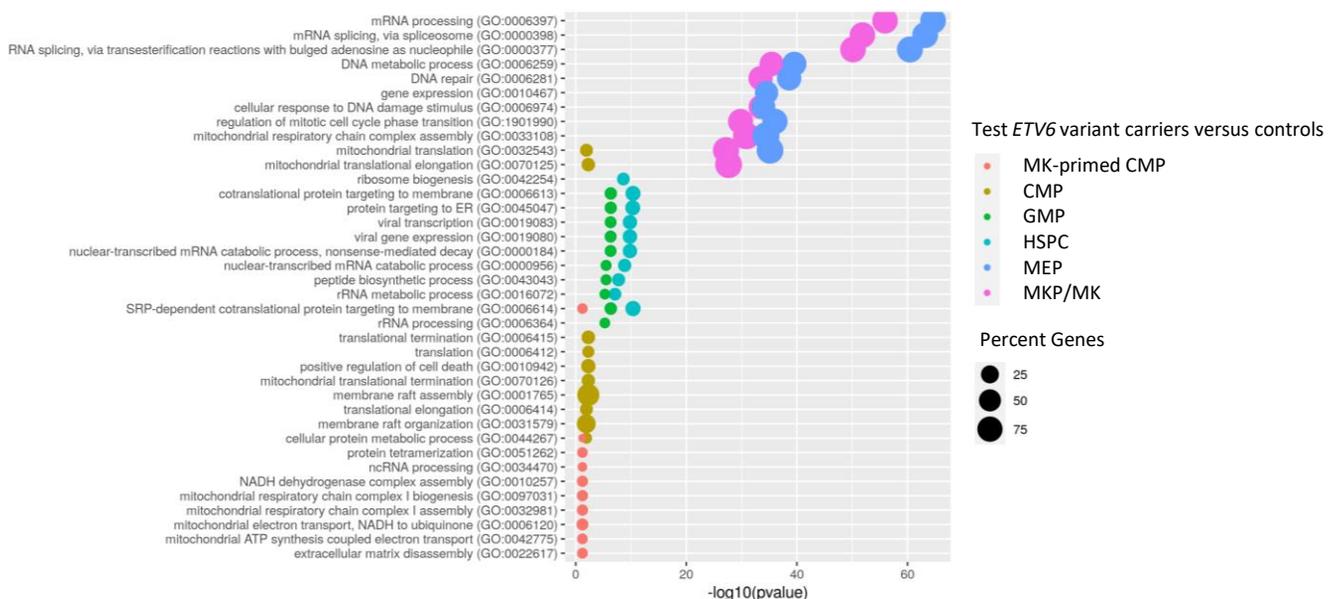
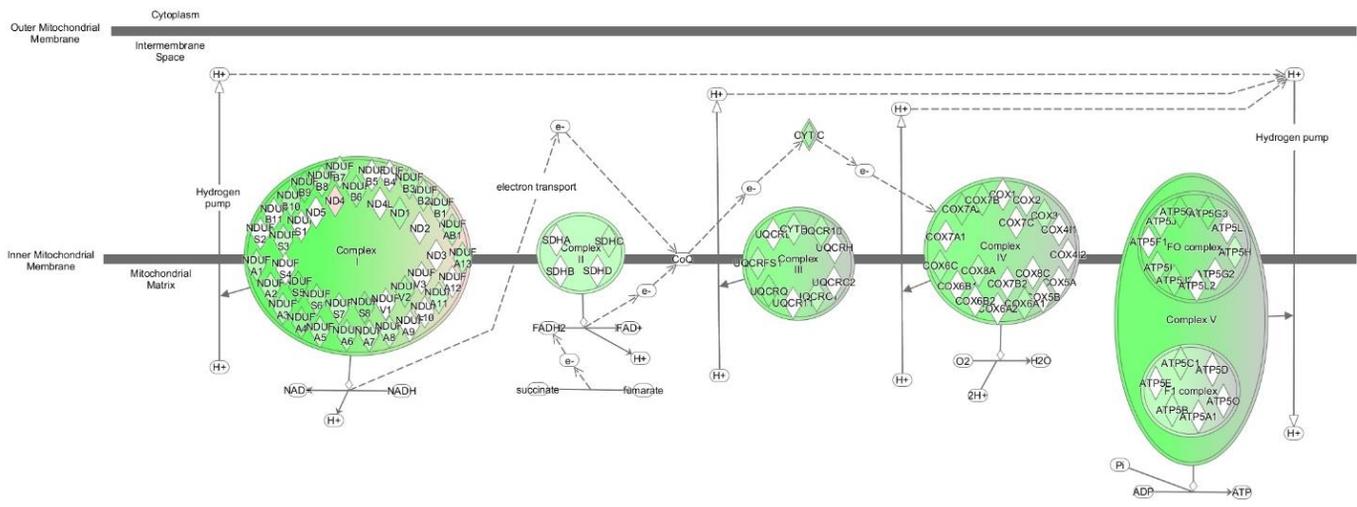


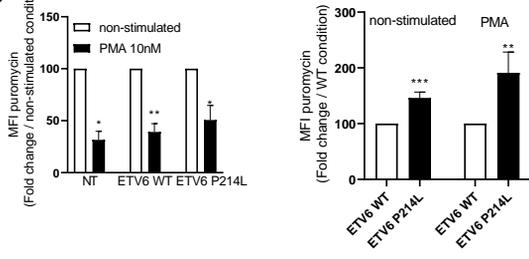
Figure 7

A



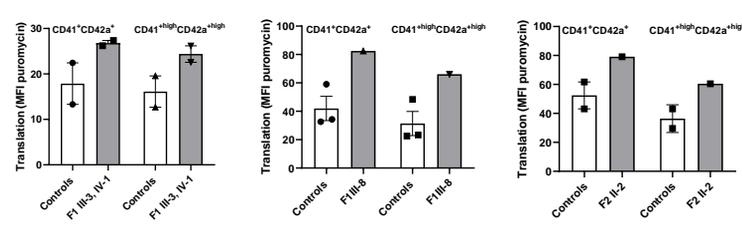
B

HEL cell line

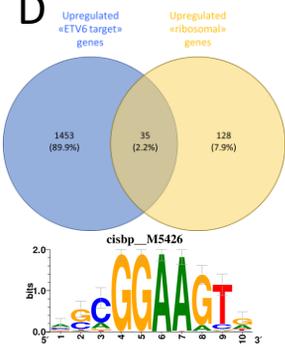


C

CD34<sup>+</sup>-derived MK

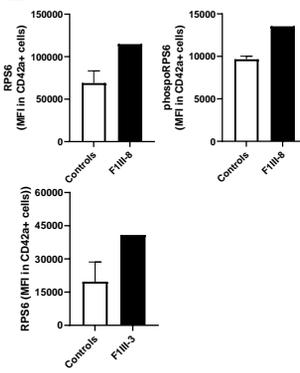


D



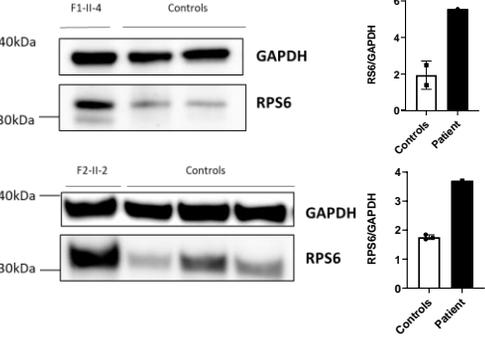
E

CD34<sup>+</sup>-derived MK



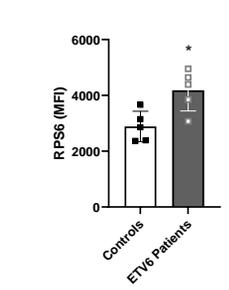
F

Platelets



G

PBMC



H

Marker Genes (DNA repair pathways)

