



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, Denis Bernot, Manal Ibrahim-Kosta, Pierre-Emmanuel Morange, 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. *Journal of Thrombosis and Haemostasis*, 2023, 21 (9), pp.2528-2544. 10.1016/j.jtha.2023.04.007 . hal-04275159

HAL Id: hal-04275159

<https://hal.science/hal-04275159v1>

Submitted on 8 Nov 2023

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.

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

3 Timothée Bigot^{1,†}, Elisa Gabinaud^{1,†}, Laurent Hannouche¹, Véronique Sbarra¹, Elisa
4 Andersen¹, Delphine Bastelica¹, Céline Falaise², Denis Bernot¹, Manal Ibrahim-Kosta¹,
5 Morange Pierre-Emmanuel¹, Marie Loosveld², Paul Saultier¹, Dominique Payet-Bornet^{3,*,}
6 Marie-Christine Alessi^{1,2,*,}, Delphine Potier^{2,*,} and Marjorie Poggi^{1,*,#}

7 ¹Aix Marseille Univ, INSERM, INRAe, C2VN, Marseille, France

8 ²APHM, C.H.U. Timone, French Reference Center on Inherited Platelet Disorders, Marseille, France

9 ³Aix-Marseille Univ, CNRS, INSERM, CIML, Marseille, France

10 ^{†, *} These authors contributed equally to this work.

11 # Corresponding author.

12 ***Running Head:*** Single cell analysis of ETV6-megakaryopoesis

13 ***Correspondence:*** Marjorie Poggi, C2VN, Faculté de Médecine, 27 Boulevard Jean Moulin,
14 13385 Marseille, France; E-mail: marjorie.poggi@univ-amu.fr; Phone: +33 491 324 452.

15 **Abstract: 225; Main text: 4169; Figures: 7; Tables: 0; Supplementary data: 10 Figures**
16 **and 8 Tables.**

17 **Essentials:**

- 18 • ETV6-related thrombocytopenia mechanisms remain poorly understood
- 19 • Megakaryopoiesis was studied by single-cell RNA sequencing and defects were
20 confirmed ex-vivo
- 21 • ETV6 deficiency led to the development of aberrant haemopoietic cells as early as the
22 MEP stage
- 23 • Upregulation of ribosome biogenesis were documented in patient megakaryocytes and
24 platelets

25 **Abstract**

26 **Background.** Germline mutations in the ETV6 transcription factor gene are responsible for
27 familial thrombocytopenia and leukemia predisposition syndrome. Although previous studies
28 have shown that ETV6 plays an important role in megakaryocyte (MK) maturation and
29 platelet formation, the mechanisms by which ETV6 dysfunction promotes thrombocytopenia
30 remain unclear.

31 **Objective and methods.** Presuming that ETV6 mutations result in selective effects at a
32 particular cell stage, we applied single-cell RNA sequencing to understand gene expression
33 changes during megakaryopoiesis in peripheral CD34+ cells from healthy controls and
34 patients with ETV6-related thrombocytopenia.

35 **Results.** Analysis of gene expression and regulon activity revealed distinct clusters partitioned
36 into seven major cell stages: hematopoietic stem/progenitor cells (HSPCs), common-myeloid
37 progenitors (CMPs), MK-primed CMPs, granulocyte-monocyte progenitors, megakaryocyte-
38 erythroid progenitors (MEPs), progenitor MKs/mature MKs (MKPs/MKs) and platelet-like
39 particles (PLPs). We observed a differentiation trajectory in which MEPs developed directly
40 from HSPCs and bypassed the CMP stage. ETV6 deficiency led to the development of
41 aberrant cells as early as the MEP stage, which intensified at the MKP/MK stage, with a
42 highly deregulated core “ribosome biogenesis” pathway. Indeed, increased translation levels
43 have been documented in patient CD34+-derived MKs with overexpression of ribosomal
44 protein S6 (RPS6) and phosphorylated-RPS6 in both CD34+-derived MKs and platelets.
45 Treatment of patient MKs with the ribosomal biogenesis inhibitor CX-5461 resulted in an
46 increase in PLPs.

47 **Conclusions.** These findings provide novel insight into both megakaryopoiesis and the link
48 between ETV6, translation and platelet production.

49 **Keywords**

50 ETV6, Megakaryopoiesis, Ribosomal Protein S6, Single-cell RNA sequencing,
51 Thrombocytopenia

52

53 **Introduction**

54 ETV6-related thrombocytopenia (ETV6-RT) is a highly penetrant form of inherited
55 thrombocytopenia with autosomal dominant inheritance [1]. The common phenotype
56 observed in ETV6-RT includes moderate thrombocytopenia sometimes associated with
57 bleeding and predisposition to acute T- or B-cell lymphoblastic leukemia. The precise role
58 that *ETV6* plays in megakaryocyte (MK) differentiation and pro-platelet formation currently
59 remains poorly understood. Studying *Etv6* function in murine models is challenging because
60 complete loss of the gene is lethal [2] and heterozygous *Etv6* mice have unperturbed
61 hematopoiesis [3]. Human induced pluripotent stem cells (iPSCs) harboring a pathogenic
62 heterozygous *ETV6* mutation do not give rise to an increase in hematopoietic progenitor cells
63 and MKs. However, iPSCs carrying the homozygous *ETV6* mutation give rise to an increase
64 in hematopoietic progenitor cells and immature MKs, as observed in heterozygous patients
65 using an *in vitro* model of CD34⁺-derived MKs [4,5]. Therefore, CD34⁺-derived MK
66 differentiation appears to be a valuable model to understand defective megakaryopoiesis
67 associated with aberrant ETV6. This model has been successfully used to characterize
68 acquired and inherited thrombocytopenia [6], investigate novel mechanisms in MK
69 differentiation and platelet function [7–10], and evaluate the effect of infectious diseases or
70 therapeutic agents on megakaryoblast differentiation [11,12].

71 Using CD34⁺-derived cells from controls and patients harboring the *ETV6* variants, we
72 examined the transcriptome of every cell type in the differentiation pathway from
73 hematopoietic stem/progenitor cells (HSPCs) to MKs. We provide detailed insight into
74 peripheral CD34⁺ megakaryopoiesis. Furthermore, to fully characterize the megakaryocytic
75 differentiation stages, regulon signatures and the corresponding regulatory networks were
76 assessed. We found that ETV6 deficiency leads to the development of aberrant
77 megakaryocyte-erythroid progenitor (MEP) and MK populations with some hyperactive

78 regulons and a highly deregulated core “ribosome biogenesis” pathway involving ribosomal
79 protein S6 (RPS6). Our findings provide insight into the pathology of *ETV6* variants and
80 potential targets for both diagnostic purposes and therapeutic approaches to restore pro-
81 platelet formation.

82

83 **Methods**

84 **Patients and DNA sequencing**

85 Patients were recruited at the Center for the Investigation of Hemorrhagic and Thrombotic
86 Pathologies (University Hospital of Marseille) after informing written consent was obtained,
87 in accordance with protocols approved by the national review process (authorization number
88 20200T2-02). Two high-throughput sequencing methods were used to identify *ETV6* variants
89 as previously described [5,13]. Two *ETV6* variants were evaluated in this study, the new
90 variant pF417LTer4 (NM_001987, c.1251del) and a previously described variant p.P214L
91 (NM_001987, c.C641T). Blood samples were collected from five different patients, on
92 several occasions (at least twice for each, with an interval of six months minimum).

93 ***In vitro* Megakaryocyte Differentiation**

94 After density gradient separation (Eurobio), circulating CD34⁺ cells were purified using
95 positive selection with magnetic beads (Miltenyi-Biotec) and then cultured in StemSpan
96 Serum-Free Expansion Medium II supplemented with Megakaryocyte Expansion Supplement
97 (Stemcell Technologies) as previously described [13]. In some experiments, at day 11 or day
98 13, cells were pre-incubated for 48 hours with CX-5461 (Sigma Aldrich) or cycloheximide
99 (Merck) as previously described [14,15]

100 **Cell Preparation and Single-Cell RNA Sequencing**

101 CD34⁺-differentiated cells from two healthy controls (Control 1: male, born in 1996 and
102 control 2: female, born in 1980) and two patients harboring *ETV6* variants (a male; born in
103 1985 and female, born in 1986) were harvested from culture at days 6 and 11 (Figure 1A).
104 Cell concentrations and viability were assessed after trypan blue staining with an automated
105 cell counter (Eve™ NanoEntek). The cell samples from each volunteer were labeled with a

106 distinct hashtag oligo (TotalSeq, Biolegend) and pooled. Single-cell isolation was then carried
107 out with the 10x Genomics Technology using the Chromium Next GEM Single Cell 5' Kit v2
108 (ref 1000263) according to the manufacturer's protocol. Single-cell cDNA synthesis and
109 sequencing libraries (days 6 and 11) were prepared with a single cell 5' Library and Gel Bead
110 kit. Libraries were sequenced applying a 75-bp paired end reads format with a Next-seq500
111 (GBiM platform) (parameters, read 1: 26 cycles, i7: 8 cycles, read 2: 57 cycles).

112 **Statistical Analyses**

113 Quantitative variables are expressed as the mean \pm SEM. Experimental data analyses were
114 performed using GraphPad Prism software. Statistical differences were determined using one-
115 way ANOVA with Kruskal-Wallis post-hoc test, Mann Whitney test or one-sample t-test
116 when appropriate. P-value <0.05 was considered statistically significant.

117 **Results**

118 **Characterization of differentiation in control CD34⁺-cells**

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

123 Using unsupervised clustering, we identified a total of 15 clusters (Figure 1C), which were
124 present in both controls (Supplementary figure 1A-C). We performed a detailed
125 characterization of each cell cluster based on known gene sets (signatures) [14–18] (Figure
126 1D and supplementary Figure 1F) and the top differentially expressed genes (DEGs) (Figure
127 1E, Supplementary Figure 2, Supplementary table 1).

128 Cluster 10 was designated HSPC based on the expression profiles of *PROM1*, *CRHBP*, *FLT3*,
129 *HOPX* and *AVP*. Similarly, clusters 11 and 7 were designated CMP based on the expression
130 profiles of *CPA3*, *PRG2*, *CLC* and *TPSAB1*. Cluster 11 clearly differed from cluster 7 as it
131 shared marker genes with the MKP/MK clusters (*VWA5A*, *LMNA*, *CAVIN2*, *LAT*, *CD9* and
132 *ITGA2B*) and expressed the genes *KIT*, *KRT1*, *HPGDS* and *TPSB2*. We thus defined this
133 cluster 11 as MK-primed CMP. Cluster 13 was designated GMP because it expressed specific
134 markers such as *MPO*, *ELANE* and *PRTN3*. Clusters 9, 5, 4, 0 and 2 were designated MEP.
135 Although these clusters shared a strong common MEP signature (*CD38*, *TFRC*, *DEPTOR*,
136 *KLF1*, *TFR2*), we deepened the characterization of these five subpopulations. Cluster 9 was
137 the most immature population and was designated “early MEP” based on the remaining
138 expression levels of HSPC and CMP genes. Subsetting of cluster 9 confirms the presence of
139 transitional steps with different levels of engagement towards MEP (HSPC-MEP and CMP-
140 MEP; data not shown). Cluster 5 was designated a MEP-ERP subpopulation based on the

141 genes associated with erythroid progenitors (ERPs) such as *HBB*, *TFR2* and *ANK1*. Clusters
142 4, 0 and 2 were designated a MEP-MK subpopulation based on the expression levels of MK
143 genes, such as *GP9*, *GP6*, *GP1BA* and *MPIG6B*, and the gradual increase in *MYH9*
144 expression levels between clusters 4, 0 and 2 (Supplementary figure 1D). Clusters 1, 6, 3, 8
145 and 12 were primarily observed at day 11 and were designated MK based on the expression of
146 genes associated with classical MK. Specific MK markers (*GP6*, *GP9*, *PF4*, *P2RY1*)
147 gradually increased from clusters 1 to 12 (Figure 1D, Supplementary figures 1D). Of these
148 five clusters, cluster 1 was the most immature stage (MKP) and cluster 12 was the most
149 mature. Cluster 14 was mainly observed at day 11 and was designated platelet-like particles
150 (PLPs) based on the low number of genes expressed (mean \pm SD: 1,418 \pm 347 vs. 4,208
151 \pm 1,399), the low overall RNA count (mean \pm SD: 2,243 \pm 751 vs. 23,855 \pm 14,826) as compared
152 with the MKP/MK populations (Figures 1F-G, Supplementary figure 1D). Also, their
153 proportion increases from day 6 to day 11 (0.04 % to 1.55 % of total cells). Furthermore,
154 cluster 14 expressed the same genes as cluster 12, with higher expression levels of *STXBP5*
155 and *RAB27b* involved in platelet granule metabolism and *RGS18*, *GP6* and *ITGA2B* (Figures
156 1D-E, Supplementary figures 1D-F). For each cell type, we proposed new enriched gene
157 signatures that included known cell-type specific genes and top-ranked DEGs (Supplementary
158 table 2). Lineage signature scores were then computed in each cell and cell types were defined
159 accordingly (Figure 1H-I).

160 **Inference of Megakaryopoiesis Regulon Activity in Control Cells**

161 The cell state transitions in megakaryopoiesis are tightly controlled by transcription factors
162 (TFs). Regulons are inferred groups of genes controlled as a unit by the same repressor or
163 activator TF [19]. Cell type-specific regulons provide a great opportunity to identify key
164 regulators of cell fate decisions and to improve signatures. Analysis of our control data set

165 using the single-cell regulatory network inference and clustering (SCENIC) workflow
166 provided new insight into cell type-specific regulons that drive cellular heterogeneity and
167 megakaryocytic differentiation. For each regulon, the activity score was calculated in each
168 cell based on the number of transcripts for all genes included in the regulon (TFs and targets).
169 Comparing each cell type, we analyzed the five most active regulons to isolate key regulons at
170 each stage of differentiation and identify a regulon signature (Figure 2 A-B, Supplementary
171 table 2). Using Cytoscape (v.3.9.1), we then constructed the TF gene regulatory network (TF-
172 GRN) based on the top five regulons of each cell type and their direct TF targets (Figure 2C).
173 Next, we established a list of specific and common active TFs for each cell type
174 (Supplementary table 3). Among the hyperactive TFs that characterize the MK-primed CMP-
175 biased population, we unexpectedly observed a TF that has never been associated with MK
176 differentiation (ZMAT4; zinc finger matrin, type 4) and two highly specific TFs, AhR (aryl
177 hydrocarbon receptor) and BATF (basic leucine zipper ATF-like transcription factor).

178 **Characterization of the ETV6 Variants**

179 We performed a functional study of the novel *ETV6* variant pF417Lter4 compared with the
180 previously described p.P214L [5] (Figure 3A). The nonsense mutation p.F417Lter4 is located
181 in the ETS domain, while the missense mutation p.P214L is localized in the linker domain
182 (Figure 3B, lower panel). The clinical and laboratory characteristics of the patients are shown
183 in Supplementary table 4. We analyzed the repressive activity of the two variants using a
184 dual-luciferase reporter assay. Co-transfection of the reporter plasmid containing the ETS-
185 binding site along with the expression of a plasmid encoding wild type (WT) *ETV6* resulted in
186 almost 90% inhibition of luciferase activity. Substitution of WT *ETV6* with any of the *ETV6*
187 variants led to a significant reduction in repressive activity (85% to 100%) (Figure 3B, left
188 panel). Western-blot analysis showed that mutant *ETV6* protein was expressed in the

189 GripTite293 macrophage scavenger receptor (MSR) cell line (figure 3B). Subcellular
190 fractionation of GripTite293 MSR cells showed increased ETV6 protein levels in the
191 cytoplasmic fraction and decreased levels or absence of ETV6 in the nuclear fraction in cells
192 expressing the p.P1214L or p.F417Lter4 variants compared with cells expressing the WT
193 protein (Supplementary figure 3A). Microscopy confirmed that WT ETV6 concentrated
194 primarily in cell nuclei, whereas both ETV6 variants were predominantly located in the
195 cytoplasm (Supplementary figure 3B).

196 **Single-cell transcriptional profiling of *ETV6*-variant CD34⁺ cells during** 197 **megakaryopoiesis**

198 We performed UMAP dimensional reduction to analyze cell transcriptome heterogeneity and
199 clusters. The day 6 and day 11 transcriptome profiles of patient cells overlapped considerably
200 more than those of control cells, thereby suggesting that differentiation was delayed with
201 accumulation of early-stage cells in *ETV6*-variant carriers (Figure 3C, left panel). Using
202 unsupervised clustering, a total of 11 clusters were identified (Figure 3C, right panel). As
203 previously, we performed a detailed characterization of each cell cluster based on known gene
204 sets and the top DEGs (Figure 3D-E, Supplementary Figure 4, Supplementary table 5).

205 The cell types found in controls were also observed in patients. Clusters 3 and 7 corresponded
206 to HSPC; cluster 4 corresponded to CMP; cluster 10 corresponded to MK-primed CMP;
207 cluster 9 corresponded to GMP; clusters 5, 2, 6, 8 and 0 corresponded to MEP; and cluster 1
208 corresponded to MKP/MK. No PLP cluster was observed (figure 3D-E, Supplementary figure
209 4). The proposed lineage genetic signatures (Supplementary table 2, left panel) were
210 computed in each cell and cell types were assigned (Figures 3F-G).

211 **Developmental Trajectory of Megakaryocyte Differentiation**

212 We then assessed the single-cell transcriptome for pseudotemporal ordering of differentiation
213 stages during megakaryopoiesis in controls and patients. Using Slingshot, we inferred
214 differentiation trajectories in cells from controls and *ETV6*-variant carriers independently. For
215 each condition, we determined the overall trajectory structure of each lineage (rooted tree) by
216 generating a transcriptomic distance matrix using the manually set root cluster (HSPCs;
217 cluster 10 for controls and cluster 3 for patients). We observed an unexpected lineage of MK
218 differentiation that bypassed the CMP cell type and differentiated directly from HSPCs into
219 MEPs (Figures 4A).

220 Analyzing the differentiation trajectory in control cells, stem cell markers such as *CD34*,
221 *CD38* and *HLA-DRA* displayed decreased expression, while the MK markers *GP9*, *PF4* and
222 *ITGA2B* were upregulated (Figure 4B). The *GATA1* and *FLII* TFs displayed increased
223 expression, while *GATA2* displayed high expression levels in immature cells and reduced
224 expression in later stages. During differentiation, *ETV6* and *RUNX1* expression levels were
225 low and stable (Figure 4B).

226 By contrast, *ETV6*-deficient cells displayed stable expression of *CD34*, *HLA-DRA* and *GATA2*
227 (Figure 4B) during MK differentiation; these genes remained highly expressed at the
228 MKP/MK stage (MKP/MK: fold change patients vs. controls = 2.5 (*CD34*), 1.5 (*HLA-DRA*),
229 1.6 (*GATA2*), adjusted p-value <0.05). The expression levels of *GATA1* and *FLII* were lower
230 than that observed in controls. Analyzing *GP9* and *PF4*, we observed lower expression levels
231 and a delayed response time (Figure 4B). These results indicate that *ETV6*-variant carriers
232 exhibit a delay in differentiation.

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

234 The full data set (including both control and patients) was analyzed to compare controls and
235 patients. The cell types identified in control and patient data sets were transferred to the full
236 data set using the cell identity barcode (Figure 5A). In the early cell stages (HSPCs, CMPs,
237 MK-primed CMPs and GMPs), the transcriptome profiles of controls and patients were
238 markedly similar. By contrast, distinct gene expression profiles were observed as early as the
239 MEP stage, which intensified at the MKP/MK stage.

240 Cell type distribution differed between controls and patients (Figures 5B). Patients harboring
241 *ETV6* variants displayed an increased proportion of HSPC (D6: 17.6 ± 6.1 vs $5.2 \pm 1.1\%$; D11:
242 16 ± 8 vs $0.6 \pm 0.04\%$) and a decline in MKP/MK (D11: 37 ± 1 vs $78 \pm 0.6\%$) compared with
243 controls (Figure 5B).

244 We applied the SCENIC workflow to further characterize the aberrant MEP and MKP/MK
245 populations. The UMAP reduction using the 312 combined regulon activity scores in each
246 cell is shown in Figure 5C. Compared to the early stages, we observed marked differences
247 between patients and controls at the MEP and MKP/MK stages. The top 20 differentially
248 active regulons are shown in Figure 5D. Some of the top 20 regulons were common for the
249 two cell stages and demonstrated hyperactivity in patient cells (Supplementary Table 3). After
250 analyzing the regulons using tree hierarchical algorithms with Cytoscape software, we found
251 that SPI1 and GATA2 hyperactivity represented initial events in the modification of regulon
252 activity observed in patients (Supplementary figure 5). Overall, these results indicate that
253 *ETV6* mutations alter megakaryopoiesis, resulting in the development of aberrant MEP and
254 MK populations as well as the disruption of key TF activity, with GATA2 and SPI1
255 hyperactivity representing initial events.

256 **Deregulated “DNA repair” and “translation” pathways in *ETV6*-variant cells**

257 When comparing controls and patients, the number of DEGs (fold change > 1.3, adjusted p-
258 value < 0.05) increased over the course of cell differentiation, especially at the MEP and MK
259 stages (HSPC, DEGs=38, 23 up-regulated genes and 15 down-regulated genes; CMP,
260 DEGs=56, 29 up-regulated genes and 27 down-regulated genes; MK-primed CMP, DEGs=30,
261 15 up-regulated genes and 15 down-regulated genes; GMP, DEGs=32, 21 up-regulated genes
262 and 11 down-regulated genes; MEP, DEGs=239, 100 up-regulated genes and 139 down-
263 regulated genes; and MKP/MK, DEGs=942, 339 up-regulated genes and 603 down-regulated
264 genes) (Figure 6A). Focusing on MEP and MKP/MK populations, DEGs were enriched in
265 various biological pathways in multiple gene set databases (GO Biological process (Figures
266 6B), KEGG (Supplementary figure 6) and Reactome (Supplementary figure 7). Top 10 GO
267 deregulated pathways are available in the supplementary tables 6 and 7.

268 Several pathways associated with DNA repair and cellular response to DNA damage were
269 downregulated in patient MEPs and MKs (Supplementary table 8). Some genes involved in
270 major DNA repair pathways displayed reduced expression levels: *FEN1* (base excision
271 repair), *MGMT* (direct reversal of DNA damage), *RAD23A* and *RAD23B* (nucleotide excision
272 repair), and *XRCC6* and *PRKDC* (non-homologous end joining) [20]. Caretaker genes
273 indirectly involved in maintaining genomic stability (e.g., *TTK*, *NUDT1*, *DUT* and *UBE2V2*)
274 were also downregulated in MKs harboring *ETV6* variants (Figure 6C).

275 The top 10 upregulated pathways in *ETV6*-variant MKs were associated with “translation
276 genes” (Supplementary table 7). The most upregulated genes, *RPS* and *RPL*, code for
277 ribosomal protein small subunits and ribosomal protein large subunits, respectively
278 (Supplementary figure 8).

279 **Increased Translation and RPS6 in Patients Harboring an *ETV6* Variant**

280 Translation was the strongly dominant upregulated pathway. Translation levels were
281 evaluated in HEL cells transduced with P214L or WT *ETV6* before and after PMA-induced
282 differentiation. As expected, PMA stimulation led to reduced translation in all samples
283 (Figure 7A). Translation levels were significantly higher in *ETV6*-deficient HEL cells at the
284 non-stimulated basal state and after PMA stimulation (Figure 7A). Compared with control
285 cells, higher translation levels were also observed in CD34⁺-derived MK from *ETV6*-variant
286 carriers (Figure 7B).

287 Among all upregulated genes in *ETV6*-variant MKPs/MKs, 1,488 genes contained 2,189
288 sequences that correspond to the canonical *ETV6*-binding site (i-cisTarget: normalized
289 enrichment score 4.85, rank 40e/512, Supplementary table 9), including RPS6, which is a
290 cornerstone of translation [21] (Figure 7C). Furthermore, the *ETV6*-binding site in RPS6 is
291 highly conserved across seven species (2nd/2,189 detected *ETV6*-binding regions). Analyzing
292 the differentiation trajectory in control cells, we found that RPS6 mRNA levels gradually
293 decreased, with the lowest levels observed in the most mature MKs (cluster 12). By contrast,
294 *ETV6*-deficient cells displayed stable RPS6 expression levels throughout the differentiation
295 pathway (Figure 7D).

296 Our data suggest that *ETV6* regulates *RPS6* expression. Therefore, we obtained *ETV6*
297 chromatin immunoprecipitation (ChIP)-sequence data on the GM12876 cell line from the
298 Gene Expression Omnibus database (GSM2574795, GSM2534228, GSM2574796). The peak
299 enrichment on chromosome 9, which includes the *RPS6* locus, was extracted and displayed in
300 the UCSC genome browser (Figure 7E). This analysis showed one enriched binding region
301 for *ETV6*. The binding peak is located in the promoter.

302 The expression of total RPS6 and phosphorylated RPS6 (p-RPS6, Ser235/236) were analyzed
303 by western blot and flow cytometry in washed platelets, CD34⁺-derived MKs from *ETV6*-
304 variant carriers, and the respective age- and gender-matched controls. RPS6 and p-RPS6 were
305 overexpressed in native platelets issued from *ETV6*-variant carriers compared with controls
306 (Figures 7F, left panel and Supplementary figure 9). We confirmed RPS6 and p-RPS6
307 overexpression in CD34⁺-derived MKs issued from three *ETV6*-variant carriers using flow
308 cytometry (Figures 7F, right panel).

309 To investigate the consequences of increased translation, we treated control and patient
310 CD34⁺-derived cell cultures with CX-5461, a specific RNA Pol I inhibitor. Healthy control
311 CD34⁺-derived MKs (day 11) treated with CX-5461 exhibited reduced translation and p-
312 RPS6 levels, whereas the expression of total RPS6 was unaffected (Figure 7G). In CD34⁺-
313 derived cells issued from three *ETV6*-variant carriers, CX-5461 treatment increased the
314 production of PLPs by CD42a⁺ MKs in all cases (Figure 7G). In addition, cells exposed to
315 cycloheximide (CHX) to inhibit protein synthesis, exhibited reduced translation, RPS6 and
316 phospho-RPS6 levels. CHX treatment was able to increase the production of PLPs
317 (Supplementary Figure 10).

318

319 **Discussion**

320 Using single-cell transcriptome profiling of MK cell cultures derived from peripheral CD34⁺
321 cells, we established a specific signature for each cell stage based on both distinct gene
322 expression profiles and regulon activities. Cells harboring *ETV6* variants displayed distinct
323 transcriptomic/regulon profiles starting at the MEP stage and intensifying at the MKP/MK
324 stage, with SPI1 and GATA2 regulon hyperactivity representing early events. Translation
325 pathways were upregulated in patient MEPs and MKs. Puromycin incorporation results as
326 well as RPS6 and p-RPS6 immunoblots were consistent with these findings. Overall, *ETV6*
327 deficiency affects key cellular processes that may represent promising mechanistic targets.

328 In CD34⁺-derived MK differentiation, we observed a biased megakaryopoiesis pathway that
329 may be due to stress-driven hematopoiesis as a result of the culture conditions [22]. HSPCs
330 directly gave rise to MEPs without passing through the CMP stage, as previously suggested
331 [23]. Furthermore, it has been hypothesized that MK-biased hematopoietic stem cells (HSCs)
332 represent the hierarchical apex with MEPs developing directly from HSCs [24], although this
333 theory has been challenged. We did not observe a MK-biased HSPC population [25] because
334 either the analysis time points were too late to detect these cell types or mRNA did not enable
335 detection of MK-biased HSPCs, as previously reported [26]. However, we identified a MK-
336 primed CMP population that co-expressed megakaryocytic genes and displayed high KIT
337 expression levels (27% of the CMP population). High KIT expression levels represented a
338 good marker of MK-biased CMPs. Previous murine *in vitro* and *in vivo* functional studies
339 have demonstrated that HSCs with higher levels of c-Kit signaling preferentially differentiate
340 into MKs [27]. In human bone marrow, MK-primed CMPs likely represented the major
341 megakaryopoiesis pathway independent of the canonical MEP lineage [28]. This *in vitro*
342 model is a valuable tool to investigate these biased pathways.

343 Hematopoiesis is a progressive process controlled by an elaborate network of TFs. We
344 characterized each cell type by analyzing the activity of specific regulons, to better define
345 hematopoietic populations and establish a TF-GRN of megakaryopoiesis. HSPCs were
346 characterized by the expression of the regulons HLF [29], HOXA9 [18] and TCF4 [30],
347 which are involved in stem cell maintenance and myeloid commitment. C/EBP family
348 members are critical for myelopoiesis [31,32]. Accordingly, we observed noticeable C/EBP
349 activity at the CMP and GMP stages. Interestingly, ZMAT4 was newly identified as an active
350 TF in MK-primed CMP populations. ZMAT4 is a TF of the zinc finger protein family and is
351 associated with “programed cell death” [33], innate lymphoid cells and natural killer cells
352 [34]. Although copy number variants of ZMAT4 are associated with leukemia [35], the role
353 that ZMAT4 plays in myeloid differentiation remains unclear. The MK-primed CMP
354 population was also characterized by AhR expression, which is involved in MK maturation
355 [36–38] and BATF hyperactivation. BATF hyperactivation was evidenced as a marker of MK
356 commitment. We also observed KLF1 activity in MEPs, which is involved in MEP lineage
357 decisions and commitment [39]. Finally, NFE2 and GATA1 were expected markers of MKs
358 [40,41]. Overall, these findings demonstrate that characterizing TF activity at the single-cell
359 level is useful to phenotype cell types and decipher unexpected key functions of new TFs in
360 each cell type.

361 Applying the lineage signature used in control cells, we found that ETV6-deficient cells
362 displayed the same cell types with a higher proportion of HSPCs, a lower proportion of MKs
363 and a lack of PLPs, which correlates with the thrombocytopenia phenotype observed in
364 patients harboring an *ETV6* variant. We also detected marked differences in regulon activities.
365 Several regulons displayed higher activity levels in the MEP and MK stages. This finding
366 may be due to a loss of ETV6 repressor activity as documented in all *ETV6*-mutant carriers

367 [5]. We observed GATA2 hyperactivity in MKs from *ETV6*-mutant carriers, directly
368 regulating SPI1 activity. *SPI1* has previously been described as a common downstream target
369 gene of GATA2 in human and murine embryonic stem cells^{52,53}. High GATA2 expression
370 levels have been described in REH cells, a B-cell acute lymphoblastic leukemia cell line
371 harboring the *ETV6/RUNX* fusion transcript, which is the most frequent structural
372 cytogenetic abnormality in children with acute lymphoblastic leukemia [42]. Beyond this
373 observation, the connection between *ETV6* deficiency and GATA2 hyperactivity remains
374 poorly understood. This relationship merits further investigation as deficient or excess
375 GATA-2 expression are both pathogenetic [43].

376 Analysis of several gene enrichment databases displayed significant differences. Assessing
377 the ten highest scores and after grouping the pathways with common genes, we observed
378 deregulation of pathways linked with translation and DNA repair in MEP and MKP/MK
379 patient cells. Translation was the most upregulated pathway. Altered expression levels of
380 numerous ribosomal protein genes were observed, including upregulation of more than 30
381 ribosomal proteins. Recent data has emphasized the unexpected contribution of ribosomal
382 biogenesis in hematopoiesis and megakaryopoiesis. Additionally, ribosome biogenesis is
383 involved in mediating the transition between proliferation and differentiation of erythroid
384 progenitors [44]. GTPase-Dynamin-2 deletion in platelets and MKs induces severe
385 thrombocytopenia and bleeding diathesis in mice and results in the upregulation of genes
386 involved in ribosome biogenesis in erythroblasts [45].

387 Furthermore, increased translation was confirmed at the functional level. Protein synthesis
388 was increased in CD34⁺-derived MKs in patients and hematopoietic cell lines transduced with
389 an *ETV6* variant. Analysis of the *RPS6* gene revealed a putative *ETV6*-binding site, which
390 was confirmed via ChIP-sequencing [46]. As *Rps6*-deficient mice display features of

391 megakaryocytic dysplasia with thrombocytosis, this gene likely plays a role in thrombopoiesis
392 [47,48]. Moreover, overexpression of RPS6 or p-RPS6 in murine Sertoli cells cultured *in vitro*
393 has been shown to disrupt cell function through changes in the organization of actin and
394 microtubule-based cytoskeletons via AKT inhibition [49,50]. In *ETV6*-variant carriers,
395 overexpression of RPS6 was observed in platelets and may act on pro-platelet formation,
396 which is dependent on massive reorganization of the actin cytoskeleton. Translation inhibition
397 led to reduced RPS6 phosphorylation and increased numbers of PLPs in *ETV6*-deficient
398 cultures, thereby suggesting that overexpression of p-RPS6 in patient cells causes a decrease
399 in actin polymerization rates and pro-platelet formation.

400 Overall, our study demonstrates the heterogeneity of human CD34⁺ cell-induced MK
401 differentiation *in vitro* and a major differentiation trajectory in which HSPCs develop into
402 MEPs, bypassing the CMP stage. We found that *ETV6* variations result in aberrant MEP and
403 MK populations, which are associated with GATA2 hyperactivity and deregulated translation
404 pathways, with RPS6 and p-RPS6 overexpression in CD34⁺-derived MKs and platelets. These
405 findings provide novel insight into megakaryopoiesis and *ETV6* function that may be applied
406 to develop targeted therapeutic strategies to alleviate platelet defects.

407

408 **Competing interests**

409 The authors declare that they have no competing interests.

410 **Authorship contributions**

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

415 **Acknowledgments**

416 The authors acknowledge the members of the French Reference Center for Inherited
417 Hereditary Platelet Disorders for their contribution regarding clinical analyses, the high-
418 performance computing center of Aix-Marseille for granting access to its high-performance
419 computing resources, the GBiM platform for the sequencing and Dr Argüello for providing
420 reagents to study translation. This work was supported by Aix-Marseille University
421 (AMIDEX “Emergence et innovation” ngSUMMIT), the Agence Nationale de la Recherche
422 (JCJC MOST) and the Institut National de la Santé et de la Recherche Médicale (PIA Biofit).

423

424 References

- 425 1 Di Paola J, Porter CC. ETV6-related thrombocytopenia and leukemia predisposition. *Blood* 2019;
426 **134**: 663–7.
- 427 2 Wang LC, Kuo F, Fujiwara Y, Gilliland DG, Golub TR, Orkin SH. Yolk sac angiogenic defect and
428 intra-embryonic apoptosis in mice lacking the Ets-related factor TEL. *EMBO J* 1997; **16**: 4374–83.
- 429 3 Zhou C, Uluisik R, Rowley JW, David C, Jones CL, Scharer CD, Noetzli L, Fisher MH, Kirkpatrick GD,
430 Bark K, Boss JM, Henry CJ, Pietras EM, Di Paola J, Porter CC. Germline ETV6 mutation promotes
431 inflammation and disrupts lymphoid development of early hematopoietic progenitors.
432 *Experimental Hematology* 2022; **112–113**: 24–34.
- 433 4 Borst S, Nations CC, Klein JG, Pavani G, Maguire JA, Camire RM, Drazer MW, Godley LA, French
434 DL, Poncz M, Gadue P. Study of inherited thrombocytopenia resulting from mutations in ETV6 or
435 RUNX1 using a human pluripotent stem cell model. *Stem Cell Reports* 2021; **16**: 1458–67.
- 436 5 Poggi M, Canault M, Favier M, Turro E, Saultier P, Ghalloussi D, Baccini V, Vidal L, Mezzapesa A,
437 Chelghoum N, Mohand-Oumoussa B, Falaise C, Favier R, Ouwehand WH, Fiore M, Peiretti F,
438 Morange PE, Saut N, Bernot D, Greinacher A, et al. Germline variants in ETV6 underlie reduced
439 platelet formation, platelet dysfunction and increased levels of circulating CD34+ progenitors.
440 *Haematologica* 2017; **102**: 282–94.
- 441 6 Grodzielski M, Goette NP, Glembotsky AC, Constanza Baroni Pietto M, Méndez-Huergo SP,
442 Pierdominici MS, Montero VS, Rabinovich GA, Molinas FC, Heller PG, Lev PR, Marta RF. Multiple
443 concomitant mechanisms contribute to low platelet count in patients with immune
444 thrombocytopenia. *Sci Rep* 2019; **9**: 2208.
- 445 7 Kollmann K, Warsch W, Gonzalez-Arias C, Nice FL, Avezov E, Milburn J, Li J, Dimitropoulou D,
446 Biddie S, Wang M, Poynton E, Colzani M, Tijssen MR, Anand S, McDermott U, Huntly B, Green T.
447 A novel signalling screen demonstrates that CALR mutations activate essential MAPK signalling
448 and facilitate megakaryocyte differentiation. *Leukemia* 2017; **31**: 934–44.
- 449 8 Zeddies S, Jansen SBG, Summa F di, Geerts D, Zwaginga JJ, Schoot CE van der, Lindern M von,
450 Thijssen-Timmer DC. MEIS1 regulates early erythroid and megakaryocytic cell fate.
451 *Haematologica* 2014; **99**: 1555–64.
- 452 9 Basak I, Bhatlekar S, Manne BK, Stoller M, Hugo S, Kong X, Ma L, Rondina MT, Weyrich AS,
453 Edelstein LC, Bray PF. miR-15a-5p regulates expression of multiple proteins in the megakaryocyte
454 GPVI signaling pathway. *J Thromb Haemost* 2019; **17**: 511–24.
- 455 10 Song B, Miao W, Cui Q, Shi B, Zhang J, Qiu H, Zhang L, Han Y. Inhibition of ferroptosis promotes
456 megakaryocyte differentiation and platelet production. *Journal of Cellular and Molecular*
457 *Medicine* 2022; **26**: 3582–5.
- 458 11 Lin G-L, Chang H-H, Lien T-S, Chen P-K, Chan H, Su M-T, Liao C-Y, Sun D-S. Suppressive effect of
459 dengue virus envelope protein domain III on megakaryopoiesis. *Virulence* 2017; **8**: 1719–31.
- 460 12 Hu L, Yin X, Zhang Y, Pang A, Xie X, Yang S, Zhu C, Li Y, Zhang B, Huang Y, Tian Y, Wang M, Cao W,
461 Chen S, Zheng Y, Ma S, Dong F, Hao S, Feng S, Ru Y, et al. Radiation-induced bystander effects
462 impair transplanted human hematopoietic stem cells via oxidative DNA damage. *Blood* 2021;
463 **137**: 3339–50.

- 464 13 Saultier P, Vidal L, Canault M, Bernot D, Falaise C, Pouymayou C, Bordet J-C, Saut N, Rostan A,
465 Baccini V, Peiretti F, Favier M, Lucca P, Deleuze J-F, Olaso R, Boland A, Morange PE, Gachet C,
466 Malergue F, Fauré S, et al. Macrothrombocytopenia and dense granule deficiency associated
467 with FLI1 variants: ultrastructural and pathogenic features. *Haematologica* 2017; **102**: 1006–16.
- 468 14 Schmidt EK, Clavarino G, Ceppi M, Pierre P. SUnSET, a nonradioactive method to monitor protein
469 synthesis. *Nat Methods* 2009; **6**: 275–7.
- 470 15 Adamik J, Munson PV, Hartmann FJ, Combes AJ, Pierre P, Krummel MF, Bendall SC, Argüello RJ,
471 Butterfield LH. Distinct metabolic states guide maturation of inflammatory and tolerogenic
472 dendritic cells. *Nat Commun* 2022; **13**: 5184.
- 473 16 Psaila B, Wang G, Rodriguez-Meira A, Li R, Heuston EF, Murphy L, Yee D, Hitchcock IS, Sousos N,
474 O’Sullivan J, Anderson S, Senis YA, Weinberg OK, Calicchio ML, NIH Intramural Sequencing
475 Center, Iskander D, Royston D, Milojkovic D, Roberts I, Bodine DM, et al. Single-Cell Analyses
476 Reveal Megakaryocyte-Biased Hematopoiesis in Myelofibrosis and Identify Mutant Clone-Specific
477 Targets. *Mol Cell* 2020; **78**: 477-492.e8.
- 478 17 Chen L, Kostadima M, Martens JHA, Canu G, Garcia SP, Turro E, Downes K, Macaulay IC, Bielczyk-
479 Maczynska E, Coe S, Farrow S, Poudel P, Burden F, Jansen SBG, Astle WJ, Attwood A, Bariana T,
480 de Bono B, Breschi A, Chambers JC, et al. Transcriptional diversity during lineage commitment of
481 human blood progenitors. *Science* 2014; **345**: 1251033.
- 482 18 Estevez B, Borst S, Jarochoa DJ, Sudunagunta VS, Gonzalez M, Garifallou J, Hakonarson H, Gao P,
483 Tan K, Liu PP, Bagga S, Holdreith N, Tong W, Speck NA, French DL, Gadue P, Poncz M. RUNX1
484 haploinsufficiency causes a marked deficiency of megakaryocyte-biased hematopoietic
485 progenitor cells. *Blood* 2021; .
- 486 19 Pellin D, Loperfido M, Baricordi C, Wolock SL, Montepeloso A, Weinberg OK, Biffi A, Klein AM,
487 Biasco L. A comprehensive single cell transcriptional landscape of human hematopoietic
488 progenitors. *Nat Commun* 2019; **10**.
- 489 20 Roy A, Wang G, Iskander D, O’Byrne S, Elliott N, O’Sullivan J, Buck G, Heuston EF, Wen WX, Meira
490 AR, Hua P, Karadimitris A, Mead AJ, Bodine DM, Roberts I, Psaila B, Thongjuea S. Transitions in
491 lineage specification and gene regulatory networks in hematopoietic stem/progenitor cells over
492 human development. *Cell Reports* 2021; **36**: 109698.
- 493 21 Aibar S, González-Blas CB, Moerman T, Huynh-Thu VA, Imrichova H, Hulselmans G, Rambow F,
494 Marine J-C, Geurts P, Aerts J, van den Oord J, Atak ZK, Wouters J, Aerts S. SCENIC: single-cell
495 regulatory network inference and clustering. *Nat Methods* 2017; **14**: 1083–6.
- 496 22 Chatterjee N, Walker GC. Mechanisms of DNA damage, repair and mutagenesis. *Environ Mol*
497 *Mutagen* 2017; **58**: 235–63.
- 498 23 Chauvin C, Koka V, Nouschi A, Mieulet V, Hoareau-Aveilla C, Dreazen A, Cagnard N, Carpentier W,
499 Kiss T, Meyuhas O, Pende M. Ribosomal protein S6 kinase activity controls the ribosome
500 biogenesis transcriptional program. *Oncogene* Nature Publishing Group; 2014; **33**: 474–83.
- 501 24 Psaila B, Mead AJ. Single-cell approaches reveal novel cellular pathways for megakaryocyte and
502 erythroid differentiation. *Blood* 2019; **133**: 1427–35.

- 503 25 Adolfsson J, Månsson R, Buza-Vidas N, Hultquist A, Liuba K, Jensen CT, Bryder D, Yang L, Borge O-
504 J, Thoren LAM, Anderson K, Sitnicka E, Sasaki Y, Sigvardsson M, Jacobsen SEW. Identification of
505 Flt3+ Lympho-Myeloid Stem Cells Lacking Erythro-Megakaryocytic Potential: A Revised Road Map
506 for Adult Blood Lineage Commitment. *Cell* 2005; **121**: 295–306.
- 507 26 Sanjuan-Pla A, Macaulay IC, Jensen CT, Woll PS, Luis TC, Mead A, Moore S, Carella C, Matsuoka S,
508 Bouriez Jones T, Chowdhury O, Stenson L, Lutteropp M, Green JCA, Facchini R, Boukarabila H,
509 Grover A, Gambardella A, Thongjuea S, Carrelha J, et al. Platelet-biased stem cells reside at the
510 apex of the haematopoietic stem-cell hierarchy. *Nature* 2013; **502**: 232–6.
- 511 27 Noetzli LJ, French SL, Machlus KR. New Insights into the Differentiation of Megakaryocytes from
512 Hematopoietic Progenitors. *Arterioscler Thromb Vasc Biol* 2019; **39**: 1288–300.
- 513 28 Haas S, Hansson J, Klimmeck D, Loeffler D, Velten L, Uckelmann H, Wurzer S, Prendergast ÁM,
514 Schnell A, Hexel K, Santarella-Mellwig R, Blaszkiewicz S, Kuck A, Geiger H, Milsom MD, Steinmetz
515 LM, Schroeder T, Trumpp A, Krijgsveld J, Essers MAG. Inflammation-Induced Emergency
516 Megakaryopoiesis Driven by Hematopoietic Stem Cell-like Megakaryocyte Progenitors. *Cell Stem*
517 *Cell* 2015; **17**: 422–34.
- 518 29 Shin JY, Hu W, Naramura M, Park CY. High c-Kit expression identifies hematopoietic stem cells
519 with impaired self-renewal and megakaryocytic bias. *J Exp Med* 2014; **211**: 217–31.
- 520 30 Miyawaki K, Iwasaki H, Jiromaru T, Kusumoto H, Yurino A, Sugio T, Uehara Y, Odawara J, Daitoku
521 S, Kunisaki Y, Mori Y, Arinobu Y, Tsuzuki H, Kikushige Y, Iino T, Kato K, Takenaka K, Miyamoto T,
522 Maeda T, Akashi K. Identification of unipotent megakaryocyte progenitors in human
523 hematopoiesis. *Blood* 2017; **129**: 3332–43.
- 524 31 Lehnertz B, Chagraoui J, MacRae T, Tomellini E, Corneau S, Mayotte N, Boivin I, Durand A,
525 Gracias D, Sauvageau G. HLF expression defines the human hematopoietic stem cell state. *Blood*
526 2021; **138**: 2642–54.
- 527 32 In 't Hout FEM, van Duren J, Monteferrario D, Brinkhuis E, Mariani N, Westers TM, Chitu D,
528 Nikoloski G, van de Loosdrecht AA, van der Reijden BA, Jansen JH, Huls G. TCF4 promotes
529 erythroid development. *Exp Hematol* 2019; **69**: 17-21.e1.
- 530 33 Radomska HS, Huettner CS, Zhang P, Cheng T, Scadden DT, Tenen DG. CCAAT/enhancer binding
531 protein alpha is a regulatory switch sufficient for induction of granulocytic development from
532 bipotential myeloid progenitors. *Mol Cell Biol* 1998; **18**: 4301–14.
- 533 34 Shyamsunder P, Shanmugasundaram M, Mayakonda A, Dakle P, Teoh WW, Han L, Kanojia D, Lim
534 MC, Fullwood M, An O, Yang H, Shi J, Hossain MZ, Madan V, Koeffler HP. Identification of a novel
535 enhancer of CEBPE essential for granulocytic differentiation. *Blood* 2019; **133**: 2507–17.
- 536 35 Apostolidis PA, Lindsey S, Miller WM, Papoutsakis ET. Proposed megakaryocytic regulon of p53:
537 the genes engaged to control cell cycle and apoptosis during megakaryocytic differentiation.
538 *Physiological Genomics* American Physiological Society; 2012; **44**: 638–50.
- 539 36 Allan DSJ, Cerdeira AS, Ranjan A, Kirkham CL, Aguilar OA, Tanaka M, Childs RW, Dunbar CE,
540 Strominger JL, Kopcow HD, Carlyle JR. Transcriptome analysis reveals similarities between human
541 blood CD3⁺ CD56^{bright} cells and mouse CD127⁺ innate lymphoid cells. *Sci Rep* Nature Publishing
542 Group; 2017; **7**: 3501.

- 543 37 Wan J, Gao Y, Zhao X, Wu Q, Fu X, Shao Y, Yang H, Guan M, Yu B, Zhang W. The association
544 between the copy-number variations of ZMAT4 and hematological malignancy. *Hematology*
545 2011; **16**: 20–3.
- 546 38 Lindsey S, Papoutsakis ET. The aryl hydrocarbon receptor (AHR) transcription factor regulates
547 megakaryocytic polyploidization. *Br J Haematol* 2011; **152**: 469–84.
- 548 39 Strassel C, Brouard N, Mallo L, Receveur N, Mangin P, Eckly A, Bieche I, Tarte K, Gachet C, Lanza
549 F. Aryl hydrocarbon receptor-dependent enrichment of a megakaryocytic precursor with a high
550 potential to produce proplatelets. *Blood* 2016; **127**: 2231–40.
- 551 40 Smith BW, Rozelle SS, Leung A, Ubellacker J, Parks A, Nah SK, French D, Gadue P, Monti S, Chui
552 DHK, Steinberg MH, Frelinger AL, Michelson AD, Theberge R, McComb ME, Costello CE, Kotton
553 DN, Mostoslavsky G, Sherr DH, Murphy GJ. The aryl hydrocarbon receptor directs hematopoietic
554 progenitor cell expansion and differentiation. *Blood* 2013; **122**: 376–85.
- 555 41 Tallack MR, Perkins AC. Megakaryocyte-erythroid lineage promiscuity in EKLf null mouse blood.
556 *Haematologica* 2010; **95**: 144–7.
- 557 42 Tsang AP, Fujiwara Y, Hom DB, Orkin SH. Failure of megakaryopoiesis and arrested erythropoiesis
558 in mice lacking the GATA-1 transcriptional cofactor FOG. *Genes Dev* 1998; **12**: 1176–88.
- 559 43 Levin J, Peng J-P, Baker GR, Villeval J-L, Lecine P, Burstein SA, Shivdasani RA. Pathophysiology of
560 Thrombocytopenia and Anemia in Mice Lacking Transcription Factor NF-E2. *Blood* 1999; **94**:
561 3037–47.
- 562 44 Gaine ME, Sharpe DJ, Smith JS, Colyer HAA, Hodges VM, Lappin TR, Mills KI. GATA2 regulates the
563 erythropoietin receptor in t(12;21) ALL. *Oncotarget Impact Journals*; 2017; **8**: 66061–74.
- 564 45 Katsumura KR, Bresnick EH, the GATA Factor Mechanisms Group. The GATA factor revolution in
565 hematology. *Blood* 2017; **129**: 2092–102.
- 566 46 Le Goff S, Boussaid I, Floquet C, Raimbault A, Hatin I, Andrieu-Soler C, Salma M, Leduc M, Gautier
567 E-F, Guyot B, d’Allard D, Montel-Lehry N, Ducamp S, Houvert A, Guillonneau F, Giraudier S,
568 Cramer-Bordé E, Morlé F, Diaz J-J, Hermine O, et al. p53 activation during ribosome biogenesis
569 regulates normal erythroid differentiation. *Blood* 2021; **137**: 89–102.
- 570 47 Eaton N, Boyd EK, Biswas R, Lee-Sundlov MM, Dlugi TA, Ramsey HE, Zheng S, Burns RT, Sola-
571 Visner MC, Hoffmeister KM, Falet H. Endocytosis of the thrombopoietin receptor Mpl regulates
572 megakaryocyte and erythroid maturation in mice. *Frontiers in Oncology* 2022; **12**.
- 573 48 ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human
574 genome. *Nature* 2012; **489**: 57–74.
- 575 49 McGowan KA, Pang WW, Bhardwaj R, Perez MG, Pluvineau JV, Glader BE, Malek R, Mendrysa
576 SM, Weissman IL, Park CY, Barsh GS. Reduced ribosomal protein gene dosage and p53 activation
577 in low-risk myelodysplastic syndrome. *Blood* 2011; **118**: 3622–33.
- 578 50 Keel SB, Phelps S, Sabo KM, O’Leary MN, Kirn-Safran CB, Abkowitz JL. Establishing Rps6
579 hemizygous mice as a model for studying how ribosomal protein haploinsufficiency impairs
580 erythropoiesis. *Exp Hematol* 2012; **40**: 290–4.

- 581 51 Mok K-W, Chen H, Lee WM, Cheng CY. rpS6 regulates blood-testis barrier dynamics through
582 Arp3-mediated actin microfilament organization in rat sertoli cells. An in vitro study.
583 *Endocrinology* 2015; **156**: 1900–13.
- 584 52 Li L-X, Wu S-W, Yan M, Lian Q-Q, Ge R-S, Cheng CY. Regulation of blood-testis barrier dynamics
585 by the mTORC1/rpS6 signaling complex: An in vitro study. *Asian J Androl* 2019; **21**: 365–75.
- 586

587 **Figure Legends**

588 **Figure 1. Single-cell RNA sequencing analysis in megakaryopoiesis cell stages in**
589 **CD34⁺-cells isolated from healthy volunteers.**

- 590 **A-** Schematic diagram of the experimental design. Fresh CD34⁺ cells from human
591 peripheral blood were isolated via density gradient and magnetic cell sorting (n=2
592 healthy volunteers and n=2 patients). CD34⁺ cells were differentiated into
593 megakaryocytes using serum-free StemSpan SFEM II medium with megakaryocyte
594 expansion supplement (SCF, IL6, TPO and IL9) (STEMCELL Technologies). Cells
595 from different samples are incubated with DNA-barcoded antibodies recognizing
596 ubiquitous cell surface proteins. Distinct barcodes (referred to as hashtag-oligos, HTO)
597 on the antibodies allow pooling of multiple samples into one single-cell RNA (scRNA)
598 sequencing experiment. The cells were analyzed at day 6 and day 11 of culture using
599 10X genomics technology.
- 600 **B-** UMAP of the control cells at days 6 and 11 after merging the two data sets. Day 6 cells
601 are shown in orange and day 11 cells are shown in blue.
- 602 **C-** UMAP of the control cells after clustering performed using a resolution of 1.2. Each
603 cluster is indicated with a distinct color.
- 604 **D-** Dot plot showing the expression levels of known specific hematopoietic cell marker
605 genes in each cluster of controls. The color is representative of relative average
606 expression of the cluster compared to all others. Circle size is associated with the
607 percentage of cells expressing the gene in the cluster.
- 608 **E-** Dot plot of the top five differentially expressed genes (DEGs) in each cluster compared
609 to all others in controls. The color is representative of relative average expression of
610 the cluster compared to all others. Circle size is associated with the percentage of cells
611 expressing the gene in the cluster.

- 612 **F-** UMAP showing the number of expressed genes detected in hematopoietic cells derived
613 from normal CD34⁺ cells. The gradient of color is associated to the number of
614 expressed genes.
- 615 **G-** UMAP showing the number of overall mRNA counts detected in hematopoietic cells
616 derived from normal CD34⁺ cells. The gradient of color is associated to the number of
617 overall mRNA count.
- 618 **H-** UMAP of control cells after assignation to hematopoietic cell types. Each cell type is
619 distinguished from another by color. Clusters from figure 1C are superimposed over
620 cell types.
- 621 **I-** UMAP of control cells after computation and scoring of proposed signatures
622 (Supplementary table 2). Each plot represents the computation for one cell type
623 signature. Color gradient is associated to the relative score of the signature.

624 **Figure 2. Inference of genes regulatory network during megakaryopoiesis: control**
625 **CD34⁺-cell culture.**

- 626 **A-** Heatmap of top 5 differentially active regulons for each cell type in control CD34⁺-cell
627 culture. Color gradient is representative of the relative activity of regulons. Each cell
628 type is distinguished from another by color. Black rectangles stake out the top 5
629 regulons of the corresponding cell type.
- 630 **B-** UMAP of control cells after computation and scoring of proposed regulon signatures
631 (Supplementary table 2, right panel). Each plot represents the computation for one cell
632 type regulon signature. Color gradient is associated to the relative activity score of the
633 signature.
- 634 **C-** Transcription factors gene regulatory network (TF-GRN) of control CD34⁺ cell culture.
635 Each hematopoietic cell type is represented by a colored rectangle. Each arrow
636 represents known (solid) or supposed (dashed) differentiation steps. Following a cross-

637 reading of active regulons in populations, we were able to propose lists of specific (in
638 rectangles) or shared (outside rectangles) TFs.

639 **Figure 3. Single-cell RNA sequencing of CD34⁺-cells from patients with an *ETV6***
640 **variant.**

641 A- Upper panel: schematic pedigrees of *ETV6*-variant inheritance. Squares represent
642 males, circles represent females, and slashes represent deceased family members.
643 Solid black symbols represent family members with thrombocytopenia. The F1 family
644 carried the *ETV6* p.P214L variant, while the F2 family carried the p.F417LTer4.
645 Lower panel: schematic diagram of the *ETV6* protein structure. The functional N-
646 terminal pointed domain (PNT) and C-terminal ETS DNA-binding domain (ETS) are
647 depicted. The positions of the *ETV6* mutations are indicated.

648 B- GripTite 293 MSR cells were co-transfected with an empty vector, wild type (WT) or
649 *ETV6* variant construct with the luciferase reporter plasmid containing three tandem
650 copies of the ETS binding site upstream of the HSV TK promoter (E743tk80Luc) and
651 the pGL4.73 Renilla luciferase control vector. Firefly:Renilla luminescence ratios
652 (Fluc/Rluc) were calculated to assess transfection efficiency and are expressed as fold
653 change relative to the empty vector. The data represent the mean \pm SEM of four
654 independent experiments; **P-value <0.01 vs. empty vector (one-way ANOVA with
655 Kruskal-Wallis post-hoc test). Representative western blot analysis of *ETV6*
656 expression in GripTite 293 MSR cells transfected with an empty vector, WT or *ETV6*
657 variant constructs using a goat anti-*ETV6* antibody. GAPDH was used as a protein
658 loading control. Quantification of band intensity is shown on the right.

659 C- Left panel: UMAP of the patient cells at days 6 and 11 after merging the two data sets
660 (patients D6 and patients D11, respectively). Day 6 cells are shown in orange and day

661 11 cells are shown in blue. Right panel: UMAP of the patient cells after clustering
662 performed using a resolution of 0.8. Each cluster is indicated with a distinct color.

663 D- Dot plot showing the expression levels of known specific hematopoietic cell marker
664 genes in each cluster of patients. The color is representative of relative average
665 expression of the cluster compared to all others. Circle size is associated with the
666 percentage of cells expressing the gene in the cluster.

667 E- Dot plot of the top five DEGs in each cluster compared to all others in patients. The
668 color is representative of relative average expression of the cluster compared to all
669 others. Circle size is associated with the percentage of cells expressing the gene in the
670 cluster.

671 F- UMAP of the patient's cells after computation and scoring of proposed signatures
672 (Supplementary table 2). Each plot represents the computation for one cell type
673 signature. Color gradient is associated to the relative score of the signature.

674 G- UMAP of the patient's cells after assignment to hematopoietic cell types. Each cell
675 type is distinguished from another by color. Clusters from figure 3C are superimposed
676 over cell types.

677 **Figure 4. Megakaryocyte differentiation in healthy controls and *ETV6* patients:**
678 **trajectory inference.**

679 A- Slingshot trajectory inference in control and patient cells. Cell types are distinguished
680 by color. Black lines represent the differentiation trajectory from an early stage
681 (HSPC) to late stages (PLP or MKP/MK) in both control and patient data sets.

682 B- Violin plots of expression of specific genes (*CD34*, *CD38*, *HLA-DRA*, *GP9*, *PF4*,
683 *ITGA2B*, *FLI1*, *GATA1*, *GATA2*, *ETV6*, *RUNX1*) in control and patient cells. Cells are
684 regrouped by cluster and colored according to their assigned cell type. Box plots are

685 superimposed over violin plots. Upper limits of box plots correspond to quantile 75 %
686 and lower limits are associated to quantile 25 %. Solid black lines correspond to the
687 median. Black dots are outlier values.

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

690 A- UMAP of hematopoietic cells derived from control and *ETV6*-variant CD34⁺ cells on
691 day 6 and day 11. Cells are colored by cell type (HSPC, CMP, MK-primed CMP,
692 GMP, MEP, MKP/MK and PLP).

693 B- Pie charts of hematopoietic populations distribution in controls and patients at day 6
694 and day 11. Cell types are associated to distinct colors. Percentages were calculated
695 for each condition independently. Control cells on day 6: HSPC = 5.16 %, CMP =
696 11.71 %, MK-primed CMP = 3.80 %, GMP = 2.83 %, MEP = 72.57 %, MKP/MK =
697 3.91 %, PLP = 0.04 %. Control cells on day 11: HSPC = 0.58 %, CMP = 2.14 %, MK-
698 primed CMP = 2.06 %, GMP = 0.42 %, MEP = 15.21 %, MKP/MK = 78.05 %, PLP =
699 1.55 %. Patient cells on day 6: HSPC = 17.60 %, CMP = 10.29 %, MK-primed CMP =
700 1.67 %, GMP = 3.42 %, MEP = 60.03 %, MKP/MK = 7.00 %, PLP = 0.00 %. Patient
701 cells on day 11: HSPC = 15.98 %, CMP = 8.18 %, MK-primed CMP = 6.37 %, GMP
702 = 3.89 %, MEP = 28.67 %, MKP/MK = 36.92 %, PLP = 0.00 %.

703 C- UMAP of hematopoietic cells derived from control and *ETV6*-variant CD34⁺ cells on
704 day 6 and 11. Cells are colored by cell type (HSPC, CMP, MK-primed CMP, GMP,
705 MEP, MKP/MK and PLP). Cells are located on the UMAP using regulon matrix.

706 D- Heatmap of the top 20 differentially active regulons (patients against controls) in MEP
707 and MKP/MK populations. The color gradient is associated to the scaled activity of
708 the regulon.

709 **Figure 6. Differentially expressed genes and deregulated pathways in *ETV6*-variant**
710 **carriers.**

711 A- Volcano plots of up regulated (red) and down regulated (blue) genes in *ETV6* patients
712 vs. controls for each cell type (HSPC, CMP, MK-primed CMP, GMP, MEP and
713 MKP/MK), $FC > 1.3$ [i.e., $\log_2 FC > 0.37$], adjusted p-value < 0.05).

714 B- Bubble plot of the top enriched GO biological processes based on differential gene
715 expression by cell type (classified by p values). The upper panel shows the
716 upregulated pathways, while the lower panel shows the down-regulated pathways.
717 Cells are color-coded by cell types (HSPC, CMP, MK-primed CMP, GMP, MEP,
718 MKP/MK)

719 C- Dot plot showing the expression levels of known specific “DNA repair” marker genes
720 in MEP and MKP/MK of patients and controls. The color is representative of relative
721 average expression in the population compared to all others. Circle size is associated
722 with the percentage of cells expressing the gene in the population.

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

724 A- Translation levels in HEL in basal conditions and after PMA (phorbol 12 myristate 13
725 acetate) stimulation (NT=non-transduced, *ETV6*-WT=transduced with wild-type
726 *ETV6*, *ETV6*-P214L=transduced with *ETV6*-P214L variant). Translation levels were
727 evaluated via flow cytometry after the incorporation of puromycin and staining with
728 anti-puromycin antibody in three independent experiments involving 4 patients.
729 * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, one-sample t-test.

730 B- Translation levels in $CD34^+$ -derived MK on day 11 for P214L variant (F1-III3, IV1,
731 III8) and day 14 for F417Lter4 variant cells (F2 II-2). Translation levels were
732 evaluated via flow cytometry after the incorporation of puromycin and staining with

733 anti-puromycin antibody in three independent experiments involving 4 patients and
734 their respective control.

735 C- Canonical ETV6 binding site contained in RPS6 gene. The list of up-regulated genes
736 in ETV6-MKP/MK was obtained with cistarget (NES 4.85, rank 40e/512).

737 D- Violin plot showing the RPS6 expression levels for each cluster on control and ETV6-
738 variant cells.

739 E- Visualization of RPS6 regulatory region with ETV6 binding peak identified via
740 chromatin immunoprecipitation-sequencing. Part of the *DENND4C* gene is shown
741 after the RPS6 gene using the UCSC genome browser (chr9: 19,372,807 – 19,387,833
742 using GRCh38/hg38).

743 F- The left panel: Western-blot analysis of RPS6 and phosphorylated RPS6 (p-RPS6)
744 expression in washed platelets from controls and ETV6-variant carriers. GAPDH was
745 used as a protein loading control. Quantification of band intensity is shown on the
746 right. Right panel: Flow-cytometry assessment of RPS6 and p-RPS6 in CD34⁺-derived
747 MK from controls and ETV6-variant carriers at day 13 (F1-III3) or 14 (F1-III8, F2-
748 II2) of differentiation. Each patient was compared to matched control subject on age
749 and sex. For each patient and matched control, blood samples were collected at the
750 same time and manipulated in parallel to avoid any bias.

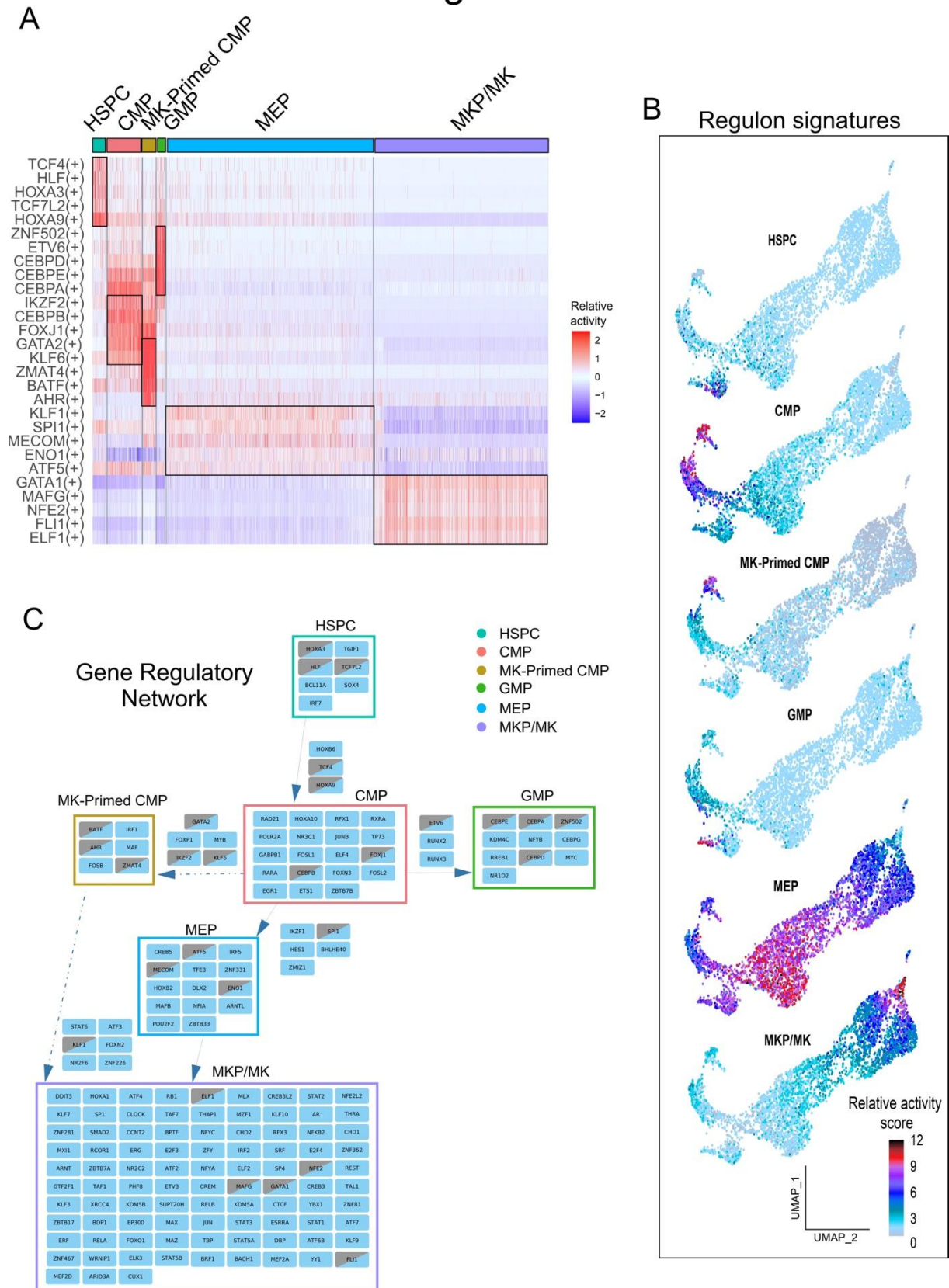
751 G- Treatment of CD34⁺-derived MK with CX-5461 at day 11. Effects of treatment were
752 analyzed after 48h. The left panel: translation levels, RPS6 and p-RPS6 were analyzed
753 by flow cytometry in healthy cells. The right panel: PLP and p-RPS6 were quantified
754 by flow cytometry in *ETV6*-variant carriers. PLP were analyzed in supernatant in
755 “platelet gate” (FSC low SSC low) and by expression of CD42a and MitoGreen, and
756 were normalized on MK number (FS high SSC high, and CD42a expression).

757

758

759

Figure 2



761

762

Figure 4

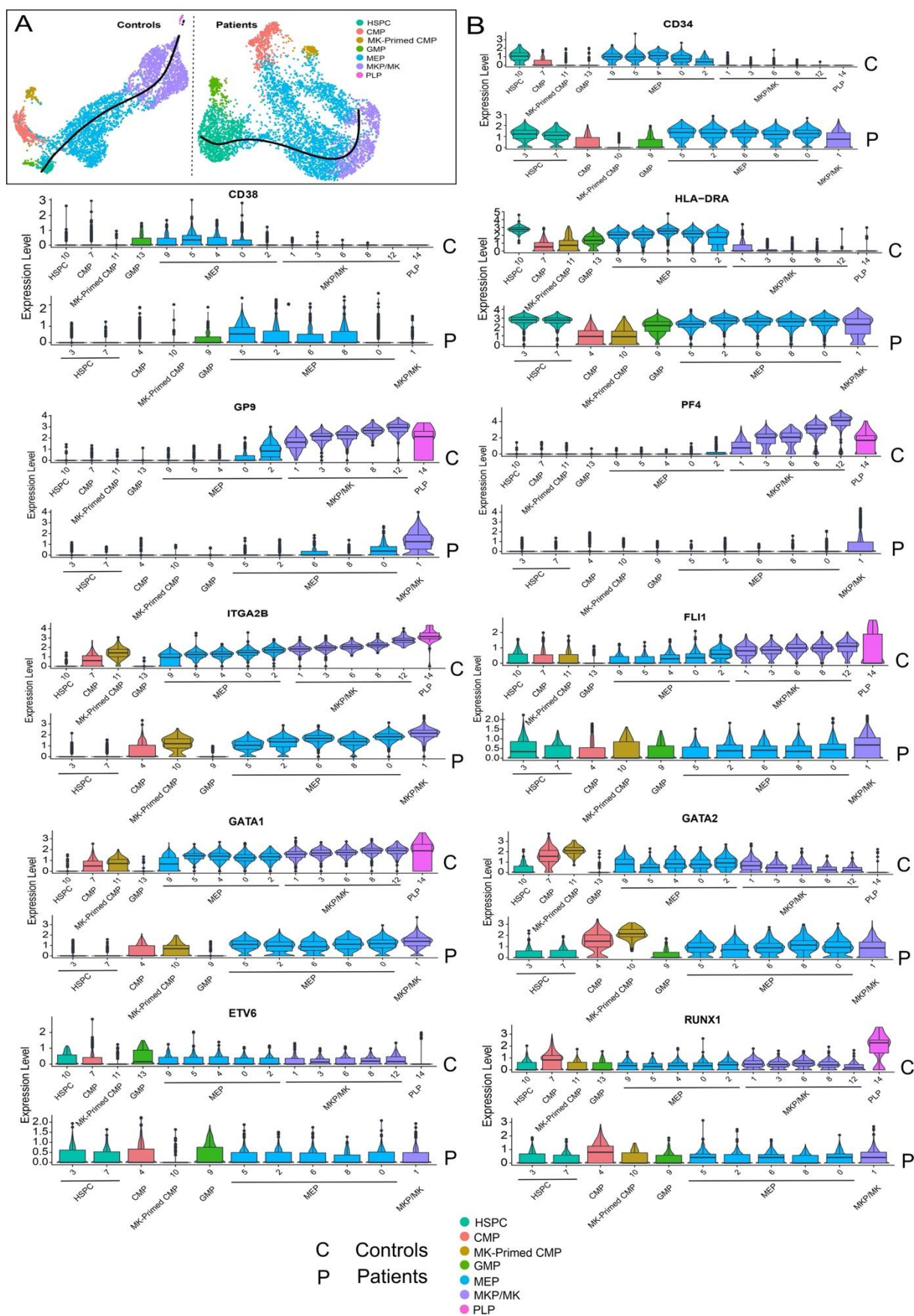
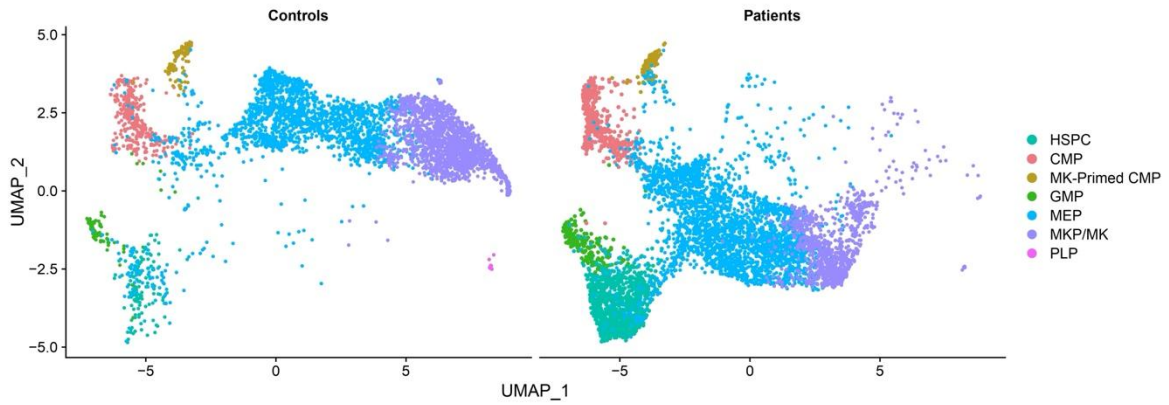
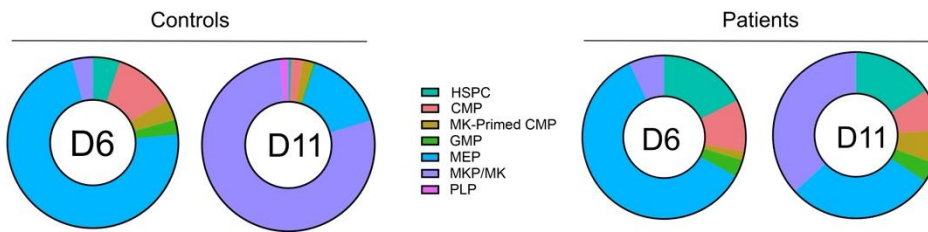


Figure 5

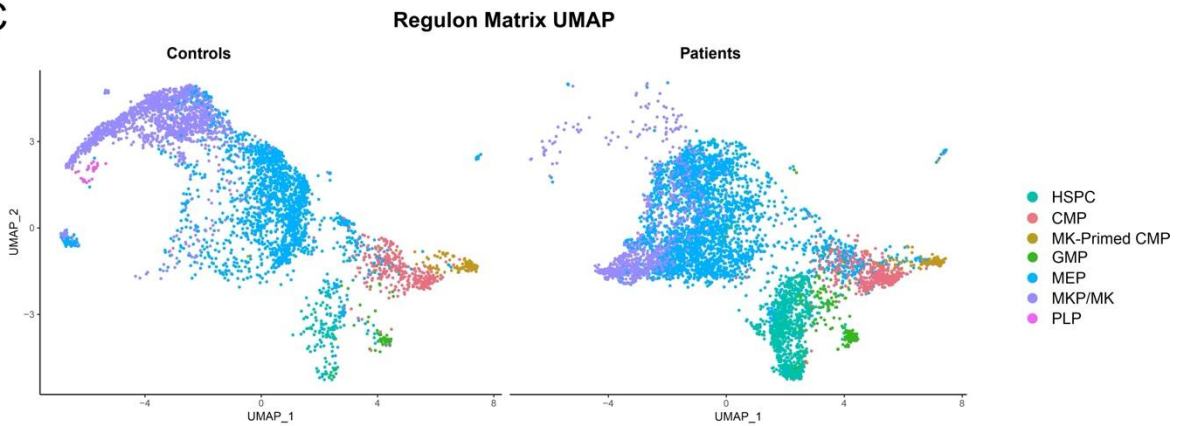
A



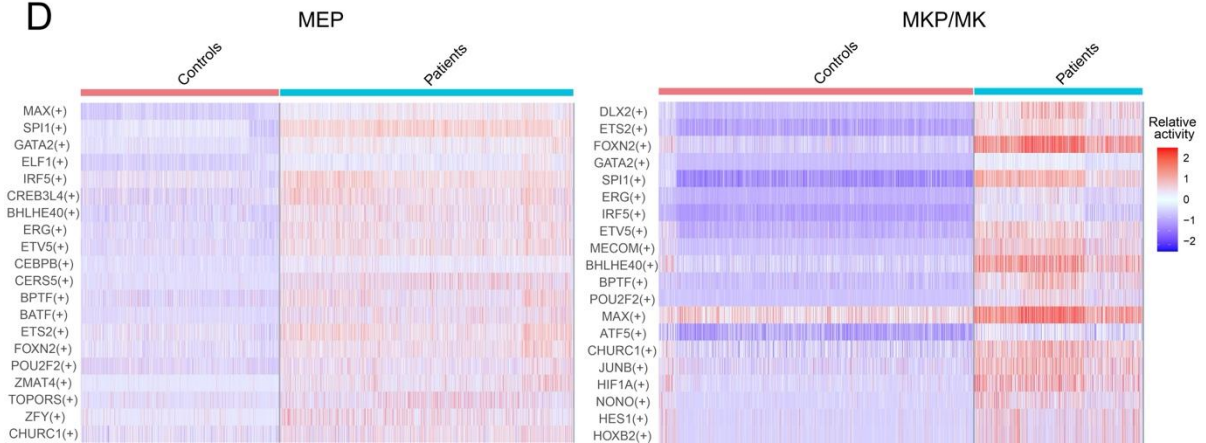
B



C



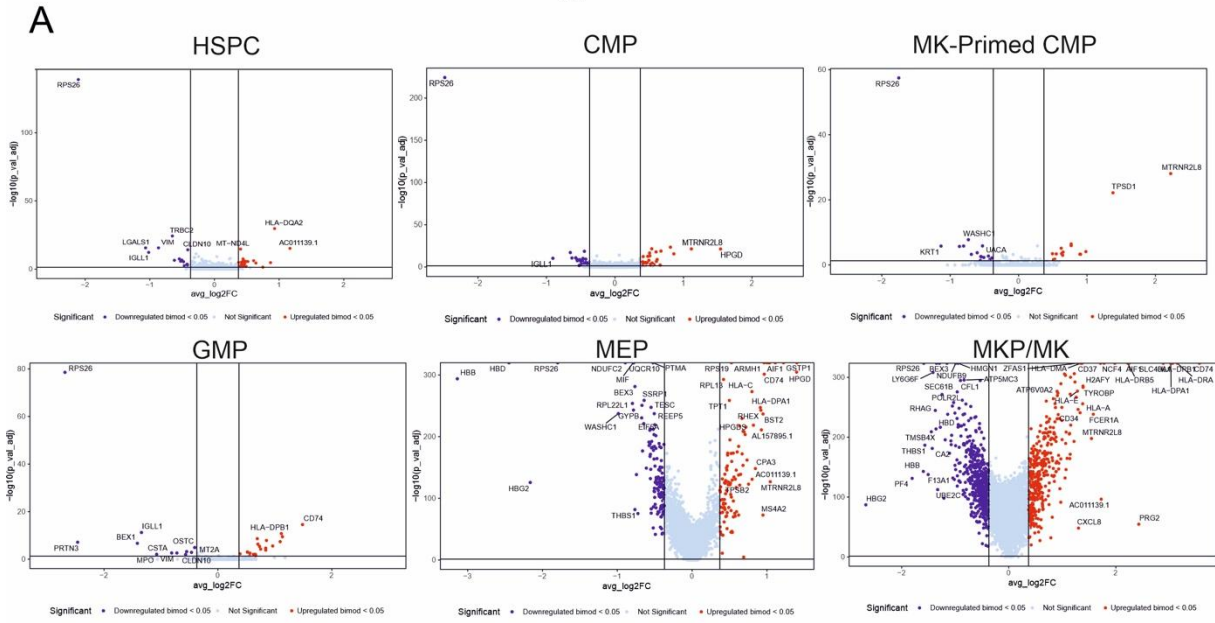
D



766

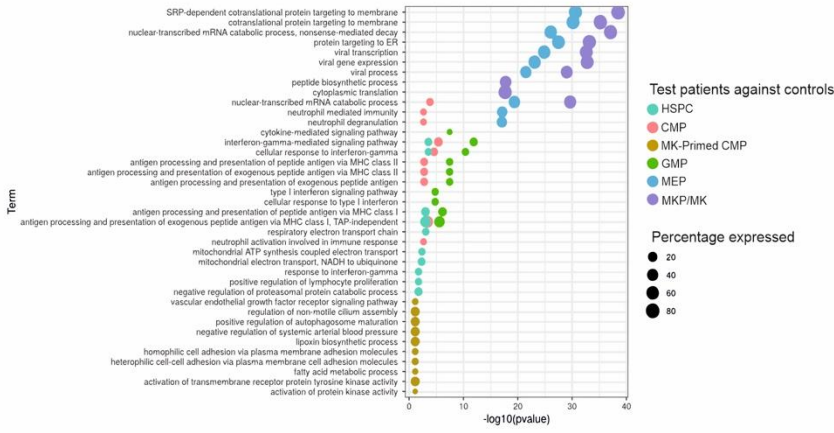
767

Figure 6

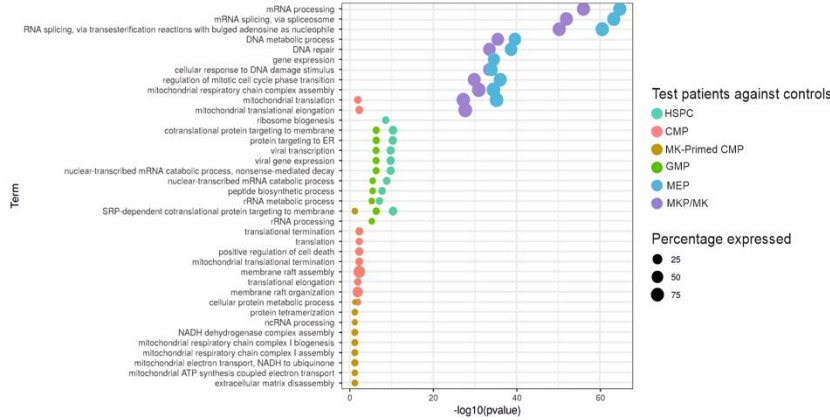


B

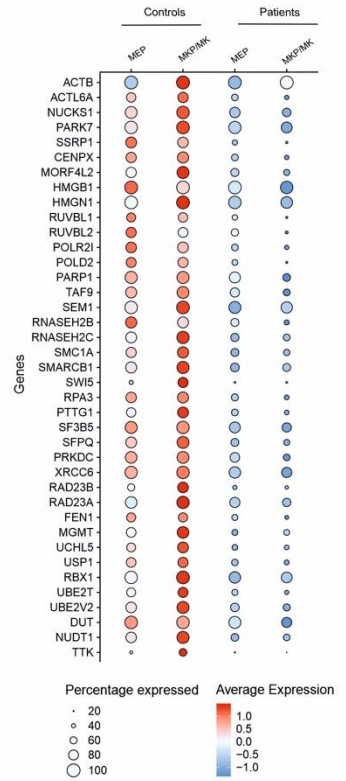
Enrichment analysis of GO biological process upregulated in patient cells



Enrichment analysis of GO biological process downregulated in patient cells



C

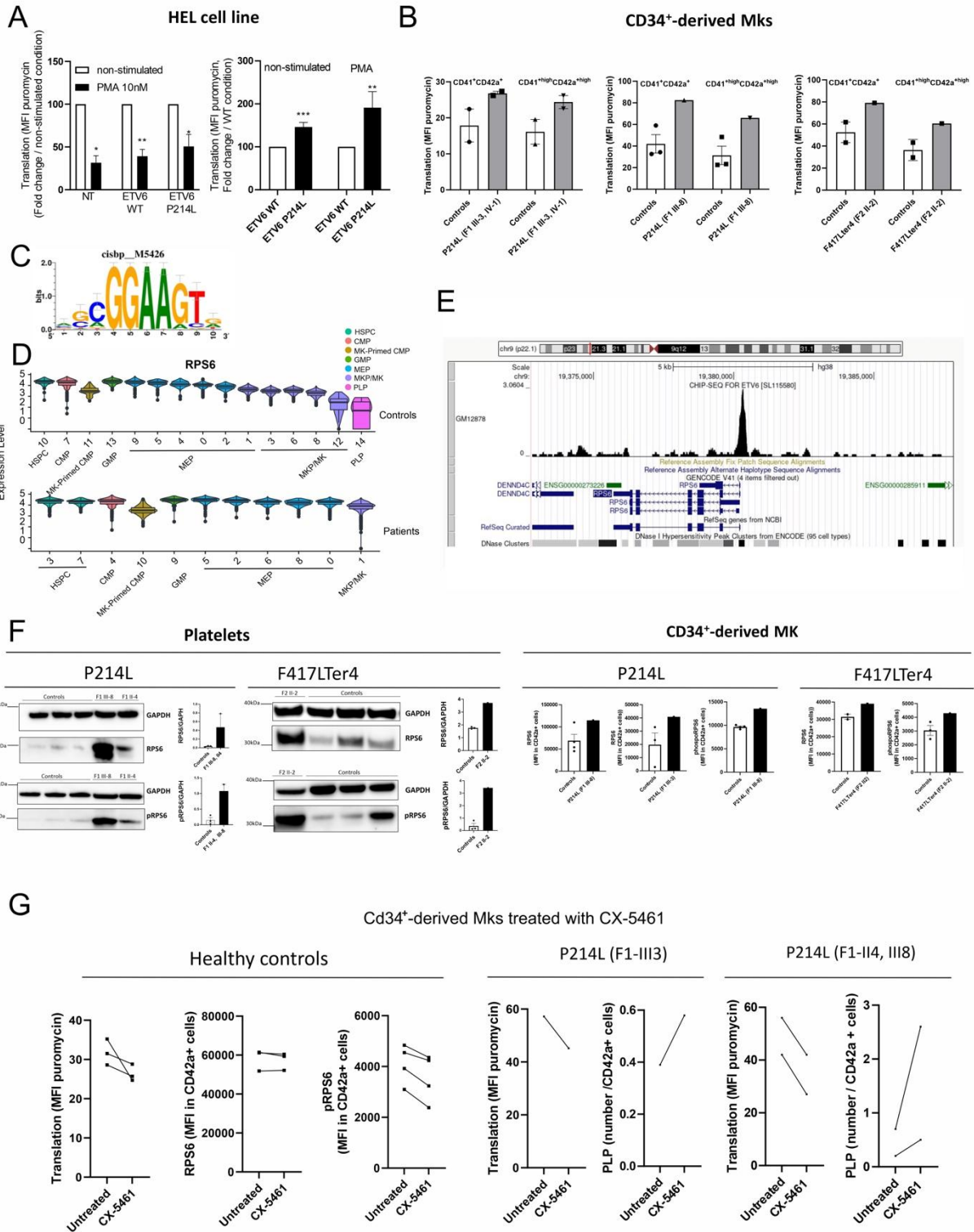


768

769

770

Figure 7



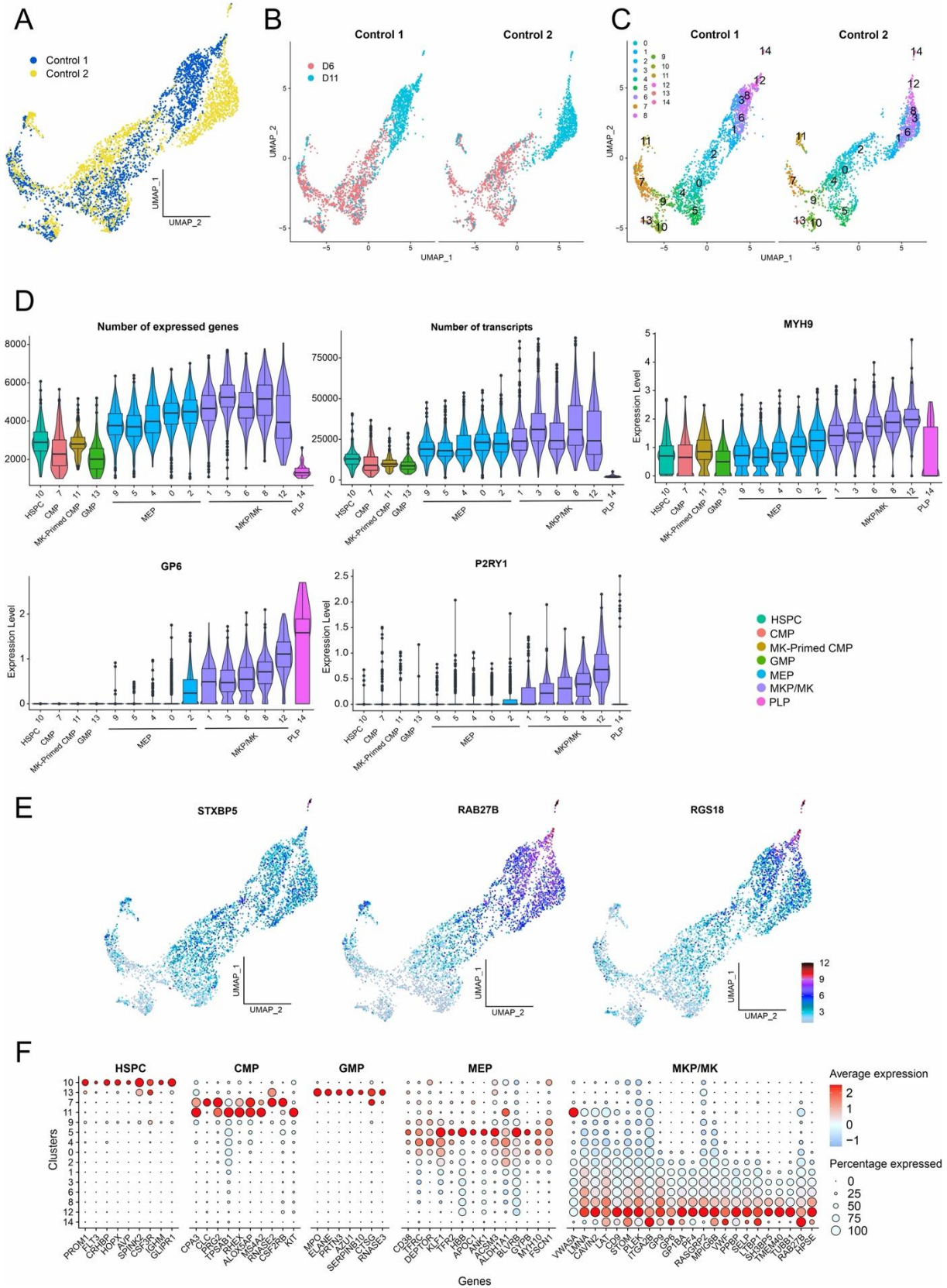
771

772

773

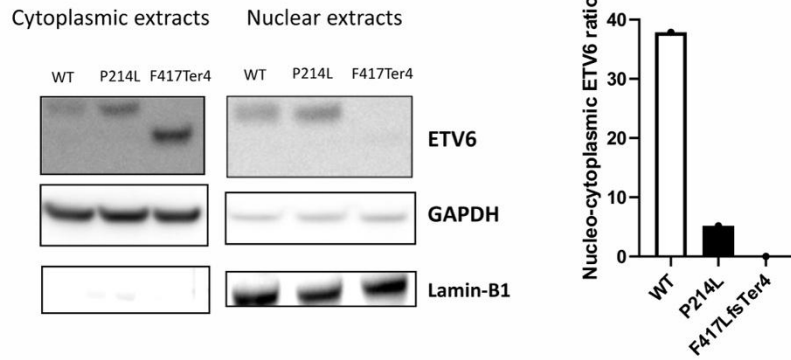
774

Supp. Fig 1

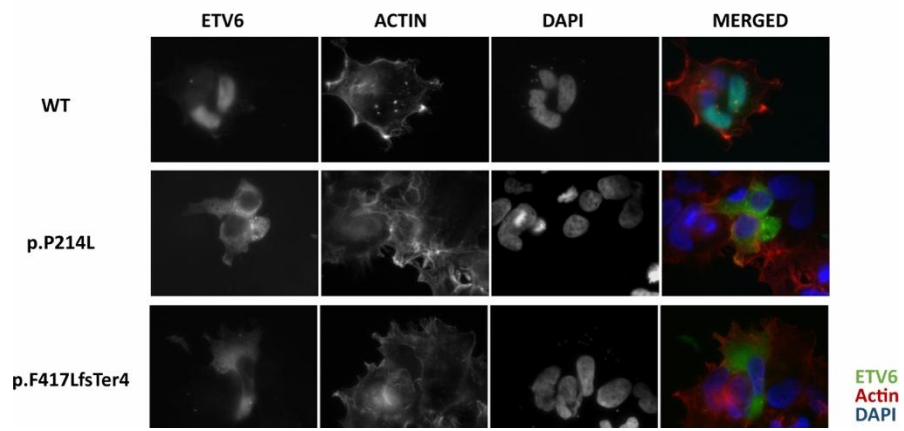


Supp. Fig 3

A

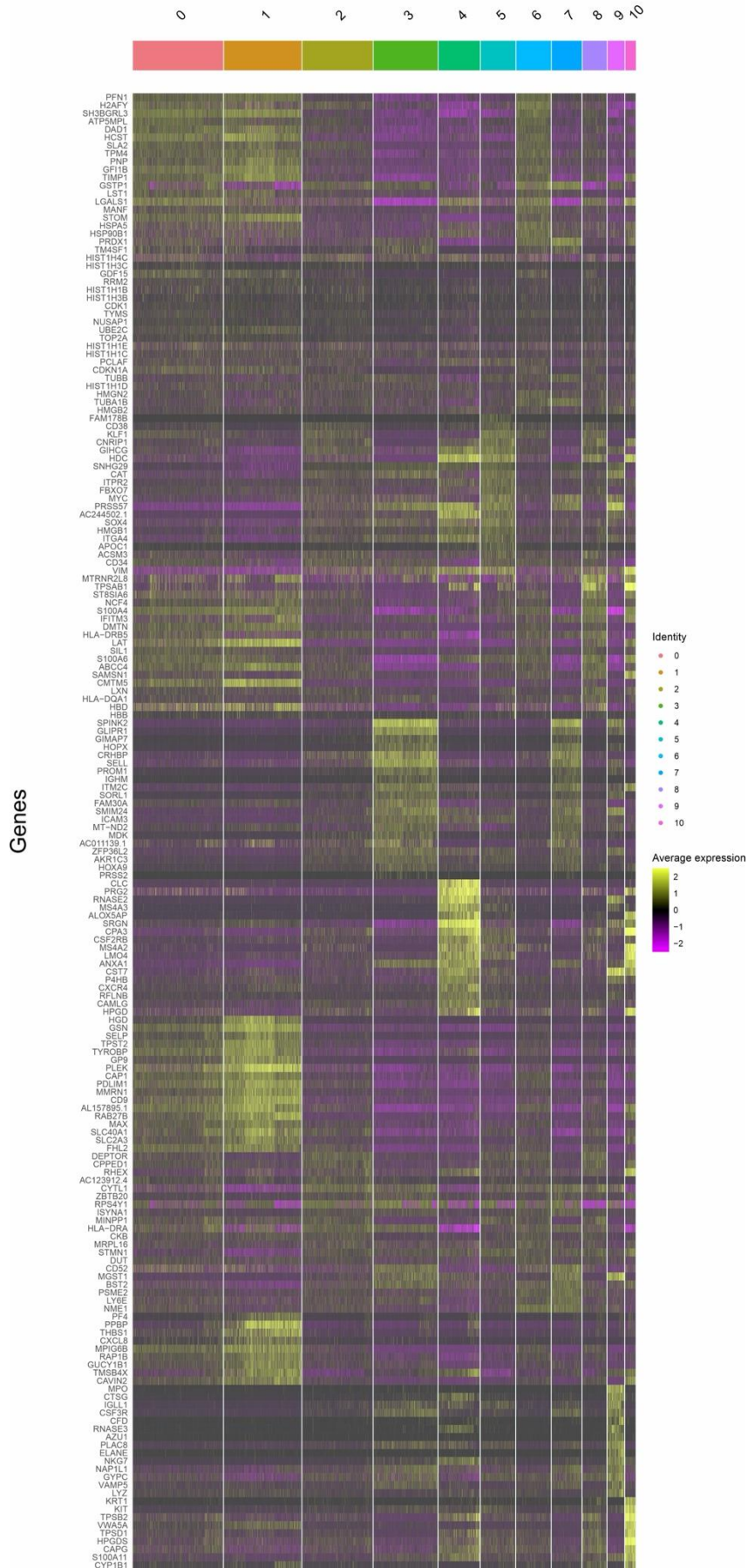


B

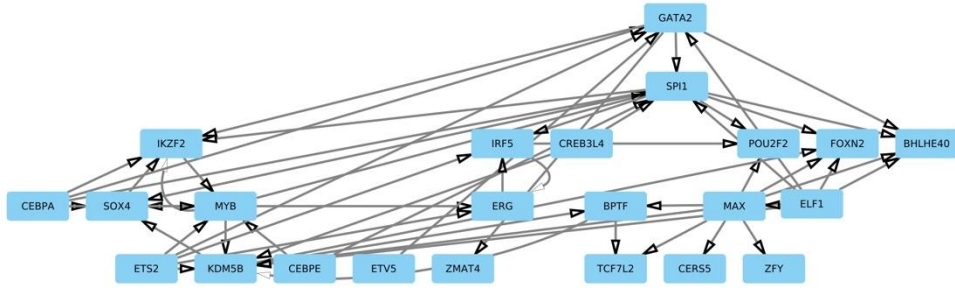


777

Supp. Fig 4



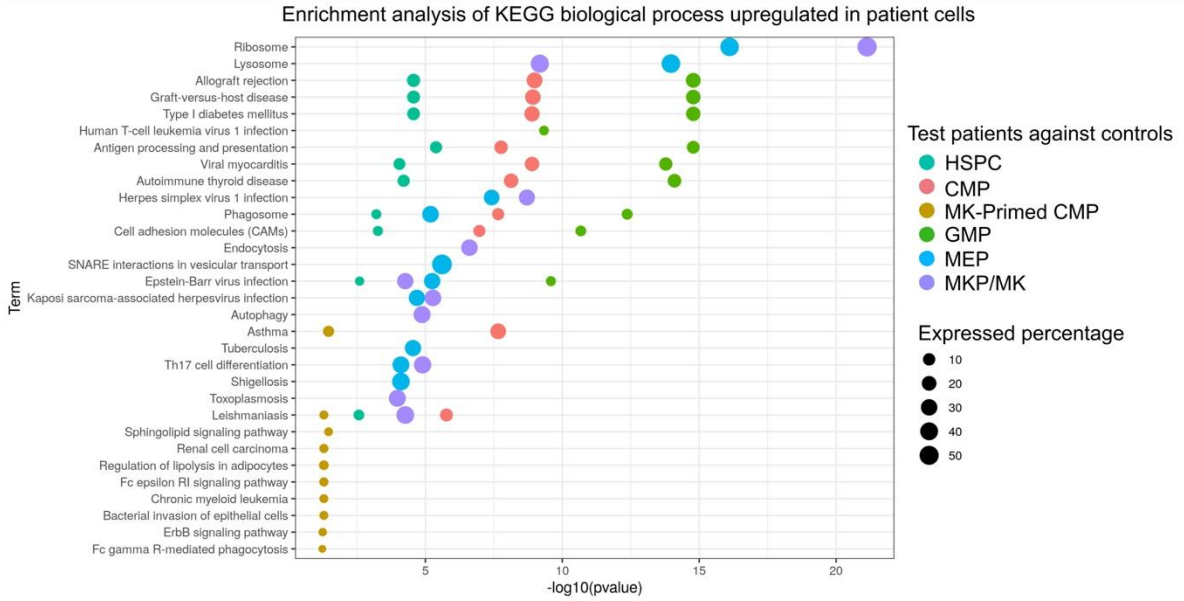
Supp. Fig 5



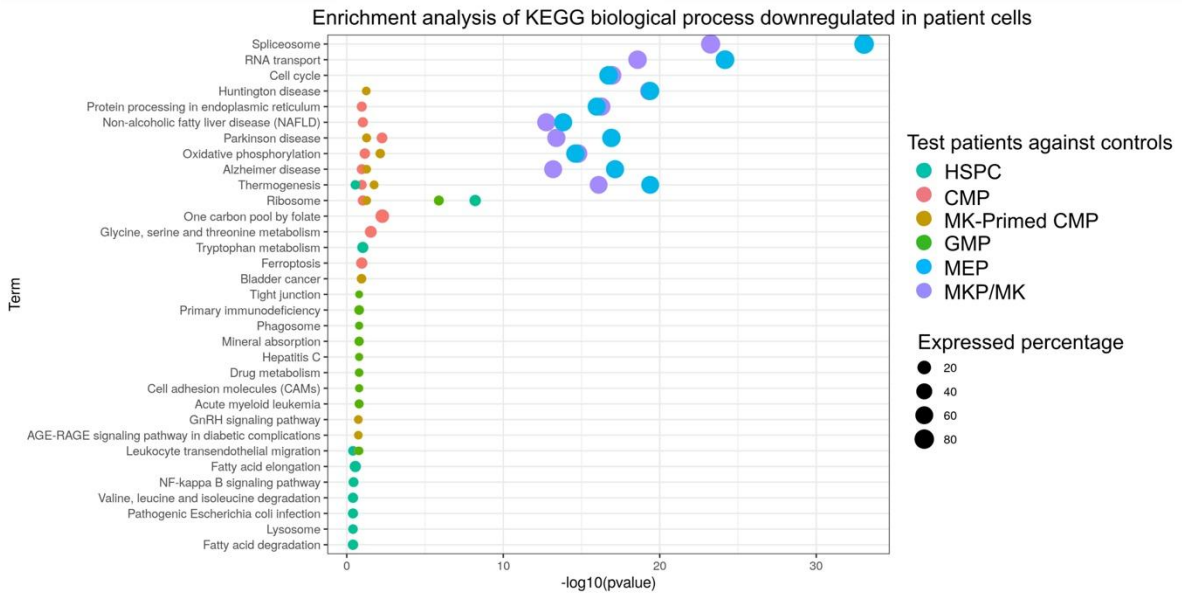
779

Supp. Fig 6

A

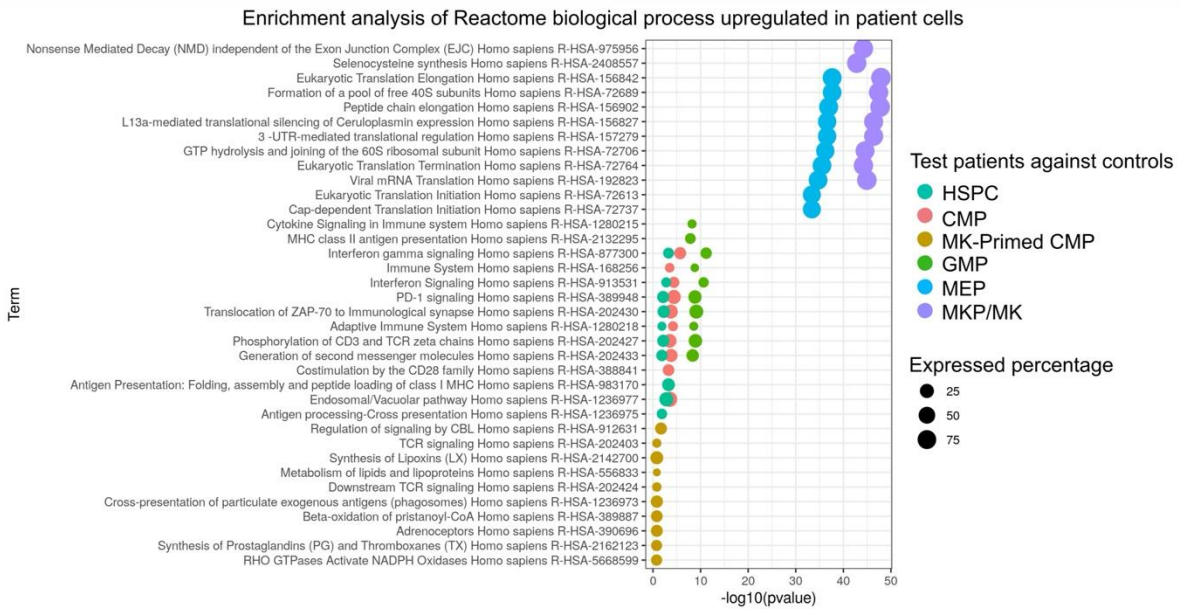


B

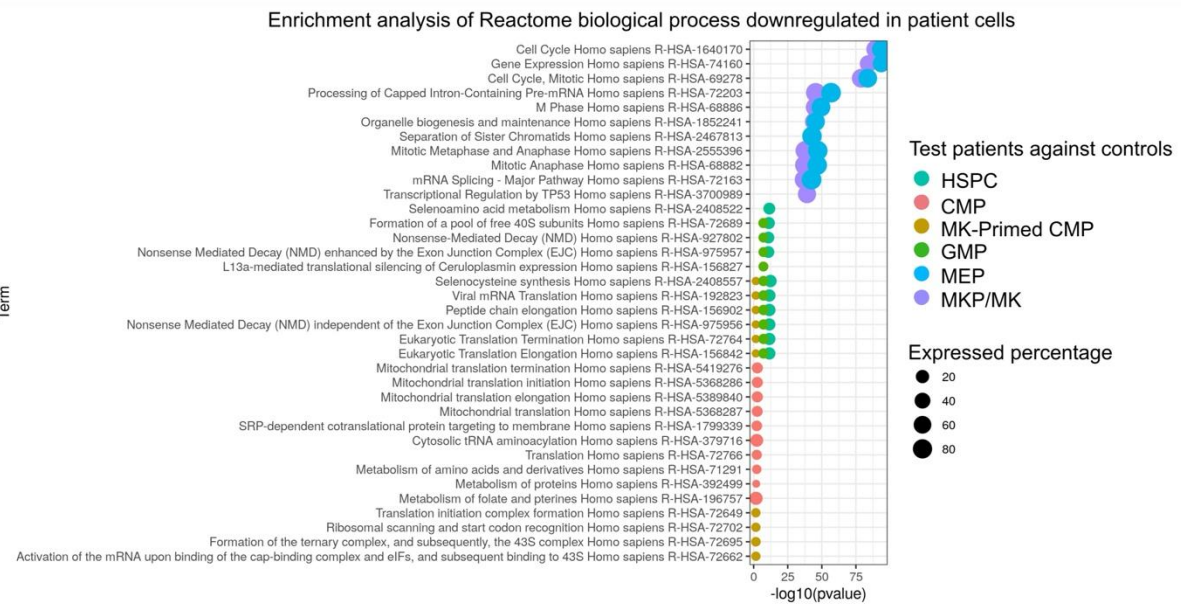


Supp. Fig 7

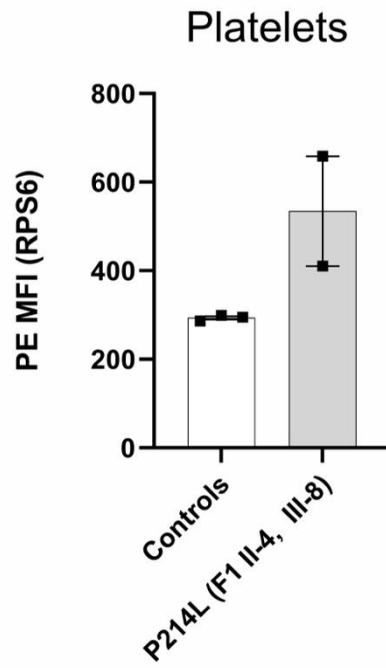
A



B



Supp. Fig 9



783

784

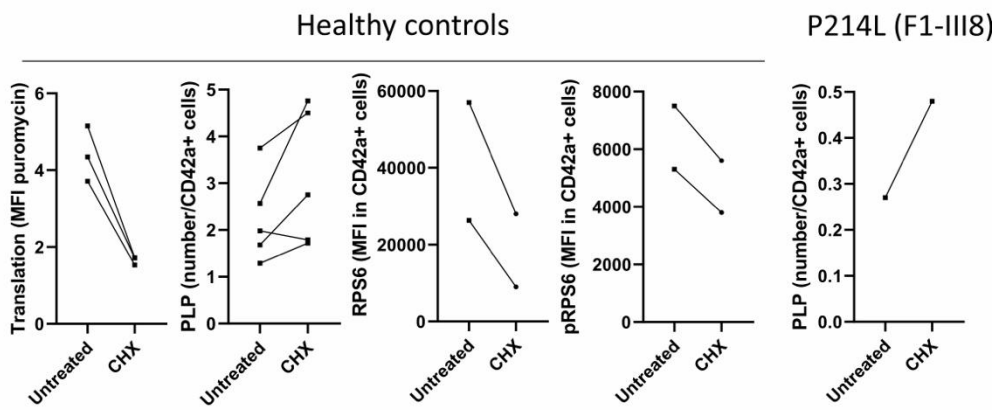
785

786

787

788

Supp. Fig 10



789