Enhanced hemato-endothelial specification during human embryonic differentiation through developmental cooperation between *AF4-MLL* and *MLL-AF4* fusions

Clara Bueno[1,6]*, Fernando J Calero-Nieto[2], Xiaonan Wang[2], Rafael Valdés-Mas[3],

Francisco Gutiérrez-Agüera[1], Heleia Roca-Ho[1], Veronica Ayllon[4], Pedro J Real[4], David Arambilet[5],

Lluis Espinosa[5,6], Raul Torres-Ruiz[1], Antonio Agraz-Doblas[1,7], Ignacio Varela[7], Jasper de Boer[8],

Anna Bigas[5,6], Bertie Gottgens[2], Rolf Marschalek[9], Pablo Menendez[1,6,10]*

[1]Josep Carreras Leukemia Research Institute and Department of Biomedicine, School of Medicine. University of Barcelona. Barcelona. Spain. [2]Department of Hematology, Cambridge Institute for Medical Research and Wellcome Trust-Medical Research Council Cambridge Stem Cell Institute, University of Cambridge, United Kingdom. [3]Dreamgenics S.L. Oviedo. Spain. [4]GENyO, Centre for Genomics and Oncological Research, Pfizer/University of Granada/Andalusian Regional Government and University of Granada, Department of Biochemistry and Molecular Biology, Granada, Spain. [5]Programa de Cáncer. Instituto Hospital del Mar de Investigaciones Médicas. Barcelona. Spain. [6]Centro de Investigación Biomédica en Red de Cáncer (CIBER-ONC), ISCIII, Barcelona, Spain. [7]Instituto de Biomedicina y Biotecnología de Cantabria (CSIC-UC-Sodercan), Departamento de Biología Molecular, Universidad de Cantabria, Santander, Spain. [8]Cancer Section, UCL Great Ormond Street Institute of Child Health, London, United Kingdom. [9]Institute of Pharmaceutical Biology, Goethe-University. Frankfurt. Germany. [10]Instituciò Catalana de Recerca i Estudis Avançats (ICREA). Barcelona. Spain.

Running Title: AF4-MLL in human embryonic hematopoietic development

Key words: AF4-MLL, MLL-AF4, HOX cluster, human ESC, hemato-endothelial specification.

*Correspondence should be addressed to:

Clara Bueno PhD & Pablo Menendez PhD
Josep Carreras Leukemia Research Institute
School of Medicine. University of Barcelona.
Casanova 143, 08036. Barcelona. Spain.
cbueno@carrerasresearch.org ; pmenendez@carrerasresearch.org

ABSTRACT

The t(4;11)(q21;q23) translocation is associated with high-risk infant pro-B-cell acute lymphoblastic leukemia and arises prenatally during embryonic/fetal hematopoiesis. The developmental/pathogenic contribution of the t(4;11)-resulting *MLL-AF4* (MA4) and *AF4-MLL* (A4M) fusions remains unclear; MA4 is always expressed in t(4;11)+ B-cell acute lymphoblastic leukemia patients, but the reciprocal fusion A4M is expressed in only half of the patients. Because prenatal leukemogenesis manifests as impaired early hematopoietic differentiation, we took advantage of well-established human embryonic stem cell-based hematopoietic differentiation models to study whether the A4M fusion cooperates with MA4 during early human hematopoietic development. Co-expression of A4M and MA4 strongly promoted the emergence of hemato-endothelial precursors, both endothelial- and hemogenic-primed. Double fusion-expressing hemato-endothelial precursors specified into significantly higher numbers of both hematopoietic and endothelial-committed cells, irrespective of the differentiation protocol used and without hijacking survival/proliferation. Functional analysis of differentially expressed genes and differentially enriched H3K79me3 genomic regions by RNA-seq and H3K79me3 ChIP-seq, respectively, confirmed a hematopoietic/endothelial cell differentiation signature in double fusion-expressing hemato-endothelial precursors. Importantly, ChIP-seq analysis revealed a significant enrichment of H3K79 methylated regions specifically associated with *HOX-A* cluster genes in double fusion-expressing differentiating hematopoietic cells. Overall, these results establish a functional and molecular cooperation between MA4 and A4M fusions during human hematopoietic development.

INTRODUCTION

The mixed-lineage leukemia (*MLL*) gene encodes for an H3K4 histone methyltransferase important in hematopoietic development[1]. The human *MLL* gene is frequently rearranged in acute leukemia and typically confers a dismal outcome[2,3]. Of particular interest is the translocation t(4;11)(q21;q23), which encodes the fusion proteins *MLL-AF4* (*MA4*) and *AF4-MLL* (*A4M*), and is associated with infant B-cell acute lymphoblastic leukemia (B-ALL). This t(4;11)+ infant leukemia is characterized by a very brief latency, raising the question of how it evolves so quickly[4]. Moreover, the exceptionally high concordance rate of t(4;11)+ B-ALL in monozygotic twin infants[5,6] suggests that all the necessary (epi)genetic events required for leukemogenesis are accomplished prenatally, during embryonic/fetal hematopoietic development[7]. However, our understanding of t(4;11)-mediated developmental effects is limited, at least in part, due to the variety of phenotypes and long latency observed in currently available t(4;11) mouse models[2,8-17]. These different phenotypes likely result from targeting a cell in the wrong developmental stage, or not addressing the impact of secondary hits, leaving open questions about the developmental impact of the t(4;11) translocation during early human development.

The functional and molecular contribution of the reciprocal fusion genes resulting from the derivative translocated chromosomes remains obscure in cancer. The MA4 fusion is always expressed in t(4;11)+B-ALL patients, whereas the reciprocal fusion A4M is expressed in only half of the patients[18-20]. Importantly, t(4;11)+ cell lines display addiction to MA4 but not to A4M[21,22], and although A4M was not sufficient to initiate leukemia in cord blood-derived CD34+ cells[23], it was nevertheless capable of initiating B-ALL in mice without the requirement of MA4, indicating that it contributes to t(4;11)-driven leukemogenesis[11,24,25]. Strikingly, a very recent clinical study have unraveled an independent prognostic value for MA4 expression in t(4;11)+ infant B-ALL, thus adding a new piece to the puzzle[19]. Thus, the developmental/pathogenic contribution of the t(4;11)-resulting reciprocal fusion A4M remains enigmatic.

87 Human embryonic stem cells (hESCs) represent a powerful tool for modeling different developmental aspects

88 of human disease that cannot otherwise be addressed by patient sample analyses or mouse models[7,26,27].

89 Because prenatal leukemogenesis manifests as impaired early hematopoietic differentiation, modeling

90 hematopoietic differentiation in hESCs may represent a promising *in vitro* approach to study the onset of

91 hematopoiesis and the mechanisms underlying early human hematopoietic development[7]. During hESC

92 differentiation, a primitive population of CD45- hemato-endothelial precursors (HEPs) arises and further

93 differentiates into CD45+ hematopoietic and mature endothelial cells[28-30]. Beyond their pathogenic role in

94 acute leukemias, the *MLL* gene has also been implicated in endothelial cell maturation[31], and endothelial

95 dysfunction was recently linked to disease outcomes in childhood leukemias[32]. We previously reported that

96 MA4 favors the emergence of endothelial-primed HEPs but not hemogenic HEPs from hESCs[10]. Here, we

97 took advantage of well-established hESC-based differentiation systems to study whether the A4M fusion

98 cooperates with MA4 during early human hematopoietic and endothelial development. We report a functional

99 and molecular cooperation between MA4 and A4M fusions, which results in an enhanced hemato-endothelial

100 output during human embryonic development.

101

102

103

104

105

106

107

108

109

110

4

METHODS

Vector construction and lentiviral transduction

The cDNAs for MA4 and A4M were subcloned into the pRRL-EF1α-PGK-NEO vector[11, 16]. Both fusions have been described previously (**FigS1A**)[11,23]. We used the following lentivectors containing either neomycin or dTo for cell selection: pRRL-EF1α-PGK-NEO (empty vector;EV), pRRL-EF1α-MA4-PGK-NEO (MA4) and pRRL-EF1α-A4M-PGK-dTo (A4M). VSV-G-pseudotyped lentiviral particles were generated in 293T cells using standard transfection protocols and concentrated by ultracentrifugation[33]. hESCs were infected overnight with concentrated EV or MA4 lentivirus plus 8 μg/ml polybrene. Viral supernatants were washed away the next day, and EV- and MA4-transduced hESCs were then selected with G418 (50-100μg/ml) for three weeks. For dual transduction of MA4 and A4M fusions, G418-resistant MA4-expressing hESCs were infected with A4M-expressing viruses. EV/G418-selected hESCs were also transduced with A4M alone. Transgene expression was confirmed for all the genotypes (**Fig1**).

Human ESC culture and characterization of transgenic human ESC lines

hESCs (AND1 line) were maintained undifferentiated on a layer of irradiated-hMSCs (iMSC) in complete KO-DMEM medium containing 20% knockout serum replacement and 8 ng/mL basic fibroblast growth factor (bFGF)[34, 35]. The medium was changed daily, and cells were passaged weekly by dissociation with 1:1 collagenase IV:Dispase. Cultures were visualized daily by phase contrast microscopy. Approval for hESC work was obtained from the Spanish National Embryo Ethical Committee. Pluripotency of transgenic hESCs was characterized by flow cytometry using antibodies against SSEA-3, SSEA-4 TRA-1-60 and TRA-1-81 (BD Biosciences)[36]. Expression of the pluripotency-associated transcription factors *OCT4*, *NANOG*, *SOX2*, *CRIPTO*, and *DNMT3B* as well as transgene expression (*MA4* and *A4M*) were analyzed by qRT-PCR. **TableS1** shows the primers and PCR conditions used[23,37,38].

Hematopoietic differentiation from human ESCs by embryoid body formation.

Undifferentiated hESCs were treated with collagenase IV:Dispase for 1 hr at 37°C. For embryoid body (EB) formation, cells were transferred to low-attachment plates and incubated overnight in differentiation medium (DM; KO-DMEM supplemented with 20% FBS, 1% nonessential amino acids, 1 mmol/L L-glutamine, and 0.1 mmol/L β-mercaptoethanol). The medium was changed next day to the same DM supplemented with the following hematopoietic cytokines: 300 ng/mL stem cell factor (SCF), 300 ng/mL Flt3L, 10 ng/mL interleukin (IL)-3, 10 ng/mL IL-6, 50 ng/mL granulocyte-colony stimulating factor (G-CSF) and 25 ng/mL bone morphogenetic protein 4 (BMP-4) (all from R&D)[9,29,39-41]. EBs were dissociated at different time points during development using collagenase B and enzyme-free Cell Dissociation Buffer (Invitrogen). Dissociated cells were stained with anti-CD34-PE, anti-CD31-FITC, anti-CD45-APC or anti-CD34-PE-Cy7, CD31-BV510, anti-Glycophorin A, anti-CD43-FITC, anti-CD45-APC antibodies (all from BD Biosciences) and 7-actinomycin D, and analyzed using a FACS Canto flow cytometer[9,29,39-41]. Colony-forming unit (CFU) assays were performed at day 10 and 15 of EB differentiation by plating 60×10^4 EB cells onto serum-free methylcellulose H4435 (Stem Cell Technologies). Colonies were scored after 12 days[9,29,42-44].

Cell cycle and apoptosis analysis

For cell cycle analysis of hESC-derived HEPs and CD45+ cells, day 15 EBs were dissociated and harvested cells were fixed overnight in 70% ice-cold ethanol. Cells were then washed in PBS and incubated with anti-CD31-FITC, anti-CD34-PE-Cy7 and anti-CD45-APC antibodies for 15 min. Cells were then suspended in propidium iodide-containing buffer and acquired-analysed on a FACS Canto-II using Modfit LT4.0 software, discriminating between quiescent cells (G0/G1), cycling cells (S-phase) and G2/M cells[45,46]. Apoptosis was assessed with the Annexin-V apoptosis detection kit (BD Biosciences)[16].

6

Human ESC-OP9 co-cultures

hESC-OP9 co-cultures were performed as described[47,48]. OP9 stroma was prepared by plating OP9 cells in gelatin-coated dishes, and allowing them to overgrow as a monolayer. hESCs were prepared as a suspension of small aggregates using collagenase IV:Dispase. One-tenth of this suspension was plated on top of the 8-day overgrown OP9 stroma. Media was replaced on the next day and one-half volume media changes were performed every other day thereafter. Hematopoietic differentiation was assessed by flow cytometry at day 9 of co-culture. Accordingly, hESC-OP9 co-cultures were treated with collagenase IV/TrypLE and cells were dissociated and filtered through a 70-µm strainer. Cell suspensions were stained with anti-mouse CD29-FITC and anti-human CD34-PE and CD45-APC antibodies. The proportion of HEPs (CD34$^+$CD31$^+$CD45$^-$), and total blood cells (CD45$^+$) was analyzed within the CD29$^-$ human ESC-derived cell population. Hemogenic and endothelial HEPs were distinguished based on CD34 and CD43 expression[40].

Culture of FACS-isolated HEPs in MS5 stroma or liquid culture

Day 9 human hESC-OP9 co-cultures were dissociated as above and both CD45+ cells and HEPs were analyzed. FACS-purified HEPs (CD29$^-$CD34$^+$CD31$^+$CD45$^-$) were plated onto MS5 stroma or in liquid culture for 30 or 16 days, respectively, in DM with hematopoietic cytokines (50 ng/mL SCF, 50 ng/mL Flt3L, 10 ng/mL IL-3,20 ng/mL IL-7). The medium was changed every 7 days, and the emergence of CD45+ hematopoietic cells was analyzed by FACS.

Endothelial differentiation of HEPs

HEPS (2×10^4) from day 9 human hESC-OP9 co-cultures were seeded onto 0.1% gelatin-coated plates in complete EGM-2 medium with microvasculature supplements (Lonza) for 7 days. Cells were then fixed, permeabilized and stained with rabbit anti-human VE-cadherin (Cayman), mouse anti-human eNOS (BD Biosciences), and mouse anti-human vWF (DAKO) followed by Alexa488-conjugated anti-rabbit or Cy3-

181 conjugated anti-mouse (Jakson Immunoresearch) antibodies. Nuclei were counterstained with DAPI. Images

182 were obtained using an inverted fluorescence microscope. Day 7 differentiating cells were trypsinized and

183 cell suspensions were stained with anti-human CD31-FITC and CD144-PerCP-Cy5.5 antibodies.

184

185 **Mouse transplantation and analysis of hematopoietic-endothelial engraftment**

186 NOD/LtSz-scid IL-2Rγ$^{-/-}$ (NSG) mice were housed under sterile conditions. The Animal Care Committee

187 approved all mouse protocols. Briefly, cord bood-derived CD34$^+$ HSPCs (3×10^4 cells) or cells from day 15

188 EBs (5×10^5 cells) were intra-bone marrow (BM)-transplanted as described[49]. Animal health was monitored

189 throughout the entire experiment. Mice were killed 10 weeks after transplantation and cell suspensions were

190 analyzed by FACS for human chimerism using anti-HLA-ABC-FITC, anti-CD31-PE, CD144-PerCP-Cy5.5,

191 and anti-CD45-APC antibodies.

192

193 **Statistical analysis**

194 All data are expressed as mean±SEM. Statistical comparisons were performed using the GraphPad Prism

195 software with the nonparametric Mann-Whitney test, two-tailed P-value (95% confidence interval). Statistical

196 significance was defined as a p-value<0.05.

197

198 **Online methods** show detailed information about RNA- and Chromatin Immunoprecipitation sequencing and

199 analysis.

200

RESULTS

Co-expression of A4M and MA4 does not hijack pluripotency

We showed very recently that only 45% of t(4;11)+ B-ALL patients express the reciprocal fusion A4M, whereas MA4 is consistently expressed in all t(4;11)+ B-ALL patients (**Fig1A**)[18-20]. Here, we generated transgenic hESC lines expressing "MA4 alone", "A4M alone" or MA4+A4M (double fusion, **Fig1B, S1B**). EV (control)- and MA4-hESCs were established by G418 selection[9]. G418-resistant EV- or MA4-expressing hESCs were then transduced with A4M/dTo-expressing lentiviruses and >90% transduction efficiency was achieved. Transgenic hESC lines were maintained for >50 passages and retained human ESC-like morphology (**Fig1B**, left), transgene expression (**Fig1B**, right), and expression of pluripotency-associated transcription factors (**Fig1C**) and surface markers (**Fig1D**). All hESC genotypes formed teratomas in NSG mice (data not shown)[9, 50]. Thus, (co-)expression of A4M and/or MA4 is compatible with hESC pluripotency.

A4M and MA4 co-operate to promote HEPs emergence and enhance blood production

Hematopoietic differentiation was assessed using two distinct and well-established differentiation systems: EB formation[43,47] (**Fig2**) and OP9 co-culture[47,48] (**Fig3**). During differentiation, a population of primitive HEPs arises, which is responsible for further hematopoietic and endothelial commitment[10,30] (**Fig2A,3A**). We investigated whether co-expression of A4M and MA4 impacts hESC-derived hematopoiesis by analyzing the emergence of HEPs during EB development in hESCs individually expressing the single fusions or the double fusion. We observed a pronounced (~5-10-fold; $p<0.05$) increase in HEPs at day 7 and 10 of development in EBs expressing the double A4M and MA4 fusion over single fusions (**Fig2B**, upper-left panel). We next assessed whether co-expression of A4M and MA4 influences subsequent hematopoietic commitment of HEPs. The kinetics of emergence and output of both total CD45+ hematopoietic cells and CD45+CD34+ hematopoietic progenitors was faster (EB day 10) from double fusion-expressing hESCs than from equivalent single fusion-expressing cells, achieving a 2-3-fold higher hematopoietic output by day 15 of EB development

9

225 (**Fig2B**). Furthermore, double fusion-expressing HEPs massively accelerated (EB day10) the emergence of

226 clonogenic hematopoietic progenitors as compared to single fusion-expressing HEPs (**Fig2B**, bottom-right

227 panel). According to our previous work, if the kinetics of hEB differentiation are extended allowing for a

228 continuum HEP-to-blood transition, MA4-expressing hEBs display an enhanced HEP production coupled to

229 an impaired blood output (EB day 20, **Fig S2A**) and clonogenic potential (EB day 15, **Fig 2B**). We confirmed

230 stable expression of ectopic MA4 and A4M upon EB differentiation, supporting the link between genotype

231 and phenotype (**Fig2C**).

232

233 We also investigated hematopoietic differentiation using the OP9 differentiation system (**Fig3A,B**), and by

234 plating FACS-sorted HEPs in either hematopoietic liquid culture (**Fig3C**) or onto MS5 feeders (**Fig3D**). After

235 10 days on OP9 stroma, double fusion-expressing hESCs yielded a 10-fold higher number of CD45+

236 hematopoietic cells than did single fusion-expressing hESCs (**Fig3B**). Moreover, when HEPs were FACS-

237 sorted from day 9 OP9 co-cultures and allowed to differentiate into CD45+ blood cells, the yield of CD45+

238 cells was up to 60-fold higher in double fusion-expressing HEPs than in single fusion-expressing HEPs.

239 (**Fig3C,D**). Encouraged by these results, we next investigated whether ectopic expression of both A4M and

240 MA4 confers *in vivo* engraftment capacity to hESC-derived hematopoietic derivatives. To do this, we

241 transplanted 5×10^5 hESC hematopoietic derivatives from each genotype into myeloablated NSG mice[4,43,47],

242 finding that, despite regulating hematopoietic development *in vitro*, double fusion-expression did not confer

243 *in vivo* engraftment to hESC hematopoietic derivatives (**Fig2D**).

244

245 The increased hematopoietic output of double fusion-expressing hESCs might be the consequence of

246 transgene-mediated proliferation/survival of the emerging HEPs or CD45+ cells. To address this, we

247 analyzed cell cycle distribution (**FigS2B**) and apoptosis (**FigS2C**) within both HEPs and the CD45+ cell

248 population. No differences in the proportion of either cycling HEPs or CD45+ cells were detected between

249 genotypes (25–36% for HEPs and 35–41% for CD45+ cells; **FigS2B**). Apoptotic levels were similarly low

250 between the different genotypes of HEPs (6–8%) and CD45+ cells (5–7%) (**FigS2C**). Collectively, these

251 results show that A4M cooperates with MA4 to induce HEP specification and blood commitment, without

252 hijacking proliferation or survival of HEPs.

253

254 **A4M and MA4 co-operate to enhance endothelial cell fate from HEPs**

255 We next addressed the developmental impact of A4M in endothelial maturation from HEPs[10,47]. We

256 hypothesized that co-expression of A4M and MA4 in HEPs may (i) concomitantly promote subsequent

257 endothelial and hematopoietic commitment or (ii) skew the hemato-endothelial commitment in favor of

258 hematopoiesis. To test this, we analyzed the ability of HEPs to differentiate into mature endothelial cells.

259 OP9-hESC co-cultures were dissociated on day 9 of development and HEPs were FACS-sorted and cultured

260 for a week in endothelial-promoting conditions (**Fig4A**). The expression of the mature endothelial markers

261 VE-cadherin (CD144), vWF, eNOS and CD31 was then analyzed. Irrespective of the genotype, HEPs

262 cultured in endothelial conditions attached, became spindle-shaped, and formed VE-Cad+ endothelial-like

263 structures co-expressing eNOS, vWF and CD31 (**Fig4B,C**, top panel). However, double fusion-expressing

264 HEPs were more prone to differentiate into mature endothelial cells than single fusion-expressing HEPs.

265 Accordingly, they yielded a 20-fold higher number of VE-Cad+ endothelial-like structures (**Fig4C**, top panel)

266 and CD144+CD31+ endothelial cells (**Fig4C**, bottom panel). Interestingly, endothelial cells

267 (HLA.ABC⁺CD31⁺CD34⁺CD144⁺CD45⁻CD43⁻) were found in the BM of mice transplanted with double fusion-

268 expressing hESC blood derivatives at levels ~4-fold higher than in mice transplanted with single fusion-

269 expressing cells (**Fig4D**).

270

271 Within CD34⁺CD31⁺CD45⁻ HEPs, two subpopulations of phenotypically and functionally distinct HEPs can be

272 distinguished based on the expression of CD34 and CD43: hemogenic HEPs (CD34^low/+CD43⁺CD45⁻) and

11

273 endothelial HEPs (CD34^{++}CD43$^-$CD45$^-$) (**Fig5A**)[40,48,51]. We thus analyzed the contribution of both HEP

274 populations to the superior hematopoietic and endothelial differentiation observed in double fusion-

275 expressing HEPs. Co-expression of A4M and MA4, but not single fusions, robustly enhanced the emergence

276 of both endothelial and hemogenic HEPs (**Fig5B,C**). The identity of hemogenic and endothelial HEPs was

277 confirmed by the specific expression of early hematopoietic and endothelial master genes (**Fig5D**). Thus,

278 A4M cooperates with MA4 to promote hematopoietic and endothelial cell fate.

279

280 **Genome-wide transcriptomic and H3K79 methylation profiles support the developmental cooperation**

281 **between A4M and MA4**

282 To identify patterns of gene expression that might molecularly explain the functional co-operation between

283 A4M and MA4 in hematopoietic specification, we performed RNA-seq analysis on FACS-purified EV-, MA4,

284 A4M- and double fusion-expressing HEPs from day 15 hEBs. **Fig6A** shows a heatmap representation of the

285 hierarchical clustering of the 335 genes differentially expressed between the four genotypes (**TableS2**).

286 There is a clear transcriptomic transition towards a hematopoietic/endothelial gene signature from EV-HEPs

287 to double-fusion-expressing HEPs. Single fusion-expressing HEPs clustered interspersed between EV and

288 double-fusion HEPs. The biological functions affected by genes differentially expressed in MA4-, A4M- and

289 double fusion-expressing HEPs relative to EV were classified by IPA software[47,52] and among the top

290 significant enriched functional categories were "hematological system development and function", "cancer"

291 and "hematological disease" (**Fig6B**). Statistical (-log(p-value)) power shows distant effects of MA4 and A4M;

292 however, co-expression of both fusions seems to establish a molecular balance/developmental cooperation

293 in promoting blood-endothelial specification from hPSCs. Strikingly, biofunctions specifically associated to

294 hematological lymphoid malignancies (not with non-hematological cancer) were predicted to be activated

295 (positive z-score) exclusively in double fusion-expressing HEPs, further suggesting a molecular cooperation

296 between MA4 and A4M in development and infant leukemia. (**Fig6C**).

12

297   The C-terminal-partners of MLL fusions normally interact with the histone methyltransferase DOT1L, which

298   is the sole histone methyltransferase catalyzing histone 3 lysine 79 (H3K79) methylation, a chromatin

299   modification widely associated with the dysregulated expression of *HOX-A* cluster genes in MLL leukemias

300   [13, 53]. We thus performed genome-wide ChIP-seq analysis of the H3K79 trimethylation (H3K79me3) profiles

301   in control, MA4-, A4M- and double fusion-expressing hESC-derived blood derivatives (**Fig7,S3A, TableS3**).

302   In line with the RNA-seq data, functional analysis of the differential H3K79me3 peaks specific for double

303   fusion-expressing cells revealed significant GO functional categories associated with "definitive

304   hematopoiesis", "myeloid and erythroid differentiation/homeostasis" and "endothelial cell development"

305   (**Fig7A,S3B**). This further supports the developmental co-operation between A4M and MA4 in promoting

306   hemato-endothelial specification.

307

308   Finally, we analysed the H3K79me3 profiles at genomic loci of the classical MLL target genes reported by

309   Guenther *et al*[54]. Non-*HOX-A* classical MLL targets such as *RUNX1, LMO2, ADMA10, and KDM6A* showed

310   a slight although non-significant, enrichment of H3K79me3 in MA4-expressing hESCs, validating our ChIP-

311   seq approach (**Fig7B**). However, enrichment of H3K79me3 in *HOX-A* cluster genes was observed

312   exclusively in A4M-expressing cells although it was statistically significant only in double fusion-expressing

313   differentiating hESCs (FDR<0.1) (**Fig7C**). As such, *HOX-A* genes were up-regulated in double-fusion-

314   expressing hematopoietic clonogenic progenitors (**FigS3C**). No differential enrichment of the repressive

315   H3K4me2/3 mark was observed in either *HOX-A* or non-*HOX* genes in double fusion-expressing cells

316   (**FigS4**). Collectively, these data suggest that the deregulated expression of *HOX-A* genes in MLL leukemias

317   may be imposed by the reciprocal A4M fusion through a H3K79 methyltransferase activity. In support of this,

318   a recent RNA-seq study performed in 42 infants with t(4;11)+ B-ALL enrolled in the Interfant treatment

319   protocol, reveal that 45% of t(4;11)+ patients express the A4M fusion, and that *HOX-A* cluster genes are

320 exclusively expressed in this *AF4-MLL*-expressing subgroup of t(4;11)+ patients, who in fact display a

321 significant more favorable clinical outcome[19].

322

323 **DISCUSSION**
324
325 From an etiological and pathogenesis standpoint, infant cancer is distinct to adult cancer and should be

326 studied as a developmental disease[4,7,16]. A biologically and clinically intriguing infant cancer is the t(4;11)+

327 B-ALL, which is associated with a dismal outcome[4,21,23]. Evidence in support of its prenatal origin comes from

328 studies in monozygotic twins and archived blood spots, providing compelling evidence of a single prenatal

329 cell as the origin for t(4;11)[5], and also from recent genome-wide studies demonstrating that this infant

330 leukemia has one of the lowest frequencies of somatic mutations of any sequenced cancer[55]. The stable

331 genome of these patients suggests that in infant developmental cancer, one "big-hit" might be sufficient for

332 overt disease, supporting a key contribution of the prenatal cell-of-origin during a critical developmental

333 window of stem cell vulnerability in leukemogenesis. However, despite its aggressiveness and short latency,

334 our current understanding about its etiology, pathogenesis and cellular origin is still limited[2,4,16,52]. Importantly,

335 a recently developed xenograft model which represents the most *bona fide* model for t(4;11)+ B-ALL so far,

336 has revealed the instructive role of MLL-Af4 in cord blood-derived CD34+ cells[14].

337

338 Studies using primary cells from t(4;11)+ B-ALL patients are incapable of addressing the developmental

339 genesis of the hematopoietic system. Recent data suggest that fetal liver lymphoid-primed multipotent

340 progenitor may provide the developmental prerequisites for the initiation of t(4;11)+/MLL-AF4 infant

341 leukemia[56]. Because leukemogenesis manifests as a blockage or altered cell differentiation, the

342 hematopoietic differentiation of hESCs may represent a promising *in vitro* model to study the onset of

343 hematopoiesis and the earliest events leading to the specification of the hematopoietic cells[36]. Previous

344 studies have addressed the oncogenic role of leukemic fusion genes in hESC-derived hematopoiesis[57,5859].

345   We previously explored the developmental impact of the pre-natal fusion MA4 in hESC hemato-endothelial

346   development[10], and found that MA4 expression promotes the emergence of endothelial-primed HEPs and

347   further endothelial commitment, but hijacks the specification of hemogenic-primed HEPs, impairing

348   hematopoietic output[10].

349

350   The functional and molecular contribution of the reciprocal fusion genes resulting from the derivative

351   translocated chromosomes remains obscure in cancer[18,60]. The MA4 fusion is always expressed in t(4;11)+B-

352   ALL patients, whereas the reciprocal fusion A4M is expressed in only 45-50% of the patients[18,60,19,20]. Here,

353   we took advantage of well-established hESC-based differentiation systems to study whether the A4M fusion

354   cooperates with MA4 during early human hematopoietic and endothelial development. Co-expression of A4M

355   and MA4 strongly promoted the emergence of HEPs, both endothelial-primed and hemogenic-primed.

356   Moreover, the double fusion-expressing HEPs specified into significantly higher numbers of both

357   hematopoietic and endothelial cells, irrespective of the *in vitro* differentiation protocol used and without

358   affecting survival or proliferation, indicating a functional (developmental) co-operation between MA4 and A4M

359   fusions during human hematopoietic development. This notion was confirmed by genome-wide

360   transcriptomic analysis of differentiating HEPs. These developmental biology studies support previous

361   evidence suggesting that A4M contributes to the pathobiology of t(4;11)+ B-ALL. Accordingly, Bursen[11]

362   reported that A4M-transduced murine hematopoietic stem cells (HSCs) developed pro-B-ALL, whereas co-

363   transduction with MA4 and A4M resulted in mixed lineage leukemia. Moreover, studies from Milne's

364   laboratory demonstrated that RUNX1 is directly activated by MA4 and the RUNX1 protein interacts with the

365   A4M protein, suggesting a mechanism of co-operation between the two fusion genes at the molecular level[25].

366

367   In the embryo, definitive hematopoiesis cannot occur in the absence of endothelial cell development.

368   Definitive HSCs in both mouse and human emerge from the hemogenic endothelium by a process known as

369 endothelial-to-hematopoietic transition[61]. Hematopoietic differentiation of hESCs occurs through the

370 generation of HEPs, from which then originate both endothelial and hematopoietic cells. Here, co-expression

371 of A4M and MA4 in HEPs concomitantly promoted endothelial and hematopoietic commitment rather than

372 skewing the hemato-endothelial commitment in favor of one lineage over the other. This finding is important

373 because beyond their pathogenic role in acute leukemias, the *MLL* gene is implicated in endothelial cell

374 maturation, and endothelial dysfunction was recently linked to disease outcome in childhood leukemias[31,32].

375 Furthermore, other leukemia fusion oncogenes as well as lymphoma-specific genetic aberrations have been

376 found in endothelial cells from chronic myeloid leukemia and B-cell lymphoma patients[10,43,44], suggesting that

377 endothelial cells may be part of the neoplastic clone. Also, BM-derived mesenchymal stem cells (BM-MSCs)

378 from infant t(4;11)+ B-cell ALL were recently found to harbor and express the t(4;11) translocation[33]. The

379 existence of a common embryonic precursor for MSCs and endothelial cells has been recently demonstrated

380 by hESC-directed differentiation[10,45]. The finding of such a common embryonic precursor, and the presence

381 of t(4;11) in both leukemic blasts and BM-MSCs of infant patients, suggests that the t(4;11) translocation

382 arises and has a developmental impact on early pre-hematopoietic precursors. As a technical caveat, it is

383 important to emphasize that MA4 and A4M were sequentially transduced to allow for antibiotic selection and

384 homogeneously-transduced hESC cultures. However, in double-fusion-expressing differentiating blood cells,

385 MA4 was never individually expressed in the absence of A4M. When hematopoietic differentiation of hESCs

386 was induced by EB formation both fusions were readily co-expressed, ruling out a biased functional

387 phenotype driven by the sequential expression of each transgene in distinct developmental windows.

388

389

390 Mechanistically, a putative function of A4M is to activate chromatin, rendering a chromatin landscape similar

391 to that during stem cell development. It is currently unknown how A4M is able to reprogram chromatin, but it

392 does contain the SET domain disrupted from its "specification domain", the N-terminal portion of MLL, which

393 binds to MEN1 and LEDGF, shaping the gene targeting module of the MLL gene. When A4M is expressed,

394 the N-terminal portion is substituted by the AF4 N-terminus, which is the crucial domain (AF4N) that binds to

395 and strongly activates RNA polymerase II (RNAP II) for transcriptional elongation. Overexpression of either

396 AF4, AF4N or the fusion protein A4M induces robust RNAP II-dependent gene transcription by overwriting

397 the elongation control process in a dominant fashion[62-64]. Since gene transcription *per se* and in particular

398 "sterile" transcription is a powerful mechanism for chromatin activation, A4M could potentiate MA4 to skew

399 normal and leukemic hematopoietic cell fate decisions. This also explains why MA4 has a more prominent

400 role in the disease than the reciprocal A4M. If A4M functioned to initiate this process only by itself, then it

401 would become obsolete after fulfilling the "chromatin opening job". However, transcription factors like MA4,

402 RUNX1 or others then establish the transcriptional program leading to leukemogenesis. This is reflected in

403 the enhanced hematopoietic potential of double fusion-expressing hESCs and the enriched H3K79me3

404 activation mark in *HOX-A* cluster genes exclusively when MA4 and A4M are co-expressed. Thus, A4M

405 prepares other transcription factors to become oncoproteins.

406

407 Molecularly, C-terminal-partners of MLL fusions (AF4, AF9, ENL) interact with DOT1L, which is the sole

408 histone methyltransferase catalyzing H3K79 methylation, a chromatin modification widely associated with the

409 dysregulated expression of *HOX-A* gene cluster in MLL-rearranged leukemias[13,53]. Here, ChIP-seq analysis

410 of differentially enriched H3K79me3 genomic regions confirmed a hematopoietic/endothelial cell

411 differentiation signature in double fusion-expressing HEPs, and revealed a significant enrichment of H3K79

412 methylated regions specifically associated with *HOX-A* cluster MLL target genes (but not to non-*HOX-A* MLL

413 targets) in double fusion-expressing differentiating hematopoietic cells. This is in line with the recently found

414 significant positive correlation between the upregulation of the *HOX-A* gene cluster and the expression of

415 A4M in primary t(4;11)+ infant B-cell ALL samples, and with previous studies identifying that approximately

416 one-half of t(4;11)+ patients do not have an activated *HOX-A* signature[20,65,66]. This may explain why MA4

417 failed recently to bind to *HOX-A* genes to regulate *HOXA* gene expression[14]. Collectively, MA4 and A4M

418 might cooperate through a complex molecular interaction to control *HOX-A* gene regulation[25]. In sum, we

419 describe a functional and molecular cooperation between MA4 and A4M fusions during human hematopoietic

420 development, and demonstrate how hESC-based hematopoietic differentiation represents a promising

421 system to explore the developmental impact of the chimeric proteins resulting from chromosomal

422 translocations, which remains obscure in human leukemia.

423 **AUTHOR CONTRIBUTION:** C.B. conceived the study, designed and performed experiments and analyzed

424 data. P.M. conceived the study, designed experiments, analyzed data and wrote the manuscript. F.C.N.,

425 X.W., R.V.M., F.G-A., H.R.H., V.A., P.J.R., D.A., L.E., R-T-R., A.A-D., J.dB. performed experiments and

426 analyzed data. I.V., A.B., B.G., and R.M. contributed intellectually and financially.

427

439

440 **CONFLICT OF INTEREST DISCLOSURE:** The authors have nothing to disclose.

441

442

443

444

445

446

447 **REFERENCES**

448 1. Milne TA. Mouse models of MLL leukemia: recapitulating the human disease. Blood.
449 2017;129(16):2217-23.
450
451 2. Montes R, Ayllon V, Prieto C, Bursen A, Prelle C, Romero-Moya D, et al. Ligand-independent FLT3
452 activation does not cooperate with MLL-AF4 to immortalize/transform cord blood CD34+ cells. Leukemia.
453 2014;28(3):666-74.
454
455 3. Stam RW, den Boer ML, Schneider P, Nollau P, Horstmann M, Beverloo HB, et al. Targeting FLT3
456 in primary MLL-gene-rearranged infant acute lymphoblastic leukemia. Blood. 2005;106(7):2484-90.
457
458 4. Sanjuan-Pla A, Bueno C, Prieto C, Acha P, Stam RW, Marschalek R, et al. Revisiting the biology of
459 infant t(4;11)/MLL-AF4+ B-cell acute lymphoblastic leukemia. Blood. 2015;126(25):2676-85.
460
461 5. Ford AM, Ridge SA, Cabrera ME, Mahmoud H, Steel CM, Chan LC, et al. In utero rearrangements
462 in the trithorax-related oncogene in infant leukaemias. Nature. 1993;363(6427):358-60.
463
464 6. Greaves M. Infection, immune responses and the aetiology of childhood leukaemia. Nat Rev Cancer.
465 2006;6(3):193-203.
466
467 7. Bueno C, Montes R, Catalina P, Rodriguez R, Menendez P. Insights into the cellular origin and
468 etiology of the infant pro-B acute lymphoblastic leukemia with MLL-AF4 rearrangement. Leukemia.
469 2011;25(3):400-10.
470
471 8. Barrett NA, Malouf C, Kapeni C, Bacon WA, Giotopoulos G, Jacobsen SE, et al. Mll-AF4 Confers
472 Enhanced Self-Renewal and Lymphoid Potential during a Restricted Window in Development. Cell Rep.
473 2016;16(4):1039-54.
474
475 9. Bueno C, Ayllon V, Montes R, Navarro-Montero O, Ramos-Mejia V, Real PJ, et al. FLT3 activation
476 cooperates with MLL-AF4 fusion protein to abrogate the hematopoietic specification of human ESCs. Blood.
477 2013;121(19):3867-78, S1-3.
478

479 10.   Bueno C, Montes R, Melen GJ, Ramos-Mejia V, Real PJ, Ayllon V, et al. A human ESC model for
480 MLL-AF4 leukemic fusion gene reveals an impaired early hematopoietic-endothelial specification. Cell Res.
481 2012;22(6):986-1002.
482
483 11.   Bursen A, Schwabe K, Ruster B, Henschler R, Ruthardt M, Dingermann T, et al. The AF4.MLL fusion
484 protein is capable of inducing ALL in mice without requirement of MLL.AF4. Blood. 2010;115(17):3570-9.
485
486 12.   Chen W, Li Q, Hudson WA, Kumar A, Kirchhof N, Kersey JH. A murine Mll-AF4 knock-in model
487 results in lymphoid and myeloid deregulation and hematologic malignancy. Blood. 2006;108(2):669-77.
488
489 13.   Krivtsov AV, Feng Z, Lemieux ME, Faber J, Vempati S, Sinha AU, et al. H3K79 methylation profiles
490 define murine and human MLL-AF4 leukemias. Cancer Cell. 2008;14(5):355-68.
491
492 14.   Lin S, Luo RT, Ptasinska A, Kerry J, Assi SA, Wunderlich M, et al. Instructive Role of MLL-Fusion
493 Proteins Revealed by a Model of t(4;11) Pro-B Acute Lymphoblastic Leukemia. Cancer Cell. 2016;30(5):737-
494 49.
495
496 15.   Metzler M, Forster A, Pannell R, Arends MJ, Daser A, Lobato MN, et al. A conditional model of MLL-
497 AF4 B-cell tumourigenesis using invertor technology. Oncogene. 2006;25(22):3093-103.
498
499 16.   Montes R, Ayllon V, Gutierrez-Aranda I, Prat I, Hernandez-Lamas MC, Ponce L, et al. Enforced
500 expression of MLL-AF4 fusion in cord blood CD34+ cells enhances the hematopoietic repopulating cell
501 function and clonogenic potential but is not sufficient to initiate leukemia. Blood. 2011;117(18):4746-58.
502
503 17.   Tamai H, Miyake K, Takatori M, Miyake N, Yamaguchi H, Dan K, et al. Activated K-Ras protein
504 accelerates human MLL/AF4-induced leukemo-lymphomogenicity in a transgenic mouse model. Leukemia.
505 2011;25(5):888-91.
506
507 18.   Kowarz E, Burmeister T, Lo Nigro L, Jansen MW, Delabesse E, Klingebiel T, et al. Complex MLL
508 rearrangements in t(4;11) leukemia patients with absent AF4.MLL fusion allele. Leukemia. 2007;21(6):1232-
509 8.
510
511 19.   Agraz-Doblas A. Bueno C., Bashford-Rogers R., Anindita R., Schneider P., Bardini M., Ballerini P.,
512 Cazzaniga G., Moreno T., Revilla C., Gut M., De Lorenzo P., Valsecchi MG, Roberts I, Pieters R, Varela I.,
513 Menendez P., Stam R.W. Unravelling the cellular origin and clinical prognosis markers of infant B-cell acute
514 lymphoblastic leukemia using genome-wide analysis. Haematologica. 2019;in press.
515
516 20.   Trentin L, Giordan M, Dingermann T, Basso G, Te Kronnie G, Marschalek R. Two independent gene
517 signatures in pediatric t(4;11) acute lymphoblastic leukemia patients. Eur J Haematol. 2009;83(5):406-19.
518
519 21.   Kumar AR, Yao Q, Li Q, Sam TA, Kersey JH. t(4;11) leukemias display addiction to MLL-AF4 but not
520 to AF4-MLL. Leuk Res. 2011;35(3):305-9.
521
522 22.   Sanders DS, Muntean AG, Hess JL. Significance of AF4-MLL reciprocal fusion in t(4;11) leukemias?
523 Leuk Res. 2011;35(3):299-300.
524

523   23.   Prieto C, Marschalek R, Kuhn A, Bursen A, Bueno C, Menendez P. The AF4-MLL fusion transiently
526   augments multilineage hematopoietic engraftment but is not sufficient to initiate leukemia in cord blood
527   CD34(+) cells. Oncotarget. 2017;8(47):81936-41.

529   24.   Rego EM, Pandolfi PP. Reciprocal products of chromosomal translocations in human cancer
530   pathogenesis: key players or innocent bystanders? Trends Mol Med. 2002;8(8):396-405.

532   25.   Wilkinson AC, Ballabio E, Geng H, North P, Tapia M, Kerry J, et al. RUNX1 is a key target in t(4;11)
533   leukemias that contributes to gene activation through an AF4-MLL complex interaction. Cell Rep.
534   2013;3(1):116-27.

536   26.   Menendez P, Bueno C, Wang L. Human embryonic stem cells: A journey beyond cell replacement
537   therapies. Cytotherapy. 2006;8(6):530-41.

539   27.   Romero-Moya D, Santos-Ocana C, Castano J, Garrabou G, Rodriguez-Gomez JA, Ruiz-Bonilla V,
540   et al. Genetic Rescue of Mitochondrial and Skeletal Muscle Impairment in an Induced Pluripotent Stem Cells
541   Model of Coenzyme Q10 Deficiency. Stem Cells. 2017;35(7):1687-703.

543   28.   Menendez P, Vargas A, Bueno C, Barrena S, Almeida J, De Santiago M, et al. Quantitative analysis
544   of bcl-2 expression in normal and leukemic human B-cell differentiation. Leukemia. 2004;18(3):491-8.

546   29.   Ramos-Mejia V, Melen GJ, Sanchez L, Gutierrez-Aranda I, Ligero G, Cortes JL, et al. Nodal/Activin
547   signaling predicts human pluripotent stem cell lines prone to differentiate toward the hematopoietic lineage.
548   Mol Ther. 2010;18(12):2173-81.

550   30.   Wang L, Li L, Shojaei F, Levac K, Cerdan C, Menendez P, et al. Endothelial and hematopoietic cell
551   fate of human embryonic stem cells originates from primitive endothelium with hemangioblastic properties.
552   Immunity. 2004;21(1):31-41.

554   31.   Diehl F, Rossig L, Zeiher AM, Dimmeler S, Urbich C. The histone methyltransferase MLL is an
555   upstream regulator of endothelial-cell sprout formation. Blood. 2007;109(4):1472-8.

557   32.   Hatzipantelis ES, Athanassiou-Metaxa M, Gombakis N, Tzimouli V, Taparkou A, Sidi-Fragandrea V,
558   et al. Thrombomodulin and von Willebrand factor: relation to endothelial dysfunction and disease outcome in
559   children with acute lymphoblastic leukemia. Acta Haematol. 2011;125(3):130-5.

561   33.   Menendez P, Catalina P, Rodriguez R, Melen GJ, Bueno C, Arriero M, et al. Bone marrow
562   mesenchymal stem cells from infants with MLL-AF4+ acute leukemia harbor and express the MLL-AF4 fusion
563   gene. J Exp Med. 2009;206(13):3131-41.

565   34.   Ramos-Mejia V, Fernandez AF, Ayllon V, Real PJ, Bueno C, Anderson P, et al. Maintenance of
566   human embryonic stem cells in mesenchymal stem cell-conditioned media augments hematopoietic
567   specification. Stem Cells Dev. 2012;21(9):1549-58.

569   35.   Sanchez L, Gutierrez-Aranda I, Ligero G, Martin M, Ayllon V, Real PJ, et al. Maintenance of human
570   embryonic stem cells in media conditioned by human mesenchymal stem cells obviates the requirement of
571   exogenous basic fibroblast growth factor supplementation. Tissue Eng Part C Methods. 2012;18(5):387-96.

572

573 36.   Bueno C, Montes R, Martin L, Prat I, Hernandez MC, Orfao A, et al. NG2 antigen is expressed in
574 CD34+ HPCs and plasmacytoid dendritic cell precursors: is NG2 expression in leukemia dependent on the
575 target cell where leukemogenesis is triggered? Leukemia. 2008;22(8):1475-8.

576

577 37.   Castano J, Menendez P, Bruzos-Cidon C, Straccia M, Sousa A, Zabaleta L, et al. Fast and efficient
578 neural conversion of human hematopoietic cells. Stem Cell Reports. 2014;3(6):1118-31.

579

580 38.   Munoz-Lopez A, Romero-Moya D, Prieto C, Ramos-Mejia V, Agraz-Doblas A, Varela I, et al.
581 Development Refractoriness of MLL-Rearranged Human B Cell Acute Leukemias to Reprogramming into
582 Pluripotency. Stem Cell Reports. 2016;7(4):602-18.

583

584 39.   Bueno C, Sardina JL, Di Stefano B, Romero-Moya D, Munoz-Lopez A, Ariza L, et al. Reprogramming
585 human B cells into induced pluripotent stem cells and its enhancement by C/EBPalpha. Leukemia.
586 2016;30(3):674-82.

587

588 40.   Giorgetti A, Castano J, Bueno C, Diaz de la Guardia R, Delgado M, Bigas A, et al. Proinflammatory
589 signals are insufficient to drive definitive hematopoietic specification of human HSCs in vitro. Exp Hematol.
590 2017;45:85-93 e2.

591

592 41.   Menendez P, Wang L, Chadwick K, Li L, Bhatia M. Retroviral transduction of hematopoietic cells
593 differentiated from human embryonic stem cell-derived CD45(neg)PFV hemogenic precursors. Mol Ther.
594 2004;10(6):1109-20.

595

596 42.   Bueno C, Roldan M, Anguita E, Romero-Moya D, Martin-Antonio B, Rosu-Myles M, et al. Bone
597 marrow mesenchymal stem cells from patients with aplastic anemia maintain functional and immune
598 properties and do not contribute to the pathogenesis of the disease. Haematologica. 2014;99(7):1168-75.

599

600 43.   Ramos-Mejia V, Navarro-Montero O, Ayllon V, Bueno C, Romero T, Real PJ, et al. HOXA9 promotes
601 hematopoietic commitment of human embryonic stem cells. Blood. 2014;124(20):3065-75.

602

603 44.   Toscano MG, Navarro-Montero O, Ayllon V, Ramos-Mejia V, Guerrero-Carreno X, Bueno C, et al.
604 SCL/TAL1-mediated transcriptional network enhances megakaryocytic specification of human embryonic
605 stem cells. Mol Ther. 2015;23(1):158-70.

606

607 45.   Bueno C, Montes R, Menendez P. The ROCK inhibitor Y-27632 negatively affects the
608 expansion/survival of both fresh and cryopreserved cord blood-derived CD34+ hematopoietic progenitor
609 cells: Y-27632 negatively affects the expansion/survival of CD34+HSPCs. Stem Cell Rev. 2010;6(2):215-23.

610

611 46.   Rubio R, Garcia-Castro J, Gutierrez-Aranda I, Paramio J, Santos M, Catalina P, et al. Deficiency in
612 p53 but not retinoblastoma induces the transformation of mesenchymal stem cells in vitro and initiates
613 leiomyosarcoma in vivo. Cancer Res. 2010;70(10):4185-94.

614

615 47.   Ayllon V, Bueno C, Ramos-Mejia V, Navarro-Montero O, Prieto C, Real PJ, et al. The Notch ligand
616 DLL4 specifically marks human hematoendothelial progenitors and regulates their hematopoietic fate.
617 Leukemia. 2015;29(8):1741-53.

618

619 48. Vodyanik MA, Slukvin, II. Hematoendothelial differentiation of human embryonic stem cells. Curr
620 Protoc Cell Biol. 2007;Chapter 23:Unit 23 6.

622 49. Bueno C, Montes R, de la Cueva T, Gutierrez-Aranda I, Menendez P. Intra-bone marrow
623 transplantation of human CD34(+) cells into NOD/LtSz-scid IL-2rgamma(null) mice permits multilineage
624 engraftment without previous irradiation. Cytotherapy. 2010;12(1):45-9.

626 50. Gutierrez-Aranda I, Ramos-Mejia V, Bueno C, Munoz-Lopez M, Real PJ, Macia A, et al. Human
627 induced pluripotent stem cells develop teratoma more efficiently and faster than human embryonic stem cells
628 regardless the site of injection. Stem Cells. 2010;28(9):1568-70.

630 51. Vodyanik MA, Bork JA, Thomson JA, Slukvin, II. Human embryonic stem cell-derived CD34+ cells:
631 efficient production in the coculture with OP9 stromal cells and analysis of lymphohematopoietic potential.
632 Blood. 2005;105(2):617-26.

634 52. Prieto C, Stam RW, Agraz-Doblas A, Ballerini P, Camos M, Castano J, et al. Activated KRAS
635 Cooperates with MLL-AF4 to Promote Extramedullary Engraftment and Migration of Cord Blood CD34+
636 HSPC But Is Insufficient to Initiate Leukemia. Cancer Res. 2016;76(8):2478-89.

638 53. Deshpande AJ, Deshpande A, Sinha AU, Chen L, Chang J, Cihan A, et al. AF10 regulates
639 progressive H3K79 methylation and HOX gene expression in diverse AML subtypes. Cancer Cell.
640 2014;26(6):896-908.

642 54. Guenther MG, Lawton LN, Rozovskaia T, Frampton GM, Levine SS, Volkert TL, et al. Aberrant
643 chromatin at genes encoding stem cell regulators in human mixed-lineage leukemia. Genes Dev.
644 2008;22(24):3403-8.

646 55. Andersson AK, Ma J, Wang J, Chen X, Gedman AL, Dang J, et al. The landscape of somatic
647 mutations in infant MLL-rearranged acute lymphoblastic leukemias. Nat Genet. 2015;47(4):330-7.

649 56. Malouf C, Ottersbach K. The fetal liver lymphoid-primed multipotent progenitor provides the
650 prerequisites for the initiation of t(4;11) MLL-AF4 infant leukemia. Haematologica. 2018.

652 57. Peters DG, Klucher KM, Perlingeiro RC, Dessain SK, Koh EY, Daley GQ. Autocrine and paracrine
653 effects of an ES-cell derived, BCR/ABL-transformed hematopoietic cell line that induces leukemia in mice.
654 Oncogene. 2001;20(21):2636-46.

656 58. Ji J, Risueno RM, Hong S, Allan D, Rosten P, Humphries K, et al. Brief report: ectopic expression of
657 NUP98-HOXA10 augments erythroid differentiation of human embryonic stem cells. Stem Cells.
658 2011;29(4):736-41.

660 59. Tan YT, Ye L, Xie F, Beyer AI, Muench MO, Wang J, et al. Respecifying human iPSC-derived blood
661 cells into highly engraftable hematopoietic stem and progenitor cells with a single factor. Proc Natl Acad Sci
662 U S A. 2018;115(9):2180-5.

664 60. Marschalek R. Mechanisms of leukemogenesis by MLL fusion proteins. Br J Haematol.
665 2011;152(2):141-54.

666

667   61.   Boisset JC, van Cappellen W, Andrieu-Soler C, Galjart N, Dzierzak E, Robin C. In vivo imaging of
668   haematopoietic cells emerging from the mouse aortic endothelium. Nature. 2010;464(7285):116-20.
669

670   62.   Ahmad K, Scholz B, Capelo R, Schweighofer I, Kahnt AS, Marschalek R, et al. AF4 and AF4-MLL
671   mediate transcriptional elongation of 5-lipoxygenase mRNA by 1, 25-dihydroxyvitamin D3. Oncotarget.
672   2015;6(28):25784-800.
673

674   63.   Benedikt A, Baltruschat S, Scholz B, Bursen A, Arrey TN, Meyer B, et al. The leukemogenic AF4-
675   MLL fusion protein causes P-TEFb kinase activation and altered epigenetic signatures. Leukemia.
676   2011;25(1):135-44.
677

678   64.   Muck F, Bracharz S, Marschalek R. DDX6 transfers P-TEFb kinase to the AF4/AF4N (AFF1) super
679   elongation complex. Am J Blood Res. 2016;6(3):28-45.
680

681   65.   Driessen EM, van Roon EH, Spijkers-Hagelstein JA, Schneider P, de Lorenzo P, Valsecchi MG, et
682   al. Frequencies and prognostic impact of RAS mutations in MLL-rearranged acute lymphoblastic leukemia in
683   infants. Haematologica. 2013;98(6):937-44.
684

685   66.   Kuhn A, Loscher D, Marschalek R. The IRX1/HOXA connection: insights into a novel t(4;11)- specific
686   cancer mechanism. Oncotarget. 2016;7(23):35341-52.
687

688

689

FIGURE LEGENDS

691 **Figure 1. Characterization of transgenic human ESCs expressing the reciprocal fusion A4M together**

692 **with MA4.** (**A**) RNA-seq and qRT-PCR validation revealed that ~45% (11/25) of the patients with t(4;11)+ B-

693 ALL do not express the reciprocal fusion A4M[18]. (**B**) *Left*, Phase-contrast morphology of representative

694 colonies from each transgenic hESC line. *Right*, RT-PCR confirming expression of both fusions in

695 undifferentiated hESCs. (**C**) qRT-PCR expression of the pluripotency genes *OCT4, SOX2, NANOG,*

696 *CRIPTO*, and *DNMT3B*. (**D**) Representative FACS data confirming expression of the pluripotency surface

697 makers SSEA-3, SSEA-4, TRA-1-60, and TRA-1-81.

698

699 **Figure 2. A4M cooperates with MA4 to accelerate human ESC/EB specification towards HEPs and**

700 **subsequent hematopoietic differentiation.** (**A**) Schematic of EB hematopoietic differentiation of hESCs

701 and end-point analyses. (**B**) *Upper left,* specification into HEPs (CD31$^+$CD34$^+$CD45$^-$) is accelerated in double

702 fusion-expressing hESCs. Subsequent differentiation of HEPs into hematopoietic progenitors (*upper right*)

703 and mature CD45$^+$ blood cells (*bottom left*) is enhanced in double fusion-expressing HEPs. *Bottom right*, CFU

704 read-out and scoring (pie charts) confirming an accelerated and enhanced hematopoietic progenitor potential

705 from double fusion-expressing blood derivatives. (**C**) RT-PCR confirming stable expression of MA4 and A4M

706 upon EB differentiation. (**D**) Neither MA4- nor double fusion-expressing blood derivatives display *in vivo*

707 hematopoietic engraftment potential in irradiated NSG mice. Data are presented as mean±SEM from at least

708 three independent experiments. *p<0.05.

709

710 **Figure 3. Co-expression of MA4 and A4M enhances hematopoietic differentiation of human ESCs in**

711 **OP9 co-culture.** (**A**) Experimental design of OP9-based hESC differentiation towards HEPs and further

712 hematopoietic commitment of HEPs maintained in either liquid culture for 16 days or in MS5 co-culture for

713 30 days. (**B**) Frequency of total CD45+ blood cells after 9 days in OP9 co-culture. (**C,D**) CD45-CD31+CD34+

714 HEPs were FACS-purified at day 9 of OP9 co-culture and allowed to differentiate into CD45+ cells in liquid

715 culture (**C**) or in MS5 co-culture (**D**). Data represent mean±SEM from independent experiments.

716

717 **Figure 4. Enhanced endothelial cell fate from HEPs co-expressing MA4 and A4M.** (**A**) Scheme of HEPs

718 endothelial differentiation and phenotypic characterization. (**B**) FACS-sorted HEPs from day 9 human ESC-

719 OP9 co-cultures were cultured in EGM2 medium for 5 days and analyzed by immunofluorescence for VE-

720 cadherin, eNOS and vWF. (**C**) *Top*, Endothelial-like structures were identified and quantified based on VE-

721 cadherin staining (white dotted-lined areas in B, *top panel*). *Bottom*, Frequency of CD45$^-$CD31$^+$CD144$^+$

722 endothelial cells quantified by flow cytometry. (**D**) *In vivo* endothelial engraftment potential

723 (HLA.ABC$^+$CD31$^+$CD144$^+$CD45$^-$) analyzed in bone marrow of NSG mice 8 weeks after transplantation of

724 HEPs. Data presented as mean±SEM from 5 independent experiments. *$p<0.05$.

725

726 **Figure 5. Co-expression of MA4 and A4M significantly enhances the emergence of both endothelial**

727 **and hemogenic HEPs.** (**A**) Representative flow cytometry analysis of HEPs with hemogenic (CD45$^-$

728 CD31$^+$CD43$^+$CD34$^{dim/+}$) and endothelial (CD45$^-$CD31$^+$CD43$^-$CD34^{++}) potential. (**B,C**) A4M cooperates with

729 MA4 to boost the emergence of both endothelial (**B**) and hemogenic (**C**) HEPs. Data presented as

730 mean±SEM from 3 independent experiments. (**D**) Expression of *RUNX1c* and *Ve-Cad* in hemogenic and

731 endothelial HEPs. * $p<0.05$

732

733 **Figure 6: Transcriptional transition towards a hematopoietic/endothelial gene signature in double**

734 **fusion-expressing HEPs.** (**A**) Heatmap representation of hierarchical clustering of genes differentially

735 expressed between EV-, single fusions- and double fusion-expressing HEPs. Each column represents a

736 technical replicate from three independent experiments. (**B,C**) Statistically significant functional categories

737 (**B**) and cancer/leukemia-associated biofunctions (**C**) identified using IPA on genes differentially expressed

738   in single fusions-, and double fusion-expressing HEPs relative to EV. They are ranked by z-score. Functional

739   categories associated with "hematological system development and function" and "cardiovascular system

740   development" are shown in bold. All significant biofunctions are associated with blood cell differentiation,

741   homeostasis and migration/movement.

742

743   **Figure 7: H3K79 methylation profiles at genomic loci of MLL targets in MA4-, A4M- and double fusion-**

744   **expressing human ESC-derived blood derivatives.** (**A**) GO enrichment of differential H3K79me3 peaks

745   specific for double fusion-expressing cells. (**B-C**) Representative profiles for ChIP-seq using anti-H3K79me3

746   antibody at genomic regions of typical non-HOXA (**B**) and *HOXA* MLL targets (**C**).

747

748

749

750

751

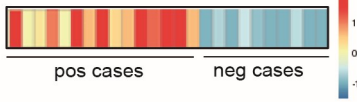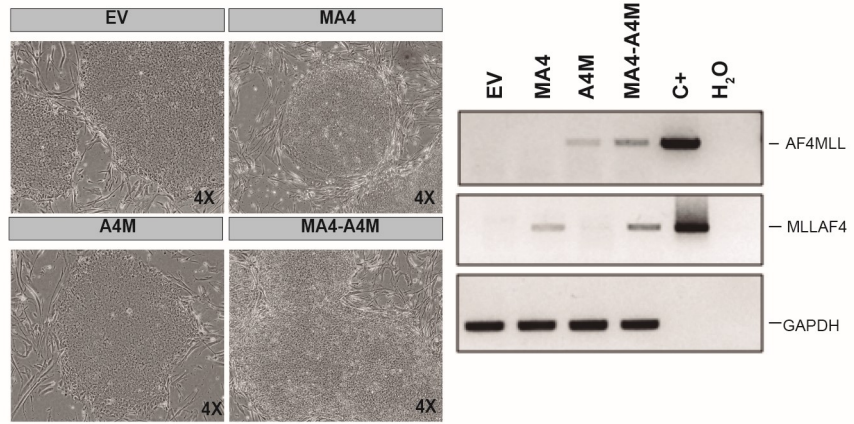752

753

754

755

756

757

758

759

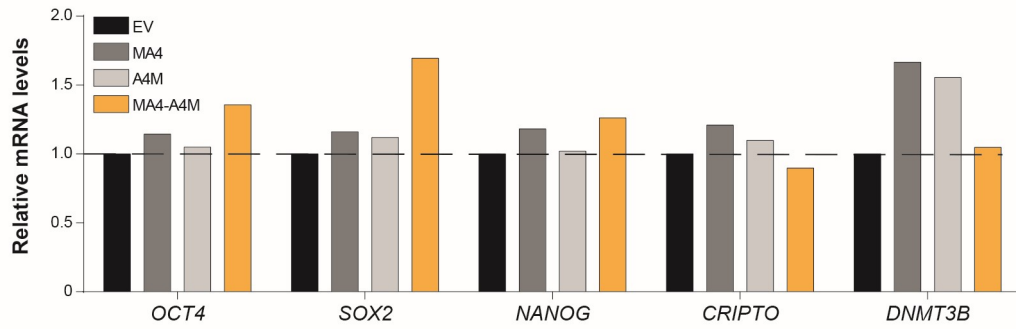760

761

A

Infant BCP-ALL
t(4;11)/MA4⁺

RNA-Seq A4M expression
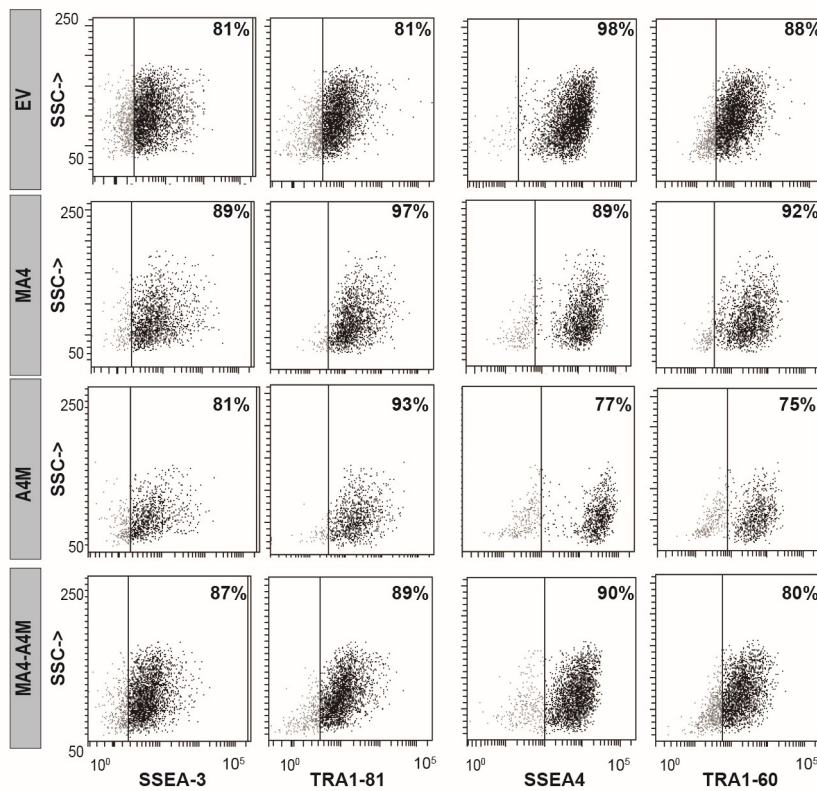
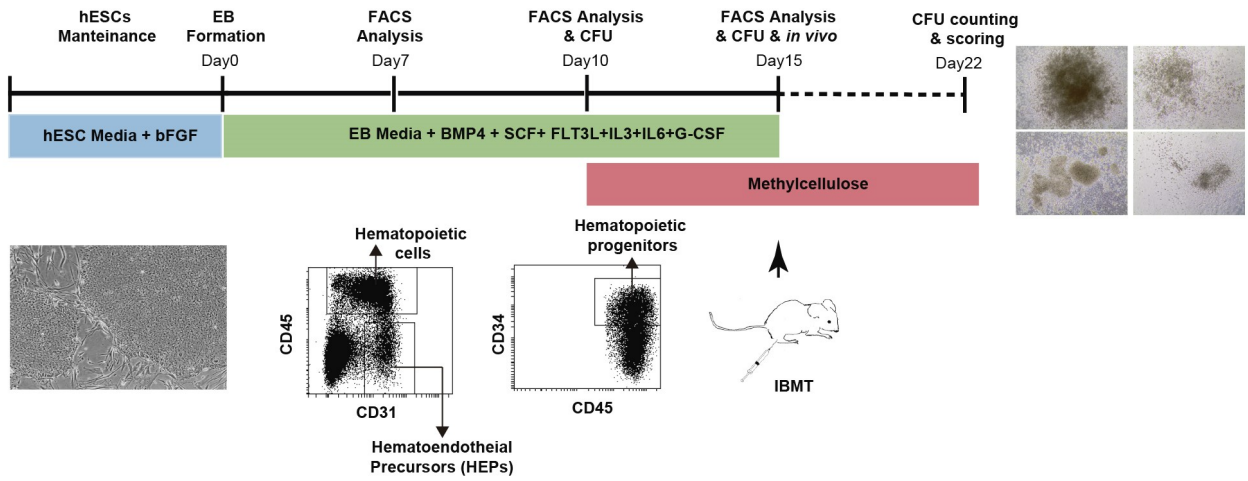pos cases   neg cases

█ % patients A4M⁺
█ % patients A4M⁻

B

C

D

A

B

C

D

HUMAN HEMATOPOIETIC ENGRAFTMENT

Bueno C *et al.* Figure 3

A

Human ESCs Manteinance

hESC Media + bFGF

Day0

OP9 Co-culture + Ascorbic Ac. + MTG

Day5

OP9 CO-CULTURE

Day9

FACS ANALYSIS & HEPs SORTING

Gate on CD29⁻ (human cells)

CD45

CD34

CD45⁺ hematopoietic cells

CD34⁺CD31⁺CD45 (HEPs)

LIQUID CULTURE
STEMSPAM MEDIA
+ SCF+FLT3L+IL6+IL3+G-CSF

Day16

FACS ANALYSIS

CD45

CD34

CD45⁺ hematopoietic cells

MS5 CO-CULTURE
MS5 CO-CULTURE IL7+SCF+IL3+FLT3L

Day30

FACS ANALYSIS

Gate on CD29 (human cells)

SSC->

CD45

CD45⁺ hematopoietic cells

B

% Total hematopoietic cells (CD45⁺ cells)

OP9 CO-CULTURE

***

EV
MA4
A4M
MA4-A4M

C

% Total hematopoietic cells (CD45⁺ cells)

LIQUID CULTURE

*
**
*

EV
MA4
A4M
MA4-A4M

D

% Total hematopoietic cells (CD45⁺ cells)

MS5 CO-CULTURE

**

EV
MA4
A4M
MA4-A4M

A

hESCs Manteinance

Day0 Day5

FACS HEPs SORTING
Day9

FACS ANALISIS & IMMUNOFLUORESCENCE
Day13

hESC Media + bFGF

OP9 Coculture

EGM2 MEDIA

Gate on CD29⁻ (human cells)

CD45 / CD34

CD34⁺CD31⁺CD45⁻ (HEPs)

SSC-> / CD144

B

EV   MA4   A4M   MA4-A4M

VE-CADHERIN

eNOS DAPI

VE-CADHERIN

vWF DAPI

C

N. Endothelial-like structure

*   *   *

EV
MA4
A4M
MA4-A4M

%CD31+CD144+ Cells

5-fold

5-fold

D

HUMAN ENDOTHELIAL ENGRAFTMENT

huCD45 / hu HLA ABC

huCD31 / huCD45

huCD31 / huCD144

% HLA ABC⁺ CD31⁺ CD144⁺ CD45⁻ cells

3-4-fold

EV
MLLAF4
AF4MLL
MLLAF4-AF4MLL
CB-derived CD34⁺ Cells

φ

A

B

Functional Categories

Categories (top to bottom):
- Cellular Movement
- **Hematological System Development and Function**
- Cell-To-Cell Signaling and Interaction
- Inflammatory Response
- Organismal Functions
- <u>Cancer</u>
- Organismal Injury and Abnormalities
- **Immune Cell Trafficking**
- Connective Tissue Disorders
- Developmental Disorder
- <u>**Hematological Disease**</u>
- Hereditary Disorder
- **Tissue Morphology**
- Cellular Function and Maintenance
- Tissue Development
- Organismal Survival
- **Cardiovascular Disease**
- **Cardiovascular System Development and Function**

x-axis: -log (p-value)

Legend: MA4, A4M, MA4-A4M

C

Cancer Functions

	MA4	A4M	MA4-A4M
Lymphatic system tumor	-0.313		2.025
Lymphatic cancer	-0.313		2.025
Lymphocytic neoplasm	-0.313		2.025
Lymphoid cancer	-0.313		2.025
Neoplasia of leukocytes	-0.313		2.025
Hematologic cancer of cells	-0.313		2.016
Lymphoma	-0.313		1.977
Hematologic cancer	-0.537		1.886
Lymphoreticular neoplasm	-0.014		1.854
Neoplasia of cells	-1.116		1.498
Hematopoietic neoplasm	-0.788		1.472
Carcinoma	-1.526	1.464	-0.445
Non-hematological neoplasm	-1.19	1.662	-0.45
Nonhematologic malignantneoplasm	-0.859	1.714	-0.61
Digestive organ tumor	-1.541	-1.231	-1.026

Haematological Disease Functions

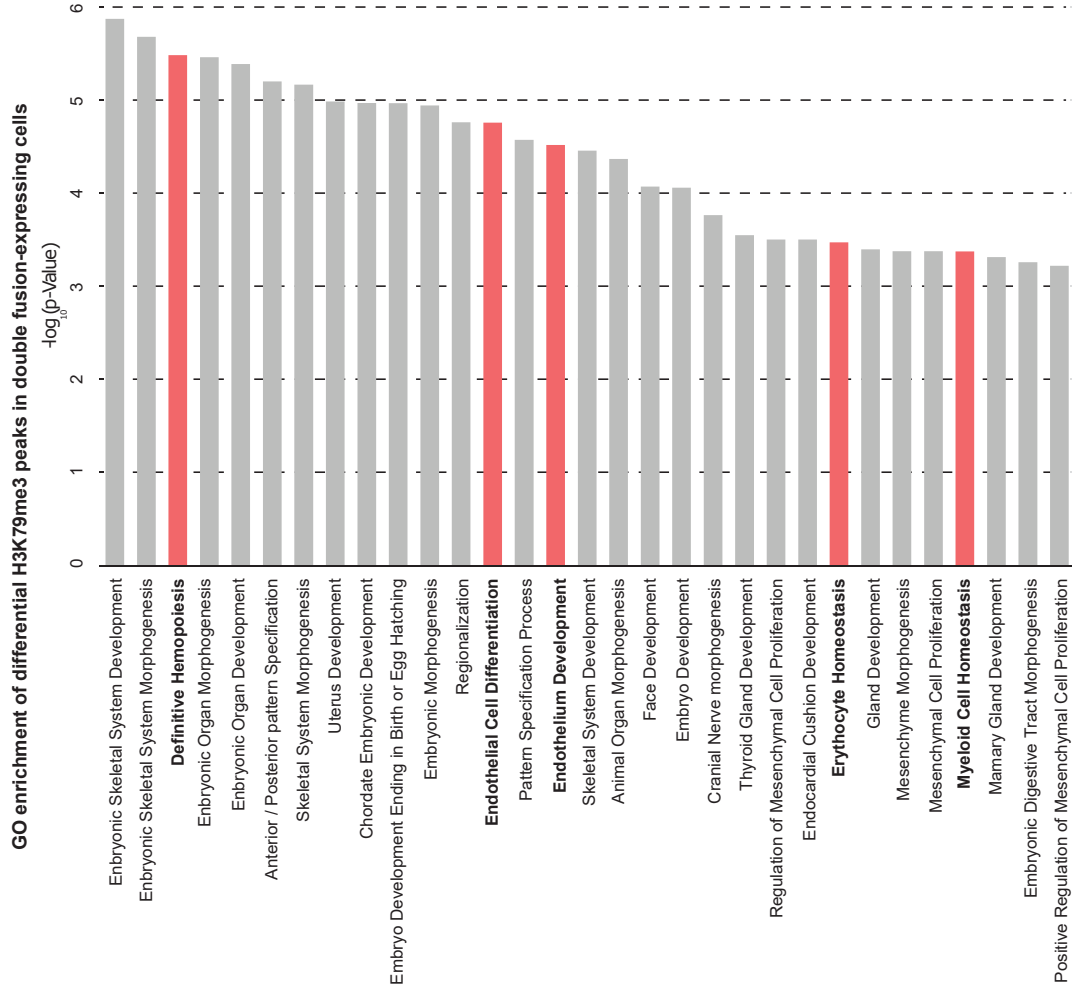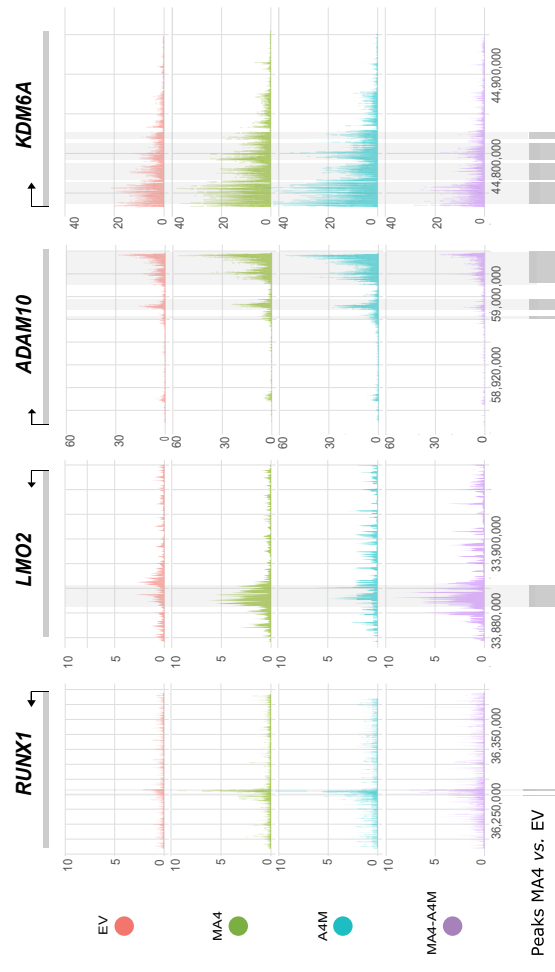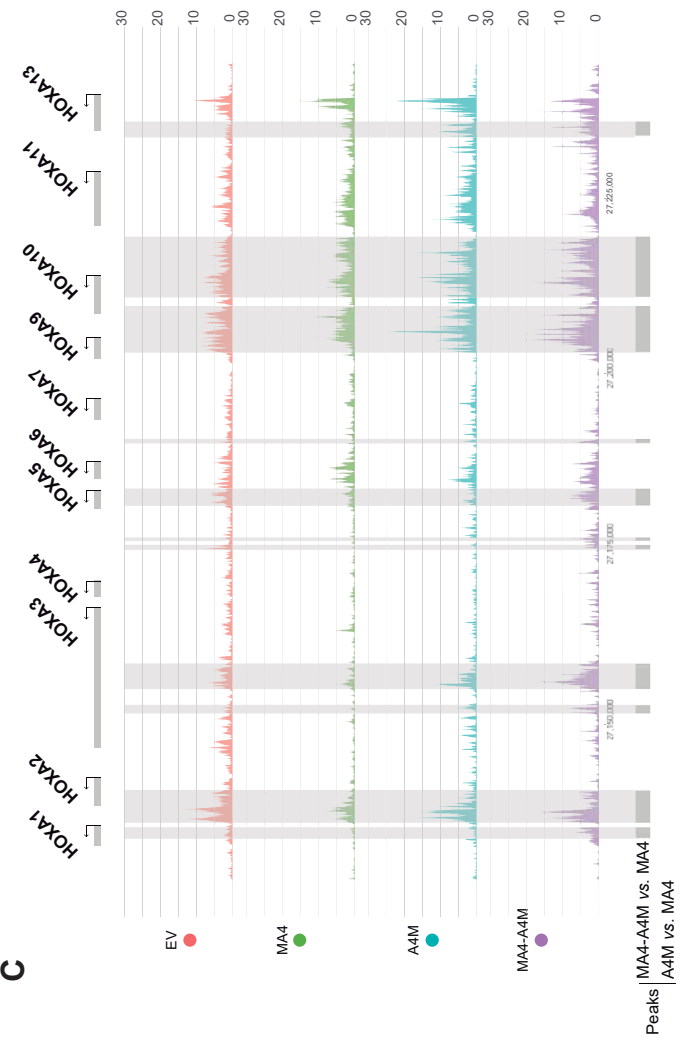|	MA4	A4M	MA4-A4M
Thrombus	1.714		0.23
Formation of thrombus	1.067		-0.277
Hematologic cancer of cells	-0.313		2.016
Lymphocytic cancer	-0.313		2.025
Lymphocytic neoplasm	-0.313		2.025
Lymphoid cancer	-0.313		2.025
Lymphoma	-0.313		1.977
Lymphoproliferative disorder	-0.313		2.025
Neoplasia of leukocyte	-0.313		2.025
Hematologic cancer	-0.537		1.886
Hematopoietic neoplasm	-0.788		1.472
Cytopenia	-2.66		
Thrombocytopenia	-2.831		
Hemorrhagic disease	-3.137		

Bueno C *et al.* Figure 7

SUPPLEMENTARY LEGENDS AND METHODS

Chromatin immunoprecipitation Sequencing

One million day 15 hematopoietic differentiating hEB-derived cells were subjected to Chromatin Immunoprecipitation (ChIP) as described [1]. Briefly, formaldehyde crosslinked cell extracts were sonicated using a Bioruptor (Diagenode) with a 0.5-min interval protocol to obtain DNA fragments of 200–500 bp. The chromatin fraction was divided into three parts and incubated overnight with 2 μg of anti-H3K4me3, anti-H3K79me3, or anti-H3 antibodies (Abcam) in RIPA buffer, and precipitated with protein A/G-Sepharose (Amersham). Cross-linkage of the co-precipitated DNA-protein complexes was reversed, and the DNA was sequenced at the CRG-CNAG Genomics Facility. Paired-end 50 bp ChIP-seq data reads were generated using a HiSeq 2500 Illumina sequencer. Alignment and peak detection was performed using the ENCODE (phase-3) histone ChIP-seq pipeline specifications. Reads were aligned to the human reference genome (assembly hg19) using BWA[2], removing all reads with a quality score <30. Peaks were called using MACS2[3] and whole cell extract samples as background samples. Differential analysis of histone mark peaks between experimental groups and the EV control group was performed using DiffBind[4], with a False Discovery Rate (FDR) cut-off of 0.1. Peaks annotation and Gene Ontology (GO) enrichment were analyzed with ChIP-Enrich package in R[4], using an FDR threshold of 0.1. ChIP-Seq data is deposited and available at Gene Expression Omnibus (GEO, accession number GSE111263).

Low-input RNA-seq and bulk RNA-seq on t(4;11)+B-ALL patients

For low-input RNAseq, 50 HEPs were FACS-purified from day 15 EBs into individual wells of a 96-well PCR plate containing lysis buffer (0.2% Triton-X100 and 2.3 U of SUPERase-In RNase Inhibitor; Ambion). Three independent replicates (50 cells/replicate) per condition were analyzed. RNA/cDNA was obtained and amplified as per the SMARTSEQ2 protocol[5,6]. Libraries were prepared for

1

sequencing using the Illumina Nextera XT DNA preparation kit. RNA-seq data was generated using a 50 single-end sequencing protocol. Pooled libraries were sequenced on the Illumina HiSeq 4000 platform. Reads were mapped to the human genome (Ensembl 81) and the ERCC sequences using STAR (version 2.4.2a) with default parameters. HTseq-count[7] was used to count the number of reads mapped to each gene with -s no (non-strand specific mode). Data were normalized for sequencing depth using the size factor from the DESeq2 package (version 1.12.2)[8]. Low-input RNA-seq data is deposited and available at GEO (GSE118947).

Bulk RNA-seq data on t(4;11)+ B-ALL patients was generated on the Illumina HiSeq 2500 platform using a 76 paired-end sequencing protocol and aligned to the human reference genome (hg19) using Tophat[19]. A gene was considered differentially expressed in MA4-A4M+ HEPs when it was >2-fold regulated (p<0.01) compared with the equivalent gene in EV-HEPs. Analysis of functional categories and biofunctions was performed with the genes differentially expressed using Ingenuity Pathway Analysis software (Ingenuity Inc., Redwood City, CA). Potential read pairs supporting the expression of MA4 and A4M fusion transcripts in patient samples were identified manually and with in-house written scripts as pairs in which each read aligned unambiguously to a different gene and which were compatible with the genomic breakpoints that were determined at the DNA level in all the samples. Additionally, RT-PCR with primers corresponding to *AF4* exon3 and *MLL* exon 12 (exons contained in all predicted fusions) was performed in all the samples to confirm the expression of any potential isoform of the reciprocal fusion gene. The identity of the isoform was confirmed by capillary sequencing. Finally, the expression of each fusion gene was determined by qRT-PCR using primer pairs spanning the breakpoint between genes. RNA-seq data is deposited and available at GEO (accession number GSE111263).

2

SUPPLEMENTARY REFERENCES

1.   Guiu J, Bergen DJ, De Pater E, et al. Identification of Cdca7 as a novel Notch transcriptional target involved in hematopoietic stem cell emergence. *J Exp Med*. 2014;211(12):2411-2423.

2.   Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25(14):1754-1760.

3.   Zhang Y, Liu T, Meyer CA, et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol*. 2008;9(9):R137.

4.   Welch RP, Lee C, Imbriano PM, et al. ChIP-Enrich: gene set enrichment testing for ChIP-seq data. *Nucleic Acids Res*. 2014;42(13):e105.

5.   Picelli S, Bjorklund AK, Faridani OR, Sagasser S, Winberg G, Sandberg R. Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat Methods*. 2013;10(11):1096-1098.

6.   Picelli S, Faridani OR, Bjorklund AK, Winberg G, Sagasser S, Sandberg R. Full-length RNA-seq from single cells using Smart-seq2. *Nat Protoc*. 2014;9(1):171-181.

7.   Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics*. 2015;31(2):166-169.

8.   Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550.

3

SUPPLEMENTARY LEGENDS

Figure S1. Coding sequence of the A4M fusion used in this study. (A) The full-length cDNA was sequenced-verified. Fusion breakpoint is located in AF4 exon 2 and MLL exon 10. (**B**) Schematic of lentivectors used in this study.

Figure S2. Co-expression of MA4 and A4M promotes specification rather than survival/proliferation of HEPs. (**A**) Day 20 MA4-expressing hEBs display a significant higher HEP production coupled to an impaired blood (CD45+ and CD45+CD34+) formation. (**B**) *Left*, representative FACS analysis of the identification of HEPs and CD45$^+$ hematopoietic cells and their cell-cycle distribution. *Right*, frequency of proliferating (S+G$_2$M phases) CD45+ blood cells and HEPs. (**C**) *Left*, representative FACS analysis of apoptotic CD45$^+$ hematopoietic cells and HEPs. *Right*, frequency of apoptotic (Annexin V+) HEPs and CD45+ hematopoietic cells for the indicated genotypes.

Figure S3. H3K79me3 ChIP-seq profiles for the different genotypes. (**A**) Correlation heatmap for both H3K79me3 and H3K4me3 using affinity scores based on read counts in consensus peaks distinguishes between the four experimental groups. (**B**) Venn diagrams depicting the number of differentially enriched genomic regions for H3K79me3 between experimental groups relative to EV. (**C**) Quantitative expression of the indicated *HOX-A* genes in CFUs from day 15 EV- and double-fusion-expressing HEPs.

Figure S4. H3K4me2 profiles at genomic loci of MLL targets identified by Guenther *et al*. Representative profiles for ChIP-seq using an anti-H3K4me^3 antibody at genomic sites of HOX-A (**A**) and non-HOX (**B**) MLL targets in the indicated genotypes of human ESC-differentiating derivatives.

4

A

AF4-MLL (8739 bp)

Red: AF4 exon 1-3

Blue: MLL exon 10 to 36

ATG :Start Codon

TAA :Stop Codon

ATGGCAGCCCAGTCAAGTTTGTACAATGACGACAGAAACCTGCTTCGAATTAGAGAGAAGGAA
AGACGCAACCAGGAAGCCCACCAAGAGAAAGAGGCATTTCCTGAAAAGATTCCCCTTTTTGGA
GAGCCCTACAAGACAGCAAAAGGTGATGAGCTGTCTAGTCGAATACAGAACATGTTGGGAAAC
TACGAAGAAGTGAAGGAGTTCCTTAGTACTAAGTCTCACACTCATCGCCTGGATGCTTCTGAAA
ATAGGTTGGGAAAGCCGAAATATCCTTTAATTCCTGACAAAGGGAGCAGCATTCCATCCAGCT
CCTTCCACACTAGTGTCCACCACCAGTCCATTCACACTCCTGCGTCTGGACCACTTTCTGTTGG
CAACATTAGCCACAATCCAAAGATGGCGCAGCCAAGAACTGAACCAATGCCAAGTCTCCATGC
CAAAAGCTGCGGCCCACCGGACAGCCAGCACCTGACCCAGGATCGCCTTGGTCAGGAGGGG
TTCGGCTCTAGTCATCACAAGAAAGGTGACCGAAGAGCTGACGGAGACCACTGTGCTTCGGT
GACAGATTCGGCTCCAGAGAGGGAGCTTTCTCCCTTAATCTCTTTGCCTTCCCCAGTTCCCCC
TTTGTCACCTATACATTCCAACCAGCAAACTCTTCCCCGGACGCAAGGAAGCAGCAAGGTTCA
TGGCAGCAGCAATAACAGTAAAGGCTATTGCCCAGCCAAATCTCCCAAGGACCTAGCAGTGAA
AGTCCATGATAAAGAGACCCCTCAAGACAGTTTGGTGGCCCCTGCCCAGCCGCCTTCTCAGAC
ATTTCCACCTCCCTCCCTCCCCTCAAAAAGTGTTGCAATGCAGCAGAAGCCCACGGCTTATGT
CCGGCCCATGGATGGTCAAGATCAGGCCCCTAGTGAATCCCCTGAACTGAAACCACTGCCGG
AGGACTATCGACAGCAGACCTTTGAAAAAACAGACTTGAAAGTGCCTGCCAAAGCCAAGCTCA
CCAAACTGAAGATGCCTTCTCAGTCAGTTGAGGAGGATTGTGAAGCAGAAATGTGTGGGAGA
TGGGAGGCTTAGGAATCTTGACTTCTGTTCCTATAACACCCAGGGTGGTTTGCTTTCTCTGTGC
CAGTAGTGGGCATGTAGAGTTTGTGTATTGCCAAGTCTGTTGTGAGCCCTTCCACAAGTTTTGT
TTAGAGGAGAACGAGCGCCCTCTGGAGGACCAGCTGGAAAATTGGTGTTGTCGTCGTTGCAA
ATTCTGTCACGTTTGTGGAAGGCAACATCAGGCTACAAAGCAGCTGCTGGAGTGTAATAAGTG
CCGAAACAGCTATCACCCTGAGTGCCTGGGACCAAACTACCCCACCAAACCCACAAAGAAGAA
GAAAGTCTGGATCTGTACCAAGTGTGTTCGCTGTAAGAGCTGTGGATCCACAACTCCAGGCAA
AGGGTGGGATGCACAGTGGTCTCATGATTTCTCACTGTGTCATGATTGCGCCAAGCTCTTTGC
TAAAGGAAACTTCTGCCCTCTCTGTGACAAATGTTATGATGATGATGACTATGAGAGTAAGATG
ATGCAATGTGGAAAGTGTGATCGCTGGGTCCATTCCAAATGTGAGAATCTTTCAGGTACAGAA
GATGAGATGTATGAGATTCTATCTAATCTGCCAGAAAGTGTGGCCTACACTTGTGTGAACTGTA
CTGAGCGGCACCCTGCAGAGTGGCGACTGGCCCTTGAAAAAGAGCTGCAGATTTCTCTGAAG
CAAGTTCTGACAGCTTTGTTGAATTCTCGGACTACCAGCCATTTGCTACGCTACCGGCAGGCT
GCCAAGCCTCCAGACTTAAATCCCGAGACAGAGGAGAGTATACCTTCCCGCAGCTCCCCCGA
AGGACCTGATCCACCAGTTCTTACTGAGGTCAGCAAACAGGATGATCAGCAGCCTTTAGATCT
AGAAGGAGTCAAGAGGAAGATGGACCAAGGGAATTACACATCTGTGTTGGAGTTCAGTGATGA
TATTGTGAAGATCATTCAAGCAGCCATTAATTCAGATGGAGGACAGCCAGAAATTAAAAAAGCC
AACAGCATGGTCAAGTCCTTCTTCATTCGGCAAATGGAACGTGTTTTTCCATGGTTCAGTGTCA
AAAAGTCCAGGTTTTGGGAGCCAAATAAAGTATCAAGCAACAGTGGGATGTTACCAAACGCAG
TGCTTCCACCTTCACTTGACCATAATTATGCTCAGTGGCAGGAGCGAGAGGAAAACAGCCACA
CTGAGCAGCCTCCTTTAATGAAGAAAATCATTCCAGCTCCCAAACCCAAAGGTCCTGGAGAAC
CAGACTCACCAACTCCTCTGCATCCTCCTACACCACCAATTTTGAGTACTGATAGGAGTCGAGA
AGACAGTCCAGAGCTGAACCCACCCCCAGGCATAGAAGACAATAGACAGTGTGCGTTATGTTT
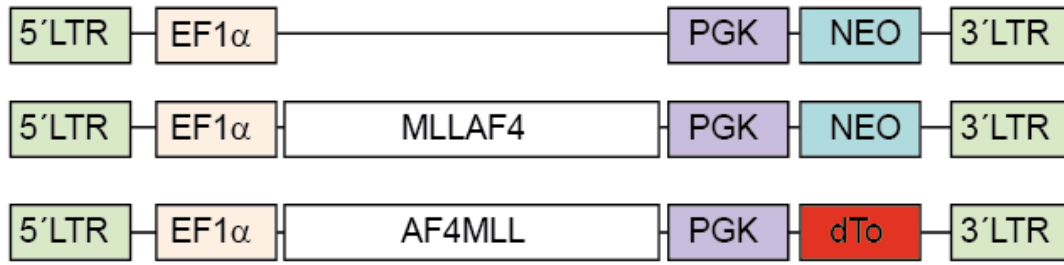GACTTATGGTGATGACAGTGCTAATGATGCTGGTCGTTTACTATATATTGGCCAAAATGAGTGG

```
ACACATGTAAATTGTGCTTTGTGGTCAGCGGAAGTGTTTGAAGATGATGACGGATCACTAAAGA
ATGTGCATATGGCTGTGATCAGGGGCAAGCAGCTGAGATGTGAATTCTGCCAAAAGCCAGGA
GCCACCGTGGGTTGCTGTCTCACATCCTGCACCAGCAACTATCACTTCATGTGTTCCCGAGCC
AAGAACTGTGTCTTTCTGGATGATAAAAAAGTATATTGCCAACGACATCGGGATTTGATCAAAG
GCGAAGTGGTTCCTGAGAATGGATTTGAAGTTTTCAGAAGAGTGTTTGTGGACTTTGAAGGAAT
CAGCTTGAGAAGGAAGTTTCTCAATGGCTTGGAACCAGAAAATATCCACATGATGATTGGGTCT
ATGACAATCGACTGCTTAGGAATTCTAAATGATCTCTCCGACTGTGAAGATAAGCTCTTTCCTAT
TGGATATCAGTGTTCCAGGGTATACTGGAGCACCACAGATGCTCGCAAGCGCTGTGTATATAC
ATGCAAGATAGTGGAGTGCCGTCCTCCAGTCGTAGAGCCGGATATCAACAGCACTGTTGAACA
TGATGAAAACAGGACCATTGCCCATAGTCCAACATCTTTTACAGAAAGTTCATCAAAAGAGAGT
CAAAACACAGCTGAAATTATAAGTCCTCCATCACCAGACCGACCTCCTCATTCACAAACCTCTG
GCTCCTGTTATTATCATGTCATCTCAAAGGTCCCCAGGATTCGAACACCCAGTTATTCTCCAAC
ACAGAGATCCCCTGGCTGTCGACCGTTGCCTTCTGCAGGAAGTCCTACCCCAACCACTCATGA
AATAGTCACAGTAGGTGATCCTTTACTCTCCTCTGGACTTCGAAGCATTGGCTCCAGGCGTCA
CAGTACCTCTTCCTTATCACCCCAGCGGTCCAAACTCCGGATAATGTCTCCAATGAGAACTGG
GAATACTTACTCTAGGAATAATGTTTCCTCAGTCTCCACCACCGGGACCGCTACTGATCTTGAA
TCAAGTGCCAAAGTAGTTGATCATGTCTTAGGGCCACTGAATTCAAGTACTAGTTTAGGGCAAA
ACACTTCCACCTCTTCAAATTTGCAAAGGACAGTGGTTACTGTAGGCAATAAAAACAGTCACTT
GGATGGATCTTCATCTTCAGAAATGAAGCAGTCCAGTGCTTCAGACTTGGTGTCCAAGAGCTC
CTCTTTAAAGGGAGAGAAGACCAAAGTGCTGAGTTCCAAGAGCTCAGAGGGATCTGCACATAA
TGTGGCTTACCCTGGAATTCCTAAACTGGCCCCACAGGTTCATAACACAACATCTAGAGAACTG
AATGTTAGTAAAATCGGCTCCTTTGCTGAACCCTCTTCAGTGTCGTTTTCTTCTAAAGAGGCCC
TCTCCTTCCCACACCTCCATTTGAGAGGGCAAAGGAATGATCGAGACCAACACACAGATTCTA
CCCAATCAGCAAACTCCTCTCCAGATGAAGATACTGAAGTCAAAACCTTGAAGCTATCTGGAAT
GAGCAACAGATCATCCATTATCAACGAACATATGGGATCTAGTTCCAGAGATAGGAGACAGAA
AGGGAAAAAATCCTGTAAAGAAACTTTCAAAGAAAAGCATTCCAGTAAATCTTTTTTGGAACCTG
GTCAGGTGACAACTGGTGAGGAAGGAAACTTGAAGCCAGAGTTTATGGATGAGGTTTTGACTC
CTGAGTATATGGGCCAACGACCATGTAACAATGTTTCTTCTGATAAGATTGGTGATAAAGGCCT
TTCTATGCCAGGAGTCCCCAAAGCTCCACCCATGCAAGTAGAAGGATCTGCCAAGGAATTACA
GGCACCACGGAAACGCACAGTCAAAGTGACACTGACACCTCTAAAAATGGAAAATGAGAGTCA
ATCCAAAAATGCCCTGAAAGAAAGTAGTCCTGCTTCCCCTTTGCAAATAGAGTCAACATCTCCC
ACAGAACCAATTTCAGCCTCTGAAAATCCAGGAGATGGTCCAGTGGCCCAACCAAGCCCCAAT
AATACCTCATGCCAGGATTCTCAAAGTAACAACTATCAGAATCTTCCAGTACAGGACAGAAACC
TAATGCTTCCAGATGGCCCCAAACCTCAGGAGGATGGCTCTTTTAAAAGGAGGTATCCCCGTC
GCAGTGCCCGTGCACGTTCTAACATGTTTTTTGGGCTTACCCCACTCTATGGAGTAAGATCCTA
TGGTGAAGAAGACATTCCATTCTACAGCAGCTCAACTGGGAAGAAGCGAGGCAAGAGATCAGC
TGAAGGACAGGTGGATGGGGCCGATGACTTAAGCACTTCAGATGAAGACGACTTATACTATTA
CAACTTCACTAGAACAGTGATTTCTTCAGGTGGAGAGGAACGACTGGCATCCCATAATTTATTT
CGGGAGGAGGAACAGTGTGATCTTCCAAAAATCTCACAGTTGGATGGTGTTGATGATGGGACA
GAGAGTGATACTAGTGTCACAGCCACAACAAGGAAAAGCAGCCAGATTCCAAAAAGAAATGGT
AAAGAAAATGGAACAGAGAACTTAAAGATTGATAGACCTGAAGATGCTGGGGAGAAAGAACAT
GTCACTAAGAGTTCTGTTGGCCACAAAAATGAGCCAAAGATGGATAACTGCCATTCTGTAAGCA
GAGTTAAAACACAGGGACAAGATTCCTTGGAAGCTCAGCTCAGCTCATTGGAGTCAAGCCGCA
GAGTCCACACAAGTACCCCCTCCGACAAAAATTTACTGGACACCTATAATACTGAGCTCCTGAA
ATCAGATTCAGACAATAACAACAGTGATGACTGTGGGAATATCCTGCCTTCAGACATTATGGAC
TTTGTACTAAAGAATACTCCATCCATGCAGGCTTTGGGTGAGAGCCCAGAGTCATCTTCATCAG
AACTCCTGAATCTTGGTGAAGGATTGGGTCTTGACAGTAATCGTGAAAAAGACATGGGTCTTTT
TGAAGTATTTTCTCAGCAGCTGCCTACAACAGAACCTGTGGATAGTAGTGTCTCTTCCTCTATC
TCAGCAGAGGAACAGTTTGAGTTGCCTCTAGAGCTACCATCTGATCTGTCTGTCTTGACCACC
CGGAGTCCCACTGTCCCCAGCCAGAATCCCAGTAGACTAGCTGTTATCTCAGACTCAGGGGA
GAAGAGAGTAACCATCACAGAAAAATCTGTAGCCTCCTCTGAAAGTGACCCAGCACTGCTGAG
CCCAGGAGTAGATCCAACTCCTGAAGGCCACATGACTCCTGATCATTTTATCCAAGGACACAT
```
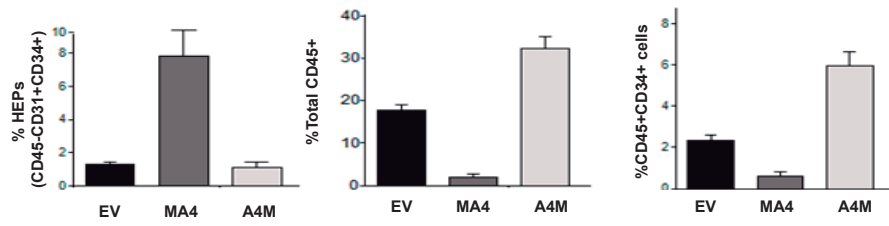
GGATGCAGACCACATCTCTAGCCCTCCTTGTGGTTCAGTAGAGCAAGGTCATGGCAACAATCA
GGATTTAACTAGGAACAGTAGCACCCCTGGCCTTCAGGTACCTGTTTCCCCAACTGTTCCCAT
CCAGAACCAGAAGTATGTGCCCAATTCTACTGATAGTCCTGGCCCGTCTCAGATTTCCAATGCA
GCTGTCCAGACCACTCCACCCCACCTGAAGCCAGCCACTGAGAAACTCATAGTTGTTAACCAG
AACATGCAGCCACTTTATGTTCTCCAAACTCTTCCAAATGGAGTGACCCAAAAAATCCAATTGA
CCTCTTCTGTTAGTTCTACACCCAGTGTGATGGAGACAAATACTTCAGTATTGGGACCCATGGG
AGGTGGTCTCACCCTTACCACAGGACTAAATCCAAGCTTGCCAACTTCTCAATCTTTGTTCCCT
TCTGCTAGCAAAGGATTGCTACCCATGTCTCATCACCAGCACTTACATTCCTTCCCTGCAGCTA
CTCAAAGTAGTTTCCCACCAAACATCAGCAATCCTCCTTCAGGCCTGCTTATTGGGGTTCAGCC
TCCTCCGGATCCCCAACTTTTGGTTTCAGAATCCAGCCAGAGGACAGACCTCAGTACCACAGT
AGCCACTCCATCCTCTGGACTCAAGAAAAGACCCATATCTCGTCTACAGACCCGAAAGAATAAA
AAACTTGCTCCCTCTAGTACCCCTTCAAACATTGCCCCTTCTGATGTGGTTTCTAATATGACATT
GATTAACTTCACACCCTCCCAGCTTCCTAATCATCCAAGTCTGTTAGATTTGGGGTCACTTAATA
CTTCATCTCACCGAACTGTCCCCAACATCATAAAAAGATCTAAATCTAGCATCATGTATTTTGAA
CCGGCACCCCTGTTACCACAGAGTGTGGGAGGAACTGCTGCCACAGCGGCAGGCACATCAAC
AATAAGCCAGGATACTAGCCACCTCACATCAGGGTCTGTGTCTGGCTTGGCATCCAGTTCCTC
TGTCTTGAATGTTGTATCCATGCAAACTACCACAACCCCTACAAGTAGTGCGTCAGTTCCAGGA
CACGTCACCTTAACCAACCCAAGGTTGCTTGGTACCCCAGATATTGGCTCAATAAGCAATCTTT
TAATCAAAGCTAGCCAGCAGAGCCTGGGGATTCAGGACCAGCCTGTGGCTTTACCGCCAAGTT
CAGGAATGTTTCCACAACTGGGGACATCACAGACCCCCTCTACTGCTGCAATAACAGCGGCAT
CTAGCATCTGTGTGCTCCCCTCCACTCAGACTACGGGCATAACAGCCGCTTCACCTTCTGGGG
AAGCAGACGAACACTATCAGCTTCAGCATGTGAACCAGCTCCTTGCCAGCAAAACTGGGATTC
ATTCTTCCCAGCGTGATCTTGATTCTGCTTCAGGGCCCCAGGTATCCAACTTTACCCAGACGGT
AGACGCTCCTAATAGCATGGGACTGGAGCAGAACAAGGCTTTATCCTCAGCTGTGCAAGCCAG
CCCCACCTCTCCTGGGGGTTCTCCATCCTCTCCATCTTCTGGACAGCGGTCAGCAAGCCCTTC
AGTGCCGGGTCCCACTAAACCCAAACCAAAAACCAAACGGTTTCAGCTGCCTCTAGACAAAGG
GAATGGCAAGAAGCACAATGTTTCCCATTTGCGGACCAGTTCTTCTGAAGCACACATTCCAGA
CCAAGAAACGACATCCCTGACCTCAGGCACAGGGACTCCAGGAGCAGAGGCTGAGCAGCAG
GATACAGCTAGCGTGGAGCAGTCCTCCCAGAAGGAGTGTGGGCAACCTGCAGGGCAAGTCGC
TGTTCTTCCGGAAGTTCAGGTGACCCAAAATCCAGCAAATGAACAAGAAAGTGCAGAACCTAA
AACAGTGGAAGAAGAGGAAAGTAATTTCAGCTCCCCACTGATGCTTTGGCTTCAGCAAGAACA
AAAGCGGAAGGAAAGCATTACTGAGAAAAAACCCAAGAAAGGACTTGTTTTTGAAATTTCCAGT
GATGATGGCTTTCAGATCTGTGCAGAAAGTATTGAAGATGCCTGGAAGTCATTGACAGATAAAG
TCCAGGAAGCTCGATCAAATGCCCGCCTAAAGCAGCTCTCATTTGCAGGTGTTAACGGTTTGA
GGATGCTGGGGATTCTCCATGATGCAGTTGTGTTCCTCATTGAGCAGCTGTCTGGTGCCAAGC
ACTGTCGAAATTACAAATTCCGTTTCCACAAGCCAGAGGAGGCCAATGAACCCCCCTTGAACC
CTCACGGCTCAGCCAGGGCTGAAGTCCACCTCAGGAAGTCAGCATTTGACATGTTTAACTTCC
TGGCTTCTAAACATCGTCAGCCTCCTGAATACAACCCCAATGATGAAGAAGAGGAGGAGGTAC
AGCTGAAGTCAGCTCGGAGGGCAACTAGCATGGATCTGCCAATGCCCATGCGCTTCCGGCAC
TTAAAAAAGACTTCTAAGGAGGCAGTTGGTGTCTACAGGTCTCCCATCCATGGCCGGGGTCTT
TTCTGTAAGAGAAACATTGATGCAGGTGAGATGGTGATTGAGTATGCCGGCAACGTCATCCGC
TCCATCCAGACTGACAAGCGGGAAAAGTATTACGACAGCAAGGGCATTGGTTGCTATATGTTC
CGAATTGATGACTCAGAGGTAGTGGATGCCACCATGCATGGAAATGCTGCACGCTTCATCAAT
CACTCGTGTGAGCCTAACTGCTATTCTCGGGTCATCAATATTGATGGGCAGAAGCACATTGTCA
TCTTTGCCATGCGTAAGATCTACCGAGGAGAGGAACTCACTTACGACTATAAGTTCCCCATTGA
GGATGCCAGCAACAAGCTGCCCTGCAACTGTGGCGCTAAGAAATGCCGGAAGTTCCTAAAC<mark>TA</mark>
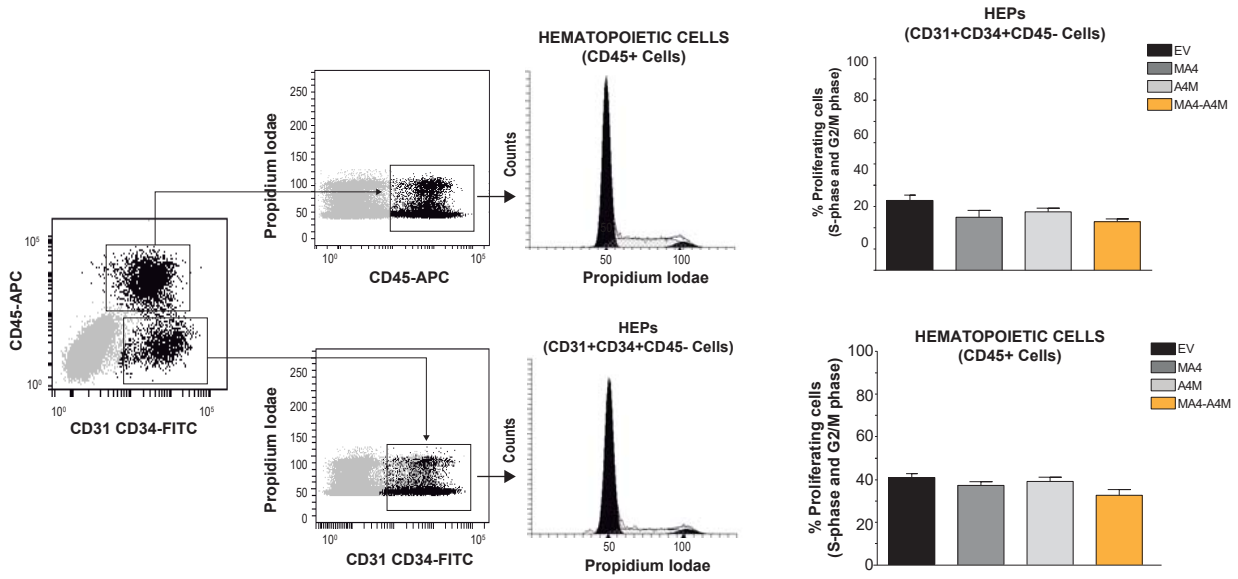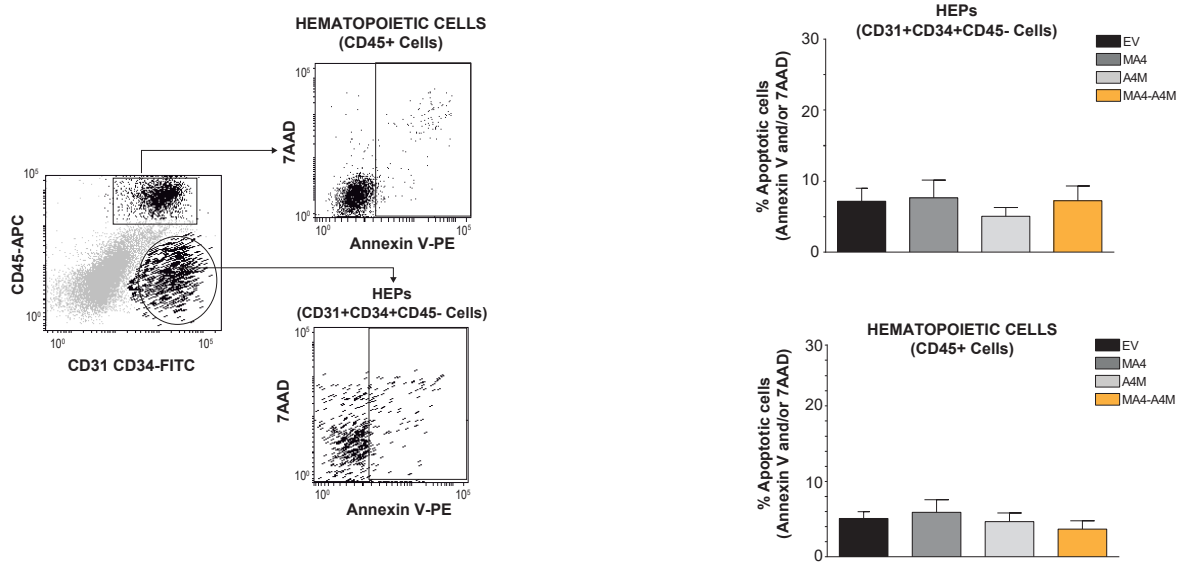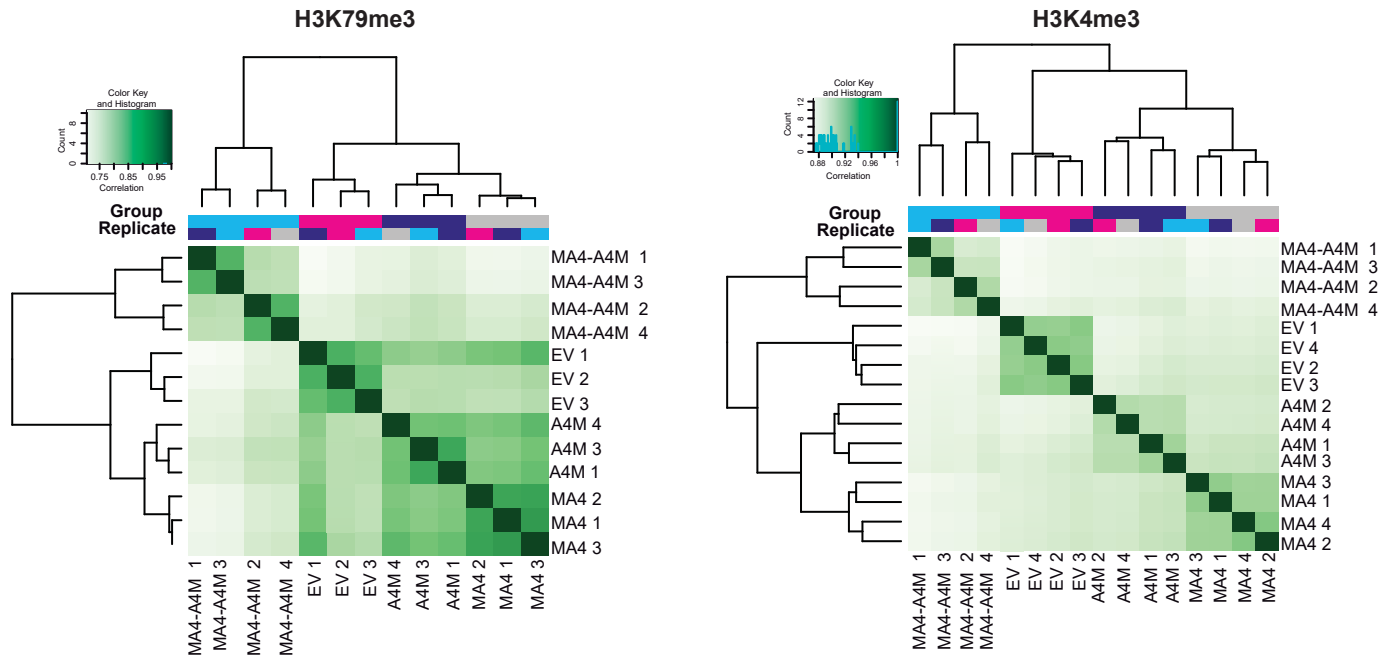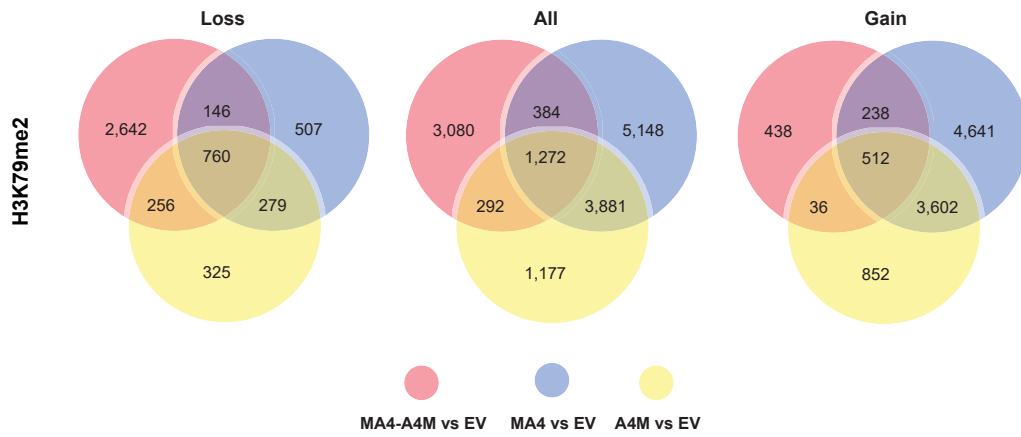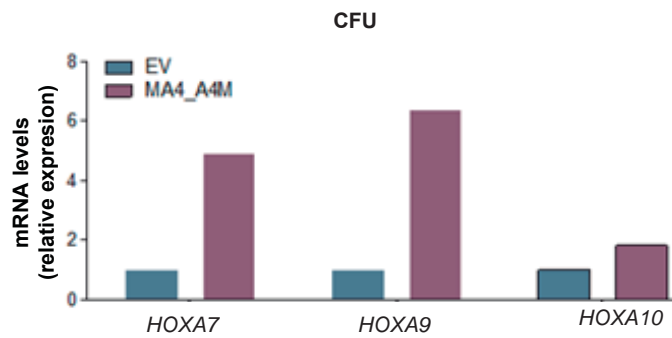<mark>A</mark>

B

A

B

C

A

B

C

A

H3K4me3 HOXA cluster

B