Ontogenic Changes in Hematopoietic Hierarchy Determine Pediatric Specificity and Disease Phenotype in Fusion Oncogene–Driven Myeloid Leukemia

Cécile K. Lopez[1,2,3,4], Esteve Noguera[1,2,4], Vaia Stavropoulou[5], Elie Robert[1,2,4], Zakia Aid[1,2,4], Paola Ballerini[6], Chrystèle Bilhou-Nabera[7], Hélène Lapillonne[8], Fabien Boudia[1,2,4,9], Cécile Thirant[1,2,4], Alexandre Fagnan[1,2,4,9], Marie-Laure Arcangeli[10], Sarah J. Kinston[11], M'Boyba Diop[2], Bastien Job[2], Yann Lecluse[2], Erika Brunet[12], Loélia Babin[12], Jean Luc Villeval[1,2], Eric Delabesse[13], Antoine H.F.M. Peters[14,15], William Vainchenker[1,2], Muriel Gaudry[1,2], Riccardo Masetti[16], Franco Locatelli[17,18], Sébastien Malinge[1,2,3,4], Claus Nerlov[19], Nathalie Droin[1], Camille Lobry[1], Isabelle Godin[1,2], Olivier A. Bernard[1,2,3,4], Berthold Göttgens[11], Arnaud Petit[6], Françoise Pflumio[10], Juerg Schwaller[5], and Thomas Mercher[1,2,4,9]

ABSTRACT

Fusion oncogenes are prevalent in several pediatric cancers, yet little is known about the specific associations between age and phenotype. We observed that fusion oncogenes, such as *ETO2–GLIS2*, are associated with acute megakaryoblastic or other myeloid leukemia subtypes in an age-dependent manner. Analysis of a novel inducible transgenic mouse model showed that *ETO2–GLIS2* expression in fetal hematopoietic stem cells induced rapid megakaryoblastic leukemia whereas expression in adult bone marrow hematopoietic stem cells resulted in a shift toward myeloid transformation with a strikingly delayed *in vivo* leukemogenic potential. Chromatin accessibility and single-cell transcriptome analyses indicate ontogeny-dependent intrinsic and *ETO2–GLIS2*-induced differences in the activities of key transcription factors, including ERG, SPI1, GATA1, and CEBPA. Importantly, switching off the fusion oncogene restored terminal differentiation of the leukemic blasts. Together, these data show that aggressiveness and phenotypes in pediatric acute myeloid leukemia result from an ontogeny-related differential susceptibility to transformation by fusion oncogenes.

SIGNIFICANCE: This work demonstrates that the clinical phenotype of pediatric acute myeloid leukemia is determined by ontogeny-dependent susceptibility for transformation by oncogenic fusion genes. The phenotype is maintained by potentially reversible alteration of key transcription factors, indicating that targeting of the fusions may overcome the differentiation blockage and revert the leukemic state.

See related commentary by Cruz Hernandez and Vyas, p. 1653.

[1]INSERM U1170, Gustave Roussy, Villejuif, France. [2]Gustave Roussy, Villejuif, France. [3]Université Paris-Saclay, Villejuif, France. [4]Equipe labellisée Ligue Nationale Contre le Cancer, Paris, France. [5]University Children's Hospital Beider Basel (UKBB) and Department of Biomedicine, University of Basel, Basel, Switzerland. [6]Hôpital Trousseau, AP-HP, Paris, France. [7]Sorbonne Université, CRSA–Unité INSERM, AP-HP, Hôpital Trousseau, Paris, France. [8]Hôpital Saint Antoine, AP-HP, Paris, France. [9]Université Paris Diderot, Paris, France. [10]Unité Mixte de Recherche 967 INSERM, CEA/DRF/IBFJ/IRCM/LSHL, Université Paris-Diderot-Université Paris-Sud, Equipe labellisée Association Recherche Contre le Cancer, Fontenay-aux-roses, France. [11]Wellcome and MRC Cambridge Stem Cell Institute and the Cambridge Institute for Medical Research, University of Cambridge, Cambridge, United Kingdom. [12]Genome Dynamics in the Immune System Laboratory, Institut Imagine, INSERM, Université Paris Descartes, Sorbonne Paris Cité, Equipe Labellisée Ligue Contre le Cancer, Paris, France. [13]INSERM U1037, Team 16, Center of Research of Cancerology of Toulouse, Hematology Laboratory, IUCT-Oncopole, France. [14]Friedrich Miescher Institute for Biomedical Research (FMI), Basel, Switzerland. [15]Faculty of Sciences, University of Basel, Basel, Switzerland. [16]Department of Pediatrics, "Lalla Seràgnoli," Hematology-Oncology Unit, Sant'Orsola-Malpighi Hospital, University of Bologna, Bologna, Italy. [17]Department of Pediatrics, Sapienza, University of Rome, Rome, Italy. [18]Hematology-Oncology-IRCCS Ospedale Bambino Gesù, Rome, Italy. [19]MRC Molecular Hematology Unit, MRC Weatherall Institute of Molecular Medicine, Radcliffe Department of Medicine, University of Oxford, Oxford, United Kingdom.

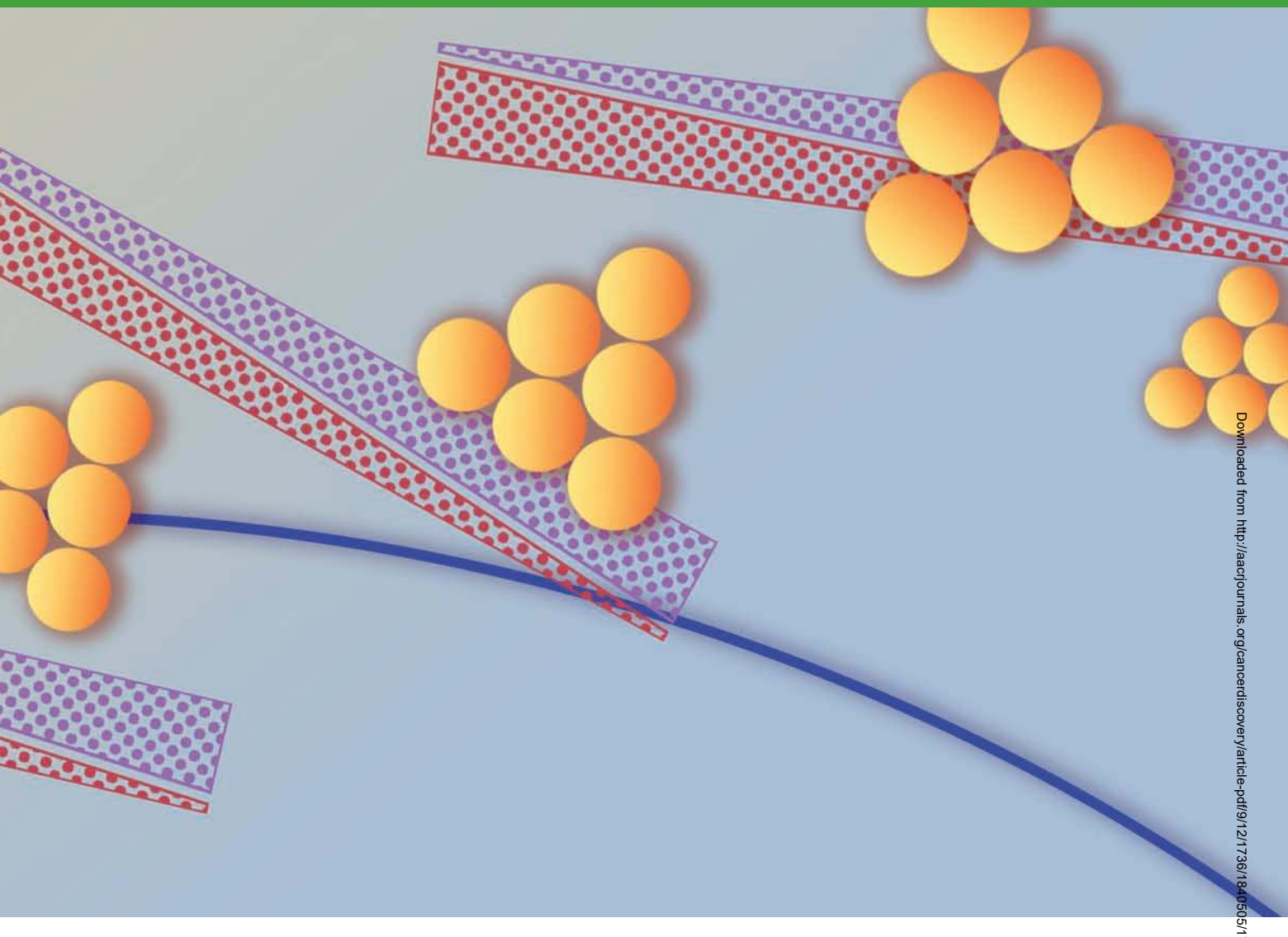
Note: Supplementary data for this article are available at Cancer Discovery Online (http://cancerdiscovery.aacrjournals.org/).

E. Noguera, V. Stavropoulou, and E. Robert are the co-second authors of this article.

Corresponding Authors: Thomas Mercher, Institut Gustave Roussy, INSERM U1170, 39 rue Camille Desmoulins, Villejuif 94800, France. Phone: 331-4211-4483; E-mail: thomas.mercher@inserm.fr; and Juerg Schwaller, Department of Biomedicine (DBM), University Children's Hospital beider Basel (UKBB), University of Basel, ZLF-Lab 202, Hebelstrasse 20, Basel CH-4031, Switzerland. Phone: 416-1265-3504; E-mail: j.schwaller@unibas.ch

Cancer Discov 2019;9:1736–53

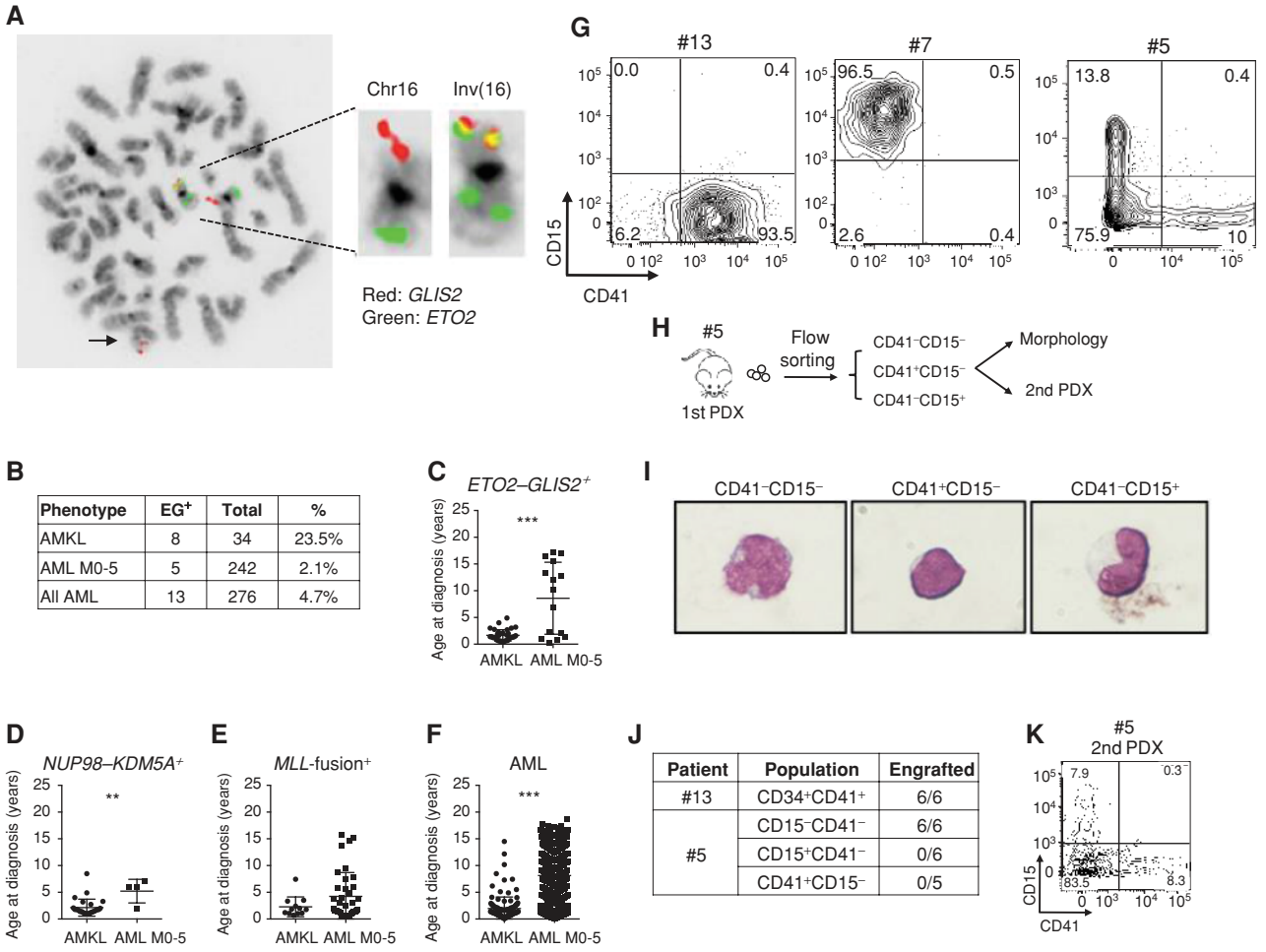
doi: 10.1158/2159-8290.CD-18-1463



INTRODUCTION

Increasing evidence supports a correlation between age, driver oncogene, and disease phenotype in pediatric cancers (1, 2). The differences between adult and pediatric cancers may stem from either the distinct properties of the driver oncogenes or inherent differences between fetal and adult tissue. In acute leukemia, some fusion oncogenes are acquired *in utero* and are diagnosed almost exclusively during childhood (3–5), for which efficient leukemogenesis models are lacking (6). Childhood *de novo* acute megakaryoblastic leukemia (AMKL) is an aggressive subtype of acute myeloid leukemia (AML) that harbors recurrent fusion oncogenes (e.g., *ETO2–GLIS2, OTT–MAL, NUP98–KDM5A*; refs. 7–10). Intriguingly, several AMKL fusions have also been detected in patients with other AML subtypes (8, 11, 12), although the determinants of the genotype–phenotype association remain unknown. Similar to other aggressive childhood cancers such as Ewing sarcoma (13), these fusions are rarely associated with additional recurrent mutations in patients (9, 14) and they strongly affect the epigenome (15, 16), thereby suggesting that global transcriptional deregulation is sufficient to drive the disease and a likely candidate for direct therapeutic targeting (17, 18). Nevertheless, these cancers currently lack adequate animal models that precisely recapitulate the characteristics of the disease observed in patients. Such models could provide mechanistic insights to understand pediatric cancer specificity.

Several differences exist between fetal and adult normal hematopoiesis (19, 20). In adults, hematopoietic stem cells (HSC) are at the top of a hierarchy and differentiate toward more committed progenitors to generate the mature blood cell types. Although the precise trajectories that govern normal hematopoiesis remain a matter of debate (21–24), leukemogenic processes are strongly influenced by the type of cell targeted by fusion oncogenes (25). A recent study has proposed that fetal liver hematopoiesis primarily involves multipotent progenitors, whereas adult hematopoiesis is chiefly regulated by committed progenitors with more restricted lineage potential (26). This view is largely supported by single-cell transcriptome data showing that most adult hematopoietic progenitors present signs of lineage commitment (21, 24, 27–29).



Phenotype	EG⁺	Total	%
AMKL	8	34	23.5%
AML M0-5	5	242	2.1%
All AML	13	276	4.7%

Patient	Population	Engrafted
#13	CD34⁺CD41⁺	6/6
#5	CD15⁻CD41⁻	6/6
	CD15⁺CD41⁻	0/6
	CD41⁺CD15⁻	0/5

Figure 1. AMKL is diagnosed at a younger age than AML in pediatric patients. **A,** Representative FISH with probes detecting *ETO2* (green) and *GLIS2* (red) in leukemic cells from a patient with EG⁺ AML. The yellow signal results from the red–green overlap, thus indicating the derivative chromosome carrying the EG fusion. **B,** Frequencies of EG⁺ patients within AMKL ($n = 34$), other AML subtypes ("AML M0-5," $n = 242$), and in the entire ELAM02 cohort ("All AML," $n = 276$). **C,** Age at diagnosis of EG⁺ patients obtained from pooling data from the ELAM02 cohort and published reports (9, 12, 73–76). AMKL ($n = 33$) versus AML ($n = 15$), $P < 0.0001$. **D,** Age at diagnosis of *NUP98-KDM5A*⁺ patients, AMKL ($n = 27$) versus AML ($n = 4$), $P = 0.0017$. **E,** Age at diagnosis of patients with MLL fusions, AMKL ($n = 18$) versus AML ($n = 88$), $P = 0.095$. **F,** Age at diagnosis of patients with AMKL ($n = 132$) versus other AML ($n = 500$) subtypes, $P < 0.0001$. **G,** Flow-cytometry analyses of primary patient-derived xenografts (PDX) obtained from patients with AMKL and AML. Analyses are gated on CD34⁺ human blasts. **H,** Experimental design of flow-cytometric sorting of the three blast populations analyzed in **I–K**. **I,** Representative cell morphology of the three leukemic cell populations on cytospots. Wright–Giemsa staining. Magnification: ×100. **J,** Secondary xenotransplantation of various cell populations from patient #5 indicates that only CD41⁻CD15⁻ cells propagated the disease. **K,** CD15 and CD41 expression on CD34⁺ cells obtained from recipients of 3×10^5 CD15⁻CD41⁻ cells 1 year post-transplant. Statistical significance is indicated as *P* values (Student *t* test). *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

At the molecular level, fetal and adult hematopoiesis present differential activity of LIN28B (30–33), GATA factor–regulated Polycomb complexes (34), and CEBPA/MYCN transcription factors (35).

To better understand the genotype–phenotype association in pediatric cancers, we investigated how the fetal and adult cellular architectures (22–24, 27, 36) affect the transformation and disease phenotype driven by *ETO2–GLIS2* (hereafter abbreviated EG). We developed an inducible transgenic mouse model that efficiently phenocopied the human disease with striking differences in latency, phenotype, and molecular wiring, depending on the ontogenic stage at which the driver fusion oncogene is activated.

RESULTS

AMKL Is Diagnosed at a Younger Age than AML in Pediatric Patients

We first screened a cohort of 276 pediatric patients with leukemia (37) for the presence of EG and confirmed that EG occurred in patients with AMKL and other AML (AML-M0 to AML-M5) subtypes (12), with an overall incidence of 4.7% (Fig. 1A and B; Supplementary Table S1; Supplementary Fig. S1A). The overall survival was poor for all patients, with a trend toward a worse prognosis for those presenting with an AMKL phenotype (Supplementary Fig. S1B). In addition to the known EG⁺ MO7e megakaryoblastic cell line, we found

that the CMS leukemia cell line was also EG$^+$ and presented with a myeloid phenotype (Supplementary Fig. S1C–S1F). Strikingly, patients with EG$^+$ AMKL were diagnosed at a significantly younger age of 1.66 ± 0.18 years ($n = 33$) compared with other patients with EG$^+$ AML [8.53 ± 1.73 years ($n = 15$); $P < 0.0001$; Fig. 1C]. A similar trend was observed in patients with AMKL carrying either *NUP98–KDM5A* (Fig. 1D) or *MLL* fusions (Fig. 1E) as well as for all patients with AMKL regardless of the genetic subgroup (Fig. 1F).

We applied patient-derived xenografts (8) to expand EG$^+$ AMKL and EG$^+$ AML cells in immunodeficient mice. Overall, 3 of 5 EG$^+$ AMKL and 2 of 8 EG$^+$ AML transplants successfully engrafted (Supplementary Table S1). Whereas AMKL xenografts generally expressed the megakaryoblastic marker CD41, but not the myeloid marker CD15 (as observed for patient #13), AML blasts were CD15$^+$CD41$^-$ (e.g., patient #7; Fig. 1G). Interestingly, we identified a 6-year-old EG$^+$ patient (#5) diagnosed as AML-M2, who presented with immature blasts (CD34$^+$CD15$^-$CD41$^-$) and two smaller cellular fractions (CD34$^+$CD41$^+$CD15$^-$ or CD34$^+$CD41$^-$CD15$^+$; Fig. 1G). All cell populations presented with distinct morphologies (Fig. 1H and I) and expressed EG (Supplementary Fig. S1G). Upon injection of the three flow-purified cell populations into secondary recipients, only the immature CD34$^+$CD41$^-$CD15$^-$ cell population successfully engrafted into recipients (Fig. 1J) and regenerated the three populations (Fig. 1K), thereby indicating dual megakaryoblastic and myeloid differentiation.

Together, these results show that pediatric AMKL occurs at a significantly younger age than other AML subtypes irrespective of the driver oncogene. Importantly, the dual megakaryoblastic/myeloid phenotype in the EG$^+$ patient indicates that AMKL-associated fusion oncogenes can transform HSC or multipotent progenitors.

ETO2–GLIS2 Efficiently Induces Leukemia in Mice

To analyze the determinants of EG-driven leukemogenesis and the associated megakaryoblastic and myeloid phenotypes, we established a doxycycline-regulated transgenic mouse model (Supplementary Fig. S2A). Hereby, EG expression is controlled by a reverse tetracycline-controlled transactivator (rtTA) integrated into the *Rosa26* gene locus, and results in similar EG expression levels as observed in AMKL patient blasts (Supplementary Fig. S2B and S2C). To assess the consequences of EG expression, we continuously exposed double transgenic *iEG:rtTA* females (hereafter abbreviated iEG) to doxycycline (Dox$^+$; Fig. 2A). Although most hematologic parameters remained unchanged at 6 weeks postinduction (data not shown), iEG animals consistently presented with an abnormal bone marrow (BM) Lin$^-$CD150$^-$Kit$^+$CD41$^+$ cell population (Supplementary Fig. S2D) bearing a myeloid and megakaryoblastic clonogenic potential in methylcellulose cultures (Supplementary Fig. S2E). Later, doxycycline-treated iEG animals developed lethal hematologic malignancies with a prevalence of 96.9% (31/32 mice; Supplementary Fig. S2F), which was classified into two groups based on the most abundant blast populations (Fig. 2B and C; Supplementary Fig. S2G). Approximately 20% of the mice (4/19, Group 1) presented with a CD41$^+$ (megakaryoblastic) phenotype with a median latency of 164 days (Fig. 2D). Group 2 presented with heterogenous phenotypes ranging from predominantly Kit$^+$ (immature) with low numbers of CD11b$^+$Gr1$^+$ (myeloid) cells to predominantly maturing Kit$^-$CD11b$^+$Gr1$^+$ myeloid features after a median latency of 251 days (Fig. 2B–D; Supplementary Fig. S2G). Histologic analyses revealed altered spleen architectures and confirmed that BM and spleens from Group 1 mice were infiltrated with immature von Willebrand factor (VWF$^+$) megakaryoblasts, whereas Group 2 presented with myeloid features in the BM and extramedullary hematopoiesis in the spleen (Fig. 2C; Supplementary Fig. S2H). All diseased mice developed hepatosplenomegaly (Fig. 2E) and anemia (Fig. 2F), and Group 1 was associated with significantly higher white blood cell (WBC) counts (Fig. 2F). The cell-autonomous nature of the disease and its dependence on EG expression was demonstrated by transplantation of BM cells into Dox$^+$ wild-type (WT) recipients that developed disease after a shorter mean latency (120 days), whereas Dox$^-$ recipients never developed any disease (Fig. 2G–J; Supplementary Fig. S2H). Collectively, this inducible transgenic model faithfully recapitulates EG$^+$ leukemia development as observed in patients.

Developmental Stage Determines the Phenotype of *ETO2–GLIS2*$^+$ Leukemia

To investigate the impact of the developmental stage, we induced iEG expression in Lin$^-$ hematopoietic stem and progenitor cells (HSPC) obtained from fetal liver at embryonic day (E) 12.5, newborn BM (1–2 weeks old), and adult BM (ABM; 8 weeks old; Fig. 3A). *In vitro*, iEG expression significantly increased the clonogenic activity with passages in all contexts (Fig. 3B). Whereas iEG$^+$ fetal liver–derived cells homogeneously expressed CD41, iEG$^+$ ABM–derived cells also consistently showed a CD11b$^+$Gr1$^+$ myeloid phenotype (Fig. 3C; Supplementary Fig. S3A). Newborn-derived iEG cells presented with a myelo/megakaryoblastic intermediate phenotype between that of fetal liver and ABM. Similar results were obtained upon retroviral expression of EG in WT cells (Supplementary Fig. S3B). These data indicate that fetal hematopoiesis favors EG-induced megakaryoblastic transformation *in vitro*.

To compare the capacity of fetal liver– and ABM-derived iEG cells to develop disease, we transplanted 1×10^6 mononuclear iEG cells from E14.5 fetal liver (FL) or ABM donors (carrying a ubiquitously expressed GFP reporter to track their progeny) into Dox$^+$ WT recipients (Fig. 3A). Mice engrafted with iEG FL cells developed lethal leukemia with median latency of 39 days (Fig. 3D) characterized by splenomegaly, high WBC, anemia, thrombocytopenia, and infiltration with CD41$^+$ megakaryoblasts in several organs, whereas mice engrafted with iEG ABM cells developed the disease after a significantly prolonged latency (median 397 days, $P = 0.0008$) associated with heterogenous phenotypes ranging from Kit$^+$CD41$^-$ immature to predominantly CD11b$^+$Gr1$^+$ myeloid features (Fig. 3D–H and data not shown). To address the consequences of EG expression on fetal progenitors *in vivo*, we exposed pregnant females (E13) to doxycycline and analyzed the embryos 36 hours later. All iEG-expressing embryos showed gross morphologic and hematologic abnormalities (Fig. 3I), with decreased erythroid and myeloid cells and increased megakaryoblastic progenitors (Fig. 3J).

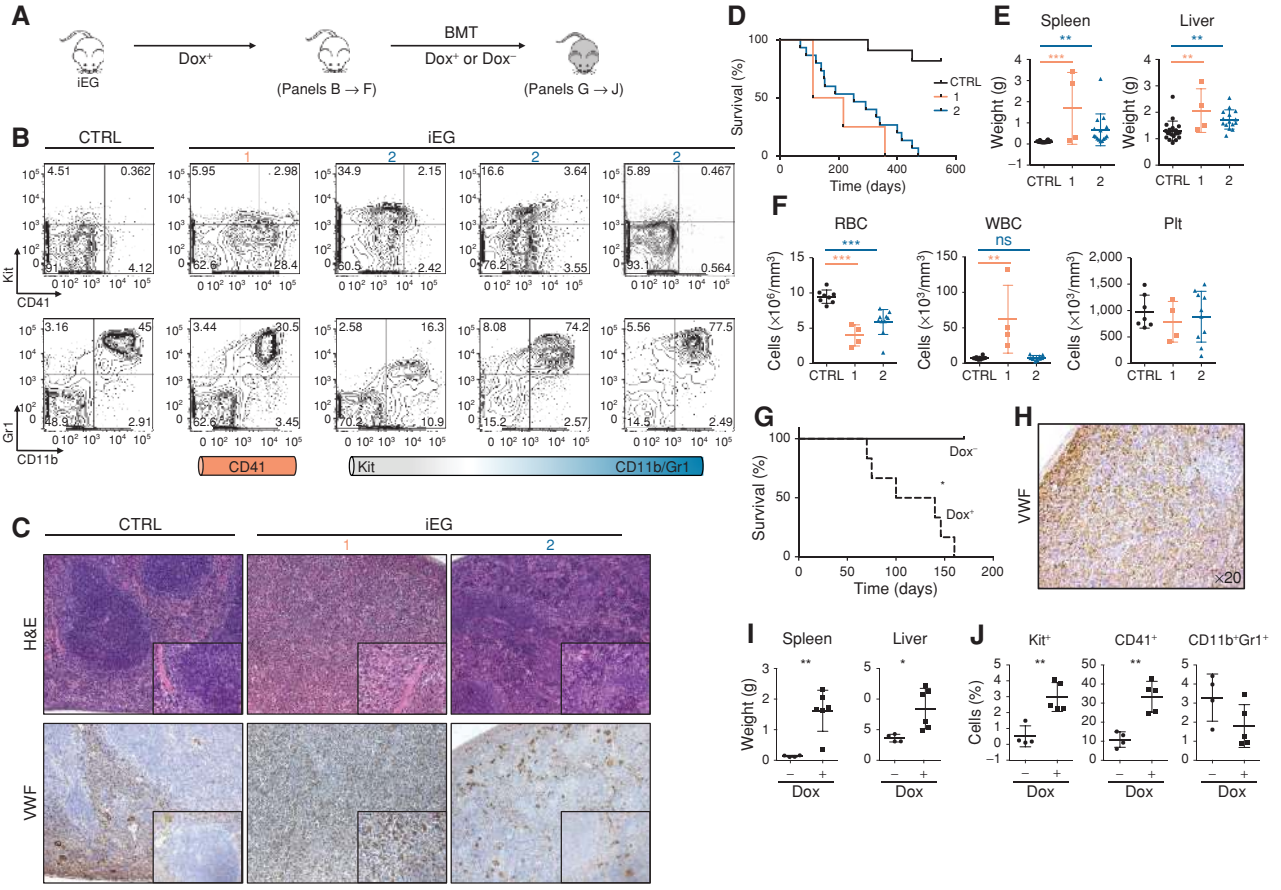


Figure 2. *ETO2–GLIS2* efficiently induces leukemia in mice. **A,** Experimental design: Disease-inducing potential was assessed by providing doxycycline (Dox) to naïve iEG female mice (primary). Cell-autonomous leukemogenesis was ascertained by transplanting 10^6 BM cells from iEG mice induced at 8 weeks of age and for 45 days into wild-type C57BL/6 recipients with or without exposure to doxycycline (secondary). BMT, bone marrow transplant. **B,** Representative flow cytometry analysis of Kit$^+$ and CD41$^+$ expression (top) or CD11b$^+$ and Gr1$^+$ expression (bottom) in BM of primary diseased iEG mice defining two phenotypes: Group 1 = megakaryoblastic (primarily CD41$^+$) and Group 2 = immature/myeloid (heterogeneous expression of Kit$^+$ and CD11b$^+$Gr1$^+$ markers; CTRL $n = 5$, iEG $n = 19$ in total). **C,** Spleen histology of primary mice (CTRL, Group 1, and Group 2). Top, Hematoxylin–Eosin–Saffron (H&E) staining, bottom: VWF staining (brown coloration). **D,** Kaplan–Meier survival plot of doxycycline-exposed iEG mice according to the two phenotypes described in **B** (CTRL $n = 11$, iEG $n = 19$ in total, CTRL vs. Group 1: $P = 0.0001$; CTRL vs. Group 2: $P < 0.0001$). Median survival: Group 1 = 164 days, Group 2 = 251 days. **E,** Spleen and liver weights in primary iEG mice according to the two phenotypes described in **B** (1 = CD41$^+$ and 2 = immature/CD11b$^+$Gr1$^+$). Spleen (CTRL vs. Group 1: $P = 0.0002$; CTRL vs. Group 2: $P = 0.0023$), liver (CTRL vs. Group 1: $P = 0.0060$; CTRL vs. Group 2: $P = 0.0016$). **F,** Peripheral blood counts of diseased iEG mice with the different phenotypes. Red blood cells (RBC; CTRL vs. Group 1: $P < 0.0001$; CTRL vs. Group 2: $P < 0.0001$), WBC (CTRL vs. Group 1: $P = 0.0068$, CTRL vs. Group 2: $P = 0.7624$), Platelet (Plt; CTRL vs. Group 1: $P = 0.3902$; CTRL vs. Group 2: $P = 0.6533$). **G,** Kaplan–Meier survival plot of secondary recipients with or without doxycycline exposure (Dox$^-$: $n = 4$, Dox$^+$: $n = 6$). $P = 0.0110$ (log-rank Mantel–Cox test). **H,** Representative spleen section of a secondary recipient stained with VWF (brown coloration) shows significant organ infiltration. **I,** Spleen and liver weights in secondary recipients with or without doxycycline exposure. Spleen: $P = 0.0025$; liver: $P = 0.0220$. **J,** Quantification of flow-cytometry analyses in the spleen of secondary recipients with or without doxycycline exposure. Kit$^+$: $P = 0.0026$; CD41$^+$: $P = 0.0015$; CD11b$^+$Gr1$^+$: $P = 0.1002$. Statistical significance is indicated as *P* values (Student *t* test except when otherwise specified). *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

Taken together, these observations show that the developmental stage determines the predominance of aggressive megakaryoblastic over myeloid leukemia upon *ETO2–GLIS2* expression.

ETO2–GLIS2 Expression in Fetal LT-HSCs Induces a Megakaryoblastic Phenotype

To address the consequence of EG on distinct populations of the hematopoietic hierarchy, we purified control (CTRL) and iEG long-term HSCs (LT-HSC) and multipotent progenitors (MPP2, MPP3, and MPP4; ref. 38) obtained from E14.5 FL or ABM (Supplementary Fig. S3C) and characterized their clonogenic and phenotypic profile in Dox$^+$ methylcellulose cultures (Fig. 4A). CTRL FL LT-HSC and MPP cells generated both myeloid Gr1$^+$ and megakaryoblastic CD41$^+$ cells (Fig. 4B; Supplementary Fig. S3D), thereby confirming that most normal mouse FL progenitors maintain megakaryoblastic potential as observed in humans (26). iEG fetal liver LT-HSC and MPP2 cells exclusively generated CD41$^+$ cells (Fig. 4B; Supplementary Fig. S3D), MPP3 cells yielded a large majority of CD41$^+$ cells and rare Gr1$^+$ myeloid cells (Fig. 4B; Supplementary Fig. S3D), and MPP4 cells did not significantly form any colony (Supplementary Fig. S3E). iEG fetal liver LT-HSC and MPP2/3 cells showed increased replating activity compared with CTRL (Fig. 4C). To define the differentiation potential of iEG-transformed

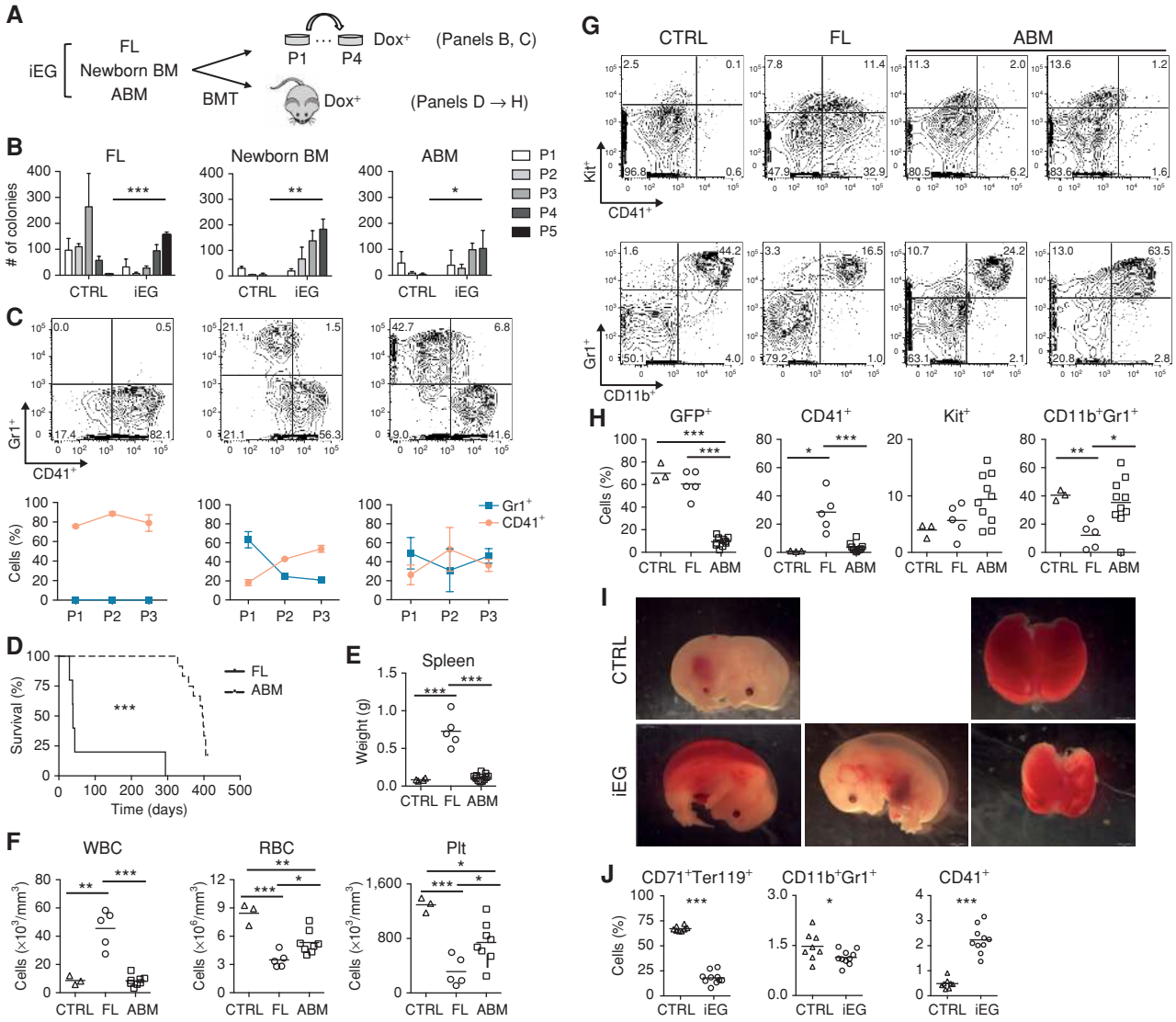


Figure 3. Developmental stage determines the phenotype of *ETO2-GLIS2*+ leukemia. **A,** Experimental design to address the impact of EG on the clonogenic potential of FL or ABM progenitors in methylcellulose shown in **B** and **C** or *in vivo* shown in **D**–**H**. **B,** Number of colonies upon culture of 2,500 FL (FL E12.5) or Lin⁻ BM cells at two weeks (newborn) or two months (ABM) postpartum in doxycycline-containing methylcellulose conditions. Serial replating was performed every 7 days. P1, first plate; P2, secondary plate; P3, tertiary plate; P4, quaternary plate; P5, fifth plate. FL P5 $P < 0.0001$; newborn P4 $P = 0.0034$; ABM P4 $P = 0.0246$. **C,** Representative flow-cytometry analysis of CD41+ and Gr1+ expression at P3 (top) and percentages of CD41+ and Gr1+ cells at P1 to P3 (bottom). **D,** Kaplan–Meier survival plot of WT doxycycline-treated recipients engrafted with 1×10^6 ABM or FL cells from double transgenic iEG+Ubiquitin-GFP+ (GFP+). Log-rank Mantel–Cox test: $P < 0.0001$. Median survival, FL = 39 days; ABM = 397 days. **E,** Spleen weights in CTRL and diseased mice at time of sacrifice (CTRL vs. FL: $P = 0.0005$; CTRL vs. ABM: $P = 0.1999$; FL vs. ABM $P < 0.0001$). **F,** Peripheral blood counts of recipients of FL and ABM. WBC (CTRL vs. FL: $P = 0.0037$; CTRL vs. ABM: $P = 0.9125$; FL vs. ABM $P < 0.0001$), RBC (CTRL vs. FL: $P = 0.0004$; CTRL vs. ABM: $P = 0.0041$; FL vs. ABM $P = 0.0133$), Platelets (Plt; CTRL vs. FL: $P = 0.0003$; CTRL vs. ABM: $P = 0.0137$; FL vs. ABM $P = 0.0181$). **G,** Representative flow-cytometry analyses of BM cells from diseased recipients at time of sacrifice. **H,** Quantification of GFP+ (CTRL vs. FL: $P = 0.2644$; CTRL vs. ABM: $P < 0.0001$; FL vs. ABM $P < 0.0001$), CD41+ (CTRL vs. FL: $P = 0.0154$; CTRL vs. ABM: $P = 0.1870$; FL vs. ABM $P < 0.0001$), and CD11b+Gr1+ (CTRL vs. FL: $P = 0.0024$; CTRL vs. ABM: $P = 0.6335$; FL vs. ABM $P = 0.0167$) in GFP+ BM cells. **I,** Images of E14.5 embryos (left and middle) and corresponding FL (right) obtained from pregnant females on doxycycline for 36 hours. **J,** Flow-cytometry analyses of FL cells obtained in **I** (CTRL $n = 8$; iEG $n = 10$). CD71+Ter119+: $P < 0.0001$; CD11b+Gr1+: $P = 0.0380$; CD41+: $P < 0.0001$. Statistical significance is indicated as *P* values (Student *t* test except when otherwise specified). *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

cells, iEG FL LT-HSC and MPP2/3 cells derived from the fourth plating were cultured without doxycycline (Fig. 4A), which resulted in myeloid and megakaryoblastic differentiation, respectively (Fig. 4D; Supplementary Fig. S3F). CD41+ and Gr1+ iEG cell lines could be derived after flow purification and long-term cultures (Supplementary Fig. S3G), and their CD41+ megakaryoblastic phenotype correlated with the high presence of doxycycline (Supplementary Fig. S3H). Consistently, higher *ETO2-GLIS2* expression was observed in AMKL versus AML patient blasts (Supplementary Fig. S1A),

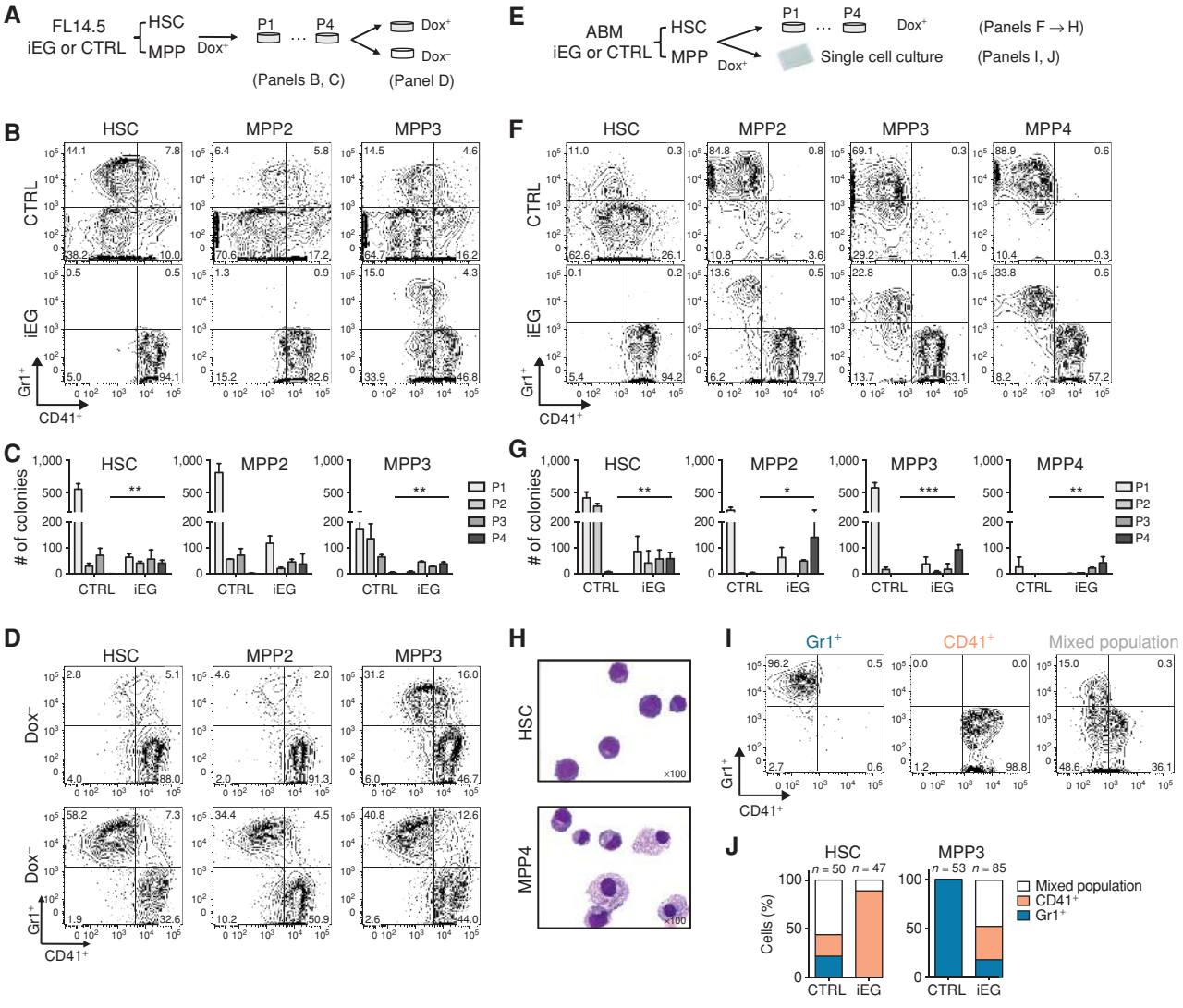


Figure 4. *ETO2-GLIS2*-induced AMKL is an HSC-derived leukemia. **A,** Experimental design to address the clonogenic potential of HSC and MPP from CTRL and iEG E14.5 embryos. Serial replating was performed every 7 days for 4 weeks (P1 to P4). **B,** Representative flow-cytometry analyses of cells obtained from the first plate (P1). **C,** Number of colonies upon serial replating. Mean ± SEM is indicated. **D,** Representative flow-cytometry analyses of cells grown in methylcellulose culture with or without doxycycline (Dox). **E,** Experimental design to address the clonogenic potential of adult (8–10 weeks old) HSC and MPP cells from iEG and CTRL mice. **F,** Representative flow-cytometry analyses of cells obtained at P1. **G,** Number of colonies upon serial replating. Mean ± SEM is indicated. **H,** Wright–Giemsa-stained cytospots from iEG ABM HSC– and MPP4-derived methylcellulose cultures at P3. Magnification: ×100. **I,** Single HSC and MPP3 ABM cells from CTRL or iEG mice were evaluated for their potential to develop myeloid (Gr1$^+$), megakaryoblastic (CD41$^+$), or mixed-cell populations (containing both Gr1$^+$CD41$^-$ and Gr1$^-$CD41$^+$ cells) in 96-well liquid cultures with doxycycline treatment. A representative flow-cytometry analysis of the three types of colonies obtained after 8 days of culture is shown. **J,** Histogram representing the percentages of CD41$^+$, Gr1$^+$, or mixed-population colonies as defined in **I** (HSC CTRL: $n = 50$; HSC iEG: $n = 47$; MPP3 CTRL: $n = 53$; MPP3 iEG: $n = 85$).

supporting the importance of EG dosage as shown for other fusion oncogenes (39).

Using ABM cells, CTRL LT-HSCs formed Gr1$^+$ and CD41$^+$ cells (Fig. 4E and F), whereas CTRL MPP2/3/4 cells generated almost exclusively Gr1$^+$ cells (Fig. 4F; Supplementary Fig. S3D), indicating that the megakaryoblastic potential is qualitatively and quantitatively lower in ABM than in fetal liver (Supplementary Fig. S3I). In contrast, iEG ABM LT-HSCs gave rise almost exclusively to CD41$^+$ cells, whereas MPP2/3/4 cells generated CD11b$^+$Gr1$^+$ myeloid cells with a higher proportion than iEG FL MPP cells (Fig. 4F; Supplementary Fig. S3D). Although iEG FL MPP4 cells did not generate colonies, iEG ABM MPP4 cells showed increasing clonogenic activity upon serial replatings and a myeloid morphology (Fig. 4F–H).

We next investigated whether iEG endows single LT-HSC or MPP3 cells with dual myeloid/megakaryoblastic phenotypic potential as observed in patient #5 (Fig. 1G). The majority (56%) of CTRL ABM LT-HSC–derived cells showed a dual phenotype, whereas others showed either megakaryoblastic or myeloid phenotypes. Upon iEG induction, ABM LT-HSCs formed primarily megakaryoblastic colonies (89%) with only a few dual phenotype colonies (11%; Fig. 4I and J; Supplementary

Fig. S3J). Whereas CTRL MPP3 cells exclusively generated myeloid cells under these conditions, 48% of iEG MPP3 cells yielded a dual megakaryoblastic and myeloid profile, thereby indicating that EG imposes megakaryoblastic as well as dual myeloid/megakaryocytic potential on single MPP3 progenitors.

Collectively, these data show that murine fetal liver HSPCs present a higher susceptibility to megakaryoblastic transformation by EG than ABM HSPCs. They also demonstrate that EG expression in ABM LT-HSCs results in an exclusively megakaryoblastic phenotype, whereas expression in further committed progenitors is required for myeloid transformation.

ETO2–GLIS2 Alters the LT-HSC Transcriptional Network in an Ontogeny-Dependent Manner

To compare the molecular consequences of EG expression in different developmental (fetal liver vs. ABM LT-HSC) or hierarchical (LT-HSC vs. MPP4 ABM cells) contexts, we performed single-cell RNA sequencing (scRNA-seq) on these purified populations after a 24-hour doxycycline induction in CTRL and iEG cells. Clustering and differential expression analyses revealed that iEG affected the FL LT-HSC transcriptome more than those from ABM LT-HSCs (3,772 vs. 151 differentially expressed genes, respectively) or MPP4 cells (only 3 differentially expressed genes; Fig. 5A; Supplementary S4A–S4C; Supplementary Tables S2–S4). Previously defined signatures of upregulated ("up") genes in patients with *ETO2–GLIS2*$^+$ AMKL (40, 41) were enriched in iEG$^+$ LT-HSCs derived from both fetal liver and ABM contexts. However, the signature of downregulated ("down") genes in patients with *ETO2–GLIS2*$^+$ AMKL was only deregulated in FL LT-HSCs (Fig. 5B; Supplementary Fig. S4D). iEG MPP4 cells showed few differentially expressed genes and no enrichment for EG$^+$ AMKL patient signatures (Supplementary Fig. S4D). These data show that the iEG signature obtained in FL LT-HSCs better recapitulates the molecular consequences observed in patients with EG$^+$ AMKL than the signatures observed in ABM LT-HSC or MPP4 cells.

Gene set enrichment analysis (GSEA) showed enrichment of stem-cell signatures in both fetal and adult iEG LT-HSCs (e.g., *Myct1, Esam*), significantly more pronounced for several genes in fetal LT-HSCs (e.g., *Vwf* and *Hlf*; Fig. 5C and D; Supplementary Fig. S4E and S4F). Expression signatures from progenitors and more mature populations (e.g., MPP3, MPP4, monocytes, neutrophils, megakaryocytes) were significantly downregulated in fetal liver LT-HSCs compared with adult LT-HSCs with a lower expression of myeloid (e.g., *Spi1*), erythroid (e.g., *Nfe2*), and megakaryocytic genes (e.g., *Mpl*; Fig. 5C and E; Supplementary Fig. S4E), supporting a more pronounced blockage of differentiation in fetal liver LT-HSCs.

To investigate related transcription factor (TF) activities, network inference and activity prediction with ARACNe and VIPER programs was used (Supplementary Fig. S5A; Supplementary Tables S5–S7). This first confirmed a higher expression level and activity of EG in iEG$^+$ cells on doxycycline (Supplementary Fig. S5B). For several key TFs, it also predicted significantly different activities without significant difference in mRNA expression (Supplementary Fig. S5C; Supplementary Tables S8 and S9). Accordingly, activity-based clustering confirmed that iEG expression affected adult LT-HSCs more than MPP4 cells (Supplementary Fig. S5D). Several TFs known to act as oncogenes or cofactors (e.g., *ERG*, *HLF*, *MECOM*, *HOXA10*, *MEIS1*) presented increased activities, whereas other key differentiation TFs were downregulated (*RBM15*, *SPI1*, *CEBPA*; Supplementary Fig. S5E).

To investigate the association between transcriptional changes and chromatin remodeling, we mapped open chromatin regions by assay for transposase-accessible chromatin with high-throughput sequencing (ATAC-seq) in fetal liver progenitors 24 hours after EG induction. Most open chromatin regions were common between CTRL and iEG (33,429 peaks), whereas some were lost/decreased (12,412 down peaks) or acquired/increased (5,975 up peaks) in iEG cells (Fig. 5F–I). Notably, 29% of the genes deregulated in scRNA-seq analyses showed differential ATAC-seq peaks around the genes' loci, suggesting that the transcriptional modifications induced by EG are partly due to chromatin remodeling (Fig. 5H). ATAC-seq also revealed differences between the CD41$^+$ and Gr1$^+$ iEG cell lines with Gr1$^+$ cells showing mostly chromatin closing as compared with WT FL (1,899 up and 16,383 down peaks), whereas CD41$^+$ cells showed more chromatin opening (34,167 up and 19,971 down peaks; Fig. 5I). Notably, the regions opened by iEG in fetal liver progenitors (5,975 up peaks) at 24 hours were also enriched in CD41$^+$ but not in Gr1$^+$ cells, suggesting a closer proximity between iEG fetal liver and CD41$^+$ than with myeloid transformed cells (Fig. 5J).

Motif analysis within the *ETO2–GLIS2*-specific gained open regions (up peaks) showed a lower representation of GATA motifs and an increased representation of ETS/ERG motifs (Fig. 5K) in fetal liver progenitors in CD41$^+$ and in Gr1$^+$ cell lines, compared with their representation within regions specific for the control cells (down peaks). Interestingly, motifs of some key myeloid transcription factors, including SPI1 and CEBP, were significantly enriched only in Gr1$^+$ transformed cells, supporting a role for these factors in the transcriptional regulation in myeloid *ETO2–GLIS2*$^+$ cells.

Together, these data show that EG imposes more prominent transcriptional changes in fetal liver–derived than in adult HSC, in part through rapid changes in chromatin accessibility to regions presenting binding sites for major hematopoietic transcription factors including GATA, ETS, and CEBP.

GATA1 and CEBPA Activities Correlate with Pediatric AML Phenotypes

HSC lineage specification involves TF expression or activity switches, including a SPI1/CEBPA module directing myeloid development (42, 43) and an ERG/GATA1 balance (44) regulating the megakaryocytic fate and altered in human AMKL (10). As these TFs are also involved in hematopoietic ontogeny (34, 35), we investigated TF expression and functional contribution to EG-induced megakaryoblastic and myeloid phenotypes. WT (CTRL) fetal liver HSPCs expressed more *Gata1, Erg*, and *Spi1* and less *Cebpa* than their ABM counterparts. Twenty-four hours after iEG induction, all genes were downregulated except *Erg* in FL HSPCs. However, iEG FL HSPCs expressed significantly higher levels of *Gata1* and *Erg* and lower levels of *Cebpa* and *Spi1* than iEG ABM HSPCs (Fig. 6A). Notably, *Spi1* was more strongly downregulated by iEG in fetal liver (4.9-fold) than in adult (2.2-fold) progenitors, consistent with scRNA-seq data performed in purified LT-HSCs (Fig. 5E).

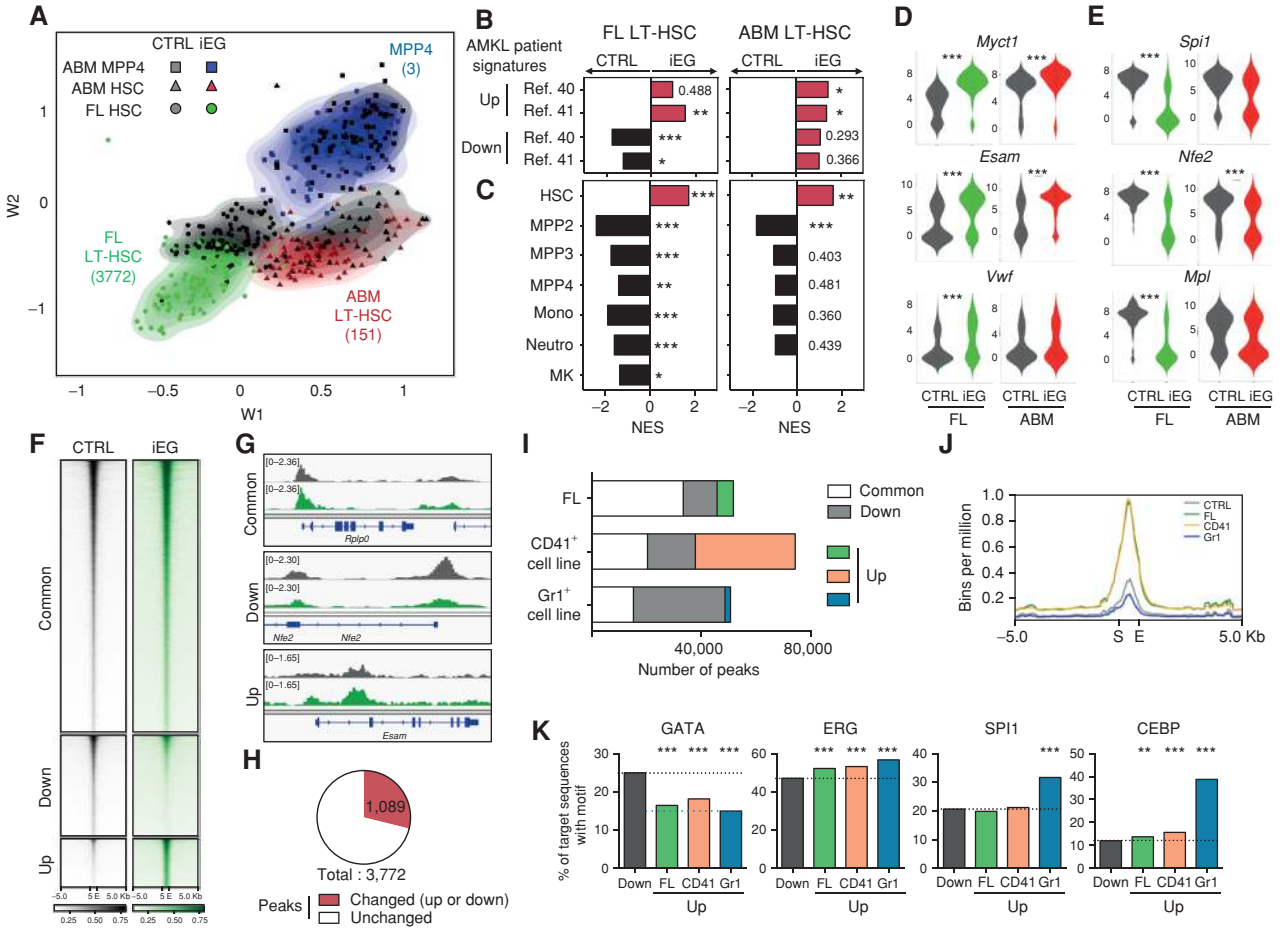


Figure 5. *ETO2-GLIS2* alters the LT-HSC transcriptional network in an ontogeny-dependent manner. **A,** Clustering of the different populations based on scRNA-seq data using Zinbwave. The number of genes significantly differentially expressed between CTRL and iEG in each population is indicated. **B,** GSEA on iEG FL LT-HSCs (left) or ABM LT-HSCs (right) versus CTRL scRNA-seq data using published EG$^+$ AMKL patients' specific gene-expression signatures from Bourquin and colleagues (40) or Gruber and colleagues (41). Normalized enrichment score (NES) in iEG versus CTRL cells are represented (positive NES indicates higher activity in iEG vs. CTRL cells; negative NES indicates lower activity in iEG vs. CTRL cells). Up, signature of genes upregulated in patients with EG$^+$ AMKL; down, signature of genes downregulated in patients with EG$^+$ AMKL. *, FDR < 0.25; **, FDR < 0.05; ***, FDR < 0.01. **C,** GSEA using published signatures from normal progenitors and mature cells (38, 77): HSC, MPP2, MPP3, MPP4. Mono, monocyte; neutro, neutrocyte; MK, megakaryocyte. **D,** Violin plots of *Myct1*, *Esam*, and *Vwf* expression in FL cells and ABM LT-HSCs. Log_2 values of the normalized read counts are represented. **E,** Violin plots of *Spi1*, *Nfe2*, and *Mpl* expression in FL cells and ABM LT-HSCs. Log_2 values of the normalized read counts are represented. **F,** Heat map showing hierarchical clustering centered on open chromatin peaks identified by ATAC-seq analysis in CTRL and iEG FL progenitors (LT-HSC+MPP1+MPP2) after 24 hours of doxycycline (Dox) induction. Common, peaks unchanged between CTRL and iEG; down, peaks lost or significantly reduced in iEG; up, peaks gained or significantly enriched in iEG. Windows of 11 kb are shown. S, peak start; E, peak end. **G,** Illustrations of ATAC-seq peaks in CTRL and iEG FL progenitors of representative common (*Rplp0*), down (*Nfe2*), and up (*Esam*) peaks using Integrative Genomics Viewer software. **H,** Pie chart illustrating the fraction of significantly deregulated genes in scRNA-seq iEG versus CTRL FL-HSCs in which a modified EG peak (either up or down) is present. The number of genes with a deregulated peak is indicated. **I,** Histogram representation of the number of common, up, and down peaks between iEG FL progenitors, CD41$^+$ cell line, or Gr1$^+$ cell line and CTRL FL progenitors. **J,** Intensity profiles of up peaks in FL progenitors described in **F** with CTRL and iEG fetal liver progenitors, CD41$^+$ and Gr1$^+$ ATACseq data. **K,** Motif analysis under the peaks enriched in CTRL versus iEG (down peaks) in fetal liver progenitor (peaks down in iEG FL progenitors, dark gray), in iEG versus CTRL (peaks up in iEG FL progenitors, green), in megakaryoblastic CD41$^+$ (CD41, orange) and in myeloid Gr1$^+$ (Gr1, blue) iEG cell lines. Chi^2 analysis: GATA: FL vs. down $P = 3.46e^{-143}$; CD41 vs. down $P = 3.67e^{-316}$; Gr1 vs. down $P = 6.77e^{-68}$; ERG: FL vs. down $P = 1.19e^{-10}$; CD41 vs. down $P = 6.5e^{-32}$; Gr1 vs. down $P = 2.31e^{-15}$; PU.1: FL vs. down $P = 0.21$; CD41 vs. down $P = 0.194$; Gr1 vs. down $P = 5.39e^{-28}$; CEBP: FL vs. down $P = 1.12e^{-3}$; CD41 vs. down $P = 1.49e^{-22}$; Gr1 vs. down $P = 2.5e^{-201}$. Statistical significance is indicated as *P* values (Student *t* test). *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

We then quantified expression of these TFs in CTRL and iEG ABM LT-HSC and MPP4 cells. *Erg* and *Gata1* were more highly expressed in iEG LT-HSC–derived cells than in iEG MPP4–derived cells, whereas *Cebpa* and *Spi1* were highly expressed in MPP4 cultures (Fig. 6B). We then functionally assessed whether increasing ERG or GATA1 activity may impose a megakaryoblastic phenotype in iEG MPP4 cells and whether increasing SPI1 or CEBPA activity biased iEG LT-HSCs toward a myeloid phenotype (Fig. 6C–E). Retroviral *GATA1*, and to a lesser extent *ERG*, expression enhanced the megakaryocyte phenotype in iEG MPP4 cells. In iEG LT-HSCs, expression of *CEBPA*, but not *SPI1*, enhanced the myeloid phenotype.

We then compared *ERG, GATA1, SPI1*, and *CEBPA* expression levels in leukemic blasts from patients with AMKL and

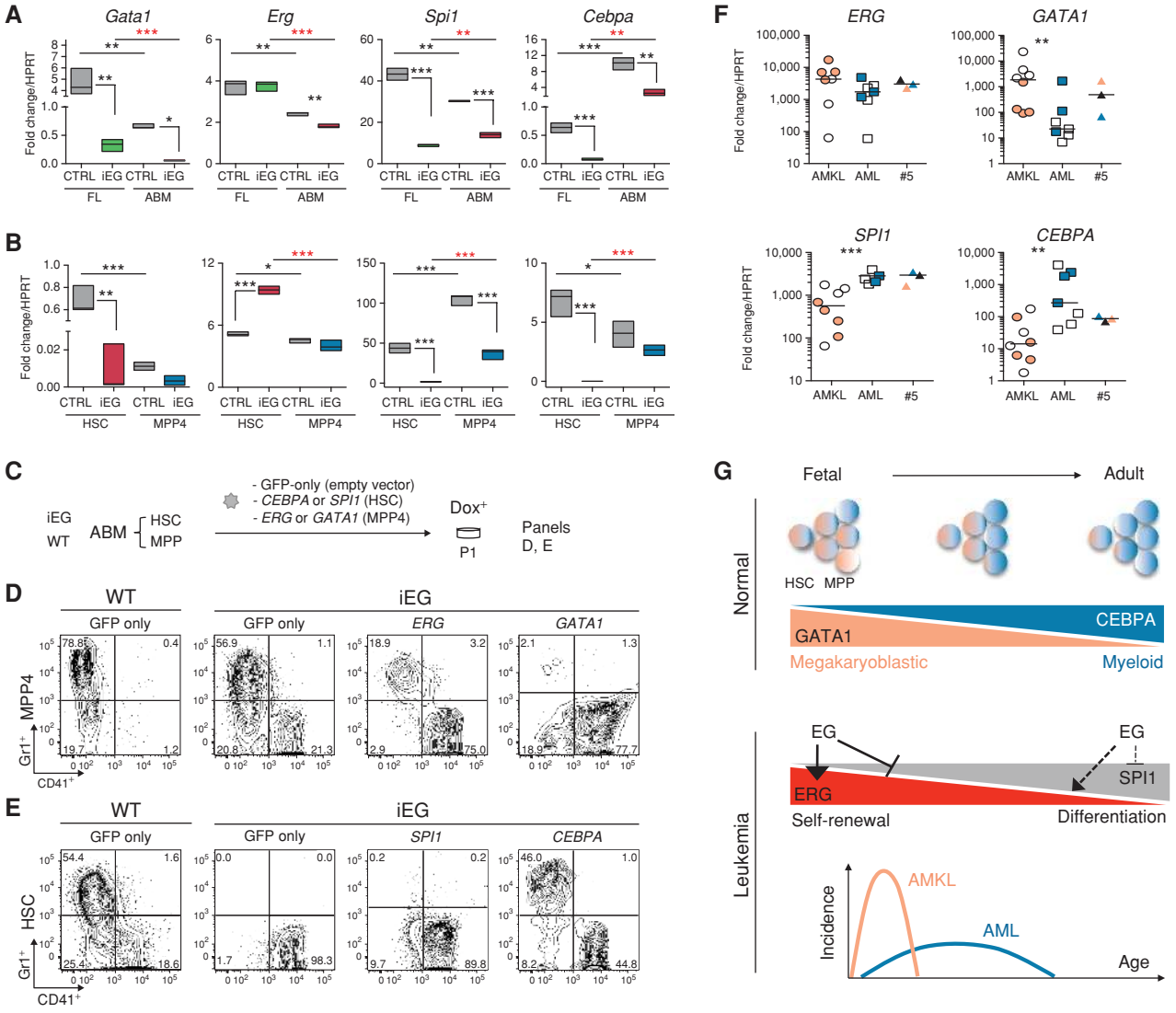


Figure 6. GATA and CEBPA activities functionally control EG-induced phenotypes. **A,** Quantitative RT-PCR analyses in Lin⁻ CTRL and iEG FL14.5 and ABM-derived cells after 24 hours in liquid culture with doxycycline. Mean ± SEM ($n = 3$) is shown. **B,** Quantitative RT-PCR analyses in CTRL and iEG ABM LT-HSC and MPP4-derived cells at P1. Mean ± SEM ($n = 3$) is shown. **C,** Experimental design to ectopically express indicated TFs in iEG ABM LT-HSC and MPP4 cells. **D,** Representative flow-cytometry analyses of cells obtained at P1 after transduction of WT and iEG MPP4 cells with *ERG*, *GATA1*, or GFP-only retroviruses. Analysis was gated on GFP⁺ cells. **E,** Representative flow-cytometry analyses of cells obtained at P1 after transduction of WT and iEG LT-HSCs with *SPI1*, *CEBPA*, or GFP-only retroviruses. Analysis was gated on GFP⁺ cells. **F,** Quantitative RT-PCR analyses in blasts from patients with AMKL ($n = 8$, EG⁻ AMKL: empty circles; EG⁺ AMKL: orange circles), AML ($n = 7$, EG− AMKL: empty squares; EG+ AMKL: blue squares) and in the different blast populations from the dual phenotype *ETO2-GLIS2*⁺ AML patient (#5) presented in Fig. 1H (CD41⁻CD15⁻ immature: black triangle; CD41⁺CD15⁻ megakaryoblastic: orange triangle; CD41⁻CD15⁺ myeloid: blue triangle). Statistical significance is indicated as *P* values (Mann–Whitney test). *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$. **G,** Schematic representation of the bases for the pediatric and phenotypic specificities associated with *ETO2-GLIS2*⁺ leukemia. Our data and published studies (22–24, 26) support the idea that a shift from a higher megakaryoblastic potential (orange color) to a higher myeloid potential (blue color) in fetal versus adult hematopoiesis is associated with a change in the GATA1 versus CEBPA activities. Here we show that *ETO2-GLIS2* imposes a stronger deregulation of self-renewal programs in fetal compared with adult hematopoiesis (including higher ERG and lower SPI1 activities). The combination of the basal transcriptional state controlling phenotypic output with a differential effect of *ETO2-GLIS2* on self-renewal programs during hematopoietic ontogeny may therefore underlie the strong prevalence of AMKL in early childhood and more generally represent a basis for the specific pediatric and phenotypic incidence of human AMKL versus other myeloid leukemia (AML; bottom). Statistical significance is indicated as *P* values (Student *t* test) except otherwise mentioned). *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

AML and in different populations from the dual AML/AMKL phenotype patient (#5; Fig. 6F; Supplementary Fig. S6A). AMKL cases showed high *ERG* and *GATA1* and low *SPI1* and *CEBPA* expression, whereas AML cases presented an opposite expression pattern. Similar profiles were observed in the human EG⁺ cell lines MO7e (AMKL) and CMS (AML; Supplementary Fig. S1F). The EG⁺ case with a dual phenotype (#5) showed an intermediate expression of *CEBPA*, *GATA1*, and *ERG* compared with patients with AMKL and AML (Fig. 6F). In this patient, the most immature CD34⁺CD41⁻CD15⁻ blasts showed the highest *ERG* expression compared with CD41⁺ and CD15⁺ blasts, while the CD41⁺ blasts showed the

highest *GATA1* expression and the CD15$^+$ blasts showed the highest *SPI1* expression levels. Notably, CD41$^+$ blasts showed an unexpected combination of *GATA1*^{hi} and *CEBPA*^{hi} expression (Supplementary Fig. S6A), the significance of which for this peculiar case of dual megakaryoblastic/myeloid phenotype remains to be understood.

Collectively, these data support the idea that the specific association between EG$^+$ AMKL and young children results from a combination of at least two parameters (Fig. 6G). First, intrinsic differences in the GATA1 and CEBPA activities observed during the normal ontogeny (FL HSC vs. ABM HSC) and in the normal HSPC hierarchy (HSC vs. MPP) are conserved upon EG expression and correlate with the megakaryoblastic/myeloid phenotypic output, respectively. Second, EG induces a stronger deregulation of self-renewal–associated genes (e.g., *ERG* and *SPI1*) in the fetal versus adult HSC.

ETO2–GLIS2 Reversibly Transforms LT-HSCs

As our cellular and molecular observations indicate that EG hijacks HSC properties, we compared the *in vivo* leukemogenic potential of purified cell populations and assessed its potential reversibility by tracking the differentiation potential of iEG-transformed cells (carrying the GFP reporter gene) after doxycycline removal.

GFP$^+$ FL LT-HSCs and MPP3 cells from E14.5 iEG embryos, or LT-HSC and MPP cells from iEG ABM, were transplanted into doxycycline-treated WT recipients (Fig. 7A). The iEG FL LT-HSC and MPP3 recipients developed fully penetrant and rapid (median: 27 days and 33 days, respectively) lethal leukemia (Fig. 7B). The iEG ABM LT-HSC recipients developed disease after a strikingly longer latency (median: 312 days, $P = 0.017$; Fig. 7B and C; Supplementary Fig. S6B). Only 1 of 9 iEG ABM MPP3 recipients developed disease with a long latency and none of the iEG ABM MPP4 recipients developed any disease (Fig. 7C; Supplementary Fig. S6B). The fetal liver iEG LT-HSC and MPP3 recipients presented with a more severe disease than the ABM iEG LT-HSC mice as assessed by greater degree of splenomegaly, higher WBC counts, higher percentage of CD41$^+$ cells and lower platelet counts, maturing myeloid populations, and lymphoid cells in BM and spleen (Fig. 7D–G). Histologic analyses showed that the BM, spleen, and liver were infiltrated with GFP$^+$ cells that were weakly VWF$^+$. The iEG ABM LT-HSC recipients displayed rare maturing VWF$^+$ megakaryocytes that were all GFP$^-$, which was not observed in the iEG FL LT-HSC counterparts (Supplementary Fig. S6C; data not shown).

To investigate the differentiation potential of LT-HSC–derived iEG leukemic cells, we purified the immature CD41$^+$CD11b$^-$Gr1$^-$ blasts from iEG fetal liver LT-HSC diseased recipients and grew them in cultures in absence of doxycycline. As compared with the immature blasts maintained with doxycycline, they generated both maturing CD41$^+$CD42$^+$ megakaryocytes and CD11b$^+$Gr1$^+$ myeloid cells (Supplementary Fig. S6D). This was associated with reduced expression of iEG and its known target *Erg*, whereas *Gata1* expression increased, consistent with broad differentiation capacities (Supplementary Fig. S6E). To assess long-term hematopoietic potential, BM cells from the ABM LT-HSC–engrafted diseased recipients were transferred into secondary doxycycline-treated recipients (Fig. 7H). Two weeks post-transplant (= D0 time point), when recipients showed the first signs of disease (Supplementary Fig. S6F and S6G), some recipients were maintained on doxycycline (group A) whereas doxycycline treatment was ceased in group B. At three days after doxycycline removal, the platelet counts were back to normal in group B; 80% of platelets originated from iEG-transformed cells as shown by GFP expression (Fig. 7I). GFP$^+$ cells were also detected later in red blood cells (RBC) and WBCs (Fig. 7I; Supplementary Fig. S6H). Notably, doxycycline-treated mice (group A) developed lethal leukemia similar to primary mice whereas group B off doxycycline did not (Dox$^+$, Fig. 7J). Analysis of group B at 6 months post doxycycline removal demonstrated that GFP$^+$ cells were still present in all analyzed hematopoietic compartments, including mature B and T cells (Fig. 7K; Supplementary Fig. S6I). Together, these data show that iEG expression in FL LT-HSCs results in a rapid and aggressive AMKL *in vivo* associated with maintenance of long-term HSC-like differentiation potential once the driver oncogene is switched off.

MLL–AF9 Expression in Fetal Liver LT-HSCs Induces Megakaryoblastic Features

In addition to *ETO2–GLIS2*, other fusion oncogenes, including *MLL–AF9*, are found in both pediatric AML and AMKL. However, published murine models only displayed *MLL–AF9*-driven myeloid, but not megakaryoblastic leukemic phenotypes. To assess whether the ontogenic stage and cell hierarchy also affect the phenotype induced by *MLL–AF9*, we retrovirally expressed *MLL–AF9* (*rMLL–AF9*) in fetal liver- and ABM-derived HSPCs from WT mice. Notably, *rMLL–AF9* resulted in an increased CD41$^+$ megakaryoblastic phenotype in fetal liver compared with ABM cells (Supplementary Fig. S7A). Next, we transduced purified fetal liver LT-HSC and MPP populations and observed increased CD41$^+$ output only with LT-HSCs (Supplementary Fig. S7B). Transplantation of *rMLL–AF9*-transduced cells revealed disease development in both cases, presenting with clear CD41$^+$ and VWF$^+$ megakaryoblastic features only in recipients of FL LT-HSCs, but not MPP4 cells (Supplementary Fig. S7C and S7D). *rMLL–AF9*-expressing CD41$^+$ blasts also expressed high *Gata1* and low *Spi1* and *Cebpa* expression, whereas MPP4-derived myeloid blasts showed a mirrored expression pattern (Supplementary Fig. S7E). Collectively, these observations indicate that ontogenic changes in the cellular hierarchy affect disease phenotypes not only for *ETO2–GLIS2*, but also for other AMKL-associated fusion oncogenes such as *MLL–AF9*.

DISCUSSION

In toto, our observations in patient samples and in inducible transgenic mouse models for pediatric leukemia–associated fusion oncogenes, including *ETO2–GLIS2* and *MLL–AF9*, provide experimental evidence that aggressive pediatric AMKL originates in fetal HSPCs and, more globally, that age- and phenotype-specific associations observed in human patients rely on changes in the cellular architecture during hematopoietic ontogeny.

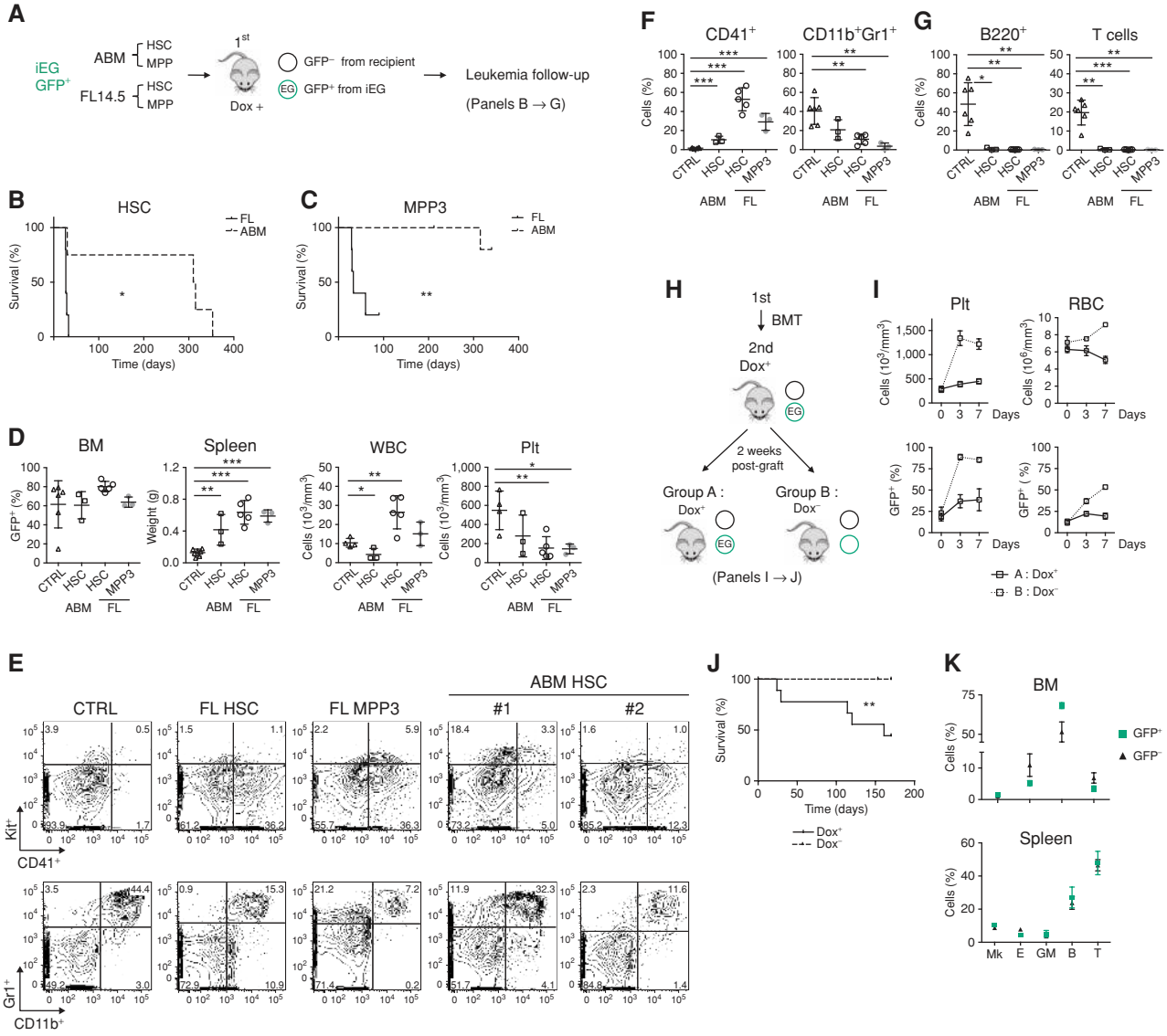


Figure 7. *ETO2-GLIS2* reversibly transforms LT-HSCs. **A,** Experimental design: purified LT-HSC and MPP cells from WT or iEG mice expressing the Ubiquitin-GFP (GFP$^+$) were transplanted into WT doxycycline (Dox)-treated recipients. Note that GFP expression is not doxycycline-dependent and therefore engraftment of all cells can be followed using GFP regardless of the doxycycline treatment. **B,** Kaplan–Meier survival curve of recipients of naïve iEG LT-HSCs from ABM or fetal liver, on doxycycline. Recipients were treated with doxycycline after transplantation. Log-rank Mantel–Cox test iEG HSC FL versus ABM recipients: $P = 0.0171$. Median survival: HSC ABM: 312 days; HSC FL: 27 days. **C,** Kaplan–Meier survival plot of recipients of naïve iEG MPP3 cells from ABM or FL kept on doxycycline. Log-rank Mantel–Cox test iEG MPP3 fetal liver versus ABM recipients: $P = 0.0010$. Median survival of MPP3 FL: 33 days. **D,** Percentages of GFP$^+$ cells in BM, spleen weights, WBC counts, and platelet (Plt) counts in CTRL and iEG ABM HSC recipients and iEG FL HSC and MPP3 recipients. Spleen (CTRL vs. FL: $P < 0.0001$; CTRL vs. ABM: $P = 0.0039$), WBC (CTRL vs. FL: $P = 0.0095$; CTRL vs. ABM: $P = 0.0309$), Plt (CTRL vs. fetal liver: $P = 0.0079$; CTRL vs. ABM: $P = 0.1536$). **E,** Representative flow-cytometry analysis of diseased recipients at time of sacrifice. **F,** Histogram of the percentages of CD41$^+$ and CD11b$^+$Gr1$^+$ cells in the BM of primary diseased recipients. CD41$^+$ (CTRL vs. FL: $P < 0.0001$; CTRL vs. ABM: $P = 0.0005$), CD11b$^+$Gr1$^+$ (CTRL vs. FL: $P = 0.0016$; CTRL vs. ABM: $P = 0.0717$). **G,** Histogram of the percentages of B220$^+$, CD4$^+$, and CD8$^+$ cells in the spleen of primary diseased recipients. B220$^+$ (CTRL vs. FL: $P = 0.012$; CTRL vs. ABM: $P = 0.0102$), T cells = CD4$^+$ or CD8$^+$ cells (CTRL vs. FL: $P < 0.0001$; CTRL vs. ABM: $P = 0.0016$). **H,** Experimental design: 1×10^6 BM cells from the diseased primary recipients at time of analysis were transplanted into secondary sublethally irradiated doxycycline-treated recipients. After two weeks, doxycycline treatment was maintained for group A (GFP$^+$ cells expressing iEG) and ceased in group B (GFP$^+$ cells lacking iEG expression). **I,** Plt and RBC counts (top) and respective percentages of GFP$^+$ cells (bottom) in secondary recipients of ABM LT-HSC–derived disease. D0: time at which doxycycline was maintained in group A and withdrawn in group B. Mean ± SEM ($n = 6$/group) is shown. **J,** Kaplan–Meier survival plot of secondary recipients transplanted with cells from diseased primary iEG ABM LT-HSC recipients. Secondary recipients were kept on doxycycline (Dox$^+$) or not (Dox$^-$). Log rank Mantel–Cox test of Dox$^-$ versus Dox$^+$ recipients: $P = 0.0086$. **K,** Flow-cytometry analyses in secondary recipients of iEG ABM LT-HSC–derived disease at 6 months after doxycycline withdrawal. Analyses were gated on GFP$^+$ (iEG donor-derived) or GFP$^-$ (WT recipient) cells. Statistical significance is indicated as P values (Student t test except when otherwise specified). *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

A novel inducible transgenic mouse model allowed us to demonstrate that expression of EG in fetal progenitors and adult LT-HSCs, but not in further committed adult progenitors, resulted in leukemogenesis *in vivo*. The previous lack of *in vivo* transformation (41, 45) could therefore be due to a poor targeting of LT-HSCs by retroviral transduction. The inducible model was also instrumental in demonstrating that HSC-derived EG$^+$ leukemic cells can maintain long-term reconstitution and differentiation properties upon withdrawal of the driver oncogene. In adult hematopoiesis, several cellular and gene expression studies have indicated the close proximity between LT-HSCs and megakaryocytic states in normal and stress conditions (22–24, 29, 46, 47), thus prompting the hypothesis that megakaryopoiesis may represent the native differentiation pathway for LT-HSCs (24). Our data indicate that EG alters master hematopoietic TF activities in LT-HSCs (e.g., ERG, HLF, MECOM, and HOX-related factors) toward a megakaryoblastic state and therefore strongly suggest that EG$^+$ AMKL is mostly an LT-HSC disease. In addition, expression of *MLL–AF9* in fetal LT-HSC led to clear megakaryoblastic features as opposed to fetal MPP or adult progenitors that generated the typical *MLL–AF9* myeloid leukemia. Also, other AMKL fusions involve transcription regulators important for both normal HSC and megakaryocyte biology (e.g., OTT1/RBM15; ref. 48). Therefore, we propose a conceptual framework for pediatric AMKL in general in which transcriptional constraints, imposed by distinct oncogenes, represent molecular switches that functionally converge on an early alteration of fetal LT-HSCs as they engage the native megakaryocytic differentiation pathway.

This study also demonstrates that the transforming potential of pediatric oncogenes is dependent on the ontogeny-related dynamics of the hematopoietic tissue architecture. Indeed, in our model, iEG expression in fetal LT-HSCs generated an aggressive and phenotypically homogeneous megakaryoblastic phenotype, whereas expression in adult LT-HSCs resulted in the development of more heterogeneous phenotypes with delayed latencies. This is consistent with recent findings that normal fetal hematopoiesis is primarily controlled by multipotent HSPCs (endowed with megakaryocytic potential), whereas adult hematopoiesis presents with a higher ratio of committed progenitors (lacking megakaryocytic potential; ref. 26). Molecularly, our data indicate that, although *GATA1* and *CEBPA* expression are both low in EG$^+$ leukemia, their relative activities may contribute to the megakaryoblastic versus myeloid phenotype of EG$^+$ leukemia, respectively. EG imposed higher *Erg* and lower *Spi1* levels in fetal than in adult LT-HSCs, but engineered expression of these factors did not drastically affect the EG-induced phenotype *in vitro*. Notably, the precise expression level of *Erg* and *Spi1* controlled self-renewal and transformation in other models (49–51). Together, we propose that the association between EG$^+$ AMKL and the young age of the patients results from the combination of (i) intrinsic differences between normal fetal and adult hematopoiesis together with (ii) a stronger alteration of fetal HSC transcriptional regulation by EG, leading to a higher susceptibility of fetal HSC to transformation by EG (Fig. 6G). The former includes a higher proportion of multipotent versus myeloid-committed progenitors and higher GATA1 versus CEBPA activity in fetal as compared with adult stages. The latter includes a more prominent deregulation by EG of a self-renewal module (e.g., ERG and SPI1). As low CEBPA activity was also associated with increased competitive advantage and proliferation (35, 52), the lower *CEBPA* expression in the EG$^+$ fetal context may also contribute to the more aggressive leukemia obtained upon EG expression in fetal HSCs.

It is likely that the importance of this higher ontogeny–related susceptibility extends to the other human megakaryocyte/platelet-related diseases (e.g., Trisomy 21, AMKL, and thrombocytopenia absent radii syndrome; refs. 53–55) but also to other pediatric leukemias, including lymphoid leukemia with *MLL–AF4* (56), whose prevalence peaks during childhood. Furthermore, multiple pediatric cancers are driven by strong epigenetic fusion oncogenes in the context of a low mutational burden, including Ewing sarcoma and rhabdomyosarcoma, for which the cellular origin is poorly defined (13), and glioblastoma, for which oncogenic drivers have been recently shown to arise in stem cells (57). Therefore, a precise characterization of the changes in the cellular architecture of the relevant tissues of origin should improve our understanding of the heterogeneous clinical presentations in other pediatric cancers and facilitate the establishment of other disease models that may serve as experimental platforms to develop more selective therapeutic strategies.

METHODS

Reagents and Resources

Supplementary Table S10 contains an extensive list of the relevant reagents.

Patient-Derived Data and Material

Patient blood and BM samples were obtained with the written informed consent of the patient or their family in accordance with the Declaration of Helsinki, and the study was approved by the Gustave Roussy Institutional Review Board. Samples were subjected to Ficoll gradient, immunophenotyped, and used fresh or frozen in FBS supplemented with 10% DMSO for subsequent use. DNA and RNA were obtained from the mononuclear cell fractions using the manufacturer's recommendations (Qiagen). cDNA from 284 patients was available to detect the EG fusion via RT-PCR. The primer sequences are provided in Supplementary Table S11.

Clinical data were obtained from the ELAM02 protocol. This study focused on 385 of 438 children treated in the ELAM02 trial (Treating Patients with Childhood Acute Myeloid Leukemia with Interleukin-2; ClinicalTrials.gov NCT00149162; ref. 37). Patient selection was based on the availability of genomic DNA and RNA upon AML diagnosis. Children ages 0 to 18 years newly diagnosed with AML were enrolled between March 2005 and December 2011. Patients diagnosed with acute promyelocytic leukemia, therapy-related AML, or Down syndrome were excluded from the ELAM02 trial. The study was approved by the Ethics Committee of Saint-Antoine Paris University Hospital (Assistance Publique-Hôpitaux de Paris) and by the Institutional Review Board of the French Regulatory Agency. The study was conducted in accordance with the declaration of Helsinki. Data concerning age at diagnosis was also collected from the ELAM02 protocol and ref. 9, including 87 patients with AMKL.

Patient-Derived Xenotransplantation

Patient-derived xenograft protocols have been described previously (8).

FISH

The DNA probe set obtained from Empire Genomics (CBFA2T3 and GLIS2 FISH probes, AmpliTech) hybridizes to chromosome 16p13.3 at *CBFA2T3* locus (red) and 16q24.3 at *GLIS2* (green) in normal metaphase spreads or interphase nuclei. FISH analysis was performed following the manufacturer's protocol. Interphase nuclei and metaphases were scored on Imager Z2 fluorescence microscope (Zeiss) using Isis/Metafer/Relosys MetaSystems Software (MetaSystems). In cells with the EG fusion, the overlap between the two probes generates a yellow signal indicative of the fusion oncogene.

Cell Lines

The human AMKL MO7e cells (ref. 58; derived from a 6-month-old girl) were cultured in α-MEM supplemented with 20% FBS, penicillin (100 U/mL)–streptomycin (100 μg/mL), 2 mmol/L L-glutamine (Gibco), and 5 ng/mL of human GM-CSF (PeproTech).

The human CMS cells (ref. 59; derived from a 2-year-old girl) were cultured in RPMI-1640 supplemented with 10% FBS, penicillin (100 U/mL)–streptomycin (100 μg/mL), and 2 mmol/L L-glutamine (Gibco).

The human AML-M6 HEL 92.1.7 (HEL) cells (ref. 60; derived from a 30-year-old patient) were cultured in RPMI-1640 supplemented with 10% FBS, penicillin (100 U/mL)–streptomycin (100 μg/mL), and 2 mmol/L L-glutamine (Gibco).

All cell lines were obtained from the laboratory of O. Bernard in 2009, were karyotyped for authentication in 2018, and were tested for *Mycoplasma* in 2017, and the described experiments were performed within 3 months after thawing.

Mice

To generate the *iETO2–GLIS2:rtTA* (iEG) model, the human *EG* cDNA was cloned into the *p2Lox* vector and electroporated into A2Lox-Cre embryonic stem (ES) cells, a generous gift from Michael Kyba (61). ES cell clones were selected for the correct integration of EG at the *Hprt* locus on chromosome X and used to obtain germline-transmitting chimeras that were subsequently back-crossed into the C57BL/6 background. The rtTA and the *EG* transgene were genotyped using the primers indicated in Supplementary Table S11.

To obtain GFP-expressing iEG mice and track the iEG cells upon *in vivo* transplantations, we crossed *iETO2–GLIS2:rtTA* mice with the *Ubiquitin-GFP* reporter line (UBI-GFP) mice (62). In these mice, GFP expression is controlled by a ubiquitin promoter and not controlled by doxycycline. Therefore, engraftment of all CTRL and iEG donor cells can be tracked using GFP regardless of the doxycycline treatment status or iEG expression.

The transgenic mouse line for inducible expression of *MLL–AF9* has been described previously (25).

The *iETO2–GLIS2* and *iMLL–AF9* transgenes were induced with doxycycline provided in impregnated food pellets (doxycycline food pellets, 545 mg/kg, provider: SSNIF), which is predicted to correspond to an approximate dose of 2.2 mg/day.

Mice were maintained at the Gustave Roussy preclinical facility and all experiments were approved by the French National Animal Care and Use Committee (CEEA 26: #2012-017, #201712180922196_v2 and #2017122111548235_v2).

Bone Marrow Transplantation

Transplantation of BM (10^6) or sorted progenitor cells was performed through intravenous injection in lethally (9.5 Gy) or sublethally (5 Gy) irradiated 8-week-old C57BL/6 recipient mice.

Morphologic analysis of peripheral blood, BM, and spleen cells as well as histologic analyses were performed according to standard procedures.

Flow Cytometry

Cells analyzed by flow cytometry were antibody-stained in PBS 1× supplemented with 2% FBS for 30 minutes at 4°C with washing prior to FACS analysis. Whole BM and spleen cells from mice were collected, subjected to RBC lysis, and stained for FACS analysis.

To obtain HSPCs, BM cells were depleted from the mature hematopoietic cells (Lin^-) with the Becton Dickinson (BD) Mouse Hematopoietic Progenitor (Stem) Cell Enrichment Set (558451).

To obtain ABM precise progenitors, populations were further purified by FACS according to the following phenotypes: LT-HSCs (HSCs) were defined as $Lin^-Kit^+Sca1^+CD135^-CD34^-CD48^-CD150^+$, MPP2 cells were defined as $Lin^-Kit^+Sca1^+CD135^-CD34^+CD48^+CD150^+$, MPP3 cells were defined as $Lin^-Kit^+Sca1^+CD135^-CD34^+CD48^+CD150^-$, and MPP4 cells were defined as $Lin^-Kit^+Sca1^+CD135^+CD34^+CD48^+CD150^-$.

To obtain fetal liver HSPCs, cells were depleted with biotin-conjugated antibodies against CD3, Ter119, B220, Gr1, NK1.1, F4/80, and CD19.

To obtain FL HSC and progenitor populations, cells were purified by FACS according to the following phenotypes: LT-HSCs were defined as $Lin^-Kit^+Sca1^+CD135^-CD48^-CD150^+$, whereas MPP2, MPP3, and MPP4 cells were defined with the same markers as ABM cells.

Flow cytometry analyses were performed with ARIAII, ARIAIII, CANTO-II, or CANTO-X instruments (BD Biosciences), and data were analyzed using FlowJo software (version Flowjo 9.3.2).

Serial Replating Assays

Total cells and/or progenitors from ABM or fetal liver at E12.5 or E14.5 were isolated from iEG mice and plated in 3 mL methylcellulose (Methocult M3434; StemCell Technologies), supplemented with IL3, IL6, SCF, EPO, and 300 ng/mL doxycycline. Colonies were scored under the microscope after 5 to 7 days, harvested, and replated (10^4 cells) for four rounds.

Single-Cell Cultures

Single cells were seeded in 96-well plates in 100 μL RPMI supplemented with 10% FBS, cytokines (mIL3, mIL6, mIL11, mSCF, mTPO, mEPO, mGM-CSF, mFLT3l) and 100 ng/mL doxycycline. Wells were scored for clonal growth after four days of culture, and frequency of megakaryoblastic and myeloid cells was assessed by surface marker analysis using a previously described method (63) after 5 to 10 days.

RNA Extraction and Quantitative RT-PCR Analysis

mRNA was isolated using an RNeasy Mini/Micro Kit (Qiagen) and quantified using a NanoDrop (Thermo Fisher Scientific). Reverse transcription was carried out with SuperScript II (Invitrogen). qPCR was performed with SYBR Select Master Mix or Taqman Gene Expression Master Mix (Applied Biosystems) using a 7500HT Fast Real-Time PCR System (Applied Biosystems) according to the manufacturer's recommendations. The primer sequences are provided in Supplementary Table S11.

ETO2–GLIS2 Genomic DNA Fusion Point

DNA were isolated using an All Prep DNA/RNA Mini/Micro Kit (Qiagen) and quantified using a NanoDrop (Thermo Fisher Scientific). PCR was performed using a GeneAmp PCR System 9700. The primer sequences are provided in Supplementary Table S11.

Immunoblotting

Total cell lysates were prepared in lysate buffer containing 50 mmol/L Tris-HCl (pH 8), 150 mmol/L NaCl, 1% NP40, 1 mmol/L EDTA, 0.1% SDS, 0.5% sodium deoxycholate, and protease inhibitors (phenylmethylsulfonyl fluoride 1×, NaF 1×, Sodium orthovanadate 1×, complete 1×). Western blotting was then performed

using standard procedures using anti-ETO2 (ab33072, Abcam), anti-HSC70 (sc7298, Santa Cruz Biotechnology), and anti-RPLP0 (ab192866, Abcam) antibodies.

Production of Retroviral Particles and Transduction

For retroviral transduction experiments, we constructed murine stem cell virus backbone constructs expressing GFP and encoding human *ETO2–GLIS2, ERG, CEBPA, SPI1,* and *GATA1* open reading frame. Briefly, 293T cells were transfected using Xtreme GENE transfection reagent according to the manufacturer's instructions (Roche). Retroviral particles were harvested 24 and 48 hours after transfection. For retroviral transduction, murine progenitor cells were spinoculated twice with retroviral particles in culture medium containing 5 μg/mL polybrene and 7.5 mmol/L HEPES buffer for 90 minutes at 2,500 rpm and 33°C.

scRNA-seq

Progenitor cells were isolated from iEG mice and cultured for 24 hours in RPMI supplemented with 10% FBS, cytokines (mIL3, mIL6, mSCF, mTPO, mFLT3l) and 100 ng/mL doxycycline. scRNA-seq analysis was performed as described previously (28, 64). Single cells were individually sorted in 96-well plates containing lysis buffer (0.2% Triton X-100 and 500 U/mL Superase-In RNase Inhibitor). The Illumina Nextera XT DNA Preparation Kit was used to prepare libraries. Pooled libraries were sequenced using the Illumina HiSequation 4000 system (single-end 125 base pair reads).

Statistical Analyses

Statistical analyses were performed using the GraphPad Prism software (version Prism 6-2) except otherwise mentioned.

Bioinformatics Analyses

Reads Quality Control and Alignment. Quality control of reads was performed using FastQC 0.11.7 and multiQC 1.5.dev0. The reads were aligned to the reference genome mm10 GRCm38.p4 81 with GSNAP (v 2015-09-29) with the following parameters: -A sam -B 5 -t 23 -n 1 -Q -N 1. Finally, gene expression was quantified using HTSeq 0.6.0

Counts Preprocessing, Quality Controls, and Normalization. Prior to data normalization, cells were filtered to remove low-quality samples and to fulfil the following conditions: have at least 4,000 detectable genes, have less than 20% of counts mapping to mitochondrial genes, have less than 50% of reads mapping to spike-in controls and at least 2^5 counts that map to nuclear genes. Normalization was next performed by library size with the deconvolution of size factors method described by Lun and colleagues using scran and scater R libraries (65, 66). Preclustering of cells and five pool sizes from 20 to 100 (by an increment of 20) were used as normalization parameters. Highly variable gene (HVG) selection was performed as described previously (67) for feature selection. Those HVGs were used for t-SNE to explore possible underlying substructures in the data.

Differential Expression Analysis and GSEA. For detection of differences at the transcriptomic level between groups, differential expression analysis was performed with Wilcoxon test over the $\log_2$ counts values and *P* value correction by Benjamini–Hochberg method (FDR cutoff at 0.05). GSEA was also performed to detect potential overexpressed signatures in patients with default parameters except for –collapse false –permute phenotype (68).

Gene Regulatory Network Inference. ARACNe-AP software was used to infer a Gene Regulatory Network (69) using iEG scRNA-seq data to predict a list of target genes of each TF in the context of cells expressing *ETO2–GLIS2* (called "EG network," Supplementary Table S6). For that purpose, iEG cells with $\log_2$ counts < 1 for *ETO2–GLIS2* were not considered. ARACNe was run over the $\log_2$ normalized counts in bootstrap mode (100 iterations), with a *P*-value threshold of 1e−8 and a custom curated list of 2,171 TFs. Similarly, we predicted the target genes of each TF in a normal hematopoietic context from published scRNA-seq data (70) obtained with the same protocol (called "normal network"; Supplementary Table S7). Therefore, the activity of each TF was computed using two independent lists of target genes, and consistent prediction of TF activity levels were obtained for key factors (Fig. 5E).

Transcription Regulator Activity Prediction and Master Regulator Analysis. Networks were used to infer activity (NES) of the transcriptomic regulator with R library viper (71), as described in the package manual. NESs were used to test differential activity by *t* test and *P* value correction by Benjamini–Hochberg (FDR cutoff at 0.05). Master Regulator Analysis was also carried out as described in the manual, with calculation of bootstrapped null model and shadow correction (enrichment of a regulator driven by similar targets of another regulator) of top 25 regulators.

ATAC-seq

ATAC-seq analysis has been adapted from ref. 72. Briefly, cells were isolated from iEG mice and cultured for 24 hours in RPMI supplemented with 10% FBS, cytokines (mIL3, mIL6, mSCF, mTPO, mFLT3l) and 100 ng/mL doxycycline. After cell lysis, transposition, and purification step, the transposed DNA fragments were amplified by PCR between 12 and 18 times depending on the number of cells at the beginning (6,000 to 50,000) using adapters from the Nextera Index Kit (Illumina). PCR purification was performed using Agencourt AMPure XP magnetic beads (Beckman Coulter A63880) to remove large fragments and remaining primers. Library quality was assessed using an Agilent 2100 Bioanalyzer using a High Sensitivity DNA Chip (Agilent Technologies 5067-4626). Libraries were sequenced using NovaSeq-6000 sequencer (Illumina; 50 bp paired-end reads).

Quality control of reads was performed using FastQC 0.11.7 and multiQC 1.5. The reads were aligned to the reference genome mm10 with bwa (aln 0.7.17). After alignment, we removed reads mapping to the mitochondrial genome, PCR duplicate reads, and reads with a mapping quality lower than 20 using samtools (v 1.9). Final read counts for all mouse datasets ranged from 37 to 128 million reads. Mapped reads were normalized to bins per million and were converted to bigwig format using deeptools (v3.2.0). Peak calling, differential analysis, annotation, and motif analysis were performed using macs2 (V 2.1.2), Diffbind R package [v 2.8.0 in R-3.5.1 with threshold $\log_2(1.5)$], and homer (v4.10.4, annotatePeak.pl and findMotifsGenome.pl), respectively.

Data Availability

scRNA-seq data were deposited into EBI Array-Express under the accession number E-MTAB-7213 and E-MTAB-8360.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors' Contributions

Conception and design: C.K. Lopez, F. Locatelli, B. Gottgens, J. Schwaller, T. Mercher

Development of methodology: C.K. Lopez, V. Stavropoulou, N. Droin, B. Gottgens, T. Mercher

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): C.K. Lopez, V. Stavropoulou, Z. Aid, P. Ballerini, C. Bilhou-Nabera, H. Lapillonne, F. Boudia, C. Thirant, S.J. Kinston, Y. Lecluse, J.L. Villeval, E. Delabesse, A.H.F.M. Peters, M. Gaudry, R. Masetti, F. Locatelli, C. Nerlov, I. Godin, O.A. Bernard, A. Petit, F. Pflumio, J. Schwaller, T. Mercher

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): C.K. Lopez, E. Noguera, E. Robert, M. Diop, B. Job, W. Vainchenker, F. Locatelli, N. Droin, C. Lobry, I. Godin, B. Gottgens, A. Petit, J. Schwaller, T. Mercher
Writing, review, and/or revision of the manuscript: C.K. Lopez, E. Noguera, Z. Aid, H. Lapillonne, F. Boudia, A. Fagnan, M.-L. Arcangeli, A.H.F.M. Peters, M. Gaudry, R. Masetti, F. Locatelli, S. Malinge, N. Droin, C. Lobry, I. Godin, O.A. Bernard, A. Petit, F. Pflumio, J. Schwaller, T. Mercher
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): C.K. Lopez, E. Noguera, S.J. Kinston, E. Brunet, L. Babin, J.L. Villeval
Study supervision: W. Vainchenker, B. Gottgens, J. Schwaller, T. Mercher
Other (FISH analysis): C. Bilhou-Nabera

Acknowledgments

We are grateful to Michael Kyba for the generous gift of the *A2Lox-Cre* ES cells and to Peter D. Aplan and Olivier Delattre for critical reading of the manuscript. We thank Clarisse Thiollier, Olivier Bluteau, Lou Le Mouël, Claude Preudhomme, Lila Diamanti, Nassera Abermil, Julie Chaumeil, and Sabine Juge for their expertise and helpful discussions. We thank the FMI transgenic core facility (J.F. Spetz and P. Kopp) for their help in establishing iETO2–GLIS2 transgenic mice, Gustave Roussy institute facilities for mouse care (Patrick Gonin and Karine Ser-Leroux) and bioinformatics support (Philippe Dessen and Daniel Gautheret), and Paule Zanardo for excellent administrative assistance. This work was supported by the Association Laurette Fugain (ALF-2015/13, to T. Mercher), SIRIC-SOCRATE (INCa-DGOS-INSERM 6043, to T. Mercher), Fondation pour la Recherche Médicale (FRM-ING20150532273, to Z. Aid and C.K. Lopez), Fondation de France (FdF-00057925, to C. Thirant and T. Mercher), Lady Tata Foundation (R14120LL, to C. Thirant), Cancéropôle Ile de France (to C.K. Lopez and Emergence 2015), Gustave Roussy Genomic Core Facility – TA2016 (to C.K. Lopez), Institut National Du Cancer (PLBIO-2014-176 and PLBIO-2018-169, to T. Mercher), PAIR-Pédiatrie/CONECT-AML (COllaborative Network for Children and Teenagers with Acute Myeloblastic Leukemia: INCa-ARC-LIGUE_11905 and Association Laurette Fugain), and Associazione Italiana Ricerca sul Cancro (to F. Locatelli). A.H.F.M. Peters was supported by the Novartis Research Foundation. S.J. Kinston and B. Gottgens were supported by a Bloodwise Specialist Programme grant (12029) and infrastructure funding from the Wellcome to the Wellcome & MRC Cambridge Stem Cell Institute. J. Schwaller was supported by the Swiss National Science Foundation (31003A_173224/1), the Swiss Cancer League (KFS-3487-08-2014), the San Salvatore Foundation (Lugano, Switzerland), the Novartis Foundation for Biomedical Research, and the Gertrude von Meissner Foundation (Basel, Switzerland). T. Mercher is an Equipe Labellisée LIGUE principal investigator.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received December 13, 2018; revised August 5, 2019; accepted September 23, 2019; published first October 29, 2019.

REFERENCES

1. Gröbner SN, Worst BC, Weischenfeldt J, Buchhalter I, Kleinheinz K, Rudneva VA, et al. The landscape of genomic alterations across childhood cancers. Nature 2018;555:321–7.
2. Bolouri H, Farrar JE, Triche T, Ries RE, Lim EL, Alonzo TA, et al. The molecular landscape of pediatric acute myeloid leukemia reveals recurrent structural alterations and age-specific mutational interactions. Nat Med 2018;24:103–12.
3. Greaves M. In utero origins of childhood leukaemia. Early Hum Dev 2005;81:123–9.
4. Hong D, Gupta R, Ancliff P, Atzberger A, Brown J, Soneji S, et al. Initiating and cancer-propagating cells in TEL-AML1-associated childhood leukemia. Science 2008;319:336–9.
5. Slany RK. The molecular biology of mixed lineage leukemia. Haematologica 2009;94:984–93.
6. Sanjuan-Pla A, Bueno C, Prieto C, Acha P, Stam RW, Marschalek R, et al. Revisiting the biology of infant t(4;11)/MLL-AF4+ B-cell acute lymphoblastic leukemia. Blood 2015;126:2676–85.
7. Creutzig U, Büchner T, Sauerland MC, Zimmermann M, Reinhardt D, Döhner H, et al. Significance of age in acute myeloid leukemia patients younger than 30 years: a common analysis of the pediatric trials AML-BFM 93/98 and the adult trials AMLCG 92/99 and AMLSG HD93/98A. Cancer 2008;112:562–71.
8. Thiollier C, Lopez CK, Gerby B, Ignacimouttou C, Poglio S, Duffourd Y, et al. Characterization of novel genomic alterations and therapeutic approaches using acute megakaryoblastic leukemia xenograft models. J Exp Med 2012;209:2017–31.
9. de Rooij JDE, Branstetter C, Ma J, Li Y, Walsh MP, Cheng J, et al. Pediatric non-Down syndrome acute megakaryoblastic leukemia is characterized by distinct genomic subsets with varying outcomes. Nat Genet 2017;49:451–6.
10. Lopez CK, Malinge S, Gaudry M, Bernard OA, Mercher T. Pediatric acute megakaryoblastic leukemia: multitasking fusion proteins and oncogenic cooperations. Trends Cancer 2017;3:631–42.
11. Gough SM, Slape CI, Aplan PD. NUP98 gene fusions and hematopoietic malignancies: common themes and new biologic insights. Blood 2011;118:6247–57.
12. Masetti R, Pigazzi M, Togni M, Astolfi A, Indio V, Manara E, et al. CBFA2T3-GLIS2 fusion transcript is a novel common feature in pediatric, cytogenetically normal AML, not restricted to FAB M7 subtype. Blood 2013;121:3469–72.
13. Grünewald TGP, Cidre-Aranaz F, Surdez D, Tomazou EM, de Álava E, Kovar H, et al. Ewing sarcoma. Nat Rev Dis Primer 2018;4:5.
14. Bardini M, Spinelli R, Bungaro S, Mangano E, Corral L, Cifola I, et al. DNA copy-number abnormalities do not occur in infant ALL with t(4;11)/MLL-AF4. Leukemia 2010;24:169–76.
15. Krivtsov AV, Feng Z, Lemieux ME, Faber J, Vempati S, Sinha AU, et al. H3K79 methylation profiles define murine and human MLL-AF4 leukemias. Cancer Cell 2008;14:355–68.
16. Thirant C, Ignacimouttou C, Lopez CK, Diop M, Le Mouël L, Thiollier C, et al. ETO2-GLIS2 hijacks transcriptional complexes to drive cellular identity and self-renewal in pediatric acute megakaryoblastic leukemia. Cancer Cell 2017;31:452–65.
17. Dawson MA, Prinjha RK, Dittmann A, Giotopoulos G, Bantscheff M, Chan W-I, et al. Inhibition of BET recruitment to chromatin as an effective treatment for MLL-fusion leukaemia. Nature 2011;478:529–33.
18. Erb MA, Scott TG, Li BE, Xie H, Paulk J, Seo H-S, et al. Transcription control by the ENL YEATS domain in acute leukaemia. Nature 2017;543:270–4.
19. Babovic S, Eaves CJ. Hierarchical organization of fetal and adult hematopoietic stem cells. Exp Cell Res 2014;329:185–91.
20. Elagib KE, Brock AT, Goldfarb AN. Megakaryocyte ontogeny: Clinical and molecular significance. Exp Hematol 2018;61:1–9.
21. Laurenti E, Göttgens B. From haematopoietic stem cells to complex differentiation landscapes. Nature 2018;553:418–26.
22. Sanjuan-Pla A, Macaulay IC, Jensen CT, Woll PS, Luis TC, Mead A, et al. Platelet-biased stem cells reside at the apex of the haematopoietic stem-cell hierarchy. Nature 2013;502:232–6.
23. Carrelha J, Meng Y, Kettyle LM, Luis TC, Norfo R, Alcolea V, et al. Hierarchically related lineage-restricted fates of multipotent haematopoietic stem cells. Nature 2018;554:106–11.
24. Rodriguez-Fraticelli AE, Wolock SL, Weinreb CS, Panero R, Patel SH, Jankovic M, et al. Clonal analysis of lineage fate in native haematopoiesis. Nature 2018;553:212–6.

25. Stavropoulou V, Kaspar S, Brault L, Sanders MA, Juge S, Morettini S, et al. MLL-AF9 expression in hematopoietic stem cells drives a highly invasive AML expressing EMT-related genes linked to poor outcome. Cancer Cell 2016;30:43–58.
26. Notta F, Zandi S, Takayama N, Dobson S, Gan OI, Wilson G, et al. Distinct routes of lineage development reshape the human blood hierarchy across ontogeny. Science 2016;351:aab2116.
27. Paul F, Arkin Y, Giladi A, Jaitin DA, Kenigsberg E, Keren-Shaul H, et al. Transcriptional heterogeneity and lineage commitment in myeloid progenitors. Cell 2015;163:1663–77.
28. Wilson NK, Kent DG, Buettner F, Shehata M, Macaulay IC, Calero-Nieto FJ, et al. Combined single-cell functional and gene expression analysis resolves heterogeneity within stem cell populations. Cell Stem Cell 2015;16:712–24.
29. Velten L, Haas SF, Raffel S, Blaszkiewicz S, Islam S, Hennig BP, et al. Human haematopoietic stem cell lineage commitment is a continuous process. Nat Cell Biol 2017;19:271–81.
30. Yuan J, Nguyen CK, Liu X, Kanellopoulou C, Muljo SA. Lin28b reprograms adult bone marrow hematopoietic progenitors to mediate fetal-like lymphopoiesis. Science 2012;335:1195–200.
31. Copley MR, Babovic S, Benz C, Knapp DJHF, Beer PA, Kent DG, et al. The Lin28b-let-7-Hmga2 axis determines the higher self-renewal potential of fetal haematopoietic stem cells. Nat Cell Biol 2013;15:916–25.
32. Bluteau O, Langlois T, Rivera-Munoz P, Favale F, Rameau P, Meurice G, et al. Developmental changes in human megakaryopoiesis. J Thromb Haemost 2013;11:1730–41.
33. Rowe RG, Wang LD, Coma S, Han A, Mathieu R, Pearson DS, et al. Developmental regulation of myeloerythroid progenitor function by the Lin28b-let-7-Hmga2 axis. J Exp Med 2016;213:1497–512.
34. Xu J, Shao Z, Li D, Xie H, Kim W, Huang J, et al. Developmental control of polycomb subunit composition by GATA factors mediates a switch to non-canonical functions. Mol Cell 2015;57:304–16.
35. Ye M, Zhang H, Amabile G, Yang H, Staber PB, Zhang P, et al. C/EBPa controls acquisition and maintenance of adult haematopoietic stem cell quiescence. Nat Cell Biol 2013;15:385–94.
36. Pimkin M, Kossenkov AV, Mishra T, Morrissey CS, Wu W, Keller CA, et al. Divergent functions of hematopoietic transcription factors in lineage priming and differentiation during erythro-megakaryopoiesis. Genome Res 2014;24:1932–44.
37. Petit A, Ducassou S, Leblanc T, Pasquet M, Rousseau A, Ragu C, et al. Maintenance therapy with interleukin-2 for childhood AML: results of ELAM02 phase III randomized trial. HemaSphere 2018;2:e159.
38. Cabezas-Wallscheid N, Klimmeck D, Hansson J, Lipka DB, Reyes A, Wang Q, et al. Identification of regulatory networks in HSCs and their immediate progeny via integrated proteome, transcriptome, and DNA methylome analysis. Cell Stem Cell 2014;15:507–22.
39. Chen W, Kumar AR, Hudson WA, Li Q, Wu B, Staggs RA, et al. Malignant transformation initiated by Mll-AF9: gene dosage and critical target cells. Cancer Cell 2008;13:432–40.
40. Bourquin J-P, Subramanian A, Langebrake C, Reinhardt D, Bernard O, Ballerini P, et al. Identification of distinct molecular phenotypes in acute megakaryoblastic leukemia by gene expression profiling. Proc Natl Acad Sci U S A 2006;103:3339–44.
41. Gruber TA, Larson Gedman A, Zhang J, Koss CS, Marada S, Ta HQ, et al. An Inv(16)(p13.3q24.3)-encoded CBFA2T3-GLIS2 fusion protein defines an aggressive subtype of pediatric acute megakaryoblastic leukemia. Cancer Cell 2012;22:683–97.
42. Giladi A, Paul F, Herzog Y, Lubling Y, Weiner A, Yofe I, et al. Single-cell characterization of haematopoietic progenitors and their trajectories in homeostasis and perturbed haematopoiesis. Nat Cell Biol 2018;20:836–46.
43. Hoppe PS, Schwarzfischer M, Loeffler D, Kokkaliaris KD, Hilsenbeck O, Moritz N, et al. Early myeloid lineage choice is not initiated by random PU.1 to GATA1 protein ratios. Nature 2016;535:299–302.
44. Bresnick EH, Lee H-Y, Fujiwara T, Johnson KD, Keles S. GATA switches as developmental drivers. J Biol Chem 2010;285:31087–93.
45. Dang J, Nance S, Ma J, Cheng J, Walsh MP, Vogel P, et al. AMKL chimeric transcription factors are potent inducers of leukemia. Leukemia 2017;31:2228–34.
46. Takano H, Ema H, Sudo K, Nakauchi H. Asymmetric division and lineage commitment at the level of hematopoietic stem cells: inference from differentiation in daughter cell and granddaughter cell pairs. J Exp Med 2004;199:295–302.
47. Haas S, Hansson J, Klimmeck D, Loeffler D, Velten L, Uckelmann H, et al. Inflammation-induced emergency megakaryopoiesis driven by hematopoietic stem cell-like megakaryocyte progenitors. Cell Stem Cell 2015;17:422–34.
48. Niu C, Zhang J, Breslin P, Onciu M, Ma Z, Morris SW. c-Myc is a target of RNA-binding motif protein 15 in the regulation of adult hematopoietic stem cell and megakaryocyte development. Blood 2009;114:2087–96.
49. Knudsen KJ, Rehn M, Hasemann MS, Rapin N, Bagger FO, Ohlsson E, et al. ERG promotes the maintenance of hematopoietic stem cells by restricting their differentiation. Genes Dev 2015;29:1915–29.
50. Rosenbauer F, Wagner K, Kutok JL, Iwasaki H, Le Beau MM, Okuno Y, et al. Acute myeloid leukemia induced by graded reduction of a lineage-specific transcription factor, PU.1. Nat Genet 2004;36:624–30.
51. Carmichael CL, Metcalf D, Henley KJ, Kruse EA, Di Rago L, Mifsud S, et al. Hematopoietic overexpression of the transcription factor Erg induces lymphoid and erythro-megakaryocytic leukemia. Proc Natl Acad Sci U S A 2012;109:15437–42.
52. Zhang P, Iwasaki-Arai J, Iwasaki H, Fenyus ML, Dayaram T, Owens BM, et al. Enhancement of hematopoietic stem cell repopulating capacity and self-renewal in the absence of the transcription factor C/EBP alpha. Immunity 2004;21:853–63.
53. Roberts I, Izraeli S. Haematopoietic development and leukaemia in Down syndrome. Br J Haematol 2014;167:587–99.
54. Albers CA, Newbury-Ecob R, Ouwehand WH, Ghevaert C. New insights into the genetic basis of TAR (thrombocytopenia-absent radii) syndrome. Curr Opin Genet Dev 2013;23:316–23.
55. Liu Z-J, Sola-Visner M. Neonatal and adult megakaryopoiesis. Curr Opin Hematol 2011;18:330–7.
56. Barrett NA, Malouf C, Kapeni C, Bacon WA, Giotopoulos G, Jacobsen SEW, et al. Mll-AF4 confers enhanced self-renewal and lymphoid potential during a restricted window in development. Cell Rep 2016;16:1039–54.
57. Lee JH, Lee JE, Kahng JY, Kim SH, Park JS, Yoon SJ, et al. Human glioblastoma arises from subventricular zone cells with low-level driver mutations. Nature 2018;560:243–7.
58. Avanzi GC, Lista P, Giovinazzo B, Miniero R, Saglio G, Benetton G, et al. Selective growth response to IL-3 of a human leukaemic cell line with megakaryoblastic features. Br J Haematol 1988;69:359–66.
59. Sato T, Sekine H, Kakuda H, Miura N, Sunohara M, Fuse A. HIV infection of megakaryocytic cell lines. Leuk Lymphoma 2000;36:397–404.
60. Martin P, Papayannopoulou T. HEL cells: a new human erythroleukemia cell line with spontaneous and induced globin expression. Science 1982;216:1233–5.
61. Iacovino M, Bosnakovski D, Fey H, Rux D, Bajwa G, Mahen E, et al. Inducible cassette exchange: a rapid and efficient system enabling conditional gene expression in embryonic stem and primary cells. Stem Cells 2011;29:1580–8.
62. Schaefer BC, Schaefer ML, Kappler JW, Marrack P, Kedl RM. Observation of antigen-dependent CD8+ T-cell/dendritic cell interactions *in vivo*. Cell Immunol 2001;214:110–22.
63. Månsson R, Hultquist A, Luc S, Yang L, Anderson K, Kharazi S, et al. Molecular evidence for hierarchical transcriptional lineage priming in fetal and adult stem cells and multipotent progenitors. Immunity 2007;26:407–19.
64. Picelli S, Faridani OR, Björklund AK, Winberg G, Sagasser S, Sandberg R. Full-length RNA-seq from single cells using Smart-seq2. Nat Protoc 2014;9:171–81.
65. Yip SH, Sham PC, Wang J. Evaluation of tools for highly variable gene discovery from single-cell RNA-seq data. Brief Bioinform 2018 Feb 21 [Epub ahead of print].
66. Lun ATL, Bach K, Marioni JC. Pooling across cells to normalize single-cell RNA sequencing data with many zero counts. Genome Biol 2016;17:75.

67. Brennecke P, Anders S, Kim JK, Kołodziejczyk AA, Zhang X, Proserpio V, et al. Accounting for technical noise in single-cell RNA-seq experiments. Nat Methods 2013;10:1093–5.
68. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005;102:15545–50.
69. Lachmann A, Giorgi FM, Lopez G, Califano A. ARACNe-AP: gene network reverse engineering through adaptive partitioning inference of mutual information. Bioinforma 2016;32:2233–5.
70. Nestorowa S, Hamey FK, Pijuan Sala B, Diamanti E, Shepherd M, Laurenti E, et al. A single-cell resolution map of mouse hematopoietic stem and progenitor cell differentiation. Blood 2016;128:e20–31.
71. Alvarez MJ, Shen Y, Giorgi FM, Lachmann A, Ding BB, Ye BH, et al. Functional characterization of somatic mutations in cancer using network-based inference of protein activity. Nat Genet 2016;48:838–47.
72. Buenrostro JD, Wu B, Chang HY, Greenleaf WJ. ATAC-seq: a method for assaying chromatin accessibility genome-wide. Curr Protoc Mol Biol 2015;109:21.29.1–9.
73. van Zutven LJCM, Onen E, Velthuizen SCJM, van Drunen E, von Bergh ARM, van den Heuvel-Eibrink MM, et al. Identification of NUP98 abnormalities in acute leukemia: JARID1A (12p13) as a new partner gene. Genes Chromosomes Cancer 2006;45:437–46.
74. Struski S, Lagarde S, Bories P, Puiseux C, Prade N, Cuccuini W, et al. NUP98 is rearranged in 3.8% of pediatric AML forming a clinical and molecular homogenous group with a poor prognosis. Leukemia 2017;31:565–72.
75. Hara Y, Shiba N, Ohki K, Tabuchi K, Yamato G, Park M-J, et al. Prognostic impact of specific molecular profiles in pediatric acute megakaryoblastic leukemia in non-Down syndrome. Genes Chromosomes Cancer 2017;56:394–404.
76. de Rooij JDE, Hollink IHIM, Arentsen-Peters STCJM, van Galen JF, Berna Beverloo H, Baruchel A, et al. NUP98/JARID1A is a novel recurrent abnormality in pediatric acute megakaryoblastic leukemia with a distinct HOX gene expression pattern. Leukemia 2013;27:2280–8.
77. Aran D, Hu Z, Butte AJ. xCell: digitally portraying the tissue cellular heterogeneity landscape. Genome Biol 2017;18:220.