



HAL
open science

The epigenetic regulator RINF (CXXC5) maintains SMAD7 expression in human immature erythroid cells and sustains red blood cells expansion.

Audrey Astori, Gabriel Matherat, Isabelle Munoz, Emilie-Fleur Gautier, Didier Surdez, Yaël Zermati, Frédérique Verdier, Sakina Zaidi, Vincent Feuillet, Amir Kadi, et al.

► To cite this version:

Audrey Astori, Gabriel Matherat, Isabelle Munoz, Emilie-Fleur Gautier, Didier Surdez, et al.. The epigenetic regulator RINF (CXXC5) maintains SMAD7 expression in human immature erythroid cells and sustains red blood cells expansion.. *Haematologica*, 2020, pp.0 - 0. 10.3324/haematol.2020.263558 . hal-03224573

HAL Id: hal-03224573

<https://hal.science/hal-03224573>

Submitted on 11 May 2021

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

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

The epigenetic regulator RINF (CXXC5) maintains SMAD7 expression in human immature erythroid cells and sustains red blood cells expansion

by Audrey Astori, Gabriel Matherat, Isabelle Munoz, Emilie-Fleur Gautier, Didier Surdez, Yaël Zermati, Frédérique Verdier, Sakina Zaidi, Vincent Feuillet, Amir Kadi, Evelyne Lauret, Oliver Delattre, Carine Lefèvre, Michaela Fontenay, Evelyne Ségal-Bendirdjian, Isabelle Dusanter-Fourt, Didier Bouscary, Olivier Hermine, Patrick Mayeux, and Frédéric Pendino

Haematologica 2020 [Epub ahead of print]

Citation: Audrey Astori, Gabriel Matherat, Isabelle Munoz, Emilie-Fleur Gautier, Didier Surdez, Yaël Zermati, Frédérique Verdier, Sakina Zaidi, Vincent Feuillet, Amir Kadi, Evelyne Lauret, Oliver Delattre, Carine Lefèvre, Michaela Fontenay, Evelyne Ségal-Bendirdjian, Isabelle Dusanter-Fourt, Didier Bouscary, Olivier Hermine, Patrick Mayeux, and Frédéric Pendino. The epigenetic regulator RINF (CXXC5) maintains SMAD7 expression in human immature erythroid cells and sustains red blood cells expansion.

Haematologica. 2020; 105:xxx

doi:10.3324/haematol.2020.263558

Publisher's Disclaimer.

E-publishing ahead of print is increasingly important for the rapid dissemination of science. Haematologica is, therefore, E-publishing PDF files of an early version of manuscripts that have completed a regular peer review and have been accepted for publication. E-publishing of this PDF file has been approved by the authors. After having E-published Ahead of Print, manuscripts will then undergo technical and English editing, typesetting, proof correction and be presented for the authors' final approval; the final version of the manuscript will then appear in print on a regular issue of the journal. All legal disclaimers that apply to the journal also pertain to this production process.

The epigenetic regulator RINF (CXXC5) maintains *SMAD7* expression in human immature erythroid cells and sustains red blood cells expansion.

Audrey Astori,^{1,2,3#} Gabriel Matherat,^{1,2,3#} Isabelle Munoz,^{1,2,3} Emilie-Fleur Gautier,^{1,2,3} Didier Surdez,^{3,4,5} Yaël Zermati,^{1,2} Frédérique Verdier,^{1,2} Sakina Zaidi,^{3,4,5} Vincent Feuillet,¹ Amir Kadi,¹ Evelyne Lauret,^{1,3} Olivier Delattre,^{3,4,5} Carine Lefèvre,^{1,2} Michaela Fontenay,^{1,2,6} Evelyne Ségal-Bendirdjian,⁷ Isabelle Dusanter-Fourt,^{1,3} Didier Bouscary,^{1,3} Olivier Hermine,^{2,3,8} Patrick Mayeux,^{1,2,3} and Frédéric Pendino.^{1,2,3}

¹ Université de Paris, Institut Cochin, INSERM, CNRS, F-75014 Paris, France;

² Laboratory of Excellence GR-ex, Paris, France;

³ Equipe Labellisée Ligue Nationale Contre le Cancer (LNCC), Paris, France;

⁴ PSL Research University, Institut Curie Research Center, INSERM U830, Paris, France;

⁵ SIREDO: Care, Innovation and Research for Children, Adolescents and Young Adults with Cancer, Institut Curie, Paris, France;

⁶ Service d'Hématologie Biologique, Hôpital Cochin, Assistance Publique-Hôpitaux de Paris, Centre-Université de Paris, Paris, France;

⁷ INSERM UMR-S 1124, Team: Cellular Homeostasis Cancer and Therapies, Université de Paris, Paris, France;

⁸ Université de Paris, Institut Imagine, INSERM, CNRS, F-75015, Paris, France.

AA and GM contributed equally to this work as cofirst authors and are ranked by alphabetical order.

Short Title: *RINF maintenance of SMAD7 sustains human erythropoiesis*

Abstract word count: **244**. Figures: **7**. Tables: **1**.

All supplementary data (Supplementary Methods, Legends, and References, Tables (1), and Figures (7)) are included into a single PDF document.

Correspondence: Frédéric PENDINO (Ph.D., CRCN Inserm, H.D.R).

✉: Institut Cochin, 22 rue Méchain, 75014 Paris, France.

☎: +33140516403 📧: frederic.pendino@inserm.fr

Abstract

The gene *CXXC5*, encoding a Retinoid-Inducible Nuclear Factor (RINF), is located within a region at 5q31.2 commonly deleted in myelodysplastic syndrome (MDS) and adult acute myeloid leukemia (AML). RINF may act as an epigenetic regulator and has been proposed as a tumor suppressor in hematopoietic malignancies. However, functional studies in normal hematopoiesis are lacking, and its mechanism of action is unknown. Here, we evaluated the consequences of *RINF* silencing on cytokine-induced erythroid differentiation of human primary CD34+ progenitors. We found that *RINF* is expressed in immature erythroid cells and that *RINF*-knockdown accelerated erythropoietin-driven maturation, leading to a significant reduction (~45%) in the number of red blood cells (RBCs), without affecting cell viability. The phenotype induced by *RINF*-silencing was TGF β -dependent and mediated by SMAD7, a TGF β -signaling inhibitor. RINF upregulates *SMAD7* expression by direct binding to its promoter and we found a close correlation between *RINF* and *SMAD7* mRNA levels both in CD34+ cells isolated from bone marrow of healthy donors and MDS patients with del(5q). Importantly, *RINF* knockdown attenuated *SMAD7* expression in primary cells and ectopic *SMAD7* expression was sufficient to prevent the *RINF* knockdown-dependent erythroid phenotype. Finally, RINF silencing affects 5'-hydroxymethylation of human erythroblasts, in agreement with its recently described role as a Tet2-anchoring platform in mouse. Altogether, our data bring insight into how the epigenetic factor RINF, as a transcriptional regulator of *SMAD7*, may fine-tune cell sensitivity to TGF β superfamily cytokines and thus play an important role in both normal and pathological erythropoiesis.

Key points:

- 1) Retinoid-Induced Nuclear Factor (RINF) binds to the promoter of *SMAD7* and transactivates its gene expression
- 2) *CXXC5/RINF* silencing sensitizes human erythroid progenitors to TGF β -induced growth inhibition

Introduction

In earlier work, we demonstrated that *RINF* loss-of-function affects granulopoiesis in normal and tumoral human hematopoietic cells (1). In keeping with its gene location at chromosome 5q31.2 (*Supplementary Figure S1*), a commonly deleted region associated with high risk in myeloid neoplasms (2), *RINF* loss of expression has been proposed to contribute to development or progression of (pre)leukemia, such as MDS (1, 3-5). We and others have demonstrated that *RINF* mRNA expression is an unfavorable (6, 7) and independent (3) prognostic factor in acute myeloid leukemia (AML) as well as in solid tumors (8, 9). However, its role in normal hematopoiesis has hitherto been poorly investigated and its contribution to the erythroid lineage and RBC expansion is unknown.

RINF contains a Nuclear Localization Signal (NLS) that has been functionally validated (10) and in most studies, its subcellular localization is reported to be mainly or exclusively nuclear and to act as a transcriptional cofactor (1, 3, 8, 10-17). *RINF* associates strongly with chromatin (1) through its conserved zinc-finger domain (CXXC) that plays an essential role in providing the capacity to bind CpG islands (18, 19). Interestingly, this domain is almost identical to the one harbored by TET1 and TET3, two epigenetic modulators involved in the erasure of DNA-methylation marks (20), pointing to the possibility that *RINF* might interfere with TET-activities, hydroxymethylation, and gene transcription, as recently demonstrated in mouse (12, 15). *RINF* has also been reported to bind ATM, mediate DNA-damage induced-activation of TP53 (3, 10) and inhibit the WNT- β -catenin signaling pathway (3, 21-24) through a cytoplasmic interaction with Disheveled proteins DVL and DVL2 (21).

Transforming growth factor β (TGF β) is a powerful and widespread cell growth inhibitor in numerous mammalian tissues (25, 26). In the hematopoietic system, TGF β is known to regulate hematopoietic stem and progenitor cells (HSPCs) and is also described as a potent inducer of erythroid differentiation and inhibitor of cell proliferation (27-30). TGF β signals through cell surface serine/threonine kinase receptors, mainly TGF β RI and TGF β RII. Activated TGF β RI phosphorylates SMAD2 and SMAD3 that translocate into the nucleus and form complexes that regulate transcription of target genes. TGF β can also elicit its biological effects by activation of SMAD-independent pathways (31, 32). Inhibitory SMADs (SMAD6 and SMAD7)

inhibit TGF β signaling. Importantly, a reduced *SMAD7* expression sensitizes cells to the antiproliferative effects of TGF β and contributes to anemia in patients suffering from myelodysplastic syndrome (MDS), suggesting, firstly, that identifying transcriptional regulators of *SMAD7* could enlighten our understanding of erythropoiesis and, second, that inhibiting TGF β signaling could be a therapeutic strategy that would mitigate ineffective hematopoiesis in disease states (33-35).

In the present work, we use primary human CD34⁺ cells to demonstrate that RINF knockdown affects human erythropoiesis and mitigates RBC production through a mechanism that is mediated by *SMAD7*, the main inhibitor of TGF β signaling.

Methods

Primary culture of hematopoietic cells. Mononuclear cells were isolated from cord blood (CRB Saint-Louis Hospital, Paris) or adult bone marrow donors with written informed consent for research use, in accordance with the Declaration of Helsinki. The ethics evaluation committee of Inserm, the *Institutional Review Board* (IRB00003888), approved our research project (N°16-319). Mononucleated cells were separated by Ficoll-Paque (Life Technologies) and purified with human CD34 MicroBead Kit (Cat.N°130100453, Miltenyi Biotec). A two-step culture method for cell expansion and erythroid differentiation was used to obtain highly stage-enriched erythroblast populations (36, 37). Briefly, CD34⁺ cells (purity 94-98%) were cultured 7 days with IL6, IL3 (10 ng/mL), and SCF (100 ng/mL) to expand hematopoietic progenitors (37). CD36⁺ erythroid progenitors were purified using magnetic microbeads (CD36 FA6.152 from Beckman Coulter and anti-mouse IgG1 MicroBeads from Miltenyi Biotec). Then, erythropoietin (EPO) was added for 10-18 days to allow erythroid differentiation. For Colony-Forming Cells (CFC, counted at 11-14 days) assays, CD34⁺ hematopoietic stem cells were seeded in methylcellulose (H4034, StemCell Technologies). TGF β RI inhibitor SB431542 was from Selleckchem (Cat.N°S1067) (38). TGF β was from Peprotech (Cat.N°100-21).

Flow Cytometry and differentiation analysis. To monitor end-stage erythroid differentiation of primary cells, cells were labeled with anti-human PE/Cy7-conjugated

glycophorin-A (GPA/CD235a) (Cat.N°A71564), PE-conjugated CD71 (IM2001U), and APC-conjugated CD49d (Cat.N°B01682) from Beckman Coulter, PE-conjugated CD233/Band3 from IBGRL (Cat.N°9439-PE). FACS (fluorescence-activated cell sorting) analysis was performed on an Accuri™ C6 Flow Cytometer (Becton-Dickinson). For morphological characterization cells were cytopun and stained with May-Grünwald-Giemsa method (MGG). For benzidine assay, 2.10^5 cells were incubated for 5 minutes with 0.01% benzidine and H₂O₂ at 0.3% (Sigma-Aldrich).

Culture of cell lines and treatments. K562 (ATCC#CCL243) and UT7_{5.3}(39) cells were cultured in RPMI 1640 medium and minimum essential medium (MEM α) supplemented with 2.5 ng/mL of GM-CSF (Myltenyi Biotech), respectively. Both medium were supplemented with 10% of inactivated fetal bovine serum (FBS), 2 mM L-Glutamine, 50 units/ml penicillin G and 50 μ g/mL streptomycin (Life Technologies). To induce hemoglobin production, K562 cells were treated with hemin at 40 μ M (Sigma-Aldrich). Erythroid maturation of UT7_{5.3} was triggered by replacing GM-CSF with EPO 5U/ml.

Lentiviral and Retroviral transduction of hematopoietic cells. For shRNA-mediated *RINF* knockdown, we used the pTRIPDU3/GFP lentiviral vector (40) which drives the same shRNA sequences downstream of H1 promoter than our previously described pLKO.1/puroR vector (1). For *RINF* overexpression, the retroviral vector MigR/IRES-GFP was used with experimental conditions as previously described (1). Green fluorescent protein (GFP) sorting was performed on a FACSAriaIII (BD Biosciences). For *SMAD7* overexpression, the retroviral vector pBABE-puro-*SMAD7-HA* (Addgene plasmid#37044) was used as previously described (1, 41). For doxycycline-inducible lentiviral expression, *SMAD7-HA* cDNA was inserted in pINDUCER21 vector by gateway technology (Addgene plasmid #46948) and shRNA sequences were inserted in the Tet-pLKO-GFP vector (modified from Tet-pLKO-puro, Addgene plasmid#21915, deposited by Dmitri Wiederschain).

Quantitative RT-PCR. Cells were collected and directly stored at -80°C for RNA preparation with TRIzol (Life technologies) extraction protocol as previously described (1). First-strand cDNA synthesis (RT) were carried out using Transcriptor First Strand cDNA Synthesis Kit (Cat.N°489703000, Roche). *RINF* mRNA expression

was detected using Lightcycler® 480 ProbesMaster kit (Cat.N°4707494001, Roche). Relative mRNA expressions were normalized to *RLPL2*, *PPIA* or *ACTB* gene expression in a two-color duplex reaction. For *SMAD7*, *PU.1* and *c-KIT* mRNAs detection, quantitative PCRs were performed using SYBRGreen. Q-RT-PCR were performed on a Light Cycler 480 machine (Roche) and gene expression calculated by with the $2^{-\Delta\Delta CT}$ method. Primer sequences are available in a supplementary file.

Western-blot analysis. Total cell extracts and western blotting were carried out as previously described (1). Blots were incubated with primary polyclonal antibody and then with appropriate peroxidase-conjugated secondary antibody (anti-rabbit and anti-mouse IgG HRP antibodies from Cell Signaling, Cat.N°7074S, Cat.N°7076S) and anti-goat IgG HRP antibody from Southern Biotech (Cat.N°6160-05). Detection of proteins was performed using chemiluminescent detection system (clarity western ECL, Cat.N°170-5060, Biorad). Rabbit polyclonal antibodies were previously described for RINF (1), P85/PI3KR1 (42), or commercially purchased for anti-RINF, anti-SMAD7 (Cat.N°16513-1-AP, Cat.N°25840-1-AP, ProteinTech), anti- β -ACTIN (ACTB, Cat.N°A1978, Sigma-Aldrich), anti-phospho-SMAD2/3 (Cat.N°8828, Cell Signaling), and anti-HSC70 mouse monoclonal antibody (SC-7298, Santa Cruz Biotechnology).

Statistical analyses. All analyses were performed using GraphPad Prism software. When distribution was normal, the 2-tailed Student *t* test was used for group comparisons. Contingency tables were performed by using the Fisher's exact test, and Pearson's correlation coefficient was used to determine the correlation between the normally distributed *RINF* mRNA and *SMAD7* mRNA expressions. Statistics were carried out on a minimum of 3 independent experiments. Significance of P values are indicated in Figures. Not significant (ns): $p > 0.05$; *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$. Error bars represent confidence intervals at 95% or standard deviations (SD) for q-PCR analyses.

Methods for “Flow cytometric cell sorting of MEP cells”, “Immunofluorescence”, “Chromatin Immunoprecipitation (ChIP) experiments”, and “Primer sequences used for quantitative RT-PCR and ChIP-qPCR” are described in the Supplemental Method section.

Results

RINF is expressed during early human erythropoiesis. We employed cytokine-induced erythroid differentiation of CD34+ progenitors isolated from human cord blood or adult bone marrow to model human erythropoiesis *in vitro*. To obtain highly enriched populations of differentiating erythroblasts, cells were grown in a two-phase liquid culture method (experimental design in Figure 1A) as previously described (36, 37). We found that RINF protein was present at the earliest stages of erythroid differentiation and reached its highest level in progenitors and proerythroblasts (ProE) (Figure 1B). After that, RINF protein level decreased in the basophilic erythroblast stage and was barely detectable in late erythroblasts (i.e. polychromatic and orthochromatic erythroblast stages). Similarly, *RINF* mRNA levels were high in erythroid progenitors but decreased at the basophilic stage and onwards (Figure 1C). This temporal expression pattern mirrors *RINF* expression data extracted from three independent transcriptome datasets (microarray or RNA-seq) performed by Merryweather-Clarke *et al.* (43), An *et al.* (44), and Keller *et al.* (45) with adults or cord-blood CD34+ cells (*Supplementary Figure S2*). Taken together, our data shows that *RINF* expression peaks in erythroid progenitors and proerythroblasts during cytokine-induced erythropoiesis.

***RINF* silencing results in diminished RBC production without impacting cell viability.** We next investigated consequences of *RINF* knockdown on erythropoiesis and cell viability (see experimental design in Figure 2A). Knockdown experiments were performed with two previously validated short-hairpin RNAs (*shRINF#4* and *shRINF#3*) sequences (1) driven by the lentiviral pTRIPDU3/GFP vector (Figures S3A and B). We found that RINF was extinguished as early as two days post-transduction (i.e. at day 0 of EPO) with *shRINF#4* at both mRNA (Figure 2B) and protein (Figure 2C) levels, without affecting cell viability (assessed by trypan blue cell counting, Figure 2D) or cell growth up to day 10 of EPO (as illustrated on Figure 2E, *left panel*). However, in the last days of *ex vivo* cultures (day 11-17), we noticed a reduction in total number of RBCs produced (average diminution of ~45%) at 17 days with EPO (Figure 2E, *right panel*), particularly marked for 4 donors out of 6 performed.

To investigate how RINF knockdown impacts clonogenic capacities of HSPCs, we next performed comparative colony forming assays for development and maturation of erythroid; burst-forming unit erythroid progenitors (BFU-E), myeloid; (CFU-G, CFU-M, or CFU-GM), and mixed colonies; (CFU-GEMM). No statistically significant change was noticed in the total number of colonies or the number of CFU-M, CFU-GM, or CFU-GEMM (*Supplementary Figure S3C*). However, the proportion of granulocytic colonies (CFU-G) and BFU-Es, were slightly reduced or increased, respectively ($p < 0.01$, Fisher's exact test). The number of MEP (megakaryocyte-erythroid progenitors)-sorted cells were also increased under RINF knockdown conditions (*Supplementary Figure S3E*), and this increase corresponded to an increase in small (more mature) BFU-Es, findings that suggest an accelerated maturation. A more careful analysis of colony size revealed that the median size of BFU-Es derived from CD34+ cells dropped by about 30% (i.e. from 0.101 to 0.071 mm², unpaired t-test, $P = 0.007$, $n = 3$ donors) after *RINF* knockdown (Figure 2F). Concomitantly with the size reduction of large BFU-Es (most immature ones), we noticed a slight but statistically significant increase in small-BFU-Es (most mature ones), in agreement with an accelerated maturation (Figure 2G).

Erythroid maturation is impacted by RINF. We next investigated if the reduced expansion observed from day 10 of EPO could be the consequence of an accelerated maturation. Indeed, we observed that erythroid differentiation was accelerated under conditions of RINF knockdown using benzidine staining quantification (Figure 3A), morphological analysis of *MGG*-stained cells (Figure 3B), and flow cytometry for erythroid markers GPA and CD71 (Figure 3C, *upper panel*), and CD49d and Band3 (Figure 3C, *lower panel*). These differences were statistically significant ($p < 0.001$) for both cord blood and adult bone marrow derived CD34+ cells. The shRNA-RINF-induced acceleration of erythropoiesis was also noticed by cell size reduction (not shown) and by a more rapid downregulation of *c-KIT* and *PU.1* mRNAs in *RINF* knockdown cells (Figure 3D). The accelerated maturation was also characterized by a pronounced reduction of ProE cells (~2.8-fold less) enumerated at day 11 (Figure 3B). Despite a weaker efficiency of *shRINF#3* to knockdown *RINF* expression compared to *shRINF#4* (*Supplementary Figure S3B and Figure 3E, left panel*), we also confirmed that a similar phenotype (i.e. accelerated maturation (as noticed by an increased benzidine staining, Figure 3E, *right panel*) and a reduction in

the total number of RBC produced at day 17, Figure 3F) was obtained with another shRNA sequence targeting *RINF*.

RINF controls erythroid maturation and RBC expansion in a TGF β -dependent manner. We then wondered if TGF β , a well-known inducer of erythroid maturation and inhibitor of RBC expansion, could be involved in the RINF-dependent phenotype. In agreement with an important role for autocrine TGF β signaling in the early stages of erythropoiesis, both TGF β 1 and TGF β RI were highly expressed in erythroid progenitors (Figure 4A). To functionally validate an involvement of the TGF β signaling pathway during *RINF* knockdown dependent acceleration of erythropoiesis, we performed liquid cell culture experiments of primary CD34+ cells in the presence of SB431542, a potent and selective inhibitor of TGF β RI (38) (Figure 4B). Strikingly, in the presence of SB431542, the maturation of RINF knockdown cells (evidenced by a faster acquisition of erythroid cell surface markers such as CD49d/Band3 at day 10 of EPO) was not accelerated and SMAD2 protein was not phosphorylated/activated after 3 days of EPO (Figure 4C, lanes 5 and 6), supporting the efficiency of the inhibitor. Conversely, in absence of inhibitor, RINF-silencing gives rise to a faster phosphorylation rate of SMAD2 protein, here noticed at 8 hours of EPO (compare lanes 1 and 2, or the kinetics of pSMAD2 waves in both conditions, *right panel*), indicating a higher sensitivity to autocrine TGF β -signaling that would be mediated through SMAD2, as previously reported in hematopoietic cells (35). A weaker RBC expansion was also noticed for RINF knockdown cells, even when TGF β 1 (5ng/mL) was added as late as 11 days of EPO treatment (Figure 4D). Moreover, even for the two donors where efficiency of *RINF* silencing was more subtle (donors 1 and 2, Figure 2E), *RINF* knockdown cells were more sensitive than control cells to TGF β 1 (Figures 4E-F). This was especially pronounced at low-doses (0.1 ng/mL), where we found a faster acquisition of GPA (at day 6 of EPO) and a reduced RBC production at 17 days of EPO (Figure 4F). Taken together, these data suggest that RINF knockdown cells were more sensitive to TGF β and that the level of RINF in erythroid progenitors impacts the upcoming RBC production, at least in presence of TGF β .

Identification of *SMAD7* as a RINF target gene candidate. To identify molecular mechanisms by which RINF might modulate TGF β -responsiveness, we reanalyzed

our previously published gene expression microarray datasets of K562 and UT7_{5.3} cells transduced with *shRINF#4* or *shControl* (7). A relatively discrete set of 193 gene candidates (*Supplementary Table S1*) were downregulated by *RINF* knockdown in both cell lines. Strikingly, ingenuity pathway analyses (IPA) revealed that this gene list was enriched in genes belonging to the TGF β signaling pathway (*Supplementary Figure S4C*). Considering its well-known inhibitory function on TGF β -signaling and its low expression in MDS, *SMAD7* appeared as a promising candidate for functional investigation (33, 34). Encouragingly, *SMAD7* was also downregulated by *RINF* knockdown in a third hematopoietic cell line, MV4-11 (*Supplementary Figure S4E, right panel*) and the downregulation of *SMAD7* was confirmed by quantitative RT-PCR analyses in K562 cells, even though the knockdown appeared moderate (~45%, $p < 0.001$) in these independent experiments (*Supplementary Figure S4D, right panel*).

We next measured *RINF* levels in two cell line models of erythroid maturation, the K562 cell line treated with hemin and the UT7_{5.3} cell line treated with EPO (39) after GM-CSF withdrawal (*Supplementary Figures S4A-B*). In agreement with our findings in primary cells, *RINF* expression was downregulated at the protein level upon treatment-induced erythroid differentiation, as early as after ten hours of treatment with hemin for K562 cells, and after two days with EPO for UT7_{5.3} cells (*Supplementary Figure S4A*). Moreover, *RINF* knockdown accelerated the erythroid maturation program (*Supplementary Figure S4B*) triggered by hemin in K562 cells (noticed after one day of treatment), or EPO in UT7_{5.3} cells (at four days of treatment). We also asked if *RINF* overexpression could delay hemoglobinization. To this end, we used the retroviral system MigR/IRES-GFP vector (1) (*Supplementary Figure S5A, upper panel*) to driving ectopic expression from full-length *RINF* cDNA. We found high constitutive expression of *RINF* at the protein level for both cell lines (*Supplementary Figure S5A*), without cell toxicity or consequences on cell proliferation (not shown). When we evaluated erythroid differentiation using benzidine staining (*Supplementary Figure S5B*) we found that *RINF* overexpression reduced hemoglobinization in both cell lines ($p < 0.001$), consistent with our knockdown experiments (Figures S4A-B).

RINF directly controls SMAD7 transcription. Our q-RT-PCR analyses demonstrated a robust increase (approximately five-fold) of *SMAD7* mRNA levels in cells overexpressing *RINF* (Figure 5A). To investigate if *RINF*-mediated induction of

SMAD7 mRNA was direct and occurred at the transcriptional level, we performed ChIP-experiments with anti-RINF antibodies. To this end, five sets of primers were designed for regions encompassing the *SMAD7* promoter (Figure 5B). As shown in Supplementary Figure S5C, we found that RINF binds directly to several regions (R1 to R4) of the *SMAD7* promoter and that this binding was increased in cells overexpressing *RINF*. Moreover, H3K4me3 histone marks (quantified by ChIP-qPCR, Figure 5C) were increased by RINF overexpression in four proximal regions (R2-R4) of the *SMAD7* transcription start site, indicating that chromatin was in a more active transcriptional state.

To investigate functional relationships between RINF and *SMAD7*, we next performed “rescue” experiments in K562 cells. We employed retroviral transduction to express *SMAD7* ectopically (Figure 5D). Our data demonstrated that *SMAD7* overexpression did not affect endogenous expression of RINF mRNA or protein (Figures 5E-F), but reduced hemoglobinization (Figure 5G, lane 3), mimicking the effect of ectopic RINF (Supplementary Figure S5B). We did not detect any change in cell viability or proliferation in the absence of hemin. Importantly, under conditions when *SMAD7* was ectopically expressed, RINF knockdown did not result in acceleration of hemoglobinization (Figure 5G). Taken together, these data indicate that RINF directly binds *SMAD7* promoter to control its expression, and that *SMAD7* is an effector downstream of RINF during erythroid maturation.

***RINF* and *SMAD7* mRNA expression are correlated in bone marrow from adult donors and patients with myelodysplastic syndrome.** We asked if *RINF* and *SMAD7* mRNA expression would correlate in primary human CD34+ cells from healthy donors and donors with myelodysplastic syndrome. To this end, we performed q-RT-PCR on CD34+ cells isolated from adult bone marrow donors (n=11) and found a highly significant correlation between *RINF* mRNA and *SMAD7* mRNA (Pearson, rho=0.684, p=0.02) (Figure 6A). We confirmed these data by analyzing an independent microarray dataset performed by Pellagatti *et al.* (46). As shown in Figure 6B, *RINF* and *SMAD7* mRNA levels strongly correlated in CD34+ cells isolated from healthy donors (Pearson’s rho=0.784, p<0.001, n=17). Interestingly, this correlation also reached statistical significance in CD34+ cells from MDS-5q patients (Pearson’s rho=0.603, p<0.001, n=47) but not in other MDS (without del(5q) or 5q-). We then compared the intensity of this correlation with the other 5q genes

(n=48 genes) from the CDR associated with high-risk, and the *CXXC5* probe sets were the one correlating the best with *SMAD7* probe set in both the healthy control group ($\rho=0.819$, and $\rho=0.784$) and MDS patients' cohort with del(5q) (*Supplementary Figure S6A*). Reinforcing the previously reported relevance of *SMAD7* in MDS pathophysiology (33, 34), *SMAD7* expression was not only extinguished in CD34+ MDS samples compared to normal samples (in the microarray dataset from Gerstung *et al.* (47), *Supplementary Figure S6C*), but patients with the lowest *SMAD7* expression have a significantly shorter overall survival (*Supplementary Figure S6D*).

RINF-silencing alters genome-wide hydroxymethylation. Since murine Rinf has recently been described as a necessary platform for Tet2 activity (15, 48), we wondered if RINF-silencing could impact the genome-wide 5hmC of human immature erythroid cells. As shown on Figure 6C, we observed a statistically significant loss of 5hmC detected by immunofluorescence or flow cytometry (Figure 6D) in knockdown primary cells.

A RINF/SMAD7 axis controls erythroid maturation and RBC expansion in primary cells. We next asked whether *RINF* knockdown would alter *SMAD7* expression in primary erythroid progenitors. Indeed, as shown in Figures 7A and B, a 60-70% knockdown of *RINF* led to a statistically significant knockdown of *SMAD7* (by 40-45%) in CD34 positive cells isolated from cord blood (n=3 donors, paired-T-test student $p<0.001$). Accordingly, a 30% knockdown of *SMAD7* protein level was detected by immunofluorescence in primary erythroid progenitors ($p<0.0001$) (Figure 7C). To directly demonstrate the relevance of this finely tuned regulation of *SMAD7* mediated by RINF protein in primary HSPCs, we next executed experiments in which we knocked down *RINF* and induced *SMAD7* expression ectopically. To this end, we first transduced CD34+ cells with the doxycycline-inducible vector pInducer21/*SMAD7*-HA (Figure 7D), and then with pTRIPUD3 vector that drives *shRNA/RINF* or *shRNA/Control* expression. Notably, in the presence of doxycycline, *SMAD7* expression (Figures 7E-F) robustly prevented both the *RINF* knockdown dependent accelerated maturation (Figure 7G) and reduction of RBC numbers (Figure 7H).

Discussion

An increasing body of work suggest that *RINF* is involved in the maturation, function, or development of various cell types such as neural stem cells (21), endothelial cells (49), myoblasts (14), myofibroblasts (24), osteoblasts (23), as well as in kidney development (22), wound healing (24), and hair-regrowth (50). Here, we report that the epigenetic factor RINF is a transcriptional regulator of *SMAD7* that fine-tunes TGF β sensitivity of erythroid progenitors, findings that bring insight into molecular barriers that may prevent effective erythropoiesis.

Our *in vitro* erythropoiesis experiments demonstrate that *RINF* is expressed in human erythroid progenitors (i.e. BFU-E and CFU-E) and proerythroblasts (ProE) but not in last stages. Loss of RINF expression does not affect cell viability or cell proliferation but accelerates erythroid maturation, and noticeably reduces RBC expansion. The average reduction in RBC number produced was estimated at ~45% (in the 6 experiments, Figure 2F) or one population doubling (i.e. 1.03, for the 4 donors #3-6), suggesting that RINF knockdown cells could skip one division (in presence of TGF β). Our data indicate that RINF-dependent erythroid maturation is also dependent on TGF β signaling at physiological levels. This cytokine, a potent inducer of erythroid maturation and growth inhibition (27, 29, 51, 52), is known to be enriched in the bone marrow microenvironment (53) and we surmise that RINF/*SMAD7* regulation axis that we describe here is likely to exert a more pronounced effect *in vivo* than in our serum-free culture conditions (i.e. without exogenous TGF β). Moreover, since the mechanism of action is *SMAD7*-dependent, *RINF* downregulation may also sensitize hematopoietic cells to other cytokines of the TGF β superfamily that may act on late-stage erythropoiesis *in vivo* (such as activins and GDF11) (54-57). In this manner, RINF could be one of the players regulating the dynamic balance between long-term expansion of the erythroid pool versus fast production of red blood cells for immediate physiological needs.

Two recent studies suggest that murine Rinf could act as an anchoring platform necessary for Tet2 (that lacks CXXC domain) and its activity at CpG sites (5'-hydroxymethylation), in plasmacytoid dendritic cells (pDCs)(15) and mouse embryonic stem cells (mESCs) (48). Conversely, an earlier study indicate that Rinf inhibits Tet2 and hydroxymethylation of genomic DNA in a caspase-dependent manner (in mESCs), and another one indicates that RINF binds to unmethylated (and

not methylated) CpGs (18), probably reflecting a more complex mechanism of action (12). We here provide the first experimental evidence (at our knowledge) that RINF can regulate genome-wide hydroxymethylation in human cells. The role for RINF that we describe here could pave the way to future studies that will be necessary to better understand the complex molecular epigenetic events that occur during normal and ineffective erythropoiesis (12, 15, 48, 58) and that may involve TET enzymes (59).

Several correlative studies have implicated *RINF* as a candidate in pathogenesis of MDS/AML (1, 3, 5, 60). MDS is a complex and heterogeneous malignant hemopathy in which the loss of several genes is suspected to lead to the pathophysiology (4). *CXXC5* is located at chromosome 5q31.2, in a commonly deleted region (CDR) associated with high-risk in MDS and AML (*Supplementary Figure S1*) (1, 4). Even though it is uncertain that *RINF* loss of expression in and of itself would be causative for development of preleukemia, our functional data in human CD34+ cells indicate that loss of *RINF* may exacerbate the cytopenia of the erythroid lineage. Of importance, the functional contribution of *Cxxc5/Rinf* loss of function to MDS development was recently described in a mouse model, in which its invalidation (by random insertional mutation) has been demonstrated to cooperate with *Egr1* haploinsufficiency to promote MDS (4). This study indicates that *Cxxc5* gene could act as a tumor suppressor gene contributing to del(5q) myeloid neoplasms. Although *RINF/CXXC5* is not frequently mutated in patients with MDS/AML (60), several recurrent chromosomal anomalies or gene mutations could lead to *RINF* loss of expression and contribute in the long-term to ineffective erythropoiesis and/or myeloid transformation (legend of *Supplementary Figure S7*).

Our findings point to a molecular mechanism in which RINF functions as a negative modulator of TGF β signaling by upregulating and maintaining SMAD7. Consequently, *RINF* loss of expression will sensitize human erythroid progenitors to autocrine/paracrine TGF β -induced growth inhibition as a mechanism to prevent expansion of the erythroid pool. Our data also suggest that this signaling would be at least partly mediated through SMAD2 activation and phosphorylation (*Figure 4C*), in agreement with previous report in human bone marrow (35). The *in vivo* relevance of these findings is strengthened by correlated *RINF* and *SMAD7* mRNA expression in human BM CD34+ cells. This correlation was also observed in MDS patients with del(5q) or 5q- but not in other MDS, in which *SMAD7* was silenced in most of patient

samples, whatever the *RINF* expression level. This extinction of *SMAD7* in other MDS could be explained by various mechanisms such as miRNA21 expression (33). However, since *RINF* knockdown leads to global loss of 5hmC, it is tempting to speculate that the loss of *RINF* could lead to hypermethylation of *SMAD7* promoter in HSPCs, a process that could have gone unnoticed in previous studies. Interestingly, *SMAD7* promoter is known to be silenced by hypermethylation in non-hematopoietic tissues, and hypomethylating agents such as azacytidine (AZA) can revert this hypermethylation and alleviates TGF β -induced diseases in several pathological models such as atherosclerosis (61), as well as renal- (62) and liver-fibrosis (63). Thus, further studies should investigate whether *SMAD7* promoter is demethylated after treatment with hypomethylating agents (AZA or decitabine) in humans HSPCs isolated from MDS patients, and to what extent this mechanism could contribute to the efficacy of these epigenetic drugs to sustain erythropoiesis on the long-term way. Supporting this hypothesis, *TET2* mutation, is a good indicator of response to azacytidine treatment (64).

Previous studies have underscored the importance of *SMAD7* regulation in MDS (33, 34). Our present data complement this work by demonstrating, *firstly*, a direct transcriptional mechanism by which *SMAD7* is regulated in hematopoietic cells and, *second*, that the residual level of *SMAD7* in MDS is associated with the severity of the disease and patient survival (*Supplementary Figure S6D*). These data are particularly relevant in the context of understanding how TGF β responsiveness in the bone marrow microenvironment is fine-tuned (28), and in a context-specific manner (30). Even though this mechanism is unlikely to be solely responsible for TGF β -hypersensitivity, our findings pave the way for novel research avenues in hemopathies characterized by ineffective erythropoiesis such as Myelofibrosis (65), Fanconi Anemia (66), β -Thalassemia (57, 67), or MDS (55), and for which novel therapies based on TGF β -inhibitors are particularly promising (56, 68). Finally, beyond its role during erythropoiesis, the *RINF/SMAD7* regulation axis here described, could at least partly contribute to the pleiotropic effects of *RINF* described during wound healing (24), fibrosis (61-63, 69), immunity (15), and tumor biology (1, 3, 6-9), and deserves to be investigated in these physiological processes.

Acknowledgements

This work was supported by INSERM, Paris-Descartes University, Ligue Nationale Contre le Cancer (LNCC), Cochin Institute, and the Laboratory of Excellence GR-EX. AA was supported by LNCC, Société Française d'Hématologie, Fondation pour la Recherche Médicale, and Boehringer Society (travel grant). GM was supported by fellowship grant from the Ministère de l'Enseignement supérieur et de la Recherche and Société Française d'Hématologie. E.S-B is supported by the National Center for Scientific Research (CNRS) and her team was supported by LNCC and Fondation de France. We are also grateful for the support from the French-Norwegian exchange program (Aurora to F.P.). Authors thanks Pr. Jérôme Larghero and Thomas Domet from *Centre de Ressources Biologiques / Banque de Sang de Cordon de l'AP-HP* (Saint-Louis Hospital) for providing cord blood units, Anne Dubart-Kupperschmitt for the pTRIPDU3/GFP vector (40), Johan Lillehaug, Øystein Bruserud, Camille Humbert, Pierre de la Grange (Genosplice), Eric Nguyen, Sabrina Bondu, Alexandre Artus, Amandine Houvert, Nadège Bercovici, Emmanuel Donnadiou, Franck Letourneur, Brigitte Izac, Sebastien Jacques, and Alain Trautmann for technical assistance and scientific discussions, as well as *Life Science Editors* for reading and editing the manuscript. We are grateful to the staff of the IMAG'IC, CYBIO, and GENOM'IC facilities of the Cochin Institute.

Authorship Contributions

A.A. and G.M. performed most of the experiments. D.S and S.Z performed the ChIP experiments. Y.Z., V.F., I.M., and E-F.G. performed some experiments and discussed the data. A.A., G.M., F.V., E.L., M.F., A.K., O.D., and C.L. discussed the data. E.S-B, I.D-F, D.B., O.H., and P.M. supervised the study. F.P. conceived the study and supervised the experiments. A.A., G.M., and F.P. analyzed the data, made the Figures and wrote the manuscript.

Conflicts of Interest Disclosures

The authors declare no potential conflict of interest. Frederic Pendino is the main coinventor of patents describing methods for *RINF* mRNA detection (WO2009/151337 and WO2012/010661).

References

1. Pendino F, Nguyen E, Jonassen I, et al. Functional involvement of RINF, retinoid-inducible nuclear factor (CXXC5), in normal and tumoral human myelopoiesis. *Blood*. 2009;113(14):3172-3181.
2. Jerez A, Gondek LP, Jankowska AM, et al. Topography, clinical, and genomic correlates of 5q myeloid malignancies revisited. *J Clin Oncol*. 2012;30(12):1343-1349.
3. Kuhn A, Valk PJ, Sanders MA, et al. Downregulation of the Wnt inhibitor CXXC5 predicts a better prognosis in acute myeloid leukemia. *Blood*. 2015;125(19):2985-2994.
4. Stoddart A, Qian Z, Fernald AA, et al. Retroviral insertional mutagenesis identifies the del(5q) genes, CXXC5, TIFAB and ETF1, as well as the Wnt pathway, as potential targets in del(5q) myeloid neoplasms. *Haematologica*. 2016;101(6):e232-236
5. Treppendahl MB, Mollgard L, Hellstrom-Lindberg E, Cloos P, Gronbaek K. Downregulation but lack of promoter hypermethylation or somatic mutations of the potential tumor suppressor CXXC5 in MDS and AML with deletion 5q. *Eur J Haematol*. 2013;90(3):259-260.
6. Astori A, Fredly H, Aloysius TA, et al. CXXC5 (retinoid-inducible nuclear factor, RINF) is a potential therapeutic target in high-risk human acute myeloid leukemia. *Oncotarget*. 2013;4(9):1438-1448.
7. Bruserud O, Reikvam H, Fredly H, et al. Expression of the potential therapeutic target CXXC5 in primary acute myeloid leukemia cells - high expression is associated with adverse prognosis as well as altered intracellular signaling and transcriptional regulation. *Oncotarget*. 2015;6(5):2794-2811.
8. Benedetti I, De Marzo AM, Geliebter J, Reyes N. CXXC5 expression in prostate cancer: implications for cancer progression. *Int J Exp Pathol*. 2017;98(4):234-243.
9. Knappskog S, Myklebust LM, Busch C, et al. RINF (CXXC5) is overexpressed in solid tumors and is an unfavorable prognostic factor in breast cancer. *Ann Oncol*. 2011;22(10):2208-2215.
10. Zhang M, Wang R, Wang Y, et al. The CXXC finger 5 protein is required for DNA damage-induced p53 activation. *Sci China C Life Sci*. 2009;52(6):528-538.
11. Aras S, Pak O, Sommer N, et al. Oxygen-dependent expression of cytochrome c oxidase subunit 4-2 gene expression is mediated by transcription factors RBPJ, CXXC5 and CHCHD2. *Nucleic Acids Res*. 2013;41(4):2255-2266.
12. Ko M, An J, Bandukwala HS, et al. Modulation of TET2 expression and 5-methylcytosine oxidation by the CXXC domain protein IDAX. *Nature*. 2013;497(7447):122-126.
13. L'Hote D, Georges A, Todeschini AL, et al. Discovery of novel protein partners of the transcription factor FOXL2 provides insights into its physiopathological roles. *Hum Mol Genet*. 2012;21(14):3264-3274.
14. Li G, Ye X, Peng X, et al. CXXC5 regulates differentiation of C2C12 myoblasts into myocytes. *J Muscle Res Cell Motil*. 2014;35(5-6):259-265.
15. Ma S, Wan X, Deng Z, et al. Epigenetic regulator CXXC5 recruits DNA demethylase Tet2 to regulate TLR7/9-elicited IFN response in pDCs. *J Exp Med*. 2017;214(5):1471-1491.
16. Marshall PA, Hernandez Z, Kaneko I, et al. Discovery of novel vitamin D receptor interacting proteins that modulate 1,25-dihydroxyvitamin D3 signaling. *J Steroid Biochem Mol Biol*. 2012;132(1-2):147-159.
17. Yasar P, Ayaz G, Muyan M. Estradiol-Estrogen Receptor alpha Mediates the Expression of the CXXC5 Gene through the Estrogen Response Element-Dependent Signaling Pathway. *Sci Rep*. 2016;6:37808.
18. Ayaz G, Razizadeh N, Yasar P, et al. CXXC5 as an unmethylated CpG dinucleotide binding protein contributes to estrogen-mediated cellular proliferation. *Sci Rep*. 2020;10(1):5971.
19. Long HK, Blackledge NP, Klose RJ. ZF-CxxC domain-containing proteins, CpG islands and the chromatin connection. *Biochem Soc Trans*. 2013;41(3):727-740.
20. Melamed P, Yosefzon Y, David C, Tsukerman A, Pnueli L. Tet Enzymes, Variants, and Differential Effects on Function. *Front Cell Dev Biol*. 2018;6:22.
21. Andersson T, Sodersten E, Duckworth JK, et al. CXXC5 is a novel BMP4-regulated modulator of Wnt signaling in neural stem cells. *J Biol Chem*. 2009;284(6):3672-3681.

22. Kim MS, Yoon SK, Bollig F, et al. A novel Wilms tumor 1 (WT1) target gene negatively regulates the WNT signaling pathway. *J Biol Chem.* 2010;285(19):14585-14593.
23. Kim HY, Yoon JY, Yun JH, et al. CXXC5 is a negative-feedback regulator of the Wnt/beta-catenin pathway involved in osteoblast differentiation. *Cell Death Differ.* 2015;22(6):912-920.
24. Lee SH, Kim MY, Kim HY, et al. The Dishevelled-binding protein CXXC5 negatively regulates cutaneous wound healing. *J Exp Med.* 2015;212(7):1061-1080.
25. Pardali K, Moustakas A. Actions of TGF-beta as tumor suppressor and pro-metastatic factor in human cancer. *Biochim Biophys Acta.* 2007;1775(1):21-62.
26. Siegel PM, Massague J. Cytostatic and apoptotic actions of TGF-beta in homeostasis and cancer. *Nat Rev Cancer.* 2003;3(11):807-821.
27. Zermati Y, Varet B, Hermine O. TGF-beta1 drives and accelerates erythroid differentiation in the epo-dependent UT-7 cell line even in the absence of erythropoietin. *Exp Hematol.* 2000;28(3):256-266.
28. Blank U, Karlsson S. TGF-beta signaling in the control of hematopoietic stem cells. *Blood.* 2015;125(23):3542-3550.
29. Krystal G, Lam V, Dragowska W, et al. Transforming growth factor beta 1 is an inducer of erythroid differentiation. *J Exp Med.* 1994;180(3):851-860.
30. Ruscetti FW, Akel S, Bartelmez SH. Autocrine transforming growth factor-beta regulation of hematopoiesis: many outcomes that depend on the context. *Oncogene.* 2005;24(37):5751-5763.
31. Derynck R, Zhang YE. Smad-dependent and Smad-independent pathways in TGF-beta family signalling. *Nature.* 2003;425(6958):577-584.
32. Shi Y, Massague J. Mechanisms of TGF-beta signaling from cell membrane to the nucleus. *Cell.* 2003;113(6):685-700.
33. Bhagat TD, Zhou L, Sokol L, et al. miR-21 mediates hematopoietic suppression in MDS by activating TGF-beta signaling. *Blood.* 2013;121(15):2875-2881.
34. Zhou L, McMahon C, Bhagat T, et al. Reduced SMAD7 leads to overactivation of TGF-beta signaling in MDS that can be reversed by a specific inhibitor of TGF-beta receptor I kinase. *Cancer Res.* 2011;71(3):955-963.
35. Zhou L, Nguyen AN, Sohal D, et al. Inhibition of the TGF-beta receptor I kinase promotes hematopoiesis in MDS. *Blood.* 2008;112(8):3434-3443.
36. Freyssinier JM, Lecoq-Lafon C, Amsellem S, et al. Purification, amplification and characterization of a population of human erythroid progenitors. *Br J Haematol.* 1999;106(4):912-922.
37. Gautier EF, Ducamp S, Leduc M, et al. Comprehensive Proteomic Analysis of Human Erythropoiesis. *Cell Rep.* 2016;16(5):1470-1484.
38. Inman GJ, Nicolas FJ, Callahan JF, et al. SB-431542 is a potent and specific inhibitor of transforming growth factor-beta superfamily type I activin receptor-like kinase (ALK) receptors ALK4, ALK5, and ALK7. *Mol Pharmacol.* 2002;62(1):65-74.
39. Millot GA, Vainchenker W, Dumenil D, Svinarchuk F. Differential signalling of NH2-terminal flag-labelled thrombopoietin receptor activated by TPO or anti-FLAG antibodies. *Cell Signal.* 2004;16(3):355-363.
40. Amsellem S, Ravet E, Fichelson S, Pflumio F, Dubart-Kupperschmitt A. Maximal lentivirus-mediated gene transfer and sustained transgene expression in human hematopoietic primitive cells and their progeny. *Mol Ther.* 2002;6(5):673-677.
41. Papageorgis P, Lambert AW, Ozturk S, et al. Smad signaling is required to maintain epigenetic silencing during breast cancer progression. *Cancer Res.* 2010;70(3):968-978.
42. Lecoq-Lafon C, Verdier F, Fichelson S, et al. Erythropoietin induces the tyrosine phosphorylation of GAB1 and its association with SHC, SHP2, SHIP, and phosphatidylinositol 3-kinase. *Blood.* 1999;93(8):2578-2585.
43. Merryweather-Clarke AT, Atzberger A, Soneji S, et al. Global gene expression analysis of human erythroid progenitors. *Blood.* 2011;117(13):e96-108.

44. An X, Schulz VP, Li J, et al. Global transcriptome analyses of human and murine terminal erythroid differentiation. *Blood*. 2014;123(22):3466-3477.
45. Keller MA, Addya S, Vadigepalli R, et al. Transcriptional regulatory network analysis of developing human erythroid progenitors reveals patterns of coregulation and potential transcriptional regulators. *Physiol Genomics*. 2006;28(1):114-128.
46. Pellagatti A, Cazzola M, Giagounidis AA, et al. Gene expression profiles of CD34+ cells in myelodysplastic syndromes: involvement of interferon-stimulated genes and correlation to FAB subtype and karyotype. *Blood*. 2006;108(1):337-345.
47. Gerstung M, Pellagatti A, Malcovati L, et al. Combining gene mutation with gene expression data improves outcome prediction in myelodysplastic syndromes. *Nat Commun*. 2015;6:5901.
48. Ravichandran M, Lei R, Tang Q, et al. Rinf Regulates Pluripotency Network Genes and Tet Enzymes in Embryonic Stem Cells. *Cell Rep*. 2019;28(8):1993-2003.
49. Kim HY, Yang DH, Shin SW, et al. CXXC5 is a transcriptional activator of Flk-1 and mediates bone morphogenic protein-induced endothelial cell differentiation and vessel formation. *FASEB J*. 2014;28(2):615-626.
50. Lee SH, Seo SH, Lee DH, Pi LQ, Lee WS, Choi KY. Targeting of CXXC5 by a Competing Peptide Stimulates Hair Regrowth and Wound-Induced Hair Neogenesis. *J Invest Dermatol*. 2017;137(11):2260-2269.
51. Akel S, Petrow-Sadowski C, Laughlin MJ, Ruscetti FW. Neutralization of autocrine transforming growth factor-beta in human cord blood CD34(+)CD38(-)Lin(-) cells promotes stem-cell-factor-mediated erythropoietin-independent early erythroid progenitor development and reduces terminal differentiation. *Stem Cells*. 2003;21(5):557-567.
52. Gao X, Lee HY, da Rocha EL, et al. TGF-beta inhibitors stimulate red blood cell production by enhancing self-renewal of BFU-E erythroid progenitors. *Blood*. 2016;128(23):2637-2641.
53. Thompson NL, Flanders KC, Smith JM, Ellingsworth LR, Roberts AB, Sporn MB. Expression of transforming growth factor-beta 1 in specific cells and tissues of adult and neonatal mice. *J Cell Biol*. 1989;108(2):661-669.
54. Carrancio S, Markovics J, Wong P, et al. An activin receptor IIA ligand trap promotes erythropoiesis resulting in a rapid induction of red blood cells and haemoglobin. *Br J Haematol*. 2014;165(6):870-882.
55. Bewersdorf JP, Zeidan AM. Transforming growth factor (TGF)-beta pathway as a therapeutic target in lower risk myelodysplastic syndromes. *Leukemia*. 2019;33(6):1303-1312.
56. Fenaux P, Kiladjian JJ, Platzbecker U. Luspatercept for the treatment of anemia in myelodysplastic syndromes and primary myelofibrosis. *Blood*. 2019;133(8):790-794.
57. Suragani RN, Cawley SM, Li R, et al. Modified activin receptor IIB ligand trap mitigates ineffective erythropoiesis and disease complications in murine beta-thalassemia. *Blood*. 2014;123(25):3864-3872.
58. Lio CJ, Yuita H, Rao A. Dysregulation of the TET family of epigenetic regulators in hematopoietic malignancies. *Blood*. 2019;134 (18): 1487-1497.
59. Yan H, Wang Y, Qu X, et al. Distinct roles for TET family proteins in regulating human erythropoiesis. *Blood*. 2017;129(14):2002-2012.
60. Gelsi-Boyer V, Trouplin V, Adelaide J, et al. Mutations of polycomb-associated gene ASXL1 in myelodysplastic syndromes and chronic myelomonocytic leukaemia. *Br J Haematol*. 2009;145(6):788-800.
61. Wei L, Zhao S, Wang G, et al. SMAD7 methylation as a novel marker in atherosclerosis. *Biochem Biophys Res Commun*. 2018;496(2):700-705.
62. Yang Q, Chen HY, Wang JN, et al. Alcohol promotes renal fibrosis by activating Nox2/4-mediated DNA methylation of Smad7. *Clin Sci (Lond)*. 2020;134(2):103-122.
63. Bian EB, Huang C, Wang H, et al. Repression of Smad7 mediated by DNMT1 determines hepatic stellate cell activation and liver fibrosis in rats. *Toxicol Lett*. 2014;224(2):175-185.

64. Itzykson R, Kosmider O, Cluzeau T, et al. Impact of TET2 mutations on response rate to azacitidine in myelodysplastic syndromes and low blast count acute myeloid leukemias. *Leukemia*. 2011;25(7):1147-1152.
65. Yue L, Bartenstein M, Zhao W, et al. Efficacy of ALK5 inhibition in myelofibrosis. *JCI insight*. 2017;2(7):e90932.
66. Zhang H, Kozono DE, O'Connor KW, et al. TGF-beta Inhibition Rescues Hematopoietic Stem Cell Defects and Bone Marrow Failure in Fanconi Anemia. *Cell Stem Cell*. 2016;18(5):668-681.
67. Piga A, Perrotta S, Gamberini MR, et al. Luspatercept improves hemoglobin levels and blood transfusion requirements in a study of patients with beta-thalassemia. *Blood*. 2019;133(12):1279-1289.
68. Verma A, Suragani RN, Aluri S, et al. Biological basis for efficacy of activin receptor ligand traps in myelodysplastic syndromes. *J Clin Invest*. 2020;130(2):582-589.
69. Cheng W, Wang F, Feng A, Li X, Yu W. CXXC5 Attenuates Pulmonary Fibrosis in a Bleomycin-Induced Mouse Model and MLFs by Suppression of the CD40/CD40L Pathway. *Biomed Res Int*. 2020;2020:7840652.

Table 1.

GENE NAME	Probeset	Pearson's (correl coeff.)	Gene location at 5q
CXXC5	222996_s_at	0,819	High-Risk CDR
CXXC5	233955_x_at	0,784	High-Risk CDR
HSPA9	200690_at	0,729	High-Risk CDR
CXXC5	224516_s_at	0,705	High-Risk CDR
ETF1	201574_at	0,640	High-Risk CDR
IL17B	220273_at	0,554	Low-Risk CDR
SPARC	212667_at	0,532	Low-Risk CDR
ETF1	201573_s_at	0,508	High-Risk CDR
MATR3	238993_at	0,452	High-Risk CDR
CD74	1567627_at	0,424	Low-Risk CDR
TGFBI	201506_at	0,422	High-Risk CDR
PDGFRB	202273_at	0,391	Low-Risk CDR
EGR1	201694_s_at	0,373	High-Risk CDR
SIL1	218436_at	0,355	High-Risk CDR
PSD2	223536_at	0,352	High-Risk CDR
UBE2D2	201345_s_at	0,350	High-Risk CDR
AFAP1L1	1555542_at	0,343	Low-Risk CDR
SH3TC2	233561_at	0,337	Low-Risk CDR
CSF1R	203104_at	0,335	Low-Risk CDR
TMEM173	224916_at	0,312	High-Risk CDR
PAIP2	222983_s_at	0,289	High-Risk CDR
HNRNPA0	229083_at	0,271	High-Risk CDR
CDX1	206430_at	0,259	Low-Risk CDR
EGR1	201693_s_at	0,214	High-Risk CDR
FAM53C	218023_s_at	0,193	High-Risk CDR
AFAP1L1	226955_at	0,184	Low-Risk CDR
MZB1	223565_at	0,180	High-Risk CDR
CARMN	231987_at	0,180	Low-Risk CDR
CDC23	223651_x_at	0,178	High-Risk CDR
CTNNA1	200764_s_at	0,176	High-Risk CDR
REEP2	205331_s_at	0,167	High-Risk CDR
SPATA24	238027_at	0,137	High-Risk CDR
RPS14	208645_s_at	0,134	Low-Risk CDR
LRRTM2	206408_at	0,128	High-Risk CDR
NRG2	242303_at	0,122	High-Risk CDR
GFRA3	229936_at	0,114	High-Risk CDR
BRD8	242265_at	0,096	High-Risk CDR
CDC25C	216914_at	0,083	High-Risk CDR
SH3TC2	240966_at	0,052	Low-Risk CDR
HNRNPA0	201055_s_at	0,041	High-Risk CDR
KDM3B	210878_s_at	0,041	High-Risk CDR
GRPEL2	238427_at	0,040	Low-Risk CDR

TMEM173	224929_at	0,028	High-Risk CDR
EGR1	227404_s_at	0,013	High-Risk CDR
DNAJC18	227166_at	-0,018	High-Risk CDR
FAM13B	218518_at	-0,025	High-Risk CDR
TRPC7	208589_at	-0,035	High-Risk CDR
ABLIM3	205730_s_at	-0,068	Low-Risk CDR
NME5	206197_at	-0,081	High-Risk CDR
KLHL3	1555110_a_at	-0,094	High-Risk CDR
WNT8A	224259_at	-0,116	High-Risk CDR
SPATA24	1558641_at	-0,137	High-Risk CDR
LECT2	207409_at	-0,143	High-Risk CDR
IL9	208193_at	-0,157	High-Risk CDR
ECSCR	227780_s_at	-0,169	High-Risk CDR
KLHL3	221221_s_at	-0,179	High-Risk CDR
PKD2L2	221118_at	-0,187	High-Risk CDR
ZNF300	228144_at	-0,191	Low-Risk CDR
SMAD5	225223_at	-0,221	High-Risk CDR
TRPC7	202363_at	-0,241	High-Risk CDR
SH3TC2	219710_at	-0,315	Low-Risk CDR
SLC23A1	223732_at	-0,319	High-Risk CDR
CDC25C	205167_s_at	-0,351	High-Risk CDR
MYOT	219728_at	-0,386	High-Risk CDR
UBE2D2	201344_at	-0,538	High-Risk CDR
PCYOX1L	218953_s_at	-0,559	Low-Risk CDR
GRPEL2	226881_at	-0,614	Low-Risk CDR
KIF20A	218755_at	-0,786	High-Risk CDR

Table 1. List of genes located at 5q and their correlation with *SMAD7* expression in primary human CD34+ isolated from bone marrow of healthy donors (n=17). All the genes were selected because they are located within one of the two Commonly Deleted Region (CDRs) at chromosome 5q. Their mRNA expression level was extracted from GDS3795 microarray study performed by Pellagatti *et al.*(46) For each gene, a Pearson's correlation coefficient was calculated by comparing its expression to the one of *SMAD7*, and the list was ranked by descending order.

Figure legends

Figure 1. *RINF* is expressed in early erythropoiesis and downregulated during maturation of human HSPCs. (A) Schematic representation of cell culture experiments. CD34⁺ cells (from cord blood or adult bone marrow) were first amplified for 7 days (D-7 to D0, in black) in presence of SCF, IL3 and IL6. Cells were sorted for CD36 expression and EPO was added to the medium in order to trigger the late-stage of erythropoiesis, a phase lasting 11-17 days (in red). (B) For morphological characterization, cells were cytopun and stained with May-Grünwald-Giemsa (MGG) and the repartition of cells among differentiation stages indicated in percentage (*right panel*): early progenitor (Prog), proerythroblast (ProE), early and late basophilic erythroblast (Baso), polychromatophilic erythroblast (Poly), orthochromatic erythroblast (Ortho), and Reticulocyte (Retic). Scale bar represents 20 μ m. The experiment presented here is representative of four distinct experiments which were performed on cells isolated from human cord blood. *RINF* protein was detected by western-blot analysis (*lower panel*). HSC70 protein was used as a loading control. For each lane, 40 μ g of protein was loaded. Bands intensities were quantified with the Multigauge Software and the relative expression of *RINF* protein (normalized to the loading control) is indicated above the image in percentage of Day 3. (C) *CXXC5* mRNA expression was measured by quantitative RT-PCR. *CXXC5* mRNA expression is downregulated from the basophilic erythroblast stage. Here, CD34⁺ cells were isolated from adult human bone marrow, but a similar expression profile was observed with cells isolated from cord blood.

Figure 2. *RINF* knockdown leads to a reduction in total number of red blood cells generated from human HSPCs. (A) Schematic representation of shRNA-mediated *RINF* loss of expression experiment. CD34⁺ cells isolated from cord-blood were transduced at day 5 of the amplification step with the lentiviral vector pTRIPDU3/GFP expressing a shRNA targeting *RINF* mRNA expression (*shRNA/RINF*) or a non-target sequence (*shRNA/Control*). GFP-positive cells were sorted by FACS two days after transduction. At that time, cell culture kinetics with less than 80% of CD36⁺ cells or more than 5% of cell death (trypan blue) were not pursued. EPO was added after sorting and maintained for 17-18 days. (B) *RINF*

knockdown efficiency was estimated at the mRNA level by quantitative RT-PCR at D0 of EPO (i.e. right after GFP-sorting). Here, a representative experiment out of seven independent experiments performed. A student T-test was performed. **(C)** RINF knockdown efficiency was estimated at the protein level (specific band at 33 KDa) by Western-Blot analysis (see Methods). P85 (PIK3R1) was used as loading control (42). In order to obtain enough protein material, CD34+ cells were here amplified from a pool of five cord-blood units (5 donors). For each lane, protein extracted from 3.10^5 cells was loaded right after sorting, at D0 of EPO. Bands intensities were quantified with the Multigauge software and the relative expression of RINF protein (normalized to the loading control) is indicated above the image in percentage of the *shControl* condition **(D)** Cell viability was assessed by trypan blue exclusion method at 2 or 3 days after GFP-sorting. Histogram represents the percentage of trypan blue negative cells. **(E)** Cell culture growth was monitored in six independent experiments (here, cord-blood donors). The proliferation curve, expressed in cumulative population doublings, is shown for one representative experiment (donor#6, *left panel*). For most of experiments (or “donors”), CD34+ cells from several cord blood units (2-4) were pooled. Cell expansion fold (*right panel*) was calculated by dividing the absolute output number of expanded RBC after 17-18 days of culture with EPO by the respective number on day 0 (of treatment with EPO). An average reduction of ~45% was noticed in *RINF* knockdown condition (paired T-test, $p=0.04$, $n=6$). For two out of the six donors (Donors 1 and 2, in grey), the *shRNA-RINF#4*-induced reduction in the total number of RBC was not obvious in our serum-free conditions (and absence of TGF β). These two unresponsive donors were also treated with TGF β , and data are presented on Figures 4E-F **(F)** Photographs of the five larger BFU-E colonies for the two conditions, here for one representative donor at day 11 (*left panel*). Two days after transduction (day 5 of expansion), GFP-expressing cells were sorted by flow cytometry and five hundred cells per well were seeded into methycellulose containing cytokines (IL3, SCF, G-CSF, GM-CSF, and EPO) enabling the formation of erythroid colonies (BFU-Es). BFU-E colonies were imaged between 11 and 14 days of culture numerated from three donors ($n=103$ and $n=110$, respectively). The size of each BFU-E colony was determined by imaging analysis by using ImageJ/Fiji. Average size of colonies was smaller for *shRINF#4*-transduced conditions (unpaired two-tailed T-test, $P<0.01$). **(G)** CD34+ were isolated from bone marrow of healthy donors ($n=4$). Colony Forming Cells (CFC) were

counted between 11 and 14 days of culture. Histograms give the mean number of small and large BFU-E-derived colonies. The error bars indicate donor-to-donor variability for each vector (\pm SEM). To assess the effect of shRNA-mediated *RINF* silencing, a paired comparison was performed by using a Fisher's exact test.

Figure 3. *RINF* silencing accelerates erythroid maturation of human primary erythroblasts. Erythroid differentiation was monitored for six separate experiments performed with CD34+ cells isolated either from cord-blood (n=3) or adult bone marrow (n=3), by using different techniques. Despite a donor-to-donor variability in the maturation kinetics, we consistently observed an accelerated maturation for *RINF* knockdown conditions. Of note, data presented on Figures 3B and 3C correspond to two distinct donors. **(A)** Hemoglobin production of primary cells was estimated by benzidine assay. CD34+ cells isolated from cord-blood were transduced with the lentiviral vector pTRIPDU3/GFP expressing a shRNA targeting *RINF* mRNA expression (*shRNA/RINF*) or a non-target sequence (*shRNA/Control*). GFP-positive cells were sorted by FACS two days after transduction. The black arrow at day 6 indicates the kinetic point selected for one representative picture on the right panel. **(B)** Morphological analysis has been performed after May-Grünwald-Giemsa (MGG) staining. In order to know whether the observed acceleration of maturation was statistically significant, a Fisher's exact test was performed. For this, enumerated erythroid cells (n=454) were segregated into either early (Progenitors, ProE, Baso) or late (Polychro, Ortho, Retic) group. P value <0.0001. **(C)** Erythroid maturation was also monitored by flow cytometry by using anti-CD71 and anti-glycophorin-A (GPA/CD235a) labelled antibodies (*upper panel*) or using anti-CD49d and anti-Band3 antibodies (*lower panel*). Scattergrams of two representative experiments are shown (n=6), indicating a more mature population of cells for *shRNA-RINF* conditions either at day 6 of EPO (GPA positive, *upper panel*), or at day 10 of EPO (Band3 positive, *lower panel*). **(D)** *RINF* loss of expression in human CD34+ is also associated with an accelerated downregulation of *cKIT* and *PU.1*. Histogram of relative mRNA expression determined by quantitative RT-PCR for *RINF* (*left panel*), *cKIT* (*middle panel*) and *PU.1* (*right panel*). *PU.1* and *c-KIT* mRNA are known to be downregulated upon erythroid maturation and are here used as molecular markers. In *RINF* knockdown cells (black bars), the downregulation of *c-KIT* and *PU.1* mRNA preceded of at least three days the one observed in control cells, in agreement with an

accelerated maturation. **(E) and (F)** In order to validate our data with a second shRNA sequence targeting *RINF*, CD34+ cells were transduced with the lentiviral vector pLKO-Tet-ON/shRNA vector, allowing the doxycycline-inducible expression of shRNA sequences targeting *RINF* expression (*shRINF#3* or *shRINF#4*). GFP-positive cells were sorted by FACS two days after transduction and cultured in presence of doxycycline at 0.2 µg/mL **(E)**. The left histogram represents the level of *RINF* mRNA obtained with the two sequences targeting *RINF* (*shRINF#3* or *shRINF#4*) and detected by quantitative RT-PCR. Values are expressed in percentage of the *shControl* condition. The right histogram represents the percentage of benzidine positive cells obtained with the three shRNA sequences (*shControl*, *shRINF#3*, or *shRINF#4*). **(F)**. Cell culture growth was monitored in one additional experiment (here, a pool from 2 adult donors) and the cumulative population doublings is indicated (*left panel*) as well as the fold expansion between D0 and D17 of EPO (*right panel*).

Figure 4. *RINF* knockdown-induced phenotype is TGFβ-dependent in human primary cells. **(A)** TGFβ ligand and TGFBR1 expression profiles were extracted from our proteomic analyses of cord-blood derived CD34+ cells (Gauthier *et al.* (37)). TGFβ1 and TGFBR1 were highly expressed at the Prog1-Prog2 stage (i.e. D1-D3 of EPO), suggesting an important autocrine signaling in early steps of erythroid maturation in our serum-free culture conditions. **(B)** Cord-blood derived CD34+ cells were transduced at day 5 of the amplification step, GFP-sorted (2 days post-transduction), and EPO was added for 8 to 12 days, in presence of TGFβ inhibitor (SB431542, at 10µM) or vehicle control (DMSO). Histograms (*left panels*) represent FACS analyses of the whole cell suspension for CD49d and Band3 antigens (here at 10 days of EPO). In order to know whether the observed acceleration of maturation was statistically significant, student-T tests were performed. Scattergrams (*right panel*), presents one representative experiment out of three separate experiments performed. **(C)** *RINF* (*left panel*) and phosphorylated SMAD2 (*right panel*) protein levels were estimated at 8h and 72h of treatment with EPO by Western-Blot analysis (see Methods), on the same kinetics than the one presented on Figure 2C (pool of five cord-blood units). As expected, the specific band for phospho-SMAD2 (Ser465/467), indicated by an arrow, was not detected in presence of TGFβ inhibitor

(SB431542, at 10 μ M). P85 (PIK3R1) was used as loading control (*lower panel*) from the same blot, and at a lower exposition time than on Figure 2C (42). For each lane, protein extracted from 3.10⁵ cells was loaded. Bands intensities were quantified with multigaugue software and the relative phospho-SMAD2 expression are shown (*right panel*). **(D)** TGF β (5 ng/mL) or vehicle control (DMSO) was added at 11 days of EPO and the proliferation monitored during one additional week, until day 18 of EPO. For both treated (in red) and untreated conditions (in black), the RBC expansion was lower for *RINF* silenced cells. The kinetics shown corresponds to Donor#4 of Figure 2E (*right panel*). Cell culture growth is represented in population doublings (*left panel*) or cumulative cell number at D18 of EPO (*right panel*). **(E and F)** The two donors (1 and 2) that seemed not sensitive to *RINF* knockdown in the absence of TGF β (Figure 2E) were treated with variable dose of TGF β 1 (ranging from 0.1 to 2.5 ng/mL), and scattergrams (*upper panel*) represent FACS analyses of the whole cell suspension labelled for both GPA and CD71 antigens and the corresponding dose-effect curves are shown (*lower panel*). One representative scattergram at the lowest concentration (0.1 ng/mL) of exogenous TGF β 1 is shown. **(F)** Histogram plots represent GPA-acquisition (*upper panel*) that is more advanced in *RINF* knockdown cells (here at day 6 of EPO) in presence of low doses of TGF β 1 (here at 0.1 ng/mL). Cell growth was also monitored, and the total number of RBC numerated at day 17 of EPO are presented on histograms (*lower panel*).

Figure 5. SMAD7 is a direct transcriptional target of RINF. **(A)** The MigR1 retroviral system was used to constitutively express *RINF* in K562 cell line. A few days after transduction, cells were sorted for GFP expression and relative expression of *RINF* and *SMAD7* mRNA levels were measured by quantitative RT-PCR. **(B)** Schematic representation of *SMAD7* gene structure (deduced from ENCODE database). CpG islands and primer set regions (R0 to R5) used for ChIP-experiments are also indicated as well as the distance to transcription start site (TSS) of *SMAD7*. **(C)** ChIP-qPCR bars indicate the percentage of *SMAD7* promoter co-immunoprecipitated by anti-*RINF* (*left panel*) or anti-H3K4me3 (*right panel*) antibody in K562 cell line. **(D)** Schematic representation of the pBABE retroviral vector used for *SMAD7* overexpression or empty cassette control. **(E)** Histograms represent the *RINF* or *SMAD7* mRNA levels detected by quantitative RT-PCR. Values are

expressed in percentage of the *shControl* condition. A student T test (unpaired) was performed to assess whether relative expressions were statistically significant of the *shControl* condition. **(F)** Relative expression of *RINF* and *SMAD7* was measured by Western-Blot analysis after retroviral transduction of K562 cells with pBABE/*SMAD7* or pBABE/Empty and puromycin selection (at 1 μ g/mL) for five days. HSC70 was used as a loading control. **(G)** K562/*SMAD7* and K562/Empty cells were then transduced with pTRIPDU3 lentiviral vector. Cells were sorted for GFP expression two days after transduction. For hemoglobin production analysis, the benzidine test was used at 24h of hemin. Percentage of benzidine positive cells was established by counting at least 300 cells per sample. Pictures represent one of three independent experiments. For statistical analysis a student test was performed for qPCR experiments, and hemoglobin production analysis (**: $p=0.0026$; ns: $p >0.05$).

Figure 6. *RINF* and *SMAD7* mRNA closely correlate in primary human CD34+ bone marrow cells isolated from healthy donors or MDS patients with del(5q) and *RINF*-silencing leads to global loss of 5-hydroxymethylation. (A) CD34+ cells were isolated from adult bone marrow (n=11). *RINF* and *SMAD7* mRNA levels were measured by q-RT-PCR. In order to assess whether the expression of these two genes correlated in donors' samples, a correlation analysis was performed with the Pearson's (for non-parametric analysis) correlation coefficient method: rho ($r=0.684$), $P(\text{two-tailed}) < 0.05$ **(B)** *RINF* and *SMAD7* gene expression data were extracted from a previously described microarray dataset performed by Pellagatti *et al.* (46). In this study, CD34+ cells were isolated from adult bone marrow (n=17). For mRNA detection, the 233955_x_at and 204790_at probesets were used for *CXXC5* and *SMAD7*, respectively. A correlation analysis was performed with the Pearson's rho (for non-parametric analysis). A strong correlation coefficient was noticed for normal healthy controls (rho = 0.784), and MDS patients with del(5q), (rho=0.603), that where both highly significant $P(\text{two-tailed}) < 0.001$. **(C)** CD34+ cells isolated from cord blood or healthy adult donors were transduced with pTripizi vector expressing a shRNA targeting *RINF* expression (*shRNA/RINF#4*) or a "non target" shRNA (*shRNA/Control*). A couple of days later (during the amplification phase, see Figure1A), cells were GFP-sorted and 5hmC was detected by immunofluorescence (*left panel*) or flow cytometry (*right panel*) by using an anti-5hmC antibodies labeled with a fluorescent dye. The intensity of the labeling was also quantified by cell

imaging analysis by using Fiji. Representative images from three independent experiments are shown (here, CD34+ cells from an adult donor at 4 days of expansion (2 days post-transduction)). Scale bars = 50 μ m.

Figure 7. *RINF* knockdown-induced phenotype is SMAD7-dependent in human primary cells. (A) Histogram represents *RINF* (left panel) and *SMAD7* (right panel) mRNA expression levels of shRNA-transduced CD34+ cells isolated from 3 cord-blood units. In these specific experiments, lentiviral transduction was performed after magnetic sorting of CD34+ cells (i.e. Day 0 of the amplification step) and cells were GFP-sorted between 2 and 5 days of the amplification step. Data shown are means \pm SD for each donor (q-RT-PCR performed in triplicates). To assess whether the mRNA level was statistically different between *shRNA-RINF* and *shRNA-Scramble* samples, a student T test was performed for each donor. Considering the three biological replicates, both *RINF* and *SMAD7* mRNA levels were statistically downregulated by *shRNA-RINF* (Paired Student T test, $p < 0.001$). (B) This histogram represents the level of *SMAD7* mRNA obtained with the two sequences targeting *RINF* (*shRINF#3* or *shRINF#4*) and detected by quantitative RT-PCR. Values are expressed in percentage of the *shControl* condition. *RINF* mRNA levels can be visualized on Supplementary Figure S3B. (C) Since detection of SMAD7 protein level requires a large quantity of cells for western blot analysis and the active endogenous level of this protein can be below the detection threshold, SMAD7 protein levels was estimated by immunofluorescence. Images of stained cells were acquired on a wide-field Nikon Eclipse microscope through a $\times 20$ objective, with an Eclipse TE2000, a cascade CDD camera (Photometrics). Images of SMAD7 had to be acquired with a binning of one, and two images were averaged for each condition. For cell imaging analysis the quantification of the labeling, performed on two images after background subtraction, the mean fluorescent intensity was measured in 100–120 cells for the two conditions by using Fiji. Representative images from three independent experiments are shown. Scale bar = 50 μ m. (D) Schematic representation of the pInducer21/SMAD7 lentiviral vector used for rescue experiments. This vector drives the expression of SMAD7/HA under a Tetracycline Response Element (TRE) promoter that is activated after doxycycline treatment. To perform rescue experiments, cord-blood derived CD34+ cells ($n=3$ donors) were first transduced with the pInducer21/SMAD7 lentiviral vector, sorted for a low GFP expression, and

transduced again with the pTRIPDU3/shRNA vectors targeting RINF or control. After a second step GFP-sorting (this time based on a high-GFP expression to sort cells transduced with both lentiviral vectors), cells were cultured during two more days before adding (or not) doxycycline that induced SMAD7 expression. **(E)** Histograms represent the *RINF* or *SMAD7* mRNA levels detected by quantitative RT-PCR for cells. Values are expressed in percentage of the *shControl* condition. A student T test (unpaired) was performed to assess if the observed relative difference was statistically significant. **(F)** RINF and SMAD7 were detected by Western-Blot analysis after doxycycline induction (here after 3 days of treatment at 0.2 µg/mL). ACTB was used as a loading control. **(G)** Cell maturation was determined by benzidine assay (upper panel histogram). **(H)** Cumulative RBC number at day 17 of EPO is presented in lower panel histogram.

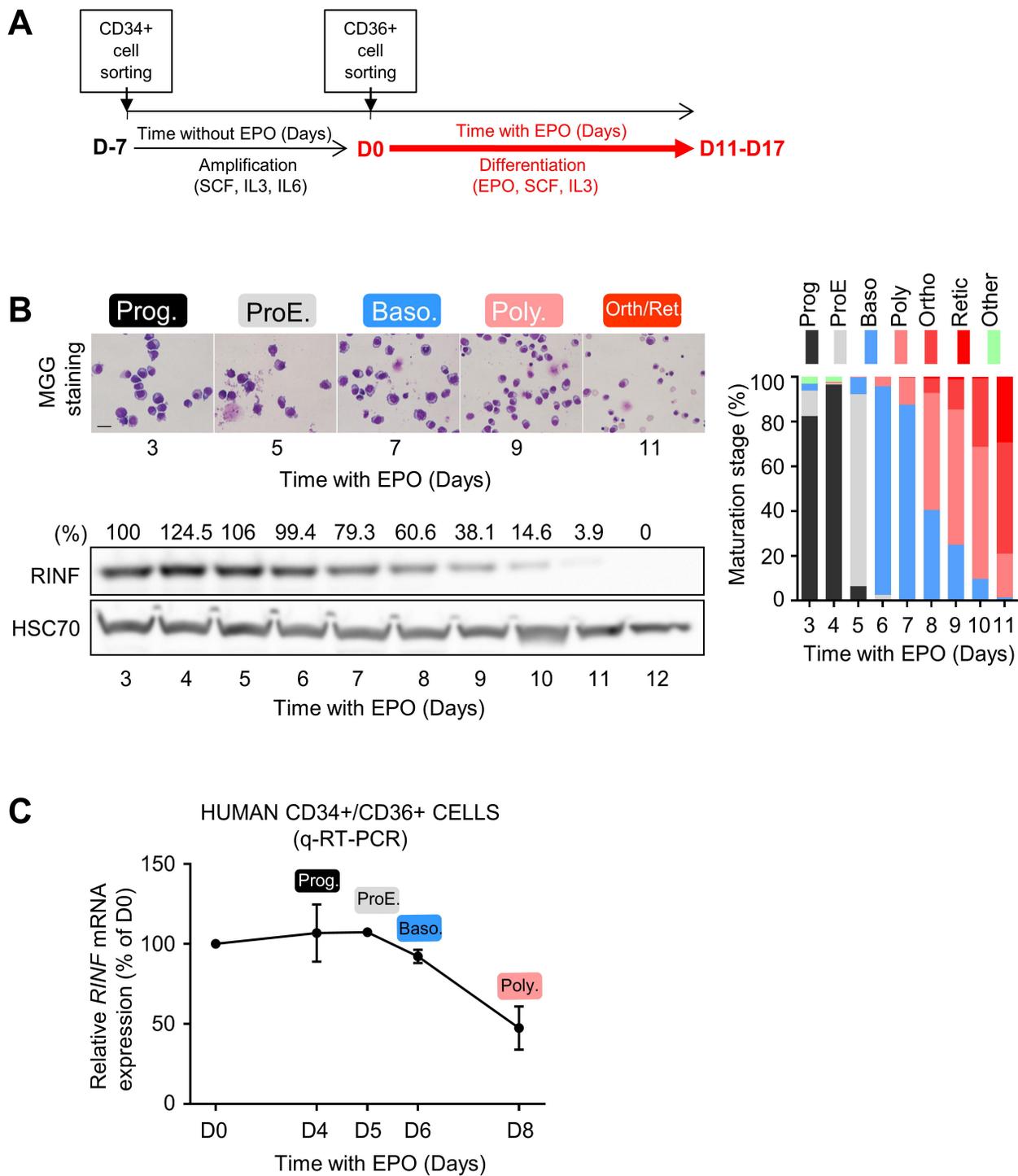


Figure 1

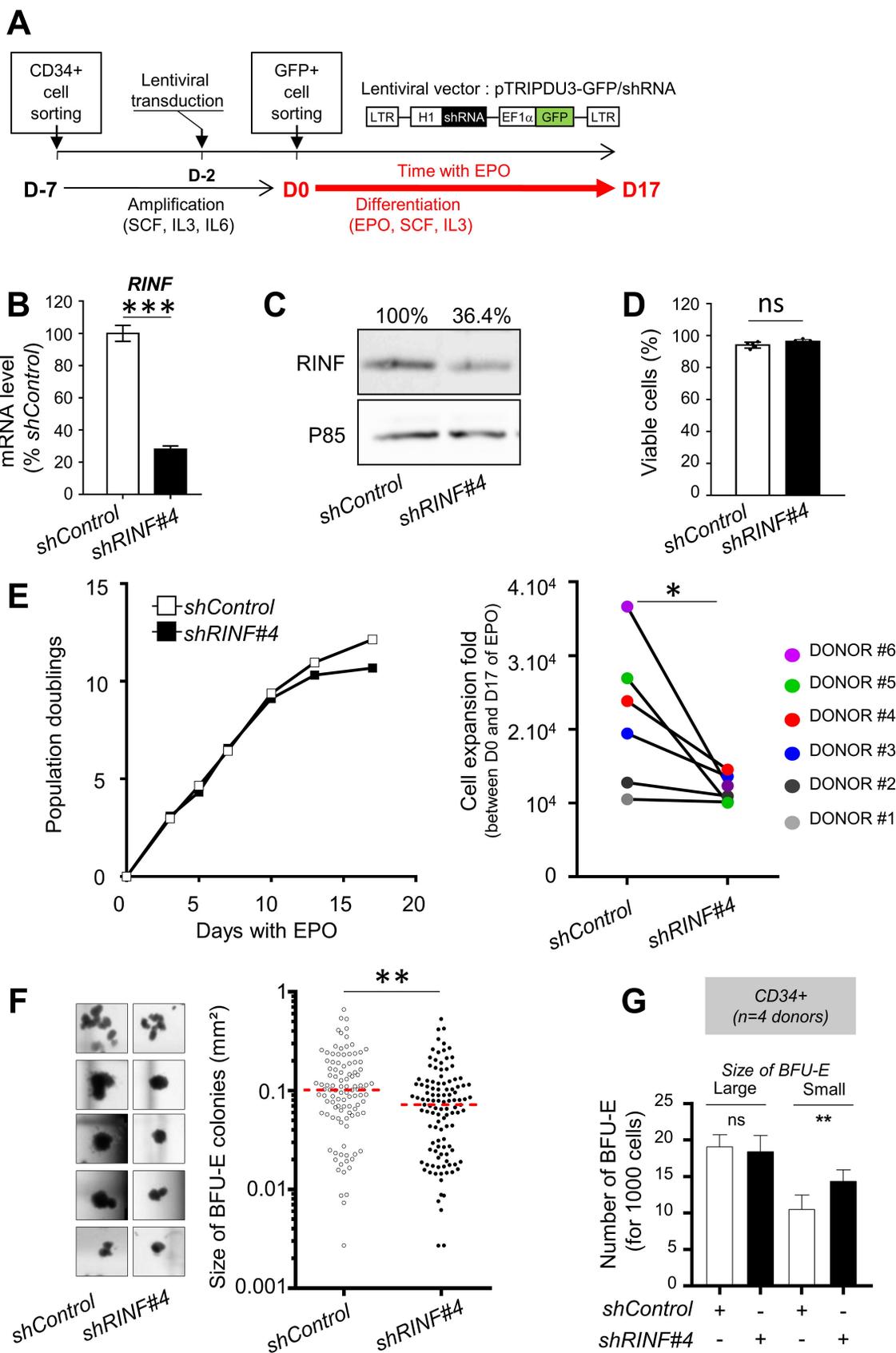


Figure 2

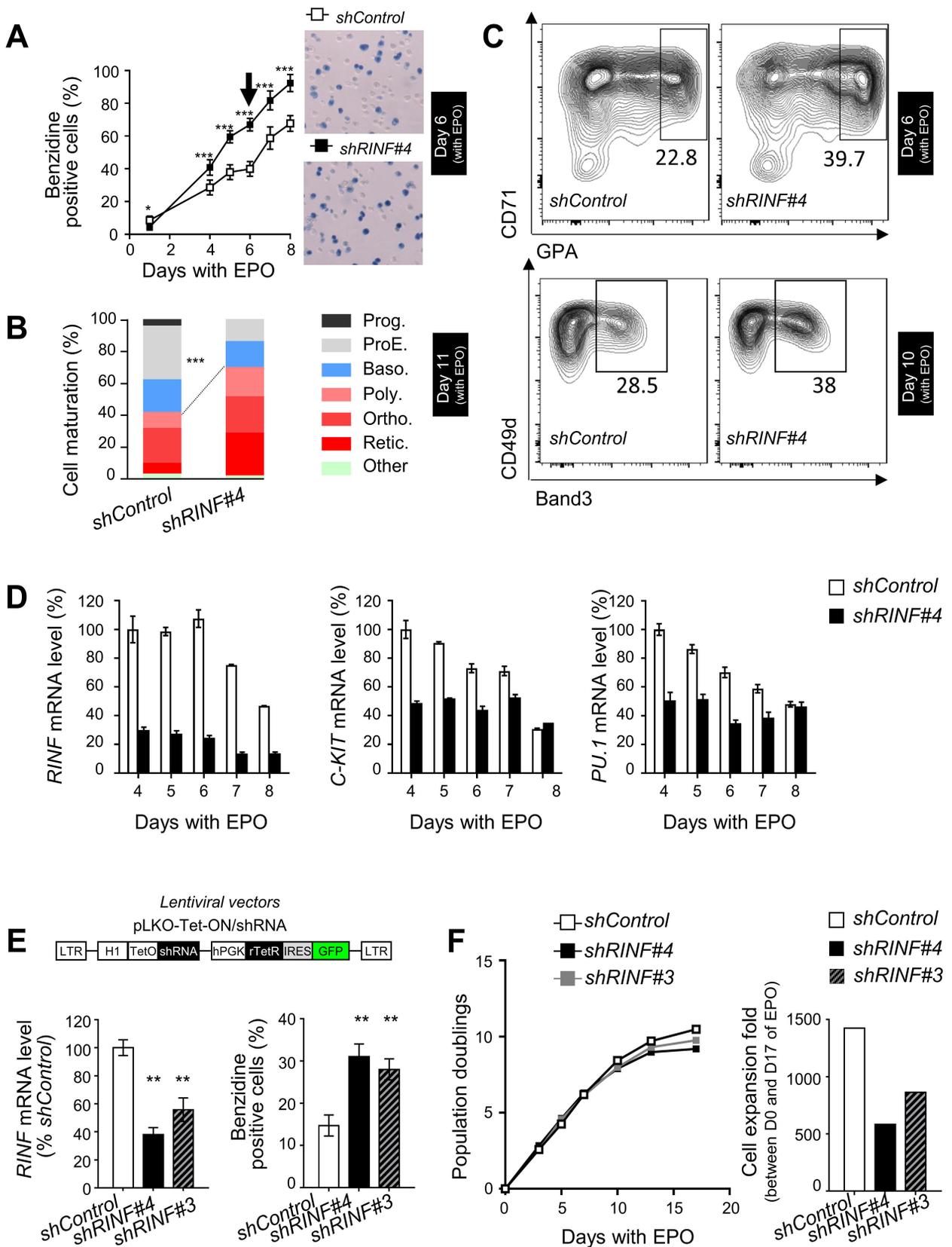


Figure 3

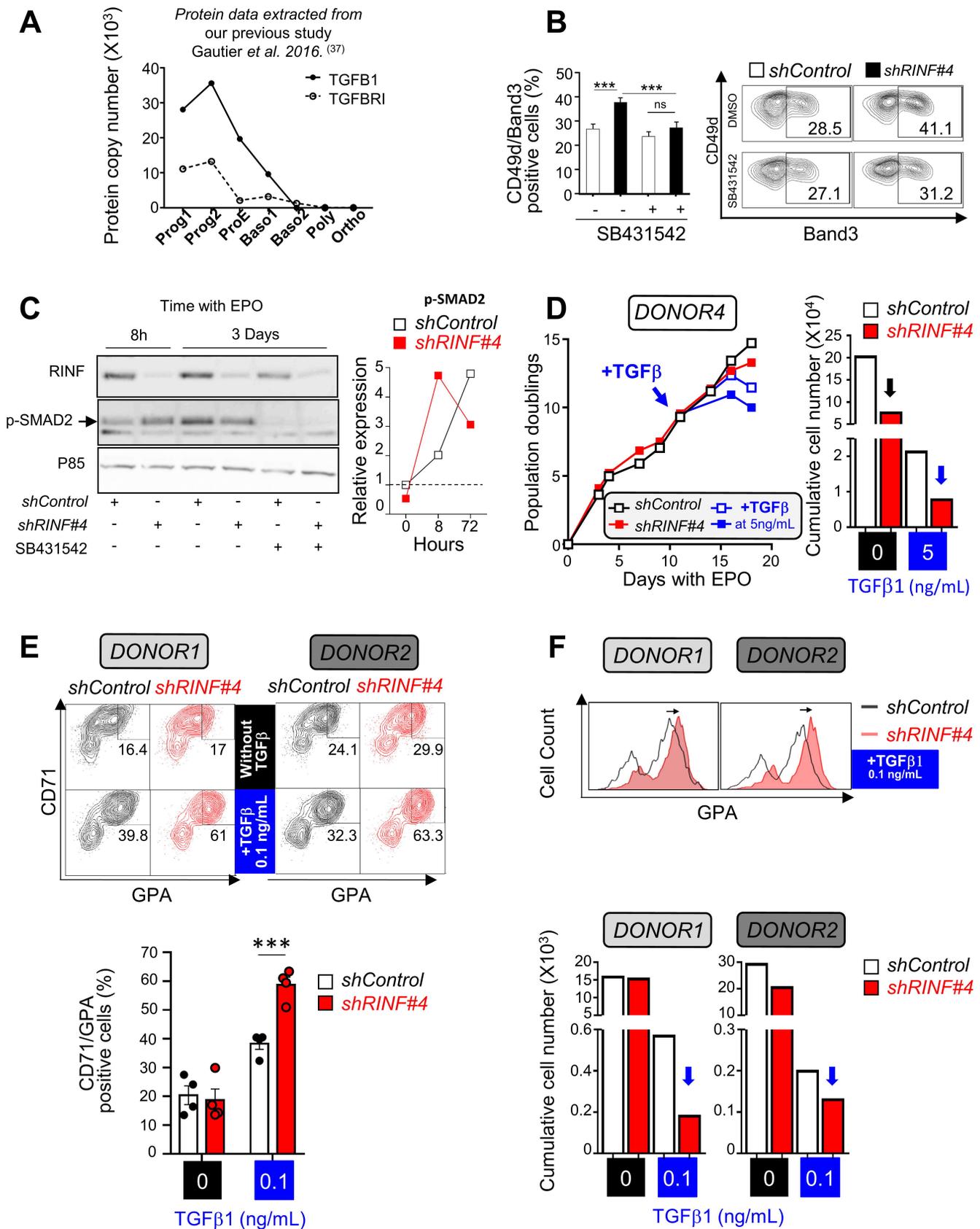


Figure 4

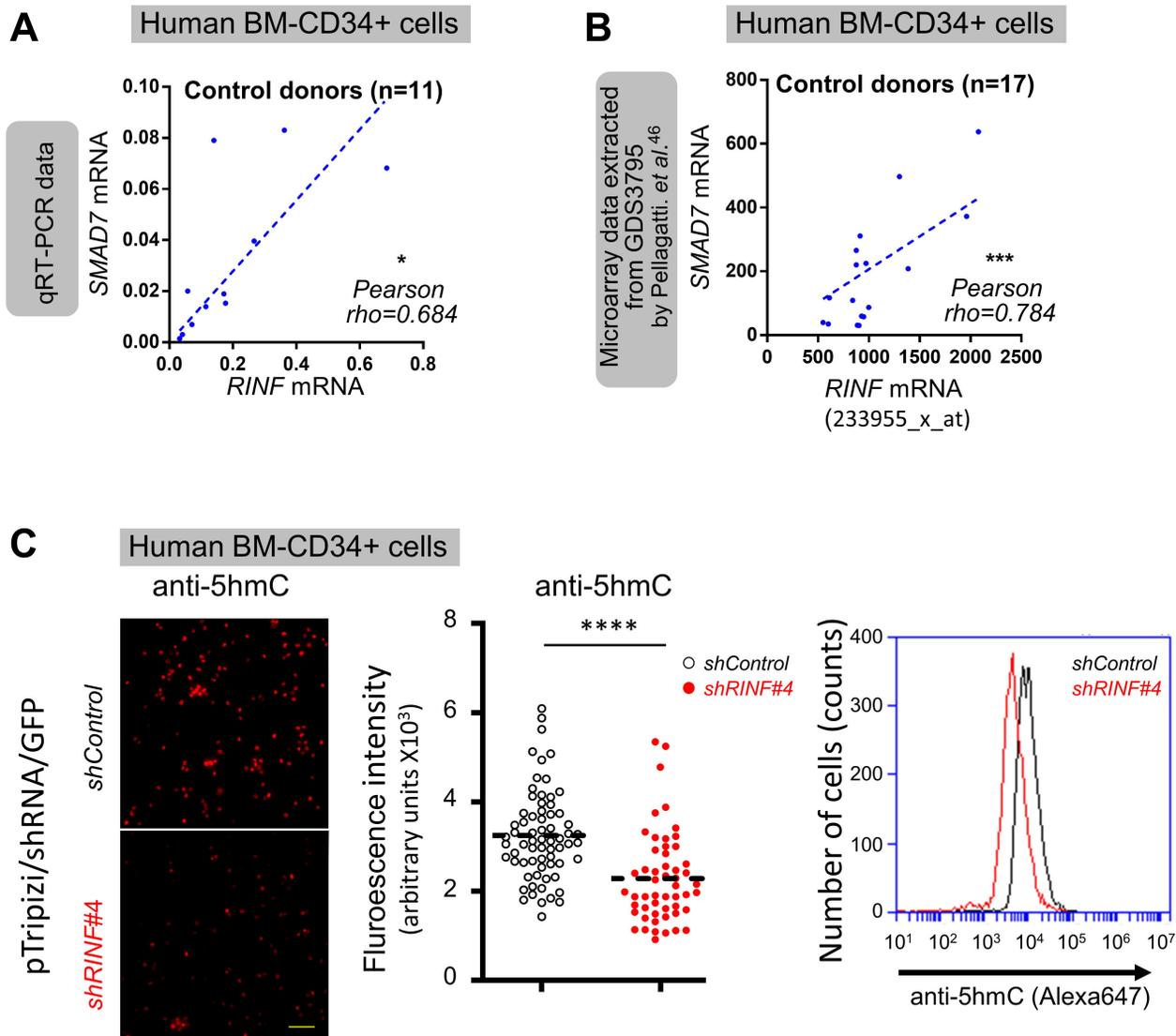


Figure 6

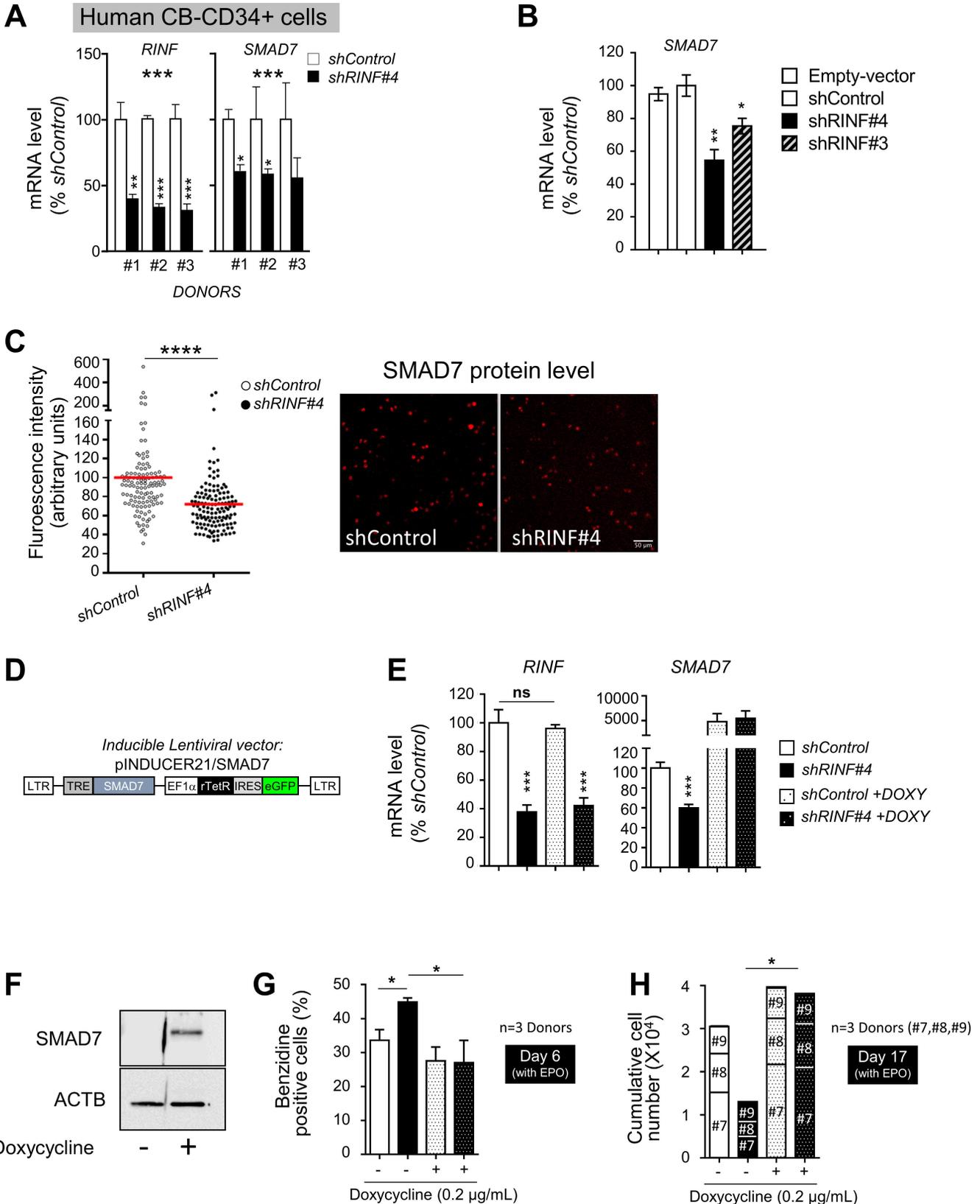


Figure 7

SUPPLEMENTAL METHODS

Flow cytometric cell sorting of MEP cells

Freshly isolated CD34⁺ were labeled with the following labeled antibodies directed against the following markers: CD34 APC-Cy7 (clone 581), CD123 APC (6H6) from BioLegend; CD38 PE (clone HIT2) and CD45RA-FITC (HI100) from BD Biosciences, CD10 PE-Cy5 (eBioscience; clone CD-CALLA). Cells were sorted using BD FACSAriaIII (BD Biosciences) according to the gating strategy described by Manz *et al.*(1) Purification was obtained to at least 98% and only erythroid colonies were generated in CFC assay (i.e. no G-, M-, or GM-derived colony were noticed).

Chromatin Immunoprecipitation (ChIP) experiments

ChIP was performed with anti-RINF (ProteinTech, Cat.N°16513-1-AP) or anti-H3K4me3 (Diagenode, Cat.N°C15410003), antibodies using iDeal ChIP-seq Kit for Transcription Factors (Diagenode) following manufacturer's instructions and using the following options: K562 cells were crosslinked with 1% formaldehyde for 10 minutes in fresh media at room temperature and quenched with 0.2 M glycine for 5 additional minutes. Chromatin was sonicated using Diagenode Bioruptor (Power High, 20 cycles of 30''on-30''off and by changing ice cold water in the bath every 5 cycles). Sonication was validated by agarose gel electrophoresis (average DNA size ~150-250 bp) using an aliquot of purified sheared chromatin (proteinase K treated followed by and phenol-chloroform extraction). Quantitative-PCR was performed using Power SYBR Green PCR Master Mix (Thermo Fisher Scientific, Cat.N°4367659) using pairs of primers encompassing six regions (R0-R5) of *SMAD7* gene. R0 and R5 regions were used as negative controls. The efficiency of chromatin immunoprecipitation at these loci was determined as the recovery of that locus calculated as the percentage of the input (the relative amount of immunoprecipitated DNA compared to input DNA). The amount used for the input was 1% of the amount used for ChIP, the recovery was calculated as follows: % recovery = $2^{(Ct^{input} - Ct^{sample})}$.

Primer sequences used for quantitative RT-PCR and ChIP-qPCR

Human *PPIA* (Cyclophilin A) endogenous control was obtained from Thermo Fisher Scientific (Cat.N°4333763T). The other primer sequences (5'-3' orientation) used for RT-qPCR and ChIP-qPCR are indicated in the following table.

Q-RT-PCR (Hydrolysis)	Forward	Reverse	Probe sequence (5'-3')
<i>RINF</i>	TCCGCTGCTCTGGAGAAG	CACACGAGCAGTGACATTGC	FAM-aacccaaagctgccctctc-c-BBQ
<i>RPLP2</i>	GACCGGCTCAACAAGGTTAT	CCCCACCAGCAGGTACAC	Cy5-agctgaatggaaaaaaca-tgaagacgtc-BBQ
Q-RT-PCR (SybrGreen)	Forward	Reverse	
<i>cKIT</i>	TATGCTCTCGCACCTTTCCA	TCTCAATGAAGTGCCCTGA	
<i>PU1</i>	CACTGGAGGTGTCTGACGG	GCGGATCTTCTTCTTGCTGC	
<i>SMAD7</i>	CAATGACCACGAGTTTATGCA	GTTGAAGATGACCTCTAGCCA	
<i>Murine Rinf</i>	TCCGCTGCTCTGGAGAAG	CACGCGGGCAGTGACATTGC	
<i>murineSmad7</i>	TGCTGTGCAAAAGTGTTCAGGTG	CCATCGGGTATCTGGAGTAAGGA	
<i>murineGapdh</i>	TGTGTCCGTCGTCGATCTGA	TTGCTGTTG AAGTCGCAGGAG	
ChIP-qPCR (SybrGreen)	Forward	Reverse	Distance to SMAD7 TSS
<i>SMAD7 R0 (neg.)</i>	ATGGGTGTTTCAAGAGTGGAG	GAAGAGGAAAGCCGTGTAGAG	-4.7kb
<i>SMAD7 R1</i>	CTAGCGCTTCATTCATTGGTTT	TCGGTCCAGTCCGGTATAA	-1962-1840
<i>SMAD7 R2</i>	GCAAACAACAGATCGGGTTTC	CAATCCATTCTGGGAGCTTCT	-1695 -1556
<i>SMAD7 R3</i>	GCCTCGGCTTCTACATGGA	CTCCCCACCCCAAATTAAG	-243 -122
<i>SMAD7 R4</i>	CACAGCCTTGTGACTTCCGAG	TGCTCCTACTCGCTCCCCT	+737 +819
<i>SMAD7-R5 (neg.)</i>	CATCCCATCTATTCTCACGCC	AGACAGCCCCCTTAAAGTTAGC	+23.7kb

Target sequences of the shRNAs used to target *RINF* knockdown

shRNA name	Target mRNA Sequence
<i>shControl</i>	CCUGGGCAAAGAAUGGACAAU
<i>shRINF#4</i>	GAAUGGACAAUCAGUUUCCUU
<i>shRINF#3</i>	CCUUUGAUUCUUUCCGACCAU

Detection of SMAD7 protein or 5hmC by immunofluorescence.

CD34+ cells were washed once (5 minutes in ice-cold PBS 1X), and fixed in PFA 2% (8 minutes at room temperature), permeabilized with Triton X100 0.1% (5 minutes at 4°C), incubated 1 hour in blocking buffer (PBS 1X, 5% Fetal Bovine Serum, 2% BSA) before overnight staining with primary mouse monoclonal anti-SMAD7 antibody (1:200, Santa-Cruz, #sc-365846). For 5hmC detection, two additional steps were performed after the permeabilization step, one DNA denaturation step in 3M HCl (5 minutes at

RT) followed by a five-minute neutralization step in TE buffer (10 mM Tris, 1 mM EDTA) at pH8. The primary anti-5hmC polyclonal rabbit antibody (1:2000) was purchased from Active motif (Cat.N°39769). Two secondary antibodies (1:1.000) coupled with AlexaFluor 647 fluorescent dye were purchased from Lifetechnologies, a goat anti-rabbit antibody (Cat # A32733) for 5hmC detection and a goat anti-mouse antibody (Cat # A32723) for SMAD7 detection. Images of 5hmC and SMAD7 staining were acquired on a wide-field Nikon Eclipse TE2000 microscope through a ×20 objective, with a Cascade CDD camera (Photometrics). The quantification of intensity was performed after background subtraction (Fiji). The fluorescent intensity was measured in 100–120 cells for the two conditions.

Cell culture with AML cell lines

Human leukemic cell lines were purchased from DSMZ (MV4-11; Braunschweig, Germany) and from the American Type Culture Collection (K562; Molsheim, France). UT7 5.3 cells were kindly provided by Isabelle Dusanter-Fourt (Cochin Institute, Paris, France). K562 and MV4-11 were cultured in RPMI 1640 medium supplemented with 10% fetal calf serum (FCS), 2 mM L-Glutamine, 50 U/ml penicillin G and 50 µg/ml streptomycin (Life Technologies, Saint-Aubin, France). UT7 5.3 cells were cultured in minimum essential medium (MEM) α medium containing 10% of FCS, 2 mM L-Glutamine, 50 U/ml penicillin G and 50 µg/ml streptomycin (Life Technologies, Saint-Aubin, France) and 2,5 ng/µl of GM-CSF (Miltenyi Biotec, France). Cells were grown in mycoplasma-free conditions, regularly tested.

Microarray and data analysis

Analysis of global gene expression profiles. After validation of RNA quality with Bioanalyzer 2100 (using Agilent RNA6000 nano chip kit), 400 ng of total RNA was reverse transcribed following the Genechip WT plus Reagent kit (Affymetrix). Briefly, the resulting double strand cDNA was used for in vitro transcription with T7 RNA pol. After purification, 15 µg of cRNA was used for reverse transcription with random primers. The cDNA obtained was then purified and fragmented. After control of fragmentation using Bioanalyzer 2100, cDNA was end labeled with biotin using Terminal Transferase (using the WT terminal labeling kit of Affymetrix). cDNA was then hybridized to GeneChip® Human Transcriptome Analysis 2.0 (Affymetrix) at 45°C for

17 hours. After overnight hybridization, chips were washed on the fluidic station FS450 following specific protocols (Affymetrix) and scanned using the GCS3000 7G. The image was then analyzed with Expression Console software (Affymetrix) to obtain raw data (cel files) and metrics for Quality Controls. Minimum information about a microarray experiment (MIAME)–compliant documentation of the microarray experiments have been deposited in Gene Expression Omnibus under the accession number GSE140770.

The Affymetrix HTA2 dataset analysis was performed by GenoSplice technology (www.genosplice.com). Data were normalized using quantile normalization. Background corrections were made with antigenomic probes and probes were selected according to their %GC, cross-hybridization status and potential overlap with repeat region as previously described (2,3). Only probes targeting exons and exon-exon junctions annotated from FAST DB® transcripts (release fastdb_2014_1) were selected (4,5). Only genes expressed in at least one compared condition were analyzed. To be considered to be expressed, the DABG P-value had to be ≤ 0.05 for at least half of the gene probes. We performed an unpaired Student's t-test to compare gene intensities between *shRINF* and *shControl* cells. Genes were considered significantly regulated when fold-change was ≥ 1.5 , a fold-change noticed for *CXXC5* genes for which we have validated the knock-down by q-RT-PCR and western-blot analysis.

Supplementary Table S1.

ID	Symbol	Entrez Gene Name	Location	Type(s)
AATK	AATK	apoptosis associated tyrosine kinase	Cytoplasm	kinase
EMR3	ADGRE3	adhesion G protein-coupled receptor E3	Plasma Membrane	G-protein cou
ADM5	ADM5	adrenomedullin 5 (putative)	Other	other
ADRA2C	ADRA2C	adrenoceptor alpha 2C	Plasma Membrane	G-protein cou
AGAP4	AGAP6 (includes ArfGAP with GTPase domain, ankyrin repeat and PH domain 5		Other	other
AKAP7	AKAP7	A-kinase anchoring protein 7	Plasma Membrane	other
ANXA2R	ANXA2R	annexin A2 receptor	Plasma Membrane	other
B3GNT4	B3GNT4	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 4	Plasma Membrane	enzyme
BEND5	BEND5	BEN domain containing 5	Cytoplasm	other
C15orf59	C15orf59	chromosome 15 open reading frame 59	Other	other
C1QB	C1QB	complement C1q B chain	Extracellular Space	peptidase
C2orf66	C2orf66	chromosome 2 open reading frame 66	Other	other
C3orf80	C3orf80	chromosome 3 open reading frame 80	Other	other
C8orf74	C8orf74	chromosome 8 open reading frame 74	Other	other
C9orf131	C9orf131	chromosome 9 open reading frame 131	Other	other
CACNG5	CACNG5	calcium voltage-gated channel auxiliary subunit gamma 5	Plasma Membrane	ion channel
CALML5	CALML5	calmodulin like 5	Cytoplasm	other
CARS	CARS	cysteinyl-tRNA synthetase	Cytoplasm	enzyme
CCDC144NL	CCDC144NL	coiled-coil domain containing 144 family, N-terminal like	Other	other
CDH4	CDH4	cadherin 4	Plasma Membrane	other
CGB8	CGB3 (includes chorionic gonadotropin beta subunit 3		Extracellular Space	other
CHADL	CHADL	chondroadherin like	Extracellular Space	other
CNTN4	CNTN4	contactin 4	Plasma Membrane	enzyme
COL18A1	COL18A1	collagen type XVIII alpha 1 chain	Extracellular Space	other
COL4A2	COL4A2	collagen type IV alpha 2	Extracellular Space	other
COL6A1	COL6A1	collagen type VI alpha 1	Extracellular Space	other
COX6B2	COX6B2	cytochrome c oxidase subunit 6B2	Cytoplasm	enzyme
CPOX	CPOX	coproporphyrinogen oxidase	Cytoplasm	enzyme
CRHR1	CRHR1	corticotropin releasing hormone receptor 1	Plasma Membrane	G-protein cou
CRP	CRP	C-reactive protein, pentraxin-related	Extracellular Space	other
CSDC2	CSDC2	cold shock domain containing C2	Cytoplasm	other
CXXC5	CXXC5	CXXC finger protein 5	Nucleus	other
CYP4F30P	CYP4F30P	cytochrome P450 family 4 subfamily F member 30, pseudogene	Other	other
DAZ2	DAZ2	deleted in azoospermia 2	Cytoplasm	translation reg
DLGAP3	DLGAP3	DLG associated protein 3	Cytoplasm	other
DNAJC5G	DNAJC5G	DnaJ heat shock protein family (Hsp40) member C5 gamma	Other	other
DND1	DND1	DND microRNA-mediated repression inhibitor 1	Cytoplasm	other
DRGX	DRGX	dorsal root ganglia homeobox	Nucleus	other
DUSP4	DUSP4	dual specificity phosphatase 4	Nucleus	phosphatase
EPN2-IT1	EPN2-IT1	EPN2 intronic transcript 1	Other	other
FAM153C	FAM153C	family with sequence similarity 153, member C	Other	other
FAM229A	FAM229A	family with sequence similarity 229 member A	Other	other
FAM26D	FAM26D	family with sequence similarity 26 member D	Other	other
FGD5P1	FGD5P1	FYVE, RhoGEF and PH domain containing 5 pseudogene 1	Other	other
GATA3	GATA3	GATA binding protein 3	Nucleus	transcription r
GJC1	GJC1	gap junction protein gamma 1	Plasma Membrane	ion channel
GPIHBP1	GPIHBP1	glycosylphosphatidylinositol anchored high density lipoprotein binding protein 1	Plasma Membrane	transporter
GPR139	GPR139	G protein-coupled receptor 139	Plasma Membrane	G-protein cou
GPR55	GPR55	G protein-coupled receptor 55	Plasma Membrane	G-protein cou
GPRIN2	GPRIN2	G protein regulated inducer of neurite outgrowth 2	Other	other
GRASPOS	GRASPOS	GRP1-associated scaffold protein opposite strand	Other	other
GSX2	GSX2	GS homeobox 2	Nucleus	transcription r
HAR1B	HAR1B	highly accelerated region 1B (non-protein coding)	Other	other
HCAR2	HCAR2	hydroxycarboxylic acid receptor 2	Plasma Membrane	G-protein cou
HCG27	HCG27	HLA complex group 27 (non-protein coding)	Other	other
HHIPL1	HHIPL1	HHIP like 1	Other	other
HNRNPKP3	HNRNPKP3	heterogeneous nuclear ribonucleoprotein K pseudogene 3	Other	other
HOXB5	HOXB5	homeobox B5	Nucleus	transcription r
HOXC9	HOXC9	homeobox C9	Nucleus	transcription r
HSPA6	HSPA6	heat shock protein family A (Hsp70) member 6	Nucleus	enzyme
IGFL1	IGFL1	IGF like family member 1	Extracellular Space	other
IL23A	IL23A	interleukin 23 subunit alpha	Extracellular Space	cytokine
IL36A	IL36A	interleukin 36, alpha	Extracellular Space	cytokine
IRF2BP1	IRF2BP1	interferon regulatory factor 2 binding protein 1	Nucleus	transcription r
IRS4	IRS4	insulin receptor substrate 4	Plasma Membrane	other
KCNA1	KCNA1	potassium voltage-gated channel subfamily A member 1	Plasma Membrane	ion channel
KCNC1	KCNC1	potassium voltage-gated channel subfamily C member 1	Plasma Membrane	ion channel
KCNK17	KCNK17	potassium two pore domain channel subfamily K member 17	Plasma Membrane	ion channel
KRTAP2-1	KRTAP2-1	keratin associated protein 2-1	Other	other
KRTAP6-3	KRTAP6-3	keratin associated protein 6-3	Other	other
KRTAP9-8	KRTAP9-8	keratin associated protein 9-8	Other	other
LCE2A	LCE2A	late cornified envelope 2A	Other	other

LCE2C	LCE2C/LCE2D	late cornified envelope 2D	Other	other
LEMD1	LEMD1	LEM domain containing 1	Other	other
LIN28A	LIN28A	lin-28 homolog A	Cytoplasm	other
LINC00641	LINC00641	long intergenic non-protein coding RNA 641	Other	other
LINC00927	LINC00927	long intergenic non-protein coding RNA 927	Other	other
LINC00968	LINC00968	long intergenic non-protein coding RNA 968	Other	other
LINC00987	LINC00987	long intergenic non-protein coding RNA 987	Other	other
LINC00996	LINC00996	long intergenic non-protein coding RNA 996	Other	other
LINC01010	LINC01010	long intergenic non-protein coding RNA 1010	Other	other
LMOD3	LMOD3	leiomodrin 3	Other	other
LRRC32	LRRC32	leucine rich repeat containing 32	Plasma Membrane	other
LRRC4B	LRRC4B	leucine rich repeat containing 4B	Plasma Membrane	other
LRRC52	LRRC52	leucine rich repeat containing 52	Other	other
MACROD2	MACROD2	MACRO domain containing 2	Nucleus	enzyme
MAFG-AS1	MAFG-AS1	MAFG antisense RNA 1 (head to head)	Other	other
MAP4K1	MAP4K1	mitogen-activated protein kinase kinase kinase kinase 1	Cytoplasm	kinase
MIR103A1	mir-103	microRNA 107	Cytoplasm	microRNA
MIR1257	mir-1257	microRNA 1257	Cytoplasm	microRNA
MIR1288	mir-1288	microRNA 1288	Cytoplasm	microRNA
MIR129-2	mir-129	microRNA 129-2	Cytoplasm	microRNA
MIR1301	mir-1301	microRNA 1301	Cytoplasm	microRNA
MIR152	mir-148	microRNA 148a	Cytoplasm	microRNA
MIR149	mir-149	microRNA 149	Cytoplasm	microRNA
MIR377	mir-154	microRNA 494	Other	microRNA
MIR1972-1	mir-1972	microRNA 1972-1	Cytoplasm	microRNA
MIR211	mir-204	microRNA 204	Cytoplasm	microRNA
MIR29C	mir-29	microRNA 29a	Cytoplasm	microRNA
MIR371A	mir-290	microRNA 372	Other	microRNA
MIR3065	mir-3065	microRNA 3065	Cytoplasm	microRNA
MIR320B2	mir-320	microRNA 320a	Cytoplasm	microRNA
MIR3689C	mir-3689	microRNA 3689b	Cytoplasm	microRNA
MIR4654	mir-4654	microRNA 4654	Cytoplasm	microRNA
MIR4803	mir-4803	microRNA 4803	Cytoplasm	microRNA
MIR489	mir-489	microRNA 489	Cytoplasm	microRNA
MIR550A1	mir-550	microRNA 550a-1	Cytoplasm	microRNA
MIR569	mir-569	microRNA 569	Cytoplasm	microRNA
MIR581	mir-581	microRNA 581	Cytoplasm	microRNA
MIR615	mir-615	microRNA 615	Cytoplasm	microRNA
MIR647	mir-647	microRNA 647	Cytoplasm	microRNA
MIR200A	mir-8	microRNA 200c	Cytoplasm	microRNA
MIR4291	MIR4291	microRNA 4291	Cytoplasm	microRNA
MIR4440	MIR4440	microRNA 4440	Cytoplasm	microRNA
MIR4692	MIR4692	microRNA 4692	Cytoplasm	microRNA
MRGPRG	MRGPRG	MAS related GPR family member G	Plasma Membrane	G-protein cou
MUC6	MUC6	mucin 6, oligomeric mucus/gel-forming	Extracellular Space	other
MYADML	MYADML	myeloid associated differentiation marker like (pseudogene)	Other	other
NEURL2	NEURL2	neuralized E3 ubiquitin protein ligase 2	Cytoplasm	other
NGB	NGB	neuroglobin	Cytoplasm	transporter
NHSL2	NHSL2	NHS like 2	Other	other
NKX6-3	NKX6-3	NK6 homeobox 3	Other	other
NOXA1	NOXA1	NADPH oxidase activator 1	Other	other
NPIP7	NPIP7	nuclear pore complex interacting protein family member B7	Other	other
NRGN	NRGN	neurogranin	Other	other
NRTN	NRTN	neurturin	Extracellular Space	growth factor
NUDT16L1	NUDT16L1	nudix hydrolase 16 like 1	Cytoplasm	other
OGFR	OGFR	opioid growth factor receptor	Plasma Membrane	other
OLFML3	OLFML3	olfactomedin like 3	Extracellular Space	other
OMD	OMD	osteomodulin	Extracellular Space	other
OR2A25	OR2A25	olfactory receptor family 2 subfamily A member 25	Plasma Membrane	G-protein cou
OR2Y1	OR2Y1	olfactory receptor family 2 subfamily Y member 1	Plasma Membrane	G-protein cou
OR5C1	OR5C1	olfactory receptor family 5 subfamily C member 1	Plasma Membrane	G-protein cou
OR5R1	OR5R1	olfactory receptor family 5 subfamily R member 1 (gene/pseudogene)	Plasma Membrane	G-protein cou
OXCT2	OXCT2	3-oxoacid CoA-transferase 2	Cytoplasm	other
P2RY13	P2RY13	purinergic receptor P2Y13	Plasma Membrane	G-protein cou
PARVG	PARVG	parvin gamma	Cytoplasm	other
PAX7	PAX7	paired box 7	Nucleus	transcription r
PCDHA1	PCDHA1	protocadherin alpha 1	Plasma Membrane	other
PF4V1	PF4V1	platelet factor 4 variant 1	Extracellular Space	cytokine
PHF13	PHF13	PHD finger protein 13	Nucleus	other
POTEKP	POTEKP	POTE ankyrin domain family member K, pseudogene	Cytoplasm	other
POU5F1P3	POU5F1P3	POU class 5 homeobox 1 pseudogene 3	Other	other
PRR21	PRR21	proline rich 21	Other	other
PRRT3-AS1	PRRT3-AS1	PRRT3 antisense RNA 1	Other	other
PRSS41	PRSS41	protease, serine 41	Plasma Membrane	peptidase
PTPRF	PTPRF	protein tyrosine phosphatase, receptor type F	Plasma Membrane	phosphatase
PWRN1	PWRN1	Prader-Willi region non-protein coding RNA 1	Other	other
RAMP2-AS1	RAMP2-AS1	RAMP2 antisense RNA 1	Other	other
RARRES2	RARRES2	retinoic acid receptor responder 2	Plasma Membrane	transmembran

RBM1E	RBM1A1 (includ	RNA binding motif protein, Y-linked, family 1, member A1	Nucleus	other
C15orf60	REC114	REC114 meiotic recombination protein	Other	other
RFPL3	RFPL1/RFPL3	ret finger protein like 3	Other	other
RHBDL2	RHBDL2	rhomboid like 2	Plasma Membrane	peptidase
RPH3AL	RPH3AL	rabphilin 3A-like (without C2 domains)	Plasma Membrane	other
S1PR1	S1PR1	sphingosine-1-phosphate receptor 1	Plasma Membrane	G-protein cou
SCARA3	SCARA3	scavenger receptor class A member 3	Plasma Membrane	transmembran
SCGB3A1	SCGB3A1	secretoglobin family 3A member 1	Extracellular Space	cytokine
SDC1	SDC1	syndecan 1	Plasma Membrane	enzyme
SIAH3	SIAH3	siah E3 ubiquitin protein ligase family member 3	Cytoplasm	other
SIGLEC16	SIGLEC16	sialic acid binding Ig like lectin 16 (gene/pseudogene)	Plasma Membrane	other
SIX3	SIX3	SIX homeobox 3	Nucleus	transcription r
SIX6	SIX6	SIX homeobox 6	Nucleus	transcription r
SLC17A6	SLC17A6	solute carrier family 17 member 6	Plasma Membrane	transporter
SLC52A1	SLC52A1	solute carrier family 52 member 1	Plasma Membrane	other
SMAD7	SMAD7	SMAD family member 7	Nucleus	transcription
SMIM18	SMIM18	small integral membrane protein 18	Other	other
SMIM22	SMIM22	small integral membrane protein 22	Other	other
SMPD3	SMPD3	sphingomyelin phosphodiesterase 3	Cytoplasm	enzyme
SMTNL1	SMTNL1	smoothelin like 1	Cytoplasm	other
SNX21	SNX21	sorting nexin family member 21	Cytoplasm	transporter
SOWAHD	SOWAHD	sosondowah ankyrin repeat domain family member D	Other	other
SOX2	SOX2	SRY-box 2	Nucleus	transcription r
SPATA21	SPATA21	spermatogenesis associated 21	Other	other
SPRR2D	SPRR2D	small proline rich protein 2D	Cytoplasm	other
SSTR1	SSTR1	somatostatin receptor 1	Plasma Membrane	G-protein cou
SYNGR4	SYNGR4	synaptogyrin 4	Plasma Membrane	other
TAAR5	TAAR5	trace amine associated receptor 5	Plasma Membrane	G-protein cou
TEPP	TEPP	testis, prostate and placenta expressed	Other	other
TFAP2E	TFAP2E	transcription factor AP-2 epsilon	Nucleus	transcription r
TIE1	TIE1	tyrosine kinase with immunoglobulin like and EGF like domains 1	Plasma Membrane	kinase
TMED6	TMED6	transmembrane p24 trafficking protein 6	Cytoplasm	other
TMPRSS4-AS	TMPRSS4-AS1	TMPRSS4 antisense RNA 1	Other	other
TOLLIP-AS1	TOLLIP-AS1	TOLLIP antisense RNA 1 (head to head)	Other	other
TPSD1	TPSD1	tryptase delta 1	Extracellular Space	peptidase
TTC30A	TTC30A	tetratricopeptide repeat domain 30A	Other	other
TUSC1	TUSC1	tumor suppressor candidate 1	Other	other
USP3-AS1	USP3-AS1	USP3 antisense RNA 1	Other	other
WFDC2	WFDC2	WAP four-disulfide core domain 2	Extracellular Space	other
ZDHHC15	ZDHHC15	zinc finger DHHC-type containing 15	Cytoplasm	enzyme
ZMIZ1	ZMIZ1	zinc finger MIZ-type containing 1	Nucleus	other
ZNF423	ZNF423	zinc finger protein 423	Nucleus	transcription r
ZNF836	ZNF836	zinc finger protein 836	Other	other

Supplementary Table S1. List of genes commonly down-regulated in K562 and UT7_{5.3} knockdown cells (n=193). For the microarray experiments, we used the GeneChip™ Human Transcriptome Array 2.0 (HTA2).

LEGENDS TO SUPPLEMENTAL FIGURES

Figure S1. The *CXXC5/RINF* gene is located at chromosome 5q31.2 within the CDR associated with higher risk in MDS and AML. Genes, as well as sequence tag site markers, are listed in the order of their physical positions (megabase, Mb) along chromosome 5q from centromere (upper) to telomere (bottom), according to the most recent human genome assembly at <http://ensembl.org> (Human, GRCh38.p12). *CXXC5* gene is located at less than 20kb from *UBE2D2*, the more distal gene of the high-risk CDR (grey) investigated by Liu *et al.* in 2007 (6). According to a SNP-array study published in 2012 by Jerez *et al.* (7), and according to the GRCH37/hg19 genome assembly version (released in 2009) that was probably used for this study, the *CXXC5* gene is located inside the refined-CDR associated with high-risk.

Figure S2. The pattern of *RINF* mRNA expression profile during erythroid maturation of human hematopoietic progenitors cultured *in vitro* was confirmed in three independent datasets. (A) *CXXC5* mRNA expression profile was obtained by querying the Human Erythroblast Maturation Database described by Merryweather-Clarke *et al.*(8) In that study, PBMC-derived erythroblasts from 38 adult donors were grown *in vitro* and cells at the CFU-E, Pro-E, Int-E and Late-E stages were sorted by FACS according to CD36, CD71 and glycophorin A expression. (B) *CXXC5* mRNA expression data were extracted from an RNA-seq dataset performed by An *et al.*(9) CD34+ cells from cord blood were cultured in a 3-phases culture system to produce primary human erythroid cells at different stages of terminal differentiation. Erythroblastic populations were isolated through sorting by FACS according to expression of cell surface markers glycophorin A, band 3, and $\alpha 4$ -integrin. (C) *CXXC5* mRNA expression was extracted from a public microarray dataset performed by Keller *et al.*(10) Briefly, CD34+ cells isolated from adult PBMC were cultures in a serum-free medium supplemented with IL3, SCF, and EPO for the indicated number of days.

Figure S3. In human CD34+ cells, *RINF* silencing did not affect the total number of colonies. Human CD34+ cells were isolated from bone marrow of healthy donors and transduced with a lentiviral vector that drives shRNA expression (pTRIPDU3). Cells were then seeded (500 cells per well) into methylcellulose plates containing

cytokines (IL3, SCF, G-CSF, GM-CSF, and EPO) enabling the formation of erythroid colonies (both large and small BFU-Es), myeloid colonies (CFU-G, CFU-M, or CFU-GM), or mixed colonies (CFU-GEMM). Colony Forming Cells were counted at 14 days of culture. **(A)** Schematic representation of the pTRIPDU3 (*lower panel*) lentiviral vector used to express a short hairpin RNA. The specificity and the efficacy of the sequences were previously described (11). The pTRIPDU3 vector drives the expression of the shRNA under a H1 promoter and the GFP expression gene under an EF1 α promoter. After two days of expansion in (IL3, SCF, IL6), human CD34+ cells were transduced with the pTRIPDU3 vector. Of note, pTRIPDU3-transduced cells were FACS-sorted for GFP expression at two days after transduction. **(B)** This histogram represents the level of *RINF* mRNA obtained with the two sequences targeting *RINF* (*shRINF#3* or *shRINF#4*) and detected by quantitative RT-PCR. Values are expressed in percentage of the *shControl* condition. (student T-test, * $p < 0.05$) **(C)** CD34+ cells were isolated from bone marrow of healthy donors (n=4). The histogram represents the total number of colonies enumerated (*left panel*) and the relative proportion of each type of colony (*right panel*). For each of the four donors, the proportions are also visualized by pairs of red dots. (student T-test, * $p < 0.05$) **(D)** CD34+ were isolated from bone marrow of healthy donors (n=4, *left panel*) or cord-blood (n=6, *right panel*). Histograms give the mean number of small and large BFU-E-derived colonies. The error bars indicate donor-to-donor variability for each vector (\pm SEM). The trend was not significant (student t-test). **(E)** Histogram represents the number of BFU-E-derived colonies enumerated from 1000 MEP (Megakaryocyte–Erythroid Progenitor)-enriched seeded cells (*right panel*). Data shown are means \pm SD (n=3). The increased number of BFU-Es enumerated in the shRNA-*RINF* condition was statistically significant (student T-test, * $p < 0.05$). The gating strategy used for MEP-enrichment is shown (*right panel*). The protocol is adapted from Manz *et al.*(1) For this, CD34+ cells were isolated from primary human bone marrow and stained with CD34, CD38, CD10, CD123 and CD45RA (see Supplemental Methods). We have verified that these MEP cells only generate erythroid colonies (BFU-Es) in our experimental conditions.

Figure S4. *RINF* silencing alters the TGF β signaling pathway in K562 and UT7_{5.3} cells and accelerates cell hemoglobinization triggered by hemin and EPO, respectively. The pTRIPDU3 lentiviral system was used to constitutively express a

shRNA targeting *RINF* expression in K562 (*upper panels*) and UT7_{5.3} (*lower panels*) cell line. Two days after transduction, cells positive for GFP expression were sorted. **(A and B)** Cells were then respectively treated for the indicated time with hemin (40μM) for K562, or EPO (5UI/ml) for UT7_{5.3}. **(A)** The relative expression of *RINF* was measured by Western-Blot analysis (see Methods). ACTIN (*ACTB*) was used as a loading control. **(B)** For hemoglobin production analysis, the benzidine test was used. The percentage of benzidine positive cells was established by counting at least 300 cells per sample. **(C)** Transcriptomic analysis was performed with HTA 2.0 Human Transcriptome Array 2.0 GeneChip (Affymetrix), and previously described (12). Venn diagram representing the number of genes downregulated after *RINF* knockdown. 193 genes were downregulated in K562 and UT7_{5.3} cell lines with at least a 1.25-fold-change (log 2), corresponding to a non-stringent condition. In absolute value, this cut-off corresponds to a 2.38 expression fold-change. We applied QIAGEN's Ingenuity® Pathway Analysis (IPA®, QIAGEN Redwood City, www.qiagen.com/ingenuity) tool to this gene list. IPA revealed an enrichment in genes of the TGFβ signaling pathway and *SMAD7* as a potential *RINF* target. **(D)** Histograms represent the *RINF* or *SMAD7* mRNA levels detected by quantitative RT-PCR for cells. Values are expressed in percentage of the *shControl* condition. A student T test (unpaired) was performed to assess if the observed relative difference was statistically significant. **(E)** Among the list of 193 genes, 39 were also commonly downregulated in MV4-11 cell line. This gene list is presented as well as the heatmap (*right panel*) of the 39 genes representing the relative mRNA expression fold-change (log2) between *shRNA-RINF* and *shRNA-Control* expressing cell lines. *CXXC5* and *SMAD7* are indicated in red.

Figure S5. Ectopic overexpression of *RINF* delays hemoglobinization of human hematopoietic cell lines K562 and UT7_{5.3} triggered by hemin and EPO, respectively. The MigR1 retroviral system (schematized in *upper panel A*) was used to constitutively express *RINF* in two cell lines. A few days after transduction, cells were sorted for GFP expression. Cells were then respectively treated with hemin (40μM) for K562 or EPO (5UI/ml) for UT7_{5.3}. **(A)** The relative expression of *RINF* was measured by Western-Blot analysis using a polyclonal rabbit antibody (see Methods) that detects a specific band at 33 kDa in total extracts. ACTIN (*ACTB*) was used as a loading control. **(B)** For hemoglobin production analysis, the benzidine test was used.

Percentage of benzidine positive cells was established by counting about 300 cells per sample (Student T-test). Pictures represent one of three independent experiments.

Figure S6. *SMAD7* gene expression correlates with *CXXC5* in CD34+ cells isolated from healthy donors or MDS patients with del(5q) and is a favorable prognostic indicator in MDS. For these analyses, two microarray datasets (Affymetrix GeneChip Human Genome U133 Plus 2.0 arrays) respectively performed by Pellagatti *et al.* (183 patients and 17 healthy control donors) or Gerstund *et al.* (159 patients and 17 healthy control donors) were downloaded from the Gene Expression Omnibus website (<http://www.ncbi.nlm.nih.gov/geo/>) by using the followings accession numbers: GSE19429 or GSE58831. **(A)** Heatmap of Pearson correlation coefficient matrix represents, in red, the gene probesets that positively correlate with *SMAD7* (and in blue, the one that negatively correlate with *SMAD7*). To create this heatmap, gene expression data of the high-risk CDR genes (n=43) were extracted from GSE19429 performed by Pellagatti *et al.* (13) and correlated with *SMAD7* mRNA expression (204790_at). **(B)** Genes of the high risk CDR, as well as sequence tag site markers, are listed in the order of their physical positions (megabase, Mb) along chromosome 5q from centromere (upper) to telomere (bottom), according to the most recent human genome assembly at <http://ensembl.org> (Human, GRCh38.p12). **(C)** Histogram represents the *SMAD7* mRNA levels (probeset 204790_at) detected in the microarray dataset performed by Gerstund *et al.*(14) Values are expressed in arbitrary units. A student T test (unpaired) was performed to assess if the observed relative difference was statistically significant between the 3 groups. **(D)** Overall survival data were available for 123 patients, and a tertile stratification was performed. For this, the patients have been classified in 3 equivalent groups according to a high (n=41), an intermediate (n=41), or a low (n=41) *SMAD7* mRNA expression level. The Kaplan-Meier curves (for survival analysis) and the log-rank test were performed by using GraphPad Prism 7.0. P value (log-rank test) of the comparison of the various groups of patients is indicated. In green, the survival curve of the third of patients with the *lowest SMAD7* expression (n=41). The median survival was significantly shorter for this group than for the group with *intermediary* or *high SMAD7* expression (n=82). Note that the median *SMAD7* mRNA level was still significantly lower in the MDS group with “*high*” *SMAD7* than in the healthy control donors (Student t-Test, P=0.005).

Figure S7. Schematic representation of the RINF/SMAD7 signaling pathway that occurs during human erythropoiesis. The black box recapitulates the main erythropoietic features obtained from immature erythroid cells (i.e. progenitors and proerythroblasts) with a high *RINF* mRNA expression, and the red box, the ones with attenuated/knocked-down *RINF* mRNA expression. *RINF* directly binds to *SMAD7* promoter to maintain its expression. *RINF* knockdown cells have a moderate loss of *SMAD7*, that is enough to sensitize erythroid cells to TGF β -signaling and mitigate RBC expansion. This signaling could contribute to the inefficient erythropoiesis observed in malignant hemopathies. The *RINF* gene is rarely mutated in hematopoietic malignancies but several cytogenetic abnormalities (i.e. del(5q)(11,15-17), t(15;17)(11) MLL rearrangements, and t(8;21)(15) and gene mutations (such as GATA2 (15), WT1(15, 18) have been shown to affect *RINF* expression in MDS or AML.

Supplemental References

1. Manz MG, Miyamoto T, Akashi K, Weissman IL. Prospective isolation of human clonogenic common myeloid progenitors. *Proceedings of the National Academy of Sciences of the United States of America* 2002;99(18):11872-11877
2. Gandoura S, Weiss E, Rautou PE, *et al.* Gene- and exon-expression profiling reveals an extensive LPS-induced response in immune cells in patients with cirrhosis. *J Hepatol* 2013;58(5):936-948
3. Wang E, Aslanzadeh V, Papa F, Zhu H, de la Grange P, Cambi F. Global profiling of alternative splicing events and gene expression regulated by hnRNPH/F. *PLoS One* 2012;7(12):e51266
4. de la Grange P, Dutertre M, Correa M, Auboeuf D. A new advance in alternative splicing databases: from catalogue to detailed analysis of regulation of expression and function of human alternative splicing variants. *BMC Bioinformatics* 2007;8:180
5. de la Grange P, Dutertre M, Martin N, Auboeuf D. FAST DB: a website resource for the study of the expression regulation of human gene products. *Nucleic acids research* 2005;33(13):4276-4284
6. Liu TX, Becker MW, Jelinek J, *et al.* Chromosome 5q deletion and epigenetic suppression of the gene encoding alpha-catenin (CTNNA1) in myeloid cell transformation. *Nature medicine* 2007;13(1):78-83
7. Jerez A, Gondek LP, Jankowska AM, *et al.* Topography, clinical, and genomic correlates of 5q myeloid malignancies revisited. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2012;30(12):1343-1349.

8. Merryweather-Clarke AT, Atzberger A, Soneji S, et al. Global gene expression analysis of human erythroid progenitors. *Blood*. 2011;117(13):e96-108.
9. An X, Schulz VP, Li J, et al. Global transcriptome analyses of human and murine terminal erythroid differentiation. *Blood*. 2014;123(22):3466-3477.
10. Keller MA, Addya S, Vadigepalli R, et al. Transcriptional regulatory network analysis of developing human erythroid progenitors reveals patterns of coregulation and potential transcriptional regulators. *Physiol Genomics*. 2006;28(1):114-128.
11. Pendino F, Nguyen E, Jonassen I, et al. Functional involvement of RINF, retinoid-inducible nuclear factor (CXXC5), in normal and tumoral human myelopoiesis. *Blood*. 2009;113(14):3172-3181.
12. Bruserud O, Reikvam H, Fredly H, et al. Expression of the potential therapeutic target CXXC5 in primary acute myeloid leukemia cells - high expression is associated with adverse prognosis as well as altered intracellular signaling and transcriptional regulation. *Oncotarget*. 2015;6(5):2794-2811.
13. Pellagatti A, Cazzola M, Giagounidis AA, et al. Gene expression profiles of CD34+ cells in myelodysplastic syndromes: involvement of interferon-stimulated genes and correlation to FAB subtype and karyotype. *Blood*. 2006;108(1):337-345.
14. Gerstung M, Pellagatti A, Malcovati L, et al. Combining gene mutation with gene expression data improves outcome prediction in myelodysplastic syndromes. *Nat Commun*. 2015;6:5901.
15. Kuhn A, Valk PJ, Sanders MA, et al. Downregulation of the Wnt inhibitor CXXC5 predicts a better prognosis in acute myeloid leukemia. *Blood*. 2015;125(19):2985-2994.
16. Stoddart A, Qian Z, Fernald AA, et al. Retroviral insertional mutagenesis identifies the del(5q) genes, CXXC5, TIFAB and ETF1, as well as the Wnt pathway, as potential targets in del(5q) myeloid neoplasms. *Haematologica*. 2016;101(6):e232-236
17. Treppendahl MB, Mollgard L, Hellstrom-Lindberg E, Cloos P, Gronbaek K. Downregulation but lack of promoter hypermethylation or somatic mutations of the potential tumor suppressor CXXC5 in MDS and AML with deletion 5q. *Eur J Haematol*. 2013;90(3):259-260.
18. Kim MS, Yoon SK, Bollig F, et al. A novel Wilms tumor 1 (WT1) target gene negatively regulates the WNT signaling pathway. *J Biol Chem*. 2010;285(19):14585-14593.

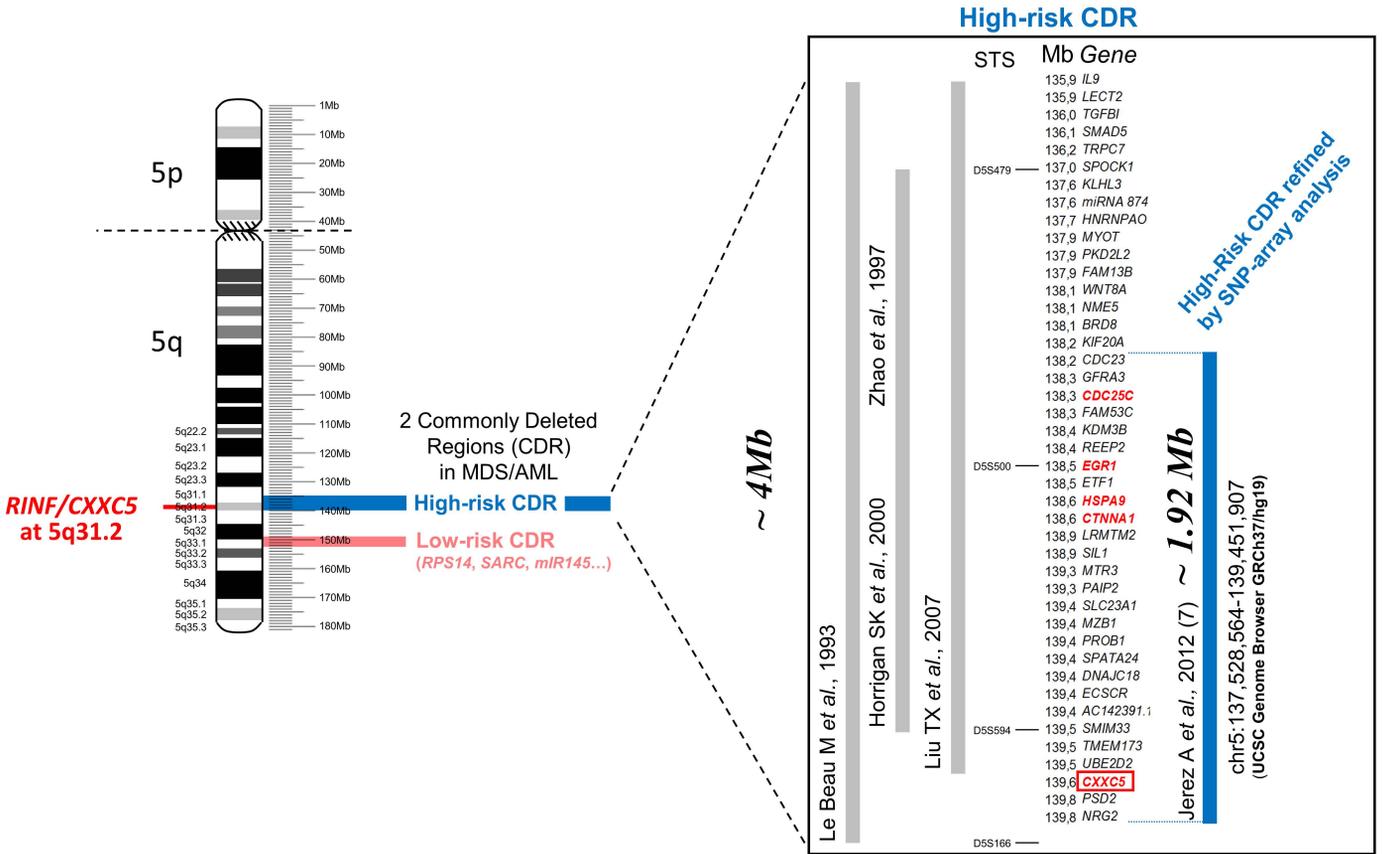


Figure S1
Astori and Matherat *et al.*

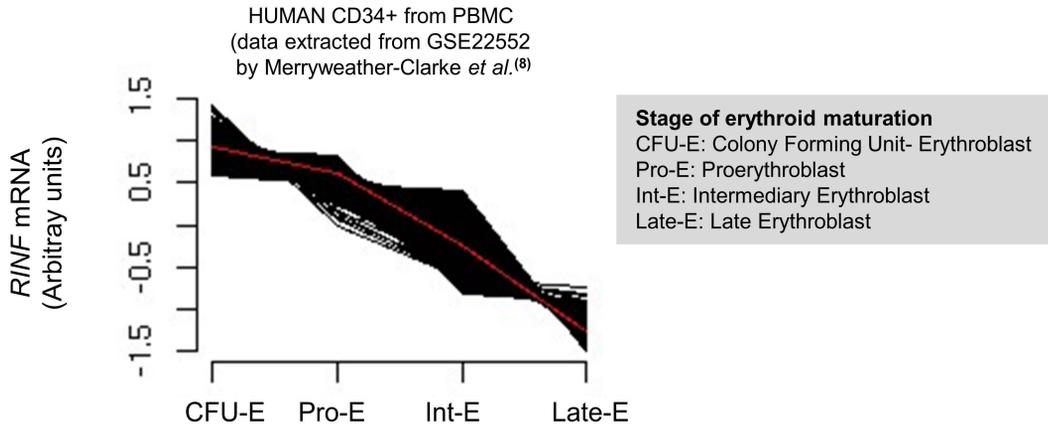
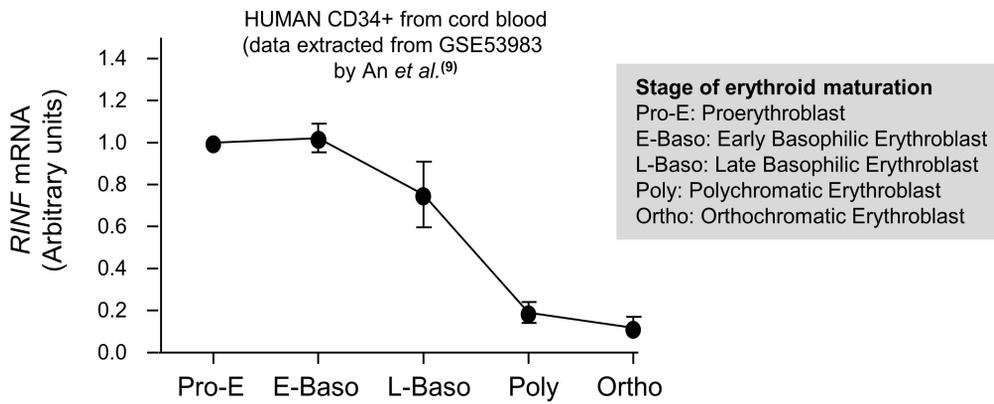
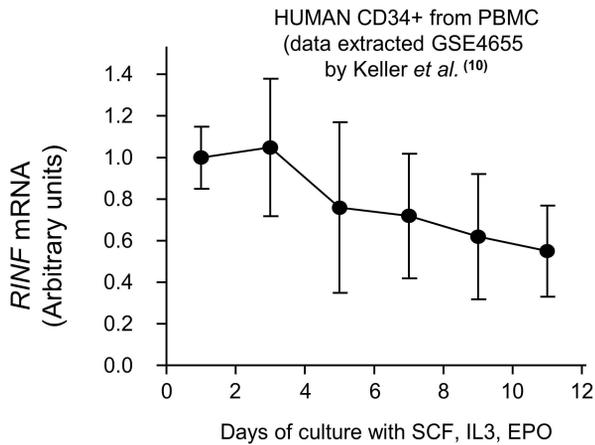
A**B****C**

Figure S2
Astori and Matherat *et al.*

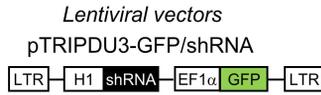
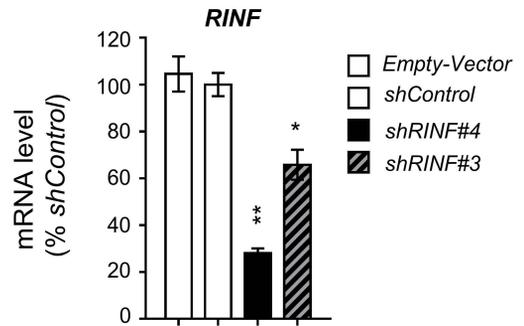
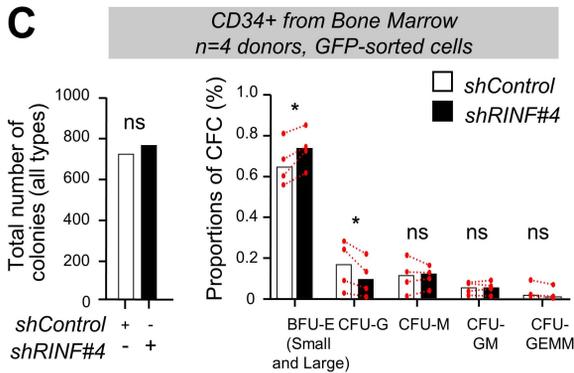
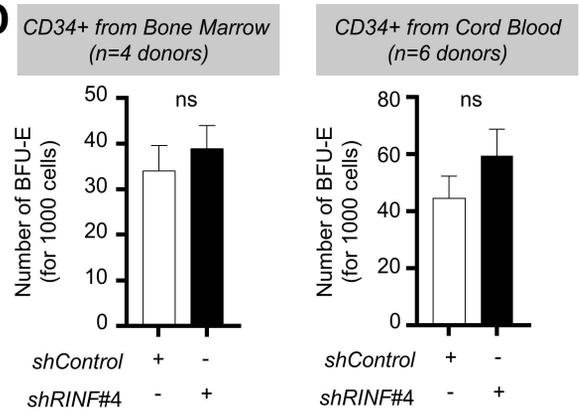
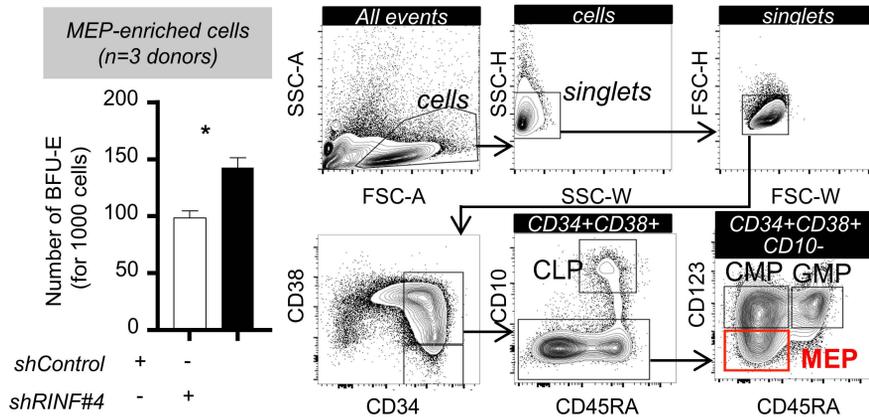
A**B****C****D****E**

Figure S3
Astori and Matherat *et al.*

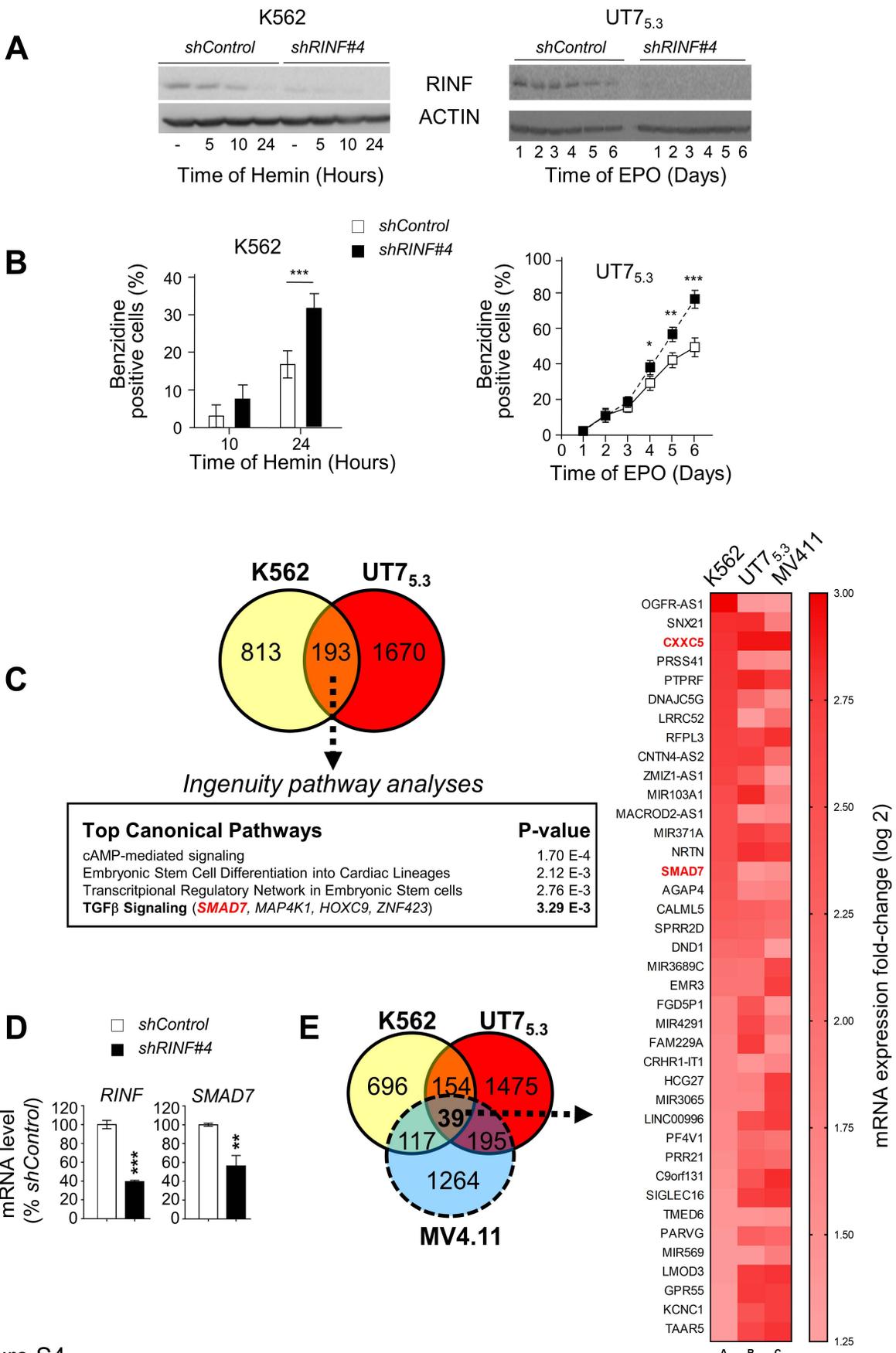


Figure S4
Astori and Matherat *et al.*

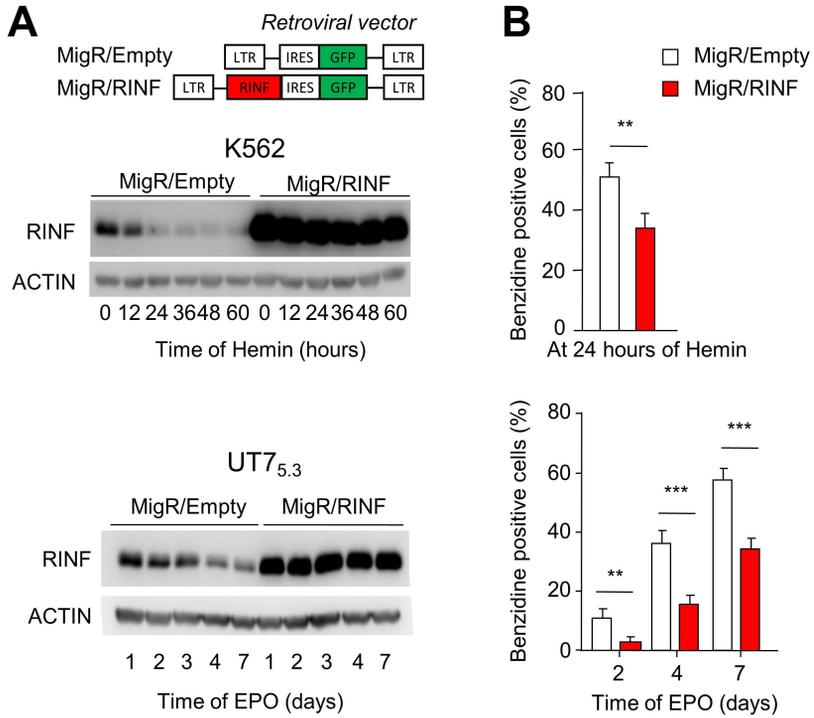


Figure S5
Astori and Matherat *et al.*

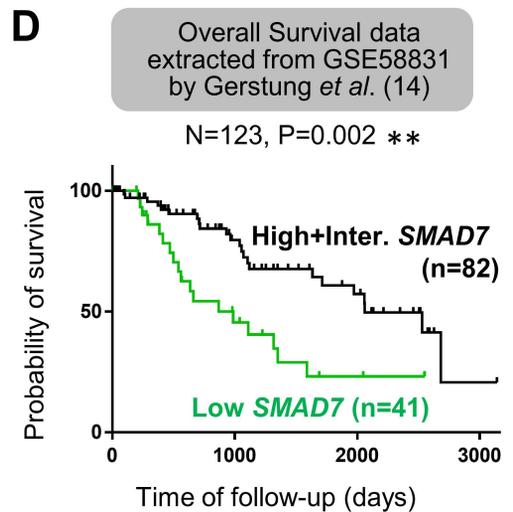
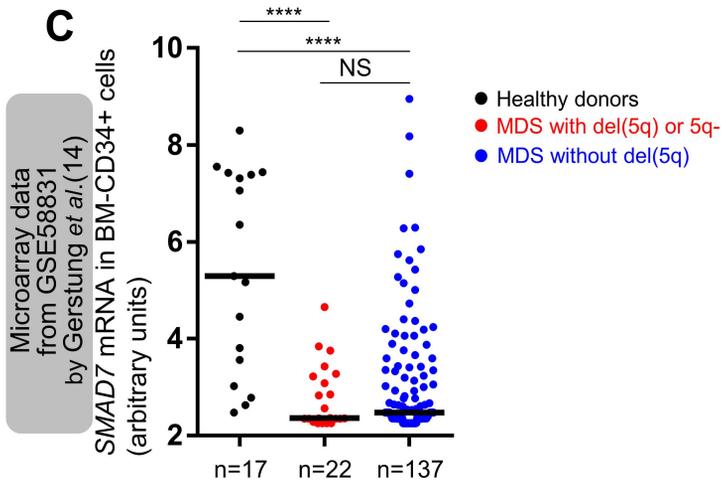
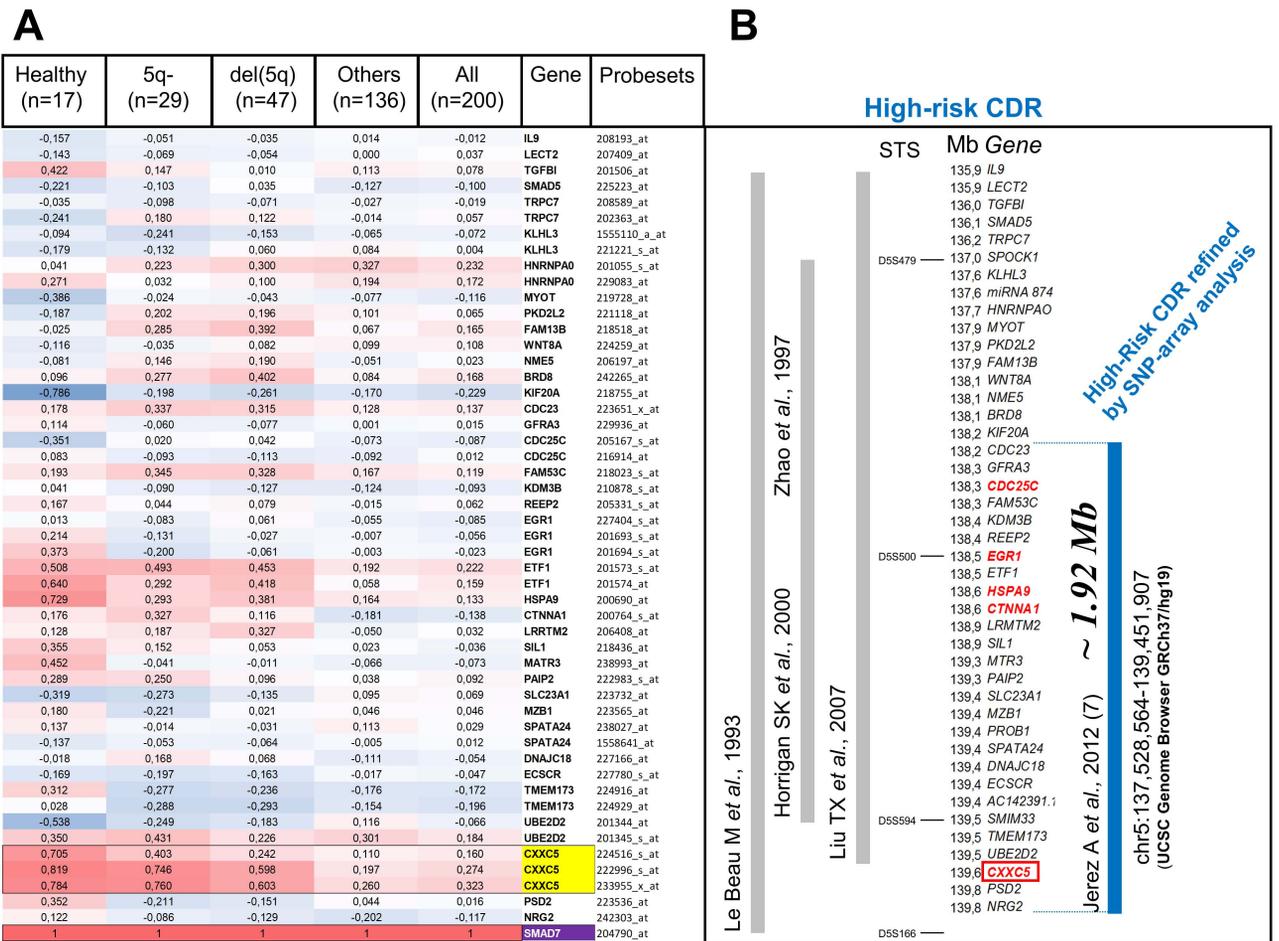


Figure S6
Astori and Matherat et al.

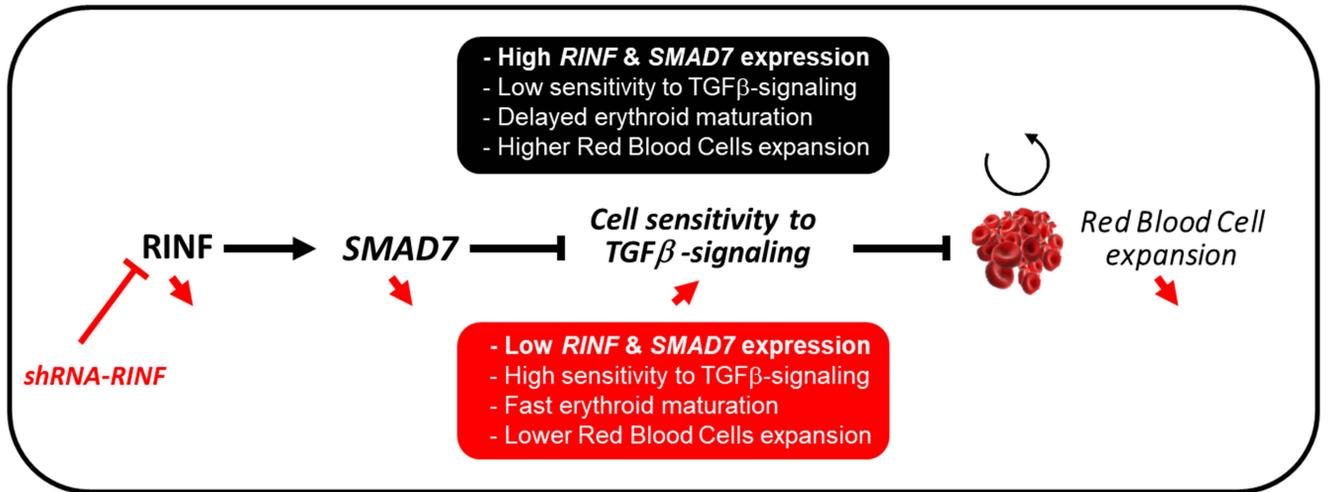


Figure S7
 Astori and Matherat *et al.*