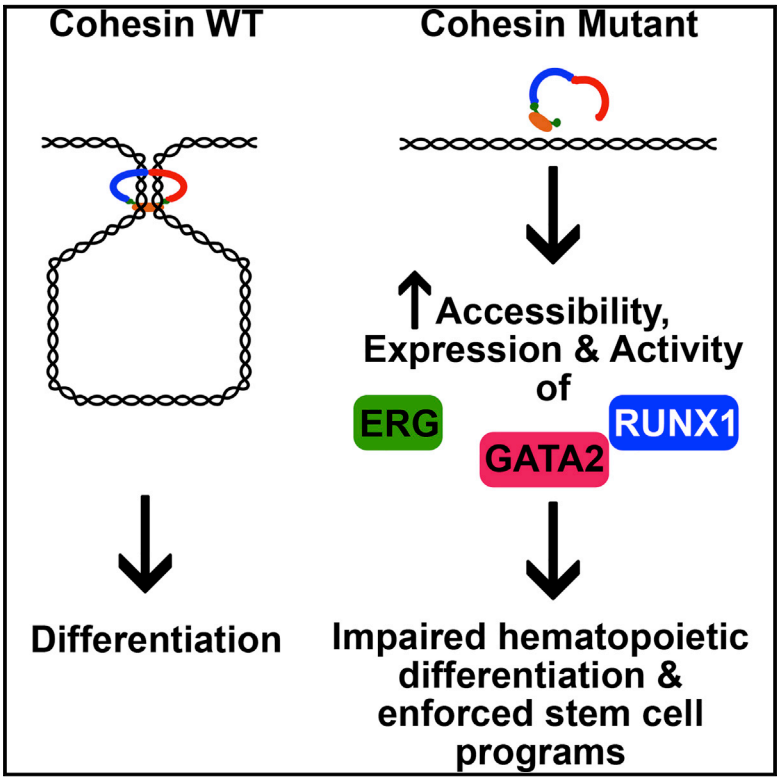
Article

Cell Stem Cell

Leukemia-Associated Cohesin Mutants Dominantly Enforce Stem Cell Programs and Impair Human Hematopoietic Progenitor Differentiation

Graphical Abstract



Authors

Claire Mazumdar, Ying Shen, Seethu Xavy, ..., Wan-Jen Hong, Howard Y. Chang, Ravindra Majeti

Correspondence

rmajeti@stanford.edu

In Brief

Mazumdar et al. investigate the effects of leukemia-associated cohesin complex mutations on human hematopoietic stem and progenitor cells. Cohesin mutants are found to impair differentiation and enforce stem cell programs through the modulation of ERG, GATA2, and RUNX1 chromatin accessibility, expression, and activity, suggesting a mechanism contributing to leukemogenesis.

Highlights

- Cohesin mutants impair differentiation and enforce stem cell programs in human HSPCs
- Effects are cell context dependent, restricted to immature HSC and MPP populations
- Mutants showed increased chromatin accessibility and binding of ERG, GATA2, and RUNX1
- Cohesin-mutant-induced stem cell programs are dependent on ERG, GATA2, and RUNX1

Accession Numbers

GSE73224
GSE73206
GSE73207

Mazumdar et al., 2015, Cell Stem Cell *17*, 675–688
December 3, 2015 ©2015 Elsevier Inc.
http://dx.doi.org/10.1016/j.stem.2015.09.017

Cell Stem Cell
Article

Leukemia-Associated Cohesin Mutants Dominantly Enforce Stem Cell Programs and Impair Human Hematopoietic Progenitor Differentiation

Claire Mazumdar,[1] Ying Shen,[2] Seethu Xavy,[1] Feifei Zhao,[1] Andreas Reinisch,[1] Rui Li,[2] M. Ryan Corces,[1] Ryan A. Flynn,[2] Jason D. Buenrostro,[2,3] Steven M. Chan,[1] Daniel Thomas,[1] Julie L. Koenig,[1] Wan-Jen Hong,[1] Howard Y. Chang,[2] and Ravindra Majeti[1,*]

[1]Department of Medicine, Division of Hematology, Cancer Institute, and Institute for Stem Cell Biology and Regenerative Medicine, Stanford University School of Medicine, Stanford, CA 94305, USA

[2]Center for Personal Dynamic Regulomes and Program in Epithelial Biology, Stanford University School of Medicine, Stanford, CA 94305, USA

[3]Department of Genetics, Stanford University School of Medicine, Stanford, CA 94305, USA

*Correspondence: rmajeti@stanford.edu

http://dx.doi.org/10.1016/j.stem.2015.09.017

SUMMARY

Recurrent mutations in cohesin complex proteins have been identified in pre-leukemic hematopoietic stem cells and during the early development of acute myeloid leukemia and other myeloid malignancies. Although cohesins are involved in chromosome separation and DNA damage repair, cohesin complex functions during hematopoiesis and leukemic development are unclear. Here, we show that mutant cohesin proteins block differentiation of human hematopoietic stem and progenitor cells (HSPCs) in vitro and in vivo and enforce stem cell programs. These effects are restricted to immature HSPC populations, where cohesin mutants show increased chromatin accessibility and likelihood of transcription factor binding site occupancy by HSPC regulators including ERG, GATA2, and RUNX1, as measured by ATAC-seq and ChIP-seq. Epistasis experiments show that silencing these transcription factors rescues the differentiation block caused by cohesin mutants. Together, these results show that mutant cohesins impair HSPC differentiation by controlling chromatin accessibility and transcription factor activity, possibly contributing to leukemic disease.

INTRODUCTION

Acute myeloid leukemia (AML) is an aggressive malignancy of the bone marrow characterized by expansion of immature myeloid cells defective in their maturation and function (Estey and Döhner, 2006; Löwenberg et al., 1999). Large-scale AML genome re-sequencing efforts have identified novel recurrently mutated genes, including the members of the cohesin complex (*RAD21*, *SMC3*, *SMC1A,* and *STAG2*), implicated in the pathogenesis of this disease (Cancer Genome Atlas Research Network, 2013). The manner in which multiple mutations accumulate in a single lineage of cells has been an area of intense investigation.

Over the last decades, a number of studies from many investigators have led to a model in which leukemic mutations serially accumulate in clones of self-renewing hematopoietic stem cells (HSCs) (reviewed in Corces-Zimmerman and Majeti, 2014; Jan and Majeti, 2013), and we provided evidence supporting this model through the identification of residual HSCs from the time of AML diagnosis harboring some, but not all, of the mutations in the patient-matched leukemia cells (Jan et al., 2012). We termed these cells pre-leukemic HSCs and further demonstrated the serial accumulation of pre-leukemic mutations in single cells as hypothesized. Subsequent analysis of a larger cohort of patients demonstrated that pre-leukemic mutations preferentially occur in genes involved in the regulation of the epigenome, including cohesin (Corces-Zimmerman et al., 2014).

Recurrent mutations in cohesin have been identified in myelodysplastic syndrome (MDS), AML, and other myeloid malignancies (Leeke et al., 2014). Cohesin mutations are almost always mutually exclusive, suggesting that an alteration in one component may be sufficient to affect the entire complex, and they collectively occur in approximately 15% of AML cases and other myeloid malignancies (Kon et al., 2013; Cancer Genome Atlas Research Network, 2013).

The cohesin complex functions to hold chromatin strands within a ring-like structure composed of the four core components (Diaz-Martinez and Clarke, 2009). Although its best-established role is to maintain the polarity of sister chromatids during mitosis, cohesin is also involved in double-stranded DNA damage repair and regulation of transcription through transcription factor (TF) recruitment and its interaction with CTCF (Panigrahi and Pati, 2012). The discovery of recurrent mutations in the cohesin complex indicates another pathway that is relevant for AML, yet the manner in which they contribute to pathogenesis has not been defined. Here, we build on the observation that cohesin mutations are often seen as early, founder mutations in pre-leukemic HSCs (Corces-Zimmerman et al., 2014) and investigated their effects on differentiation and self-renewal of human hematopoietic stem and progenitor cells (HSPCs).

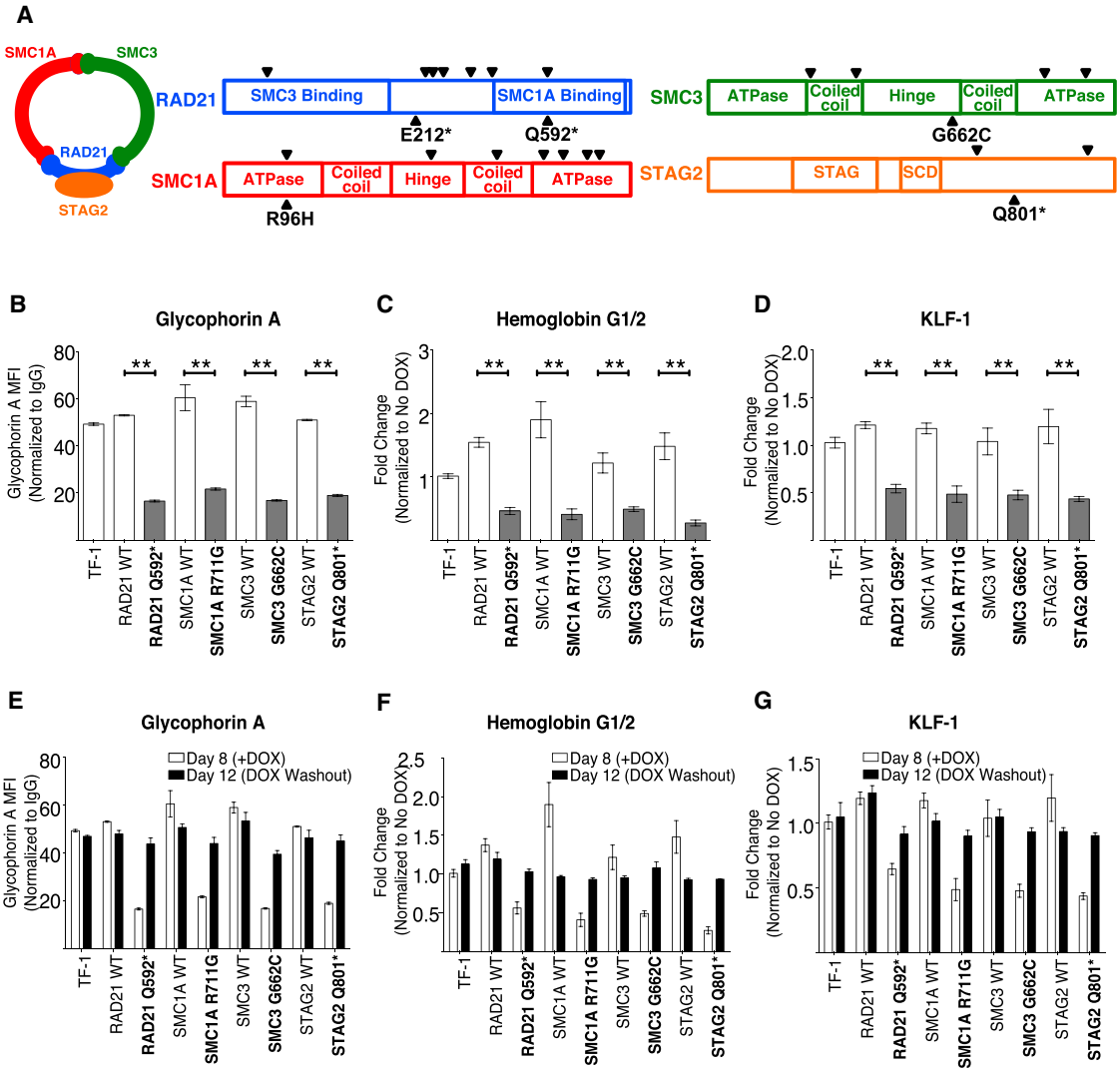


Figure 1. Cohesin Mutants Impair Hematopoietic Differentiation

(A) Schematic diagram of the cohesin complex with four components (RAD21, SMC1A, SMC3, and STAG2). Mutations found in AML (in the Cancer Genome Atlas Research Network or our own tissue bank) in each cohesin component are indicated by triangles. Specific mutations used for studies reported here are indicated. (*) indicates nonsense mutations.

(B) TF-1 cells were infected with lentiviruses encoding doxycycline (DOX)-inducible cohesin WT or mutant variants and GFP. Erythroid differentiation of parental TF-1 cells and variants was determined by flow cytometry for GPA expression after 2 initial days of DOX treatment and 8 days of EPO and DOX treatment. Relative expression is shown as mean fluorescence intensity (MFI) of GPA normalized to IgG isotype control. $^{**}p < 0.01$.

(C) Expression of fetal hemoglobin was determined by qRT-PCR for cells treated as in (B). Values are normalized to No DOX controls. $^{**}p < 0.01$.

(D) Expression of KLF-1 was determined by qRT-PCR for cells treated as in (B). Values are normalized to No DOX controls. $^{**}p < 0.01$.

(E) TF-1 cells were treated as in (B) and then washed out of DOX and maintained in EPO media for an additional 4 days. Erythroid differentiation of parental TF-1 cells and variants was determined by flow cytometry for GPA expression. Relative expression is shown as MFI of GPA normalized to IgG isotype control. No statistically significant differences were detected.

(F) Expression of fetal hemoglobin was determined by qRT-PCR for cells treated as in (E). Values are normalized to TF-1 control. No statistically significant differences were detected.

(G) Expression of KLF-1 was determined by qRT-PCR for cells treated as in (E). Values are normalized to TF-1 control. No statistically significant differences were detected.

RESULTS

Cohesin Mutants Impair Hematopoietic Differentiation

Mutations in cohesin complex genes associated with myeloid malignancies are usually heterozygous and include both missense and nonsense mutations distributed throughout their coding regions (Figure 1A), suggesting that these mutations result in haploinsufficiency and/or the generation of dominant-negative mutant proteins that disrupt cohesin functions (Kon et al., 2013). Thus, in order to investigate the effects of cohesin mutations on hematopoiesis, we engineered human AML cell lines (verified to lack endogenous mutations in all cohesin

components) to express WT or mutant cohesin components under the control of a doxycycline (DOX)-inducible promoter (Figures S1A and S1B). We investigated multiple mutations (Figure 1A) identified in primary AML samples from our own laboratory or in the Cancer Genome Atlas Research Network cohort, including both missense and nonsense mutations in all four core cohesin components (*RAD21* E212*, *RAD21* Q592*, *SMC1A* R711G, *SMC3* G662C, and *STAG2* Q801*, where [*] indicates a stop codon).

Cohesin mutants were initially investigated in the TF-1 erythroleukemia cell line due to its ability to undergo partial erythroid differentiation in response to erythropoietin (EPO). Upon exposure to EPO for 6 days, cohesion-mutant-induced cell lines exhibited a significant decrease in erythroid differentiation as determined by surface expression of glycophorin A (GPA) (Figure 1B) and RNA expression of fetal hemoglobin (HGB1/2) (Figure 1C) and KLF-1 (Figure 1D), a key erythroid TF. This suggests that cohesin mutants can act in a dominant-negative manner to impair hematopoietic differentiation. This phenotype was supported by expression of WT and mutant cohesin components in the THP-1 AML cell line, which can undergo myeloid differentiation upon addition of all-trans retinoic acid (ATRA) or phorbol 12-myristate 13-acetate (PMA). Upon exposure to these agents, cohesion-mutant-induced cell lines exhibited significantly decreased surface expression of the mature myeloid marker CD11b compared to WT or parental THP-1 cells (Figures S1C and S1D).

Although previously published reports (Kon et al., 2013) showed an effect of cohesin mutants on cell proliferation, we saw no significant changes in either proliferation (Figure S1E) or cell death (Figure S1F). Furthermore, given the role of cohesin in double-stranded DNA damage repair, we determined expression of phospho-gamma H2AX, a marker of DNA double-strand breaks. Even after 18 days of cohesin mutant expression, there were no changes in DNA damage compared to WT cells (Figure S1G), consistent with the finding that the majority of cohesin mutant AML cases are of normal karyotype (Cancer Genome Atlas Research Network, 2013).

To determine whether this impaired differentiation phenotype was dependent on continuous expression of cohesin mutants, TF-1 cells initially DOX-induced for 6 days in the presence of EPO were removed from DOX and replated in EPO-containing medium. Analysis 4 days later showed a strong induction of erythroid markers in all mutant cell lines (Figures 1E–1G), indicating that the impairment of differentiation is reversible in this model.

Cohesin Mutants Impair Myeloid, Erythroid, and Stem Cell Differentiation of Primary Human HSPCs

These cell line assays support the hypothesis that cohesin mutations contribute to disease pathogenesis by impairing hematopoietic differentiation in HSPCs. To directly investigate this hypothesis, we transduced CD34-enriched primary normal human cord blood cells with lentivirus encoding constitutive expression of either WT or mutant cohesin components with a GFP marker (Figure S2A**)**. Transduced GFP+ cells were purified by fluorescence activated cell sorting (FACS) and cultured under several conditions (Figure 2A). First, cells were cultured with cytokines designed to promote retention of HSPCs (Chan et al., 2015), and these cultures showed a significantly increased frequency of CD34-expressing cells compared to WT or empty vector controls (Figures 2B and 2C). Second, cells were cultured under conditions designed to promote myeloid differentiation (Sinha et al., 2015), and resulting cultures showed a significant decrease in CD14-expressing cells compared to controls (Figures 2D and 2E). Third, cells were cultured under conditions designed to promote erythroid differentiation, and cohesin mutant cultures showed a significant decrease in CD71 and GPA-double-positive erythroid cells (Figures 2F and 2G). As with the cell lines, expression of cohesin mutants did not affect cord blood cell proliferation or cell death (Figures S2B and S2C). No impaired differentiation was observed with FACS-purified GFP− cells from the same normal donors in these assays (Figures S2D–S2F). Together, these data suggest that cohesin mutations impart a differentiation block on primary human HSPCs.

Knockdown of RAD21 Impairs Hematopoietic Differentiation In Vitro and In Vivo

In addition to acting in a dominant-negative manner, cohesin mutations may also act through loss of function. To investigate this mechanism, we used a second model system in which we knocked down the expression of RAD21 using shRNA-encoding lentiviruses. We engineered inducible shRNA vectors containing one of three independent validated RAD21 targeting sequences, and separately, two constitutive RAD21-shRNA targeting vectors, to silence RAD21 in TF-1 cells and CD34-enriched cord blood cells. Both constitutive and inducible shRNA-mediated knockdown led to a reduction in RAD21 expression by 80%–90% in CD34+ cord blood cells (Figure S3A) and strong protein knockdown in TF-1 cells (Figure S3B). When cultured with EPO, the RAD21-knockdown TF-1 cells showed a marked reduction in erythroid differentiation as measured by expression of HGB1/2 and KLF-1 (Figures S3C and S3D). Moreover, constitutive RAD21-knockdown in cord blood HSPCs followed by culture in the HSPC-retention medium resulted in a significantly increased frequency of CD34-expressing cells compared to control (Figures S3E and S3F). Similar to the results with cohesin mutants, knockdown of RAD21 in HSPCs showed a marked decrease in CD14-expressing cells in the myeloid culture medium (Figures S3G and S3H) and a significant decrease in CD71 and GPA-double-positive erythroid cells in the erythroid culture medium (Figures S3I and S3J).

Finally, we investigated the effects of RAD21-knockdown on human HSPC engraftment and differentiation in vivo in NOD/SCID/IL2R-gamma null (NSG) mice. CD34-enriched cord blood cells were transduced with lentivirus encoding for constitutive RAD21-shRNA knockdown and then transplanted into NSG mice. After 7 weeks, there were no statistically significant differences in human hematopoietic engraftment levels between scrambled shRNA control and RAD21-shRNAs (Figure 3A). However, there was a significant increase in human CD34-expressing cells in the bone marrow (Figures 3B and 3C), indicating that cohesin mutants can impair hematopoietic differentiation in vivo. Furthermore, flow cytometry analysis of the engrafted human cells showed a significant increase in CD33+ myeloid cells compared to CD19+CD20+ B-lymphoid cells, indicating a skewing toward the myeloid lineage (Figures 3D and 3E). This

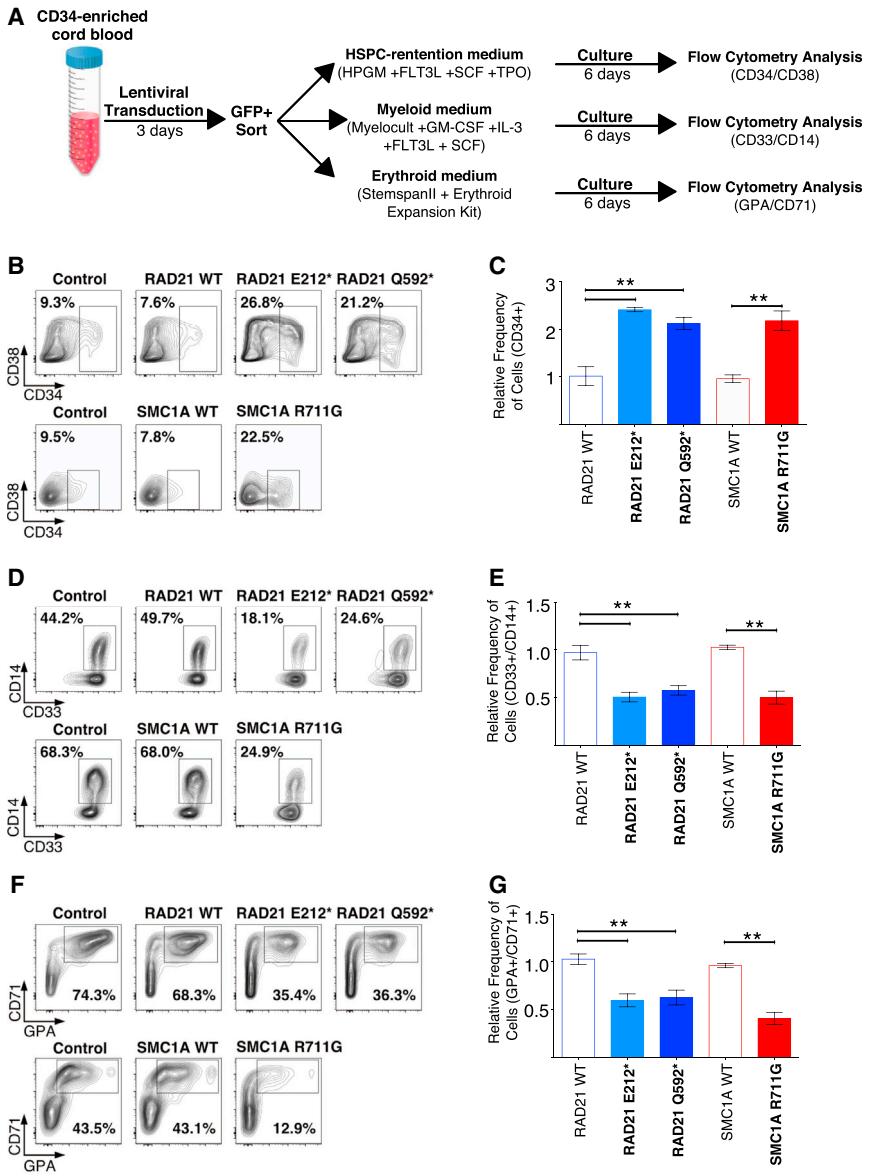


Figure 2. Cohesin Mutants Impair Myeloid, Erythroid, and Stem Cell Differentiation of Primary Human HSPCs

(A) Schematic of in vitro differentiation experiments using human CD34-enriched HSPCs from cord blood.

(B) HSPCs were infected with lentiviruses encoding GFP alone (control) or GFP in addition to the indicated cohesin variants. 72 hr post-infection, GFP+ cells were isolated by FACS and cultured in HSPC-retention medium. 6 days later, cells were analyzed for expression of progenitor markers CD34 and CD38. Representative FACS plots are shown.

(C) Summary of data from three independent experiments as described in (B); the percentage of CD34+ cells was normalized to GFP control. Unpaired Student's t test was used to determine statistical significance between WT and mutant populations. $^{**}p < 0.01$.

(D) HSPCs isolated and lentivirally transduced as in (B) were cultured in myeloid differentiation medium. 6 days later, cells were analyzed for expression of myeloid markers CD33 and CD14. Representative FACS plots are shown.

(E) Summary of data from three independent experiments as described in (D); the percentage of CD33+/CD14+ cells was normalized to GFP control. Unpaired Student's t test was used to determine statistical significance between WT and mutant populations. $^{**}p < 0.01$.

(F) HSPCs isolated and lentivirally transduced as in (B) were cultured in erythroid differentiation medium. 6 days later, cells were analyzed for expression of erythroid markers CD71 and GPA. Representative FACS plots are shown.

(G) Summary of data from three independent experiments as described in (F); the percentage of GPA+/CD71+ cells was normalized to GFP control. Unpaired Student's t test was used to determine statistical significance between WT and mutant populations. $^{**}p < 0.01$.

phenotype is consistent with the association of cohesin mutations with human myeloid malignancies.

Thus, as with overexpression of cohesin mutants, decreased expression of cohesin results in impaired hematopoietic differentiation, suggesting that these mutations can act through both haploinsufficiency and dominant-negative mechanisms. We performed whole-cell western blots on primary AML samples (Table S1) with either WT or mutant cohesin to determine cohesin component protein levels. In samples that were mutated for a member of the cohesin complex, we detected a slight reduction in the protein level of that component compared to WT samples, and in some cases, a reduction of its binding partners (Figure S4A). Next, we investigated formation of the cohesin protein complex in WT and mutant primary AML cells using co-immunoprecipitation. Cellular lysates were immunoprecipitated with an antibody to SMC1A and then blotted for protein levels of RAD21 and SMC3. The results indicated varying levels of protein complex formation, although in one of the samples (SU067, Q592*) that had a truncation in RAD21 at the SMC1A-interacting C terminus, no RAD21 was co-immunoprecipitated, indicating that in some cases, cohesin mutations have a dominant-negative effect in abrogating complex formation (Figure S4B).

Cohesin Mutants Induce Myeloid Skewing and Increased Serial Replating of Human HSPCs In Vitro

Next, we investigated the effect of mutant cohesin on human HSPC colony formation and serial replating capacity. Cord blood cells transduced with cohesin WT, mutant, or control lentivirus were plated in methylcellulose for 14 days. No significant differences in the number of colonies were observed after the first plating (Figure 4A); however, an increase in CFU-GM myeloid colonies was detected with expression of the cohesin mutants (Figure 4B). Flow cytometry analysis of the resulting cells demonstrated a statistically significant increase in CD33+ myeloid cells compared to GPA+ erythroid cells in cohesin mutant colonies, indicating a skewing of differentiation to the myeloid lineage

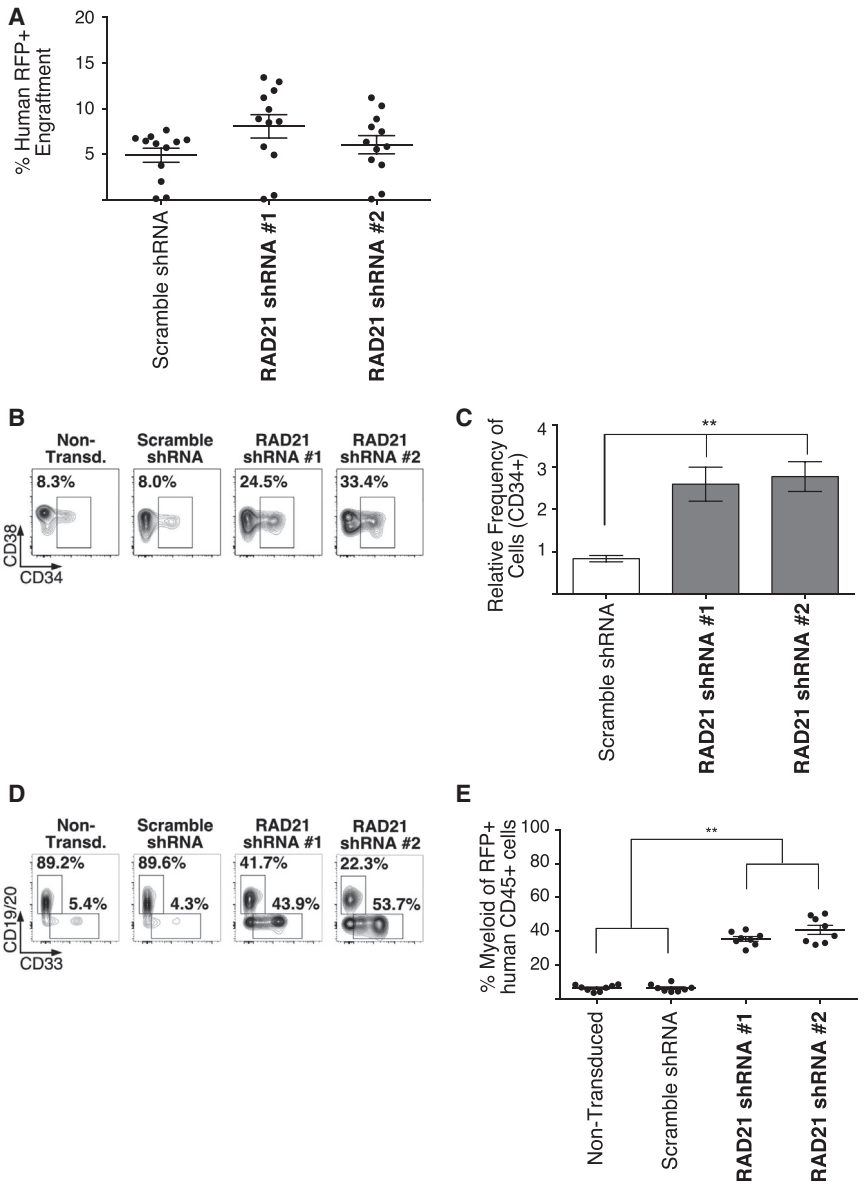


Figure 3. Knockdown of RAD21 Impairs Hematopoietic Differentiation and Induces Myeloid Skewing In Vivo

(A) Human CD34-enriched cord blood HSPCs were infected with constitutive RAD21 shRNA or scramble control RFP-encoding lentiviral vectors. 72 hr post-infection ~50,000 cells were transplanted by intrafemoral injection into NSG mice (three mice per condition). 7 weeks post-transplant, bone marrow aspirates were analyzed for human engraftment (human CD45+) and RFP+ expression. No statistically significant differences were detected.

(B) Human RFP+ engrafted cells from (A) were analyzed for expression of progenitor markers CD34 and CD38. Representative FACS plots are shown.

(C) Summary of data from 3 independent experiments as described in (B); the percentage of CD34+ cells was normalized to non-transduced control. Unpaired Student's t test was used to determine statistical significance between scramble shRNA and RAD21 shRNA populations. $^{**}p < 0.01$.

(D) Human RFP+ engrafted cells from (A) were analyzed for expression of B lymphoid (CD19 & CD20) and myeloid markers (CD33). Representative FACS plots are shown.

(E) Summary of data from three independent experiments as described in (D). The percentage of CD33+ cells is shown. Unpaired Student's t test was used to determine statistical significance between non-transduced and scramble shRNA versus RAD21 shRNA populations. $^{**}p < 0.01$.

(Figure 4C). Upon replating, cohesin mutant cells retained the ability to serially generate colonies through the fifth replating, while essentially no WT or control colonies were detected after the second plating (Figure 4A). This increased serial-replating capacity suggests that mutant cohesin enforces stem cell functions in human HSPCs.

Cohesin Mutants Enforce Expression of Hematopoietic and Leukemia Stem Cell Gene Expression Programs

To further investigate the effect of mutant cohesin on stem cell programs in human HSPCs, we conducted global transcriptional profiling by microarray analysis of cord blood cells transduced with cohesin WT, mutant, or control lentiviral vectors. Using stringent bioinformatic criteria of fold change > 3 and FDR < 0.05, 202 unique genes differentially expressed between cohesin mutant and control HSPCs (94 upregulated and 108 downregulated; Table S2) were identified. Based on hierarchical clustering by Pearson correlation, the mutant samples formed a distinct cluster from WT and control samples (Figure 4D). In the genes upregulated in cohesion-mutant-transduced cord blood, there was a strong enrichment for gene sets highly expressed in normal cord blood HSCs and curated gene sets defining AML stem cell signatures (Eppert et al., 2011; Gentles et al., 2010) (Figure 4E). Genes upregulated at least 3-fold in cohesin mutant samples included many known to be crucial for HSC function, such as homeobox genes (Spencer et al., 2015) and *MEIS1* (Wilson et al., 2010) (Table S2). Conversely, cohesin mutant samples exhibited lower expression of gene sets and genes that are crucial for myeloid differentiation such as myeloperoxidase (*MPO*) and colony stimulating factor 1 receptor (*CSF1R*) (Figure 4F). The enrichment for expression of HSC and AML leukemia stem cell gene sets supports the model that cohesin mutants enforce stem cell programs in human HSPCs.

Cohesin Mutants Impair Human HSPC Differentiation in a Cell-Context-Dependent Manner

Cohesin mutations were identified in pre-leukemic HSCs in several of the cases we investigated (Corces-Zimmerman et al., 2014; Jan et al., 2012), suggesting that they may act to impair hematopoietic differentiation and enforce stem cell

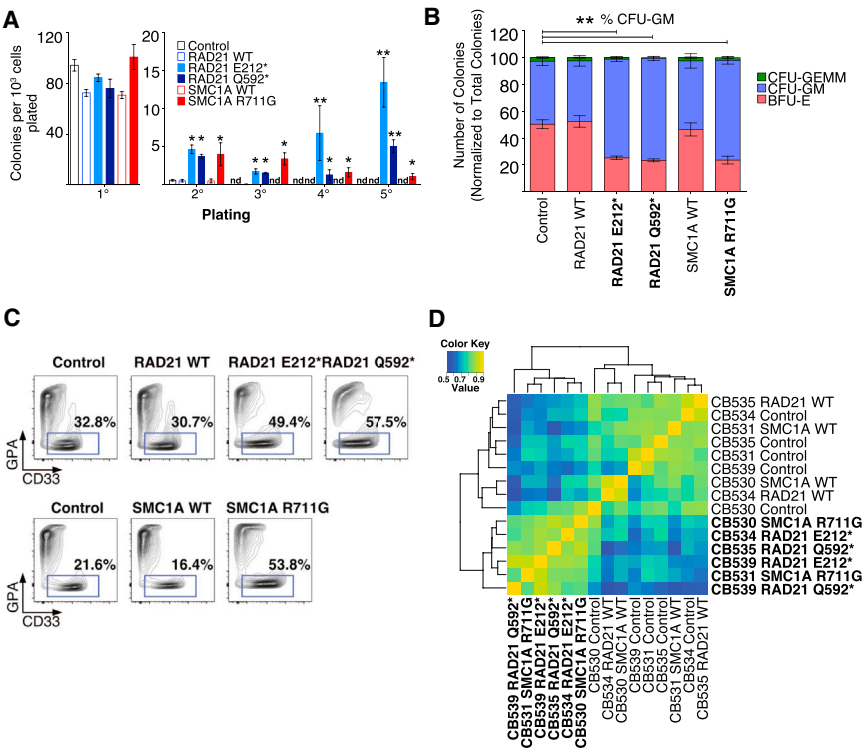


Figure 4. Cohesin Mutants Induce Myeloid Skewing, Increase Serial Replating, and Enforce Stem Cell Gene Expression Programs in Human HSPCs

(A) Human CD34-enriched HSPCs were infected with lentiviruses encoding GFP alone (control) or GFP in addition to the indicated cohesin variants. 72 hr post-infection, GFP+ cells were isolated by FACS and cultured in methylcellulose for colony-forming assays. Every 14 days, colonies were scored for morphology and cells (1,000–10,000) were replated up to five cycles. The number of colonies per 1,000 cells in the first plating is indicated on the left; the right indicates the number of colonies per 1,000 cells in the subsequent platings. nd, none detected. Unpaired Student's t test was used to determine statistical significance between WT and mutant populations. *$p < 0.05$ and **$p < 0.01$.

(B) The morphological colony-types from the primary plating of cells as described in (A) are indicated. Statistically significant differences in the percent of CFU-GM were detected. **$p < 0.01$.

(C) Cells from the methylcellulose colonies from (B) were isolated and analyzed for expression of myeloid marker (CD33) or erythroid marker (GPA). Representative FACS plots are shown. Data are representative of three independent experiments.

(D) Human CD34-enriched HSPCs were infected with control, cohesin WT, or cohesin mutant GFP-encoding lentiviruses. 72 hr later, GFP+ transduced cells were FACS purified and subject to gene expression analysis on Affymetrix microarrays. The resulting expression data was analyzed by unsupervised hierarchical clustering and the Pearson correlation dendogram is shown here.

(E) Gene expression microarray data as described in (D) was utilized in Gene Set Enrichment Analysis (GSEA) to identify gene sets enriched in cohesin mutant compared to cohesin WT and control samples in both the molecular signatures database (MSigDB) and manually curated leukemia stem cell gene sets.

(F) GSEA identified gene sets enriched in WT and control samples compared to cohesin mutants in the MSigDB.

programs in a cell-context-dependent manner. To investigate this hypothesis, six human HSPC subpopulations were isolated from human cord blood by FACS: HSCs, multipotent progenitors (MPPs), lymphoid-primed multipotent progenitors (LMPPs), common myeloid progenitors (CMPs), megakaryocyte-erythroid progenitors (MEPs), and granulocyte-monocyte progenitors (GMPs) (Goardon et al., 2011; Majeti et al., 2007) (Figure S5A). These cells were then transduced with cohesin WT, mutant, or control lentivirus and cultured in either the myeloid-differentiation promoting or erythroid-differentiation promoting conditions. Strikingly, a strong myeloid differentiation block was only observed with cohesion-mutant-transduced HSCs and MPPs, and to a lesser extent CMPs (Figures 5A and 5B). Similar results were observed in the erythroid assay (Figures 5C and 5D). Parallel studies with RAD21 shRNA-mediated knockdown also showed differentiation defects only in the most immature populations (Figures 5B and 5D, Figures S5B and S5C). Notably, cohesion-mutant-transduced and cohesin-deficient GMPs and MEPs showed no differentiation block in the myeloid and erythroid assays, respectively, indicating that the effect of mutant cohesin is context dependent and restricted to the most immature human HSPCs.

Cohesin Mutants Alter Chromatin Accessibility at Transcriptional Regulatory Elements

Cohesin functions to establish and maintain DNA accessibility, and knockdown of cohesin can lead to a decrease in chromatin accessibility at TF clustered regions (Yan et al., 2013), leading us to hypothesize that cohesin mutants impart their phenotypic effects through modulation of chromatin accessibility. To investigate this hypothesis, we employed a recently described technique known as the assay for transposase-accessible chromatin followed by sequencing (ATAC-seq) (Buenrostro et al., 2013) to assess the genome-wide accessibility to the Tn5 transposase.

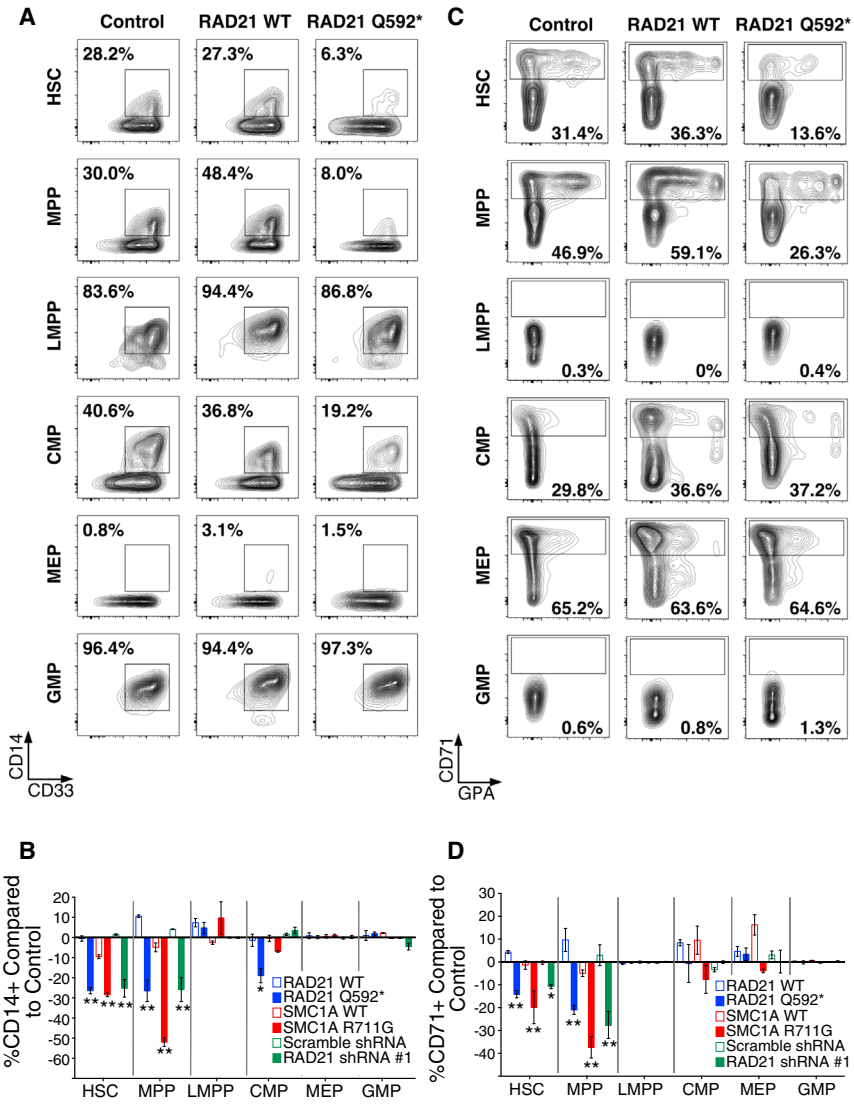


Figure 5. Cohesin Mutants Impair Human HSPC Differentiation in a Cell-Context-Dependent Manner

(A) Human CD34-enriched cord blood was sorted for HSPC subpopulations (HSCs, MPPs, LMPPs, CMPs, MEPs, and GMPs) (See Figure S5 for sorting strategy). Each subpopulation was then infected with lentiviruses encoding GFP alone (control) or GFP in addition to the indicated cohesin variants. The virally transduced cells were cultured in myeloid differentiation medium, as in Figure 2. 8 days later, the cells were analyzed by flow cytometry for expression of myeloid markers CD33 and CD14. Representative FACS plots of GFP+ cells are shown.

(B) Summary of data from three independent experiments as described in (A); the percentage of CD14+ cells was normalized to GFP control. Scramble shRNA and RAD21 shRNA transduced HSPC subpopulations (Figure S5B) are included in the summary. Unpaired Student's t test was used to determine statistical significance between WT and mutant populations. *$p < 0.05$ and **$p < 0.01$.

(C) HSPC subpopulations isolated and lentivirally transduced as in (A) were cultured in erythroid differentiation medium, as in Figure 2. 8 days later, the cells were analyzed for expression of erythroid markers CD71 and GPA. Representative FACS plots of GFP+ cells are shown.

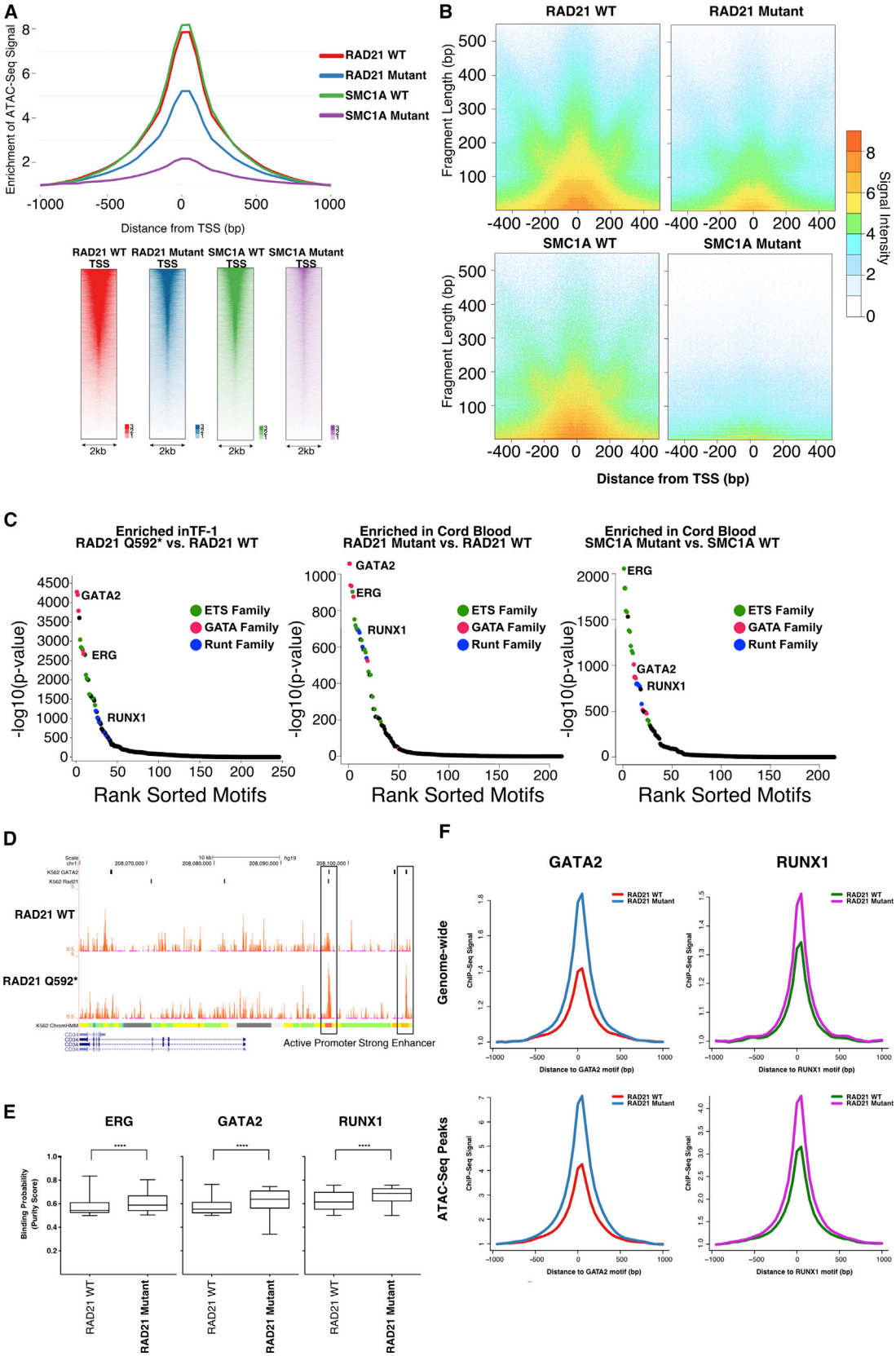
(D) Summary of data from three independent experiments as described in (C); the percentage of CD71+ cells was normalized to GFP control. Scramble shRNA and RAD21 shRNA transduced HSPC subpopulations (Figure S5C) are included in the summary. Unpaired Student's t test was used to determine statistical significance between cohesin WT and mutant populations. **$p < 0.01$.

Both cohesin WT and mutant-expressing TF-1 cell lines (RAD21 Q592* and RAD21 WT) and primary human cord blood transduced with cohesin WT or mutant lentivirus were subject to ATAC-seq. Two biological replicates in both cohesin WT and mutant variants were highly correlated to each other (Figure S6A), indicating strong reproducibility of this technique on isolated cord blood cells.

To determine whether cohesin mutants had an effect on chromatin accessibility, we first performed an unbiased analysis of ATAC-seq signal intensity near transcription start sites (TSSs) using all unique-mapped and properly paired ATAC-seq tags. A decrease of the average intensity of ATAC-seq signal was observed in RAD21 and SMC1A mutants (Figure 6A left panel) and was further supported by less intensity over all TSS regions in the cohesin mutant compared to WT samples (Figure 6A right panel). Next, we further examined the nucleosome positioning status surrounding the TSSs using a V-plot analysis (Figure 6B). In this approach, the mid-points of all fragments were centered on the x axis relative to TSSs, and fragment sizes were arranged from the smallest to the largest on the y axis. In the cohesin WT, nucleosome positioning leads to a pattern of enriched nucleosome-free fragments (<100 bp) and mono-nucleosome fragments (180–247 bp) surrounding TSSs (Figure 6B). However, the enrichment of nucleosome-free and mono-nucleosome fragments ("open-chromatin" indicators) was remarkably reduced in the cohesin mutants (Figure 6B), suggesting that appropriate nucleosome positioning with an open-chromatin status around TSSs was disrupted due to mutant RAD21 and SMC1A. In addition to the chromatin accessibility near TSSs, we further investigated chromatin accessibility globally using a functional genomic annotation based on 15 chromHMM states (Ernst and Kellis, 2012) derived from ENCODE histone markers on the K562 chronic myeloid leukemia cell line. The cohesin mutants exhibited less accessibility in active transcription regions, such as promoters and enhancers, as well as in insulator sites where CTCF is bound (Figure S6B).

Cohesin Mutants Exhibit Enrichment for Accessible HSPC TF Motifs

While cohesin mutants exhibited global reduced chromatin accessibility at transcriptional regulatory elements, increased



(legend on next page)

accessibility was detected at some sites. To characterize chromatin regions with differential accessibility in cohesin WT and cohesion-mutant-expressing cells, we performed motif searching of the gained and lost open accessible peaks with cohesin disruption (1,940 gained peaks and 1,551 lost peaks in RAD21 mutant; 5,230 gained peaks and 4,662 lost peaks in SMC1A mutants; $p < 0.01$ and fold change > 2). Motif analysis of more highly accessible sites in cohesin mutant samples identified a strong enrichment for TF binding sites including ERG, GATA2, and RUNX1, which are known to be highly expressed and critical regulators of HSPCs (Wilson et al., 2010) (Figure 6C and Table S3). RNA expression analysis using qRT-PCR demonstrated that cohesion-mutant-transduced cells exhibited increased expression of these three TFs compared to WT controls and did so at levels similar to those of immature CD34+CD38− cells freshly isolated from cord blood (Figure S6C). For instance at the *CD34* locus, the *CD34* promoter and upstream enhancer show higher accessibility in RAD21 mutant cells compared to WT; this enhancer is known to be bound by GATA2 in K562 cells (Figure 6D). Collectively, these results support a model in which mutant cohesin impairs hematopoietic differentiation and enforces stem cell programs through the modulation of TF chromatin accessibility.

Cohesin Mutants Exhibit Increased Occupancy of Select HSPC TFs

To corroborate whether increased accessibility at ERG, GATA2, and RUNX1 motifs indicated a higher likelihood of binding of these TFs, we performed footprinting analysis using protein interaction quantification (PIQ) software (Sherwood et al., 2014). The output of PIQ is the probability of occupancy for each candidate-binding site in the genome. PIQ analysis revealed a statistically significant increase in inferred binding for ERG, GATA2, and RUNX1 in RAD21 mutant cells compared to WT (Figure 6E), indicating that these TFs may be playing a significant role in enforcing stem cell programs in cohesin mutants.

To conclusively determine if there was increased binding of these TFs in the presence of mutant cohesin, we conducted ChIP-seq of GATA2 and RUNX1 in the TF-1 cells expressing RAD21 Q592* or RAD21 WT. After we normalized for sequencing depth, this analysis showed that both TFs had a stronger signal across GATA2 or RUNX1 motif instances genome-wide in RAD21 mutant cells compared to WT (Figure 6F). In addition, GATA2 and RUNX1 occupancy was enriched at ATAC-seq peaks identified in HSPCs (Figure 6F) and further enriched at ATAC-seq peaks upregulated in RAD21 mutant HSPCs (Figure S6D). Moreover, genes known to be involved in hematopoietic stem cells (*TAL1*, *CD34*, and *GATA2*) showed stronger binding of GATA2 and RUNX1 (see Figure S6E for illustrative tracks). Consistent with these results, both GATA2 and RUNX1 exhibited increased occupancy at chromatin accessible sites that were selectively gained in TF-1 cells expressing RAD21 mutants compared to RAD21 WT (Figure S6F). Moreover, SMC3 ChIP-seq revealed loss of SMC3 occupancy in TF-1 cells expressing RAD21 Q592* compared to cells expressing RAD21 WT. (Figure S6F). Thus, both ATAC-seq and ChIP-seq directly demonstrate that mutant cohesin subunits act as dominant negatives and alter chromatin accessibility, leading to increased occupancy of HSPC TFs.

Cohesin-Mutant-Induced Stem Cell Programs Are Dependent on ERG, GATA2, and RUNX1

Based on this model in which the effects of mutant cohesin are mediated by TFs exhibiting increased chromatin accessibility and binding such as ERG, GATA2, and RUNX1, we hypothesized that knockdown of these TFs would prevent the maintenance in CD34 expression observed in the in vitro HSPC retention assay (Figure 2A). Two separate shRNA lentiviral constructs targeting these TFs were validated for their knockdown effects in FACS-purified transduced CD34+ cord blood cells (Figure 7A). GATA1 and PU.1 were also targeted as control TFs expressed in HSPCs and not demonstrating increased chromatin accessibility in the ATAC-seq assays. We first confirmed TF knockdown effects, demonstrating impaired erythroid differentiation with knockdown of GATA1, while ERG, PU.1, and RUNX1 deficiency significantly reduced myeloid differentiation in in vitro culture assays (Figures S7A–S7C). CD34+ cord blood cells were then transduced with single or combination lentiviral vectors encoding a cohesin WT or mutant component (GFP+) and a TF-targeting shRNA (RFP+) (Figure 7B). As expected, expression of cohesin mutants with a scramble shRNA resulted in an increase in CD34-expressing cells; however, knockdown of ERG, GATA2, or RUNX1, but not GATA1 or PU.1, in the context of cohesin

Figure 6. Cohesin Mutants Exhibit Altered Chromatin Accessibility at Transcriptional Regulatory Elements, but Increased Accessibility and Binding at HSPC TF Motifs

(A) Human CD34-enriched HSPCs were infected with lentiviruses encoding cohesin WT or mutants in addition to GFP. 72 hr later, GFP+ cells were isolated by FACS and subject to ATAC-seq. Average diagram of genome-wide chromatin accessibility at TSS regions (2 kb window) comparing cohesin WT (RAD21 WT and SMC1A WT) and cohesin mutants (RAD21 mutant and SMC1A mutant) (left panel) is shown. Heat map of ATAC-seq signal intensity at all TSS regions (2 kb window) in cohesin WT and mutants (right panel) is also shown.

(B) V-plot analysis of ATAC-seq fragments near TSSs (1 kb window) in cohesin WT (RAD21 WT and SMC1A WT) and cohesin mutants (RAD21 mutant and SMC1A mutant). The x axis represents the distance between the centers of the fragments to the TSSs. The y axis represents the fragment length. The color (scaled from 0 to 9) represents the intensity of the ATAC-seq signal at the coordinate of this xy plane.

(C) Enrichment of TF motifs in the peaks that gain open accessibility in RAD21 mutant-expressing TF1 cells (left panel), RAD21 mutant-expressing CD34+ cord blood cells (middle panel), and SMC1A mutant-expressing CD34+ cord blood cells (right panel). The y axis is $-\log_{10}$(p value) of a motif enrichment test, which is sorted from largest to smallest. The x axis is the sorted motif rank. The ETS, GATA, and Runt families are indicated by green, red, and blue, respectively.

(D) Illustrative UCSC genome browser track of normalized ATAC-seq signal at the CD34 locus. K562 ENCODE ChIP-seq data for GATA2 and RAD21 are overlayed, as well as K562 ChromHMM indicating a higher signal at an active promoter and strong enhancer of CD34.

(E) PIQ footprinting analysis for ERG, GATA2, and RUNX1 indicates a higher likelihood of TF occupancy in RAD21 mutant HSPCs compared to RAD21 WT. ****$p < 0.0001$.

(F) TF-1 RAD21 WT and RAD21 Q592* were induced with DOX for 6 days and then subjected to ChIP-seq for GATA2 and RUNX1. Average diagrams for ChIP-seq signal (values normalized to input control) at GATA2 and RUNX1 motif are shown here. The top row indicates all genome-wide GATA2 motifs, and the bottom row indicates all GATA2 motifs in ATAC-seq peaks (from above).

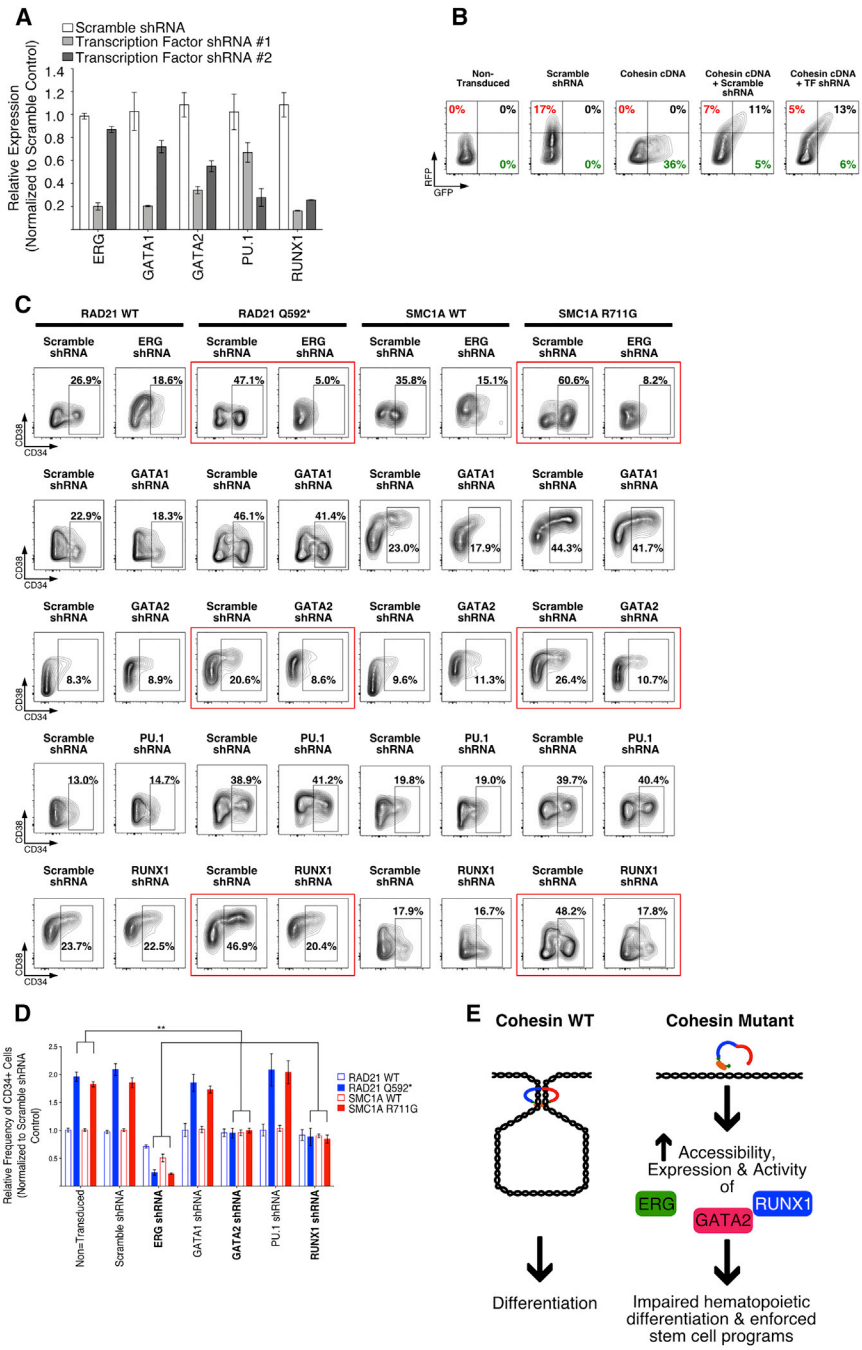


Figure 7. Cohesin-Mutant-Induced Stem Cell Programs Are Dependent on ERG, GATA2, and RUNX1

(A) TF expression was determined by qRT-PCR to determine the knockdown efficiency of two independent shRNAs for each of the indicated TFs in FACS-purified RFP+ CD34-enriched cord blood. The shRNA exhibiting the strongest knockdown for each TF was used for subsequent experiments.

(B) CD34-enriched cord blood HSPCs were double transduced with lentiviruses encoding scramble or TF-targeting shRNAs (RFP+, from A) and cohesin WT or mutants (GFP+). 6 days later, cells were analyzed for RFP and GFP expression by flow cytometry and representative plots are presented.

(C) HSPCs isolated and lentivirally transduced as in (B) were cultured in HSPC retention medium as in Figure 2. 6 days later, cells were analyzed for expression of progenitor markers CD34 and CD38. Representative FACS plots are shown. Key comparisons highlighting ERG, GATA2, and RUNX1 knockdown in the setting of mutant RAD21 or SMC1A are indicated in red boxes.

(D) Summary of data from three independent experiments as described in (C); the percentage of CD34+ cells was normalized to scramble shRNA only control. Unpaired Student's t test was used to determine statistical significance between cohesin WT and mutant populations. $^{**}p < 0.01$.

(E) Schematic model: mutant cohesin impairs hematopoietic differentiation and enforces stem cell programs through the modulation of ERG, GATA2, and RUNX1 chromatin accessibility, expression, and activity.

mutant expression completely prevented this increase in CD34-expressing cells (Figures 7C and 7D). These results are consistent with our proposed model that mutant cohesin impairs hematopoietic differentiation and enforces stem cell programs through the modulation of ERG, GATA2, and RUNX1 chromatin accessibility, expression, and activity (Figure 7E).

DISCUSSION

Recurrent mutations in the components of the cohesin complex have been identified in human AML and other myeloid malignancies and in particular have been shown to occur as pre-leukemic mutations in HSCs. Here, we sought to investigate the effects of mutant cohesin on hematopoiesis, including HSPCs. Introduction of mutant cohesin into AML cell lines and primary cord blood HSPC resulted in a differentiation block with an increased frequency of CD34+ cord blood progenitor cells. A similar phenotype was observed with knockdown of RAD21 both in vitro and in vivo, indicating that mutant cohesin can act through either haploinsufficiency or dominant-negative mechanisms. Mutant cohesin increased the serial replating ability of human HSPCs in vitro and showed enrichment for HSCs and leukemia stem cell gene expression programs, indicating an effect to enforce stem cell functions. Interestingly, the effect of mutant cohesin was found to be cell context dependent, in that these phenotypes could only be observed in the most immature HSPC populations. Using ATAC-seq, we determined that mutant cohesin lead to a state of elevated chromatin accessibility and higher predicted binding at TF binding sites for ERG, GATA2, and RUNX1. Moreover, using ChIP-seq, we formally demonstrated increased binding of GATA2 and

RUNX1 to these sites. Finally, we demonstrated that knockdown of these three TFs in human HSPCs can revert the differentiation block induced by mutant cohesin. These results support a model in which mutant cohesin impairs hematopoietic differentiation and enforces stem cell programs through the modulation of ERG, GATA2, and RUNX1 chromatin accessibility, expression, and activity (Figure 7E).

Our results indicate a link between the role of cohesin in the regulation of transcription through the modulation of chromatin accessibility and the maintenance of stem cell functions in HSPCs. Consistent with a previous report (Yan et al., 2013), we observed a global loss of open chromatin in cohesion-mutant-expressing cells, but an increase in accessibility at specific motifs for key hematopoietic TFs ERG, GATA2, and RUNX1 that in turn execute stem cell transcriptional programs and phenotypes. Several mechanisms are possible to explain these results. One model proposes that the global loss of open chromatin observed in cohesion-mutant-expressing cells enriches, either directly or indirectly, for the activity of pioneer factors that can bind chromatin in the absence of cohesin. Previous reports have suggested that GATA2 (Wang et al., 2007), RUNX1 (Lichtinger et al., 2012), and ERG (Chen et al., 2013b) can all function as pioneer factors in hematopoietic cells. This results in the enrichment of open motifs for these factors and a higher likelihood of occupancy in cohesion-mutant-expressing cells as detected here by ATAC-seq and ChIP-seq, with a decrease in accessibility for motifs of non-pioneer factors that promote hematopoietic differentiation. The result is impaired hematopoietic differentiation and enforcement of stem cell programs in HSPCs.

Another possibility is based on the role of cohesin in maintaining chromosome structure along with CTCF. Recent cohesin ChIA-PET studies in mouse ESCs have shown that many cell identity genes, including GATA2, are found within chromosome structures that are formed by the looping of two interacting CTCF sites co-occupied by cohesin (Dowen et al., 2014). These looped structures form insulated neighborhoods whose integrity is important for proper expression of local genes, and it has been suggested that loss of cohesin would lead to de-repression of these genes.

We examined the effect of cohesin mutants on hematopoietic differentiation in isolated subpopulations of human HSPCs and found that mutant cohesin imparts its differentiation block only in the most immature HSCs and MPPs. This cell context dependence may derive from the models discussed above in which cohesin is required to establish accessible chromatin for hematopoietic differentiation. However, once the differentiation program has been established, for example in GMPs for myeloid differentiation, cohesin is not required to maintain it, and loss of cohesin function does not revert these cells to a stem cell-like state or impair differentiation. This may explain our results demonstrating that cohesin mutations are often detected in pre-leukemic HSCs in human AML, as they must be acquired in HSCs to exert their leukemia-promoting effects (Corces-Zimmerman et al., 2014; Jan et al., 2012).

Our studies of pre-leukemic HSCs in AML demonstrated that pre-leukemic mutations preferentially occur in genes involved in regulating the epigenome including DNA methylation (*DNMT3A*, *TET2*, and *IDH1/2*), histone modifications (*ASXL1*), and cohesin (Corces-Zimmerman et al., 2014). A number of studies investigating the effects of these mutations on hematopoietic functions have demonstrated similar phenotypes with impaired differentiation and maintenance of stem cell programs, as reported here for cohesin mutations (Challen et al., 2012; Moran-Crusio et al., 2011; Sasaki et al., 2012). Moreover, we observed skewed differentiation toward the myeloid lineage both in vitro and in vivo, as was also detected with *TET2* (Pronier et al., 2011) and *IDH1/2* mutations (Chen et al., 2013a; Sasaki et al., 2012). Collectively, these studies support our model in which pre-leukemic mutations occur in genes that regulate the epigenome resulting in impaired hematopoietic differentiation and enforcement of stem cell programs, followed by downstream mutations in drivers of cell proliferation (*FLT3* and *RAS*), to eventually result in AML (Corces-Zimmerman and Majeti, 2014).

Interestingly, cohesin mutations have been identified in multiple cancer types, including Ewing sarcoma (Solomon et al., 2011), colorectal carcinoma (Barber et al., 2008), and bladder carcinoma (Balbás-Martínez et al., 2013; Solomon et al., 2013). Conflicting reports (Balbás-Martínez et al., 2013; Solomon et al., 2011) have reported that cohesin mutations may or may not be associated with aneuploidy in bladder cancer, implicating defects in its canonical role in maintaining sister chromatid cohesion during mitosis as a mechanism for malignant transformation. Notably, in most AML cases, cohesin mutations are not associated with karyotypic abnormalities (Cancer Genome Atlas Research Network, 2013), suggesting that defects in chromatid cohesion do not contribute to leukemogenesis. Ultimately, it will be important to determine if cohesin mutations impart the same phenotype observed in hematopoietic cells in these other tissues and cellular contexts.

Cohesin mutants impair hematopoietic differentiation and enforce stem cell programs in primary human HSPCs, and they occur as pre-leukemic and early mutations in human AML and myeloid malignancies. These observations indicate that cohesin mutants are excellent targets for novel therapeutics with the potential to change disease outcomes. Understanding the mechanisms by which these mutants act will be critical to the identification and design of such novel targeted agents, but the observation here that the effect of cohesin mutants on impaired hematopoietic differentiation is reversible suggests that it may be possible and beneficial to target mutant cohesin in AML and other myeloid malignancies.

EXPERIMENTAL PROCEDURES

AML Cell Lines

THP-1 and TF-1 cells were acquired from ATCC. THP-1 cells were cultured in RPMI-1640 medium (Life Technologies) with 10% Tet-approved FBS (Clontech) and PenStrep (Life Technologies). TF-1 cells were cultured in the same medium supplemented with GM-CSF (Peprotech) at 2 ng/ml.

Primary Human Samples

Human AML samples were obtained from patients at the Stanford Medical Center with informed consent, according to Institutional Review Board (IRB)-approved protocols (Stanford IRB no. 18329 and no. 6453). Individual case information is presented in Table S1. Cord blood samples were obtained from patients at the Lucile Packard Children's Hospital with informed consent, according to IRB-approved protocols (Stanford IRB no. 5637) and from New York Blood Center. Mononuclear cells from each sample were isolated by Ficoll (GE Healthcare Dharmacon) separation and cryopreserved in liquid nitrogen.

Lentivirus Transduction

TF-1 and THP-1 cell lines were transduced as previously described (Chan et al., 2015). Primary human CD34-enriched cord blood was transduced as follows: wells in a 96-well non-tissue-culture-treated plate were coated with 100 μl of RetroNectin per well (Clontech Laboratories) at 100 μg/ml in PBS for 2 hr at room temperature (RT). The wells were then blocked with 150 μl of a 2% (w/v) solution of BSA in PBS for 30 min at RT. The BSA solution was then aspirated, and lentiviral particles in HBSS with 25 mM HEPES were added to each well at a volume of 80–200 μl. The plate was then centrifuged at 3,700 r.p.m. for 2 hr at RT. After the supernatant was aspirated, 1,000–200,000 cells were added to each well in 150–200 μl of medium. The plate was centrifuged again at 1,300 r.p.m. for 10 min at RT and then transferred to a 37°C incubator to initiate lentiviral infection.

Generation of Inducible Cell Lines

TF-1 and THP-1 cells were transduced with lentiviral particles made using the DOX-inducible pINDUCER21 vector modified to express either WT or mutant cohesin along with constitutive expression of an eGFP marker. The cells were then sorted for eGFP expression using a FACSAria II (Becton Dickinson). Approximate heterozygous gene expression was induced with DOX at 2 μg/ml and replenished every 48 hr.

TF-1 EPO Differentiation Assay

TF-1 cells were induced with DOX for 48 hr and then differentiation was induced with EPO as previously described (Wang et al., 2013). Surface expression of GPA was analyzed using anti-human GPA-PE (CD235a) (HIR2, BD Biosciences). Expression of HBG1/2 and KLF-1 was analyzed by qRT-PCR.

For the DOX washout experiment, cells were washed five times in PBS and then resuspended in medium supplemented with EPO for an additional 4 days. The cells were then analyzed for surface expression of GPA and expression of HBG1/2 and KLF-1 as described above.

HSPC Differentiation Assays

Primary human CD34-enriched cord blood was transduced with lentivirus (pLVX EF1α-IRES-zsGreen; Clontech) overnight and sorted for GFP expression using a FACSAria II (Becton Dickinson). For progenitor media assay, cells were washed and incubated in HPGM medium (Lonza) supplemented with FLT3L, thrombopoietin, and stem cell factor (SCF). The cytokines were all purchased from Peprotech and used at 20 ng/ml. 6 days post-culture, differentiation of GFP+ cells was assessed by flow cytometry using anti-human CD34-APC (8G12) and CD38-PE-Cy7 (HB7) (BD Biosciences). For the myeloid differentiation assay, cells were cultured in Myelocult H5100 (Stem Cell Technologies) with 20 μg/ml each of IL-3, SCF, FLT3L, and GM-CSF (Peprotech) for 6 days. Myeloid differentiation was assessed by flow cytometry using anti-human CD33-PE (WM53) and CD14-APC-Cy7 (MΦP9) (BD Biosciences). For the erythroid differentiation assay, cells were cultured in Stemspan II with erythroid expansion kit (StemCell Technologies, Inc.) for 6 days. Erythroid differentiation was assessed by flow cytometry using anti-human GPA-PE (CD235a) (HIR2) and CD71-PE-Cy7 (OKT9) (BD Biosciences).

Human Xenografts and Assessment of Human Engraftment in NSG Mice

Transplantation of cord blood cells into NSG mice and engraftment of human cells within mouse bone marrow was assessed by immunophenotyping as previously described (Chan et al., 2015).

Methylcellulose Colony Assays

HSPCs were lentivirally transduced and isolated as described above (see HSPC Differentiation Assays) and then cultured for 14 days in Methocult H4435 Enriched methylcellulose (Stem Cell Technologies) according to standard protocols. For replating, methylcellulose was dissolved with the addition of DMEM at RT, and 1,000-10,000 cells were counted and replated in the same medium.

Expression Microarray Analysis

RNA was extracted from GFP+ sorted HSPC-enriched cord blood that was transduced with control vector (pLVX-IRES-zs-GREEN), cohesin WT, or cohesin mutant vectors. RNA was processed and analyzed using the Human Genome U133 Plus 2.0 Array (Affymetrix) according to the manufacturer's protocol. Array data are deposited in Gene Expression Omnibus database (accession number GEO: GSE73224).

For data analysis, CEL files were analyzed using GenePattern (Broad Institute) according to current protocols. Pearson correlation values were determined from all genes that passed normalization and filters. Gene Set Enrichment Analysis (GSEA) was conducted according to current protocols.

ChIP-Seq

ChIP-seq was performed as described previously (Calo et al., 2015). Briefly, TF-1 RAD21 Q592* and TF-1 RAD21 WT cells were induced with DOX for 6 days, after which ~100,000,000 cells were crosslinked in 1% formaldehyde for 10 min at RT. Typically, 5 million cells were used for each ChIP experiment and chromatin was sheared on a Bioruptor (Diagenode) to an average of 200–400 bp. After specific recovery of chromatin-bound proteins via immunoprecipitation, samples were reverse crosslinked at 65°C overnight and DNA was purified for subsequent library preparation. ChIP-seq libraries were prepared according to the NEBNext protocol and sequenced using Illumina NextSeq. The following antibodies were used for ChIP studies: SMC3 (clone: ab9263, Abcam), GATA2 (clone: sc-9008, Santa Cruz Biotechnology), and RUNX1 (clone: sc-365644, Santa Cruz Biotechnology). Sequencing data are deposited in GEO database (accession number GEO: GSE73207).

ATAC-Seq

TF-1 RAD21 Q592* and TF-1 RAD21 WT were induced with DOX for 6 days, and then 50,000 cells were directly transposed as previously described (Buenrostro et al., 2013). Cord blood HSPCs were transduced with control vector, cohesin WT, or cohesin mutant vectors and sorted after 3 days for GFP+ cells. Sorted cells (ranging from 1,000–20,000) were immediately transposed as previously described (Buenrostro et al., 2013). Sequencing data are deposited in GEO database (accession number GEO: GSE73206).

Statistical Analysis

Unless otherwise stated, p values comparing two means were calculated using the two-tailed unpaired Student's t test in Prism version 6 (GraphPad Software, Inc.). $p < 0.05$ was considered statistically significant. All bar graphs show mean with ± SEM.

ACCESSION NUMBERS

The accession number for the microarray studies reported in this paper is GSE73224; for the ATAC-Seq studies, it is GEO: GSE73206; and for the ChIP-Seq studies, it is GEO: GSE73207.

SUPPLEMENTAL INFORMATION

Supplemental Information for this article includes Supplemental Experimental Procedures, seven figures and three tables and can be found with this article online at http://dx.doi.org/10.1016/j.stem.2015.09.017.

AUTHOR CONTRIBUTIONS

C.M. and R.M. designed experiments. C.M., S.X., A.R., R.L., F.Z., R.A.F., M.R.C., S.M.C., D.T., J.L.K., and W.-J.H. performed the experiments. C.M., Y.S., J.D.B., H.Y.C., and R.M. analyzed and interpreted the data. C.M. and R.M. wrote the manuscript.

ACKNOWLEDGMENTS

We acknowledge the Hematology Division Tissue Bank and the patients for donating their samples. C.M. is supported by the California Institute of Regenerative Medicine Stanford Training Program (TG2-01159). This research was supported by grants from the New York Stem Cell Foundation (R.M.), Stinehart-Reed Foundation (R.M.), Ludwig Institute (R.M.), National Institutes of Health Grant R01CA188055 (R.M.), Stanford Cancer Center (H.Y.C.), and National Institutes of Health Grant P50-HG007735 (H.Y.C.). R.M. is a New York Stem Cell Foundation Robertson Investigator.

Received: April 20, 2015
Revised: August 24, 2015
Accepted: September 23, 2015
Published: October 22, 2015

REFERENCES

Balbás-Martínez, C., Sagrera, A., Carrillo-de-Santa-Pau, E., Earl, J., Márquez, M., Vazquez, M., Lapi, E., Castro-Giner, F., Beltran, S., Bayés, M., et al. (2013). Recurrent inactivation of STAG2 in bladder cancer is not associated with aneuploidy. Nat. Genet. *45*, 1464–1469.

Barber, T.D., McManus, K., Yuen, K.W.Y., Reis, M., Parmigiani, G., Shen, D., Barrett, I., Nouhi, Y., Spencer, F., Markowitz, S., et al. (2008). Chromatid cohesion defects may underlie chromosome instability in human colorectal cancers. Proc. Natl. Acad. Sci. USA *105*, 3443–3448.

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat. Methods *10*, 1213–1218.

Calo, E., Flynn, R.A., Martin, L., Spitale, R.C., Chang, H.Y., and Wysocka, J. (2015). RNA helicase DDX21 coordinates transcription and ribosomal RNA processing. Nature *518*, 249–253.

Cancer Genome Atlas Research Network (2013). Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N. Engl. J. Med. *368*, 2059–2074.

Challen, G.A., Sun, D., Jeong, M., Luo, M., Jelinek, J., Berg, J.S., Bock, C., Vasanthakumar, A., Gu, H., Xi, Y., et al. (2012). Dnmt3a is essential for hematopoietic stem cell differentiation. Nat. Genet. *44*, 23–31.

Chan, S.M., Thomas, D., Corces-Zimmerman, M.R., Xavy, S., Rastogi, S., Hong, W.J., Zhao, F., Medeiros, B.C., Tyvoll, D.A., and Majeti, R. (2015). Isocitrate dehydrogenase 1 and 2 mutations induce BCL-2 dependence in acute myeloid leukemia. Nat. Med. *21*, 178–184.

Chen, C., Liu, Y., Lu, C., Cross, J.R., Morris, J.P., 4th, Shroff, A.S., Ward, P.S., Bradner, J.E., Thompson, C., and Lowe, S.W. (2013a). Cancer-associated IDH2 mutants drive an acute myeloid leukemia that is susceptible to Brd4 inhibition. Genes Dev. *27*, 1974–1985.

Chen, Y., Chi, P., Rockowitz, S., Iaquinta, P.J., Shamu, T., Shukla, S., Gao, D., Sirota, I., Carver, B.S., Wongvipat, J., et al. (2013b). ETS factors reprogram the androgen receptor cistrome and prime prostate tumorigenesis in response to PTEN loss. Nat. Med. *19*, 1023–1029.

Corces-Zimmerman, M.R., and Majeti, R. (2014). Pre-leukemic evolution of hematopoietic stem cells: the importance of early mutations in leukemogenesis. Leukemia *28*, 2276–2282.

Corces-Zimmerman, M.R., Hong, W.-J., Weissman, I.L., Medeiros, B.C., and Majeti, R. (2014). Preleukemic mutations in human acute myeloid leukemia affect epigenetic regulators and persist in remission. Proc. Natl. Acad. Sci. USA *111*, 2548–2553.

Diaz-Martinez, L.A., and Clarke, D.J. (2009). Chromosome cohesion and the spindle checkpoint. Cell Cycle *8*, 2733–2740.

Dowen, J.M., Fan, Z.P., Hnisz, D., Ren, G., Abraham, B.J., Zhang, L.N., Weintraub, A.S., Schuijers, J., Lee, T.I., Zhao, K., and Young, R.A. (2014). Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. Cell *159*, 374–387.

Eppert, K., Takenaka, K., Lechman, E.R., Waldron, L., Nilsson, B., van Galen, P., Metzeler, K.H., Poeppl, A., Ling, V., Beyene, J., et al. (2011). Stem cell gene expression programs influence clinical outcome in human leukemia. Nat. Med. *17*, 1086–1093.

Ernst, J., and Kellis, M. (2012). ChromHMM: automating chromatin-state discovery and characterization. Nat. Methods *9*, 215–216.

Estey, E., and Döhner, H. (2006). Acute myeloid leukaemia. Lancet *368*, 1894–1907.

Gentles, A.J., Plevritis, S.K., Majeti, R., and Alizadeh, A.A. (2010). Association of a leukemic stem cell gene expression signature with clinical outcomes in acute myeloid leukemia. JAMA *304*, 2706–2715.

Goardon, N., Marchi, E., Atzberger, A., Quek, L., Schuh, A., Soneji, S., Woll, P., Mead, A., Alford, K.A., Rout, R., et al. (2011). Coexistence of LMPP-like and GMP-like leukemia stem cells in acute myeloid leukemia. Cancer Cell *19*, 138–152.

Jan, M., and Majeti, R. (2013). Clonal evolution of acute leukemia genomes. Oncogene *32*, 135–140.

Jan, M., Snyder, T.M., Corces-Zimmerman, M.R., Vyas, P., Weissman, I.L., Quake, S.R., and Majeti, R. (2012). Clonal evolution of preleukemic hematopoietic stem cells precedes human acute myeloid leukemia. Science Translational Medicine *4*, 149ra118–149ra118.

Kon, A., Shih, L.-Y., Minamino, M., Sanada, M., Shiraishi, Y., Nagata, Y., Yoshida, K., Okuno, Y., Bando, M., Nakato, R., et al. (2013). Recurrent mutations in multiple components of the cohesin complex in myeloid neoplasms. Nat. Genet. *45*, 1232–1237.

Leeke, B., Marsman, J., Sullivan, J.M.O., and Horsfield, J.A. (2014). Cohesin mutations in myeloid malignancies: underlying mechanisms. Exp. Hem. Onc. *3*, 13.

Lichtinger, M., Ingram, R., Hannah, R., Müller, D., Clarke, D., Assi, S.A., Lie-A-Ling, M., Noailles, L., Vijayabaskar, M.S., Wu, M., et al. (2012). RUNX1 reshapes the epigenetic landscape at the onset of haematopoiesis. EMBO J. *31*, 4318–4333.

Löwenberg, B., Downing, J.R., and Burnett, A. (1999). Acute myeloid leukemia. N. Engl. J. Med. *341*, 1051–1062.

Majeti, R., Park, C.Y., and Weissman, I.L. (2007). Identification of a hierarchy of multipotent hematopoietic progenitors in human cord blood. Cell Stem Cell *1*, 635–645.

Moran-Crusio, K., Reavie, L., Shih, A., Abdel-Wahab, O., Ndiaye-Lobry, D., Lobry, C., Figueroa, M.E., Vasanthakumar, A., Patel, J., Zhao, X., et al. (2011). Tet2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. Cancer Cell *20*, 11–24.

Panigrahi, A.K., and Pati, D. (2012). Higher-order orchestration of hematopoiesis: is cohesin a new player? Exp. Hematol. *40*, 967–973.

Pronier, E., Almire, C., Mokrani, H., Vasanthakumar, A., Simon, A., da Costa Reis Monte Mor, B., Massé, A., Le Couédic, J.P., Pendino, F., Carbonne, B., et al. (2011). Inhibition of TET2-mediated conversion of 5-methylcytosine to 5-hydroxymethylcytosine disturbs erythroid and granulomonocytic differentiation of human hematopoietic progenitors. Blood *118*, 2551–2555.

Sasaki, M., Knobbe, C.B., Munger, J.C., Lind, E.F., Brenner, D., Brüstle, A., Harris, I.S., Holmes, R., Wakeham, A., Haight, J., et al. (2012). IDH1(R132H) mutation increases murine haematopoietic progenitors and alters epigenetics. Nature *488*, 656–659.

Sherwood, R.I., Hashimoto, T., O'Donnell, C.W., Lewis, S., Barkal, A.A., van Hoff, J.P., Karun, V., Jaakkola, T., and Gifford, D.K. (2014). Discovery of directional and nondirectional pioneer transcription factors by modeling DNase profile magnitude and shape. Nat. Biotechnol. *32*, 171–178.

Sinha, S., Thomas, D., Yu, L., Gentles, A.J., Jung, N., Corces-Zimmerman, M.R., Chan, S.M., Reinisch, A., Feinberg, A.P., Dill, D.L., and Majeti, R. (2015). Mutant WT1 is associated with DNA hypermethylation of PRC2 targets in AML and responds to EZH2 inhibition. Blood *125*, 316–326.

Solomon, D.A., Kim, T., Diaz-Martinez, L.A., Fair, J., Elkahloun, A.G., Harris, B.T., Toretsky, J.A., Rosenberg, S.A., Shukla, N., Ladanyi, M., et al. (2011). Mutational inactivation of STAG2 causes aneuploidy in human cancer. Science *333*, 1039–1043.

Solomon, D.A., Kim, J.-S., Bondaruk, J., Shariat, S.F., Wang, Z.-F., Elkahloun, A.G., Ozawa, T., Gerard, J., Zhuang, D., Zhang, S., et al. (2013). Frequent truncating mutations of STAG2 in bladder cancer. Nat. Genet. *45*, 1428–1430.

Spencer, D.H., Young, M.A., Lamprecht, T.L., Helton, N.M., Fulton, R., Laughlin, M.O.A., Fronick, C., Magrini, V., Demeter, R.T., Miller, C.A., et al. (2015). Epigenomic analysis of the HOX gene loci reveals mechanisms that may control canonical expression patterns in AML and normal hematopoietic cells. Leukemia *29*, 1279–1289.

Wang, Q., Li, W., Liu, X.S., Carroll, J.S., Jänne, O.A., Keeton, E.K., Chinnaiyan, A.M., Pienta, K.J., and Brown, M. (2007). A hierarchical network of transcription factors governs androgen receptor-dependent prostate cancer growth. Mol. Cell *27*, 380–392.

Wang, F., Travins, J., DeLaBarre, B., Penard-Lacronique, V., Schalm, S., Hansen, E., Straley, K., Kernytsky, A., Liu, W., Gliser, C., et al. (2013). Targeted inhibition of mutant IDH2 in leukemia cells induces cellular differentiation. Science *340*, 622–626.

Wilson, N.K., Foster, S.D., Wang, X., Knezevic, K., Schütte, J., Kaimakis, P., Chilarska, P.M., Kinston, S., Ouwehand, W.H., Dzierzak, E., et al. (2010). Combinatorial transcriptional control in blood stem/progenitor cells: genome-wide analysis of ten major transcriptional regulators. Cell Stem Cell *7*, 532–544.

Yan, J., Enge, M., Whitington, T., Dave, K., Liu, J., Sur, I., Schmierer, B., Jolma, A., Kivioja, T., Taipale, M., and Taipale, J. (2013). Transcription factor binding in human cells occurs in dense clusters formed around cohesin anchor sites. Cell *154*, 801–813.

Cell Stem Cell

Supplemental Information

Leukemia-Associated Cohesin Mutants Dominantly Enforce Stem Cell Programs and Impair Human Hematopoietic Progenitor Differentiation

Claire Mazumdar, Ying Shen, Seethu Xavy, Feifei Zhao, Andreas Reinisch, Rui Li, M. Ryan Corces-Zimmerman, Ryan A. Flynn, Jason D. Buenrostro, Steven M. Chan, Daniel Thomas, Julie L. Koenig, Wan-Jen Hong, Howard Y. Chang, and Ravindra Majeti

SUPPLEMENTARY EXPERIMENTAL PROCEDURES

$CD34^+$ HSPC enrichment from cord blood.

$CD34^+$ HSPCs were enriched from freshly processed cord blood samples by magnetic separation using CD34 microbeads (Miltenyi Biotech) per manufacturer's protocol and cultured in HPGM medium (Lonza) supplemented with FLT3 ligand, thrombopoietin, and stem cell factor. The cytokines were all purchased from Peprotech and used at 20 ng/ml final concentration.

Construction of cDNA expression lentiviral vectors.

The complementary DNA (cDNA) clones of human RAD21 (Accession: BC001229), SMC3 (Accession: BC047324) and STAG2 (Accession: BC017095) were purchased from GE Healthcare Dharmacon in the pCMV-SPORT6 vector. SMC1A (Accession: BC112127) was purchased in pCR-XL-TOPO. The mutations presented in this paper (RAD21 E212*, RAD21 Q592*, SMC1A R711G, SMC3 G662C and STAG2 Q801*) were generated using the QuikChange II site-directed mutagenesis kit (Agilent Technologies) per manufacturer's instructions. The open reading frames of wild-type or mutant RAD21/SMC1A/SMC3/STAG2 were PCR amplified and cloned into the pLVX-EF1a-IRES-zsGREEN vector (Clontech) and doxycycline-inducible pINDUCER21 (IRES-eGFP) from the Elledge Lab (Addgene plasmid #46948) using the Infusion Cloning Kit per manufacturer's instructions (Clontech).

Construction of shRNA expression lentiviral vectors.

The human RAD21 (Accession: BC001229), GATA1 (Accession: NM_002049.3), GATA2 (Accession: NM_001145661.1), PU.1 (Accession: NM_001243998.1) and RUNX1 (Accession: NM_001001890.2) shRNA target sequences (**Table S4**) were selected using the BLOCK-iT RNAi Designer tool (Life Technologies). Cloning into pRSI9 DECIPHER shRNA expression vector (Cellecta) was conducted as previously described (Chan et al., 2015). Knockdown efficiency of shRNA constructs was determined using qRT-PCR using Taqman assays from Life Technology (see **Gene Expression Assays**). The scramble control sequence used in this study was 5′-GCACTACCAGAGCTAACTCAGATAGTACT-3′.

Lentivirus production.

The 293TN producer cell line (System Biosciences) was grown in Advanced DMEM (Life Technologies) with 10% FBS and 2 mM GlutaMAX. One day before transfection, 4 million cells were plated in a 150-mm tissue-culture dish. On the day of transfection, 13.5 μg of the lentiviral expression plasmid was combined with 8.8 μg of the packaging vector psPAX2 and 4.7 μg of the envelope expressing plasmid pCMV-VSV-G. The DNA mixture was diluted in Opti-MEM I medium (Life Technologies), mixed with 293fectin transfection reagent (Life Technologies) at a ratio of 3 μl per 1 μg of DNA and added to the plated 293TN cells. Viral supernatant was collected at 72 h after transfection, filtered through a 0.45-μm PVDF filter and concentrated via ultracentrifugation at 23,000rpm for 2 hours. The concentrated lentiviral particles were resuspended in HBSS with 25 mM HEPES and stored at −80 °C.

Gene Expression Assays.

Quantitative real-time PCR (qRT-PCR) was performed as previously described (Reinisch et al., 2015). Taqman gene expression assays were purchased from Life Technologies: RAD21 (Hs00366721_mH), SMC1A (Hs00196849_m1), SMC3 (Hs00271322_m1), STAG2 (Hs00198227_m1), GATA1 (Hs01085823_m1), GATA2 (Hs00231119_m1), GATA3 (Hs00231122_m1), PU.1 (Hs00162150_m1), ERG (Hs01554629_m1), RUNX1 (Hs01021970_m1), KLF-1 (Hs00610592_m1), HBG1/2 (Hs00361131_g1)

Cell Proliferation Assay

To assess the proliferation of cells we used Absolute Countbright beads (Life Technology) as per manufacturer protocol.

THP-1 Differentiation Assays

THP-1 cells were induced with DOX for 48 hours and then differentiation was induced with all-trans retinoic acid at 1 uM as previously described (Drach et al., 1993). Macrophage differentiation was monitored by surface expression of anti-human CD11b-PE-Cy5 (ICRF44) (BD Biosciences).

THP-1 cells were induced with DOX for 48 hours and then differentiation was induced with phorbol 12-myristate 13-acetate (PMA) at a concentration of 100 ng/mL for 48 hours as

previously described (Park et al., 2007). Macrophage differentiation was monitored by surface expression of anti-human CD11b-PE-Cy5 (ICRF44) (BD Biosciences).

Annexin V staining.

Staining for apoptotic cells was performed as previously described (Chan et al., 2015).

HSPC Subpopulation Sorting

A previously established panel of antibodies was used for analysis and sorting of hematopoietic stem and progenitor populations (Majeti et al., 2007). Briefly, mononuclear cells were isolated from freshly processed UCB by density gradient centrifugation (Ficoll, GE Healthcare). Thereafter, CD34+ cells were pre-enriched using MACS-technology (Miltenyi), washed, and incubated for 30 minutes, at 4ºC with monoclonal anti CD2 (RPA-2.10), CD3 (S4.1), CD4 (S3.5), CD7 (CD7-6B7); CD8 (3B5), CD11b (ICRF44), CD14 (TUK4), CD16 (3GA), CD19 (SJ25-C1), CD20 (13.6E12), CD56 (B159), GPA (GA-R2, all PE-Cy5, all BD), as well as CD45RA-BV605 (MEM56, ebioscience), CD10-APC-Cy7 (HI10a, biolegend), CD90-FITC (5E10), CD123-PE (7G3), CD38-PE-Cy7 (HIT2), CD34-APC (8G12, all BD) antibodies.

Western blotting.

Standard western blotting and co-immunoprecipitation techniques were performed. A rabbit monoclonal RAD21 antibody (clone D5Y8S; Cell Signaling) was used at a dilution of 1:1,000. A goat polyclonal antibody against SMC1A (Abcam) was used at a dilution of 1:1,000. A rabbit monoclonal antibody against SMC3 (clone D47B5, Cell Signaling) was used at a dilution of 1:2,000. A mouse monoclonal β-actin antibody (clone 8H10D10; Cell Signaling) was used at a dilution of 1:5,000. Standard secondary antibodies conjugated to HRP were used following incubation with primary antibodies. Cell lysis for western blotting and co-immunoprecipitation was conducted as previously described (Kon et al., 2013).

Intracellular flow cytometry staining.

Intracellular staining of TF-1 cells was conducted per Cell Signaling's online protocol using a phospho-gamma-H2AX antibody (Clone 20E3) conjugated to Alexa-647 (Cell Signaling) at a dilution of 1:100. A rabbit IgG isotype control antibody conjugated to Alexa Fluor-647 was used to determine background nonspecific fluorescence staining.

ATAC-Seq Analysis

Data preprocessing

ATAC-Seq paired end reads were trimmed for Illumina adapter sequences and transposase sequences using an in-house script and mapped to hg19 using Bowtie v0.12.9 (Langmead et al., 2009) with parameters –S –X2000 –m1. Duplicate reads were discarded with Samtools v0.1.18 (Li et al., 2009). Peak calling using ZINBA was as described (Rashid et al., 2011). Chromosomal regions with a posterior probability of >0.99 were identified as peaks. Overlapping peaks from all samples were merged together to a unique peak list, and number of raw reads mapped to each peak for each individual samples was quantified.

Differential Analysis and Motif Analysis of ATAC-Seq peak

Differentially accessible peaks from the union peak list were identified with edgeR (Robinson et al., 2009) using raw counts of each sample in the overlapping peak list. edgeR was run with default settings, with a fold change threshold of 2, and P-value < 0.01. Motif analysis was performed using HOMER (http://homer.salk.edu/homer/motif/) with default parameters to test the occurrence of a TF motif in peaks regions compared to that in background regions.

ATAC-Seq signal intensity around TSS

A 1kb window centered on TSS was divided into twenty 50 equal sized bins. The number of unique-mapped and properly paired ATAC-Seq tags overlapping each bin was counted. The average fragment count plotted in each bin was normalized to the average tag count in the first five bins in order to normalize the background signal among various samples. The heat maps of ATAC-Seq at all the TSS regions were generated using Java TreeView 3.0.

V-plot analysis for nucleosome positioning

V-plots for cohesin WT (RAD21WT, SMC1A WT) and cohesin mutants (RAD21 mutants, SMC1A mutants) were normalized by sequencing depth, which allows for direct visual comparison of nucleosome position near TSS region (1 kb window; 500 bp left and 500 bp right). The X-axis represents the distance between the mid-point of the fragments to TSS. The Y-axis represents the fragment length. The color represents the intensity of ATAC-Seq (the number of fragments at the coordinates).

PIQ Footprinting Analysis

The genome-wide motif footprinting analysis was performed using PIQ v1.3 (Sherwood et al., 2014) with input motif position weight matrices (PWM) from jasper database (http://jaspar.genereg.net/). For footprinting, we adjusted the read start sites to represent the center of the transposon binding event. Previous studies of the Tn5 transposase have shown that the transposon binds as a dimer and inserts two adaptors separated by 9 bp (Adey et al., 2010). Therefore, all reads mapped to the forward strand were offset by +4 bp, and all reads mapped to the reverse strand were offset -5 bp. The purity scores (indicating the likelihood of true TF binding) from PIQ algorithm were compared between cohesin WT and mutant samples by paired t-test.

ChIP-Seq Analysis

ChIP-seq single end reads were mapped to hg19 using Bowtie2 v2.1.0 (Langmead and Salzberg, 2012) with–very sensitive option. Peaks were identified using MACS2 (Zhang et al., 2008) with default parameters. The differential binding of transcription factors between ChIP-Seq samples were identified using MAnorm (Shao et al., 2012). The average diagram ChIP-Seq signal (GATA2 and RUNX1) of ChIP-Seq signal around 1kb of the TF motif in 50 bp resolution was plotted using an in-house script.

shRNA Primers	
RAD21 shRNA #1 –Fwd	ACCGGCAGCTTATAATGCCATTATCAAGAGTAATGGCATTATAAGCTGCTTTT
RAD21 shRNA #1 –Rev	CGAAAAAAGCAGCTTATAATGCCATTACTCTTGATAATGGCATTATAAGCTGC
RAD21 shRNA #2 –Fwd	ACCGGCCACTGCCTGACTTAGATTCAAGAGATCTAAGTCAGGCAGTGGCTTTT
RAD21 shRNA #2 –Rev	CGAAAAAAGCCACTGCCTGACTTAGATCTCTTGAATCTAAGTCAGGCAGTGGC
GATA1 shRNA #1 –Fwd	ACCGGCCTCTATCACAAGATGAATGTCAAGAGCATTCATCTTGTGATAGAGGCTTTT
GATA1 shRNA #1 –Rev	CGAAAAAAGCCTCTATCACAAGATGAATGCTCTTGACATTCATCTTGTGATAGAGGC
GATA1 shRNA #2 –Fwd	ACCGGCGCCTGATTGTCAGTAAACGTCAAGAGCGTTTACTGACAATCAGGCGCTTTT
GATA1 shRNA #2 – Rev	CGAAAAAAGCGCCTGATTGTCAGTAAACGCTCTTGACGTTTACTGACAATCAGGCGC
GATA2 shRNA #1 –Fwd	ACCGGGAACCGGAAGATGTCCAACATCAAGAGTGTTGGACATCTTCCGGTTCCTTTT
GATA2 shRNA #1 – Rev	CGAAAAAAGGAACCGGAAGATGTCCAACACTCTTGATGTTGGACATCTTCCGGTTCC
GATA2 shRNA #2 –Fwd	ACCGGGTGGACGTCTTCTTCAATCATCAAGAGTGATTGAAGAAGACGTCCACCTTTT
GATA2 shRNA #2 – Rev	CGAAAAAAGGTGGACGTCTTCTTCAATCACTCTTGATGATTGAAGAAGACGTCCACC
PU.1 shRNA #1 – Fwd	ACCGGCTTCGCCGAGAACAACTTCATCAAGAGTGAAGTTGTTCTCGGCGAAGCTTTT
PU.1 shRNA #1 – Rev	CGAAAAAAGCTTCGCCGAGAACAACTTCACTCTTGATGAAGTTGTTCTCGGCGAAGC
PU.1 shRNA #2 – Fwd	ACCGGCAAGAAGATGACCTACCAGATCAAGAGTCTGGTAGGTCATCTTCTTGCTTTT
PU.1 shRNA #2 – Rev	CGAAAAAAGCAAGAAGATGACCTACCAGACTCTTGATCTGGTAGGTCATCTTCTTGC
RUNX1 shRNA #1 –Fwd	ACCGGGATACAAGGCAGATCCAACCTCAAGAGGGTTGGATCTGCCTTGTATCCTTTT
RUNX1 shRNA #1 – Rev	CGAAAAAAGGATACAAGGCAGATCCAACCCTCTTGAGGTTGGATCTGCCTTGTATCC
RUNX1 shRNA #2 –Fwd	ACCGGGCTGAGCTGAGAAATGCTACTCAAGAGGTAGCATTTCTCAGCTCAGCCTTTT
RUNX1 shRNA #2 – Rev	CGAAAAAAGGCTGAGCTGAGAAATGCTACCTCTTGAGTAGCATTTCTCAGCTCAGCC
ERG shRNA #1 – Fwd	ACCGGCACTATTAAGGAAGCCTTATTCAAGAGATAAGGCTTCCTTAATAGTGCTTTT
ERG shRNA #1 – Rev	CGAAAAAAGCACTATTAAGGAAGCCTTATCTCTTGAATAAGGCTTCCTTAATAGTGC
ERG shRNA #2 – Fwd	ACCGGGAGTGGGCGGTGAAAGAATATCAAGAGTATTCTTTCACCGCCCACTCCTTTT
ERG shRNA #2 – Rev	CGAAAAAAGGAGTGGGCGGTGAAAGAATACTCTTGATATTCTTTCACCGCCCACTCC

SUPPLEMENTARY FIGURE AND TABLE LEGENDS

Figure S1. Cohesin Mutants Exhibit no Changes in Proliferation, Apoptosis or DNA Damage, Related to Figure 1.

(A) TF-1 cells were infected with lentiviruses encoding doxycycline (DOX)-inducible cohesin WT or mutant variants and green fluorescent protein (GFP). After 6 days of DOX induction, expression of cohesin complex genes (RAD21, SMC1A, SMC3, and STAG2) was determined by qRT-PCR. Values normalized to TF-1 No DOX control.

(B) Whole cell lysate Western blot analysis of cohesin component expression from TF-1 cell lines after 6 days of DOX induction. Anti-RAD21, anti-SMC1A, anti-SMC1A, and anti-STAG2 antibodies were used, and beta-actin was monitored as a protein loading control. Cohesin mutant samples are bolded.

(C) THP-1 cells were infected with lentiviruses encoding doxycycline (DOX)-inducible cohesin WT or mutant variants and green fluorescent protein (GFP). Myeloid differentiation of parental THP-1 cells and variants was determined by flow cytometry for CD11b expression after 2 initial days of DOX treatment and 4 days of 1uM all trans retinoic acid (ATRA). Relative expression shown as mean fluorescence intensity (MFI) of CD11b normalized to IgG isotype control. * indicates $p<0.05$ and ** indicates $p<0.01$.

(D) THP-1 lentivirally transduced as in (C) were induced to differentiate with 100 ng/mL phorbol myristate acetate (PMA). Myeloid differentiation of parental THP-1 cells and variants was determined by flow cytometry for CD11b expression after 2 initial days of DOX treatment and 4 days of PMA. Relative expression shown as mean fluorescence intensity (MFI) of CD11b normalized to IgG isotype control. ** indicates $p<0.01$.

(E) TF-1 cells were infected with lentiviruses encoding doxycycline (DOX)-inducible cohesin WT or mutant variants and green fluorescent protein (GFP). Proliferation by absolute cell count (trypan blue exclusion) after 2 initial days of DOX treatment was monitored every 4 days for a total of 20 days. No statistically significant differences were detected.

(F) TF-1 cells treated as in (E) were analyzed after 8 days in culture following DOX induction for Annexin V positivity and propidium iodide (PI) staining by flow cytometry. % Annexin V (-) cells is shown. No statistically significant differences were detected.

(G) TF-1 cells treated as in (E) were cultured and then fixed and permeabilized for phospho-gamma H2AX intracellular staining after the 3 indicated time points. Relative expression is shown as mean fluorescence intensity (MFI) of phospho-gamma H2AX normalized to IgG isotype control. TF-1 cells treated with 1uM etoposide was used as a positive control for phospho-gamma H2AX staining. No statistically significant differences were detected.

Figure S2. Cohesin Mutants do not Affect Proliferation and Apoptosis of Primary Human HSPC , Related to Figure 2.

(A) Human CD34-enriched cord blood HSPC were infected with lentiviruses encoding GFP alone (control) or GFP in addition to the indicated cohesin variants. 72 hours post-infection, GFP+ cells were isolated by FACS and expression of cohesin complex genes (RAD21, SMC1A, SMC3, and STAG2) was determined by qRT-PCR. Values normalized to empty vector control.

(B) HSPC treated as in (A) were cultured in HSPC-retention medium. Proliferation was monitored using absolute countbright beads by flow cytometry. The proliferation rate was calculated by normalizing to non-transduced control. No statistically significant differences were detected.

(C) HSPC treated as in (A) were analyzed after 4 days for Annexin V positivity and propidium iodide (PI) staining by flow cytometry. %Annexin V (-) cells is shown. No statistically significant differences were detected.

(D) HSPC were infected with lentiviruses encoding GFP alone (control) or GFP in addition to the indicated cohesin variants. 72 hours post-infection, GFP(-) cells were isolated by FACS and cultured in HSPC-retention medium. 6 days later, cells were analyzed for expression of progenitor markers CD34 and CD38. Representative FACS plots are shown. No statistically significant differences were detected.

(E) HSPC isolated and lentivirally transduced as in (D) were cultured in myeloid differentiation medium. 6 days later, cells were analyzed for expression of myeloid markers CD33 and CD14. Representative FACS plots are shown. No statistically significant differences were detected.

(F) HSPC isolated and lentivirally transduced as in (D) were cultured in erythroid differentiation medium. 6 days later, cells were analyzed for expression of erythroid markers CD71 and GPA. Representative FACS plots are shown. No statistically significant differences were detected.

Figure S3. Knockdown of RAD21 Impairs Myeloid, Erythroid, and Stem Cell Differentiation of Primary Human HSPC, Related to Figure 3.

(A) RAD21 expression was determined by qRT-PCR to determine the knockdown efficiency of 3 RAD21 shRNA vectors including 1 inducible vector (#3) and 2 constitutive vectors (#1 and #2) in FACS-purified RFP+ CD34-enriched cord blood. Expression is reported relative to a scrambled shRNA control.

(B) RAD21 protein expression was determined by Western blot in TF-1 cells transduced with the same 3 RAD21 shRNA vectors. Cells transduced with inducible vector #3 were analyzed in the absence and presence of DOX. Beta-actin was monitored as a protein loading control.

(C) TF-1 cells were infected with lentiviruses encoding doxycycline (DOX)-inducible RAD21 shRNA or constitutive RAD21 shRNAs and FACS-sorted for GFP+ or RFP+ cells. Erythroid differentiation of the resulting cell lines was determined by qRT-PCR for fetal hemoglobin expression after 2 initial days of DOX treatment and 8 days of EPO and DOX treatment. Values normalized to TF-1 control. ** indicates $p<0.01$.

(D) Expression of KLF-1 was determined by qRT-PCR for cells treated as in (C). Values normalized to TF-1 controls. ** indicates $p<0.01$.

(E) Human CD34-enriched cord blood HSPC were infected with constitutive RAD21 shRNA or scramble control RFP-encoding lentiviral vectors. 72 hours post-infection, RFP+ cells were FACS-purified and cultured in HSPC-retention medium as in Figure 2. 6 days later, cells were analyzed for expression of progenitor markers CD34 and CD38. Representative FACS plots are shown.

(F) Summary of data from 3 independent experiments as described in (E); the percentage of CD34+ cells was normalized to scramble shRNA control. Unpaired Student t test was used to determine statistical significance between RAD21 WT and RAD21 knockdown and mutant-expressing populations. ** indicates $p<0.01$.

(G) HSPC isolated and lentivirally transduced as in (E) were cultured in myeloid differentiation medium. 6 days later, cells were analyzed for expression of myeloid markers CD33 and CD14. Representative FACS plots are shown.

(H) Summary of data from 3 independent experiments as described in (G); the percentage of CD33+/CD14+ cells was normalized to GFP control. Unpaired Student t test was used to determine statistical significance between WT and mutant populations. ** indicates $p<0.01$.

(I) HSPC isolated and lentivirally transduced as in (E) were cultured in erythroid differentiation medium. 6 days later, cells were analyzed for expression of erythroid markers CD71 and GPA. Representative FACS plots are shown.

(J) Summary of data from 3 independent experiments as described in (I); the percentage of GPA+/CD71+ cells was normalized to GFP control. Unpaired Student t test was used to determine statistical significance between WT and mutant populations. ** indicates $p<0.01$.

Figure S4. Primary Cohesin-Mutant AML Samples Exhibit Reduced Expression of Cohesin Components and Reduced Binding to Cohesin Partners, Related to Figure 4.

(A) Whole cell lysate Western blot analysis of cohesin component expression from 7 primary patient samples and 2 TF-1 cell line samples is shown. Anti-RAD21, anti-SMC3, and anti-SMC1A antibodies were used, and beta-actin was monitored as a protein loading control. Cohesin mutant samples are bolded.

(B) Co-immunoprecipitation analysis from samples shown in (A). Whole cell lysates were first immunoprecipitated using an anti-SMC1A antibody and then blotted with anti-RAD21, SMC3, or SMC1A antibodies.

Figure S5. Knockdown of RAD21 Impairs Human HSPC Differentiation in a Cell Context-Dependent Manner, Related to Figure 5.

(A) Representative FACS plots (pre-sort) and purity analysis (post-sort) for purification of HSPC subpopulations including: HSC, MPP, LMPP, CMP, GMP, and MEP.

(B) Human CD34-enriched cord blood was sorted for HSPC subpopulations (HSC, MPP, LMPP, CMP, MEP, and GMP). Each subpopulation was then transduced with

lentiviruses encoding scramble shRNA (RFP) or RAD21-shRNA (RFP). The virally transduced cells were cultured in myeloid differentiation medium, as in Figure 2. 8 days later, the cells were analyzed by flow cytometry for expression of myeloid markers CD33 and CD14. Representative FACS plots of RFP+ cells are shown of 3 independent experiments.

(C) HSPC subpopulations isolated and lentivirally transduced as in (B) were cultured in erythroid differentiation medium, as in Figure 2. 8 days later, the cells were analyzed for expression of erythroid markers CD71 and GPA. Representative FACS plots of RFP+ cells are shown of 3 independent experiments.

Figure S6. Cohesin Mutants Exhibit Decreased Chromatin Accessibility at K562 Promoter and Enhancer Features, Increased Expression of HSPC Transcription Factors, and Increased Binding of HSPC Transcription Factors at Open Chromatin Sites, Related to Figure 6.

(A) Human CD34-enriched HSPC were infected with control, cohesin WT, or cohesin mutant GFP-encoding lentiviruses. 3 days later, GFP+ transduced cells were FACS-purified and subject to ATAC-Seq. The heat map shows the Pearson correlation between every pair of the cohesin WT and mutant samples. A dendrogram from hierarchical clustering based on the Euclidean distance of the sample correlation vectors is shown symmetrically on the top and left sides of the heat map. The color bar (1 – correlation) on the left corner is ranged from 0 to 1.

(B) The heat map of ATAC-Seq signal intensity in 15 annotated genomic features. The genomic features are annotated based on ChromHMM states learned by histone markers of K562 cell line (ENCODE data). The ATAC-Seq signal intensity is indicated by the percentage of unique-mapped and properly-paired ATAC-Seq tags in each ChromHMM state. The data in each row of the heat map were standardized for contrast in visualization. The dendrogram on top of the heat map is based on a hierarchical clustering of genome-wide ATAC-Seq tag distribution of cohesin WT, cohesin mutants, normal CD34+ CD38-, and normal CD34+ CD38+ cells.

(C) HSPC isolated and lentivirally transduced as in (A) were subjected to qRT-PCR for expression of HSPC transcription factors (ERG, GATA2, and RUNX1). Freshly isolated

normal HSPC subpopulations (CD34+/CD38-, CD34+/CD38+, and Lineage +) were analyzed in parallel.

(D) TF-1 RAD21 WT and RAD21 Q592* were induced with DOX for 6 days and then subjected to ChIP-Seq for GATA2 and RUNX1. Average diagrams for ChIP-Seq signal at GATA2 and RUNX1 motifs shown here for all ATAC-Seq peaks differentially upregulated in RAD21-mutant HSPC.

(E) Illustrative tracks of ChIP-Seq signal (normalized to input) for GATA2, TAL1, and CD34 in TF-1 RAD21 WT vs. TF-1 RAD21 Q592* cells treated as in (D).

(F) Overlap of ChIP-Seq peaks of GATA2, RUNX1, or SMC3 with chromatin accessible sites that were gained or lost in RAD21 WT or RAD21 Q592* cells. Note that GATA2 and RUNX1 ChIP-seq peaks are enriched in accessible sites gained in cells expressing mutant RAD21.

Figure S7. Knockdown of GATA1 and PU.1 reduces erythroid and myeloid differentiation, respectively, Related to Figure 7.

(A) CD34-enriched cord blood HSPC were transduced with lentiviruses encoding scramble or transcription factor-targeting shRNAs and cultured in HSPC-retention medium. 6 days later, cells were analyzed for expression of progenitor markers CD34 and CD38. Representative FACS plots are shown of 3 independent experiments.

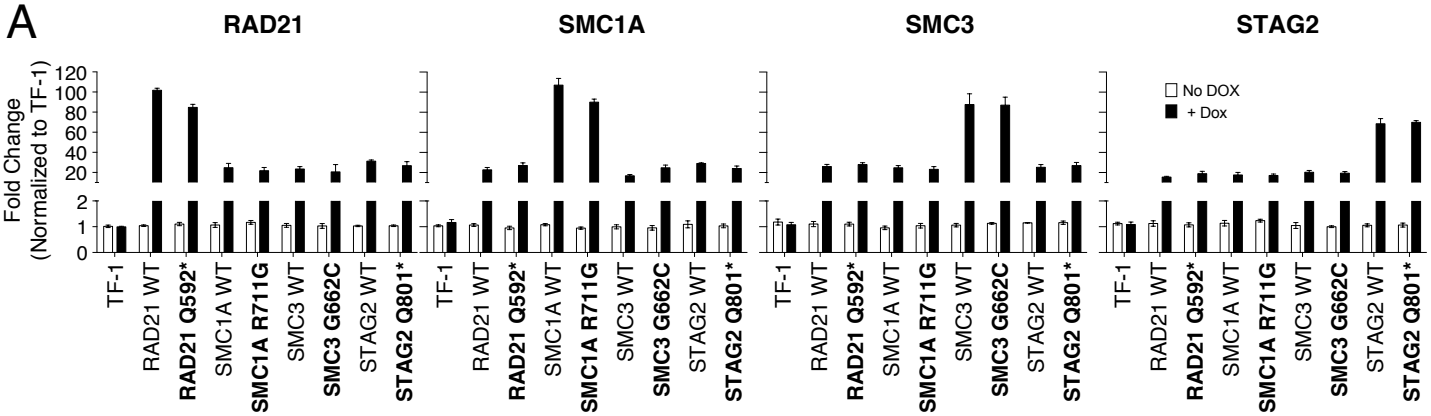
(B) HSPC isolated and lentivirally transduced as in (A) were cultured in erythroid differentiation medium. 6 days later, cells were analyzed for expression of erythroid markers CD71 and GPA. Representative FACS plots are shown of 3 independent experiments.

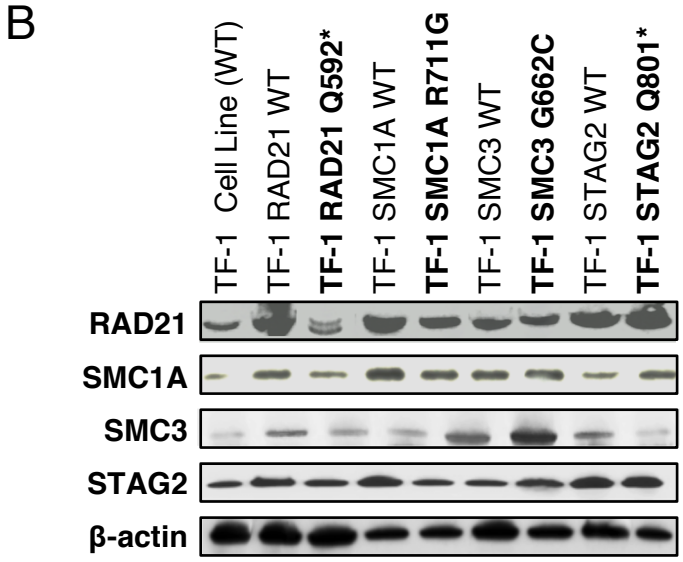
(C) HSPC isolated and lentivirally transduced as in (A) were cultured in myeloid differentiation medium. 6 days later, cells were analyzed for expression of myeloid markers CD33 and CD14. Representative FACS plots are shown of 3 independent experiments.

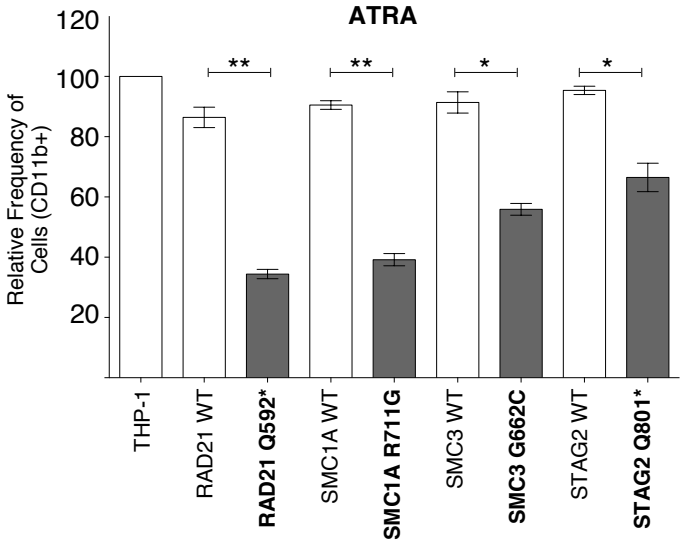
Table S1. Primary AML Sample Information, Related to Figure S4.

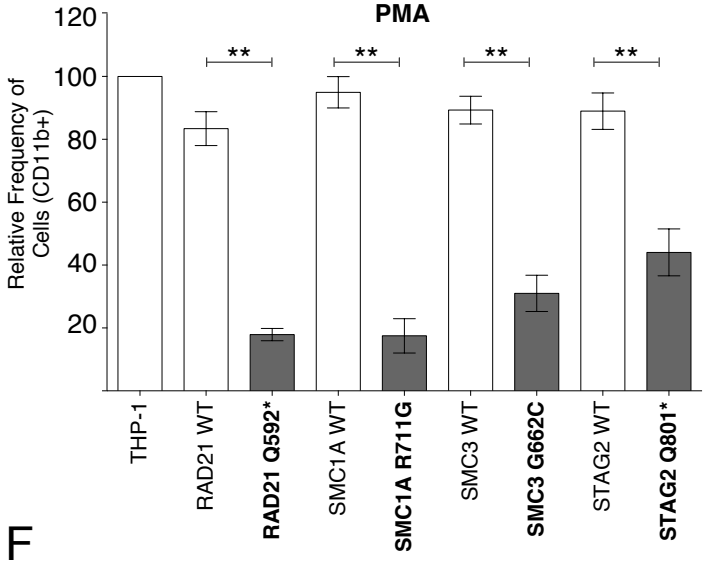
Table S2. Differentially Expressed Genes Between Cohesin Mutant and WT Cord Blood, Related to Figure 3.

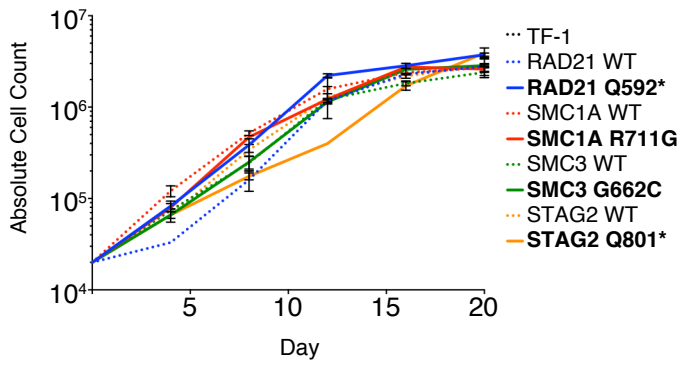
Table S3. Top 50 Motifs Upregulated in ATAC-Seq Data, Related to Figure 6.

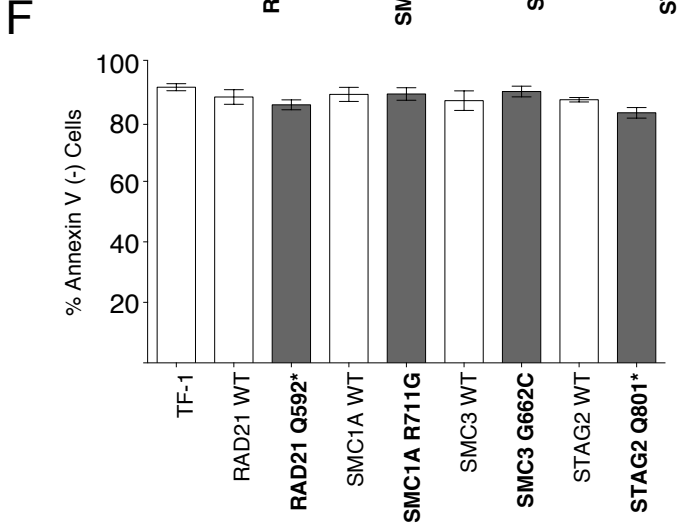


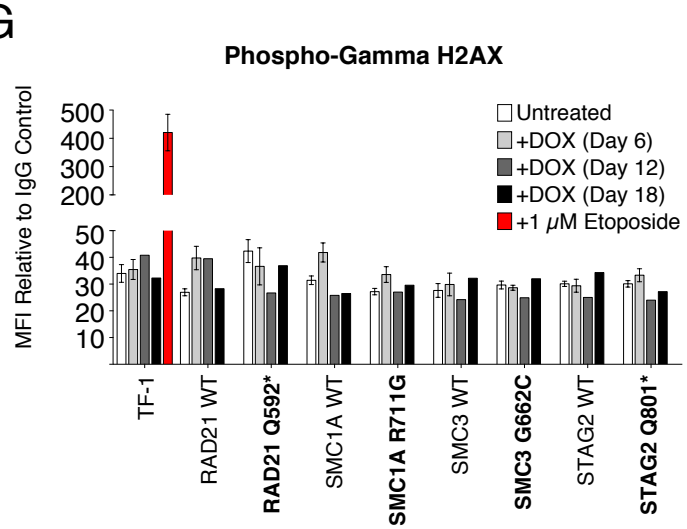




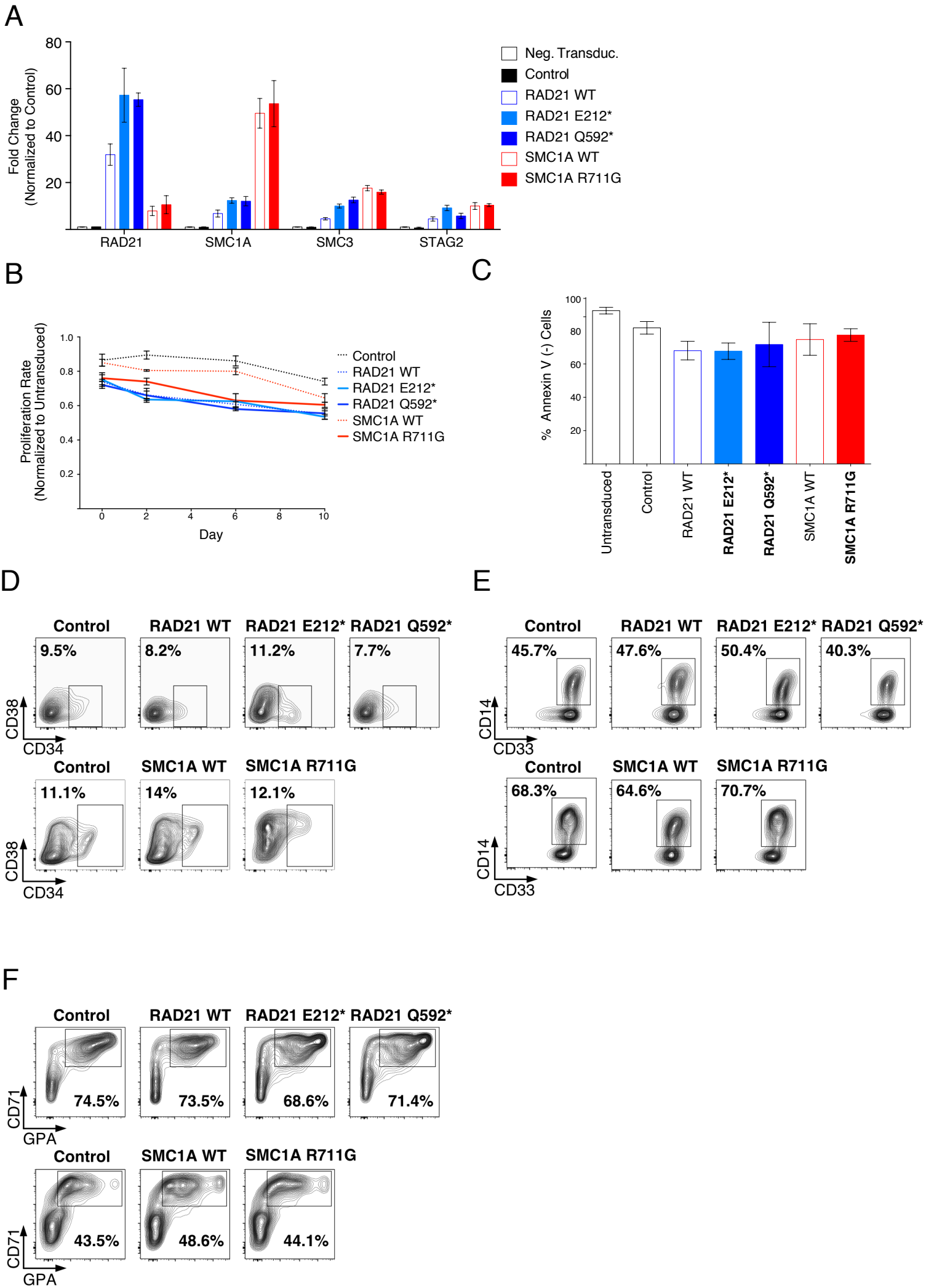




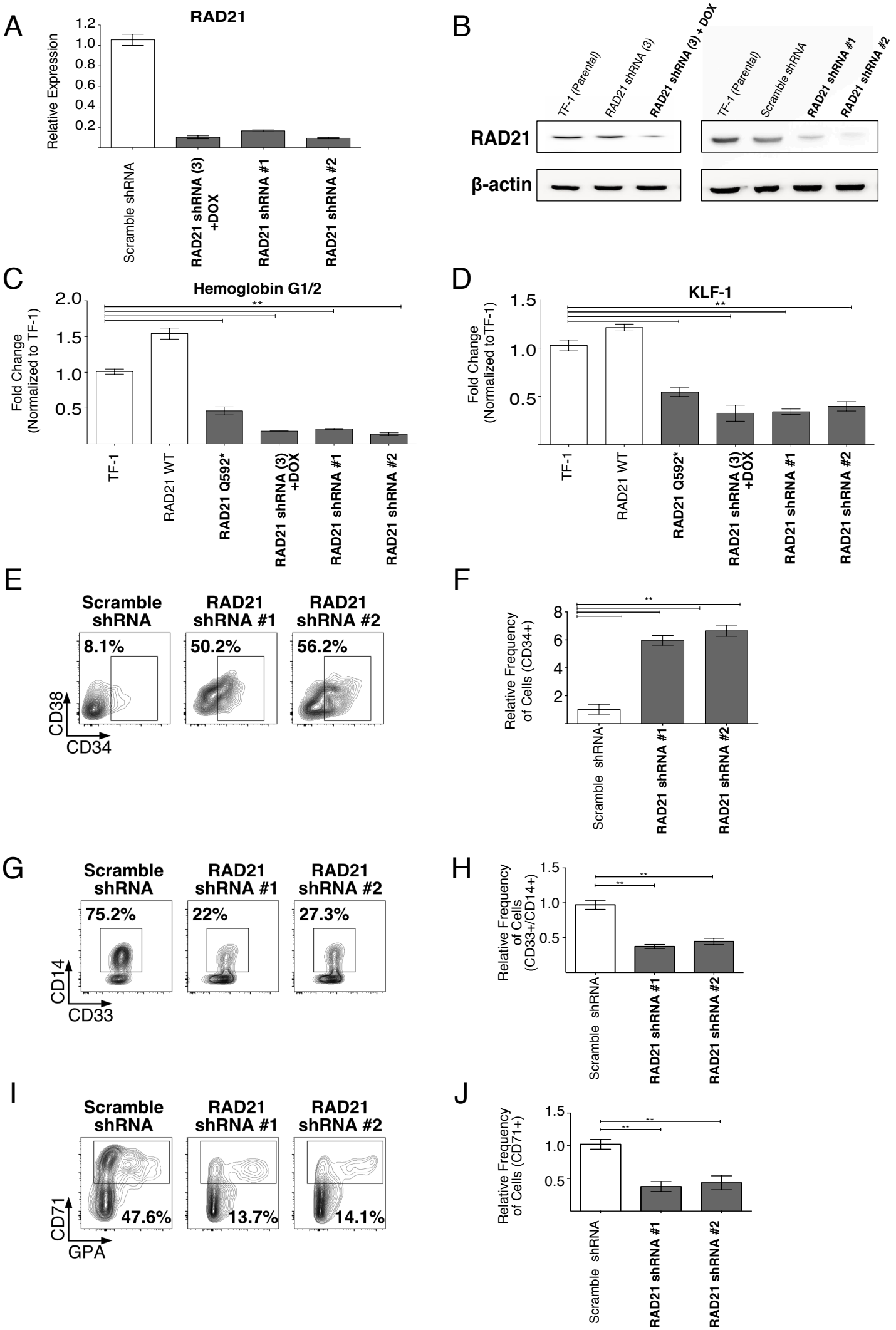




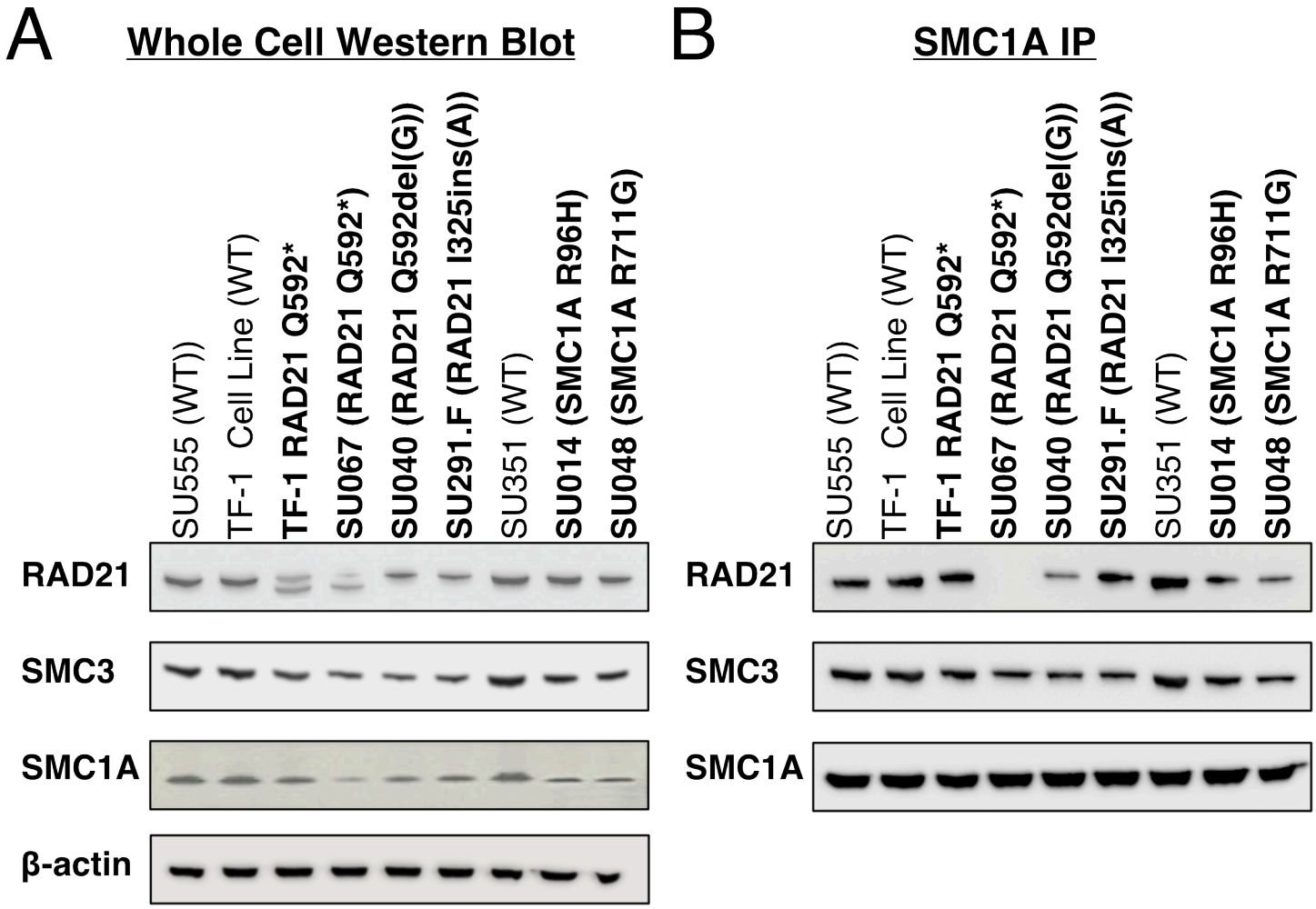
Supplementary Figure 1



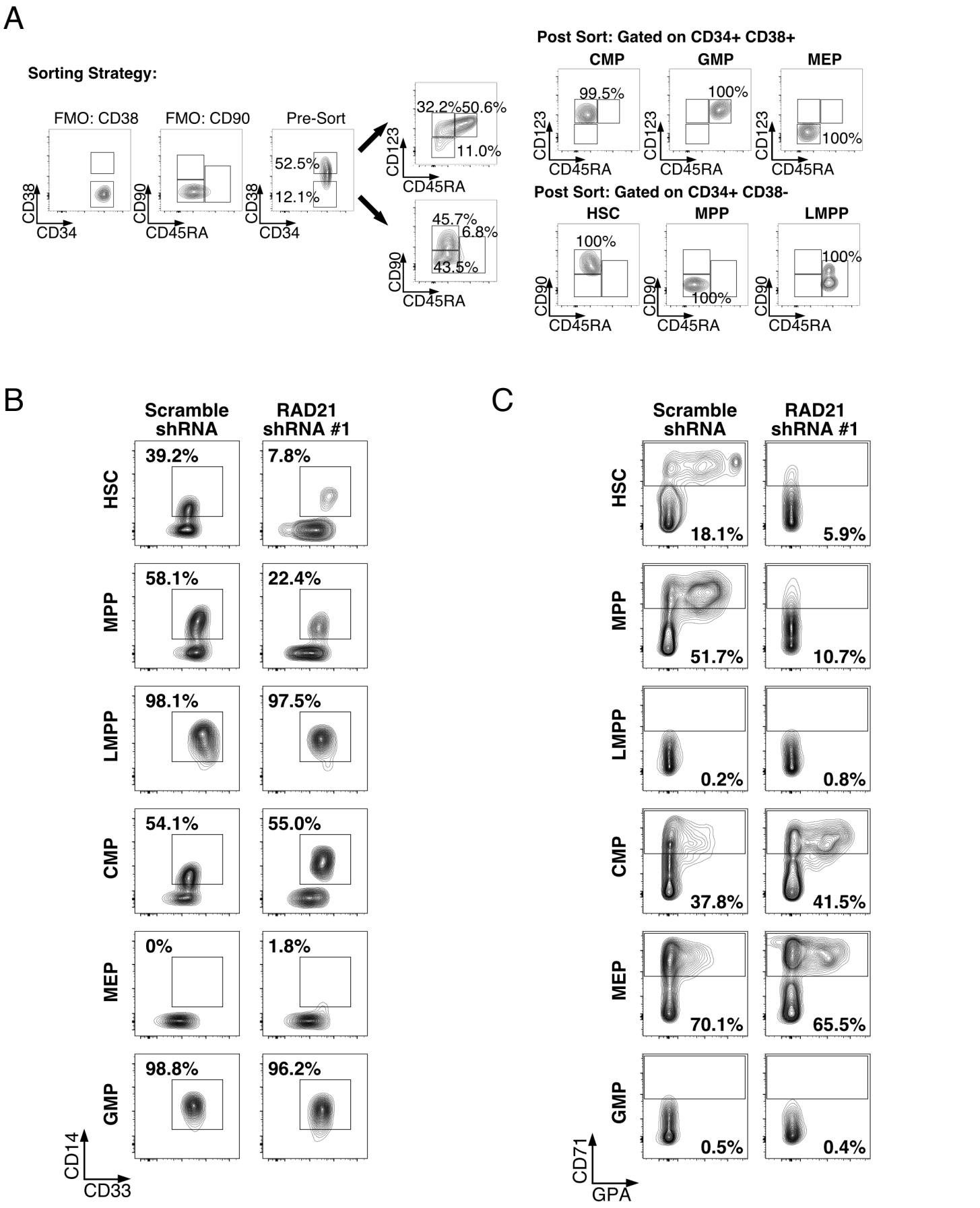
Supplementary Figure 2



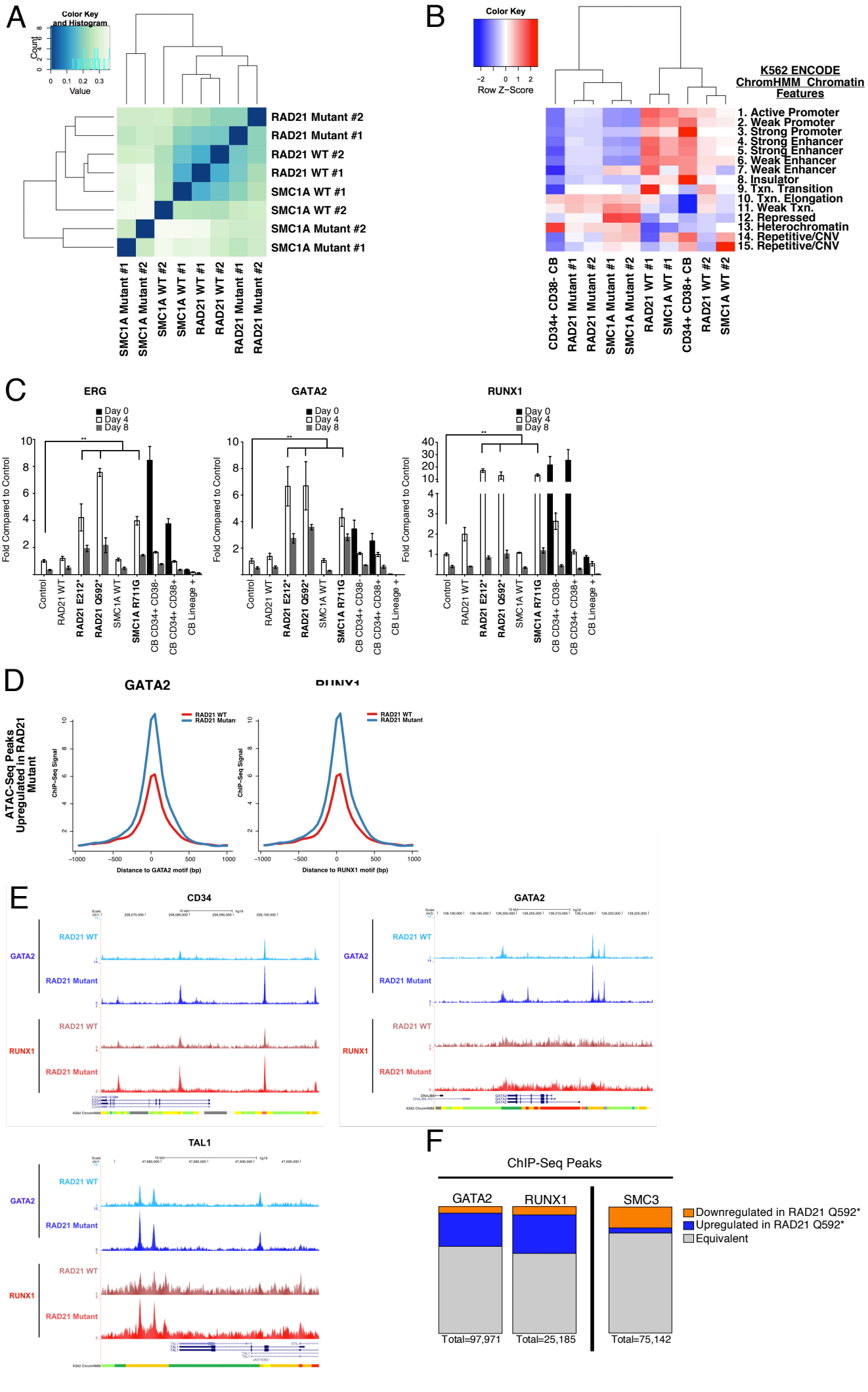
Supplementary Figure 3

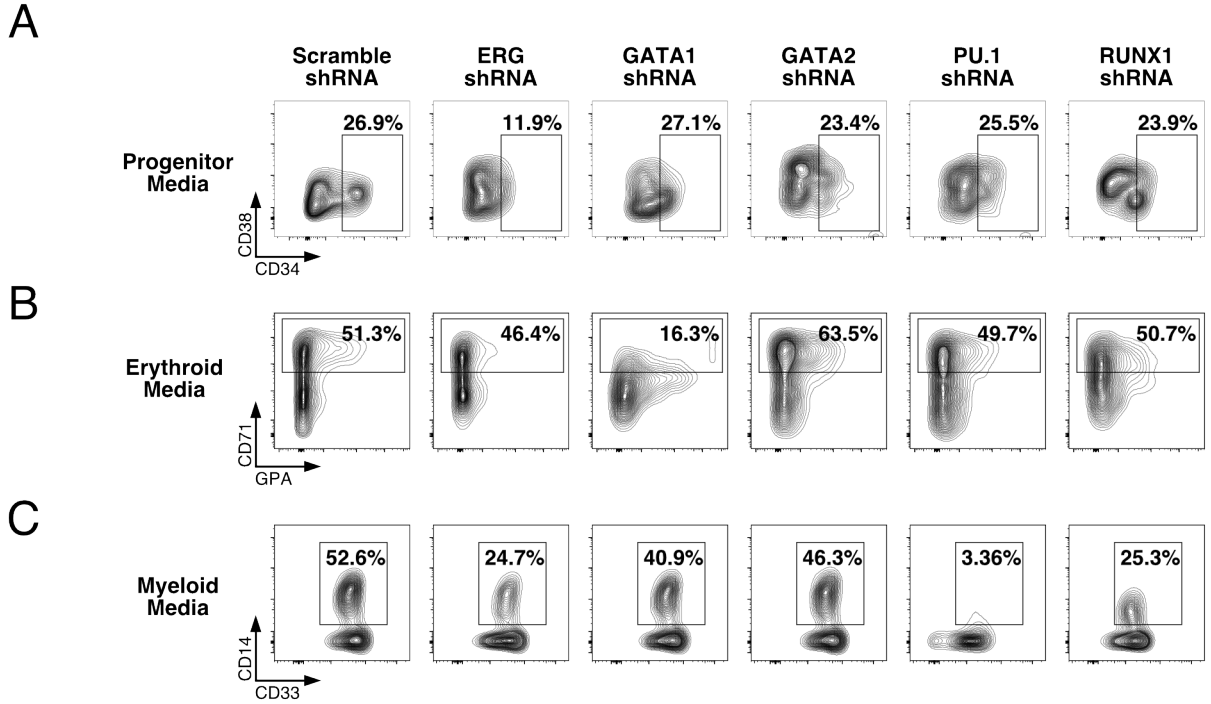


Supplementary Figure 4



Supplementary Figure 5


A
Color Key and Histogram
Count
Value
RAD21 Mutant #2
RAD21 Mutant #1
RAD21 WT #2
RAD21 WT #1
SMC1A WT #1
SMC1A WT #2
SMC1A Mutant #2
SMC1A Mutant #1
B
Color Key
Row Z-Score
K562 ENCODE ChromHMM Chromatin Features
1. Active Promoter
2. Weak Promoter
3. Strong Promoter
4. Strong Enhancer
5. Strong Enhancer
6. Weak Enhancer
7. Weak Enhancer
8. Insulator
9. Txn. Transition
10. Txn. Elongation
11. Weak Txn.
12. Repressed
13. Heterochromatin
14. Repetitive/CNV
15. Repetitive/CNV
CD34+ CD38- CB
RAD21 Mutant #1
RAD21 Mutant #2
SMC1A Mutant #1
SMC1A Mutant #2
RAD21 WT #1
SMC1A WT #1
CD34+ CD38+ CB
RAD21 WT #2
SMC1A WT #2
C
ERG
GATA2
RUNX1
Day 0
Day 4
Day 8
Fold Compared to Control
Control
RAD21 WT
RAD21 E212*
RAD21 Q592*
SMC1A WT
SMC1A R711G
CB CD34+ CD38-
CB CD34+ CD38+
CB Lineage +
D
GATA2
RUNX1
ATAC-Seq Peaks Upregulated in RAD21 Mutant
RAD21 WT
RAD21 Mutant
ChIP-Seq Signal
Distance to GATA2 motif (bp)
Distance to RUNX1 motif (bp)
E
CD34
GATA2
TAL1
RAD21 WT
RAD21 Mutant
GATA2
RUNX1
F
ChIP-Seq Peaks
GATA2
RUNX1
SMC3
Total=97,971
Total=25,185
Total=75,142
Downregulated in RAD21 Q592*
Upregulated in RAD21 Q592*
Equivalent



Supplementary Figure 7

Supplementary Table 1: Primary AML Sample Information							
Sample	**Disease Type**	**WHO Classification**	**Cytogenetics**	**Age**	**Gender**	**Cohesin Status**	**Other Mutations**
SU014	De Novo	AML not-otherwise specified	Normal	59	M	SMC1A Mutant (R96H)	IDH1-R132H, NPM1-L287ins(TCTG), ZBTB33-C496Y, FLT3-ITD
SU040	De Novo	AML not-otherwise specified	Normal	57	F	RAD21 Mutant (Q592del(G))	DNMT3A-R882H, BRWD3-R1447Q, MTCO1-A387T, FLT3-N841K, NPM1-L287ins(TCTG)
SU048	De Novo	AML not-otherwise specified	Normal	76	F	SMC1A Mutant (R711G)	TET2-E1357*, TET2-D1384V, NPM1-L287ins(TCTG), FLT3-ITD
SU067	De Novo	AML not-otherwise specified	Normal	54	F	RAD21 Mutant (Q592*)	PTPN11-A72T, NPM1-L287ins(TCTG)
SU291.F	Relapse	AML not-otherwise specified	+8	38	F	RAD21 Mutant (I325ins(A))	IDH1-R132H, FLT3-D835Y, SMG1-V2264I, , NPM1-L287ins(TCTG), DNMT3A-S535ins(21bp)
SU351	De Novo	AML with multilineage dysplasia without antecedent MDS	der(10)t(X;10),+8	75	M	Cohesin WT	IDH2-R140Q, NRAS-G13V, SI-R250H, BCORL1-G121R, SETD2-D2100H, SETD2-T2101del(A)
SU555	De Novo	AML not-otherwise specified	Normal	27	F	Cohesin WT	MTND6-T19I, LRP1B-I2264V, MTND4-L65P, PTPN11-G60V, SETD2-T2540ins(AA)

Supplementary Table 3: Top 50 Motifs Upregulated in ATAC-Seq Data			
	Upregulated in TF-1 RAD21 Q592* vs. RAD21 WT	**Upregulated in Human Cord Blood RAD21 Mutant vs. RAD21 WT**	**Upregulated in Human Cord Blood SMC1A Mutant vs. SMC1A WT**
1	Gata1(Zf)/K562-GATA1-ChIP-Seq(GSE18829)/Homer	Gata2(Zf)/K562-GATA2-ChIP-Seq/Homer	ERG(ETS)/VCaP-ERG-ChIP-Seq/Homer
2	Gata2(Zf)/K562-GATA2-ChIP-Seq(GSE18829)/Homer	Gata1(Zf)/K562-GATA1-ChIP-Seq/Homer	Fli1(ETS)/CD8-FLI-ChIP-Seq(GSE20898)/Homer
3	Gata4(Zf)/Heart-Gata4-ChIP-Seq(GSE35151)/Homer	Gata4(Zf)/Heart-Gata4-ChIP-Seq(GSE35151)/Homer	ETV1(ETS)/GIST48-ETV1-ChIP-Seq/Homer
4	CTCF(Zf)/CD4+-CTCF-ChIP-Seq(Barski et al.)/Homer	ERG(ETS)/VCaP-ERG-ChIP-Seq/Homer	ETS1(ETS)/Jurkat-ETS1-ChIP-Seq/Homer
5	Fli1(ETS)/CD8-FLI-ChIP-Seq(GSE20898)/Homer	GATA3(Zf)/iTreg-Gata3-ChIP-Seq(GSE20898)/Homer	EWS:ERG-fusion(ETS)/CADO_ES1-EWS:ERG-ChIP-Seq/Homer
6	ETV1(ETS)/GIST48-ETV1-ChIP-Seq(GSE22441)/Homer	ETV1(ETS)/GIST48-ETV1-ChIP-Seq/Homer	CTCF(Zf)/CD4+-CTCF-ChIP-Seq/Homer
7	ETS1(ETS)/Jurkat-ETS1-ChIP-Seq(GSE17954)/Homer	ETS1(ETS)/Jurkat-ETS1-ChIP-Seq/Homer	PU.1(ETS)/ThioMac-PU.1-ChIP-Seq/Homer
8	GATA:SCL/Ter119-SCL-ChIP-Seq(GSE18720)/Homer	Fli1(ETS)/CD8-FLI-ChIP-Seq(GSE20898)/Homer	GABPA(ETS)/Jurkat-GABPa-ChIP-Seq/Homer
9	GATA3(Zf)/iTreg-Gata3-ChIP-Seq(GSE20898)/Homer	RUNX1(Runt)/Jurkat-RUNX1-ChIP-Seq/Homer	Ets1-distal(ETS)/CD4+-PolII-ChIP-Seq/Homer
10	GABPA(ETS)/Jurkat-GABPa-ChIP-Seq(GSE17954)/Homer	RUNX(Runt)/HPC7-Runx1-ChIP-Seq/Homer	EWS:FLI1-fusion(ETS)/SK_N_MC-EWS:FLI1-ChIP-Seq/Homer
11	ERG(ETS)/VCaP-ERG-ChIP-Seq(GSE14097)/Homer	EWS:ERG-fusion(ETS)/CADO_ES1-EWS:ERG-ChIP-Seq/Homer	Gata1(Zf)/K562-GATA1-ChIP-Seq/Homer
12	BORIS(Zf)/K562-CTCFL-ChIP-Seq(GSE32465)/Homer	GABPA(ETS)/Jurkat-GABPa-ChIP-Seq/Homer	Gata2(Zf)/K562-GATA2-ChIP-Seq/Homer
13	ELF5(ETS)/T47D-ELF5-ChIP-Seq(GSE30407)/Homer	RUNX2(Runt)/PCa-RUNX2-ChIP-Seq(GSE33889)/Homer	BORIS(Zf)/K562-CTCFL-ChIP-Seq/Homer
14	PU.1(ETS)/ThioMac-PU.1-ChIP-Seq(GSE21512)/Homer	EWS:FLI1-fusion(ETS)/SK_N_MC-EWS:FLI1-ChIP-Seq/Homer	Gata4(Zf)/Heart-Gata4-ChIP-Seq(GSE35151)/Homer
15	EHF(ETS)/LoVo-EHF-ChIP-Seq(GSE49402)/Homer	Ets1-distal(ETS)/CD4+-PolII-ChIP-Seq/Homer	RUNX1(Runt)/Jurkat-RUNX1-ChIP-Seq/Homer
16	Ets1-distal(ETS)/CD4+-PolII-ChIP-Seq(Barski et al.)/Homer	RUNX-AML(Runt)/CD4+-PolII-ChIP-Seq/Homer	RUNX(Runt)/HPC7-Runx1-ChIP-Seq/Homer
17	ELF1(ETS)/Jurkat-ELF1-ChIP-Seq(SRA014231)/Homer	PU.1(ETS)/ThioMac-PU.1-ChIP-Seq/Homer	RUNX-AML(Runt)/CD4+-PolII-ChIP-Seq/Homer
18	Elk4(ETS)/Hela-Elk4-ChIP-Seq(GSE31477)/Homer	GATA:SCL/Ter119-SCL-ChIP-Seq/Homer	ETS:E-box/HPC7-Scl-ChIP-Seq/Homer
19	Elk1(ETS)/Hela-Elk1-ChIP-Seq(GSE31477)/Homer	Jun-AP1(bZIP)/K562-cJun-ChIP-Seq/Homer	RUNX2(Runt)/PCa-RUNX2-ChIP-Seq(GSE33889)/Homer
20	ETS(ETS)/Promoter/Homer	HIF1b(HLH)/O785-HIF1b-ChIP-Seq(GSE34871)/Homer	GATA3(Zf)/iTreg-Gata3-ChIP-Seq(GSE20898)/Homer
21	EWS:FLI1-fusion(ETS)/SK_N_MC-EWS:FLI1-ChIP-Seq(SRA014231)/Homer	AP-1(bZIP)/ThioMac-PU.1-ChIP-Seq/Homer	Jun-AP1(bZIP)/K562-cJun-ChIP-Seq/Homer
22	EWS:ERG-fusion(ETS)/CADO_ES1-EWS:ERG-ChIP-Seq(SRA014231)/Homer	Atoh1(bHLH)/Cerebellum-Atoh1-ChIP-Seq/Homer	AP-1(bZIP)/ThioMac-PU.1-ChIP-Seq/Homer
23	Jun-AP1(bZIP)/K562-cJun-ChIP-	CTCF(Zf)/CD4+-CTCF-ChIP-	HIF1b(HLH)/O785-HIF1b-ChIP-

	Seq(GSE31477)/Homer	Seq/Homer	Seq(GSE34871)/Homer
24	BATF(bZIP)/Th17-BATF-ChIP-Seq(GSE39756)/Homer	ELF1(ETS)/Jurkat-ELF1-ChIP-Seq/Homer	SPDEF(ETS)/VCaP-SPDEF-ChIP-Seq/Homer
25	Atf3(bZIP)/GBM-ATF3-ChIP-Seq(GSE33912)/Homer	Elk4(ETS)/Hela-Elk4-ChIP-Seq(GSE31477)/Homer	GATA:SCL/Ter119-SCL-ChIP-Seq/Homer
26	AP-1(bZIP)/ThioMac-PU.1-ChIP-Seq(GSE21512)/Homer	NeuroD1(bHLH)/Islet-NeuroD1-ChIP-Seq(GSE30298)/Homer	ETS(ETS)/Promoter/Homer
27	RUNX(Runt)/HPC7-Runx1-ChIP-Seq(GSE22178)/Homer	Elk1(ETS)/Hela-Elk1-ChIP-Seq(GSE31477)/Homer	ELF1(ETS)/Jurkat-ELF1-ChIP-Seq/Homer
28	RUNX1(Runt)/Jurkat-RUNX1-ChIP-Seq(GSE29180)/Homer	SPDEF(ETS)/VCaP-SPDEF-ChIP-Seq/Homer	SCL/HPC7-Scl-ChIP-Seq/Homer
29	Bach2(bZIP)/OCILy7-Bach2-ChIP-Seq(GSE44420)/Homer	BORIS(Zf)/K562-CTCFL-ChIP-Seq/Homer	Atoh1(bHLH)/Cerebellum-Atoh1-ChIP-Seq/Homer
30	SPDEF(ETS)/VCaP-SPDEF-ChIP-Seq(SRA014231)/Homer	Bach1(bZIP)/K562-Bach1-ChIP-Seq(GSE31477)/Homer	Elk1(ETS)/Hela-Elk1-ChIP-Seq(GSE31477)/Homer
31	Sp1(Zf)/Promoter/Homer	SCL/HPC7-Scl-ChIP-Seq/Homer	NeuroD1(bHLH)/Islet-NeuroD1-ChIP-Seq(GSE30298)/Homer
32	RUNX-AML(Runt)/CD4+-PolII-ChIP-Seq(Barski et al.)/Homer	ETS:E-box/HPC7-Scl-ChIP-Seq/Homer	Elk4(ETS)/Hela-Elk4-ChIP-Seq(GSE31477)/Homer
33	RUNX2(Runt)/PCa-RUNX2-ChIP-Seq(GSE33889)/Homer	NF-E2(bZIP)/K562-NFE2-ChIP-Seq/Homer	NF-E2(bZIP)/K562-NFE2-ChIP-Seq/Homer
34	BMYB(HTH)/Hela-BMYB-ChIPSeq(GSE27030)/Homer	Nrf2(bZIP)/Lymphoblast-Nrf2-ChIP-Seq(GSE37589)/Homer	ETS:RUNX/Jurkat-RUNX1-ChIP-Seq/Homer
35	MYB(HTH)/ERMYB-Myb-ChIPSeq(GSE22095)/Homer	Tcf12(HLH)/GM12878-Tcf12-ChIP-Seq/Homer	Bach1(bZIP)/K562-Bach1-ChIP-Seq(GSE31477)/Homer
36	AMYB(HTH)/Testes-AMYB-ChIP-Seq(GSE44588)/Homer	ETS(ETS)/Promoter/Homer	Nrf2(bZIP)/Lymphoblast-Nrf2-ChIP-Seq(GSE37589)/Homer
37	NFY(CCAAT)/Promoter/Homer	MyoD(HLH)/Myotube-MyoD-ChIP-Seq/Homer	Olig2(bHLH)/Neuron-Olig2-ChIP-Seq(GSE30882)/Homer
38	Bach1(bZIP)/K562-Bach1-ChIP-Seq(GSE31477)/Homer	Olig2(bHLH)/Neuron-Olig2-ChIP-Seq(GSE30882)/Homer	Foxo1(Forkhead)/RAW-Foxo1-ChIP-Seq/Homer
39	NF-E2(bZIP)/K562-NFE2-ChIP-Seq(GSE31477)/Homer	MyoG(HLH)/C2C12-MyoG-ChIP-Seq(GSE36024)/Homer	Hoxc9/Ainv15-Hoxc9-ChIP-Seq/Homer
40	ETS:RUNX/Jurkat-RUNX1-ChIP-Seq(GSE17954)/Homer	ETS:RUNX/Jurkat-RUNX1-ChIP-Seq/Homer	Tcf12(HLH)/GM12878-Tcf12-ChIP-Seq/Homer
41	Nrf2(bZIP)/Lymphoblast-Nrf2-ChIP-Seq(GSE37589)/Homer	Myf5(bHLH)/GM-Myf5-ChIP-Seq(GSE24852)/Homer	MyoG(HLH)/C2C12-MyoG-ChIP-Seq(GSE36024)/Homer
42	GFY-Staf/Promoters/Homer	MYB(HTH)/ERMYB-Myb-ChIPSeq(GSE22095)/Homer	MyoD(HLH)/Myotube-MyoD-ChIP-Seq/Homer
43	KLF5(Zf)/LoVo-KLF5-ChIP-Seq(GSE49402)/Homer	BMYB(HTH)/Hela-BMYB-ChIPSeq(GSE27030)/Homer	REST-NRSF(Zf)/Jurkat-NRSF-ChIP-Seq/Homer
44	SCL(HLH)/HPC7-Scl-ChIP-Seq(GSE13511)/Homer	EGR(Zf)/K562-EGR1-ChIP-Seq/Homer	PU.1-IRF/Bcell-PU.1-ChIP-Seq/Homer
45	PU.1-IRF(ETS:IRF)/Bcell-PU.1-ChIP-Seq(GSE21512)/Homer	Hoxc9/Ainv15-Hoxc9-ChIP-Seq/Homer	BMYB(HTH)/Hela-BMYB-ChIPSeq(GSE27030)/Homer
46	Stat3+il21(Stat)/CD4-Stat3-ChIP-Seq(GSE19198)/Homer	MafK(bZIP)/C2C12-MafK-ChIP-Seq(GSE36030)/Homer	HOXA9/HSC-Hoxa9-ChIP-Seq(GSE33509)/Homer
47	STAT4(Stat)/CD4-Stat4-ChIP-Seq(GSE22104)/Homer	MafA(bZIP)/Islet-MafA-ChIP-Seq(GSE30298)/Homer	MafK(bZIP)/C2C12-MafK-ChIP-Seq(GSE36030)/Homer
48	GFY(?)/Promoter/Homer	STAT4(Stat)/CD4-Stat4-ChIP-Seq/Homer	MafA(bZIP)/Islet-MafA-ChIP-Seq(GSE30298)/Homer
49	Maz(Zf)/HepG2-Maz-ChIP-Seq(GSE31477)/Homer	Foxo1(Forkhead)/RAW-Foxo1-ChIP-Seq/Homer	Myf5(bHLH)/GM-Myf5-ChIP-Seq(GSE24852)/Homer
50	Stat3(Stat)/mES-Stat3-ChIP-Seq(GSE11431)/Homer	GATA-IR4(Zf)/iTreg-Gata3-ChIP-Seq(GSE20898)/Homer	RFX(HTH)/K562-RFX3-ChIP-Seq/Homer

Supplemental References

Adey, A., Morrison, H.G., Asan, Xun, X., Kitzman, J.O., Turner, E.H., Stackhouse, B., MacKenzie, A.P., Caruccio, N.C., Zhang, X., and Shendure, J. (2010). Rapid, low-input, low-bias construction of shotgun fragment libraries by high-density in vitro transposition. Genome Biol. 11, R119.

Drach, J., Lopez-Berestein, G., McQueen, T., Andreeff, M., and Mehta, K. (1993). Induction of differentiation in myeloid leukemia cell lines and acute promyelocytic leukemia cells by liposomal all-trans-retinoic acid. Cancer Res. 53, 2100–2104.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 10, R25.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079.

Park, E.K., Jung, H.S., Yang, H.I., Yoo, M.C., Kim, C., and Kim, K.S. (2007). Optimized THP-1 differentiation is required for the detection of responses to weak stimuli. Inflamm. Res. 56, 45–50.

Rashid, N.U., Giresi, P.G., Ibrahim, J.G., Sun, W., and Lieb, J.D. (2011). ZINBA integrates local covariates with DNA-seq data to identify broad and narrow regions of enrichment, even within amplified genomic regions. Genome Biol. 12, R67.

Reinisch, A., Etchart, N., Thomas, D., Hofmann, N.A., Fruehwirth, M., Sinha, S., Chan, C.K., Senarath-Yapa, K., Seo, E.-Y., Wearda, T., et al. (2015). Epigenetic and in vivo comparison of diverse MSC sources reveals an endochondral signature for human hematopoietic niche formation. Blood 125, 249–260.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139–140.

Shao, Z., Zhang, Y., Yuan, G.-C., Orkin, S.H., and Waxman, D.J. (2012). MAnorm: a robust model for quantitative comparison of ChIP-Seq data sets. Genome Biol. 13, R16.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol. 9, R137.