Published in final edited form as:

Nat Med. 2015 October ; 21(10): 1172–1181. doi:10.1038/nm.3936.

Minimal PU.1 reduction induces a preleukemic state and promotes development of acute myeloid leukemia

Britta Will^{*,1,†}, Thomas O. Vogler^{1,†}, Swathi Narayanagari¹, Boris Bartholdy¹, Tihomira I. Todorova¹, Mariana da Silva Ferreira¹, Jiahao Chen¹, Yiting Yu², Jillian Mayer¹, Laura Barreyro¹, Luis Carvajal¹, Daniela Ben Neriah¹, Michael Roth¹, Johanna van Oers¹, Sonja Schaetzlein¹, Christine McMahon³, Winfried Edelmann^{1,4,5}, Amit Verma^{2,4,6,7}, and Ulrich Steidl^{1,2,4,7,*}

¹Department of Cell Biology, Albert Einstein College of Medicine, Bronx, USA

²Department of Medicine (Oncology), Division of Hemato-Oncology, Albert Einstein College of Medicine / Montefiore Medical Center, Bronx, USA

³Department of Pathology, Albert Einstein College of Medicine, Bronx, USA

⁴Albert Einstein Cancer Center, Albert Einstein College of Medicine, Bronx, USA

⁵Department of Genetics, Albert Einstein College of Medicine, Bronx, USA

⁶Department of Developmental and Molecular Biology, Albert Einstein College of Medicine, Bronx, USA

⁷Institute for Stem Cell and Regenerative Medicine Research, Albert Einstein College of Medicine, Bronx, USA

Abstract

Modest transcriptional changes caused by genetic or epigenetic mechanisms are frequent in human cancer. Although loss or near-complete loss of the hematopoietic transcription factor PU.1 induces acute myeloid leukemia (AML) in mice, a similar degree of PU.1 impairment is exceedingly rare in human AML; yet moderate PU.1 inhibition is common in AML patients. We assessed functional consequences of modest reduction of PU.1 expression on leukemia development in mice harboring DNA lesions resembling those acquired during human stem cell aging. Heterozygous deletion of an enhancer of PU.1, which resulted in 35% reduction of PU.1 expression, was sufficient to induce myeloid biased preleukemic stem cells and subsequent transformation to AML in a DNA mismatch repair-deficient background. AML progression was mediated by inhibition of expression of a PU.1 cooperating transcription factor, Irf8. Strikingly, we found significant molecular similarities with human myelodysplastic syndrome and AML. This

*Correspondence should be addressed to B.W. (britta.will@einstein.yu.edu) or U.S. (ulrich.steidl@einstein.yu.edu).

†Authors contributed equally.

AUTHOR CONTRIBUTIONS

B.W., U.S., A.V., W.E. designed the study and experiments. B.W., T.O.V., S.N., T.I.T., J.M., M.S.F., L.C., D.B.N., M.R., J.O., S.S. conducted experiments. B.B., J.C., Y.Y., L.B. and B.W. performed gene expression, large data set and pathway analyses. C.M. and A.V. performed cell pathological analyses. B.W., T.O.V., and U.S. wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interest.

study demonstrates that minimal reduction of a key lineage-specific transcription factor that commonly occurs in human disease is sufficient to initiate cancer development and provides mechanistic insight into the formation and progression of preleukemic stem cells in AML.

INTRODUCTION

Genomic studies have shown that in human cancer somatic DNA alterations often occur within the non-coding part of the genome, are enriched in gene-regulatory regions, and cause only moderate transcriptional changes. It is currently not well understood if and how such moderate gene expression changes contribute to malignant transformation.

The progression from a hematopoietic stem cell (HSC) to a fully differentiated cell is a multistep process¹. A set of key transcriptional regulators establish stable, lineage- and cell type-specific gene expression and thereby control cell fate and differentiation outcomes². One such master regulator is the Ets-family transcription factor PU.1, which is indispensable for HSC function and the differentiation of cells within the myeloid as well as lymphoid lineages^{3–5}.

Acute myeloid leukemia (AML) is the most frequent acute leukemia in adults with a median age of 67 years at diagnosis⁶; it develops through a multi-step transformation process originating in HSCs. Initial genetic or epigenetic aberrations lead to the formation of preleukemic stem cells with altered function and an increased propensity for subsequent progression to AML⁷. AML consists of transplantable “leukemia-initiating cells” and a tumor bulk of myeloid cells incapable of terminal differentiation (“leukemic blasts”) accumulating in peripheral blood and bone marrow⁸. Genes encoding transcription factors are frequently mutated, rearranged, or otherwise deregulated in human AML, and mouse models of leukemia have demonstrated roles for several deregulated lineage-determining transcriptional master regulators, including PU.1, in the initiation of AML^{9–12}. Reduction of PU.1 expression by 80%–100% induces AML in mice, whereas PU.1 haploinsufficiency causes subtle changes in hematopoietic differentiation, but is not sufficient to induce leukemia^{3, 9, 11, 13, 14}. The greatly diminished PU.1 levels required to induce AML in mice do not resemble the relatively moderate reduction in PU.1 levels frequently observed in human AML. Several molecular mechanisms through which PU.1 expression or its activity is impaired in human AML cells have been described but while common, their effects on PU.1 are relatively modest^{15–20}. Homozygous mutations or deletions of the *PU.1* gene have not been observed in human AML; only some rare cases with heterozygous mutations or heterozygous deletions have been reported^{21, 22}. We hypothesized that minimal reduction in PU.1 expression can be a founding event for myeloid transformation, specifically in the context of acquired mutations accumulating during aging.

The exact mechanisms of how HSCs and preleukemic stem cells in AML acquire disease-relevant mutations is currently not well resolved, but several lines of evidence support a role of impaired DNA mismatch repair (MMR) in leukemogenesis^{23–25}. Mice lacking *Msh2*, the key component of both the MutS α and MutS β complexes that mediate DNA MMR, display a genetic phenotype that closely mimics the spectrum of mutations found in aging human HSCs and in patients with myeloid leukemia (transition mutations and small insertions and

deletions)^{26, 27}. We therefore generated a mouse model carrying a heterozygous deletion of an upstream enhancer of *PU.1* and a homozygous deletion of *Msh2* to evaluate the role of minimal PU.1 reduction in the context of acquired mutations.

RESULTS

Minimal reduction of PU.1 expression leads to AML

To assess the effects of minimal PU.1 inhibition in the context of an elevated number of point mutations, in particular C/G>T/A transitions and small insertions/deletions resembling the mutations acquired in aging human individuals and patients with AML, we crossed mice with a heterozygous deletion of a regulatory element 14 kb upstream of the transcriptional start site of *PU.1* (URE^{het})⁹ with *Msh2*^{-/-} mice²⁸. $URE^{het}Msh2^{-/-}$ mice were born at Mendelian frequencies. PU.1 expression in hematopoietic multipotent stem and progenitor cells sorted from URE^{het} mice exhibited a significant ($P < 0.05$), but very modest reduction of *PU.1* expression compared to wild type (WT) littermates ($37 \pm 8\%$ in Lin⁻Sca-1⁺cKit⁺ (LSK) cells, $33\% \pm 4\%$ in common myeloid progenitors (CMP), and $26\% \pm 20\%$ in granulocytic/monocytic progenitors (GMP)) (Fig. 1a and Supplementary Fig. 1a,b). Western blotting confirmed minimal impairment of PU.1 at the protein level (by 36% in myeloid progenitor cells, and 21% in mature neutrophils; Supplementary Fig. 1c). As previously reported⁹, $URE^{-/-}$ mice showed a much greater reduction of *PU.1* levels ($97\% \pm 2\%$ reduction in LSK, $92\% \pm 3\%$ in CMP, and $76\% \pm 5\%$ reduction in GMP) (Fig. 1a and Supplementary Fig. 1a,b).

In agreement with previous reports we observed that *Msh2*^{-/-} mice develop T cell lymphomas with 100% penetrance and a late onset²⁹ (Fig. 1b). Minimal reduction of PU.1 in combination with *Msh2* deficiency had a two-pronged effect: (1) reducing median survival time from 7.6 months for $URE^{+/+}Msh2^{-/-}$ mice to 4.6 months for $URE^{het}Msh2^{-/-}$ mice ($P = 0.0026$; Log-rank test); and (2) altering the disease phenotype to AML in more than two thirds of $URE^{het}Msh2^{-/-}$ mice (Fig. 1c–j). AML was never observed in $URE^{+/+}Msh2^{-/-}$ mice. To exclude the possibility that the observed myeloid phenotype was caused by further inhibition of PU.1 expression or function, we examined expression of known PU.1 target genes in different stem and progenitor cells and found only minimal reduction in their expression level in $URE^{het}Msh2^{+/+}$ and no further reduction in leukemic $URE^{het}Msh2^{-/-}$ mice (Supplementary Fig. 1d). In mice with AML, we also ensured that the PU.1 encoding gene or its regulatory regions had not acquired mutations that further reduced PU.1 activity (except for one mouse out of 20 examined that was excluded from further analyses) (Supplementary Fig. 1e).

Macroscopically, $URE^{het}Msh2^{-/-}$ leukemic mice presented with severe splenomegaly (Fig. 1c) and hepatomegaly (data not shown). Moribund $URE^{het}Msh2^{-/-}$ AML mice showed elevated white blood cell counts (WBC) and reduced red blood cell and platelet counts, in comparison to 4–12 week old (preleukemic) mice of the same genotype or age-matched wild type animals (Fig. 1d–f). We assessed the time course of WBCs of $URE^{het}Msh2^{-/-}$ mice, and found that WBCs rapidly increased shortly before death (Fig. 1g), indicating an acute disease. Aberrant myeloid blasts were present in the peripheral blood, bone marrow and spleen, stained positive for the myeloid marker myeloperoxidase, and disrupted the tissue

architecture of spleen, liver and bone marrow (Fig. 1h and Supplementary Fig. 1f–h) Cytogenetic examination of AML cells showed evidence for clonality (Supplementary Table 1).

Bone marrow and spleen of URE^{het}*Msh2*^{-/-} mice with AML showed an expanded cKit⁺CD8a⁻CD4⁻B220⁻ cell population also expressing high levels of CD44 (Supplementary Fig. 1i–k). Further analysis revealed phenotype variations (Fig. 1i). In 17% of mice with AML (termed “AML (immature)”), aberrant populations appeared morphologically immature (Supplementary Fig. 1l), and did not express or lowly coexpressed the mature myeloid marker gene CD11b on cKit⁺ cells (Supplementary Fig. 1j, second panel from the top). 33% of the URE^{het}*Msh2*^{-/-} mice with AML (termed “AML (mature)”) displayed aberrant cKit⁺CD44⁺CD8a⁻CD4⁻B220⁻ cells coexpressing CD11b (Supplementary Fig. 1j, third panel from the top). Blasts in these mice resembled more mature metamyelocyte-like neutrophils with ring-shaped nuclei (Supplementary Fig. 1l); and 50% of the URE^{het}*Msh2*^{-/-} mice with AML (termed “AML and ALL (mixed lineage)”), harbored an aberrant myeloid cKit⁺CD44⁺CD8a⁻CD4⁻B220⁻ population, and a separate, expanded CD8a⁺ and/or CD4⁺ morphologically immature cell population within hematopoietic organs (Supplementary Fig. 1i,j bottom panels, and Supplementary Fig. 1m–o). Kaplan-Meier survival analysis of URE^{het}*Msh2*^{-/-} mice further revealed that mice succumbing to mixed lineage leukemia or T cell lymphoma had a significantly faster disease progression compared to mice developing AML alone ($P < 0.001$, Log-rank test) (Fig. 1j). These data show that minimal reduction of PU.1 expression leads to development of AML in 70% of URE^{het}*Msh2*^{-/-} animals, which show phenotype variations resembling human disease.

Leukemia-initiating cells in URE^{het}*Msh2*^{-/-} AML

We characterized self-renewal and differentiation capacity of the aberrant CD8a⁻CD4⁻B220⁻ (Lymph⁻) cKit⁺ population from URE^{het}*Msh2*^{-/-} mice with AML. Compared to wild type cKit⁺Lymph⁻ cells, URE^{het}*Msh2*^{-/-} cKit⁺Lymph⁻ cells from mice with immature AML formed 57% more total colonies, while cells from mice with mature AML gave rise to 74% less colonies in the initial plating of colony-forming assays (Fig. 2a). However, URE^{het}*Msh2*^{-/-} cells of either AML subtype showed greater clonogenic capacity in the 2nd to 5th plating and continued to give rise to aberrant blast-like colonies after the wild type colony-initiating cells were exhausted. The number of myeloid colonies with normal morphology was found reduced in immature and mature AML in the first plating (Fig. 2a). URE^{het}*Msh2*^{-/-}-derived aberrant colonies were comprised of differentiation-impaired myeloid cells with blast morphology and became the sole colony type after the initial cell plating (Fig. 2a–c). Cells isolated from the blast colonies maintained the same immunophenotype as the primary AML bulk cell population (Fig. 2d, Supplementary Fig. 1i–k). Next, we transplanted purified cKit⁺Lymph⁻ cells and stem cell-enriched LSK populations into NOD-SCID IL2R γ null mice. Recipient animals displayed a massive expansion of a donor-derived bulk tumor population with the same phenotype as observed in the primary tumors (Fig. 2e–i and Supplementary Fig. 2a–d). All recipients died of AML within 3–11 weeks demonstrating that the disease is transplantable from the stem cell-containing cell compartment (Fig. 2g and Supplementary Fig. 2d).

Gene expression analysis of leukemia-initiating cKit⁺Lymph⁻ cells from URE^{het}Msh2^{-/-} mice revealed alterations significantly affecting several cellular networks. Comparison with cKit⁺Lymph⁻ cells from age-matched healthy URE^{+/+}Msh2^{+/+} mice identified 587 annotated, differentially expressed genes (Supplementary Table 2), which separated URE^{het}Msh2^{-/-} AML from wild type cells in a hierarchical cluster analysis (Fig. 2i). Pathway analysis showed that differentially expressed genes were enriched for molecules involved in several key cellular processes (Fig. 2j and Supplementary Fig. 2e).

Role for Irf8 downregulation in AML induction

Among the differentially expressed genes found in cKit⁺Lymph⁻ cells from URE^{het}Msh2^{-/-} AML we identified *Irf8* (also known as *ICSBP*) to be significantly reduced compared to cells from age-matched URE^{+/+}Msh2^{+/+} (WT) controls ($P = 0.025$, T test) (Fig. 3a). As Irf8 cooperatively regulates the expression of target genes together with PU.1³⁰⁻³², we examined whether a subset of the deregulated genes in URE^{het}Msh2^{-/-} AML are common targets of PU.1 and Irf8. Utilizing published ChIP-seq data sets, we identified genes with PU.1 or combined PU.1/Irf8 occupancy within their promoter regions that were also differentially expressed in URE^{het}Msh2^{-/-} AML (Supplementary Tables 3, 4). Both subsets of genes showed a significant enrichment of genes involved in AML-relevant pathways (Fig. 3b and Supplementary Fig. 3a). Motif enrichment analysis designated a PU.1/IRF composite DNA binding motif containing a PU-box and an adjacent IRF consensus site as the top enriched motif (Fig. 3c). Intersection of genes occupied by PU.1 and Irf8 in their promoter regions with differentially expressed genes in URE^{het}Msh2^{-/-} AML cells (compared to URE^{+/+}Msh2^{+/+} cells) showed reduced expression of several co-occupied genes (Supplementary Fig. 3b). We next tested whether these could be rescued by Irf8 restoration using a retroviral vector system containing a GFP reporter. Compared to the empty vector control, transduction with a retrovirus encoding Irf8 led to restoration of Irf8 protein levels (Fig. 3d) and indeed concomitantly a significant increase ($P < 0.05$, T test) of expression of several putative PU.1/Irf8 target genes, including *Fam132a*, *Rnf13*, *Osbpl3*, and previously established PU.1/Irf8 co-regulated targets *Etv3*, *Nf1*, and *Emr1* in URE^{het}Msh2^{-/-} AML cells (Fig. 3e, Supplementary Fig. 3c,d). Irf8 expression restoration also induced myeloid differentiation of URE^{het}Msh2^{-/-} AML cells as evidenced by morphology and immunophenotyping (Fig. 3f,g). Furthermore, we observed induction of apoptosis in leukemic cells upon rescue of Irf8 expression ($P < 0.05$, T test) (Fig. 3h). Restoration of Irf8 expression also led to significantly impaired self-renewal of URE^{het}Msh2^{-/-} AML cells ($P < 0.05$) and a progressive loss of colony-initiating cKit-expressing cells in serial replating assays (Fig. 3i and Supplementary Fig. 3e). Mice injected with Irf8-rescued (GFP⁺) URE^{het}Msh2^{-/-} AML cells survived significantly longer than mice transplanted with GFP⁺ URE^{het}Msh2^{-/-} AML cells infected with an empty vector control ($P < 0.001$, Log-rank test) (Fig. 3j). Competitive transplantation of a mix of Irf8-restored (GFP⁺) and non-restored (GFP⁻) URE^{het}Msh2^{-/-} AML cells demonstrated that low Irf8 levels were critical to confer the growth advantage *in vivo*, as URE^{het}Msh2^{-/-} AML cells with restored Irf8 levels were outcompeted by GFP⁻ AML cells in the recipients' bone marrow (BM) and spleen (SP); this observation was not made with the empty vector control (Fig. 3j). Rescue of PU.1 expression led to similar differentiation and apoptosis-inducing effects as rescue of Irf8 in the murine AML cells (Supplementary Fig. 3f-h). Together, these data show that reduced

expression of *Irf8* contributes to the myeloid differentiation block, impaired apoptosis, *in vitro* self-renewal, and competitive growth advantage of leukemic cells in $URE^{het}Msh2^{-/-}$ AML.

Minimal *PU.1* expression reduction induces a preleukemic state

Irf8 was not decreased at the preleukemic stage (Fig. 4a). In contrast, *PU.1* expression although minimally reduced did not change upon progression to AML (Fig. 4a). FACS analysis of phenotypically defined HSC and lymphoid-myeloid multipotent progenitor cells (LMPP) (Supplementary Fig. 4a) revealed a significant two-fold increase of myeloid biased HSCs ($Lin^{-}IL7R\alpha^{-}cKit^{+}Sca-1^{+}CD150^{+}Flt3^{-}CD41^{+}$ ³³ and $CD150^{high}Lin^{-}IL7R\alpha^{-}Flt3^{-}cKit^{+}Sca-1^{+}CD48^{-}$ ³⁴) in URE^{het} mice (vs. age-matched $URE^{+/+}$ mice; $P < 0.01$) and a significant decrease of phenotypic lineage-unbiased HSCs ($Lin^{-}IL7R\alpha^{-}cKit^{+}Sca-1^{+}CD150^{+}Flt3^{-}CD41^{-}$ and $CD150^{medium}Lin^{-}IL7R\alpha^{-}Flt3^{-}cKit^{+}Sca-1^{+}CD48^{-}$; $P < 0.05$) (Fig. 4 b,c and Supplementary Fig. 4b). We next assessed multi-lineage reconstitution by highly purified, phenotypic lineage-unbiased HSCs or myeloid biased HSCs from $URE^{het}Msh2^{+/+}$ animals; both populations resulted in a significant increase in donor-derived myeloid cells in the peripheral blood of recipient animals compared to animals transplanted with wild type HSCs (Fig. 4d and Supplementary Fig. 4c). *In vivo* BrdU labeling revealed an increase in cycling cells within the phenotypic myeloid-biased HSC population in URE^{het} mice (Supplementary Fig. 4d). These results demonstrate that minimal *PU.1* expression reduction is sufficient to induce functional alterations in the immature hematopoietic cell compartment.

HSC-enriched cell populations derived from preleukemic $URE^{het}Msh2^{-/-}$ mice ($URE^{+/+}Msh2^{-/-}$ as a control) had an extended self-renewal capacity with colony initiation beyond the 4th plating (Fig. 4e). Using serial bone marrow aspirations and peripheral blood sampling of preleukemic $URE^{het}Msh2^{-/-}$ mice we observed a progressive increase in immature myeloid cells ($cKit^{+}CD11b^{-/lo}Lymph^{-}$) along with a gradual decrease in mature myeloid cells ($Gr-1^{++}CD11b^{+}Lymph^{-}$) (Fig. 4f,g and Supplementary Fig. 4e,f), as well as expansion of phenotypic HSC/MPP compartments (Supplementary Fig. 4g). We furthermore detected cells with dysplastic morphological features, including Pseudo Pelger-Huet cells, megablastoid erythropoiesis, eosinophilic hypergranulation, and a general mild left-shift with increased numbers of promyelocytes and neutrophilic myelocytes in preleukemic mice (Fig. 4h). Longitudinal complete blood cell count (CBC) analysis showed normal total white blood cell-(WBC), red blood cell-(RBC) and platelet (PTL) numbers in the early preleukemic phase (12–30 weeks before overt AML), a slight, but significant reduction of cells in all three lineages during the late preleukemic phase (4–6 weeks before overt AML), and an up to 25-fold increase of WBCs accompanied by a progressive reduction of RBCs and PTLs in the leukemic stage (Fig. 4i). Comparative pathway analysis of differential gene expression profiles of $cKit^{+}Lymph^{-}$ cells derived from preleukemic mice (compared to age-matched WT mice) revealed a set of differentially expressed genes (Supplementary Table 5) affecting pathways also found significantly deregulated in human myelodysplastic syndrome (MDS) ($P > 1.3 \times 10^{-4}$; Fig. 4j and Supplementary Table 6). These observations demonstrate that minimally impaired *PU.1* induces myeloid biased murine HSPCs and, in the context of accumulating mutations, increases self-renewal and leads to the development of a

preleukemic state resembling cellular and molecular features of human myelodysplasia, a preleukemic disease.

Similarities between mouse $URE^{het}Msh2^{-/-}$ and human AML

Comparative pathway analysis of differentially expressed genes in leukemic $URE^{het}Msh2^{-/-}$ cKit⁺Lymph⁻ cells (relative to age-mated WT controls) identified significantly dysregulated pathways ($P > 1.3 \times 10^{-4}$) shared with human AML (compared to healthy individuals) (Fig. 5a and Supplementary Table 7). We next evaluated whether molecular alterations seen in murine $URE^{het}Msh2^{-/-}$ -induced leukemogenesis are observed during the process of myeloid transformation in humans. As human AML development is highly age-dependent³⁵, we first tested whether changes in *PU.1* expression occur during human aging. Our analyses revealed slightly, but significantly decreased expression levels of PU.1 in highly purified HSC from healthy donors older than 65 years when compared to levels in HSCs from young (20–35 years) ($P < 0.05$; Fig. 5b). We next assessed whether *PU.1* expression reduction correlates with the expression of its transcriptional target genes in human AML. We divided patients from two published gene expression studies into two subgroups: “PU.1 lower” and “PU.1 higher” patients (Fig. 5c and Supplementary Fig. 5a). We determined differential gene expression between the two groups and identified 325 probe IDs corresponding to 219 unique annotated genes with significant expression changes that were also found to be differentially expressed in mouse AML (vs. age-matched wild type controls) (Fig. 5d and Supplementary Fig. 5b). Among differentially expressed genes we found a significant enrichment of genes harboring the PU.1 binding motif in their promoter ($p < 0.05$, Fisher’s exact T test) (Supplementary Fig. 5c), suggesting that the reduction of *PU.1* expression has an impact on expression of its target genes. To test the hypothesis that the molecular alterations seen in our murine AML model show more similarities to human AML with reduced PU.1 expression than to human AML with higher PU.1 expression, we used two different approaches. Performing comparative gene set enrichment analysis (GSEA) we found that dysregulated gene expression programs in PU.1 lower expressing AML patients significantly enriched for 1223 gene sets that were also concordantly enriched in mouse AML, while comparison of PU.1 higher expressing human AML with murine AML only revealed 464 shared gene sets (Supplementary Fig. 5d). Using a second analysis strategy we directly tested whether genes differentially regulated in murine AML are dysregulated in human PU.1 lower or higher AML. The set of genes aberrantly expressed in murine AML showed a significant enrichment in the group of aberrantly expressed genes found in PU.1 lower expressing patients, while they did not enrich in PU.1 higher expressing AML patients’ samples (Supplementary Fig. 5e). Interestingly, among the genes correlated with PU.1 higher or lower expression status in AML patients, we also found *IRF8* (Fig. 5e, Supplementary Fig. 5f). Interestingly, we found no significant *IRF8* expression change during aging (Fig. 5f), or in patients with lower risk MDS (Fig. 5g); yet, we detected significantly lower *IRF8* expression in advanced stage MDS patients (RAEB), who have a higher propensity to progressing to AML (Fig. 5g). Consistent with this observation, MDS patients with lower *IRF8* expression had a significantly worse overall survival ($P = 0.0144$) than patients with higher *IRF8* expression (Supplementary Fig. 5g). AML patients with lower *PU.1* expression showed a significant hypermethylation of the *IRF8* promoter (Fig. 5h). Lastly, we compared expression changes found in $URE^{het}Msh2^{-/-}$ leukemia-initiating

cells with differential gene expression in human leukemic stem cells (LSC) versus healthy HSC. Strikingly, while a significantly positive correlation of *PU.1* and *IRF8* expression was detected in LSC from two independent studies (Fig. 5i), *PU.1* and *IRF8* expression were not correlated in healthy HSC (Fig. 5j). Together, these data show significant molecular resemblance of murine $\text{URE}^{\text{het}}\text{Msh2}^{-/-}$ and human myeloid leukemia pathogenesis. When we restored *IRF8* expression in *IRF8* lower and *IRF8* higher expressing human AML cell lines (Supplementary Fig. 5h) we found that elevation of *IRF8* led to induction of differentiation in all four, and apoptosis in two AML cell lines with lower *IRF8* expression levels, while ectopic expression of *IRF8* in the *IRF8* higher-expressing AML cell lines neither induced differentiation nor apoptosis (Supplementary Fig. 5i).

DISCUSSION

Modest impairment of *PU.1* activity or expression is common in human AML pathogenesis, yet murine models have failed to demonstrate a functional relevance for minimal dosage alterations of this key hematopoietic regulator in AML. Here we provide genetic evidence that minimal inhibition of *PU.1* can be a founding event in leukemogenesis.

Preleukemic alterations induced by minimal *PU.1* reduction

Our data reveal that a minimal *PU.1* reduction by 35% introduces a myeloid bias in multipotent murine HSPCs. Our analysis based on phenotypic markers suggested an expanded myeloid-biased HSC population in mice with minimally reduced *PU.1*. However, cell cycle studies showed that these cells were less quiescent than their wild type counterparts, in line with a previous report³⁶; and adoptive cell transfer experiments showed that they were, while myeloid biased, still capable of multi-lineage reconstitution. Thus, although the cell surface markers define this population as myeloid-biased HSCs the phenotypic characteristics are more consistent with an early multipotent progenitor. Importantly, these myeloid lineage-biased multipotent cells did not give rise to overt myeloid leukemia, which is in line with previous observations in *PU.1* haploinsufficient mice^{37, 38}. Interestingly, we also found minimally reduced *PU.1* expression levels in aged, but otherwise healthy human HSCs, which also have been shown to harbor a myeloid lineage bias^{39, 40}.

PU.1 co-factor inhibition during AML progression

Molecularly, *PU.1* exerts many of its functions through binding of highly specific DNA sequence motifs, often in concert with other transcription factors. Among genes with significantly reduced expression in immature myeloid cells from $\text{URE}^{\text{het}}\text{Msh2}^{-/-}$ mice with AML, we identified *Irf8*, a *PU.1* co-factor^{39, 40}, which is frequently lost or impaired in human myeloid leukemia^{30, 41, 41}. *Irf8*-deficient mice show dysfunctional granulocytic/monocytic lineage determination and develop a myeloproliferative neoplasm⁴². *PU.1*-*IRF8* consensus DNA binding motifs discriminate sites with actual *PU.1* occupancy from *PU.1* consensus sites that are not bound by the transcription factor in human hematopoietic stem and progenitor cells, monocytes, and macrophages⁴³. We indeed found that *Irf8* expression restoration rescued impaired expression of genes harboring *PU.1*/*IRF* consensus binding sites, led to the loss of aberrant self-renewal, promoted myeloid differentiation, and induced

apoptosis in leukemic $URE^{het}Msh2^{-/-}$ cells, which demonstrates that *Irf8* impairment functionally cooperates with minimally reduced PU.1 expression in our model. IRF8 has also been implicated in human leukemia in which downregulation of the transcription factor caused resistance to apoptosis and was associated with disease progression⁴⁴. Our results provide evidence that minimal PU.1 reduction cooperates with IRF8 impairment also in human leukemogenesis: (1) patients with myelodysplastic syndrome (MDS) with a higher risk for the progression to AML had lower IRF8 levels, suggesting that inactivation of IRF8 promotes leukemogenesis; (2) Lower IRF8 expression was detected specifically in AML patients with reduced PU.1 levels; (3) Restoration of IRF8 expression induced differentiation in IRF8 low expressing AML cells, and (4) a positive correlation of PU.1 and IRF8 expression was found in human leukemia stem cells, but not in healthy HSCs.

Co-modeling of age-related human stem cell alterations

Aging hematopoietic stem cells progressively accumulate mutations and molecular alterations, which can impair cellular function, give rise to preleukemic stem cells, and initiate age-associated myeloid malignancies^{45, 46}. Modeling the accumulation of small DNA lesions during the aging process utilizing an *Msh2*-deficient background, we found that minimal PU.1 expression reduction triggered significant cellular and molecular alterations consistent with MDS, a preleukemic disorder frequently progressing to AML⁴⁷. Compound mutant $URE^{het}Msh2^{-/-}$ mice also frequently developed AML that, as in human AML, showed considerable phenotype heterogeneity. This variability is most likely owing to differences in acquired genetic and epigenetic alterations in *Msh2*-deficient cells⁴⁷, and further supports a role of PU.1-dependent stem cell fate dysregulation as an early preleukemic event during leukemogenesis. Recent studies in patients with AML have confirmed the functional importance and clinical significance of preleukemic stem cell populations as the cellular origin of leukemia-initiating cells in human disease^{27, 48}. However, our functional and mechanistic knowledge of preleukemic stem cells and their progression is still very limited. Our study reveals a mechanism of leukemic transformation mediated by the minimal inhibition of a key transcriptional regulator of hematopoiesis, PU.1, which alters hematopoietic stem cell fate and sensitizes them to further malignant transformation.

Genetic models have demonstrated that alteration of enhancer function can cause hematologic diseases^{49, 50}. Very recently, AML has also emerged as such an “enhanceropathy”⁵¹. Our study provides further proof that even heterozygous disruption of critical *cis*-regulatory regions can be potent drivers of malignant transformation. It is possible that other key transcriptional regulators with small expression alterations play critical roles, particularly in the early stages of tumorigenesis, and thus deserve our attention in the future.

ONLINE METHODS

Mice

PU.1 URE^{het} mice were kindly provided by Dr. D. G. Tenen. $URE^{het}Msh2^{-/-}$ were generated by mating *Msh2* deficient mice (*Msh2*^{-/-}) (C57Bl/6) to *PU.1* URE^{het} mice. NOD-

Scid IL2R γ null (NOG) and B6.SJL-Ptprca Pepcb/BoyJ mice were purchased from Jackson Labs (Bar Harbor, ME). All mice were housed in a special pathogen-free (SPF) barrier facility. All experimental procedures conducted on mice were approved by the Institutional Animal Care and Use Committee (IACUC; protocol #2013-1202). Mice with the indicated genotypes were included in the study without any further preselection or formal randomization and comprised balanced numbers from both genders; we used age and gender-matched mice. Investigators were not blinded to genotype group allocations.

Histology

Femoral bones, spleens, and livers were fixed for > 24 hrs in neutrally buffered formalin at room temperature, subjected to paraffin embedding, cut into sections using a microtome and stained with Hematoxylin and Eosin (H&E) according to standard protocols. Cytospins of single cell suspensions from bone marrow and spleen samples were prepared after erythrocyte lysis using ACK buffer. Cytospun cells were stained using a modified Giemsa stain (Thermo Scientific Shandon Kwik-Diff Stains) according to the manufacturer's recommendation. Cell and tissue morphology was evaluated using an Axiovert 200M microscope (Zeiss, Maple Grove, MN) or an EVOS FL Auto microscope (Life Technologies, Grand Island, NY).

Complete blood counts

Peripheral blood was obtained from the mouse facial vein using standard techniques and analyzed using the Forcyte Hematology Analyzer (Oxford Science Inc., Oxford CT) according to the manufacturer's instructions.

Analysis and purification of hematopoietic stem and progenitor cells

Total bone marrow cells were isolated from the tibiae, femurs and pelvic bones as previously described⁵². Isolated cells were treated with ACK buffer pH 7.4 (0.15 M NH₄Cl, 10 mM KHCO₃, 1.0 mM EDTA) to lyse red blood cells. After two washes with phosphate buffered saline (without MgCl₂) containing 2% fetal bovine serum (PBS/FBS) immature hematopoietic cells were enriched using magnetic bead negative enrichment with Dynabeads (Invitrogen, Carlsbad, CA). Cells were incubated with the following antibodies (all labeled with PE-Cy5/Tricolor): CD4 (GK1.5), CD8a (53-6.7), CD19 (1D3) (1:100 in PBS/FBS; all from eBioscience, San Diego, CA) and B220 (RA3-6B2), Gr-1 (RB6-8C5) (1:50 in PBS/FBS, both from Invitrogen, Carlsbad, CA) (further referred to as "Lin"). After immunomagnetic separation, unbound cells were washed once and stained with the following antibody cocktail for isolation of hematopoietic stem (HSC) and progenitor cells (all 1:30 in PBS/FBS) for 30 min on ice: APC Alexa 750 CD117 (eBioscience, ACK2), Pacific Blue Sca-1 (Biolegend, D7), PE-Cy7 CD127 (eBioscience, A7R34), FITC CD41 (eBioscience, MWReg30), PE CD150 (Biolegend, TC15-12F12.2), APC Flt3 (eBioscience, A2F10). Cells were washed once and immediately sorted using an Aria II Special Order flow cytometer (Beckton Dickinson, San Jose, CA). Analysis of FACS data was done using the BD FACSDiva software (Beckton Dickinson).

Metaphase preparation of primary hematopoietic cells

cKit⁺Lymph⁻ cells were isolated from ACK-treated single cell suspensions of the bone marrow or spleens of mice by FACS. To promote cell proliferation, single cell suspensions were precultured for 12hrs in M5300 containing 50 ng/ml rmSCF, 10 ng/ml rmIL-3, 10 ng/ml rmIL-6 (all from Peprotech) and 10IU/ml rhEpo (Epogen, Amgen). The cells were then exposed to colcemid (1h at 37 °C, KaryoMax10µg/mL, Invitrogen), centrifuged for 10 min at 300xg and incubated with hypotonic KCl solution (0.075 M KCl pre-warmed at 37°C, ThermoFisher). Afterwards, the cells were fixed and washed 4 times with methanol/ acetic acid solution (3:1). 40µl of the cell suspension was dropped onto a clean slide, in 48% humidity and 24 °C, de-identified for blinded analysis and then stored at 37 °C until analyzed. Chromosomes were counted in 10 metaphases for each condition.

Spectral karyotyping (SKY)

SKY was performed as described before⁵³. Briefly, the slides with metaphases were denatured with 50% FA/2xSSC at 80 °C for 1.5 min and then dehydrated with serial ethanol washing steps (ice cold 70, 90, and 100% for 3 min each). Mouse SKY paint probes (Applied Spectral Imaging) were denatured in the hybridization solution (50% dextran sulfate/2xSSC) at 85 °C for 5min and then applied to the metaphase slides. After 72 hrs the slides were then washed 3-times for 5 min with 50% formamide/2X SSC, 1X SSC and 4xSSC/0.1% Tween. Slides were dehydrated with serial ethanol washing steps and mounted with ProLong Gold antifade reagent with DAPI (Invitrogen) for imaging. Images were acquired using an Axiovert 200 microscope (Zeiss, Germany) connected to an imaging interferometer (SD200, Applied Spectral Imaging, Migdal HaEmek, Israel). Defined rearrangements with nomenclature rules from the International Committee on Standard Genetic Nomenclature for Mice were used for classification.

Transplantation of bone marrow-derived cells

Total bone marrow cells were isolated from the tibiae, femurs and pelvic bones. Single cell suspensions were generated and red blood cells were lysed using ACK-buffer treatment for 30 sec on ice. Cells were washed once with PBS and stained with antibodies. Using an ARIA II Special Order System flow cytometer (Beckton Dickinson) Lin⁻cKit⁺Sca-1⁺CD150⁺IL7Rα⁻Flt3⁻CD48⁻CD41⁻ (lineage-unbiased HSC), CD150^{high} cKit⁺Sca-1⁺Lin⁻IL7Rα⁻Flt3⁻CD48⁻ (myeloid-biased HSC) or Lin⁻Sca-1⁺cKit⁺ (LSK), cKit⁺B220⁻CD4⁻CD8a⁻ (cKit⁺Lymph⁻), or cKit⁻B220⁻CD4⁻CD8a⁻ (cKit⁻Lymph⁻) cells were isolated. Sorted LSK, cKit⁺Lymph⁻ or cKit⁻Lymph⁻ cells were resuspended in HBSS and then transplanted into sublethally irradiated 4–6 week old NOG recipient animals via retro-orbital injection 4hrs after irradiation. Total body irradiation was delivered in a single dose of 250cGy using a Shepherd 6810 sealed-source ¹³⁷Cs irradiator. Lineage-unbiased HSCs and myeloid biased HSCs were transplanted into congenic recipient mice (B6.SJL-Ptprca Pepcb/BoyJ) 4hrs after lethal irradiation (950 cGy total body irradiation) along with 2x10⁵ CD45.1/CD45.2 nucleated bone marrow cells for support. Engraftment of donor cells was monitored as indicated in the figures by analysis of CD45.2 and CD45.1 expression on peripheral blood cells. We stained ACK-treated peripheral blood cells with antibodies (all from eBioscience) against CD45.1 (A20), CD45.2 (104), CD11b (M1/70), Gr-1, CD4, B220

and Ter119 (TER-119) (1:100 in PBS/2% FBS) and analyzed their binding by flow cytometry on an ARIA II Special order System (Beckton Dickinson). Animal numbers were chosen based on previous experiences with transplanting hematopoietic cells of the parental mouse strain^{9, 16}.

In vivo HSC proliferation assay

To characterize in vivo cell cycle activity of HSCs we used a previously described bromodeoxyuridine (BrdU) incorporation protocol³⁶. Briefly, URE^{het} and wild type mice were once injected with BrdU 100 mg/kg (i.p.), and were afterwards given 0.8 mg/ml in the drinking water for 16hrs. Bone marrows were harvested, stained with antibodies against cell surface markers, and fixed. BrdU incorporation was detected by using the BrdU Flow Kit (BD Biosciences) according to the manufacturer's recommendation. BrdU incorporation was measured by flow cytometry on gated CD150^{high} Sca1⁺ckit⁺CD48⁻Lin⁻Flt3⁻ cells.

In vitro colony formation assay and serial replating assay

To characterize clonogenic capacity of cells, we plated 2,000 cKit⁺CD4⁻CD8a⁻B220⁻ or 5,000 GFP⁺ URE^{het}*Msh2*^{-/-} AML cells after transduction in MethoCult M3434 GF+ (Stem Cell Technologies, Vancouver, BC) containing IL-3, IL-6, SCF, and EPO as previously described⁵². Colonies were scored 8–10 days after plating using an Axiovert 200M microscope (Zeiss, Maple Grove, MN) or an EVOS FL Auto microscope (Life Technologies). After 1st plating, we proceeded with serial replating assays until colony formation ceased. Cells (5,000–10,000 cells/ml) were replated in M3434 MethoCult GF+ and colonies were scored again after 8–10 days.

Restoration of Irf8 or PU.1 expression

Total bone marrow was isolated from a URE^{het}*Msh2*^{-/-} mouse with AML, single cell suspension was prepared, and cells were cultured in M5300 (Stem Cell Technologies) containing 50 ng/ml rmSCF, 10 ng/ml rmIL-3, 10 ng/ml rmIL-6 and 10 IU/ml hEpo for five weeks with weekly passage. After the initial five passages, cells were grown in cytokine-free M5300 for 10 more passages. Characterization of cell surface marker expression by FACS revealed high-intermediate cKit, intermediate CD11b and Gr-1 expression, high CD44 (IM7) expression and the absence of B220, CD19, CD4, and CD8a expression. Exponentially growing URE^{het}*Msh2*^{-/-} AML cells were transduced at a cell density of 5x10⁵/ml with either a retroviral expression construct allowing for the ectopic expression of Irf8-eGFP or empty vector-eGFP alone using multiplicities of infection (MOI) ranging between 3–10 as previously described⁵⁴; a lentiviral expression construct was used to ectopically express PU.1-IRES-GFP or GFP alone as described before¹⁵. 48hrs after transduction cultures were analyzed or sorted for eGFP⁺ cells. Overexpression was verified by qRT-PCR and intracellular FACS. Cell death was assessed using the Annexin-V PE Apoptosis Detection kit according to the manufacturer's recommendations (eBioscience). For IRF8 expression rescue experiments, human AML cell lines were transduced with Irf8-eGFP or eGFP vectors⁵⁴. Briefly, exponentially growing Kasumi-3, ML-2, and Molm14 cultured in RPMI-1640 medium with 20% FBS, and HL-60, Thp-1, Nomo-1 cells cultured in RPMI-1640 medium with 10% FBS were transduced in presence of polybrene by spininfection. 72hrs after transduction cells were analyzed for GFP expression by FACS.

Assessment of protein expression

PU.1 protein levels were measured in purified myeloid progenitor cells (~22,000) and neutrophils (~300,000) by Western blotting using a rabbit polyclonal anti-PU.1 antibody (Santa Cruz, clone T-21) as previously described⁹. IRF8 protein abundance was detected by intracellular FACS analysis in murine and human AML cells using an Anti-human/mouse monoclonal IRF8 antibody conjugated with APC (eBioscience, clone V3GYWCH) following the manufacturer's recommendations. A mouse IgG1 K APC-conjugated isotype control (eBioscience, clone P3.6.2.8.1) was used to assess unspecific binding. IRF8 protein abundance was expressed as isotype/background-subtracted mean fluorescence intensity (MFI).

RNA purification, Real time PCR and hybridization to microarrays

RNA was extracted from FACS purified hematopoietic stem and progenitor cells using the RNeasy Micro kit (Qiagen). RNA quantity and quality was assessed using a 2100 Bioanalyzer (Agilent, Santa Clara, CA) device. For real time PCR, RNA was reverse-transcribed using Superscript II (Invitrogen, Carlsbad, CA). Amplification of target genes was performed using the Universal PCR Power SYBR Green mix or TaqMan Universal PCR Master Mix (both from Applied Biosystems, Carlsbad, CA). cDNA was amplified in a final volume of 15 µl in 96, or 8 µl in 384 well microtiter plates according to the manufacturer's recommendation. Primers and probes used for real-time PCR can be found in Supplementary Table S8. We performed real-time PCR using an ViiA7 instrument (Life Technologies) with 1 cycle of 50 °C (2min) and 95 °C (10min) followed by 40 cycles of 95 °C (15sec) and 60 °C (1min). Specific amplification for the target gene products was validated by melting curve analysis and Sanger sequencing. Target gene expression quantification was calculated using the Pfaffl model and normalized to GAPDH expression levels. For global gene expression analysis using microarrays, high quality RNA (RIN ≥ 8) was amplified using the WT Ovation Pico RNA amplification system (Nugen, San Carlos, CA). After labeling with the GeneChip WT terminal labeling kit (Affymetrix, Santa Clara, CA), labelled cRNA of each individual sample was hybridized to an Affymetrix Mouse Gene 1.0ST microarray (Affymetrix), stained, and scanned by GeneChip Scanner 3000 7G system (Affymetrix) according to standard protocols. Data reported in the manuscript are tabulated in the Supplementary Material Methods section and are available at GEO (GSE65671).

Analysis of microarray data

Raw data was normalized with the RMA algorithm (*oligo* package v. 1.3.0) under R (v. 3.1.2)/Bioconductor (v. 3.0). Differentially expressed genes (DEG) were determined after SAM analysis (EMA package, v. 1.4.4); FDR and fold change (FC) cutoffs as indicated. After filtering out unannotated and duplicate genes, the remaining genes were clustered by hierarchical clustering, using Euclidean distance, complete linkage clustering, using R (plots generated using *gplots* (v.2.16.0)). Gene set enrichment analysis was performed using GSEA (Broad Institute). Pathway analyses were performed using DAVID⁵⁵ and Ingenuity Pathway Analysis (IPA[®], Qiagen, Redwood City, www.qiagen.com/ingenuity).

Integrative analysis with published data sets

For direct target gene identification, we intersected binding peaks from published datasets of whole-genome chromatin immunoprecipitation (ChIP-seq) for PU.1 and Irf8 (GSE38824). We used the published bedgraph tracks GSM950325_201104_s_5_chipseq.ucsc.bedGraph (for Irf8) and GSM1031977_Tot2-IRF8_PU.1_Tags.bedGraph (for PU.1), and retained peaks with cutoffs of scores of 15 and 30, respectively. This was sufficiently stringent to eliminate background signal (from GSM950323_201104_s_4_chipseq.ucsc.bedGraph). We next identified genes with PU.1 or PU.1 and IRF8 co-occupancy within promoter regions (using several different region cut offs, as indicated in the figure legends) that were also differentially expressed in URE^{het} *Msh2*^{-/-} AML. HOMER was used to retrieve information about nearest TSS and for identification of DNA binding motifs. Bedtools (v.2.20.1) was used to find geometric overlaps. PWMScan was used to identify HOMER motifs.

Comparative pathway analyses between differentially expressed genes in murine and human AML were performed using IPA and GSEA. Briefly, to identify common pathways with differential activation in murine and human MDS or AML, we used the comparative analysis module under IPA, and compared GSEA results on differentially expressed genes between murine preleukemic or leukemic cells (versus age-matched wild type cells), and human MDS (GSE19429, excluding -7/(del)7 patients) vs. healthy controls, or AML (GSE13204, normal karyotype AML; GSE14468, excluding patients with -7/(del)7). To test whether murine AML was molecularly similar to human AML with lower PU.1 expression we generated gene sets from differentially expressed genes found in human *PU.1 lower* AML compared to healthy controls and *PU.1 higher* AML vs. healthy controls and performed GSEA analysis using genes found to be significantly lower expressed murine AML (compared to wild type control) as the gene expression data set. Significant enrichment was identified using FDR *q*-values of 5% (GSE13204) and 1% (GSE14468) as cutoff.

Statistical analysis

Statistical analysis of group comparisons was performed using Student's T test, Wilcoxon rank sum test and Log-rank test in Excel and Graph Pad Prism or R, as indicated. Statistical significance was set at $P < 5\%$. Statistical evaluation of microarray expression data was performed using the built-in functions of MeV, Ingenuity Pathway analysis, GSEA or the respective R/Bioconductor packages. Sample sizes chosen are indicated in the individual figure legends and were not based on formal power calculations to detect pre-specified effect sizes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to thank Dr. Keiko Ozato for kindly providing the Irf8 expression vector. We thank Dr. Arthur Skoultschi, Dr. Kira Gritsman and the Steidl lab members for very helpful discussions and suggestions. We also thank Guillermo Simkin and Daqian Sun of the Einstein Stem Cell Isolation and Xenotransplantation Facility (funded through NYSTEM grant #C029154), Dr. Cristina Montagna and Zhi Xia Yang from the Einstein Genome Imaging Facility, Mr. David Reynolds and Wen Tran from the Einstein Genomics Core Facility, and Peter Schultes

from the Einstein Department of Cell Biology for expert technical assistance. This work was supported by NIH R00CA131503 (U.S.), and The Gabrielle's Angel Foundation for Cancer Research (U.S.). U.S. is a Research Scholar of the Leukemia and Lymphoma Society and the Diane and Arthur B. Belfer Faculty Scholar in Cancer Research of the Albert Einstein College of Medicine.

References

- Orkin SH. Diversification of haematopoietic stem cells to specific lineages. *Nat Rev Genet.* 2000; 1:57–64. [PubMed: 11262875]
- Iwasaki H, et al. The order of expression of transcription factors directs hierarchical specification of hematopoietic lineages. *Genes Dev.* 2006; 20:3010–3021. [PubMed: 17079688]
- Iwasaki H, et al. Distinctive and indispensable roles of PU.1 in maintenance of hematopoietic stem cells and their differentiation. *Blood.* 2005; 106:1590–1600. [PubMed: 15914556]
- Dakic A, et al. PU.1 regulates the commitment of adult hematopoietic progenitors and restricts granulopoiesis. *J Exp Med.* 2005; 201:1487–1502. [PubMed: 15867096]
- Scott EW, Simon MC, Anastasi J, Singh H. Requirement of transcription factor PU. 1 in the development of multiple hematopoietic lineages. *Science.* 1994; 265:1573–1577. [PubMed: 8079170]
- Juliusson G, et al. Age and acute myeloid leukemia: real world data on decision to treat and outcomes from the Swedish Acute Leukemia Registry. *Blood.* 2009; 113:4179–4187. [PubMed: 19008455]
- Shlush LI, et al. Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia. *Nature.* 2014; 506:328–333. [PubMed: 24522528]
- Klco JM, et al. Functional heterogeneity of genetically defined subclones in acute myeloid leukemia. *Cancer Cell.* 2014; 25:379–392. [PubMed: 24613412]
- Rosenbauer F, et al. Acute myeloid leukemia induced by graded reduction of a lineage-specific transcription factor, PU. 1. *Nat Genet.* 2004; 36:624–630. [PubMed: 15146183]
- Kuo YH, et al. Cbf beta-SMMHC induces distinct abnormal myeloid progenitors able to develop acute myeloid leukemia. *Cancer Cell.* 2006; 9:57–68. [PubMed: 16413472]
- Krivtsov AV, et al. Transformation from committed progenitor to leukaemia stem cell initiated by MLL-AF9. *Nature.* 2006; 442:818–822. [PubMed: 16862118]
- Shimizu R, et al. Leukemogenesis caused by incapacitated GATA-1 function. *Mol Cell Biol.* 2004; 24:10814–10825. [PubMed: 15572684]
- Metcalf D, et al. Inactivation of PU.1 in adult mice leads to the development of myeloid leukemia. *Proc Natl Acad Sci U S A.* 2006; 103:1486–1491. [PubMed: 16432184]
- Cook WD, et al. PU.,1 is a suppressor of myeloid leukemia, inactivated in mice by gene deletion and mutation of its DNA binding domain. *Blood.* 2004; 104:3437–3444. [PubMed: 15304397]
- Steidl U, et al. Essential role of Jun family transcription factors in PU.1 knockdown-induced leukemic stem cells. *Nat Genet.* 2006; 38:1269–1277. [PubMed: 17041602]
- Steidl U, et al. A distal single nucleotide polymorphism alters long-range regulation of the PU.1 gene in acute myeloid leukemia. *J Clin Invest.* 2007; 117:2611–2620. [PubMed: 17694175]
- Yoshida H, et al. PML-retinoic acid receptor alpha inhibits PML IV enhancement of PU.1-induced C/EBPepsilon expression in myeloid differentiation. *Mol Cell Biol.* 2007; 27:5819–5834. [PubMed: 17562868]
- Mueller BU, et al. ATRA resolves the differentiation block in t(15;17) acute myeloid leukemia by restoring PU.1 expression. *Blood.* 2006; 107:3330–3338. [PubMed: 16352814]
- Cancelas JA, et al. Rac GTPases differentially integrate signals regulating hematopoietic stem cell localization. *Nat Med.* 2005; 11:886–891. [PubMed: 16025125]
- Vangala RK, et al. The myeloid master regulator transcription factor PU.1 is inactivated by AML1-ETO in t(8;21) myeloid leukemia. *Blood.* 2003; 101:270–277. [PubMed: 12393465]
- Bonadies N, Pabst T, Mueller BU. Heterozygous deletion of the PU.1 locus in human AML. *Blood.* 2010; 115:331–334. [PubMed: 19890096]
- Mueller BU, et al. Heterozygous PU.1 mutations are associated with acute myeloid leukemia. *Blood.* 2002; 100:998–1007. [PubMed: 12130514]

23. Mao G, et al. Preferential loss of mismatch repair function in refractory and relapsed acute myeloid leukemia: potential contribution to AML progression. *Cell Res.* 2008; 18:281–289. [PubMed: 18227862]
24. Zhu YM, Das-Gupta EP, Russell NH. Microsatellite instability and p53 mutations are associated with abnormal expression of the MSH2 gene in adult acute leukemia. *Blood.* 1999; 94:733–740. [PubMed: 10397740]
25. Diouf B, et al. Somatic deletions of genes regulating MSH2 protein stability cause DNA mismatch repair deficiency and drug resistance in human leukemia cells. *Nat Med.* 2011; 17:1298–1303. [PubMed: 21946537]
26. Zhang S, Lloyd R, Bowden G, Glickman BW, de Boer JG. Thymic lymphomas arising in Msh2 deficient mice display a large increase in mutation frequency and an altered mutational spectrum. *Mutat Res.* 2002; 500:67–74. [PubMed: 11890935]
27. Welch JS, et al. The origin and evolution of mutations in acute myeloid leukemia. *Cell.* 2012; 150:264–278. [PubMed: 22817890]
28. Smits R, et al. Somatic Apc mutations are selected upon their capacity to inactivate the beta-catenin downregulating activity. *Genes Chromosomes Cancer.* 2000; 29:229–239. [PubMed: 10992298]
29. Reitmair AH, et al. MSH2 deficient mice are viable and susceptible to lymphoid tumours. *Nat Genet.* 1995; 11:64–70. [PubMed: 7550317]
30. Schmidt M, Bies J, Tamura T, Ozato K, Wolff L. The interferon regulatory factor ICSBP/IRF-8 in combination with PU.1 up-regulates expression of tumor suppressor p15(Ink4b) in murine myeloid cells. *Blood.* 2004; 103:4142–4149. [PubMed: 14976051]
31. Pourcet B, et al. LXRalpha regulates macrophage arginase 1 through PU.1 and interferon regulatory factor 8. *Circ Res.* 2011; 109:492–501. [PubMed: 21757649]
32. Meraro D, Gleit-Kielmanowicz M, Hauser H, Levi BZ. IFN-stimulated gene 15 is synergistically activated through interactions between the myelocyte/lymphocyte-specific transcription factors, PU.1, IFN regulatory factor-8/IFN consensus sequence binding protein, and IFN regulatory factor-4: characterization of a new subtype of IFN-stimulated response element. *J Immunol.* 2002; 168:6224–6231. [PubMed: 12055236]
33. Gekas C, Graf T. CD41 expression marks myeloid-biased adult hematopoietic stem cells and increases with age. *Blood.* 2013; 121:4463–4472. [PubMed: 23564910]
34. Morita Y, Ema H, Nakauchi H. Heterogeneity and hierarchy within the most primitive hematopoietic stem cell compartment. *J Exp Med.* 2010; 207:1173–1182. [PubMed: 20421392]
35. Appelbaum FR, et al. Age and acute myeloid leukemia. *Blood.* 2006; 107:3481–3485. [PubMed: 16455952]
36. Staber PB, et al. Sustained PU.1 levels balance cell-cycle regulators to prevent exhaustion of adult hematopoietic stem cells. *Mol Cell.* 2013; 49:934–946. [PubMed: 23395001]
37. Dahl R, et al. Regulation of macrophage and neutrophil cell fates by the PU.1:C/EBPalpha ratio and granulocyte colony-stimulating factor. *Nat Immunol.* 2003; 4:1029–1036. [PubMed: 12958595]
38. Genik PC, et al. Leukemogenesis in heterozygous PU.1 knockout mice. *Radiat Res.* 2014; 182:310–315. [PubMed: 25076114]
39. Dykstra B, Olthof S, Schreuder J, Ritsema M, de Haan G. Clonal analysis reveals multiple functional defects of aged murine hematopoietic stem cells. *J Exp Med.* 2011; 208:2691–2703. [PubMed: 22110168]
40. Pang WW, et al. Human bone marrow hematopoietic stem cells are increased in frequency and myeloid-biased with age. *Proc Natl Acad Sci U S A.* 2011; 108:20012–20017. [PubMed: 22123971]
41. Schmidt M, Hochhaus A, Nitsche A, Hehlmann R, Neubauer A. Expression of nuclear transcription factor interferon consensus sequence binding protein in chronic myeloid leukemia correlates with pretreatment risk features and cytogenetic response to interferon-alpha. *Blood.* 2001; 97:3648–3650. [PubMed: 11369663]
42. Holtschke T, et al. Immunodeficiency and chronic myelogenous leukemia-like syndrome in mice with a targeted mutation of the ICSBP gene. *Cell.* 1996; 87:307–317. [PubMed: 8861914]

43. Pham TH, et al. Mechanisms of in vivo binding site selection of the hematopoietic master transcription factor PU.1. *Nucleic Acids Res.* 2013; 41:6391–6402. [PubMed: 23658224]
44. Hu X, et al. IRF8 regulates acid ceramidase expression to mediate apoptosis and suppresses myelogenous leukemia. *Cancer Res.* 2011; 71:2882–2891. [PubMed: 21487040]
45. Holstege H, et al. Somatic mutations found in the healthy blood compartment of a 115-yr-old woman demonstrate oligoclonal hematopoiesis. *Genome Res.* 2014; 24:733–742. [PubMed: 24760347]
46. Xie M, et al. Age-related mutations associated with clonal hematopoietic expansion and malignancies. *Nat Med.* 2014
47. Klepin HD, Rao AV, Pardee TS. Acute Myeloid Leukemia and Myelodysplastic Syndromes in Older Adults. *J Clin Oncol.* 2014
48. Jan M, et al. Clonal Evolution of Preleukemic Hematopoietic Stem Cells Precedes Human Acute Myeloid Leukemia. *Science Translational Medicine.* 2012; 4:1–10.
49. Mercola M, Wang XF, Olsen J, Calame K. Transcriptional enhancer elements in the mouse immunoglobulin heavy chain locus. *Science.* 1983; 221:663–665. [PubMed: 6306772]
50. Herranz D, et al. A NOTCH1-driven MYC enhancer promotes T cell development, transformation and acute lymphoblastic leukemia. *Nat Med.* 2014; 20:1130–1137. [PubMed: 25194570]
51. Groschel S, et al. A single oncogenic enhancer rearrangement causes concomitant EVI1 and GATA2 deregulation in leukemia. *Cell.* 2014; 157:369–381. [PubMed: 24703711]
52. Will B, et al. Satb1 regulates the self-renewal of hematopoietic stem cells by promoting quiescence and repressing differentiation commitment. *Nat Immunol.* 2013; 14:437–445. [PubMed: 23563689]
53. Montagna C, et al. The Septin 9 (MSF) gene is amplified and overexpressed in mouse mammary gland adenocarcinomas and human breast cancer cell lines. *Cancer Res.* 2003; 63:2179–2187. [PubMed: 12727837]
54. Tsujimura H, et al. ICSBP/IRF-8 retrovirus transduction rescues dendritic cell development in vitro. *Blood.* 2003; 101:961–969. [PubMed: 12393459]
55. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc.* 2009; 4:44–57. [PubMed: 19131956]

Figure 1. Minimal reduction of PU.1 expression is sufficient to trigger AML development in mice (a) Quantitative RT-PCR analysis of PU.1 expression in isolated LincKit⁺Sca-1⁺ cells (LSK) from 4–6 week old URE^{het}Msh2^{+/+}, URE⁻Msh2^{+/+} and URE^{het}Msh2^{-/-} mice ($n = 3$). Shown are averages and standard deviations expressed as fold changes compared to wild type controls. (b) Kaplan-Meier analysis of the survival of URE^{het}Msh2^{+/+} ($n = 18$), URE^{+/+}Msh2^{-/-} ($n = 16$) and URE^{het}Msh2^{-/-} ($n = 34$) mice ($P = 0.0026$, Log-rank (Mantel-Cox) test) compared to URE^{+/+}Msh2^{-/-} mice. (c) Spleen weights of moribund URE^{het}Msh2^{-/-} with acute myeloid leukemia (AML) and age-matched wild type controls. Shown are averages and standard error of the mean of $n = 5$ mice/genotype. (d–f) White blood cell counts ($\times 10^3/\mu\text{l}$) (d), red blood cell counts ($\times 10^6/\mu\text{l}$) (e) and platelet counts ($\times 10^6/\mu\text{l}$) (f) of preleukemic or AML URE^{het}Msh2^{-/-} animals, and age-matched URE^{+/+}Msh2^{+/+} mice ($n = 5/\text{group}$). (g) White blood cell counts ($\times 10^3/\mu\text{l}$) of URE^{het}Msh2^{-/-} with AML 2–4 weeks before death and at time of death ($n = 4$). (h) May-Giemsa (Diff-quick), hematoxylin-eosin (H&E) and myeloperoxidase (MPO) stained cells in peripheral blood, bone marrow, and spleen of URE^{het}Msh2^{-/-} mice with AML. Cytospins of Diff-quick stained peripheral blood, bone marrow, and spleen cells (upper row; scale bars: $8\mu\text{m}$). Tissue sections (lower row) stained with MPO (left; scale bar: $8\mu\text{m}$) or H&E (center, right; scale bars: $80\mu\text{m}$) showing infiltration with blast cells (arrow). (i) Pie charts depicting the different disease phenotypes in URE^{het}Msh2^{-/-} (right graph, $n = 35$) and URE^{+/+}Msh2^{-/-} mice (left graph, $n = 17$). (j) Kaplan-Meier survival analysis of URE^{het}Msh2^{-/-} mice with AML ($n = 18$), mixed lineage AML and ALL ($n = 10$), or T cell lymphoma ($n = 11$), and URE^{het}Msh2^{+/+} ($n = 18$). Significance values are indicated as * $P < 0.05$, ** $P < 0.01$; *** $P < 0.001$ (T test, if not specified otherwise).

Figure 2. Characterization of *URE*^{het}*Msh2*^{-/-}-induced AML

(a–d) Serial replating assay of sorted cKit⁺CD4⁻CD8a⁻B220⁻ cells from *URE*^{het}*Msh2*^{-/-} mice with AML or age-matched wild types. (a) Numbers of colony-forming units (CFU)-granulocyte or monocyte (CFU-G/M), CFU-granulocyte/monocyte (CFU-GM), burst-forming units-erythroid (BFU-E), CFU-granulocyte/erythrocyte/monocyte/megakaryocyte (CFU-GEMM), CFU-erythrocyte/ megakaryocyte (CFU-E/MK), and colonies containing undifferentiated blast-like cells (blast). (b) Colony morphology (left) and cytospin (right) of cells after 5th round of plating (scale bars: 100μm and 8μm, respectively). (c) Representative FACS analysis of CD4, CD8a, CD19, B220 (lineage markers) negative cells after 5th round of plating. (d–h) Transplantation of *URE*^{het}*Msh2*^{-/-} Lin⁻cKit⁺Sca-1⁺(LSK) cells. (d) Isolation of LSK cells from the bone marrow of a leukemic donor mouse. (e) FACS analysis of peripheral blood cells of moribund recipients. (f) Genotyping of cells derived from tail and spleen of moribund recipient animals for the presence of PU.1 *URE*⁻ and *Msh2*^{-/-} alleles by PCR. (g) Kaplan-Meier survival analysis of recipients of 5,000 leukemic *URE*^{het}*Msh2*^{-/-} LSK cells (*n* = 4). (h) Cells in bone marrow (left, scale bar: 8μm), liver (H&E stain, center, scale bar: 100μm) and recipient spleen (myeloperoxidase stain, right, scale bar: 100μm) (insets: digital magnification). (i) Differential gene expression in sorted cKit⁺CD4⁻CD8a⁻B220⁻ cells from *URE*^{het}*Msh2*^{-/-} mice with AML (*n* = 3) vs. healthy *URE*^{+/+}*Msh2*^{+/+} (*n* = 4). Heat map of differentially expressed genes (FDR<10%, FC>1.5 or <1/1.5). (j) Ingenuity Pathway analysis of differentially expressed (as in i)).

Figure 3. Role for *Irf8* downregulation in compound mutant AML cells

(a) *Irf8* mRNA expression in *cKit*⁺CD4⁻CD8a⁻B220⁻ cells from URE^{het}*Msh2*^{-/-} mice with AML and wild type controls ($n = 2/\text{genotype}$). Averages and standard deviations are shown. (b) Identification of pathways with significant enrichment by IPA of direct PU.1 and *Irf8* target genes. Subsets were identified by intersection of differentially expressed genes (all DEG) (as in 2i) with published ChIP-seq data reveals subsets of DEG bound by PU.1 alone (PU.1-bound DEG) or PU.1 and IRF8 (PU.1/IRF8-co-bound DEG) in regions spanning the transcriptional start site (TSS) to 20kb upstream. (c) Top composite motif (ETS:IRF) in HOMER analysis of PU.1 and IRF8 co-occurring ChIP-seq peaks in regions 0–5kb upstream of the TSS with high similarity to known consensus motifs (TRANSFAC). (d–j) *Irf8* expression rescue in URE^{het}*Msh2*^{-/-} AML cells by transduction with retroviral constructs expressing *Irf8*-eGFP or eGFP as control. (d) *Irf8* protein levels in rescued and control AML cells compared to primary WT *cKit*⁺FcγR16/32⁺Lymph⁻ cells (% background-subtracted mean fluorescence intensities). (e) mRNA expression changes of PU.1/*Irf8* co-regulated target genes. Shown are averages and standard deviations of technical triplicates expressed as % primary myeloid progenitor cell expression. (f) Diff-Quick stain of AML cells transduced with empty vector control (left panels) or *Irf8*-expressing (right panels) vector. Dividing cells (red asterisks), morphologically more mature neutrophil-like cells (*n*), and myelocyte and metamyelocyte (*a*) stages are indicated. Scale bar: 50μm. (g) Representative FACS plots (left) and quantification of cell surface marker positive, eGFP⁺ cells in three independent experiments (right). (h) Representative FACS density plots of Annexin-V and DAPI stains. Quantification of GFP⁺ viable (Annexin-V⁻DAPI⁻), early apoptotic (Annexin-V⁺DAPI⁻) and late apoptotic (Annexin-V⁺DAPI⁺) cells 48hrs after transduction (three

independent experiments). (i) Serial replating assay of URE^{het}Msh2^{-/-} AML cells (left graph). FACS analysis of cells in culture after serial replating of cKit⁺ URE^{het}Msh2^{-/-} AML (right graph) ($n = 4$). (j) Kaplan-Meier survival analysis of NOG mice receiving 40,000 eGFP⁺ control ($n = 8$) or Irf8-restored ($n = 8$) URE^{het}Msh2^{-/-} AML cells ($P = 0.0002$, Log-rank test) (left graph). FACS analysis of competitive transplantation of unfractionated AML cells after transduction (T) and quantification of eGFP⁺ and eGFP⁻, donor (CD45.2⁺) cells in recipients' spleen (SP) and bone marrow (BM) ($n = 4$ /group). * $P < 0.05$ (T test), if not specified otherwise.

Figure 4. Minimal *PU.1* expression reduction induces a preleukemic state

(a) qRT-PCR detecting *PU.1* or *Irf8* in cKit⁺CD4⁺CD8a⁻B220⁻ cells at time of death (Leukemic) and 12–15 weeks before development of AML (Preleukemic) from the same mice. Averages and standard deviations as fold changes compared to age-matched wild type controls (n=2/group, measured in technical triplicates). (b) Representative FACS plot of CD41 expression on CD150⁺Flt3⁻IL7Rα⁻ KSL HSC in 6–8 week old URE^{het} (right plot) or WT (left plot) animals. (c) Quantification of phenotypical CD41⁺ (left) and CD41⁻ (right) HSC (n = 3 mice/genotype). (d) Lineage output of 100 transplanted CD41⁻ HSC 90 days after competitive transplantation. Donor-derived (CD45.2⁺) T cells (CD4⁺Gr-1⁻B220⁻CD11b⁻), monocytes (CD4⁻B220⁻Gr-1^{medium}CD11b⁺), neutrophils (CD4⁻B220⁻Gr-1⁺CD11b⁺) and B cells (CD4⁻Gr-1⁻CD11b⁻B220⁺) were assessed. Shown are values from individual recipients. (e) Serial replating assay of bone marrow cells derived from 4–6 week old URE^{het}Msh2^{-/-} and URE^{+/+}Msh2^{-/-} mice (six serial replatings (1°–6°), n = 4). (f, g) Longitudinal analysis of bone marrow from URE^{het}Msh2^{-/-} mice ultimately developing AML. (f) FACS plots of viable CD4⁺CD8a⁻B220⁻ (Lymph⁻) cells 18, 12 and 6 weeks before death, and at time of death. (g) Quantification of Gr-1⁺CD11b⁺ neutrophils (top panels) and cKit⁺ cells from 5 individual URE^{het}Msh2^{-/-} mice. (h) Bone marrow cells from preleukemic URE^{het}Msh2^{-/-} mouse showing cell morphological myelodysplastic abnormalities. Scale bar: 8μm. (i) Differential peripheral blood cell counts at the early preleukemic (12–30 weeks before overt AML), late preleukemic (4–6 weeks before overt AML), and leukemic stages. Averages and standard deviations of consecutive measurements from 3 animals per time point. (j) Comparative pathway analysis of differentially expressed genes in cKit⁺Lymph⁻ cells from preleukemic URE^{het}Msh2^{-/-} mice (vs. age-matched wild

type controls, FDR<10%, FC >1.5 or <1/1.5, $n = 4$) and human patients with MDS ($n = 174$) vs. healthy controls ($n = 17$; FDR<5%, FC >1.5 or <1/1.5) (colored tiles indicate activation z-score). * $P < 0.05$, ** $P < 0.01$ (T test, if not specified otherwise).

Figure 5. Similarities between mouse $URE^{het}Msh2^{-/-}$ and human myeloid leukemogenesis

(a) Comparative pathway analysis of genes differentially expressed in $cKit^{+}CD4^{-}CD8a^{-}B220^{-}$ cells derived from $URE^{het}Msh2^{-/-}$ mice with AML (vs. age-matched wild type controls (FDR<10%, FC >1.5 or <1/1.5, $n = 4$)), and human patients with AML (GSE13204 ($n = 351$); FDR<1%, logFC >1.1 or <-1.1; GSE14468 ($n = 212$) vs. healthy controls ($n = 11$); FDR<0.2%, logFC >1.5 or <-1.5) (colored tiles indicate activation z-score). (b) *PU.1* expression in human hematopoietic stem cells during aging (GSE32719; $n(Young) = 14$, $n(Middle) = 5$, $n(Old) = 8$). (c) AML patients (GSE14468) were dichotomized into two groups based on their *PU.1* expression: PU.1 higher expressing (> 80th percentile, $n = 25$) and PU.1 lower expressing (< 50th percentile, $n = 62$), and (d) hierarchical cluster analysis was performed based on 325 probe IDs of differentially expressed genes between PU.1 higher and PU.1 lower expressing patients which also showed differential expression in $URE^{het}Msh2^{-/-}$ mice with AML ($q < 0.01$, log2FC: ± 1.2). (e–g) *IRF8* gene expression analysis in (e) AML patient- and age-matched healthy individual-derived $CD34^{+}$ cells (GSE14468) (* $P < 0.05$, Wilcoxon sum rank test). (f) *IRF8* expression in young and aged HSCs (GSE32719), (g) patients with MDS (grouped by WHO classification into patients with refractory anemia (RA; $n = 51$), refractory anemia with ringed sideroblasts (RARS; $n = 44$), refractory anemia with excess of blasts (RAEB; $n =$

79)) and age-matched healthy controls ($n = 17$). **(h)** DNA cytosine methylation analysis (GSE18700, $n = 125$) ($*P < 0.001$, T test). **(i,j)** Correlation plots of *PU.1* and *IRF8* expression in leukemia stem cells from two independent studies **(i)**, and in healthy individuals (GSE35008) **(j)**. Axes show arbitrary units of log2-transformed array values. $*P < 0.05$, $**P < 0.01$; $***P < 0.001$ (T test, if not specified otherwise).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript