

 Open access • Posted Content • DOI:10.1101/639617

## Myeloid Lineage Enhancers Drive Oncogene Synergy in CEBPA/CSF3R Mutant Acute Myeloid Leukemia — [Source link](#)

Theodore P. Braun, Mariam Okhovat, Cody Coblentz, Sarah A. Carratt ...+25 more authors

**Institutions:** [Oregon Health & Science University](#), [Fred Hutchinson Cancer Research Center](#), [University of Oxford](#), [University of Washington](#) ...+1 more institutions

**Published on:** 15 May 2019 - [bioRxiv](#) (Cold Spring Harbor Laboratory)

**Topics:** [CEBPA](#), [Myeloid leukemia](#), [Lineage \(genetic\)](#), [Acute leukemia](#) and [Granulocyte colony-stimulating factor receptor](#)

Related papers:

- [Myeloid lineage enhancers drive oncogene synergy in CEBPA/CSF3R mutant acute myeloid leukemia](#)
- [Identification of novel CEBPA double mutations capable of promoting familial AML via the suppression of myeloid differentiation.](#)
- [Recent Advances of Research on CEBPA Mutation in Acute Myeloid Leukemia](#)
- [Mutant CEBPA directly drives the expression of the targetable tumor-promoting factor CD73 in AML](#)
- [A regulatory element in the 3'-untranslated region of CEBPA is associated with myeloid/NK/T-cell leukemia](#)

Share this paper:    

View more about this paper here: <https://typeset.io/papers/myeloid-lineage-enhancers-drive-oncogene-synergy-in-cebpa-32hcbb6pmz>

ARTICLE

<https://doi.org/10.1038/s41467-019-13364-2>

OPEN

# Myeloid lineage enhancers drive oncogene synergy in CEBPA/CSF3R mutant acute myeloid leukemia

Theodore P. Braun<sup>1,2</sup>, Mariam Okhovat<sup>3</sup>, Cody Coblentz<sup>1,2</sup>, Sarah A. Carratt<sup>1,2</sup>, Amy Foley<sup>1,2</sup>, Zachary Schonrock<sup>1,2</sup>, Brittany M. Smith<sup>1,2</sup>, Kimberly Nevenon<sup>3</sup>, Brett Davis<sup>3</sup>, Brianna Garcia<sup>1,2</sup>, Dorian LaTocha<sup>1,2</sup>, Benjamin R. Weeder<sup>4</sup>, Michal R. Grzadkowski<sup>5</sup>, Joey C. Estabrook<sup>5</sup>, Hannah G. Manning<sup>5</sup>, Kevin Watanabe-Smith<sup>1,5</sup>, Sophia Jeng<sup>6</sup>, Jenny L. Smith<sup>7</sup>, Amanda R. Leonti<sup>7</sup>, Rhonda E. Ries<sup>7</sup>, Shannon McWeeney<sup>1,6</sup>, Cristina Di Genua<sup>8</sup>, Roy Drissen<sup>8</sup>, Claus Nerlov<sup>8</sup>, Soheil Meshinchi<sup>7,9</sup>, Lucia Carbone<sup>3,6,10</sup>, Brian J. Druker<sup>1,2,11</sup> & Julia E. Maxson<sup>1,2\*</sup>

Acute Myeloid Leukemia (AML) develops due to the acquisition of mutations from multiple functional classes. Here, we demonstrate that activating mutations in the granulocyte colony stimulating factor receptor (CSF3R), cooperate with loss of function mutations in the transcription factor CEBPA to promote acute leukemia development. The interaction between these distinct classes of mutations occurs at the level of myeloid lineage enhancers where mutant CEBPA prevents activation of a subset of differentiation associated enhancers. To confirm this enhancer-dependent mechanism, we demonstrate that CEBPA mutations must occur as the initial event in AML initiation. This improved mechanistic understanding will facilitate therapeutic development targeting the intersection of oncogene cooperativity.

<sup>1</sup> Knight Cancer Institute, Oregon Health & Science University, Portland, OR 97239, USA. <sup>2</sup> Division of Hematology and Medical Oncology, Oregon Health & Science University, Portland, OR 97239, USA. <sup>3</sup> Knight Cardiovascular Institute, Oregon Health & Science University, Portland, OR 97239, USA. <sup>4</sup> Program in Molecular and Cellular Biology, Oregon Health & Science University, Portland, OR 97239, USA. <sup>5</sup> Computational Biology Program, Oregon Health & Science University, Portland, OR 97239, USA. <sup>6</sup> Division of Bioinformatics and Computational Biology, Oregon Health & Science University, Portland, OR 97239, USA. <sup>7</sup> Clinical Research Division, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave N, Seattle, WA 98109, USA. <sup>8</sup> MRC Molecular Haematology Unit, MRC Weatherall Institute of Molecular Medicine, John Radcliffe Hospital, Headley Way, Oxford OX3 9DS, UK. <sup>9</sup> Division of Pediatric Hematology/Oncology, University of Washington, Seattle, WA 98195, USA. <sup>10</sup> Department of Molecular and Medical Genetics, Oregon Health & Science University, Portland, OR 97239, USA. <sup>11</sup> Howard Hughes Medical Institute, Portland, OR, USA. \*email: [maxsonj@ohsu.edu](mailto:maxsonj@ohsu.edu)

**A**cute myeloid leukemia (AML) is a deadly hematologic malignancy that results from the stepwise acquisition of genetic aberrations. The mutations that collaborate to produce AML often occur in distinct functional categories<sup>1</sup>. Class I mutations activate signaling pathways and in isolation produce disease with a myeloproliferative phenotype. Class II mutations perturb the function of transcription factors or epigenetic modifiers—and alone can cause a myelodysplastic phenotype. However, when present in combination, Class I and Class II mutations produce a highly proliferative disease with a block in differentiation, both hallmarks of AML.

The transcription factor CCAAT enhancer-binding protein alpha (CEBPA) is a master regulator of myeloid lineage commitment. CEBPA is recurrently mutated in AML and is a classic example of a Class II mutation<sup>2</sup>. The CEBPA gene comprises a single exon with an internal translational start site. Mutations in CEBPA cluster at the N- and C-terminus of the protein. N-terminal mutations typically result in a frameshift, leading to a premature stop codon and loss of expression of the long (p42) isoform of the protein but ongoing translation of the short isoform (p30). The p30 isoform lacks a crucial transactivation domain that represses cell cycle progression through a direct interaction with E2F<sup>3</sup>. In contrast, C-terminal mutations occur in the DNA-binding domain leading to loss of function and a blockade in granulocytic differentiation<sup>2</sup>. The most common pattern of mutation in AML is an N-terminal mutation on one allele and a C-terminal mutation on the other allele. This results in alterations in the ratio of the p42 to p30 CEBPA isoforms, changing the balance of differentiation and proliferation. AML with biallelic CEBPA mutations is associated with favorable prognosis, with approximately 50% of younger patients achieving a cure with chemotherapy alone. However, the precise determinants of relapse in this mutational context are unknown.

Recent studies have identified a high rate of co-occurrence of mutations in CEBPA with mutations in the granulocyte-colony stimulating factor receptor (CSF3R)<sup>4–7</sup>. Approximately 20–30% of patients with CEBPA-mutant AML harbor a cooperating mutation in CSF3R. Although CSF3R mutations are often associated with biallelic CEBPA mutations, monoallelic cases also occur. Interestingly, monoallelic C-terminal CEBPA mutations are far more likely to co-occur with mutations in CSF3R. Patients with CSF3R/CEBPA mutant AML have inferior outcomes to those with mutant CEBPA alone, arguing that the presence of Class I mutations in CSF3R may be an important determinant of chemotherapy resistance and relapse<sup>8</sup>. The mutations in CSF3R most commonly occur in the membrane proximal region, and lead to ligand independent receptor dimerization and constitutive signaling via the JAK/STAT pathway. Similar to other Class I mutations, membrane proximal CSF3R mutations produce a myeloproliferative phenotype when present in isolation and are the major oncogenic driver of the disease chronic neutrophilic leukemia<sup>9</sup>. During normal myeloid development, G-CSF signaling via CSF3R leads to proliferation of myeloid precursors and neutrophilic differentiation. CEBPA is required for CSF3R-mediated transcription of myeloid specific genes, and myeloid differentiation arrests at the level of the common myeloid progenitor when CEBPA is deleted<sup>10,11</sup>. In spite of the established functional interdependence of CSF3R and CEBPA during normal hematopoiesis, the mechanism by which oncogenic mutations in these two genes interact to drive AML remains unknown.

Herein, we demonstrate that CSF3R and CEBPA cooperate to produce a highly proliferative immature myeloid leukemia in mice that phenocopies the human disease transcriptionally and morphologically. We establish that mutant CSF3R drives both proliferation and differentiation of maturing myeloid cells. In contrast, mutant CEBPA selectively blocks myeloid differentiation through

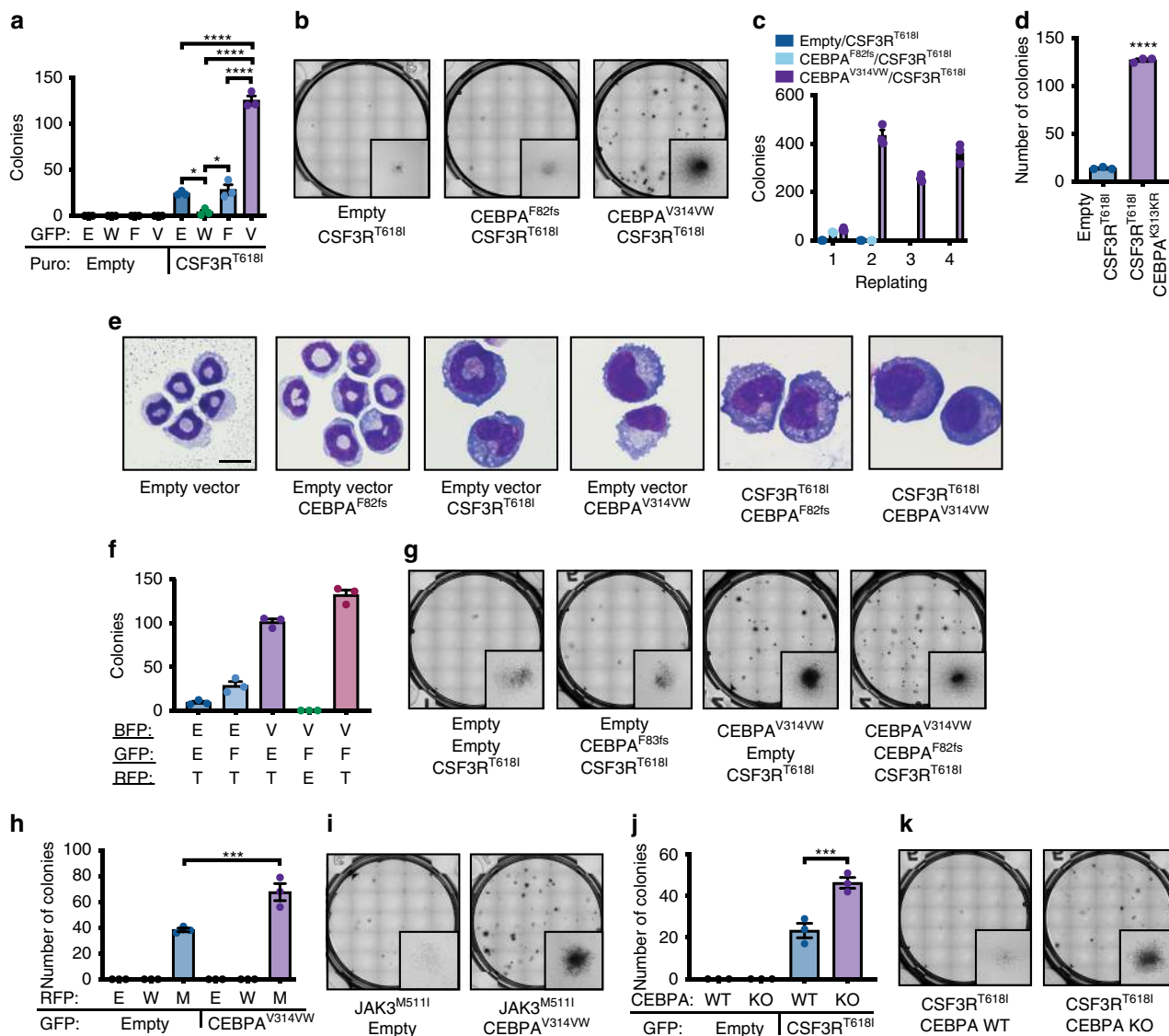
inactivation of myeloid lineage enhancers. We further demonstrate that mutant CEBPA must occur prior to mutant CSF3R in order for AML to initiate. This confirms predictions based on clinical sequencing and provides a novel model to study oncogene order.

## Results

**Cooperativity Between CSF3R<sup>T618I</sup> and CEBPA Mutations.** We elected to study a representative N-terminal and C-terminal mutation (F82fs and V314VW), in combination with mutant CSF3R, as these specific mutations are known to co-occur in AML<sup>5</sup>. When expressed in mouse bone marrow cells via retroviral transduction, neither CEBPA<sup>F82fs</sup> nor CEBPA<sup>V314VW</sup> produced colonies in cytokine-free methylcellulose (Fig. 1a, b). As previously reported, CSF3R<sup>T618I</sup> produced a modest number of colonies in isolation<sup>9</sup>. The addition of CEBPA<sup>V314VW</sup>, but not CEBPA<sup>F82fs</sup>, dramatically augmented CSF3R<sup>T618I</sup>-driven colony production and produced indefinite replating (Fig. 1c). These results were confirmed with a second C-terminal CEBPA mutation CEBPA<sup>K313KR</sup> (Fig. 1d). As CSF3R and CEBPA are both known to impact myeloid differentiation, we assessed cell morphology after 3 days of culture under cytokine-free conditions. Cells harboring empty vector and cells with CEBPA<sup>F82fs</sup> alone demonstrated neutrophilic morphology (Fig. 1e). CSF3R<sup>T618I</sup> increased the abundance of cells with larger morphology and immature granulocytic morphology. Expression of CEBPA<sup>V314VW</sup> produced a blast-like morphology, as did the combination of CSF3R<sup>T618I</sup> with either CEBPA mutation. Thus, it appears that mutant CSF3R drives proliferation of granulocytic precursors. Both N- and C-terminal mutations appeared to impact granulocytic differentiation in combination with CSF3R<sup>T618I</sup>; however, CEBPA<sup>V314VW</sup> exerted a stronger effect, consistent with observed changes in colony production.

To determine how co-expression of both N- and C-terminal CEBPA mutations would alter the effects of CSF3R<sup>T618I</sup>, we expressed all three mutations simultaneously. In this context, the addition of CEBPA<sup>F82fs</sup> mildly augmented CFU formation beyond that seen with CSF3R<sup>T618I</sup> and CEBPA<sup>V314VW</sup> (Fig. 1f, g). To establish whether the N- and C-terminal CEBPA mutations were producing the expected shifts in the CEBPA p30/p42 isoform ratio, we expressed each construct alone as well as the combination of isoforms in K562 cells. Western blotting for CEBPA in these cell lines revealed expression of the p30 isoform when CEBPA<sup>F82fs</sup> was present and expression of both the p42 and p30 isoforms when CEBPA<sup>V314VW</sup> was present (Supplementary Fig. 1A). We further validated that both mutant CSF3R and CEBPA were expressed in a cell line derived from mouse bone marrow colony-forming assay (Supplementary Fig. 1A). To determine whether the observed synergy between CSF3R<sup>T618I</sup> and CEBPA<sup>V314VW</sup> was more broadly generalizable, we looked for other mutations that co-occur with CEBPA in AML, which like CSF3R<sup>T618I</sup>, also activate the JAK/STAT pathway. We identified a patient with CEBPA-mutant AML from a recently published data set that had a previously characterized activating mutation in JAK3 (JAK3<sup>M511I</sup>)<sup>12,13</sup>. Simultaneous introduction of these mutations promoted cytokine-independent colony growth in vitro (Fig. 1h, i). Finally, we wanted to determine whether loss of CEBPA function is sufficient for enhanced oncogenesis in the context of mutant CSF3R. We therefore transduced bone marrow from CEBPA knockout mice with CSF3R<sup>T618I</sup>, and found that they produced significantly more colonies than littermate control bone marrow (Fig. 1j, k).

To characterize gene expression changes driven by CSF3R<sup>T618I</sup> and CEBPA<sup>V314VW</sup>, we performed RNA-seq on hematopoietic progenitors transduced with CSF3R<sup>T618I</sup>, CEBPA<sup>V314VW</sup>, or the combination of both. After 48 h of in vitro culture, cells expressing both oncogenes (marked by GFP and RFP) and not expressing mature lineage antigens (Lin<sup>−</sup>) were sorted by FACS prior to RNA

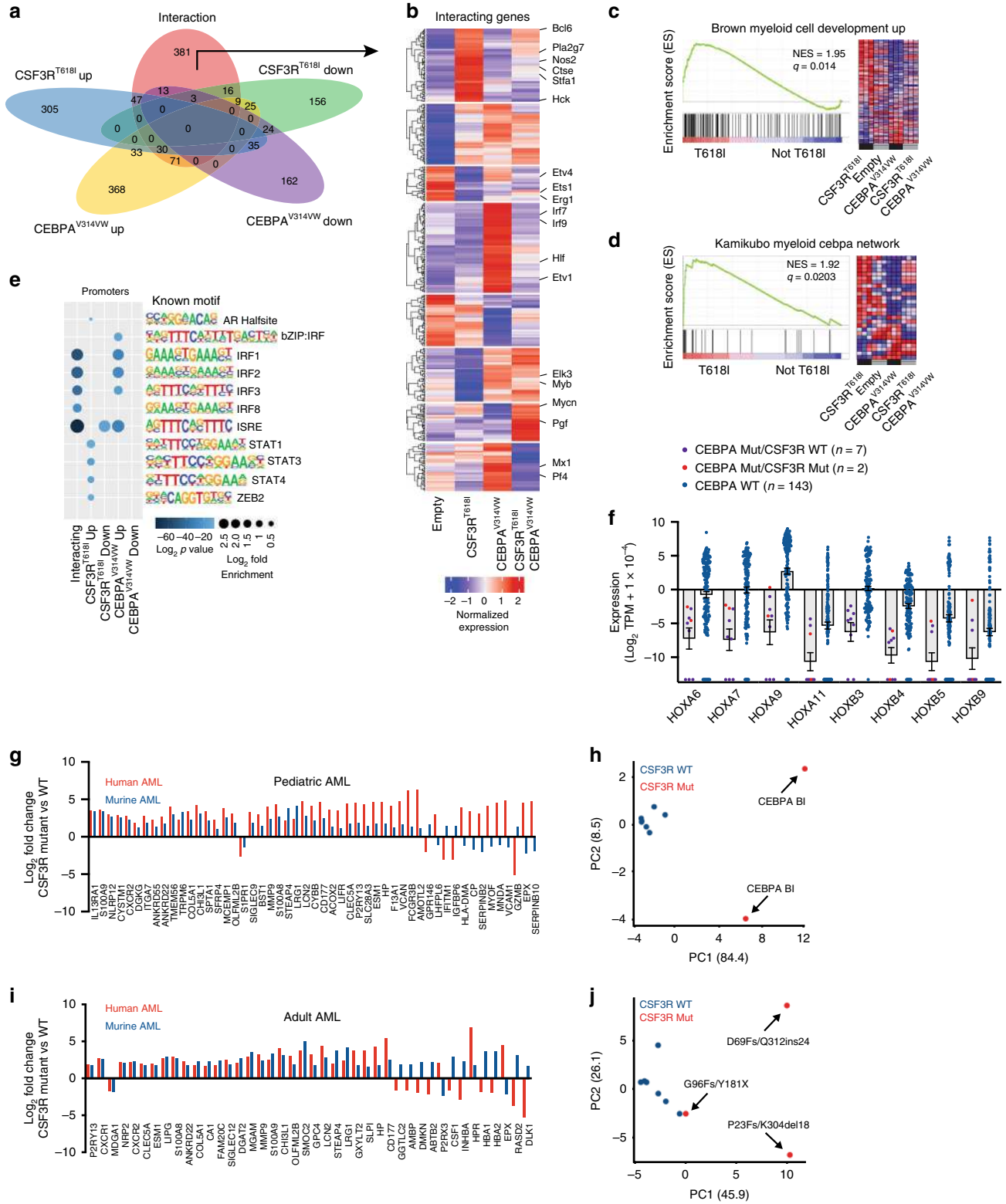


**Fig. 1** CSF3R<sup>T618I</sup> and CEBPA mutations cooperate in vitro. **a** Colony assay from mouse bone marrow transduced with CSF3R<sup>T618I</sup> and either CEBPA<sup>WT</sup> (W), CEBPA<sup>F82fs</sup> (F), or CEBPA<sup>V314VW</sup> (V) (*n* = 3/group). **b** Representative images for **a**. **c** Serial replating of colony assay in **a**. **d** Colony assay from mouse bone marrow transduced with CSF3R<sup>T618I</sup> and either empty vector or CEBPA<sup>K313KR</sup>. **e** Example images from cytopins of conditions in **a** grown for 72 h in cytokine-free liquid culture (scale bar represents 10 μm). **f** Colony-forming assay from mouse bone marrow transduced with the combination of CSF3R<sup>T618I</sup> (T), CEBPA<sup>F82fs</sup> (F), CEBPA<sup>V314VW</sup> (V), two oncogene combinations and empty vector controls. In all, 1700 cells were plated per condition. **g** Representative images for **f**. **h** Colony assay from mouse bone marrow transduced with Empty Vector (E) JAK3<sup>WT</sup> (W), JAK3<sup>M511I</sup> (M), or CEBPA<sup>V314VW</sup>. **i** Representative images of colony assay for **h**. **j** Colony assay from mouse bone marrow from CEBPA KO or Cre- littermate control bone marrow transduced with Empty Vector or CSF3R<sup>T618I</sup>. **k** Representative colony assay images for **j**. In all cases *n* = 3/group and values are represented as mean with error bars representing SEM. Significance of comparisons assessed by ANOVA with Sidak post-test \**p* < 0.05, \*\*\**p* = 0.001, \*\*\*\**p* < 0.0001). Source data are provided as a Source Data file.

seq (Supplementary Fig. 1B, C). To describe the pattern of transcriptional changes driven by each oncogene, we used a linear model with an interaction term (Fig. 2a, Supplementary Data File 1), and identified 683 genes that were significantly up or down-regulated in response to CSF3R<sup>T618I</sup> (*q* < 0.05, Log<sub>2</sub> fold change >1 or <-1). We also identified 773 genes up- or down-regulated in response to CEBPA<sup>V314VW</sup> expression. Additionally, there were 570 genes that demonstrated an interaction between CSF3R<sup>T618I</sup> and CEBPA<sup>V314VW</sup> (effect less or more than additive). These interacting genes demonstrated a variety of patterns of regulation across all four oncogene conditions as demonstrated by K-means clustering (Fig. 2b). Of particular interest was a cluster of genes that were strongly up-regulated by CSF3R<sup>T618I</sup>, but suppressed by co-expression of CEBPA<sup>V314VW</sup>, as this mirrored the pattern of

myeloid differentiation. This cluster contained genes such as *Nos2*, *Hck*, *Stf1a*, and *Pla2g7*, all of which are expressed in mature neutrophils<sup>14</sup>. Also of interest was the gene *Myb*, a known driver of myeloid oncogenesis<sup>15</sup>. CEBPA<sup>V314VW</sup> increased the expression of *Myb* irrespective of the presence of CSF3R, suggesting that *Myb* plays a role in driving CEBPA-mediated oncogenesis. To confirm this, we overexpressed *Myb* in mouse bone marrow cells with and without CSF3R<sup>T618I</sup> (Supplementary Fig. 1D, E). While *Myb* did not drive cytokine-independent colony formation in isolation, it dramatically increased colony formation in combination with CSF3R<sup>T618I</sup>, similar to that observed with CEBPA<sup>V314VW</sup>.

To identify enriched transcriptional programs, we performed Gene Set Enrichment Analysis (GSEA) comparing CSF3R<sup>T618I</sup> to all other conditions (Fig. 2c, d, Supplementary Fig. 1F). CSF3R<sup>T618I</sup>



dramatically up-regulated genes associated with myeloid differentiation, while CEBPA<sup>V314VW</sup> strongly down-regulated them. In addition, genes associated with the wild-type CEBPA network followed a similar pattern (Fig. 2d). This suggests that myeloid differentiation in response to CSF3R<sup>T618I</sup> is, at least in part, dependent on CEBPA. We next performed motif enrichment analysis at the promoters of each category of differentially expressed genes (Fig. 2e). This identified a strong enrichment of STAT-binding sites in the promoters of CSF3R<sup>T618I</sup> up-regulated genes.

This is consistent with the known role of STAT3 as a crucial transcription factor downstream of oncogenic CSF3R mutations<sup>9,16</sup>. To confirm that STAT3 activation still occurs in the context of CSF3R<sup>T618I</sup> and CEBPA<sup>V314VW</sup> co-expression, we performed a western blot on bone marrow cells harboring both mutant CSF3R<sup>T618I</sup> and CEBPA<sup>V314VW</sup> growing in liquid culture after isolation from colony assay. This confirmed an increase in phosphorylated STAT3 relative to normal bone marrow (Supplementary Fig. 1G). Given the results from the GSEA, we were



**Fig. 2** Mutant CEBPA blocks myeloid differentiation in response to CSF3R. **a** Venn diagram of differentially expressed genes from RNA-seq on lineage-negative mouse bone marrow transduced with empty vector, CSF3R<sup>T618I</sup>, CEBPA<sup>V314VW</sup> or the combination of oncogenes. **b** Hierarchical clustering of interacting genes. **c, d** Select enriched gene sets in CSF3R<sup>T618I</sup> vs. other categories NES = normalized enrichment score ( $q < 0.05$ ). GSEA  $p$  value calculated by empirical permutation test and FDR adjusted. **e** Motif enrichment at promoters of differentially expressed (DE) genes. Top five motifs per category with  $q < 0.05$  shown.  $P$  values generated via comparison to binomial distribution and FDR adjusted. **f** Expression of differentially expressed HOX genes in pediatric AML patients harboring CEBPA mutations. **g** Expression of genes differentially expressed in both murine and pediatric human CSF3R/CEBPA AML as compared with CEBPA-mutant CSF3R-WT AML. **h** PCA analysis of pediatric CEBPA mutant AML using convergent human-mouse gene set. Cases with biallelic CEBPA mutations are indicated by CEBPA-Bi. **i** Expression of genes differentially expressed in both murine and adult human CSF3R/CEBPA AML as compared with CEBPA-mutant CSF3R-WT AML. **j** PCA analysis of adult CEBPA mutant AML using convergent human-mouse gene set. Specific CEBPA mutations are indicated by the text. Source data are provided as a Source Data file.

somewhat surprised that CEBPA motifs were not detected at the promoters of any of the groups of DE genes, suggesting that CEBPA<sup>V314VW</sup> may impact CSF3R<sup>T618I</sup>-induced myeloid differentiation through binding to non-promoter regulatory regions. To determine, whether CEBPA binding was significantly associated with any of the categories of differentially expressed genes, we performed a permutation analysis using published ChIP-seq data from mouse granulocyte-macrophage progenitors (GMPs)<sup>17</sup>. This analysis revealed that the interacting genes, CSF3R<sup>T618I</sup> up genes, and CEBPA<sup>V314VW</sup> down genes were all found in closer proximity to CEBPA peaks than would be expected by random chance (Supplementary Fig. 1H).

To validate these gene findings in human AML, we examined RNA-seq data from the pediatric TARGET initiative<sup>18</sup>. A total of 152 patients were evaluated, including 7 patients with CEBPA mutations and 2 patients with both CSF3R and CEBPA mutations. Given the small number of patients, both monoallelic and biallelic CEBPA mutations were considered together. Differential gene expression revealed markedly decreased *HOX* gene expression in CEBPA mutant samples compared with CEBPA WT samples, a well-established finding in adult CEBPA mutant AML (Fig. 2f, Supplementary Data File 2)<sup>19</sup>. Comparison of CEBPA mutant/CSF3R WT and CEBPA mutant/CSF3R mutant patient samples revealed 913 differentially expressed genes (Supplementary Data File 2). Comparison of these differentially expressed genes to those identified in mouse revealed 52 ortholog pairs with differential expression driven by mutant CSF3R (Fig. 2g). Of these, 75% demonstrated concordant regulation between mouse and human. A similar analysis was performed on adult CEBPA/CSF3R mutant AML samples from the Leucegene cohort, which also demonstrated that the preponderance of orthologous gene pairs display concordant regulation (68%; Fig. 2i)<sup>4</sup>. Both human/mouse concordant gene sets were able to independently segregate CSF3R mutant from wild-type samples by PCA (Fig. 2h, j). As CEBPA is known to regulate CSF3R expression, we examined CSF3R expression in CEBPA WT and Mutant AML samples (Supplementary Fig. 1I)<sup>20</sup>. Both subsets of patients robustly expressed CSF3R and no significant difference in expression was observed on the basis of CEBPA mutational status. Thus, although wild-type CEBPA does regulate the activity of the CSF3R promoter, it appears that CSF3R expression is not dramatically altered in AML by the presence of putative loss-of-function mutations in CEBPA.

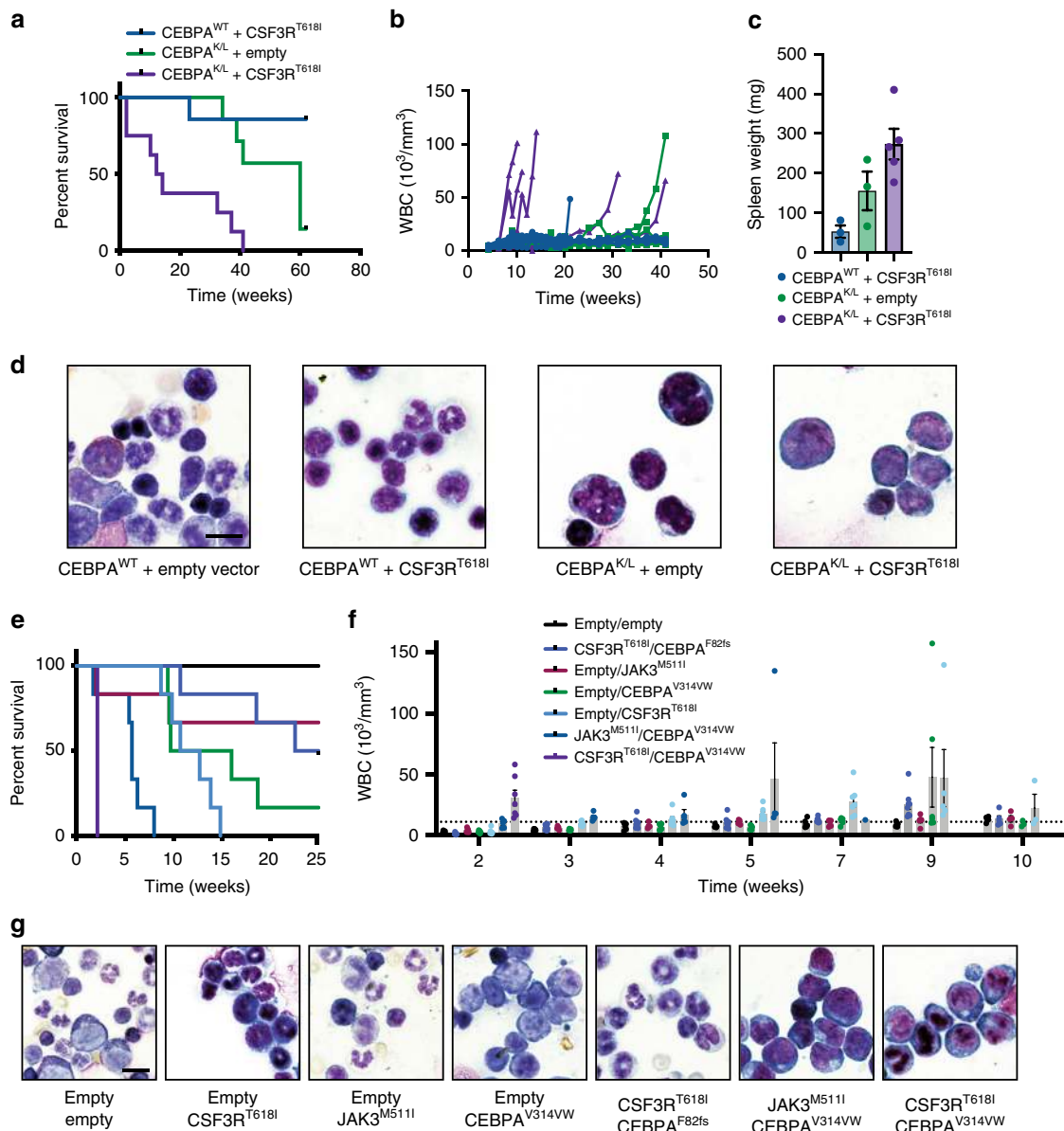
#### JAK/STAT activation and CEBPA dysregulation drive AML.

To establish whether oncogenic CSF3R and CEBPA mutations collaborate in vivo to produce AML, we performed murine bone marrow transplantation with fetal liver hematopoietic cells harboring compound heterozygous CEBPA mutations (CEBPA<sup>K/L</sup>) or wild-type CEBPA retrovirally transduced with CSF3R<sup>T618I</sup><sup>21</sup>. As previously reported, transplantation with CEBPA<sup>K/L</sup> cells alone leads to disease onset with a median latency of

approximately 1-year post transplant (Fig. 3a). In contrast, mice transplanted with CEBPA<sup>K/L</sup> cells harboring mutant CSF3R developed AML that was uniformly lethal by 1 year and occurred far more rapidly in the majority of cases. When leukemia developed, it was accompanied by leukocytosis, splenomegaly and morphologic myeloblasts in the bone marrow (Fig. 3b–d).

As CSF3R mutations appear to correlate most strongly with the presence of C-terminal CEBPA mutations in human AML, and our colony assay data presented in Fig. 1 demonstrate more potent oncogenic activity in vitro we investigated whether similar biology occurs in vivo. We therefore also performed retroviral bone marrow transplantation with retrovirally introduced CSF3R<sup>T618I</sup>, CEBPA<sup>F82fs</sup>, CEBPA<sup>V314VW</sup>, and JAK3<sup>M511I</sup> with corresponding single oncogene and empty vector controls. Mice receiving cells transduced with CSF3R<sup>T618I</sup> and CEBPA<sup>V314VW</sup> developed a myeloid leukemia that was uniformly lethal by day 14 post-transplant and was associated with leukocytosis and marked splenomegaly (Fig. 3e, f, Supplementary Fig. 2A). In the bone marrow, normal hematopoiesis was completely replaced by large cells with myeloblastic morphology (Fig. 3g, Supplementary Figs. 2B and 3A–C). Mice harboring CSF3R<sup>T618I</sup> alone or in combination with CEBPA<sup>F82fs</sup> developed disease with a long latency, associated with leukocytosis, variable splenomegaly, mature immunophenotype, and morphologic neutrophils in the peripheral blood and bone marrow (Fig. 3e, g). CEBPA<sup>V314VW</sup> alone resulted in a long-latency myeloid leukemia with immature histologic features at the time of ultimate disease development (Fig. 3e, g). Similar to CSF3R<sup>T618I</sup>, JAK3<sup>M511I</sup> in combination with CEBPA<sup>V314VW</sup> also produced a rapidly lethal myeloid leukemia with markedly reduced latency compared with expression of either mutation alone, again with the accumulation of marrow myeloblasts (Fig. 3e–g).

CSF3R mutations are associated with chronic neutrophilic leukemia (defined by an abundance of mature neutrophils) yet also accelerate AML formation in the context of mutant CEBPA. Therefore, we were interested in understanding how the presence of mutant CEBPA changes the oncogenic program of mutant CSF3R. Unfortunately, the AML that develops from CSF3R<sup>T618I</sup>/CEBPA<sup>V314VW</sup> co-expression results in a uniformly lethal AML when control mice are still pancytopenic. Comparison with untransplanted control mice reveals that CSF3R/CEBPA mutant AML cells are CD11b positive and GR-1 dim and completely replace the normal hematopoietic hierarchy (Supplementary Fig. 3A–C). However, to make a direct comparison, we transplanted syngeneic recipient mice with either 100,000 cells expressing CSF3R<sup>T618I</sup> alone or 10,000 cells expressing CSF3R<sup>T618I</sup> and CEBPA<sup>V314VW</sup> (Fig. 4a). This log-fold reduction in cell dose delayed the timing of disease onset to approximately 20 days, allowing for side by side comparison of these two groups. While CSF3R<sup>T618I</sup> alone mice demonstrated an abundance of mature neutrophils (marked by high levels of CD11b and GR-1), the leukemic blasts seen in CSF3R<sup>T618I</sup>/CEBPA<sup>V314VW</sup> mice exhibited lower levels of GR-1 staining consistent with an immature myeloid

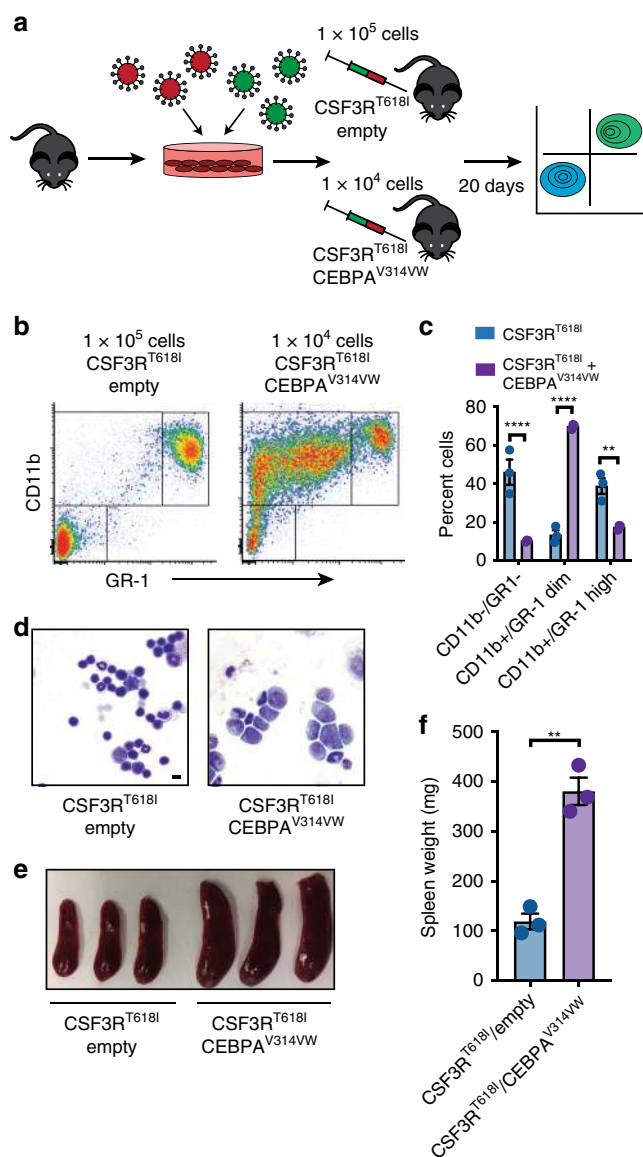


**Fig. 3** JAK/STAT activation and CEBPA mutations cooperate in vivo. **a** Survival of mice transplanted with CEBPA wild type or CEBPA<sup>K/L</sup> fetal liver cells transduced with empty vector or CSF3R<sup>T618I</sup> ( $n = 7\text{--}8/\text{group}$ ). **b** WBC counts from mice in **a**. **c** Spleen weights from moribund mice in **a** ( $n = 3\text{--}5/\text{group}$ ). **d** Example bone marrow smears for mice in **a** (scale bar represents 10  $\mu\text{m}$ ). **e** Survival of mice transplanted with 100,000 cells containing Empty Vector, JAK3<sup>M511I</sup> + Empty Vector, CEBPA<sup>V314VW</sup> + Empty Vector, JAK3<sup>M511I</sup> + CEBPA<sup>V314VW</sup>, CSF3R<sup>T618I</sup> + Empty vector, CSF3R<sup>T618I</sup> + CEBPA<sup>F82fs</sup>, or CSF3R<sup>T618I</sup> + CEBPA<sup>V314VW</sup> ( $n = 5\text{--}6/\text{group}$ ). **f** WBC counts of mice harboring mutation combinations in **e** ( $n = 5\text{--}6/\text{group}$  at experiment start). **g** Example bone marrow smears from mice in **e** (scale bar represents 10  $\mu\text{m}$ ). Values in bar graphs are represented as mean with error bars representing SEM. Source data are provided as a Source Data file.

phenotype (Fig. 4b–d, Supplementary Fig. 3D). Marked splenomegaly was observed in mice transplanted with CSF3R<sup>T618I</sup>/CEBPA<sup>V314VW</sup> cells but not in mice receiving CSF3R<sup>T618I</sup> alone cells (Fig. 4e, f). To establish whether the leukemic blasts generated by co-expression of CSF3R<sup>T618I</sup> and CEBPA<sup>V314VW</sup> were capable of indefinite self-renewal, we performed serial transplantation studies. These experiments revealed that the leukemia was capable of disease initiation in up to quaternary recipients (Supplementary Fig. 3E–G).

**CEBPA mutations disrupt myeloid lineage enhancers.** During normal hematopoietic development, CEBPA is responsible for establishing the enhancer landscape that permits myeloid differentiation<sup>22</sup>. We therefore hypothesized that CEBPA<sup>V314VW</sup>

blocks differentiation by inhibiting priming or activation of myeloid lineage enhancers. To test this, we utilized murine HoxB8-ER cells, which mimic GMPs and differentiate down the neutrophilic lineage to become CD11b and GR-1 positive after estrogen withdrawal<sup>23</sup>. In this model, expression of CSF3R<sup>T618I</sup> accelerated differentiation, while CEBPA<sup>V314VW</sup> blocked differentiation, as measured by CD11b and GR-1 expression (Supplementary Fig. 4A, B). The combination of mutations produced maturation arrest, similar to CSF3R/CEBPA mutant murine leukemia. We confirmed that neither oncogene was changing the expression of the other by qPCR (Supplementary Fig. 4C), consistent with our findings in human CEBPA mutant AML. Thus, HoxB8-ER cells expressing CSF3R<sup>T618I</sup> and CEBPA<sup>V314VW</sup> recapitulate the phenotype seen in our in vivo model.



**Fig. 4** CEBPA mutations block differentiation in response to mutant CSF3R. **a** Experimental strategy to directly compare the differentiation of cells harboring CSF3R<sup>T618I</sup> and CSF3R<sup>T618I</sup>/CEBPA<sup>V314VW</sup>. **b** Bone marrow flow cytometry comparing myeloid differentiation markers in mice transplanted under conditions in **a**. **c** Quantification of populations in **b**. **d** Example bone marrow histology from mice transplanted in **a** (scale bar represents 10 μm). **e** Spleens from mice in **a**. **f** Spleen weights from mice in **a**. Values are represented as mean with error bars representing SEM. Significance of comparisons assessed by Student's *T*-test or ANOVA with Sidak post-test as appropriate \*\**p* < 0.01, \*\*\*\**p* < 0.0001. Source data are provided as a Source Data file.

To study enhancer dynamics in the context of these mutations, we performed chromatin immunoprecipitation sequencing (ChIP-seq) for H3K4me1, H3K4me3, and H3K27ac in HoxB8-ER cells transduced with CSF3R<sup>T618I</sup>, CEBPA<sup>V314VW</sup>, both mutations, or empty vectors. We identified 42,124 active enhancers (defined as high H3K4me1, low H3K4me3, and high H3K27ac) across all four conditions (Supplementary Data File 3). To validate these findings in human AML, we mapped our enhancers to orthologous human genomic coordinates and assessed overlap of enhancers only active in the CSF3R<sup>T618I</sup>/CEBPA<sup>V314VW</sup> condition with those present in human AML

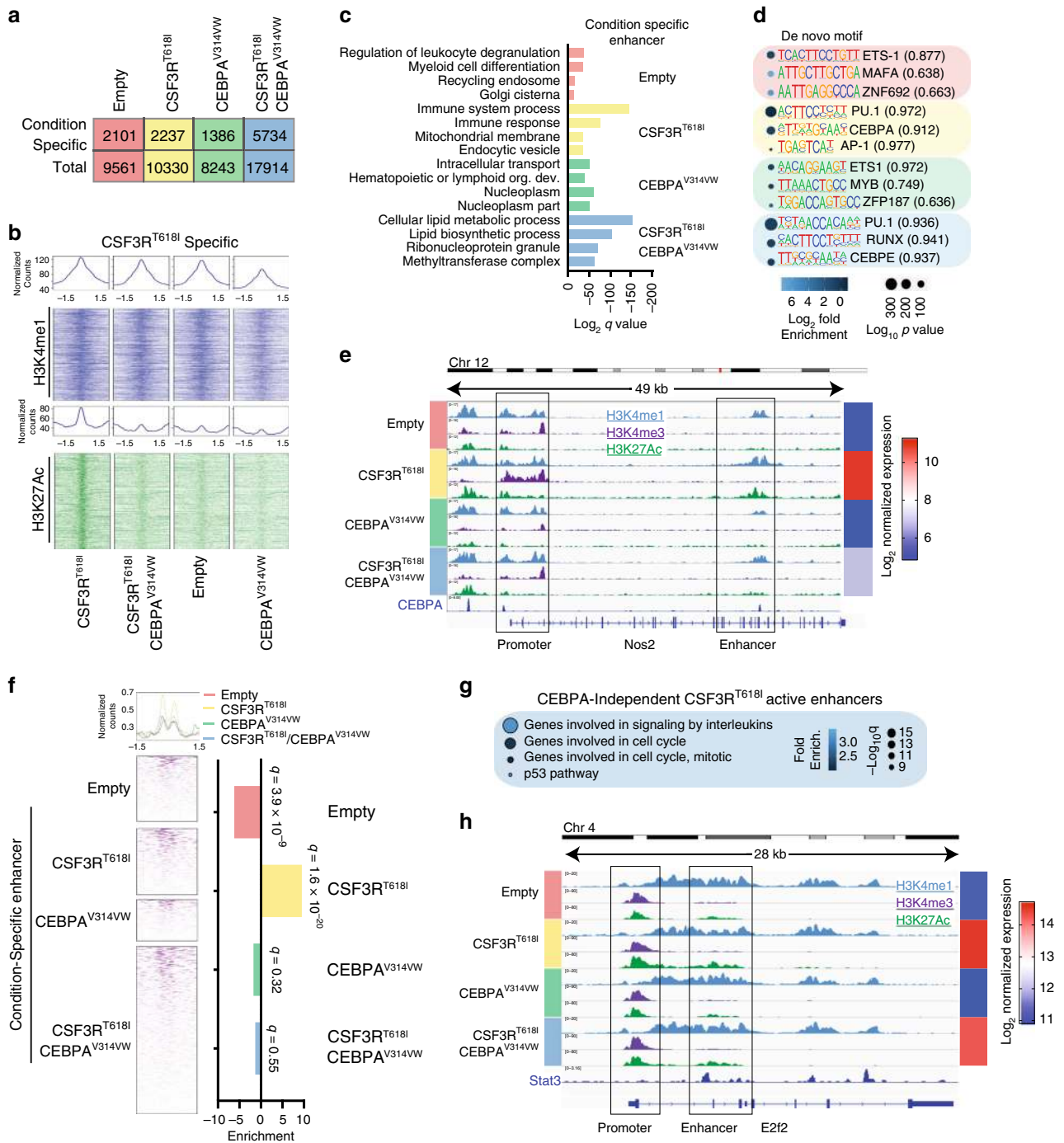
with mutant CEBPA<sup>24</sup>. We observed a significant enrichment of overlap between our mouse AML enhancers with those found in human CEBPA mutant AML, suggesting conserved biology (Supplementary Fig. 5A). Furthermore, our murine leukemic enhancers demonstrated less overlap than expected through random chance with those found in human CEBPA wild-type AML.

We next focused on enhancers that were active exclusively in one condition (Condition-Specific Enhancers, Fig. 5a, b, Supplementary Fig. 5B). Gene ontology analysis on the 2237 CSF3R<sup>T618I</sup>-specific enhancers revealed enrichment for terms associated with immune responses and phagocytosis, demonstrating that these genes are associated with the mature neutrophil phenotype (Fig. 5c, Supplementary Data File 4). To understand whether CEBPA<sup>V314VW</sup> blocks myeloid differentiation through disruption of CSF3R<sup>T618I</sup>-specific enhancers and associated gene expression, we performed microarray gene expression analysis on HoxB8-ER cells harboring each oncogene in isolation or the combination, and the expression of key genes was validated by qPCR (Supplementary Data File 5, Supplementary Fig. 5C). Genes associated with mature myeloid phenotypes, such as *Nos2*, *Hck*, *Bcl6*, and *Pla2g7*, displayed increased expression in the CSF3R<sup>T618I</sup> condition, and were repressed with co-expression of CEBPA<sup>V314VW</sup>. To globally assess whether CSF3R target genes were associated with CSF3R<sup>T618I</sup>-specific enhancers, we performed hierarchical clustering on genes that were differentially expressed in one or more treatment condition. We then examined enrichment of each condition-specific enhancer group across these clusters (Supplementary Fig. 5D, Supplementary Data File 5). This analysis demonstrated globally that activated enhancers are associated with increased expression of the nearest gene.

To identify possible regulators of enhancer activation, we performed transcription factor motif enrichment. This analysis demonstrated enrichment of CEBPA motifs in CSF3R<sup>T618I</sup>-specific enhancers, but not in any other group (Fig. 5d). This suggests that CEBPA<sup>V314VW</sup> blocks differentiation by disrupting the interaction of wild-type CEBPA with CSF3R<sup>T618I</sup>-specific enhancers. To provide further evidence for this hypothesis, we utilized published CEBPA ChIP-seq data from GMPs (the nearest normal cell to HoxB8 immortalized progenitors)<sup>17</sup>. We found strong enrichment of CEBPA peaks overlapping CSF3R<sup>T618I</sup>-specific enhancers, but not the other condition-specific enhancer groups (Fig. 5e, f). These data suggest that these differentiation-associated enhancers are regulated by direct CEBPA binding during normal hematopoiesis and the presence of mutant CEBPA prevents activation of these enhancers.

Our data suggest that differentiation in response to oncogenic CSF3R signaling requires the action of wild-type CEBPA at differentiation-associated enhancers. However, the proliferative effects of CSF3R<sup>T618I</sup> are likely independent of CEBPA. To identify the effectors of CSF3R-induced proliferation, we examined enhancers activated by CSF3R<sup>T618I</sup> irrespective of the presence of mutant CEBPA. Gene ontology analysis on this subset of enhancers identified multiple pathways associated with cell cycle progression (Fig. 5g). One likely driver of this CSF3R-induced proliferation is the transcription factor *E2f2* which is strongly induced by CSF3R<sup>T618I</sup> in the presence and absence of CEBPA<sup>V314VW</sup> (Fig. 5h). Correlation with published STAT3 ChIP seq data from murine myeloid cells reveals STAT3 binding to an enhancer at the *E2f2* locus that is activated by CSF3R<sup>T618I</sup> (Fig. 5h)<sup>25</sup>. Thus, it seems likely that key components of the proliferative program downstream of CSF3R<sup>T618I</sup> are driven via activation of enhancers associated with cell cycle progression, possibly in a STAT3-dependent manner.





**Fig. 5** CEBPA mutations disrupt activation of myeloid lineage enhancers. **a** Number of condition-specific enhancers identified in this study. **b** ChIP-seq heat maps for H3K4me1 and H3K27ac at +1.5 kb around the center of CSF3R<sup>T618I</sup>-specific enhancers, across treatment conditions. ChIP tracks display fold enrichment relative to corresponding input. **c** Gene ontology analysis for condition-specific enhancers. *P* values generated via comparison to binomial distribution and FDR adjusted. **d** De novo motif enrichment in condition-specific enhancers. Values in parenthesis correspond to motif match score. *P* values generated via comparison to binomial distribution and FDR adjusted. **e** The epigenetic landscape at the *Nos2* locus demonstrates condition-specific promoter and enhancer activation. **f** Assessment of CEBPA ChIP-seq peak overlap with condition specific enhancers by  $\chi^2$  test with Pearson residual values plotted. Adjusted *p* values (by the method of Holm-Bonferroni) are displayed above bars. **g** Gene ontology analysis for CEBPA-independent CSF3R<sup>T618I</sup> active enhancers. **h** The epigenetic landscape at the *E2f2* locus demonstrates enhancer activation in response to CSF3R<sup>T618I</sup>. Source data are provided as a Source Data file.

**CEBPA mutations must precede mutations in CSF3R.** Enhancer priming and activation precede promoter activation, suggesting that epigenetic changes at enhancers must occur as early events<sup>26</sup>. Clinical sequencing data demonstrate that CEBPA mutations frequently occur at higher variant allele frequencies

than CSF3R mutations<sup>4,5</sup>. This has led to the prediction that CEBPA mutations occur early in disease development. Conversely, however, it is possible that CEBPA mutations could be acquired late during blast-crisis transformation of chronic neutrophilic leukemia. To evaluate the impact of mutation order on

AML initiation, we developed a Cre-inducible retroviral vector. We paired this vector with hematopoietic cells from Rosa26<sup>ERT2</sup>-Cre mice, in which recombination is activated via tamoxifen administration (Fig. 6a). We co-expressed this inducible vector with a constitutive vector, enabling the study of two distinct mutational orders of acquisition: CSF3R<sup>T618I</sup>-first and CEBPA<sup>V314VW</sup>-first. We first assessed the impact of mutation order on bone marrow colony formation by plating both ordered combinations in the presence of 4-hydroxytamoxifen (4-OHT), thus delaying the expression of the second oncogene by approximately 48 h. Strikingly, we found that CSF3R<sup>T618I</sup>-first produced far fewer colonies than CEBPA<sup>V314VW</sup>-first (Fig. 6b, c). Numerically and morphologically, CSF3R<sup>T618I</sup>-first colonies were closer to CSF3R<sup>T618I</sup>-only. This finding was not due to differential expression of CEBPA from the inducible and constitutive vectors, as both showed equivalent CEBPA expression (Fig. 6d).

To establish the transcriptional profile of these two distinct mutation orders, we performed RNA-seq on lineage negative mouse bone marrow transduced with CSF3R<sup>T618I</sup>-only, CSF3R<sup>T618I</sup>-first, and CEBPA<sup>V314VW</sup>-first, cultured for 48 h with 4-OHT (Supplementary Fig. 6A, B, Supplementary Data File 6). Unsupervised clustering by Euclidean distance revealed that CSF3R<sup>T618I</sup>-only and CSF3R<sup>T618I</sup>-first cluster together and are globally distinct from CEBPA<sup>V314VW</sup>-first (Fig. 6e). The expression of differentiation-associated genes (*Nos2*, *Hck*, *Bcl6*, *Plg2g7*) in CSF3R<sup>T618I</sup>-first cells was intermediary between CSF3R<sup>T618I</sup>-only and CEBPA<sup>V314VW</sup>-first (Fig. 6f, Supplementary Data File 6). GSEA revealed enrichment for numerous signatures associated with myeloid differentiation and inflammatory responses in CSF3R<sup>T618I</sup>-first cells (Fig. 6g, Supplementary Fig. 6C). CEBPA<sup>V314VW</sup>-first cells demonstrated enrichment for signatures associated with cell cycle progression and stem/progenitor cell phenotype. Taken together, these studies demonstrate a profound effect of mutation order on gene expression and hematopoietic phenotype in vitro.

We next explored the contribution of mutation order to disease latency and phenotype in vivo. Mice were transplanted with CSF3R<sup>T618I</sup>-first or CEBPA<sup>V314VW</sup>-first cells and allowed to recover for 4-weeks post-transplant. After 4-weeks, the second oncogene was induced with tamoxifen (Fig. 7a). CEBPA<sup>V314VW</sup>-first mice succumbed to lethal myeloid leukemia with a median survival of 3.5 weeks post-tamoxifen (Fig. 7b). CEBPA<sup>V314VW</sup>-first leukemia was associated with bone marrow and peripheral blood blasts with an immature immunophenotype as well as splenomegaly (Fig. 7c–g, Supplementary Fig. 7A, B). In contrast, CSF3R<sup>T618I</sup>-first mice displayed a differentiated myeloid immunophenotype when sacrificed at an early time point. Only one CSF3R<sup>T618I</sup>-first mouse developed leukemia at 7 weeks post-tamoxifen administration. Bone marrow cytology revealed a blast-like morphology; however, flow cytometry revealed increased GR-1 staining, indicative of a higher degree of myeloid maturation (Fig. 7c, e). Collectively, these data reveal that when CEBPA mutations are introduced after mutations in CSF3R, they are unable to fully block myeloid differentiation. Importantly, this impaired ability to block differentiation disrupts the development of AML in vivo.

## Discussion

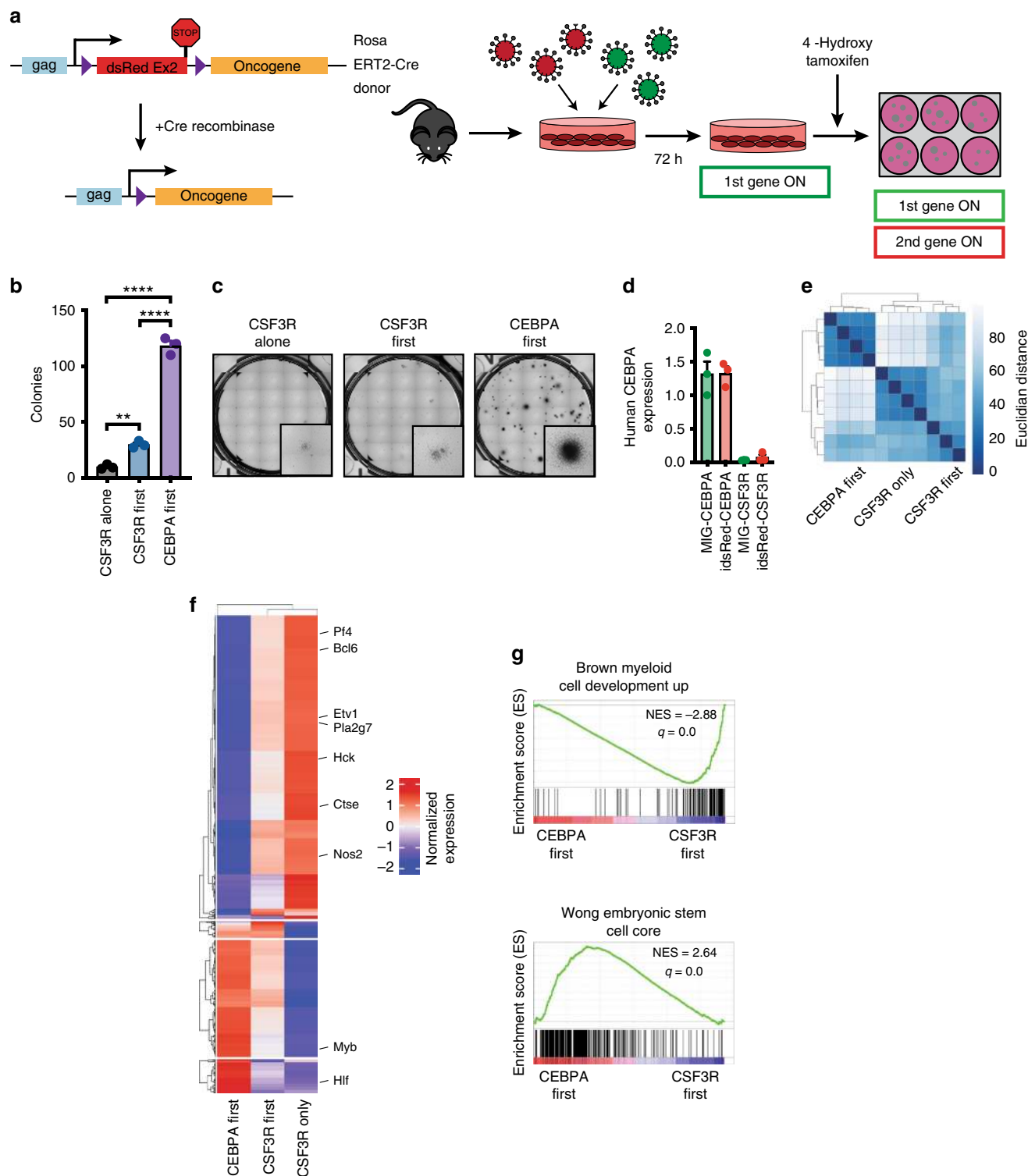
Our study adds to a growing body of data demonstrating that enhancer biology is integral to the development of hematologic malignancies. While AML is very heterogeneous from a genomic standpoint, recent work demonstrates that there are only a few epigenetic disease subtypes<sup>24,27</sup>. Although multiple global epigenetic regulators are recurrently mutated in AML, these have little impact on the organization of the epigenome<sup>24,27</sup>. Instead,

mutations in lineage determining transcription factors are a major determinant of clustering. Thus, understanding the epigenetic dysfunction associated with transcription factor mutations in AML may provide broad insight into therapeutic approaches.

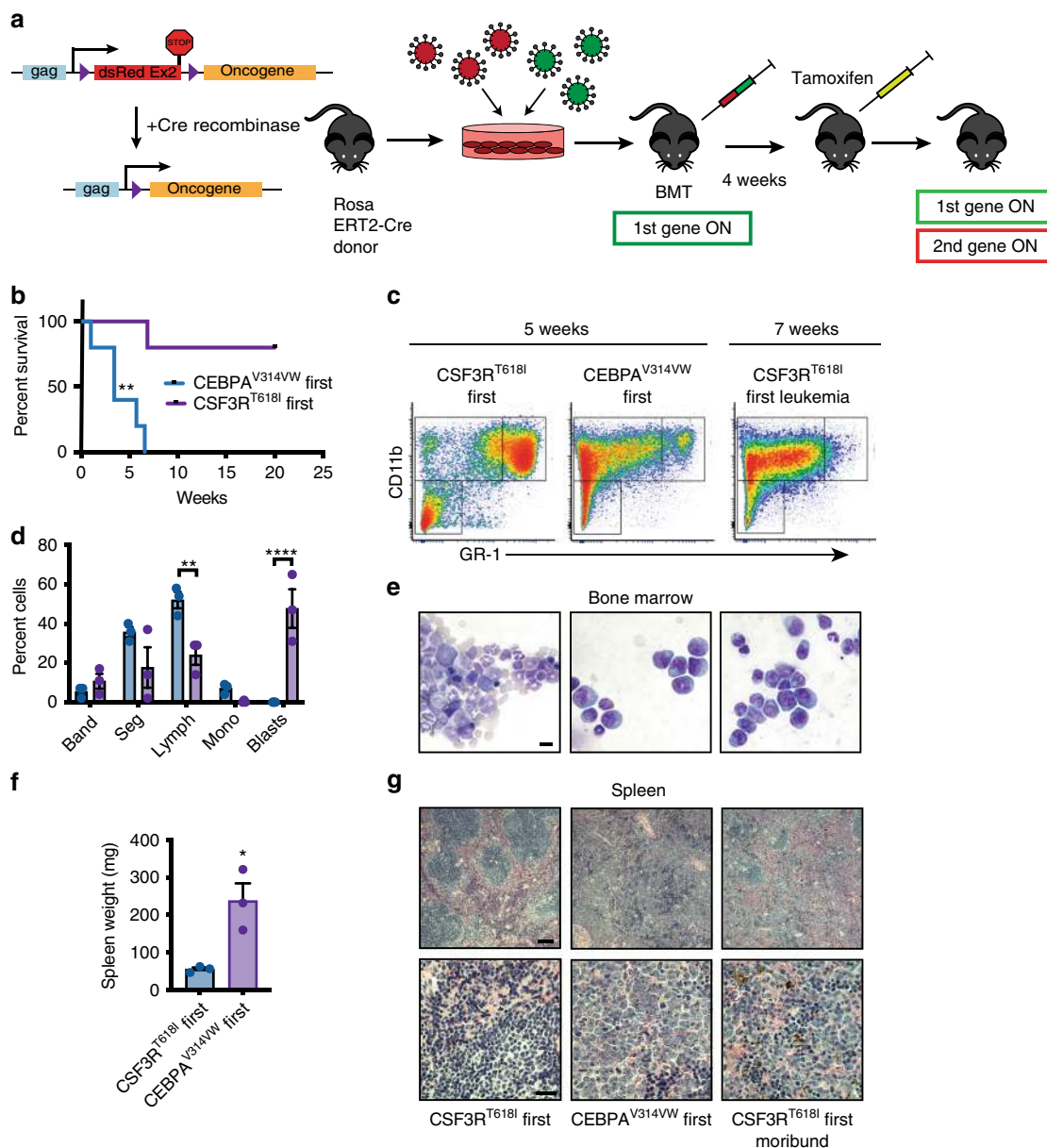
N- and C-terminal CEBPA mutations exert distinct biological roles in leukemia initiation<sup>21,28</sup>. While it is clear that reduced CEBPA activity is potentially oncogenic, CEBPA knockout mice fail to develop AML<sup>29</sup>. Similarly, mice with homozygous C-terminal mutations (putatively loss of function) develop leukemia with an exceedingly long latency<sup>21</sup>. Interestingly, the phenotype of these leukemias is erythroid rather than myeloid. This is consistent with data from MLL-rearranged leukemias, where CEBPA is requisite for entrance into the myeloid lineage and leukemia initiation<sup>30,31</sup>. As the majority of patients with CEBPA mutant AML harbor combined N- and C-terminal CEBPA mutations, it is possible that the N-terminal CEBPA mutation provides sufficient residual myeloid differentiation potential to enter the myeloid lineage and initiate AML. This relationship becomes even more complex when CSF3R mutations are considered. While the majority of CSF3R mutations occur in CEBPA-bi cases there are a number of CSF3R mutant CEBPA C-terminal monoallelic cases as well<sup>5</sup>. Our data demonstrate that CSF3R<sup>T618I</sup> potentially synergizes with biallelic CEBPA mutations to induce AML. Our retroviral model best represents monoallelic cases (with the caveats of ectopic expression), and demonstrates that C- but not N-terminal CEBPA mutations exhibit synergy with mutant CEBPA confirming the high degree of clinical association between these mutations. It is possible that mutant CSF3R provides both a myeloid commitment signal and proliferative advantage, thus rendering N-terminal CEBPA mutations less crucial for AML initiation. Going forward, it will be crucial to define the biological interaction with mutant CSF3R and N- and C-terminal mutations present in the endogenous locus to understand whether these mutant forms of CEBPA demonstrate differential binding or recruitment of cofactors to critical differentiation-associated enhancers.

Although we focused on differentiation-associated enhancers driven by the CSF3R<sup>T618I</sup>, our study revealed a second set of enhancers that were activated exclusively in the presence of both mutant CSF3R and CEBPA. Both subsets of enhancers demonstrated strong enrichment of PU.1 motifs, consistent with the known role of this transcription factor in driving myeloid development. In normal hematopoiesis, PU.1 and CEBPA cooperate to open myeloid lineage enhancers with PU.1 performing pioneering function in early progenitors and CEBPA assuming this role late<sup>22</sup>. Thus, it is likely that certain early myeloid lineage enhancers can be activated by CSF3R<sup>T618I</sup> in a CEBPA-independent manner via PU.1. Another interesting finding was the enrichment of RUNX motifs in enhancers activated only in the presence of both oncogenes. The RUNX family of transcription factors are critical to normal hematopoietic development, and in addition to being the targets of chromosomal translocations in core-binding factor AML, also frequently harbor point mutations in AML<sup>32</sup>. Interestingly, Runx1 haploinsufficiency leads to hypersensitivity to G-CSF, suggesting a negative feedback role<sup>33</sup>. The role of RUNX transcription factors in driving CEBPA/CSF3R mutant AML is an interesting area for future work.

We present the first direct evidence that the order in which oncogenic mutations occur is a major determinant of leukemia development (Fig. 8). Our finding that myeloid differentiation blockade can only occur with a distinct mutational order may also be a broadly conserved mechanism that applies to Class I and Class II mutation pairings. As nearly all prior studies have investigated mutation cooperation in the setting of simultaneous introduction, it is likely that important aspects of



**Fig. 6** CEBPA mutations must precede mutations in CSF3R to block differentiation. **a** Diagram of order of acquisition system. Oncogenes in the MIG (GFP+) vector tagged with GFP are constitutively expressed, and oncogenes in the idsRed vector are expressed only after Cre-mediated recombination. Expression of the second oncogene can be induced by culture in 4-OHT. **b** Colony assay from mouse bone marrow transduced with CEBPA First (MIG-CEBPA<sup>V314VW</sup> + idsRed-CSF3R<sup>T618I</sup>) CSF3R First (MIG-CSF3R<sup>T618I</sup> + idsRed-CEBPA<sup>V314VW</sup>) and CSF3R only (MIG-CSF3R<sup>T618I</sup> + idsRed Empty) and plated in 4-OHT ( $n = 3$ /group). **c** Representative images of colony assay. **d** Human CEBPA and CSF3R expression in mouse bone marrow expressing MIG-CEBPA<sup>V314VW</sup>, idsRed-CEBPA<sup>V314VW</sup>, MIG-CSF3R<sup>T618I</sup> or idsRed-CSF3R<sup>T618I</sup> measured by TaqMan quantitative PCR ( $n = 3$ /group). **e** Clustering by Euclidian distance for RNA sequencing performed on lineage negative mouse bone marrow expressing CEBPA-first, CSF3R-first, or CSF3R-only ( $n = 3$ -4/group). **f** K-means clustering of top 750 differentially expressed genes per pairwise comparison ( $q < 0.05$ ,  $\log_2$ Fold Change  $< -1$  or  $> 1$ ) ( $n = 3$ -4/group). **g** Representative enriched gene sets from GSEA performed on samples from **e** ( $n = 3$ -4/group). GSEA  $p$  value calculated by empirical permutation test and FDR adjusted. In all cases, values are represented as mean with error bars representing SEM. \*\* $p < 0.01$ , \*\*\*\* $p < 0.0001$ , as measured by ANOVA with Sidak's post-test. Source data are provided as a Source Data file.



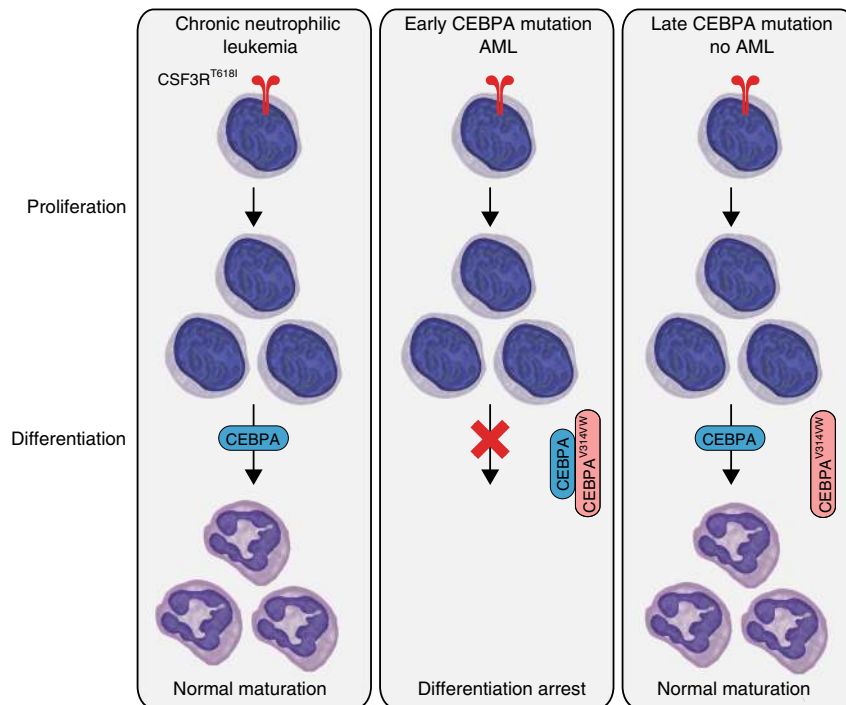
**Fig. 7** CEBPA mutations must precede mutations in CSF3R for AML initiation. **a** Diagram of order of acquisition system. Oncogenes in the MIG vector tagged with GFP are constitutively expressed and oncogenes in the idsRed vector are expressed only after Cre mediated recombination. Expression of the second oncogene can be induced by administration of tamoxifen. Mice were administered 75,000 GFP/RFP-positive cells. **b** Survival reported as time from day 1 of tamoxifen induction ( $n = 5/\text{group}$ ). Statistical significance calculated by Log Rank test,  $**p < 0.01$ . **c** Expression of CD11b and GR-1 in bone marrow at 5 weeks in moribund CEBPA<sup>V314VW</sup>-first mice, healthy CSF3R<sup>T618I</sup>-first mice or at 7-weeks in leukemic CSF3R<sup>T618I</sup>-first mice (only one animal represented by CSF3R-first leukemia, others are representative of 3/group). **d** Manual differentials of peripheral blood from mice 5 weeks after tamoxifen induction ( $n = 3/\text{group}$ ),  $**p < 0.01$ ,  $****p < 0.0001$ . **e** Representative bone marrow histology (scale bar represents 10  $\mu\text{m}$ ). **f** Spleen weight of mice sacrificed at 5 weeks post tamoxifen treatment,  $*p < 0.05$  as calculated by Student's *T*-test **g** Representative spleen histology (scale bar represents 100  $\mu\text{m}$  for upper panels and 20  $\mu\text{m}$  for lower panels). In all cases, values are represented as mean with error bars representing SEM. Source data are provided as a Source Data file.

order-dependent disease biology have not yet been discovered. If reprogramming of the lineage-specific enhancer repertoire is a common initiating event in AML, this creates even more impetus for the development of treatment strategies targeting these epigenetic pathways.

Our finding that epigenetic dysfunction is an obligate early event is likely to be generalizable to other forms of cancer. In renal cell carcinoma, loss of the tumor suppressor von Hippel-Lindau (VHL) is widely regarded as an initiating event and associated with dramatic changes in global DNA

methylation, potentially impacting subsequently acquired signaling mutations<sup>34</sup>. In breast cancer, mutation order is an important determinant of cancer phenotype, with luminal-type tumors demonstrating early loss of PTEN while basal-type tumors display early p53 mutation<sup>35</sup>. In colon cancer, the canonical APC mutations decrease DNA methylation through upregulation of demethylases, which potentially alters the impact of RAS mutations acquired later in disease evolution<sup>36</sup>. Order-dependent mutational phenotypes were recently reported in myeloproliferative neoplasms, where the order of mutations in





**Fig. 8** CEBPA mutations initiate AML through disruption of myeloid lineage enhancers. During normal myeloid differentiation, GCSF/CSF3R signaling drives both differentiation and proliferation of myeloid precursors. The differentiative program downstream of GCSF/CSF3R signaling is dependent on CEBPA acting at myeloid lineage enhancers. Mutant CEBPA blocks activation of these enhancers preventing transcription of differentiation-associated genes. When CSF3R mutations are introduced early, they activate transcription using the native enhancer repertoire initiating differentiation. Importantly this differentiation is insensitive to subsequent introduction of mutant CEBPA.

JAK2 and TET2 are important determinants of disease phenotype<sup>37</sup>. When TET2 precedes JAK2, patients are more likely to present with polycythemia vera. In contrast, patients with JAK2-first disease were more likely to have essential thrombocytosis. Thus, it appears that preceding TET2 mutation provides an epigenetic context supporting erythroid development for subsequently-acquired JAK2 mutations.

In summary, we describe the epigenetic mechanism by which mutant CEBPA and CSF3R interact to drive AML development. Our study demonstrates that a subset of differentiation-associated enhancers are dysregulated by mutant CEBPA, preventing normal myeloid maturation. Critically, this epigenetic dysregulation must occur as the initial event otherwise AML does not develop. These differentiation-associated enhancers represent a promising novel therapeutic target in this poor prognosis molecular subtype of AML.

## Methods

**Mice.** C57BL/6J mice (#000664), Balb/cJ mice (#000651), Rosa26 ERT<sub>2</sub>-Cre mice (#008463), MX-1 Cre mice (#003556), and CEBPA<sup>Flox/Flox</sup> mice (#006447) were obtained from The Jackson Laboratories. Female mice were used between 6 and 20 weeks, and age and weight matched in all experiments. Tamoxifen was dissolved in corn oil (Sigma) and injected intraperitoneally at 75 mg/kg/day for 5 days. Poly I:C (Sigma) was dissolved in PBS at 12 mg/kg on days 1, 3, and 5. All experiments were conducted in accordance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals, and approved by the Institutional Animal Care and Use Committee of Oregon Health & Science University (Protocol #TR01\_IP00000482).

**Cell Lines.** 293T17 cells were obtained from ATCC and cultured in DMEM (Gibco) with 10% fetal calf serum (FCS; HyClone). Murine HoxB8-ER cells were a generous gift from David Sykes (Massachusetts General Hospital, Boston, MA) and cultured in RPMI (Gibco) with 10% FCS and CHO-SCF cell conditioned media (final concentration ~100 ng/mL)<sup>27</sup>. Wild-type HoxB8-ER cells were differentiated in estrogen-free media<sup>23</sup>. Cell lines were tested monthly for mycoplasma

contamination and all lines utilized tested negative. Cell lines were authenticated by ATCC prior to shipment.

**Cloning and Retrovirus production.** The following plasmids were utilized: pMSCV-IRES GFP<sup>9</sup>, pMXs-IRES-Puro (Cell Biolabs Inc), pMSCV-IRES-mCherry FP (a gift from Dario Vignali, Addgene #52114), pMSCV-loxp-dsRed-loxp-eGFP-Puro-WPRE (a gift from Hans Clevers, Addgene #32702). The CSF3R<sup>T618I</sup> mutation was generated previously<sup>9</sup>. Full-length CEBPA and JAK3 cDNA was obtained from Genecopia. The CEBPA and JAK3 mutations were generated using the Quikchange Site Directed Mutagenesis Kit (Agilent) using the primers in Supplementary Table 1. To produce retrovirus, 293T17 cells were transfected with EcoPac helper plasmid (a gift from Dr. Rick Van Etten) and the appropriate transfer plasmid. Conditioned media was harvested 48–72 h after transduction.

**Western blotting.** K562 cells were transfected with lipofectamine 2000 according to the manufacturer's instructions. Normal mouse bone marrow cells were isolated fresh on the day of the experiment. CSF3R<sup>T618I</sup>/CEBPA<sup>V314VW</sup> cells were cultured in IMDM media supplemented with 20% FCS after isolation from methocult. In all cases,  $2 \times 10^6$  cells were used per condition. Cells were lysed in cell lysis buffer (Cell Signaling Technologies) containing complete mini protease inhibitor tablets (Roche). To pellet cellular debris, the lysates were spun at 12,000 r.p.m., 4 °C for 10 min, and subsequently mixed with 3× SDS sample buffer (75 mmol/L Tris (pH 6.8), 3% SDS, 15% glycerol, 8% β-mercaptoethanol, and 0.1% bromophenol blue). Samples were incubated at 95 °C for 5 min and run on Criterion 4 to 15% Tris-HCl gradient gels (Bio-Rad). Gels were transferred to PVDF membranes and blocked in Tris-buffered saline with Tween (TBST) with 5% bovine serum albumin (BSA). Blots were probed with the antibodies listed in Supplementary Table 3. Horseradish peroxidase-conjugated secondary antibodies against mouse IgG and rabbit IgG (CS) were used followed by imaging of the blots on a BioRad Touch Gel Doc (BioRad).

**Retroviral transduction.** Retroviral conditioned media was produced using 293T17 cells transduced with packaging plasmid and the appropriate transfer plasmid. Mouse bone marrow cells were harvested from healthy syngeneic donors at age 6–12 weeks and spinnoculated on two successive days with supernatant in the presence of polybrene<sup>9,16</sup>. Cells expressing GFP or GFP/mCherry were sorted by FACS using an FACS Aria III sorter (BD). For colony assay, 2500 sorted cells were used per replicate. For bone marrow transplantation experiments, donor mice were treated with 100 mg/kg 5-FU 5 days prior to harvest. In all cases, 200,000 cells were administered, with the balance of non-transduced cells as fresh syngeneic

bone marrow. Survival endpoints for mutation comparison studies included WBC count >100, weight loss >20% initial weight, and moribund appearance.

**Real-time PCR.** RNA was extracted using RNeasy or RNeasy micro kit (Qiagen). cDNA was synthesized using a High Capacity cDNA Synthesis kit (ThermoFisher). Real-time PCR was performed using the QuantStudio7 Real Time PCR System (ThermoFisher) and Taqman primer probes (ThermoFisher). For Taqman low-density arrays, array cards were custom printed by the manufacturer (ThermoFisher). Taqman primer probes used in this study are listed in Supplementary Table 2.

**Flow cytometry.** The antibodies listed in Supplementary Table 3 were utilized for FACS according to the manufacturer instructions. Stained cells were analyzed on a FACSAria III flow cytometer (BD).

**RNA sequencing.** Bone marrow from C57BL/6 mice was harvested and retrovirally transduced with the indicated oncogene combinations as described above. Cells were stained with a lineage cocktail, and lineage-negative GFP/mCherry-double-positive cells were sorted. For order of acquisition experiments, cells were cultured in cytokine-free I20 media for 48 h after sorting in the presence of 500 nM 4-Hydroxytamoxifen. RNA was extracted using the RNeasy micro kit (Qiagen). cDNA libraries were constructed using the SMARTer universal low input RNA kit (Clontech) and sequenced using a HiSeq 2500 Sequencer (Illumina) 100 bp SR.

**Microarray.** Labeled target, sscDNA was prepared from 12 RNA samples using the GeneChip™ WT-Plus protocol (ThermoFisher). Prior to cDNA synthesis and amplification, sample order was randomized. Amplified and labeled cDNA target samples were hybridized to a GeneChip Clarion S-Mouse array (ThermoFisher). Image processing was performed using Affymetrix Command Console (AGCC) v.3.1.1 software and expression analysis was performed using Affymetrix Expression Console. Differential expression analysis was performed using the Transcript Analysis Console 4.0 (ThermoFisher).

**Chromatin-immunoprecipitation sequencing (ChIP-seq).** Twenty-five million cells per condition (Empty vector, CSF3R<sup>T618L</sup>-only, CEBPA<sup>V314VW</sup>-only, and CSF3R<sup>T618L</sup> + CEBPA<sup>V314VW</sup>) were fixed with 1% formaldehyde and quenched by glycine. Cells were resuspended in lysis buffer (0.1% SDS, 0.5% Triton X-100, 20 mM Tris-HCl pH = 8.0, 150 mM NaCl, 1× Proteinase inhibitor (Roche)), and sheared using the Bioruptor Pico sonicator (Diagenode). Reactions were rotated overnight at 4 °C with antibody (H3K4me1 (ab8895; Abcam), H3K4me3 (ab8580; Abcam), and H3K27ac (ab4729; Abcam), Table S10). Next, samples were rotated with Protein A/G Magnetic Beads (Pierce). Beads were washed with TBST buffer (3×), lysis buffer, and 2× TE pH = 8. Chromatin was eluted at room temperature in 1% SDS, 0.1 M NaHCO<sub>3</sub>. Crosslinks were reversed at 65 °C overnight, digested with RNase A and proteinase K, then purified by phenol–chloroform extraction. Sequencing libraries were generated using the NEBNext Ultra II DNA Library Prep Kit for Illumina (New England Biolabs). Libraries were sequenced using SE 75 bp Illumina NextSeq.

**RNA-seq analysis: murine samples.** Raw reads were trimmed with Trimmomatic<sup>38</sup> and aligned with STAR<sup>39</sup>. Differential expression analysis was performed using DESeq2 (ref. 40). Raw *p* values were adjusted for multiple comparisons using the Benjamini–Hochberg method.

**RNA seq analysis: TARGET pediatric AML samples.** RNA sequencing was performed on RNA collected from pediatric AML samples as described in the initial TARGET AML publication<sup>18</sup>. Raw reads were aligned with Kallisto with count tables produced using Tximport<sup>41,42</sup>. Differential expression analysis was performed using DESeq2 (ref. 40). Raw *p* values were adjusted for multiple comparisons using the Benjamini–Hochberg method. For comparison with mouse RNA seq, genes with differential expression between CSF3R WT and CSF3R<sup>T618L</sup> were mapped to orthologous human genes using ensembl BioMart. Genes with differential expression in both mouse and human were considered and compared by Log2 Fold change between CSF3R<sup>mutant</sup> and CSF3R<sup>WT</sup> conditions.

**Convergent gene expression analysis: Leucegene AML.** RPKM values from adult AML samples from the Leucegene cohort were downloaded using the following SRA accession numbers: GSE49642, GSE52656, GSE62190, GSE66917, GSE67039. Mouse genes with differential expression driven by CSF3R<sup>T618L</sup> were converted to human gene symbols as above. Genes demonstrating an absolute fold change of >2 in both datasets were compared by Log2 Fold change between CSF3R<sup>mutant</sup> and CSF3R<sup>WT</sup> conditions.

**Motif enrichment analysis.** The enrichment analysis for motifs was performed using HOMER<sup>43</sup> using the -findMotifs command in a 500 bp window upstream and 200 bp downstream of the transcriptional start site. De novo motifs at enhancers were identified using the -findMotifsGenome command in HOMER in a

±1000 bp region surrounding the peak center. De novo motifs were matched to their closest known motif and displayed with the alignment score (with 1 being a perfect match). Top five motifs with *p* < 1E-10 for each group of promoters or enhancers are displayed.

**Gene set enrichment analysis.** Gene set enrichment analysis using the GSEA software<sup>44</sup>. As CSF3R<sup>T618L</sup>-only and CSF3R<sup>T618L</sup>-first groups clustered together by Euclidian distance, they were combined and compared with CEBPA<sup>V314VW</sup>-first samples. Analysis was performed using the C2 collection from MSigDB. Permutations were performed by gene set and significance was set as an FDR adjusted *p* value of <0.05.

**Permutation analysis.** CEBPA ChIP-seq peaks from GMPs<sup>17</sup> were converted from the mm9 to mm10, using the liftOver tool from the UCSC genome browser. We used BEDTools<sup>45</sup> to randomly shuffle the location of all interacting genes, within their original chromosomes. After each shuffle, we used the closest-features sub-command in BEDOPS<sup>46</sup> to calculate the distance between the new position of all shuffled genes and the closest CEBPA ChIP-seq.

**ChIP-seq analysis.** Reads were aligned to the mouse reference genome (mm10) using bwa 0.7.12 (ref. 47) with default single end settings. Low mapping alignments were removed with samtools (MAPQ <30)<sup>48</sup>. For quality control purposes, PCA was performed using Deeptools<sup>49</sup> which showed separation of signal by oncogene condition. Next, MACS2 2.1.1 (ref. 50) was used to predict significant peaks of ChIP-seq enrichment relative to the appropriate input controls and generate fold enrichment tracks. We used the ChromHMM software<sup>51</sup> to characterize and annotate the genomes of each treatment group according to six chromatin states, based on different combinations of H3K4me1, H3K4me3, and H3K27ac marks. Enhancers were identified through the presence of H3K4me1 and H3K27ac, absence of H3K4me3. Enhancers less than 500 bp apart were merged. The closest gene to each enhancer was identified using BEDOPS<sup>46</sup>. In order to remove potential false positives due to proximity to large H3K27ac peaks, we removed treatment-specific putative enhancers without a H3K27ac peak summit.

**Gene Ontology Analysis for ChIP-seq.** Gene Ontology Analysis for histone mark ChIP-seq performed by Genomic Regions Enrichment of Annotations Tool<sup>52</sup> using the basal plus extension model to annotate enhancer coordinates with nearby genes.

**Correlation of gene expression with enhancer activation.** All differentially expressed genes as assessed by microarray analysis utilized for unsupervised hierarchical clustering analysis with a *k*-means = 8. Each active enhancer was then annotated to the nearest gene. Condition-specific enhancers associated with each cluster of differentially expressed genes were counted and enrichment assessed as described under statistics.

**Overlap of CEBPA peaks with condition-specific enhancers.** We used publicly available genome-wide positions of all GMP CEBPA ChIP-seq peaks<sup>17</sup>. CEBPA peaks overlapping each group of condition-specific enhancers were identified using the mergePeaks command in HOMER with the -cobound option. CEBPA peaks associated with each group of condition-specific enhancers were counted and enrichment assessed as described below under statistics.

**Enhancer overlap using BEDTools.** Published human CEBPA AML DNase sensitivity peaks<sup>27</sup> and enhancers<sup>22</sup> were compared with mouse CSF3R/CEBPA-specific enhancers. The coordinates for mouse enhancers converted to human coordinates using the UCSC liftOver tool. Overlap was assessed using the BEDTools fisher test<sup>45</sup>. Enrichment of overlap is detected using a Fisher's exact test with a significant right-sided *p* value while depletion is detected with a significant left-sided *p* value.

**Quantification and statistical analysis.** Data are expressed as mean ± SEM. Statistical analysis was performed using Prism software (Version 7.0; Prism Software Corp.) or RStudio. Statistical analyses are described in the figure legends. All data were analyzed with either an unpaired Student's *t*-test, or ANOVA followed by post hoc analysis using a Sidak's corrected *t*-test. For Taqman-based array, microarray and RNA-seq data, *p* values were adjusted for repeated testing using a false discovery rate by the method of Benjamini–Hochberg<sup>53</sup>. For enhancer enrichment analyses, a  $\chi^2$  analysis was utilized with individual *p* values adjusted by the method of Holm–Bonferroni. Survival analysis was conducted using the method of Kaplan–Meier and statistical significance was assessed using a log rank test.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

**Data availability**

The accession numbers for all genomic data reported in this paper is GSE122166. The source data underlying Figs. 1–7 and Supplementary Figs. 1–6 are provided as a Source Data file.

Received: 14 May 2019; Accepted: 4 November 2019;

Published online: 29 November 2019

**References**

- Kelly, L. M. & Gilliland, D. G. Genetics of myeloid leukemias. *Annu. Rev. Genomics Hum. Genet.* **3**, 179–198 (2002).
- Avellino, R. & Delwel, R. Expression and regulation of C/EBP $\alpha$  in normal myelopoiesis and in malignant transformation. *Blood* **129**, 2083–2091 (2017).
- Porse, B. T. et al. Loss of C/EBP alpha cell cycle control increases myeloid progenitor proliferation and transforms the neutrophil granulocyte lineage. *J. Exp. Med.* **202**, 85–96 (2005).
- Lavallée, V.-P. et al. Chemo-genomic interrogation of CEBPA mutated AML reveals recurrent CSF3R mutations and subgroup sensitivity to JAK inhibitors. *Blood* **127**, 3054–3061 (2016).
- Maxson, J. E. et al. CSF3R mutations have a high degree of overlap with CEBPA mutations in pediatric AML. *Blood* **127**, 3094–3098 (2016).
- Su, L. et al. Mutational spectrum of acute myeloid leukemia patients with double CEBPA mutations based on next-generation sequencing and its prognostic significance. *Oncotarget* **9**, 24970–24979 (2018).
- Zhang, Y. et al. CSF3R mutations are frequently associated with abnormalities of RUNX1, CBF $\beta$ , CEBPA, and NPM1 genes in acute myeloid leukemia. *Cancer* **124**, 3329–3338 (2018).
- Su, L. et al. CSF3R mutations were associated with an unfavorable prognosis in patients with acute myeloid leukemia with CEBPA double mutations. *Ann. Hematol.* <https://doi.org/10.1007/s00277-019-03699-7> (2019).
- Maxson, J. E. et al. Oncogenic CSF3R mutations in chronic neutrophilic leukemia and atypical CML. *N. Engl. J. Med.* **368**, 1781–1790 (2013).
- Wang, W., Wang, X., Ward, A. C., Touw, I. P. & Friedman, A. D. C/EBP $\alpha$  and G-CSF receptor signals cooperate to induce the myeloperoxidase and neutrophil elastase genes. *Leukemia* **15**, 779–786 (2001).
- Iida, S., Watanabe-Fukunaga, R., Nagata, S. & Fukunaga, R. Essential role of C/EBP $\alpha$  in G-CSF-induced transcriptional activation and chromatin modification of myeloid-specific genes. *Genes Cells* **13**, 313–327 (2008).
- Degryse, S. et al. JAK3 mutants transform hematopoietic cells through JAK1 activation, causing T-cell acute lymphoblastic leukemia in a mouse model. *Blood* **124**, 3092–3100 (2014).
- Tyner, J. W. et al. Functional genomic landscape of acute myeloid leukaemia. *Nature* **562**, 526–531 (2018).
- Chen, J. et al. Intracellular PAF catabolism by PAF acetylhydrolase counteracts continual PAF synthesis. *J. Lipid Res.* **48**, 2365–2376 (2007).
- Klempnauer, K. H., Symonds, G., Evan, G. I. & Bishop, J. M. Subcellular localization of proteins encoded by oncogenes of avian myeloblastosis virus and avian leukemia virus E26 and by chicken c-myc gene. *Cell* **37**, 537–547 (1984).
- Fleischman, A. G. et al. The CSF3R T618I mutation causes a lethal neutrophilic neoplasia in mice that is responsive to therapeutic JAK inhibition. *Blood* **122**, 3628–3631 (2013).
- Hasemann, M. S. et al. C/EBP $\alpha$  is required for long-term self-renewal and lineage priming of hematopoietic stem cells and for the maintenance of epigenetic configurations in multipotent progenitors. *PLoS Genet.* **10**, e1004079 (2014).
- Bolouri, H. et al. The molecular landscape of pediatric acute myeloid leukemia reveals recurrent structural alterations and age-specific mutational interactions. *Nat. Med.* **24**, 103–112 (2018).
- Dufour, A. et al. Acute myeloid leukemia with biallelic CEBPA gene mutations and normal karyotype represents a distinct genetic entity associated with a favorable clinical outcome. *J. Clin. Oncol.* **28**, 570–577 (2009).
- Smith, L. T., Hohaus, S., Gonzalez, D. A., Dziennis, S. E. & Tenen, D. G. PU.1 (Spi-1) and C/EBP alpha regulate the granulocyte colony-stimulating factor receptor promoter in myeloid cells. *Blood* **88**, 1234–1247 (1996).
- Bereshchenko, O. et al. Hematopoietic stem cell expansion precedes the generation of committed myeloid leukemia-initiating cells in C/EBP $\alpha$  mutant AML. *Cancer Cell* **16**, 390–400 (2009).
- Pundhir, S. et al. Enhancer and transcription factor dynamics during myeloid differentiation reveal an early differentiation block in Cebpa null progenitors. *Cell Rep.* **23**, 2744–2757 (2018).
- Wang, G. G. et al. Quantitative production of macrophages or neutrophils ex vivo using conditional Hoxb8. *Nat. Methods* **3**, 287–293 (2006).
- Yi, G. et al. Chromatin-based classification of genetically heterogeneous AMLs into two distinct subtypes with diverse stemness phenotypes. *Cell Rep.* **26**, 1059–1069.e6 (2019).
- Conaway, E. A., de Oliveira, D. C., McInnis, C. M., Snapper, S. B. & Horwitz, B. H. Inhibition of inflammatory gene transcription by IL-10 is associated with rapid suppression of lipopolysaccharide-induced enhancer activation. *J. Immunol.* **198**, 2906–2915 (2017).
- Kim, Y. W., Lee, S., Yun, J. & Kim, A. Chromatin looping and eRNA transcription precede the transcriptional activation of gene in the  $\beta$ -globin locus. *Biosci. Rep.* **35**, e00179 (2015).
- Assi, S. A. et al. Subtype-specific regulatory network rewiring in acute myeloid leukemia. *Nat. Genet.* **51**, 151 (2019).
- Kirstetter, P. et al. Modeling of C/EBP $\alpha$  mutant acute myeloid leukemia reveals a common expression signature of committed myeloid leukemia-initiating cells. *Cancer Cell* **13**, 299–310 (2008).
- Zhang, P. et al. Enhancement of hematopoietic stem cell repopulating capacity and self-renewal in the absence of the transcription factor C/EBP alpha. *Immunity* **21**, 853–863 (2004).
- Ohlsson, E. et al. Initiation of MLL-rearranged AML is dependent on C/EBP $\alpha$ . *J. Exp. Med.* **211**, 5–13 (2014).
- Ye, M. et al. Hematopoietic differentiation is required for initiation of acute myeloid leukemia. *Cell Stem Cell* **17**, 611–623 (2015).
- Gaidzik, V. I. et al. RUNX1 mutations in acute myeloid leukemia are associated with distinct clinico-pathologic and genetic features. *Leukemia* **30**, 2160 (2016).
- Chin, D. W. L. et al. RUNX1 haploinsufficiency results in granulocyte colony-stimulating factor hypersensitivity. *Blood Cancer J.* **6**, e379 (2016).
- Robinson, C. M. et al. Consequences of VHL Loss on Global DNA Methylation. *Sci. Rep.* **8**, 3313 (2018).
- Yates, L. R. et al. Subclonal diversification of primary breast cancer revealed by multiregion sequencing. *Nat. Med.* **21**, 751–759 (2015).
- Hammoud, S. S., Cairns, B. R. & Jones, D. A. Epigenetic regulation of colon cancer and intestinal stem cells. *Curr. Opin. Cell Biol.* **25**, 177–183 (2013).
- Ortmann, C. A. et al. Effect of mutation order on myeloproliferative neoplasms. *N. Engl. J. Med.* **372**, 601–612 (2015).
- Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
- Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
- Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
- Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **34**, 525–527 (2016).
- Soneson, C., Love, M. I. & Robinson, M. D. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Res.* **4**, 1521 (2015).
- Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
- Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* **102**, 15545–15550 (2005).
- Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
- Neph, S. et al. BEDOPS: high-performance genomic feature operations. *Bioinformatics* **28**, 1919–1920 (2012).
- Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
- Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
- Ramirez, F., Dündar, F., Diehl, S., Grüning, B. A. & Manke, T. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.* **42**, W187–W191 (2014).
- Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
- Ernst, J. & Kellis, M. Chromatin-state discovery and genome annotation with ChromHMM. *Nat. Protoc.* **12**, 2478–2492 (2017).
- McLean, C. Y. et al. GREAT improves functional interpretation of cis-regulatory regions. *Nat. Biotechnol.* **28**, 495–501 (2010).
- Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B (Methodol.)* **57**, 289–300 (1995).

**Acknowledgements**

We are grateful to the following core facilities: Histopathology Shared Resource, Massively Parallel Sequencing Shared Resource, Gene Profiling Shared Resource, Epigenetics Consortium, Knight Cancer Institute Biostatistics Shared Resource, Flow Cytometry Shared Resource, ExaCloud Cluster Computational Resource and the Advanced Computing Center. Funding provided by an American Society of Hematology Research

Training Award for Fellows and Collins Medical Trust Award to T.P.B., OHSU School of Medicine Faculty Innovation Fund to B.J.D. and J.E.M., Howard Hughes Medical Institute to B.J.D, NCI R00-CA190605, an ASH Scholar Award, and an MRF New Investigator Grant to J.E.M. NCI U10 CA 98543-08S6, a St. Baldrick's Foundation Grant and Help the Hutch Funding to S.M., L.C. and M.O are supported by the National Science Foundation (grant #1613856) and from the National Institute of Health (R01HG010333), L.C. is also supported in part by NIH/OD P51 OD011092 to the Oregon National Primate Research Center.

### Author contributions

Conceptualization: T.P.B., B.J.D., and J.E.M. Methodology: T.P.B., M.O., and J.E.M. Software: T.P.B., M.O., B.D., S.J., and S.M. Formal analysis: T.P.B., M.O., B.D., B.R.W., M.R.G., J.C.E., H.G.M., K.W., J.L.S., A.R.L., R.E.R., and S.J. Investigation: T.P.B., C.C., S.A.C., A.F., Z.S., K.N., B.S., D.L., B.G., C.D., R.D., and J.E.M. Data curation: T.P.B., M.O., B.D., S.J., and S.M. Writing—original draft: T.P.B., L.C., and J.E.M. Writing—review & editing: T.P.B., M.O., L.C., B.J.D., and J.E.M. Supervision: T.P.B., S.M., L.C., B.J.D. and J.E.M. Funding acquisition: T.P.B., C.N., S.M., L.C., B.J.D., and J.E.M.

### Competing interests

B.J.D. potential competing interests—Consultant: Monojul, Patient True Talk; SAB: Aileron Therapeutics, ALLCRON, Cepheid, Gilead Sciences, Vivid Biosciences, Celgene & Baxalta (inactive); SAB & Stock: Aptose Biosciences, Blueprint Medicines, Beta Cat, GRAIL, Third Coast Therapeutics, CTI BioPharma (inactive); Scientific Founder & Stock: MolecularMD; Board of Directors & Stock: Amgen; Board of Directors: Burroughs Wellcome Fund, CureOne; Joint Steering Committee: Beat AML LLS; Clinical Trial Funding: Novartis, Bristol-Myers Squibb, Pfizer; Royalties: OHSU #606-Novartis exclusive license, OHSU #2573; Dana-Farber Cancer Institute #2063-Merck exclusive license. The remaining authors declare no competing interests.

### Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41467-019-13364-2>.

**Correspondence** and requests for materials should be addressed to J.E.M.

**Peer review information** *Nature Communications* thanks Peter Van Galen and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019