

# SCIENTIFIC REPORTS



OPEN

## NF-E2, FLI1 and RUNX1 collaborate at areas of dynamic chromatin to activate transcription in mature mouse megakaryocytes

Received: 12 May 2016

Accepted: 01 July 2016

Published: 26 July 2016

Chongzhi Zang<sup>1,2,\*</sup>, Annouck Luyten<sup>2,3,4,\*</sup>, Justina Chen<sup>2,3</sup>, X. Shirley Liu<sup>1,2</sup> & Ramesh A. Shivdasani<sup>2,3,4,5</sup>

Mutations in mouse and human *Nfe2*, *Fli1* and *Runx1* cause thrombocytopenia. We applied genome-wide chromatin dynamics and ChIP-seq to determine these transcription factors' (TFs) activities in terminal megakaryocyte (MK) maturation. Enhancers with H3K4me2-marked nucleosome pairs were most enriched for NF-E2, FLI and RUNX sequence motifs, suggesting that this TF triad controls much of the late MK program. ChIP-seq revealed NF-E2 occupancy near previously implicated target genes, whose expression is compromised in *Nfe2*-null cells, and many other genes that become active late in MK differentiation. FLI and RUNX were also the motifs most enriched near NF-E2 binding sites and ChIP-seq implicated FLI1 and RUNX1 in activation of late MK, including NF-E2-dependent, genes. Histones showed limited activation in regions of single TF binding, while enhancers that bind NF-E2 and either RUNX1, FLI1 or both TFs gave the highest signals for TF occupancy and H3K4me2; these enhancers associated best with genes activated late in MK maturation. Thus, three essential TFs co-occupy late-acting *cis*-elements and show evidence for additive activity at genes responsible for platelet assembly and release. These findings provide a rich dataset of TF and chromatin dynamics in primary MK and explain why individual TF losses cause thrombocytopenia.

Terminally differentiated megakaryocytes (MKs) assemble thousands of platelets *de novo* and release them into the circulation<sup>1,2</sup>. The high synthetic demand for thrombopoiesis requires coordinated expression of hundreds of MK-specific genes, reflecting cell- and stage-specific chromatin modulation and transcription factor (TF) binding. Defects in this response underlie some clinical states of platelet deficiency. The TFs FLI1 and RUNX1, for example, are causally implicated in disorders of thrombopoiesis: Paris-Trousseau syndrome (OMIM 188025) and familial platelet disorder with predisposition to acute myeloid leukemia (FDP-AML, OMIM 601399), respectively<sup>3-5</sup>.

These and other TFs are expressed in multiple blood lineages<sup>6,7</sup>, and cell-restricted transcriptional activity represents the outcome of their specific interactions with each other and with chromatin. Erythroid cells and MKs derive from a common bipotential progenitor<sup>8</sup> and share expression of several TFs such as GATA1, NF-E2, TAL1, RUNX1, and FLI1. Mice lacking any of these TFs have various defects in erythropoiesis and thrombopoiesis, and may even die in utero<sup>9,10</sup>, with each TF controlling different genes at various stages in sibling cells. *Fli1* and *Nfe2* null mice lack circulating platelets, reflecting a late MK maturation arrest, reduced  $\alpha$ -granules, and disorganized internal membranes in both cases; *Nfe2*<sup>-/-</sup> MKs elaborate no proplatelets<sup>3,11-15</sup>. *Runx1* deficiency causes mild but significant MK deficits, including arrested maturation and thrombocytopenia<sup>16,17</sup>. These defects reflect the TFs' earliest overall requirement and no transcriptional targets, alone or as a group, explain them fully. In the multipotent progenitor cell line HPC-7<sup>18</sup> and in cultured human MKs<sup>19</sup>, some of the same TFs co-occupy regulatory

<sup>1</sup>Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute and Harvard T.H. Chan School of Public Health, Boston, MA 02215, USA. <sup>2</sup>Center for Functional Cancer Epigenetics, Dana-Farber Cancer Institute, Boston, MA 02215, USA. <sup>3</sup>Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA 02215, USA. <sup>4</sup>Department of Medicine, Harvard Medical School, Boston, MA 02115, USA. <sup>5</sup>Department of Pediatric Hematology/Oncology, Boston Children's Hospital and Harvard Medical School, Boston, MA 02115, USA. \*These authors contributed equally to this work. Correspondence and requests for materials should be addressed to R.A.S. (email: ramesh\_shivdasani@dfci.harvard.edu)

regions more commonly than do solitary TFs or pairs. Five TFs in particular - GATA1, GATA2, RUNX1, FLI1 and TAL1 - co-occupy many genomic sites in immature human MKs<sup>19</sup> and analysis of cultured murine MKs reinforces the idea that these TFs prime MK genes in progenitor cells<sup>20</sup>. However, fewer than 1/3 of MK-specific genes showed nearby binding of this TF pentad, and the outcomes and chromatin states associated with TF binding in mature MKs remain unclear.

As the combinatorial basis for MK maturation and platelet release are not understood well, we reasoned that *cis*-regulatory regions activated late in MK maturation might hold useful clues. Dynamic changes in active chromatin can accurately identify enhancers that function at specific stages in cell differentiation<sup>21,22</sup> and dimethylation of Lysine 4 on Histone 3 (H3K4me2), in particular, appears at both active promoters and enhancers<sup>23–26</sup>. Moreover, we previously showed<sup>24</sup> that H3K4me2 profiles parallel those for H3K27ac, the histone modification most commonly considered to mark functional enhancers<sup>21–23</sup>. To investigate transcriptional control of platelet biogenesis, we examined genome-wide dynamics of H3K4me2-marked chromatin in young and terminally differentiated MK. In so doing, we identified significant activity of three TFs known to be necessary for platelet biogenesis *in vivo*: NF-E2, FLI1 and RUNX1. We then examined the separate and joint activities of these TFs in regulating a large proportion of genes expressed toward the end of the MK life cycle, coincident with platelet release. The sum of our original, high-quality data reveals individual TFs' activities in thrombopoiesis.

## Results

**Identification of dynamic enhancers in differentiating MK.** To investigate chromatin and TF dynamics in MK maturation, we cultured mouse fetal liver cells in thrombopoietin (TPO), collecting young committed cells after 2 days of culture (MK<sub>Imm</sub>) and large, terminally mature cells at the peak of proplatelet formation 3 days later (MK<sub>Mat</sub>). We used Ly76 (Ter119), Ly6g (Gr1) and Itgam (CD11/Mac1) antibodies (Ab) to deplete other blood lineages and 3 sedimentations over bovine serum albumin to remove (in the case of MK<sub>Imm</sub>) or enrich (for MK<sub>Mat</sub>) large, mature cells. Histochemistry and flow cytometry verified the expected differences in cell populations (Fig. 1a,b). Moreover, Ly76<sup>+</sup> erythrocytes were efficiently depleted (<2% residual) and small, unavoidable granulocyte/macrophage residuals were equal in both populations. RNA profiles confirmed lineage depletion and showed efficient separation of immature and mature MKs, revealing higher levels of 692 transcripts and lower levels of 408 genes in MK<sub>Mat</sub> (Fig. 1c), on par with results from prior studies<sup>27</sup>. Gene Ontology analysis of the increased transcripts highlighted membrane-related function, guanyl nucleotide-regulated processes and hemostasis (Fig. 1d), processes known to dominate in terminal MK differentiation<sup>27,28</sup>.

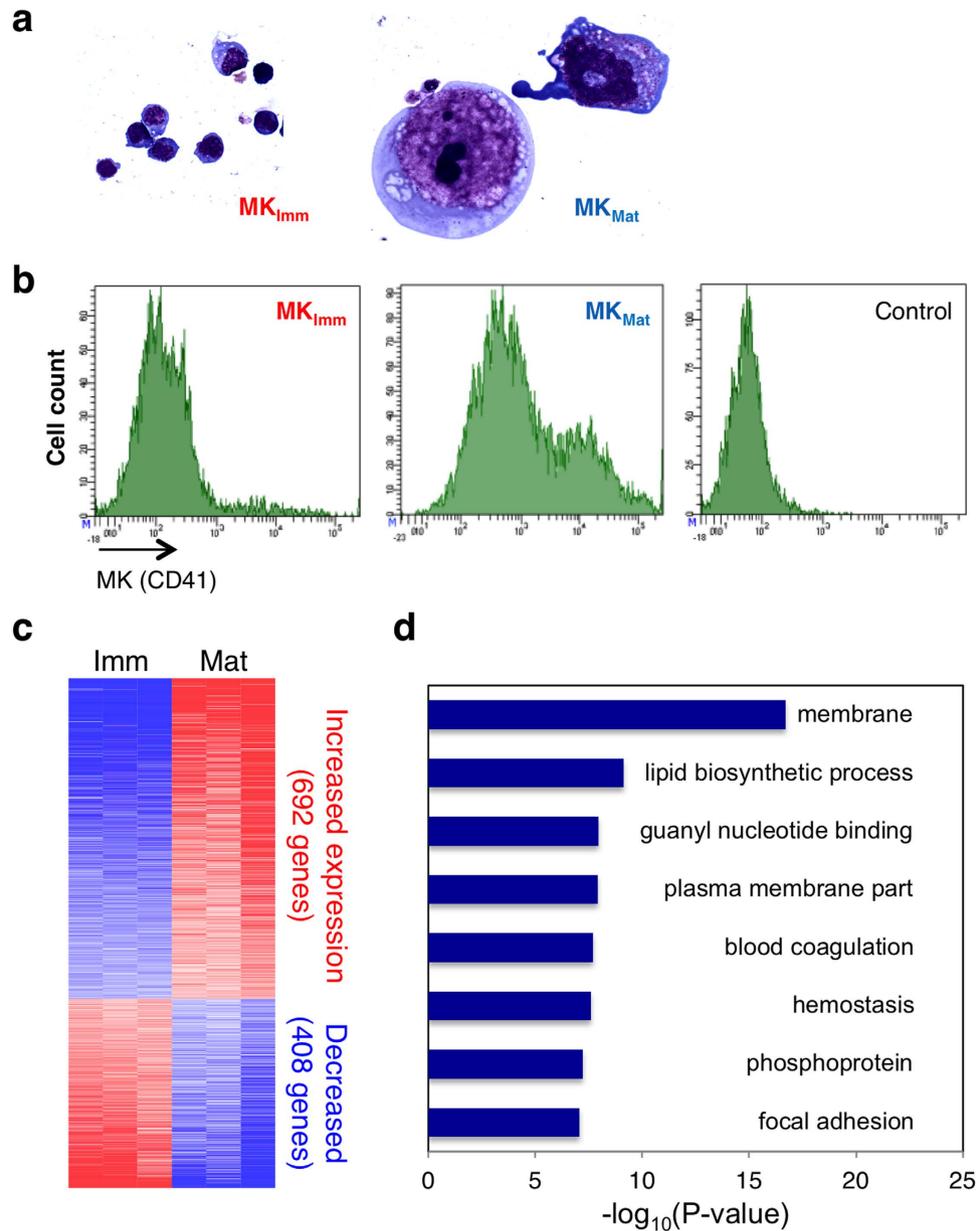
ChIP-seq for H3K4me2 on micrococcal nuclease (MNase)-digested chromatin from the two MK populations identified thousands of H3K4me2-marked nucleosomes (Suppl. Fig. 1a). Many of these were present as nucleosome pairs separated by 250–450 bp of H3K4me2-depleted chromatin (Fig. 2a,b), a configuration that denotes active, TF-occupied *cis*-regulatory elements<sup>21,22</sup>. Nearly half of all such H3K4me2-marked nucleosome pairs localized at gene promoters (<2 kb from a transcription start site, TSS, Suppl. Fig. 1b) and showed little difference between immature and mature MK; this was expected because mammalian genes are controlled largely through distant enhancers<sup>24</sup>, including in MKs<sup>26</sup>. To investigate the relation of these paired nucleosomes with gene expression, we placed all expressed genes in bins of 50 genes each and determined the average number of nucleosome pairs located 2 kb to 20 kb from their TSSs. Differentially expressed genes showed more H3K4me2-marked nucleosome pairs than genes that express at similar levels in immature and mature MK (Fig. 2c). This association was particularly evident for genes activated in MK<sub>Mat</sub> (blue dots), affirming that H3K4me2 ChIP at nucleosome resolution reveals functional enhancers.

To quantify chromatin dynamics at these sites, we assigned Nucleosome Stabilization-Destabilization (NSD) scores<sup>21,22</sup>, based on differences in H3K4me2 ChIP signals in MK<sub>Mat</sub> and MK<sub>Imm</sub>. To identify TFs that may act selectively at enhancers in immature or mature MK, we searched for DNA sequence motifs enriched between all nucleosome pairs unique to each population (Fig. 2d). The GATA motif was enriched within sites in the highest percentile of MK<sub>Imm</sub> selectivity (Fig. 2e), but not at sites enriched in MK<sub>Mat</sub>, consistent with known GATA1 requirements early in MK differentiation<sup>20,29,30</sup>. In contrast, the highest percentile of MK<sub>Mat</sub>-selective sites were enriched for sequences that bind NF-E2, ETS-family proteins, and RUNX1/AML (Fig. 2e), corresponding to three TFs - NF-E2, FLI1, and RUNX1 - whose absence arrests MK maturation, leading to thrombocytopenia *in vivo*<sup>3,11,16</sup>. Thus, a few TFs that bind hundreds of stage-restricted enhancers seem to control a large segment of the transcriptional program in late MK maturation.

### NF-E2 binds DNA at a high fraction of dynamic paired-nucleosome sites to activate MK genes.

Mice lacking NF-E2, the factor with the most enriched motif in areas of dynamic chromatin in MK<sub>Mat</sub>, are profoundly thrombocytopenic<sup>11,13</sup>. To determine NF-E2 interactions with stage-specific chromatin and gene regulation, we used p45 NF-E2 Ab for ChIP-seq (Suppl. Fig. 1a), identifying few binding sites unique to MK<sub>Imm</sub> and thousands of sites exclusive to MK<sub>Mat</sub> (Fig. 3a). The number of confident binding sites was considerably higher than identified in a previous study<sup>31</sup> and the canonical NF-E2 motif<sup>32,33</sup> was the most enriched (Z-score = -135.2) at these sites, which lay mainly in introns or intergenic regions >2 kb from TSSs and showed high phylogenetic conservation (Suppl. Fig. 1c). Individual (e.g., Fig. 3b) and aggregate (Fig. 3c) profiles revealed significant binding in selective regions with H3K4me2-marked nucleosome pairs. To examine this relationship further, we arranged all ~77,000 non-promoter nucleosome pairs into bins of 1,000, ranked by the NSD score, i.e., the magnitude of difference in H3K4me2 signal between MK<sub>Mat</sub> and MK<sub>Imm</sub> (Fig. 3d, x-axis). Most of the top-ranked bins showed frequent NF-E2 binding, which was evident in 5% to 9% of the 5,000 highest-scoring regions (Fig. 3d). Together with the large number of binding sites unique to MK<sub>Mat</sub>, these observations reveal the scope of potential NF-E2 function at enhancers that are active late in MK maturation.

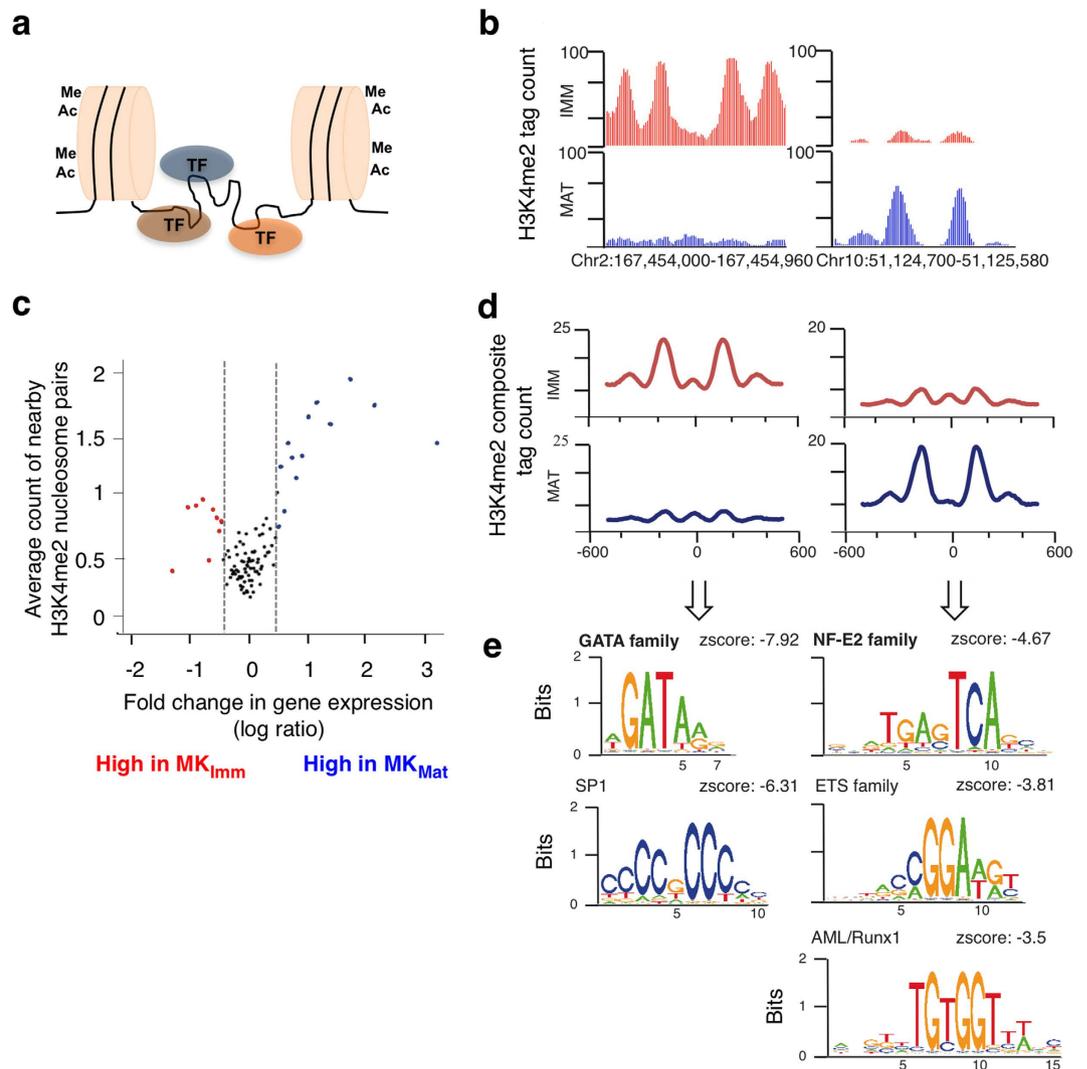
To determine if this binding is functional, we first examined NF-E2 occupancy near genes that are selectively active in MK<sub>Mat</sub>. To this end, we represented differential expression of all genes in MK<sub>Imm</sub> and MK<sub>Mat</sub> on a scatter



**Figure 1. Megakaryocyte (MK) isolation and characteristics.** (a) Representative mouse MK populations stained with May-Grünwald-Giemsa after immunomagnetic and density-gradient isolation 2 days (Imm) or 5 days (Mat) after culture of fetal liver cells in thrombopoietin. (b) Flow cytometry analysis of isolated MK<sub>Imm</sub> and MK<sub>Mat</sub>. Morphology (A) and immunophenotype (B) together verify MK maturation and effective separation. (c) Heat map of genes differentially expressed in MK<sub>Imm</sub> and MK<sub>Mat</sub>, as determined in triplicate Affymetrix microarrays (blue = low, red = high expression). This report centers on regulation of the 692 genes increased  $\geq 1.5$ -fold (FDR  $< 0.05$ ) in MK<sub>Mat</sub>. (d) Gene Ontology analysis of genes expressed selectively in MK<sub>Mat</sub>, showing high enrichment of functions classically attributed to mature MK and blood platelets.

plot in relation to the distance from the TSS to the nearest NF-E2 binding site (Fig. 3e top). This illustrated that NF-E2 binds significantly closer to genes with high expression in MK<sub>Mat</sub> than to invariant genes or those with high expression in MK<sub>Imm</sub> (also shown in Fig. 3f,  $P < 2.2 \times 10^{-16}$  by the KS test). Beyond a distance of 20 kb, similar numbers of genes fall above and below the horizontal lines in Fig. 3e, which indicates that about as many genes with distant NF-E2 binding increase as decrease expression in MK<sub>Mat</sub>. Accordingly, for the subsequent purpose of defining enhancers, we consider 20 kb an empiric boundary for a large fraction of functional binding sites.

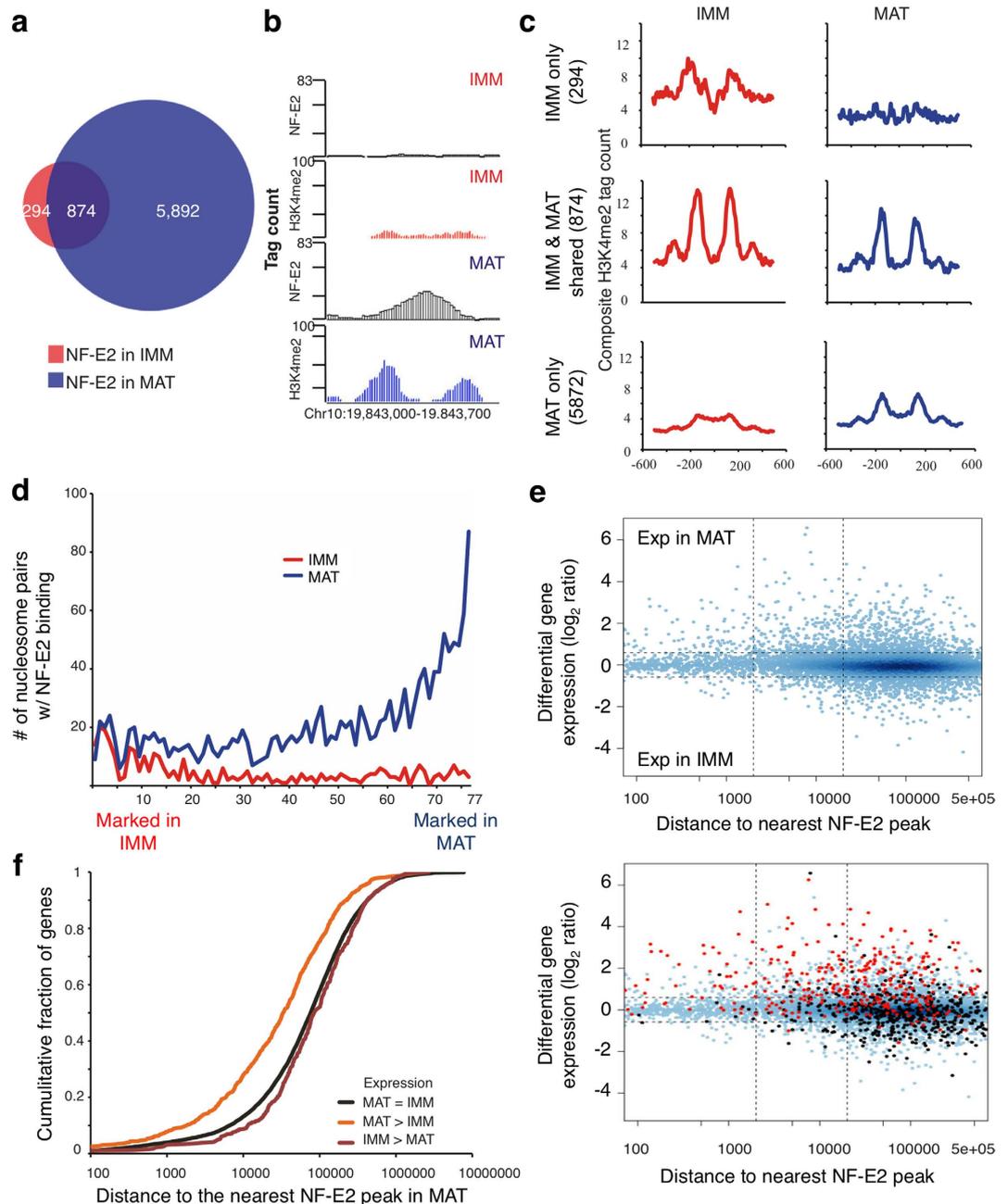
To assess gene dependence on NF-E2, we projected gene expression changes identified in *Nfe2l3*<sup>-/-</sup> MKs<sup>34,35</sup>, which were cultured and harvested similarly to our procedures<sup>34</sup>, on the scatter plot, marking genes that decrease in *Nfe2l3*<sup>-/-</sup> MKs, compared to wild-type cells, in red and those that increase in black (Fig. 3e bottom). We observed higher NF-E2 occupancy near genes with reduced transcript levels in the mutant cells, compared to those that increase or stay fixed ( $P < 2.2 \times 10^{-16}$ ). Although transcript levels in *Nfe2l3*<sup>-/-</sup> MKs again revealed 20 kb as the distance to impute enhancers with confidence, many of these NF-E2-dependent genes showed binding



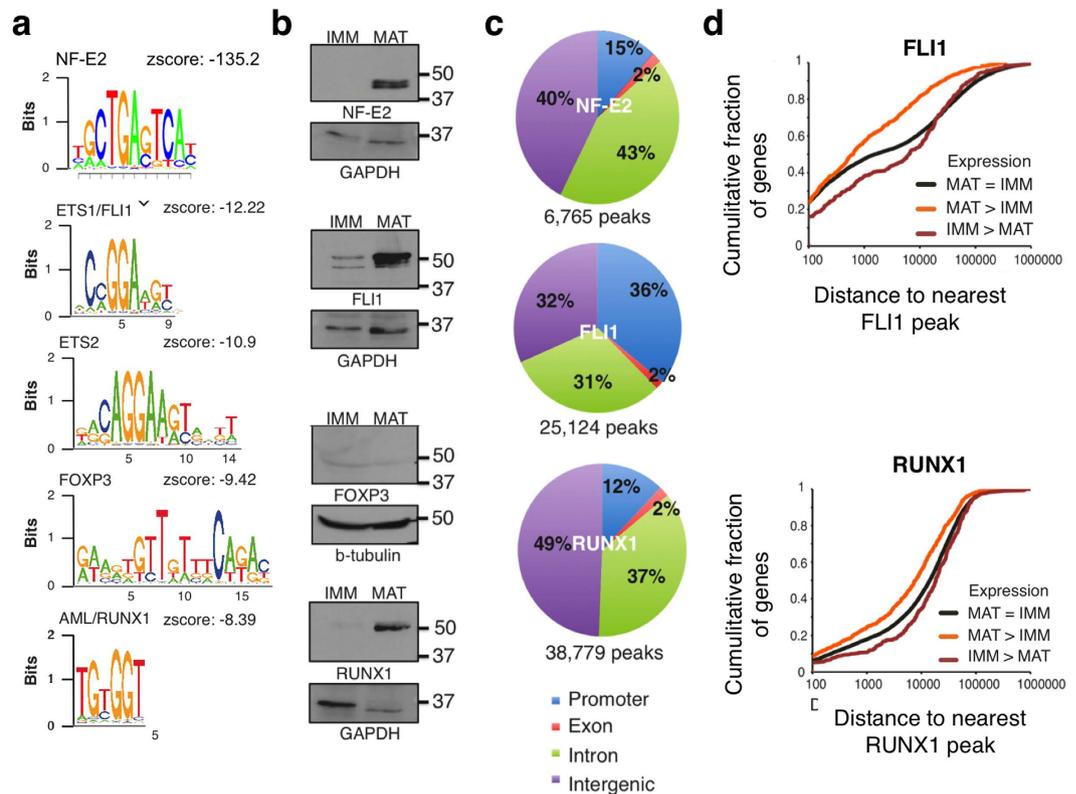
**Figure 2. Analysis of H3K4me2-marked enhancers in immature and mature MK.** (a) Diagram representing active enhancers, showing TF binding in nucleosome-depleted regions flanked by nucleosomes that carry active histone marks such as H3K4me2. (b) H3K4me2 ChIP-Seq data at two representative regions in MK<sub>Imm</sub> (red) and MK<sub>Mat</sub> (blue), illustrating differential enhancer activity in the two cell populations. (c) Differential gene expression (x-axis, bins of 50 genes each) plotted against average counts of H3K4me2-marked nucleosome pairs per bin. Dashed lines demarcate 1.5-fold differential mRNA expression. (d) Composite H3K4me2 tag counts from over 10,000 regions with differential chromatin structure in MK<sub>Imm</sub> (red) and MK<sub>Mat</sub> (blue), aligned at the center of nucleosome pairs. These plots represent the aggregate of signals such as those shown in B. (e) Transcription factor binding motifs significantly enriched (Z-score >3) near the troughs of paired nucleosomes represented in E. Similar motifs were merged and the one with highest Z-scores were selected as representative.

at even larger distances (Fig. 3e bottom). Thus, NF-E2 activates genes, with little direct effect on gene silencing, mainly through enhancers. Transcripts reduced in *Nfe2<sup>-/-</sup>* MKs could be direct transcriptional targets or merely reflect arrested cell maturation. Among reported candidate target genes<sup>27,36–39</sup>, *Tubb1*, which is highly expressed in wild-type MK<sub>Mat</sub> and absent in *Nfe2<sup>-/-</sup>* MKs<sup>40</sup>, showed no NF-E2 occupancy within 250 kb (Suppl. Fig. 1d). In contrast, *Tbxas1*, *Casp12*, *Lims12* and *Rab27b* showed nearby NF-E2 binding (Suppl. Fig. 1e and data not shown). Moreover, NF-E2 bound DNA within 20 kb of 270 out of 692 highly MK<sub>Mat</sub>-selective genes (39%, Suppl. Table 1), compared to 60 of 408 MK<sub>Imm</sub>-specific genes (14.7%,  $P = 1.3 \times 10^{-18}$  by Fisher's exact test). Together, these data confirm the scope of NF-E2 activity and can explain the profound maturation arrest in *Nfe2<sup>-/-</sup>* MKs.

**Roles for RUNX1 and FLI1 in MK maturation.** To identify TFs that may collaborate in this activity, we searched for sequence motifs enriched near NF-E2 binding sites in mature MK. Motifs corresponding to NF-E2, ETS proteins (such as ETS1 and FLI1), FOXP3, and AML/RUNX1 were the most highly represented (Fig. 4a); notably, three of these motifs were also enriched at sites of dynamic chromatin in MK<sub>Mat</sub> (Fig. 2e). Neither immunoblotting (Fig. 4c) nor qRT-PCR (data not shown) detected FOXP3 in mouse MKs, suggesting that some other TF may bind that motif. Moreover, among ETS factors, *Ets1* and *Ets2* mRNA levels fall during MK maturation



**Figure 3. NF-E2 activates genes expressed late in MK maturation.** (a) Overlap of NF-E2 binding sites in  $MK_{Imm}$  (red) and  $MK_{Mat}$  MAT (blue). (b) ChIP-seq signals for NF-E2 (black) and H3K4me2 (red and blue) at representative 1-kb regions in  $MK_{Imm}$  (red) and  $MK_{Mat}$  MAT (blue). (c) Composite H3K4me2 signals in  $MK_{Imm}$  (red) and  $MK_{Mat}$  MAT (blue) at NF-E2 binding sites present only in  $MK_{Imm}$  (top), sites common to both populations (middle row), and sites occupied only in  $MK_{Mat}$  (bottom). (d) NF-E2 shows a strong tendency to bind chromatin that is marked selectively in  $MK_{Mat}$ . H3K4me2-marked nucleosome pairs (putative enhancers, x-axis) were ranked according to the degree of differential marking in  $MK_{Imm}$  (left) or  $MK_{Mat}$  (right) and grouped in bins of 1000 pairs. Y-axis represents the number of nucleosome pairs in each bin that show NF-E2 binding. (e) NF-E2 regulation map. Each dot represents a gene, with the x-axis marking the distance from its TSS to the nearest NF-E2 binding site in  $MK_{Mat}$  and the y-axis marking the log-scaled fold change in mRNA level during MK maturation. The horizontal lines demarcate 1.5-fold cutoffs for transcript levels expressed higher in  $MK_{Imm}$  (negative numbers) or higher in  $MK_{Mat}$  (positive numbers). Dots to the left of the first vertical line represent promoter binding (<2 kb from the TSS) and dots to the right represent binding >2 kb away. In the bottom panel, genes decreased (red) or increased (black) in  $Nfe2^{-/-}$ , relative to wild-type, MK are projected on the above regulation map. Most genes diminished in the absence of NF-E2 (red dots) are expressed selectively in wild-type  $MK_{Mat}$ . (f) Cumulative distribution of the distance from each gene's TSS to its nearest NF-E2 binding site for genes expressed selectively in  $MK_{Mat}$  (orange), in  $MK_{Imm}$  (brown) or at similar levels in both stages (black). NF-E2 binds significantly closer to  $MK_{Mat}$ -specific genes ( $P < 2.2 \times 10^{-16}$  by the K-S test) and further from  $MK_{Imm}$  ( $P = 0.03$ ) genes, compared to the background for invariant genes.

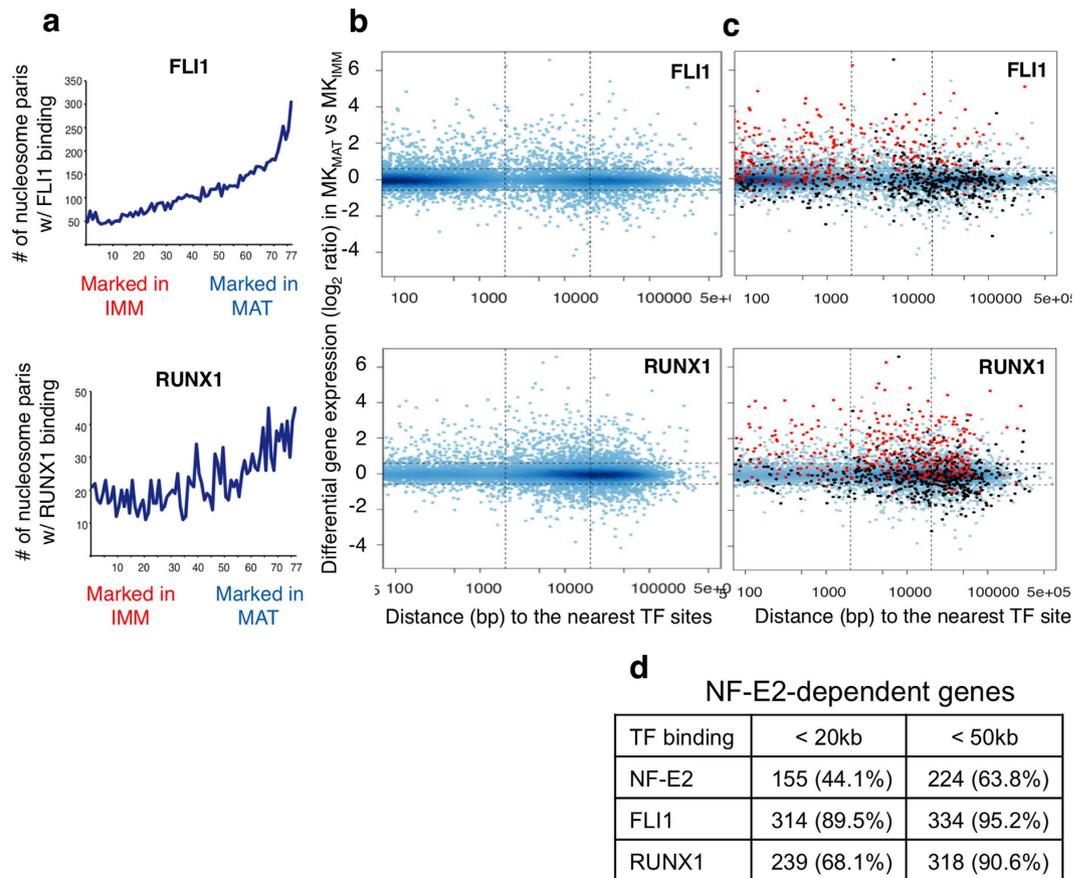


**Figure 4. Expression and binding of TFs with highly enriched recognition motifs detected near NF-E2 binding sites.** (a) Sequence motifs most enriched near sites of NF-E2 occupancy in  $MK_{Mat}$ . (b) Immunoblots showing high levels of NF-E2, FLI1 and RUNX1 in  $MK_{Mat}$  compared to  $MK_{Imm}$ . FoxP3 was not detected and GAPDH or TUBB served as loading controls. (c) Genomic distributions of NF-E2, FLI1 and RUNX1 binding sites in  $MK_{Mat}$ . (d,e) Cumulative distribution of the distance from each gene's TSS to its nearest FLI1 (d) or RUNX1 (e) binding sites for different gene sets. Both TFs bind significantly closer to  $MK_{Mat}$  genes (orange;  $P < 2.2 \times 10^{-16}$  for FLI1,  $P = 2.28 \times 10^{-13}$  for RUNX1 by the K-S test) compared to invariant genes and further from  $MK_{Imm}$  genes (brown;  $P = 1.08 \times 10^{-5}$  for FLI1,  $P = 1.67 \times 10^{-3}$  for RUNX1).

(Suppl. Fig. 2a) and *Ets1*<sup>-/-</sup> or *Ets2*<sup>-/-</sup> mice lack MK or platelet defects<sup>41</sup>. In contrast, levels of another ETS protein, FLI1, rise significantly in  $MK_{Mat}$ , similar to p45 NF-E2 (Fig. 4b), and *Fli1*<sup>-/-</sup> mice show profound dysmegakaryopoiesis<sup>14,15</sup>. Moreover, RUNX1 and FLI1 commonly co-occupy DNA in human and mouse MKs<sup>19,20</sup> and in HPC-7 cells<sup>18</sup>. NF-E2 binding in  $MK_{Mat}$  also occurred at many sites that bind FLI1 and RUNX1 in HPC-7 cells (Suppl. Fig. 2b) and NF-E2 binds the *Fli1* and *Runx1* loci in wild-type  $MK_{Mat}$  (Suppl. Fig. 2c), implying a TF network, a common feature of stable differentiated cells<sup>42</sup>. These observations collectively suggest that NF-E2, FLI1 and RUNX1 are key transcriptional regulators of terminal MK maturation.

ChIP-seq for FLI1 and RUNX1 revealed thousands of confident binding sites for the two TFs (Suppl. Fig. 1a) and encompassing nearly all previously mapped RUNX1 binding sites<sup>43</sup>. Motifs corresponding to each TF were the most significantly enriched in the respective ChIP fragments (Z-score -143.9 for FLI1, -11.7 for RUNX1), which implies their direct binding to DNA. Both factors bound mainly in intergenic regions and introns, similar to NF-E2, although about 1/3 of FLI1 binding occurred at promoters (Fig. 4c). Also resembling NF-E2, both FLI1 and RUNX1 tend to bind closer to genes that increase expression in  $MK_{Mat}$  and further from  $MK_{Imm}$  genes (Fig. 4d), suggesting that all three TFs activate genes in maturing MK. Coupled with stage-selective *cis*-regulatory regions, the binding profiles of 3 abundant, essential TFs provided information relevant to the transcriptional basis of MK maturation.

To investigate this basis, next we mapped FLI1 and RUNX1 occupancy at differentially active regulatory regions, i.e., in relation to NSD scores (differences in H3K4me2 marking). RUNX1 binding was modestly enriched among enhancers selectively marked in  $MK_{Mat}$ , whereas FLI1 binding was evident at 15% to 30% of enhancers with the greatest differential marking (Fig. 5a). The extent of FLI1 binding reflects the 4-fold excess of FLI1 over NF-E2 binding sites and suggests that FLI1 may control an especially large fraction of late MK genes. We therefore considered every gene with respect to its differential expression in  $MK_{Mat}$  vs.  $MK_{Imm}$  (y-axis in Fig. 5b) and the nearest binding of each TF (x-axis). This analysis revealed first that each TF binds DNA mainly near genes that increase expression significantly in  $MK_{Mat}$  (Fig. 5b), similar to NF-E2; this is especially the case at distances under 20 kb. Second, RUNX1 and particularly FLI1 bind the promoters of many more  $MK_{Mat}$  genes than does NF-E2, although, both TFs occupy many more distant sites than promoters, consistent with the bulk of gene regulation occurring at enhancers.

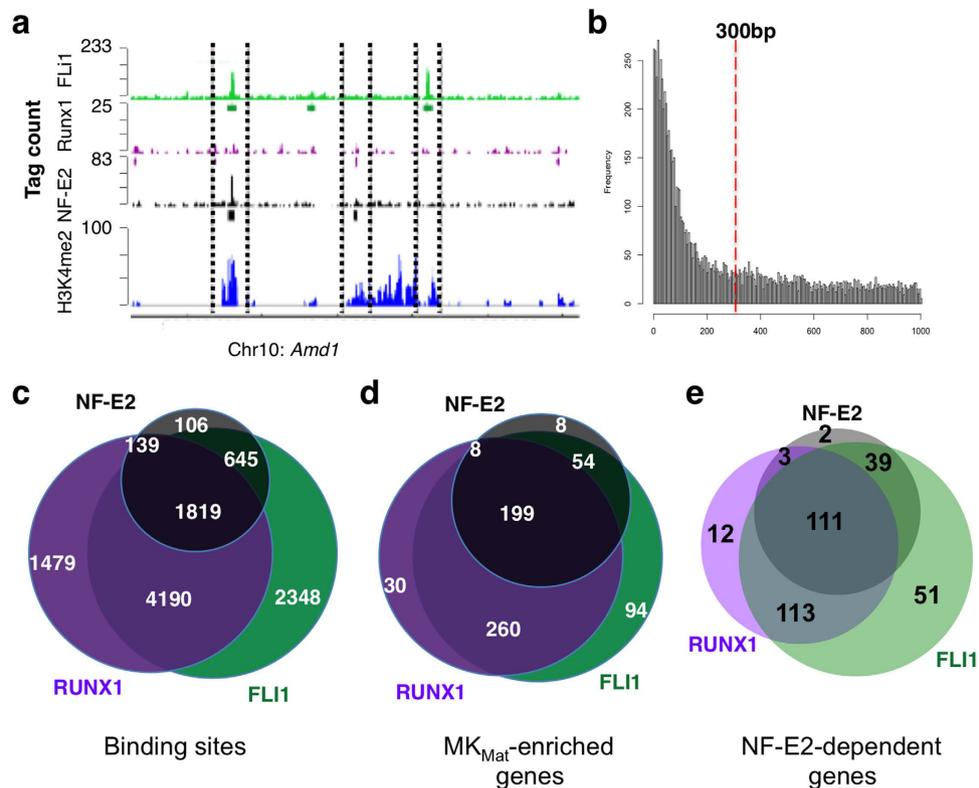


**Figure 5. Features of FLI1 and RUNX1 binding in terminally mature MK.** (a) The tendency of FLI1 (top) and RUNX1 (bottom) to bind at H3K4me2-marked, open chromatin in MK<sub>Mat</sub>. The 77,000 nucleosome pairs identified in MK (Fig. 2) are binned in groups of 1,000 (x-axis, as described for Fig. 3d) and the number of nucleosome pairs having FLI1 or RUNX1 binding sites in each bin is plotted on the y-axis. (b) Regulation maps of FLI1 (top) and RUNX1 (bottom), prepared as described for NF-E2 in Fig. 3f. Each dot represents a gene, with the x-axis marking the distance from its TSS to the nearest TF binding site and the y-axis marking the log-scaled fold-change in transcript level between MK<sub>Mat</sub> and MK<sub>Imm</sub>. (c) Genes with reduced (red dots) or elevated (black dots) expression in *Nfe2<sup>-/-</sup>*, compared to wild-type, MK<sup>35</sup> are projected onto these regulation maps to survey potential roles for FLI1 and RUNX1 in regulating NF-E2-dependent genes. (d) Frequency of TF binding sites within 20 and 50 kb of the TSSs of NF-E2-dependent genes (reduced expression in *Nfe2<sup>-/-</sup>* MK).

To impute the functions of FLI1 and RUNX1 binding with respect to NF-E2, we examined each TF's occupancy near genes that are enriched in MK<sub>Mat</sub> and depend on NF-E2 *in vivo*<sup>34</sup>. Figure 5c projects these genes in red (reduced expression in *Nfe2<sup>-/-</sup>* MK) or black (levels rise in *Nfe2<sup>-/-</sup>* MK) onto the adjoining (Fig. 5b) regulation maps. The densities of red dots in each frame reveal extensive FLI1 and RUNX1 binding near NF-E2-dependent genes, and in contrast to the few such genes with promoter (<2 kb) binding of NF-E2 or RUNX1, nearly 2/3 of these promoters showed FLI1 occupancy (Fig. 5c). FLI1 also bound about half as many enhancers as RUNX1 or NF-E2. Notably, whereas 44.1% of NF-E2-dependent genes bind NF-E2 within 20 kb and 63.8% of these genes bind NF-E2 within 50 kb, the fraction of genes that bind within 20 kb is higher for both RUNX1 (68.1%) and FLI1 (89.5%) binding (Fig. 5d). These associations indicate extensive co-regulation of MK<sub>Mat</sub> genes by these three TFs.

**Singular and combinatorial TF activity in MK maturation.** As it is increasingly clear that multiple *cis*-elements control individual genes, NF-E2, FLI1 and RUNX1 could co-regulate MK<sub>Mat</sub> genes through binding at the same or different enhancers. This is an important distinction because 2 of these 3 TFs co-occupy DNA in other contexts<sup>18,19</sup>, albeit without clear functional consequence. Moreover, the individual profiles for FLI1, RUNX1 and NF-E2 revealed all possible binding combinations: alone, in pairs, or as a trio (Fig. 6a). We therefore sought to determine the prevalence and roles of solitary and combined TF binding.

By surveying the interval distances between summits of all ChIP-seq peaks for the 3 TFs, we identified 300 bp as the upper limit that likely encompasses discrete regulatory elements (Fig. 6b). With this parameter, although each TF showed some binding without concomitant binding of another, we detected substantial co-occupancy, ranging from 74% of all FLI1 binding sites to 96% of all NF-E2 binding sites; more than 1,800 regions bound all three TFs and few regions bound only NF-E2 and RUNX1 (Fig. 6c). Considering the 692 genes expressed selectively in MK<sub>Mat</sub>, we observed TF co-occupancy in the enhancers of 84.5% of all FLI1-bound genes and 94% of all RUNX1-bound genes; indeed, only 8 of 269 NF-E2-bound regions lacked another TF (Fig. 6d). Finally,

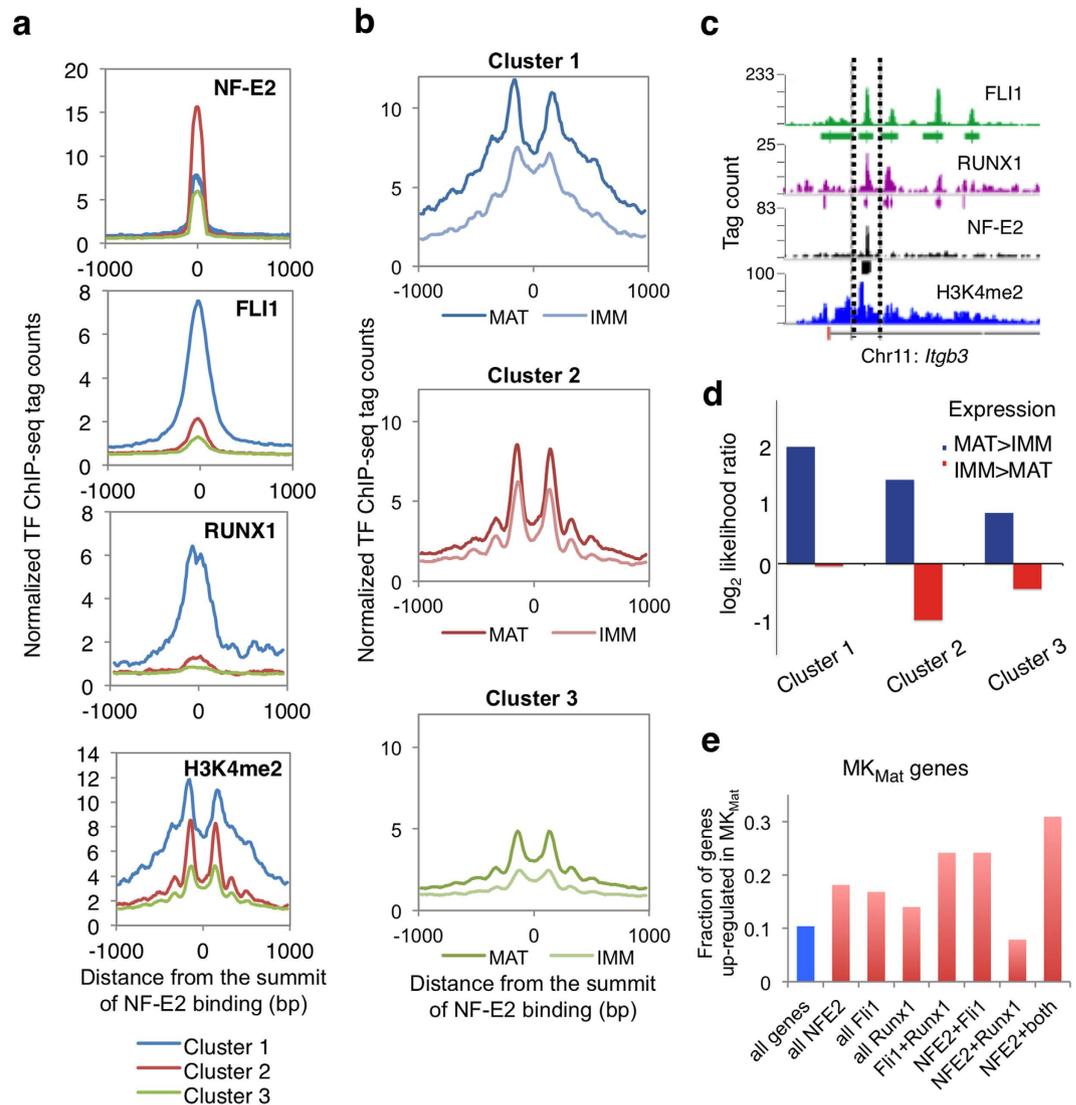


**Figure 6. Combinatorial TF activity in mature MK.** (a) ChIP-seq data traces at a representative MK<sub>Mat</sub>-specific locus, *Amd1*, showing permutations of solitary and combined TF occupancy and H3K4me2 marks within pairs of dotted lines. (b) Histogram of the distances between binding summits for NF-E2, FLI1 and RUNX1, showing the empiric basis for our choice of 300 bp as the criterion for TF co-occupancy. (c–e) Venn diagrams showing the overlap of *all* FLI1, RUNX1 and/or NF-E2 binding sites (c); the TF binding at the enhancer regions of the MK<sub>Mat</sub>-selective genes that bind any of the three TFs (d); and the co-binding of the 3 TFs at genes that *depend on NF-E2 in vivo* (e).

near genes that depend on NF-E2 *in vivo*<sup>34</sup>, we detected substantial co-occupancy with at least FLI1 or RUNX1 and commonly both (Fig. 6e). Thus, co-occupancy of TFs essential to MK function is pervasive and particularly evident in the vicinity of MK<sub>Mat</sub>-selective genes. In particular, genes frequently bind FLI1 at their promoters and show significant binding of NF-E2 and FLI1, with or without RUNX1, at distant enhancers.

To ask how combinations of NF-E2 with other TFs might affect MK gene expression, we used *k*-means clustering, an unbiased and unsupervised method, to sub-divide NF-E2 binding sites into 3 clusters based on the ChIP-seq signal and FLI1/RUNX1 co-occupancy (Fig. 7a). H3K4me2 levels in all 3 clusters of NF-E2 binding sites were increased from MK<sub>Imm</sub> to MK<sub>Mat</sub> and enhancers with high FLI1 and RUNX1 occupancy (cluster 1) showed the largest average gain in H3K4me2 during MK maturation (Fig. 7b). Among all NF-E2-bound enhancers, cluster 1 also showed the highest H3K4me2 signals in flanking nucleosomes (Fig. 7a, blue curves). These features are nicely illustrated at the *Itgb3* locus, which is strongly activated during MK maturation and binds each TF in close proximity within H3K4me2-marked intronic sites (Fig. 7c). NF-E2 binding signals were highest, and H3K4me2 signals remained robust, when only one of the other TFs was also present (cluster 2, where average signals from RUNX1 and FLI1 binding were also considerably lower); H3K4me2 signals were the lowest when NF-E2 was found alone (cluster 3) (Fig. 7a). Notably, enhancers that bound 2 or 3 TFs were associated with a significantly higher probability of selective expression of nearby genes in MK<sub>Mat</sub>. This relationship was most clear and extreme for sites that bind all three TFs (cluster 1) but was also evident for sites that bind just one TF in addition to NF-E2 (Fig. 7d). Thus, enhancers that bind 2 or all 3 TFs mark MK<sub>Mat</sub>-selective genes best.

Because the foregoing analysis centered on *binding sites*, next we considered TF occupancy with respect to the 692 genes expressed selectively in MK<sub>Mat</sub> (Suppl. Table 1). Relative to the background for all genes, those with nearby binding of NF-E2, FLI1 or RUNX1 were enriched for MK<sub>Mat</sub>-specific genes (Fig. 7e). Co-occupancy of FLI1 with NF-E2 or RUNX1 showed even better association with MK<sub>Mat</sub>-specific genes, implying that FLI1 collaborates productively with both TFs. Importantly, nearly 1/3 of genes that bind all 3 TFs in close vicinity are MK<sub>Mat</sub>-selective genes (right-most bar in Fig. 7e), indicating potent additive effects. This association is unlikely to be spurious because pairing of NF-E2 and RUNX1 is much less frequent, suggesting that the latter TF pair may have a limited regulatory role in the absence of FLI1. Furthermore, near genes that bind any of the three TFs alone or in combination, we observed no enrichment - and even some depletion - of genes expressed selectively in MK<sub>Imm</sub> (Suppl. Fig. 2d).



**Figure 7. Functional analysis of TF co-occupancy at MK enhancers.** (a) Composite ChIP-seq signals for NF-E2, FLI1, RUNX1, and H3K4me2 at 3 distinct clusters of NF-E2 binding sites, identified by K-means clustering of binding sites for the three TFs and described in the text. (b) Composite ChIP-seq signals for H3K4me2 in MK<sub>Imm</sub> and MK<sub>Mat</sub> at the 3 clusters of NF-E2 binding sites. (c) ChIP-seq data traces at a representative MK<sub>Mat</sub>-specific locus, *Itgb3*, showing all three TFs occupying a putative regulatory region marked with H3K4me2. Other sites in the locus show binding of one or two TFs. (d) Odds ratios of genes near (<20 kb) NF-E2 binding sites in each cluster showing higher expression in MK<sub>Mat</sub> (blue) or in MK<sub>Imm</sub> (red), relative to the genome background. (e) Proportion of genes with nearby binding of the various combinations of NF-E2, FLI1, and RUNX1 among genes showing higher expression in MK<sub>Mat</sub>. The data indicate the functional role of FLI1, in conjunction with NF-E2 or RUNX1 and especially with both TFs, in regulating genes expressed in mature, but not in immature (Suppl. Fig. 2d) MK.

## Discussion

Much of the current appreciation of the transcriptional control of thrombopoiesis rests on findings in knockout mice and in human pedigrees with syndromic thrombocytopenia, which point separately to three necessary TFs – NF-E2, FLI1 and RUNX1. It is, however, unclear if these are the principal transcriptional determinants of platelet assembly, as other TFs may be equally essential. Moreover, although NF-E2 is activated late in MK maturation and defects in *Nfe2* mutant mice are confined to the final steps in platelet assembly, FLI1 and RUNX1 also show activities in young MKs; thus, defective thrombopoiesis in the absence of the latter factors might reflect those early roles or additional functions in terminally mature cells. The basis for the strong functional overlap among TFs with distinct DNA-binding preferences is also unclear, particularly if they act at promoters or enhancers, and collaborate at the same late-active MK genes or regulate different transcriptional targets. We therefore studied chromatin dynamics during MK differentiation and accurately mapped TF occupancy in terminally mature MK. Our application of the genome-wide histone mark H3K4me2 in different cell states to identify TF activity

highlights the rich information contained in dynamic chromatin and the power of this information to reveal gene regulatory mechanisms in primary cells.

In MKs and other cells, histone modifications are much more dynamic at enhancers than at promoters<sup>24,26</sup>. The hundreds of distant *cis*-elements that acquire hallmarks of activation in terminally mature MKs hence represent the principal sites of relevant TF activity, and sequence motifs corresponding to NF-E2, FLI1 and RUNX1 are by far the most enriched within these regions. FLI1 and RUNX1 were also the next most enriched motifs (after NF-E2) at sites of NF-E2 binding and no other DNA sequences were significantly enriched at regions of FLI1 and RUNX1 occupancy. Although various TFs can in principle bind the same motifs, NF-E2, FLI1 and RUNX1 are the dominant family members expressed selectively in MKs and notably increased in terminal cells. When considered in the light of the thrombopoietic defects associated with mutations in each of these genes, our observations imply that NF-E2, FLI1 and RUNX1 together control much of the late MK transcriptional program, though additional TFs – such as one that binds the Fox motif enriched near NF-E2 binding sites – likely have supporting roles. Most of the ~700 genes activated late in MK maturation show nearby binding of at least one – and often 2 or all 3 – of these TFs, mainly in regions that carry the activated histone mark H3K4me2 only in mature cells. Although gains and losses in many transcripts accompany MK differentiation, genes that are selectively active in terminally mature cells show far greater binding than genes that are silenced, which indicates a predominant activating role for NF-E2, FLI1 and RUNX1, with probably little to no activity in transcriptional repression.

These features contrast with those of other TF families – GATA1/2 and TAL1/LYL1 – that are also expressed in erythroid, MK, and other blood cells. Elegant studies have highlighted activities of the latter TFs, which mainly function early in MK differentiation and activate or repress transcription, depending on the cellular context and associated protein complexes<sup>19,20,44,45</sup>. Moreover, in progenitor cells GATA and TAL proteins seem to prime enhancers for subsequent activity in specified MKs, probably in conjunction with RUNX1 and FLI1<sup>19,20</sup>, which are present at low levels in mouse MK<sub>imm</sub> (Fig. 4b). Indeed, WGATAR was the most enriched motif in enhancers that are selectively active in immature MK, consistent with known GATA1 requirements in these cells<sup>29</sup>. NF-E2, FLI1 and RUNX1 levels increase dramatically in terminal MKs and our study highlights their occupancy at newly activated enhancers.

The frequent association of NF-E2, FLI1 and RUNX1 in various combinations at or near late-MK genes is of particular note. Histones showed low H3K4me2 at enhancers occupied by single TFs and the highest levels of this activation mark in areas that bound NF-E2 and either RUNX1, FLI1 or both TFs. Co-occupied enhancers were also the best associated with genes expressed selectively in MK<sub>Mat</sub> and gave the strongest signals for TF occupancy. Taken together, these findings reveal the importance of these TFs at late-MK enhancers and provide a basis to understand why their absence compromises thrombopoiesis. Notably, loss of any member of this triad produces related cellular defects, each associated with thrombocytopenia. In particular, nearby binding of RUNX1 and/or FLI1, which is evident at most genes with reduced expression in *Nfe2*-null mice (Fig. 5c,d), seems insufficient to drive transcription. Thus, although various combinations of these TFs co-regulate many of the same target genes through discrete enhancers, their functions are not overtly redundant but complementary and individually essential.

Finally, our ChIP-seq data from primary cells will serve as a vital community resource to study epigenome regulation of platelet biogenesis. Genomic data revealing the collaborative functions of three TFs – NF-E2, FLI1 and RUNX1 – provide a foundation to uncover detailed mechanism their collaboration at dynamic enhancers to activate MK-selective genes. Those mechanisms will in turn lead to refined insights into how platelet biogenesis may be manipulated to manage disorders of platelet deficiency or excess. Although our conclusions are based on MKs cultured from mouse fetal livers, and human or adult mouse MKs may well show some differences, the overall regulatory logic of combinatorial TF activity at MK enhancers is probably conserved.

## Methods

**Cell culture and verification.** Fetal livers from the CD1 strain of mice were collected on embryonic day 14 (E14) and single-cell suspensions were prepared by filtration through a 40- $\mu$ m cell strainer, followed by successive passing through 18- to 23-gauge needles. All methods were carried out in accordance with guidelines established by the Animal Care and Use Committee of the Dana-Farber Cancer Institute and all experimental protocols were approved by this committee. After removal of erythrocytes in ammonium chloride-potassium (ACK) lysis buffer, cells were cultured in Dulbecco's Modified Eagle Medium (Invitrogen) supplemented with 10% fetal bovine serum (FBS) and thrombopoietin (TPO, 1% culture supernatant from a producer cell line<sup>46</sup>). After 2 or 5 days of culture, cells were subjected to negative selection with TER119, GR1 and CD11b antibodies (Ab, BD Pharmingen; catalog #553671, 553123 and 553308, respectively) and magnetic Dynabeads (Invitrogen, catalog #110.35), followed by positive (MK<sub>Mat</sub>) or negative (MK<sub>Imm</sub>) selection over bovine serum albumin (BSA: 4%, 3%, 1.5%) gradients. May-Grunwald Giemsa staining of cytocentrifuged cells from each culture verified proper isolation. MK cultured for 2 or 5 days were also stained with FITC-labeled CD41, APC-labeled TER119, APC-labeled GR1 or PE-labeled CD11b Ab (1:200, BD Pharmingen; catalog #553848, 557909, 553129 and 557397, respectively) for 20 min at 4°C, washed in cold phosphate-buffered saline (PBS) containing 2% FBS, incubated in Hoechst dye (1:10,000), and analyzed on a FACSCanto II flow cytometer (BD Biosciences).

**Expression, gene association, and Gene Ontology (GO) analyses.** RNA isolated from purified MK using RNeasy Mini kits (Qiagen, catalog# 74104) was processed and hybridized to Mouse Genome 430A 2.0 microarrays (Affymetrix) according to the manufacturer's instructions. Microarray experiments were done in triplicate. Data were processed using robust multiarray analysis (RMA) to normalize expression indices<sup>47</sup>. Genes with a unique RefSeq ID assigned to the probe set and called as "present" in at least 1 sample were retained for analyses. Differentially expressed genes between MK<sub>Imm</sub> and MK<sub>Mat</sub> were identified using LIMMA<sup>48</sup>, with false discovery rate (FDR) <0.05 and fold-change  $\geq$ 1.5. GO analysis was performed using DAVID tools<sup>49</sup> with default

parameters, and GO terms with FDR <0.001 were selected. cDNA for RT-PCR analysis was synthesized using the QuantiTect reverse transcription kit (Qiagen, Catalog# 205311). Down- or up-regulated genes in *Nfe2<sup>-/-</sup>* MK<sup>34</sup> were chosen on the basis of  $\geq 2$ -fold change on the microarray expression indices between wild-type and *Nfe2<sup>-/-</sup>* cells.

**Immunoblotting.** Cells were lysed in RIPA buffer and boiled in Laemmli sample buffer for 5 min before fractionation by 10% SDS-PAGE. After transfer to nitrocellulose membranes over 1.5 h at 65 V, blots were blocked with 5% milk in PBS containing 0.1% Tween-20 (pH 7.5) and incubated with p45 NF-E2 (1:1000, ref. 39.), FOXP3 (eBioscience, catalog #14-5773), RUNX1 (Abcam, catalog #23980), or FLI1 (Abcam, catalog #15289) Ab overnight at 4 °C. Blots were washed in PBS-Tween, incubated with horseradish peroxidase-conjugated goat anti-rabbit IgG (Santa Cruz, catalog #sc-2054, 1:2,000), and exposed briefly to chemiluminescence reagents (Santa Cruz, catalog #sc-2048).

**ChIP-seq for H3K4me2 at nucleosome resolution and for TFs.** ChIP on purified MK was performed as described previously<sup>22</sup>, with H3K4me2 Ab (Millipore, catalog #07-030) and input chromatin control following chromatin digestion with micrococcal nuclease, or with Ab against p45 NF-E2<sup>39</sup>, RUNX1 (Abcam, catalog #23980) or FLI1 (Abcam, catalog #15289) after chromatin was sheared by sonication. Libraries, prepared using ThruPlex-FD kits (Rubicon Genomics) and at least 30 ng DNA from 3 or more precipitates was pooled for sequencing on an Illumina Hi-Seq instrument.

**ChIP-seq data analysis.** ChIP-seq reads were mapped to mouse genome build mm9 using Bowtie with default parameters, and uniquely mapped, non-redundant reads were retained<sup>50</sup>. H3K4me2 ChIP-seq reads from MK<sub>Imm</sub> and MK<sub>Mat</sub> were merged, and H3K4me2-marked nucleosomes were identified using NPS<sup>51</sup> with default parameters from the merged ChIP-seq library. H3K4me2-marked nucleosome pairs, NSD scores during MK maturation, and sequence motifs enriched among the most dynamic nucleosome pairs were identified using BINOCh<sup>52</sup> with default parameters. NF-E2, FLI1, and RUNX1 binding sites in the genome were identified using MACS 1.4 with default parameters<sup>53</sup>. Motifs enriched at TF binding sites were identified using SeqPos<sup>54</sup> on the Cistrome analysis platform<sup>55</sup>. Whereas all reported sequence motifs have a *P*-value less than 0.001, SeqPos uses a Z-score to measure the enrichment level of a motif based on both frequency of occurrence and proximity to the peak summits (Figs 2e and 4a). ChIP-seq data were represented in wiggle format generated from NPS for H3K4me2 or from MACS for TFs, respectively, and were visualized using IGV Genome Browser (Figs 2b, 3b, 6a and 7c).

**Composite profile of ChIP-seq signal density on a set of regions.** Anchored regions were aligned by the summit locations for TF binding sites (Figs 3c and 7a,b) or center locations for nucleosome pairs (Fig. 2d), and ChIP-seq reads were tallied in non-overlapping 10-bp windows. The genomic location of a sequence read was shifted in the 3' direction by half the average ChIP DNA fragment size to represent the center of the relevant fragment, estimated by calculating the cross-correlation between all 5' and 3' reads<sup>50</sup> (i.e., 150 bp for most datasets). Total read counts were then normalized to RPKM by the total non-redundant read count for each dataset.

**Integrative bioinformatics analyses.** The summit of the MACS-identified ChIP-seq peak was considered as the location of a TF binding site. In TF regulation maps (Figs 3e and 5b,c), the distance from a gene to its nearest TF binding site is calculated by the distance on chromosomal locations between the TSS of that gene and the summit of the nearest peak, upstream or downstream, regardless of the presence of intervening genes, if any. Co-occupancies of multiple TFs were determined under the criterion that the distance between the summits of any two binding sites was lower than the inflection point in the distribution of such distances, i.e., 300 bp (Fig. 6b). *K*-means clustering of NF-E2 binding sites based on FLI1 and RUNX1 co-occupancy (Fig. 7a,b) was performed on the normalized wiggle profiles of NF-E2, FLI1, and RUNX1 ChIP-seq signal extracted from a 1-kb region centered at each NF-E2 peak summits, using the Heatmap tool on the Cistrome analysis platform. Among the different values we tested, *k* = 3 gave simple and the clearest separation based on ChIP-seq signal patterns. ChIP-seq and RNA expression profiling microarray data are deposited in the Gene Expression Omnibus (accession numbers GSE42108 and GSE42110).

## References

1. Italiano, J. E., Lecine, P., Shivdasani, R. A. & Hartwig, J. H. Blood platelets are assembled principally at the ends of proplatelet processes produced by differentiated megakaryocytes. *J Cell Biol* **147**, 1299–1312 (1999).
2. Junt, T. *et al.* Dynamic visualization of thrombopoiesis within bone marrow. *Science* **317**, 1767–1770 (2007).
3. Hart, A. *et al.* Fli-1 is required for murine vascular and megakaryocytic development and is hemizygotously deleted in patients with thrombocytopenia. *Immunity* **13**, 167–177 (2000).
4. Raslova, H. *et al.* Fli-1 monoallelic expression combined with its hemizygous loss underlies the Paris-Trousseau/Jacobsen thrombopenia. *J Clin Invest* **114**, 77–84 (2004).
5. Song, W. J. *et al.* Haploinsufficiency of CBFA2 causes familial thrombocytopenia with propensity to develop acute myelogenous leukaemia. *Nat Genet* **23**, 166–175 (1999).
6. Orkin, S. H. & Zon, L. I. Hematopoiesis: an evolving paradigm for stem cell biology. *Cell* **132**, 631–644 (2008).
7. Hattangadi, S. M., Wong, P., Zhang, L., Flygare, J. & Lodish, H. F. From stem cell to red cell: regulation of erythropoiesis at multiple levels by multiple proteins, RNAs, and chromatin modifications. *Blood* **118**, 6258–6268, doi: 10.1182/blood-2011-07-356006 (2011).
8. Debili, N. *et al.* Characterization of a bipotent erythro-megakaryocytic progenitor in human bone marrow. *Blood* **88**, 1284–1296 (1996).
9. Okuda, T., van Deursen, J., Hiebert, S. W., Grosfeld, G. & Downing, J. R. AML-1, the target of multiple chromosomal translocations in human leukemia, is essential for normal fetal liver hematopoiesis. *Cell* **84**, 321–330 (1996).
10. Wang, Q. *et al.* Disruption of the CBFA2 gene causes necrosis and hemorrhaging in the central nervous system and blocks definitive hematopoiesis. *Proc Natl Acad Sci USA* **93**, 3444–3449 (1996).

11. Shivdasani, R. A. *et al.* Transcription factor NF-E2 is required for platelet formation independent of the actions of thrombopoietin/ MGDF in megakaryocyte development. *Cell* **81**, 695–704 (1995).
12. Lecine, P. *et al.* Mice lacking transcription factor NF-E2 validate the proplatelet model of thrombocytopoiesis and show a platelet production defect that is intrinsic to megakaryocytes. *Blood* **92**, 1608–1616 (1998).
13. Onodera, K., Shavit, J. A., Motohashi, H., Yamamoto, M. & Engel, J. D. Perinatal synthetic lethality and hematopoietic defects in compound mafG::mafK mutant mice. *The EMBO journal* **19**, 1335–1345 (2000).
14. Kawada, H. *et al.* Defective megakaryopoiesis and abnormal erythroid development in Fli-1 gene-targeted mice. *Int J Hematol* **73**, 463–468 (2001).
15. Moussa, O. *et al.* Thrombocytopenia in mice lacking the carboxy-terminal regulatory domain of the Ets transcription factor Fli1. *Mol Cell Biol* **30**, 5194–5206, doi: 10.1128/MCB.01112-09 (2010).
16. Ichikawa, M. *et al.* AML-1 is required for megakaryocytic maturation and lymphocytic differentiation, but not for maintenance of hematopoietic stem cells in adult hematopoiesis. *Nat Med* **10**, 299–304, doi: 10.1038/nm997 (2004).
17. Growney, J. D. *et al.* Loss of Runx1 perturbs adult hematopoiesis and is associated with a myeloproliferative phenotype. *Blood* **106**, 494–504, doi: 10.1182/blood-2004-08-3280 (2005).
18. Wilson, N. K. *et al.* Combinatorial transcriptional control in blood stem/progenitor cells: genome-wide analysis of ten major transcriptional regulators. *Cell Stem Cell* **7**, 532–544, doi: 10.1016/j.stem.2010.07.016 (2010).
19. Tijssen, M. R. *et al.* Genome-wide analysis of simultaneous GATA1/2, RUNX1, FLI1, and SCL binding in megakaryocytes identifies hematopoietic regulators. *Dev Cell* **20**, 597–609, doi: 10.1016/j.devcel.2011.04.008 (2011).
20. Pimkin, M. *et al.* Divergent functions of hematopoietic transcription factors in lineage priming and differentiation during erythromegakaryopoiesis. *Genome Res* **24**, 1932–1944 (2014).
21. He, H. H. *et al.* Positioned nucleosomes flanking a labile nucleosome characterize transcriptional enhancers. *Nat Genet* **42**, 343–347 (2010).
22. Verzi, M. P. *et al.* Differentiation-specific histone modifications reveal dynamic chromatin interactions and partners for the intestinal transcription factor CDX2. *Dev Cell* **19**, 713–726 (2010).
23. Barski, A. *et al.* High-resolution profiling of histone methylations in the human genome. *Cell* **129**, 823–837 (2007).
24. Heintzman, N. D. *et al.* Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* **459**, 108–112 (2009).
25. Shen, Y. *et al.* A map of the cis-regulatory sequences in the mouse genome. *Nature* **488**, 116–120, doi: 10.1038/nature11243 (2012).
26. Luyten, A., Zang, C., Liu, X. S. & Shivdasani, R. A. Active enhancers are delineated *de novo* during hematopoiesis, with limited lineage fidelity among specified primary blood cells. *Genes Dev* **28**, 1827–1839 (2014).
27. Chen, Z., Hu, M. & Shivdasani, R. A. Expression analysis of primary mouse megakaryocyte differentiation and its application in identifying stage-specific molecular markers and a novel transcriptional target of NF-E2. *Blood* **109**, 1451–1459 (2007).
28. Chang, Y., Bluteau, D., Debili, N. & Vainchenker, W. From hematopoietic stem cells to platelets. *J Thromb Haemost* **5** Suppl 1, 318–327, doi: 10.1111/j.1538-7836.2007.02472.x (2007).
29. Vyas, P., Ault, K., Jackson, C. W., Orkin, S. H. & Shivdasani, R. A. Consequences of GATA-1 deficiency in megakaryocytes and platelets. *Blood* **93**, 2867–2875 (1999).
30. Dore, L. C., Chlon, T. M., Brown, C. D., White, K. P. & Crispino, J. D. Chromatin occupancy analysis reveals genome-wide GATA factor switching during hematopoiesis. *Blood* **119**, 3724–3733, doi: 10.1182/blood-2011-09-380634 (2012).
31. Fujita, R. *et al.* NF-E2 p45 is important for establishing normal function of platelets. *Mol Cell Biol* **33**, 2659–2670 (2013).
32. Ney, P. A., Sorrentino, B. P., McDonagh, K. T. & Nienhuis, A. W. Tandem AP-1-binding sites within the human b-globin dominant control region function as an inducible enhancer in erythroid cells. *Genes & Dev* **4**, 993–1006 (1990).
33. Lecine, P., Blank, V. & Shivdasani, R. Characterization of the hematopoietic transcription factor NF-E2 in primary murine megakaryocytes. *J Biol Chem* **273**, 7572–7578 (1998).
34. Motohashi, H. *et al.* NF-E2 domination over Nrf2 promotes ROS accumulation and megakaryocytic maturation. *Blood* **115**, 677–686, doi: 10.1182/blood-2009-05-223107 (2010).
35. Motohashi, H. *et al.* Molecular determinants for small Maf protein control of platelet production. *Mol Cell Biol* **31**, 151–162, doi: 10.1128/MCB.00798-10 (2011).
36. Deveaux, S. *et al.* p45 NF-E2 regulates the expression of thromboxane synthase in megakaryocytes. *The EMBO journal* **16**, 5654–5661 (1997).
37. Shiraga, M. *et al.* Primary megakaryocytes reveal a role for transcription factor NF-E2 in integrin alpha IIb beta 3 signaling. *J Cell Biol* **147**, 1419–1430 (1999).
38. Eto, K. *et al.* Megakaryocytes derived from embryonic stem cells implicate CalDAG-GEFI in integrin signaling. *Proc Natl Acad Sci USA* **99**, 12819–12824 (2002).
39. Tiwari, S. *et al.* A role for Rab27b in NF-E2-dependent pathways of platelet formation. *Blood* **102**, 3970–3979 (2003).
40. Lecine, P., Italiano, J. E. Jr, Kim, S. W., Villeval, J. L. & Shivdasani, R. A. Hematopoietic-specific beta 1 tubulin participates in a pathway of platelet biogenesis dependent on the transcription factor NF-E2. *Blood* **96**, 1366–1373 (2000).
41. Bartel, F. O., Higuchi, T. & Spyropoulos, D. D. Mouse models in the study of the Ets family of transcription factors. *Oncogene* **19**, 6443–6454, doi: 10.1038/sj.onc.1204038 (2000).
42. Howard, M. L. & Davidson, E. H. cis-Regulatory control circuits in development. *Dev Biol* **271**, 109–118 (2004).
43. Pencovich, N. *et al.* Cell-autonomous function of Runx1 transcriptionally regulates mouse megakaryocytic maturation. *PLoS one* **8**, e64248 (2013).
44. Yu, M. *et al.* Insights into GATA-1-mediated gene activation versus repression via genome-wide chromatin occupancy analysis. *Molecular cell* **36**, 682–695, doi: 10.1016/j.molcel.2009.11.002 (2009).
45. Dore, L. C. & Crispino, J. D. Transcription factor networks in erythroid cell and megakaryocyte development. *Blood* **118**, 231–239, doi: 10.1182/blood-2011-04-285981 (2011).
46. Villeval, J.-L. *et al.* High thrombopoietin production by hemopoietic cells induces a fatal myeloproliferative syndrome in mice. *Blood* **90**, 4369–4383 (1997).
47. Irizarry, R. A. *et al.* Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostat* **4**, 249–264 (2003).
48. Smyth, G. K. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* **3**, Article3, doi: 10.2202/1544-6115.1027 (2004).
49. Huang da, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols* **4**, 44–57 (2009).
50. Zang, C. *et al.* A clustering approach for identification of enriched domains from histone modification ChIP-Seq data. *Bioinformatics* **25**, 1952–1958, doi: 10.1093/bioinformatics/btp340 (2009).
51. Zhang, Y., Shin, H., Song, J. S., Lei, Y. & Liu, X. S. Identifying positioned nucleosomes with epigenetic marks in human from ChIP-Seq. *BMC Genomics* **9**, 537 (2008).
52. Meyer, C. A., He, H. H., Brown, M. & Liu, X. S. BINOCh: binding inference from nucleosome occupancy changes. *Bioinformatics* **27**, 1867–1868 (2011).
53. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137 (2008).

54. Lupien, M. *et al.* FoxA1 translates epigenetic signatures into enhancer-driven lineage-specific transcription. *Cell* **132**, 958–970 (2008).
55. Liu, T. *et al.* Cistrome: an integrative platform for transcriptional regulation studies. *Genome Biol* **12**, R83 (2011).

### Acknowledgements

The authors thank Hui Huang and Alan B. Cantor for mutant mice and valuable discussions. This work was supported by US National Institutes of Health grants R01 HL63143 to R.A.S. and R01 GM099409 to X.S.L., and a fellowship from the Leukemia & Lymphoma Society to C.Z.

### Author Contributions

R.A.S. conceived the experiments. A.L. and J.C. conducted the experiments. C.Z. and X.S.L. analyzed the data. C.Z., A.L. and R.A.S. wrote the manuscript. All authors reviewed and approved the manuscript.

### Additional Information

**Accession codes:** ChIP-seq and expression profiling data are available in the Gene Expression Omnibus (GEO) with accession numbers GSE42108 and GSE42110.

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Zang, C. *et al.* NF-E2, FLI1 and RUNX1 collaborate at areas of dynamic chromatin to activate transcription in mature mouse megakaryocytes. *Sci. Rep.* **6**, 30255; doi: 10.1038/srep30255 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>