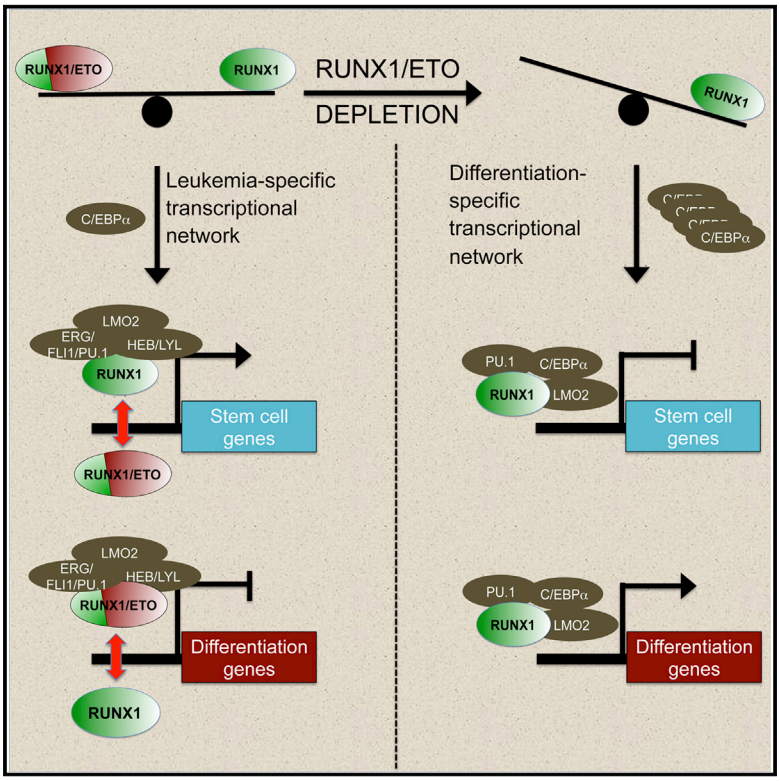
Article

Cell Reports

Identification of a Dynamic Core Transcriptional Network in t(8;21) AML that Regulates Differentiation Block and Self-Renewal

Graphical Abstract



Highlights

RUNX1/ETO drives a t(8;21)-specific transcriptional network

RUNX1/ETO and RUNX1 dynamically compete for the same genomic sites

RUNX1/ETO targets transcription factor complexes that control differentiation

RUNX1/ETO depletion activates a transcriptional network dominated by C/EBPα

Authors

Anetta Ptasinska, Salam A. Assi, ..., Olaf Heidenreich, Constanze Bonifer

Correspondence

olaf.heidenreich@ncl.ac.uk (O.H.), c.bonifer@bham.ac.uk (C.B.)

In Brief

Chromosomal rearrangements generate cancer-specific fusion genes that interfere with cell differentiation. Ptasinska et al. show that the most frequent fusion protein in acute myeloid leukemia (RUNX1/ETO) controls a cancer-propagating transcriptional network by binding to genomic sites in a dynamic equilibrium with wild-type RUNX1. Depletion of RUNX1/ETO installs a differentiation-promoting transcriptional network. Our findings demonstrate that the differentiation block in AML has a dynamic component as its core feature, which might provide a target for cancer-specific differentiation therapy.

Accession Numbers

GSE29225
GSE54478

Ptasinska et al., 2014, Cell Reports 8, 1974–1988
September 25, 2014 ©2014 The Authors
http://dx.doi.org/10.1016/j.celrep.2014.08.024

OPEN ACCESS
CellPress

Cell Reports
Article

Identification of a Dynamic Core Transcriptional Network in t(8;21) AML that Regulates Differentiation Block and Self-Renewal

Anetta Ptasinska,[1,7] Salam A. Assi,[1,3,7] Natalia Martinez-Soria,[4] Maria Rosaria Imperato,[1] Jason Piper,[2] Pierre Cauchy,[1] Anna Pickin,[1] Sally R. James,[6] Maarten Hoogenkamp,[1] Dan Williamson,[4] Mengchu Wu,[5] Daniel G. Tenen,[5] Sascha Ott,[2] David R. Westhead,[3] Peter N. Cockerill,[1] Olaf Heidenreich,[4,*] and Constanze Bonifer[1,*]

[1]School of Cancer Sciences, College of Medicine and Dentistry, University of Birmingham, Birmingham B15 2TT, UK
[2]Warwick Systems Biology Centre, University of Warwick, Coventry CV4 7AL, UK
[3]School of Molecular and Cellular Biology, Faculty of Biological Sciences, University of Leeds, Leeds LS2 9JT, UK
[4]Northern Institute for Cancer Research, University of Newcastle, Newcastle upon Tyne NE2 4HH, UK
[5]Cancer Science Institute, National University of Singapore, Republic of Singapore, Singapore 117456, Singapore
[6]Section of Experimental Haematology, Leeds Institute for Molecular Medicine, University of Leeds, Leeds LS2 9JT, UK
[7]Co-first author
*Correspondence: olaf.heidenreich@ncl.ac.uk (O.H.), c.bonifer@bham.ac.uk (C.B.)
http://dx.doi.org/10.1016/j.celrep.2014.08.024

SUMMARY

Oncogenic transcription factors such as RUNX1/ETO, which is generated by the chromosomal translocation t(8;21), subvert normal blood cell development by impairing differentiation and driving malignant self-renewal. Here, we use digital footprinting and chromatin immunoprecipitation sequencing (ChIP-seq) to identify the core RUNX1/ETO-responsive transcriptional network of t(8;21) cells. We show that the transcriptional program underlying leukemic propagation is regulated by a dynamic equilibrium between RUNX1/ETO and RUNX1 complexes, which bind to identical DNA sites in a mutually exclusive fashion. Perturbation of this equilibrium in t(8;21) cells by RUNX1/ETO depletion leads to a global redistribution of transcription factor complexes within preexisting open chromatin, resulting in the formation of a transcriptional network that drives myeloid differentiation. Our work demonstrates on a genome-wide level that the extent of impaired myeloid differentiation in t(8;21) is controlled by the dynamic balance between RUNX1/ETO and RUNX1 activities through the repression of transcription factors that drive differentiation.

INTRODUCTION

Lineage-specific cell differentiation is controlled by the establishment of specific gene-expression patterns in normal cells, and interference with this process underpins oncogenesis. Hematopoiesis is one of the best-understood developmental pathways and involves dynamic alterations in transcriptional programs, which regulate progression along the differentiation hierarchy (Pimanda and Göttgens, 2010). Individual cellular differentiation states are defined by transcriptional networks composed of combinations of transcription factors that bind to specific sets of *cis*-regulatory elements (Davidson, 2010). Therefore, experimental analysis of the binding activities of multiple factors has served as a means of identifying crucial regulators for a specific cell type (DeVilbiss et al., 2014; Tijssen et al., 2011). However, normal differentiation is impaired in cancers, leading cells to adopt a new malignant identity. Unique insights into processes that control development toward both normal and perturbed differentiation states can be gained from a detailed examination of the mechanisms utilized by leukemic transcription factors such as PML/RARA, MLL fusion proteins, and RUNX1/ETO. These factors reprogram the epigenome and thereby block the hierarchical succession of normal transcriptional networks.

Leukemias are characterized by good experimental accessibility and, compared with many carcinomas, relatively high genetic stability, which makes them very amenable to investigations of general as well as specific mechanisms of oncogenesis. Acute myeloid leukemia (AML) is the second most common leukemia and is a heterogeneous disease with impaired myeloid differentiation (Valk et al., 2004). The hallmarks of AML are multiple somatic mutations, including genetic rearrangements that affect signal transduction and gene expression. This includes mutations in genes encoding DNA methylases, chromatin modifiers, and transcription factors. Many such mutations affect transcription factors that are crucial for the development of hematopoietic stem cells or for terminal myeloid differentiation, such as RUNX1 and C/EBPα, respectively (Gaidzik et al., 2011; Michaud et al., 2002; Pabst et al., 2001b; Snaddon et al., 2003). However, the molecular details of how such mutant transcription factors cause alterations of the epigenome are still insufficiently understood. In addition, so far no experiments have defined the core transcriptional network of a specific type of AML and dissected the role of mutated transcription factors within this network.

One of the best-characterized chromosomal rearrangements found in AML is the t(8;21) translocation, which accounts for approximately 10% of all AMLs. This translocation fuses the DNA-binding domain of the hematopoietic master regulator RUNX1 to almost the entire ETO protein, which is an adaptor protein for histone deacetylase (HDAC) complexes (Miyoshi et al., 1993). The resulting RUNX1/ETO fusion protein lacks the transactivation domain of RUNX1, resulting in major differences in the biological activities of the two proteins. RUNX1 normally recruits transcriptional activators and binds to DNA as a heterodimer with core-binding factor β (CBFβ). The RUNX1/ETO fusion protein also interacts with CBFβ but functions as a RUNX1/ETO tetramer (Liu et al., 2006), and like ETO itself, it also interacts with NCOR and SIN3A corepressors (Amann et al., 2001). Consequently, this chromosomal rearrangement converts a transcriptional activator into a repressor. However, there is evidence that RUNX1 also interacts with HDACs via SIN3A and can act as a repressor (Reed-Inderbitzin et al., 2006; Taniuchi et al., 2002). Proteomic and chromatin immunoprecipitation (ChIP) analyses in t(8;21) cell lines have demonstrated the association of RUNX1/ETO with multiple hematopoietic regulators known to be involved in the regulation of hematopoietic stem cell genes (Wilson et al., 2010). The RUNX1/ETO complex consists of the E box binding transcription factors HEB and LYL1 and the bridging factors LMO2 and LDB1. In chromatin, this complex interacts with the ETS family members FLI1 and ERG, and these interactions are required for the stability of the complex and its leukemogenicity (Martens et al., 2012; Sun et al., 2013).

Genome-wide analyses in t(8;21) cell lines and in patients via ChIP sequencing (ChIP-seq) identified thousands of RUNX1/ETO-binding sites (Ben-Ami et al., 2013; Martens et al., 2012; Ptasinska et al., 2012; Saeed et al., 2012), but the role of specific binding sites within the AML-specific transcriptional network is unclear. All t(8;21) AML cells retain an intact copy of RUNX1, which is required for cell survival—a feature that has also been observed in other CBF leukemias (Ben-Ami et al., 2013; Goyama et al., 2013). RUNX1 and RUNX1/ETO each drive the expression of alternate subsets of genes (Ben-Ami et al., 2013). However, 60% of the RUNX1/ETO sites are shared with RUNX1 (Ptasinska et al., 2012), and whether there is a direct dynamic competition between RUNX1/ETO and RUNX1 for the same genomic sites remains to be investigated.

The differentiation of t(8;21) cells is blocked at an early myeloid progenitor stage and so far the core transcriptional program underlying this block has been elusive. Changes in RUNX1/ETO expression in t(8;21) AML cells are associated with both up- and downregulated genes, and individual RUNX1/ETO-bound genomic sites recruit both histone acetyltransferases (HATs) and HDACs (Follows et al., 2003; Ptasinska et al., 2012; Sun et al., 2013; Wang et al., 2011). However, we previously showed that the genome-wide loss of RUNX1/ETO binding correlates with increased histone H3 lysine 9 (H3K9) acetylation (Ptasinska et al., 2012). In addition, RUNX1/ETO depletion is associated with the upregulation of C/EBPα, a driver of myeloid and, in particular, granulocytic differentiation (Zhang et al., 1997). Moreover, RUNX1/ETO has been shown to sequester C/EBPα from its murine promoter, thereby interfering with C/EBPα expression (Pabst et al., 2001a). RUNX1/ETO knockdown causes release of the differentiation block, resulting in a gene-expression pattern that resembles that of granulocytes and monocytes (Ptasinska et al., 2012). Taken together, these results suggest that RUNX1/ETO-mediated reprogramming of the epigenome involves a complex and so far unexplored interplay of different transcription-factor and chromatin-modifying cofactor activities. To date, we have gained little insight into the nature of this reprogrammed network and the sequential order of factors required to restore normal myeloid cell functions.

In this study, we addressed these issues by investigating the dynamic changes in global transcription-factor-binding patterns that occur following depletion of RUNX1/ETO. To that end, we combined ChIP-seq for multiple factors, DNaseI footprinting, and transcriptome analysis to identify the core transcriptional network of t(8;21) AML cells, and then characterized changes in these networks upon RUNX1/ETO knockdown. These analyses revealed a dynamic equilibrium between RUNX1/ETO and RUNX1 complexes competing for identical genomic sites. Results from sequential ChIP (re-ChIP) show that the two complexes have similar accessory-factor compositions but differ in their preference for the recruitment of coactivators and corepressors. Using a digital DNaseI footprinting approach, we found that both t(8;21)-positive cell lines (Kasumi-1 and SKNO-1) and patient-derived primary AML cells with the t(8;21) translocation (patient cells) share the same pattern of binding-site occupancy. Within this core transcriptional network, RUNX1/ETO-bound loci are predominantly associated with transcriptional repression. Furthermore, loss of RUNX1/ETO establishes a differentiation-associated transcriptional network dominated by de novo binding of C/EBPα resulting from the upregulation of *CEBPA* gene expression. Our results demonstrate that the block in myeloid differentiation in t(8;21) AML results from the dynamic interference of RUNX1/ETO with *cis*-regulatory elements that normally are destined to change transcription-factor assemblies during myeloid differentiation, notably those that increase binding of RUNX1 and C/EBPα.

RESULTS

Transcription-Factor Occupancy Patterns Are Highly Comparable between t(8;21) Cell Lines and Patient Cells

To define the RUNX1/ETO-responsive core transcriptional network and monitor dynamic changes associated with alterations in RUNX1/ETO status, we utilized Kasumi-1 cells, which represent a well characterized and widely used model system for t(8;21) AML (Ben-Ami et al., 2013; Martens et al., 2012; Ptasinska et al., 2012; Sun et al., 2013). We measured the binding of multiple transcription factors in these cells using genome-wide ChIP-seq and performed perturbation experiments by transiently knocking down RUNX1/ETO expression. We then monitored the consequences using ChIP-seq and RNA sequencing (RNA-seq) analyses (Heidenreich et al., 2003; Ptasinska et al., 2012; Table S1). We used antibodies against RUNX1, the ETO moiety of RUNX1/ETO, LMO2 as a member of the RUNX1/ETO complex, RNA-Polymerase II, and acetylated histone H3 for ChIP. To obtain a more complete picture of the composition of RUNX1 and RUNX1/ETO-associated transcription-factor complexes

without RUNX1/ETO knockdown, we also analyzed publicly available data for the E box protein HEB (Martens et al., 2012; Ptasinska et al., 2012). In order to follow additional alterations in the epigenome after RUNX1/ETO knockdown, we also measured the binding of PU.1 and C/EBPα, which are both required for myeloid differentiation (Scott et al., 1994; Zhang et al., 1997). We identified high-confidence transcription-factor binding-site peaks by integrating ChIP data with DNaseI-seq data before and after RUNX1/ETO depletion, and considered only those peaks that were located within DNaseI hypersensitive sites (DHSs).

RUNX1/ETO exists as a complex with other transcription factors (Sun et al., 2013). Consistent with these findings, we observed a colocalization of RUNX1/ETO, RUNX1, HEB, LMO2, C/EBPα, and/or PU.1 binding at many DHSs in Kasumi-1 cells, as exemplified by the *LMO2* locus (Figure 1A). Closer examination of the genome-wide occupancy patterns of LMO2 and HEB revealed that a substantial overlap existed among LMO2, HEB, and RUNX1/ETO binding sites (Figure S1A). Although there was some overlap with the other factors, the PU.1 and C/EBPα binding sites did not closely cluster as a group with those for the RUNX1/ETO complexes in Kasumi-1.

We next sought to determine whether the RUNX1/ETO and RUNX1 binding patterns identified in Kasumi-1 cells were shared with patient cells. First, we performed a DHS analysis on patient cells and normal CD34+ hematopoietic stem and precursor cells (CD34+ cells) derived from the peripheral blood of healthy donors. This fraction is enriched for stem and multipotent progenitor cells. DHS mapping was complemented by RUNX1/ETO and RUNX1 ChIP analysis. However, the large quantity of material required for this approach precluded analysis of patient cells. Therefore, to determine which subsets of DHSs from patient cells overlap with sites that recruit RUNX1 and RUNX1/ETO in the cell line and in CD34+ cells, we first generated a scatter diagram of the joint DHS signal of patient cells (Ptasinska et al., 2012) compared with normal CD34+ cells (Figure S1B). We then projected the genomic coordinates from the RUNX1/ETO and RUNX1 ChIP experiments onto these sequences. These diagrams clearly show that the RUNX1- and RUNX1/ETO-bound sequences from Kasumi-1 cells projected onto the DHS peaks from patient cells, whereas RUNX1-bound sequences from CD34+ cells projected onto the DHS peaks from the CD34+ cells.

To further confirm the similarity between t(8;21) cell lines and patient cells, and to test whether we could overcome the need to conduct multiple ChIP-seq experiments, we generated additional higher-read-depth DNaseI data from two t(8;21) patients and developed a digital footprinting algorithm (Wellington). This high-resolution approach takes the chromatin structure surrounding transcription-factor motifs that are protected from DNaseI digestion into account and thus evaluates the genome-wide transcription-factor occupancy with high accuracy (Piper et al., 2013). DNaseI footprinting data obtained from one t(8;21) patient were compared with ChIP data for regions bound by RUNX1/ETO, RUNX1, HEB, and LMO2 in Kasumi-1 cells (13,584 peaks in total). This comparison demonstrated a high concordance between transcription-factor binding in Kasumi-1 cells and motif occupancy in patient cells, as defined by preferential protection against DNaseI digestion (Figure S1C). This is exemplified by the DNaseI footprints found at the *NFE2* locus (Figure 1B, gray areas), which in both patient samples reflect the pattern of binding of RUNX1/ETO, HEB, LMO2, PU.1, and RUNX1 in Kasumi-1 cells. These sites also form a DHS in normal CD34+ cells and are bound by RUNX1 in these cells, as determined by ChIP (Figure 1B, top).

In contrast to RUNX1, which interacts with a multiplicity of factors in different cell types (Scheitz and Tumbar, 2013; van Riel et al., 2012), RUNX1/ETO preferentially binds to DNA elements containing RUNX, ETS, and E box motifs, thus reflecting the composition of the RUNX1/ETO complex (Sun et al., 2013). To examine whether our footprinting analysis was able to confirm this preference of colocalizing motifs in patient cells, we conducted an unbiased pairwise clustering analysis of footprinted motifs in regions bound by RUNX1/ETO. This analysis demonstrated that motifs bound by RUNX1/ETO in Kasumi-1 cells strongly clustered with ETS (PU.1 and ERG) and E box (SCL, LYL, and HEB) motifs that are footprinted in patient cells (Figure 1C). We found a similar clustering pattern using sequences from the Kasumi-1 ChIP-seq experiments (Figure S1D), although it was less defined due to the larger peak sizes in this experimental context. In conclusion, RUNX1/ETO-positive Kasumi-1 cells show similar transcription-factor motif occupancy patterns, confirming that at this level of accuracy, digital footprinting provides a viable method for investigating transcription-factor binding-site occupancy and preferential interaction in patient cells.

RUNX1/ETO and RUNX1-Containing Complexes Compete for the Same Genomic Sites

We previously showed that more than 60% of RUNX1/ETO binding sites are shared with RUNX1 in the bulk population of cells (Ptasinska et al., 2012), with many of the footprinted sites containing multiple TGYGGT RUNX1-binding motifs (e.g., Figure 1B). Therefore, we conducted re-ChIP experiments in Kasumi-1 cells to test at known RUNX1/ETO binding sites whether the two factors co-occupy single sites or whether binding is mutually exclusive at such sites. In addition, we examined which other factors were shared between RUNX1 and RUNX1/ETO complexes. RUNX1 and RUNX1/ETO both colocalize with LMO2, HEB, and LYL1 in the Kasumi-1 cell population (Figures 1A, 2A, and S2A). However, binding of RUNX1 and RUNX1/ETO to their target sites was mutually exclusive, even at elements containing multiple RUNX motifs, such as the *NFE2* locus (Figures 1B, 2B, 2C, and S2B).

Both RUNX1/ETO and RUNX1 have been shown to interact with HDACs and the HAT p300 (also known as EP300) (Amann et al., 2001; Kitabayashi et al., 1998; Levanon et al., 1998; Reed-Inderbitzin et al., 2006; Wang et al., 2011). Using paralled re-ChIP experiments, we show that RUNX1-bound elements had a preference for binding the coactivator p300, whereas RUNX1/ETO-occupied elements preferentially bound the corepressor HDAC2 (Figures 2D–2F). We further confirmed this preferential binding and the strong association between RUNX1 and p300 by performing manual ChIP and ChIP-sequencing experiments after knockdown of RUNX1/ETO (Figure 3). These experiments demonstrated (1) that the loss of RUNX1/ETO binding led to an

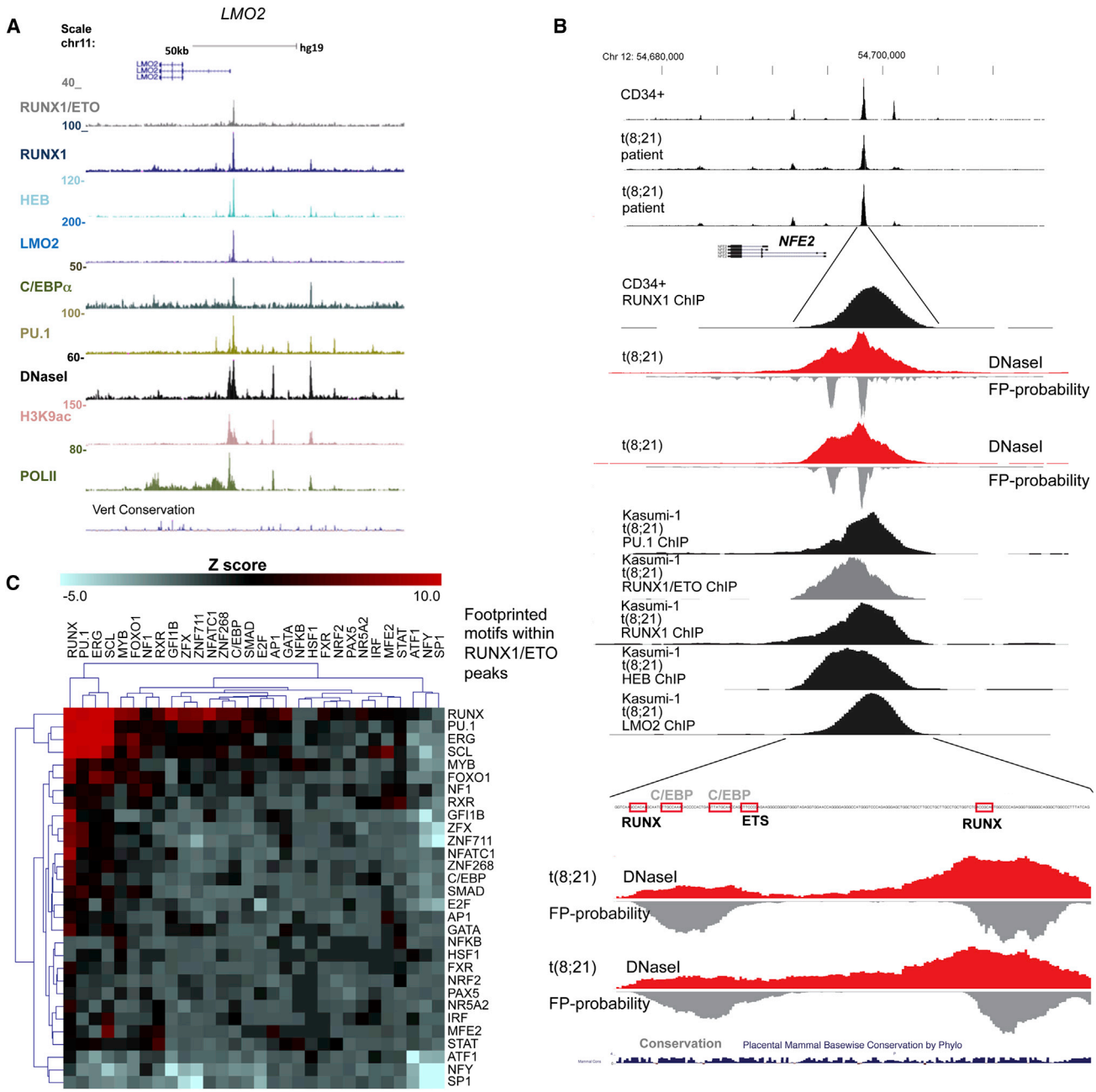


Figure 1. Transcription-Factor Occupancy Patterns Are Similar between RUNX1/ETO-Expressing Cell Lines and Patient Cells

(A) UCSC genome browser screenshot showing the binding patterns of RUNX1/ETO, RUNX1, HEB, LMO2, C/EBPα, PU.1, DHS, H3K9Ac, and RNA-Polymerase II (POLII), as well as input reads and conservation among vertebrates at the *LMO2* locus as aligned reads.

(B) UCSC genome browser screenshot of ChIP-seq and DHS data aligned with digital footprints at the *NFE2* locus within a DHS shared between two t(8;21) patients and purified normal CD34+ cells (top). It also shows the binding pattern of RUNX1 in CD34+ cells and RUNX1/ETO, RUNX1, HEB, LMO2, and PU.1 in Kasumi-1 cells as determined by ChIP. Footprint probabilities as calculated by Wellington are indicated as gray columns below the lines. The bottom indicates the location of occupied RUNX, ETS, and C/EBP motifs.

(C) Occupied RUNX, E box, and ETS motifs in patient cells cluster within DHS sites that colocalize with RUNX1/ETO binding in Kasumi-1 cells. The heatmap shows hierarchical clustering of footprinted motif co-occurrences by *Z* score within RUNX1/ETO peaks, indicating transcription factor co-occupancy. Footprint probabilities within RUNX1/ETO-bound peaks were calculated using DNaseI-seq data from t(8;21) patient 1. The motif search was done within RUNX1/ETO footprint coordinates. Red and blue colors indicate statistically over- and underrepresented motif co-occurrences, respectively. For a more detailed explanation, see the legend of Figure S1 and the Supplemental Experimental Procedures.

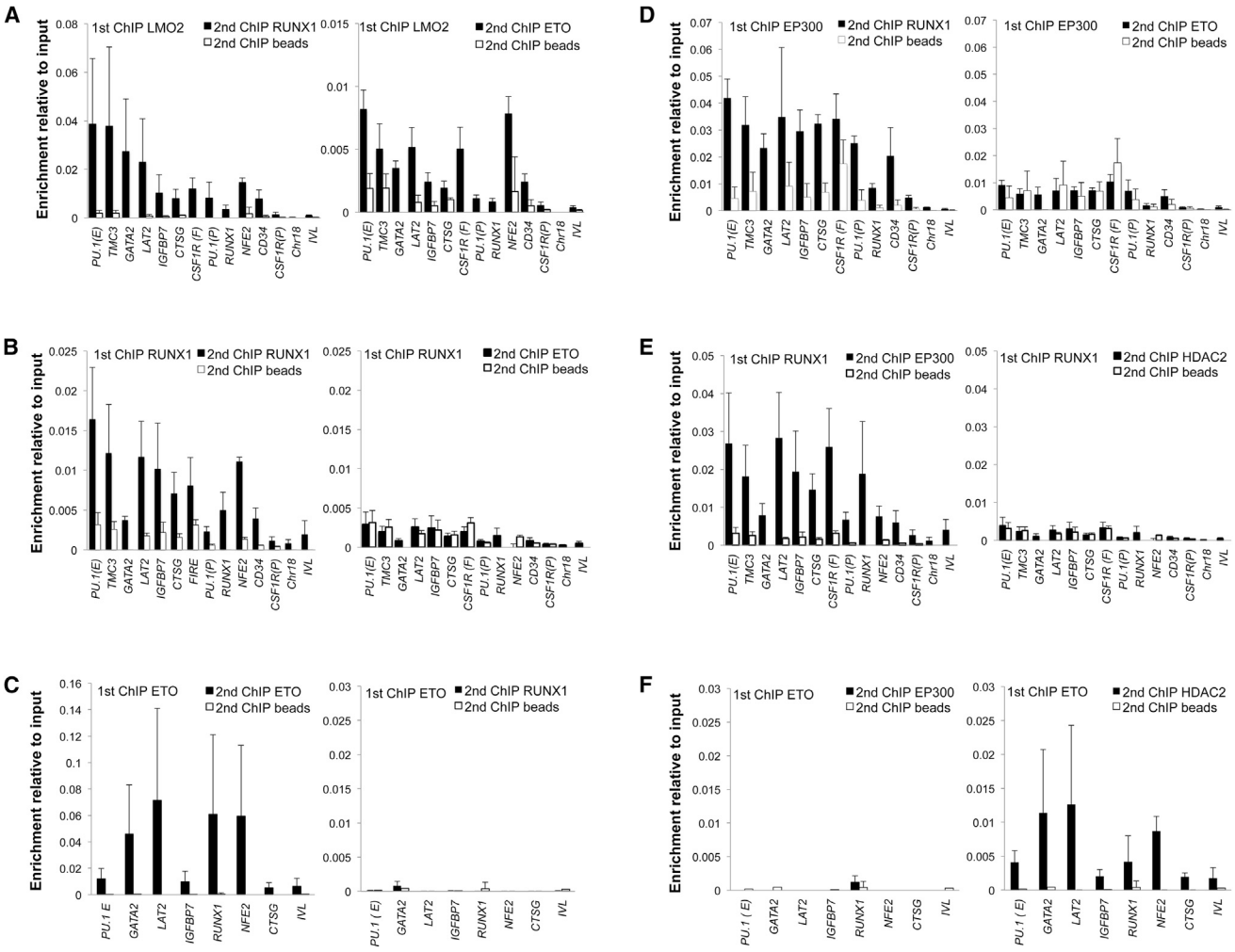


Figure 2. RUNX1 and RUNX1/ETO Complexes Differentially Interact with Coactivator and Corepressor Complexes, and Binding to the Same Sites Is Mutually Exclusive

(A–E) Multiple RUNX1/ETO binding sequences and control sequences (*IVL*, Chr18) were selected and validated for factor binding by a first round of ChIP followed by a second round with a different antibody or with just beads as indicated. All of the chosen binding sites contain several RUNX1 motifs (data not shown).

(A) LMO2 associates with both RUNX1 and RUNX1/ETO.

(B and C) RUNX1 and RUNX1/ETO binding is mutually exclusive. Control ChIPs were performed with the same antibody.

(D) EP300 associates with RUNX1, but not RUNX1/ETO.

(E and F) RUNX1 preferentially binds p300, whereas RUNX1/ETO preferentially associates with HDAC2. For additional amplicons, see Figure S2B. qPCR data represent the mean ± SD of at least three independent experiments.

increase in RUNX1 binding at the same sites, and (2) there was an increased recruitment of p300 without a concomitant increase in the expression of these factors (Figures 3 and S3A), providing an explanation for the increased histone H3 lysine 9 acetylation at such sites that we observed previously (Figure S3B; Ptasinska et al., 2012). In contrast, knockdown of RUNX1/ETO led to a reduction of HDAC2 binding to these target sites (Figure 3C). Taken together, these data show that RUNX1/ETO and RUNX1 (1) compete for the same genomic sites and (2) colocalize with the same transcription factors but have distinct preferences for histone-modifying cofactors, with RUNX1 associated complexes preferring to interact with p300 and RUNX1/ETO complexes preferring to recruit HDACs, including HDAC2.

The Core Transcriptional Network Bound by RUNX1/ETO Is Predominantly Associated with Repressed Genes

We next analyzed our ChIP-seq data sets to identify the core transcriptional network that characterizes the cellular identity of t(8;21) cells by determining overrepresented combinatorial binding patterns for the transcription factors RUNX1/ETO, C/EBPα, HEB, LMO2, PU.1, and RUNX1 (Tijssen et al., 2011). ChIP sequences in RUNX1/ETO-positive cells were enriched for just 11 of the 63 possible different binding patterns, which included six significantly enriched combinatorial patterns containing RUNX1/ETO and five patterns that did not (Figure 4A, marked by asterisks). Two possible binding patterns (111010 and 110011) were not observed. We then associated such

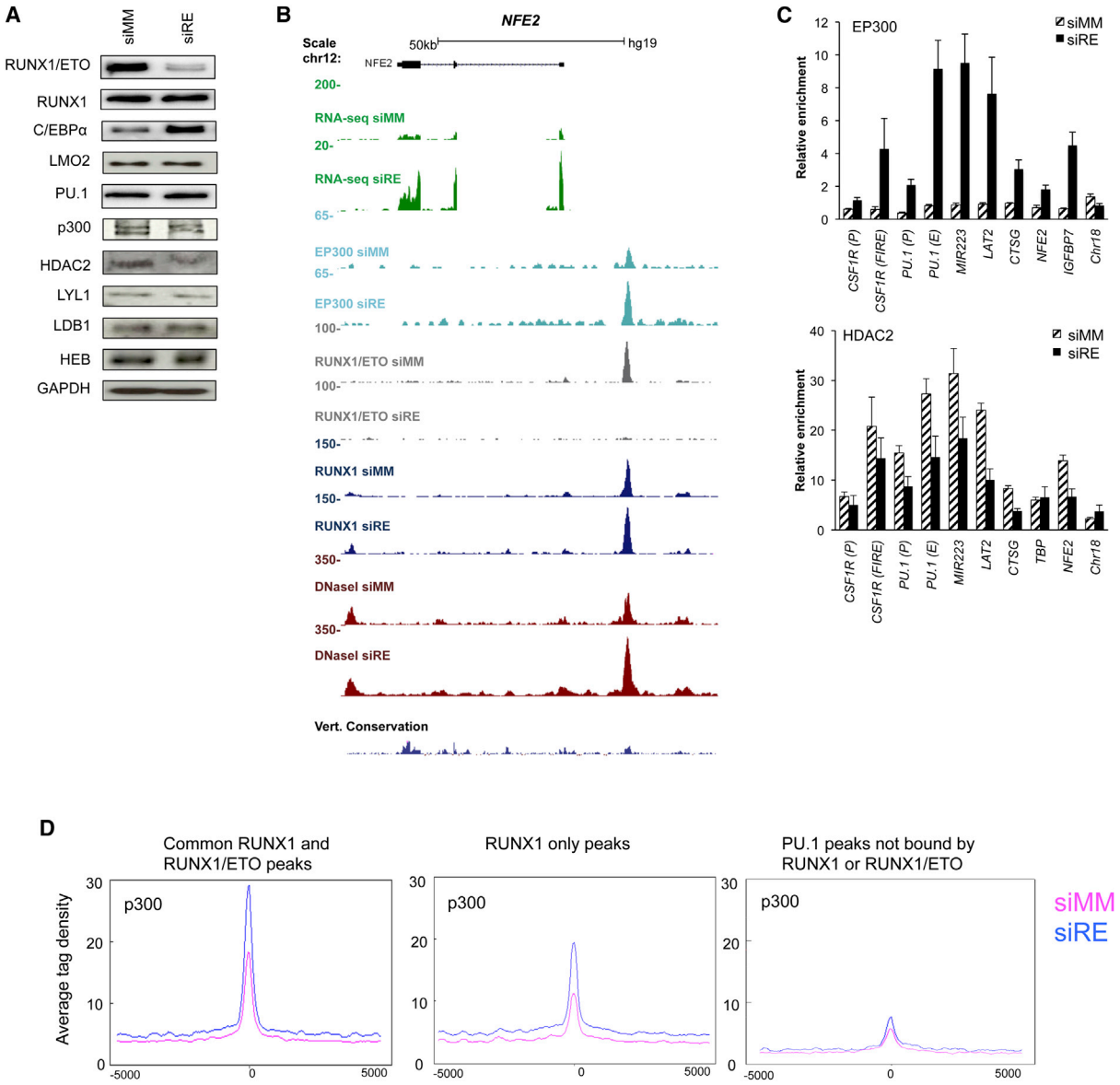


Figure 3. Dynamic Alterations in Cofactor Binding upon RUNX1/ETO Knockdown

(A) Western blot detecting RUNX1/ETO, RUNX1, C/EBPα, LMO2, PU.1, p300, HDAC2, LYL1, LDB1, and HEB protein in Kasumi-1 cells treated for 48 hr with mismatch control siRNA (siMM) and with RUNX1/ETO siRNA (siRE). GAPDH served as the loading control.

(B) UCSC genome browser screenshot of the *NFE2* locus showing changes in the RNA expression and binding pattern of p300, RUNX1/ETO (R/E), RUNX1, and DHS upon RUNX/ETO knockdown in Kasumi-1 cells.

(C) Increase of p300 binding and decrease of HDAC2 binding upon RUNX1/ETO knockdown.

(D) Global changes of p300 binding peaks shared between RUNX1/ETO and RUNX, peaks exclusively bound by RUNX1, and PU.1 peaks not associated with RUNX1/ETO or RUNX1 binding. qPCR data represent the mean ± SD of three to five independent experiments. For other control analyses, see Figure S3B.

elements with the nearest genes and performed a gene set enrichment analysis (GSEA) using gene-expression data sets derived from a time course of RUNX1/ETO knockdown in two different t(8;21) cell lines (Figures S4A and S4B; Ptasinska et al., 2012). In addition, we compared these gene signatures with a RNA-seq-based gene-expression data set derived from a 4-day RUNX1/ETO knockdown in Kasumi-1 cells (Figures 4B and S4C). This analysis demonstrated that all overrepresented RUNX1/ETO-containing binding patterns were associated with the upregulation of gene expression upon knockdown (Figure 4B, red asterisks), whereas loci that do not bind RUNX1/ETO were enriched in genes that were downregulated after RUNX1/ETO knockdown (green asterisks). The very same genes behaved similarly when assayed after knockdown of RUNX1/ETO in patient cells, confirming the similarity between cell lines and primary cells (Figure 4C).

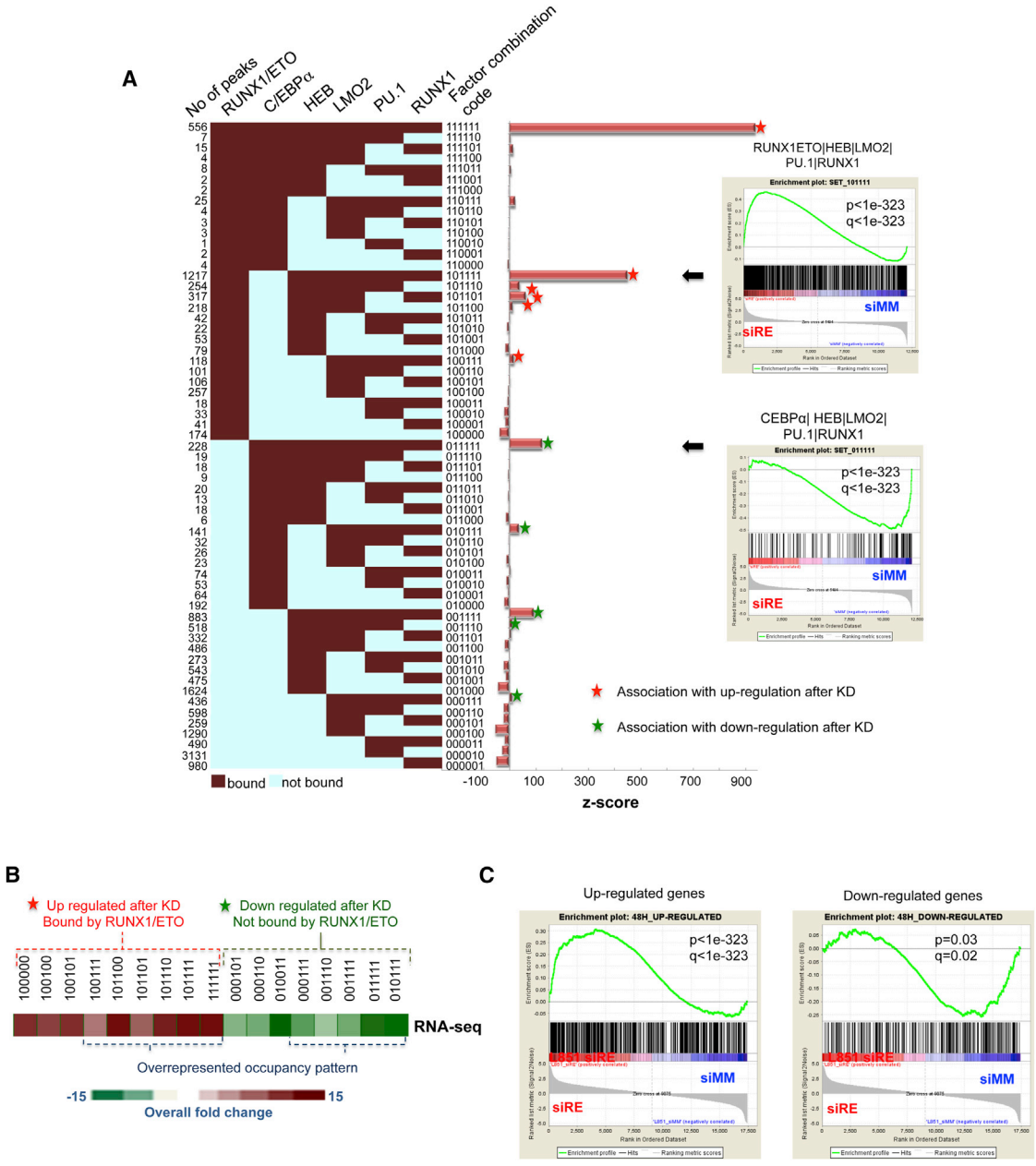


Figure 4. Specific Transcription-Factor Binding Patterns in t(8;21) Cells Correlate with the Response to RUNX1/ETO Knockdown

Genes bound by RUNX1/ETO are preferentially upregulated, whereas genes not bound by RUNX1/ETO are preferentially downregulated.

(A) Analysis of combinatorial binding identifies prevalent patterns in Kasumi-1 cells. The numbers of peaks are shown on the left of the heatmap for 61 factor-binding combinations (red: bound, scored as 1; blue: not bound, scored as 0 with the order of factors as depicted on top of the heatmap). *Z* scores on the right indicate the significance of deviation between observed and expected instances for all 61 combinatorial binding patterns. We identified 11 overrepresented binding patterns, which we analyzed further when each was associated with more than 100 genes. GSEA of selected large groups of genes (indicated by arrows) shows a highly significant enrichment of genes upregulated (upper left) or downregulated (lower left) after 4 days of RUNX1/ETO knockdown.

(B) Heatmap showing the RNA-seq overall fold change in Kasumi-1 cells 4 days after RUNX1/ETO knockdown.

(C) GSEA plots showing enrichment for up- or downregulated genes associated with dominant binding patterns in patient cells subjected to RUNX1/ETO knockdown, demonstrating that changes in gene expression were concordant between Kasumi-1 and patient cells after RUNX1/ETO knockdown. Note that in patient cells, RUNX1/ETO was only depleted for 48 hr and it takes about 4 days for the majority of genes to be downregulated (Ptasinska et al., 2012), thus explaining the lower p value seen with downregulated genes.

See also Figure S4.

Using the different overrepresented binding patterns, we constructed an interacting transcriptional network (Figure S4D). Most genes were regulated by a single binding pattern (node), and only some of these genes were associated with *cis* elements that bound different factor combinations (depicted as located between nodes). This specific binding pattern is of biological relevance because the genes that occupied the different network nodes clustered by overlapping but distinct Gene Ontology (GO) terms and KEGG pathways (Figures S4D and S4F; Table S2), indicating that they perform different functions. For example, *cis*-regulatory elements that bind RUNX1/ETO and all other factors (pattern 111111) are associated with genes involved in myeloid differentiation and hematopoiesis (Figure S4E; Table S2). Among the genes without RUNX1/ETO binding (pattern 011111) that were downregulated after RUNX1/ETO knockdown, we found the transcription factor genes *ERG* and *ETV6* (*TEL1*) (Figure S4F; Table S2), both of which are important for stem cell function and maintenance (Taoudi et al., 2011; Wang et al., 1998) but also have been implicated in AML (Diffner et al., 2013). ERG has also been shown to be important for stabilization of the RUNX1/ETO complex (Martens et al., 2012). Another downregulated transcription factor gene was *MEF2C*, which encodes a transcription factor that modulates myeloid fate and has oncogenic activity when overexpressed (Schwieger et al., 2009).

In summary, our analysis of the RUNX1/ETO-responsive core transcriptional network in t(8;21) cells highlights the predominantly repressive role of RUNX1/ETO within this network. Moreover, our analysis identified distinct classes of genes, with repressed genes involved in myeloid differentiation and active genes forming part of the stem cell signature.

Knockdown of RUNX1/ETO Leads to a Dynamic Reorganization of Transcription-Factor Binding

We next examined how the t(8;21) core transcriptional network changed 2 days after RUNX1/ETO depletion. Depletion had no immediate influence on the expression levels of any of the other factors studied above, with the notable exception of C/EBPα (Figure 3A). Nevertheless, loss of RUNX1/ETO had a profound effect on the binding of these transcription factors (Figure S5A). As exemplified by the *CEBPE* locus, depletion led to increased RUNX1 occupancy at several thousand sites, confirming that RUNX1/ETO and RUNX1 binding are in equilibrium (Figures 5A, top left, 5B, S5B, and S5C). Furthermore, increased RUNX1 occupancy, including RUNX1 sites that were not previously bound by RUNX1/ETO, was associated with a strong increase in p300 binding (Figure 3D). In contrast, more than 3,000 LMO2 binding sites were lost, mainly outside the regions bound by RUNX1/ETO and RUNX1 (Figures 5A, bottom-right panel, and S5C). Furthermore, whereas 80% of all PU.1 binding sites remained unchanged, the number of sites bound by C/EBPα increased 4-fold. Interestingly, 65% of all C/EBPα de novo sites colocalized with PU.1 (Figures 5A, top left, S5B, and S5D). In agreement with these results, C/EBPα binding sites clustered more strongly with both RUNX1 and PU.1 sites upon depletion of RUNX1/ETO (Figure S5E).

The changes in RUNX1 and C/EBPα binding, however, were not reflected by major global changes in DHS patterns. The comparison of DHS profiles before and after 2 days of RUNX1/ETO knockdown revealed that the majority of DHSs were unchanged (Figure 5C). Both C/EBPα and RUNX1 mainly associated with DHSs that were already present before RUNX1/ETO depletion. Only 20% of sites showed increased DNaseI sensitivity or arose de novo following RUNX1/ETO knockdown coinciding with de novo RUNX1 and C/EBPα binding (Figures S5F and S5G).

In summary, knockdown of RUNX1/ETO led to immediate genome-wide alterations in transcription-factor binding after 48 hr. Although a small fraction of binding sites arose de novo, this reprogramming occurred predominantly within preexisting transcription-factor assemblies.

The Dynamic Reorganization of the Leukemic Transcriptional Network after RUNX1/ETO Depletion Is Driven by C/EBPα

Many transcription factors upregulate the expression of their own gene, with *PU.1* (*SPI1*) being a prominent example (Leddin et al., 2011; Staber et al., 2013). However, of all the transcription factors examined, only C/EBPα was found to be significantly increased after RUNX1/ETO depletion (Figure 3A). Similarly to PU.1, C/EBPα upregulates its own expression in murine cells, and it was previously suggested that RUNX1/ETO interferes with C/EBPα expression by sequestering it from its promoter and thereby suppressing autoactivation (Pabst et al., 2001a). Our data demonstrate binding of C/EBPα to an element about 40 kb downstream of its own gene, a site that is also occupied by RUNX1/ETO, suggesting a more direct mechanism of repression (Ptasinska et al., 2012). C/EBPα is absolutely essential for terminal myeloid differentiation (Zhang et al., 1997) and occupies a large number of binding sites in mature macrophages (Heinz et al., 2010). However, *CEBPA* is not the only direct target gene of the *CEBP* family that responds to RUNX1/ETO: *CEBPE* and *CEBPD* are upregulated as well (Ptasinska et al., 2012), indicating that these factors may be part of a wider network of C/EBP proteins that control myeloid gene expression.

To test whether increased expression of C/EBPα was crucially involved in shifting the transcriptional network after RUNX1/ETO depletion, we defined overrepresented binding patterns for C/EBPα, PU.1, RUNX1, and LMO2 after RUNX1/ETO knockdown. Loss of RUNX1/ETO resulted in the formation of a transcriptional network dominated by C/EBPα-containing binding patterns, all of which were predominantly associated with upregulated genes in RUNX1/ETO-depleted Kasumi-1 and patient cells (Figures 6A–6C, S6A, and S6B; Table S3). Different patterns were again indicative of different classes of genes in terms of both GO and pathway analyses, with differentiation and signal transduction pathways being prominently featured (Figures S6C, S6D, and S7A). However, increased C/EBPα binding was also observed with a subset of genes that were downregulated (Figure 6D). Previous studies have shown that in addition to C/EBPα's role in driving myeloid differentiation, low levels of C/EBPα are required for stem cell maintenance, as upregulation of C/EBPα represses genes required for stem-cell self-renewal (Zhang et al., 2004, 2013). Therefore, we identified genes that (1) were downregulated after RUNX1/ETO knockdown and (2) showed increased C/EBPα binding (a total of 145 genes met

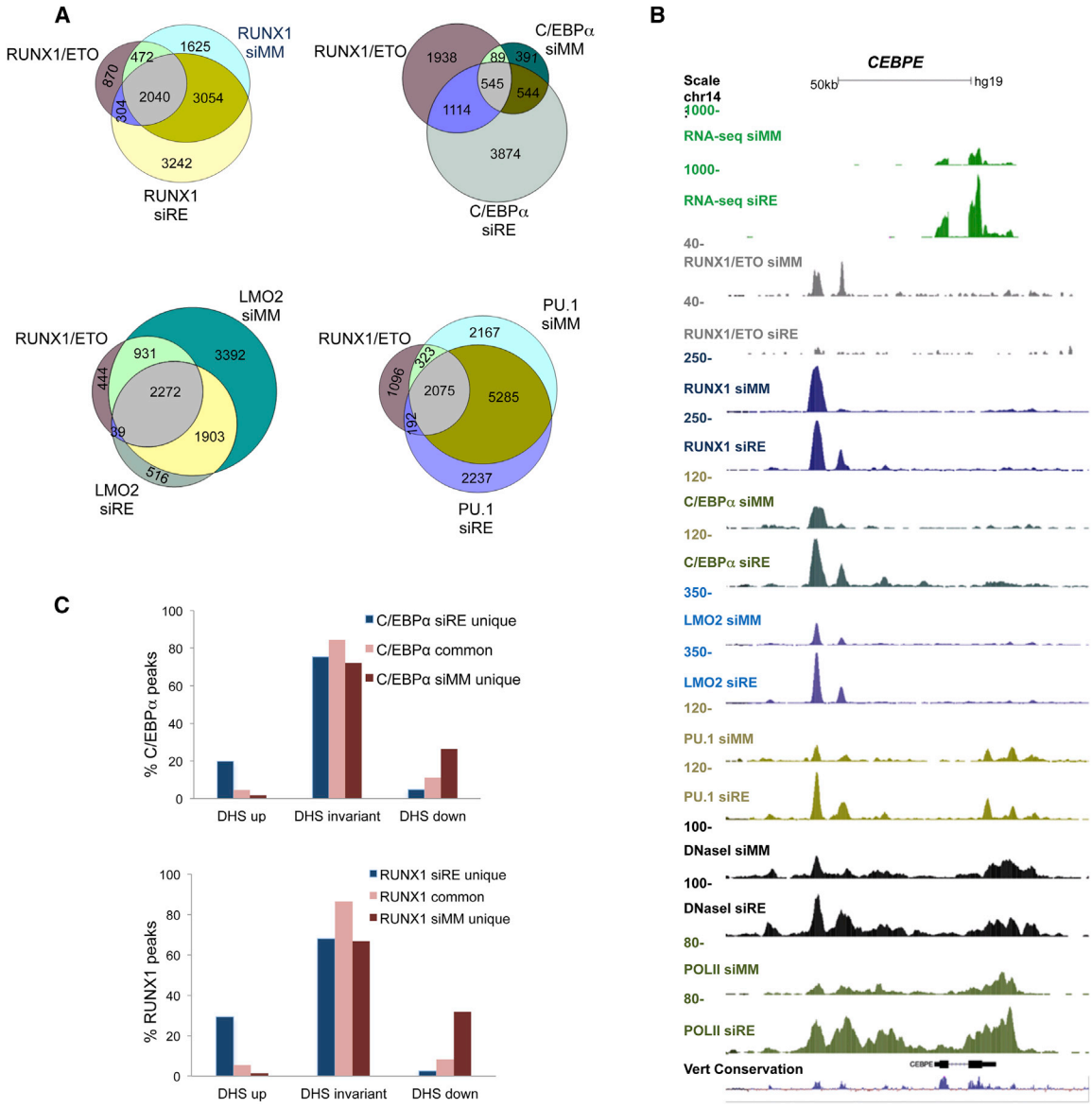


Figure 5. Knockdown of RUNX1/ETO Leads to a Reorganization of Transcription-Factor Assemblies within Preexisting Open Chromatin Regions

(A) Three-way Venn diagrams showing the overlap between RUNX1/ETO and RUNX1 (top left), CEBPα (top right), LMO2 (bottom left), and PU.1 (bottom right) in Kasumi-1 cells treated for 48 hr with control (siMM) and with RUNX1/ETO siRNA (siRE).

(B) UCSC genome browser screenshot showing the binding pattern of the indicated factors at the *CEBPE* locus in Kasumi-1 cells treated for 48 hr with control siRNA (siMM) and with RUNX1/ETO siRNA (siRE).

(C) Binding of de novo (siRE unique), common, and lost (siMM unique) transcription factors (C/EBPα (top) and RUNX1 (bottom) to regions of increased (DHS up), unchanged (DHS invariant), or reduced DNaseI hypersensitivity.

See also Figure S5.

the latter criterion; Figure 6D). This category included stem cell genes such as *ERG* and *CD34* (Figures S6F and S6G), as well as a large number of genes encoding for signaling molecules that are involved in regulating proliferation and differentiation, such as *DUSP6* or *PTK2* (Figure S6G).

We next evaluated whether C/EBPα was required for the upregulation of repressed RUNX1/ETO target genes. For this purpose, we depleted RUNX1/ETO with and without a concomitant C/EBPα knockdown. Knockdown of RUNX1/ETO led to a 2-fold increase in *C/EBPα* expression (Figures 3A, 7A, and 7B) and increases in expression of the direct RUNX1/ETO target genes, including *MS4A3*, *NKG7*, and *RNASE2*, which all show increased C/EBPα binding upon RUNX1/ETO depletion (Figures 7C and 7D; data not shown). Codepletion of C/EBPα diminished the induction of the three target genes in both Kasumi-1 and SKNO-1 cells (Figures 7D and S7B–S7D). These data indicate that

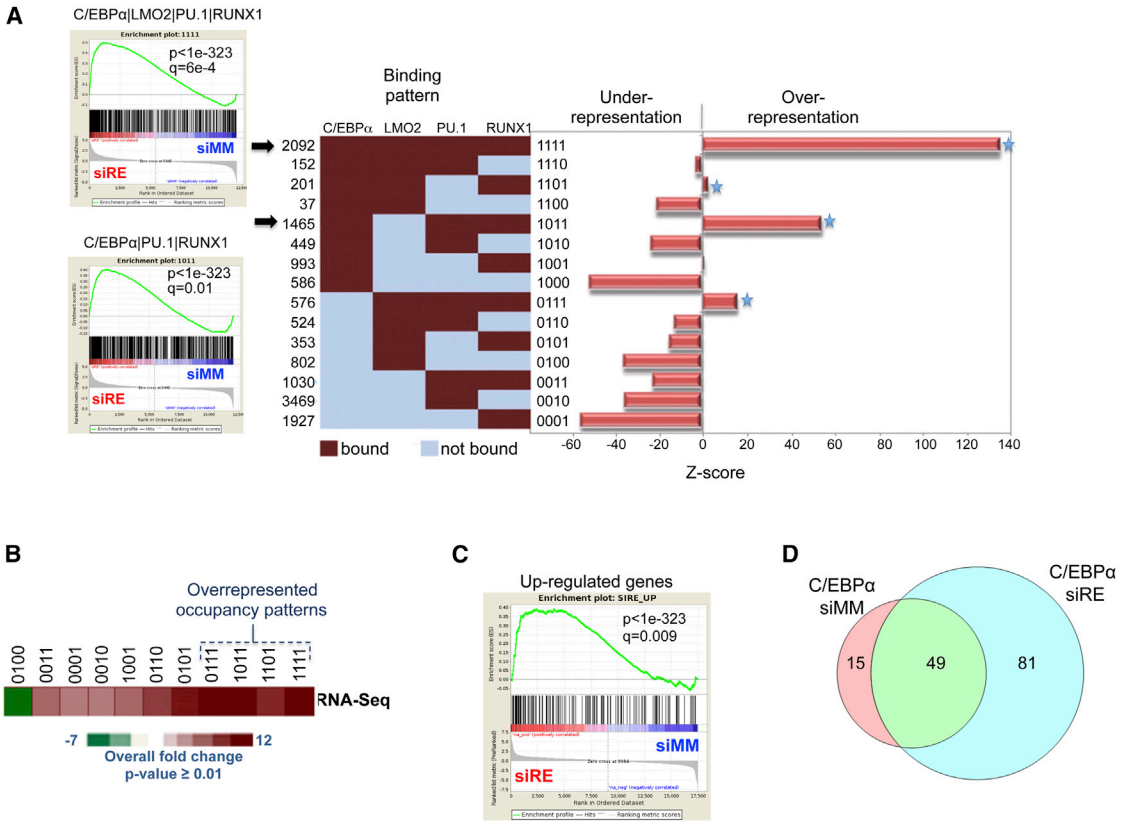


Figure 6. Transcriptional Network after RUNX1/ETO Depletion Is Enriched for C/EBPα Target Genes

(A) The transcription-factor binding state for CEBPα, LMO2, PU.1, and RUNX1 after RUNX1/ETO knockdown is characterized by an overrepresentation of four dominant occupancy patterns. The number of peaks for all 15 factor combinations is shown on the left of the heatmap (red: bound, scored as 1; blue: not bound, scored as 0). *Z* scores on the right indicate the significance of deviation between observed and expected instances for all 15 binding patterns. Left: GSEAs of genes associated with the two most enriched dominant occupancy patterns (indicated by arrows) show highly significant enrichment of upregulated genes after RUNX1/ETO knockdown.

(B) Genes associated with specific occupancy patterns that significantly change expression as measured by RNA-seq 4 days after RUNX1/ETO knockdown. The heatmap shows the RNA-seq overall fold change in Kasumi-1 cells 4 days after RUNX1/ETO knockdown.

(C) GSEAs showing that genes associated with dominant occupancy patterns that are upregulated in Kasumi-1 cells behave similarly in patient cells.

(D) Venn diagram depicting the number of genes bound by C/EBPα that are downregulated after RUNX1/ETO knockdown and show increased C/EBPα binding. See also Figure S6.

derepression of C/EBPα caused by RUNX1/ETO depletion is required for the full upregulation of a number of RUNX1/ETO target genes. However, we cannot rule out a similar function for other C/EBP members and in particular C/EBPδ and C/EBPε, which are both upregulated upon RUNX1/ETO knockdown (Figure 5B and data not shown). Nevertheless, our data confirm that C/EBPα plays an important role in orchestrating a transcriptional network that drives myeloid differentiation downstream of the original RUNX1/ETO network (Figure 7E).

DISCUSSION

The study presented here shows that expression of the oncogenic transcription factor RUNX1/ETO interferes with the hierarchical succession of transcriptional networks required for myeloid differentiation. Binding of RUNX1/ETO to key regulatory elements inhibits the expression of genes that drive differentiation. Moreover, we show that the establishment of a stable leukemic state not only depends on a static interaction of transcription factor complexes but also contains a dynamic competitive component as its key feature. We demonstrate that the transcriptional network controlled by RUNX1/ETO depends on a dynamic equilibrium between RUNX1/ETO and RUNX1 complexes, whose binding to their target sites is mutually exclusive. Although these complexes share the factors LMO2, HEB, and LYL1, they differ in their preferences for histone modifiers. RUNX1 can also act as a repressor (Levanon et al., 1998; Reed-Inderbitzin et al., 2006; Taniuchi et al., 2002), but in this factor context it preferentially recruits the HAT p300, whereas RUNX1/ETO recruits histone deacetylases, including HDAC2. RUNX1/ETO shares almost three-quarters of its binding sites with RUNX1, suggesting that the equilibrium between these two complexes results in a finely tuned modulation of expression for a wide range of genes. Thus, the leukemic phenotype requires the downmodulation of genes associated with differentiation, but may not tolerate their complete suppression.

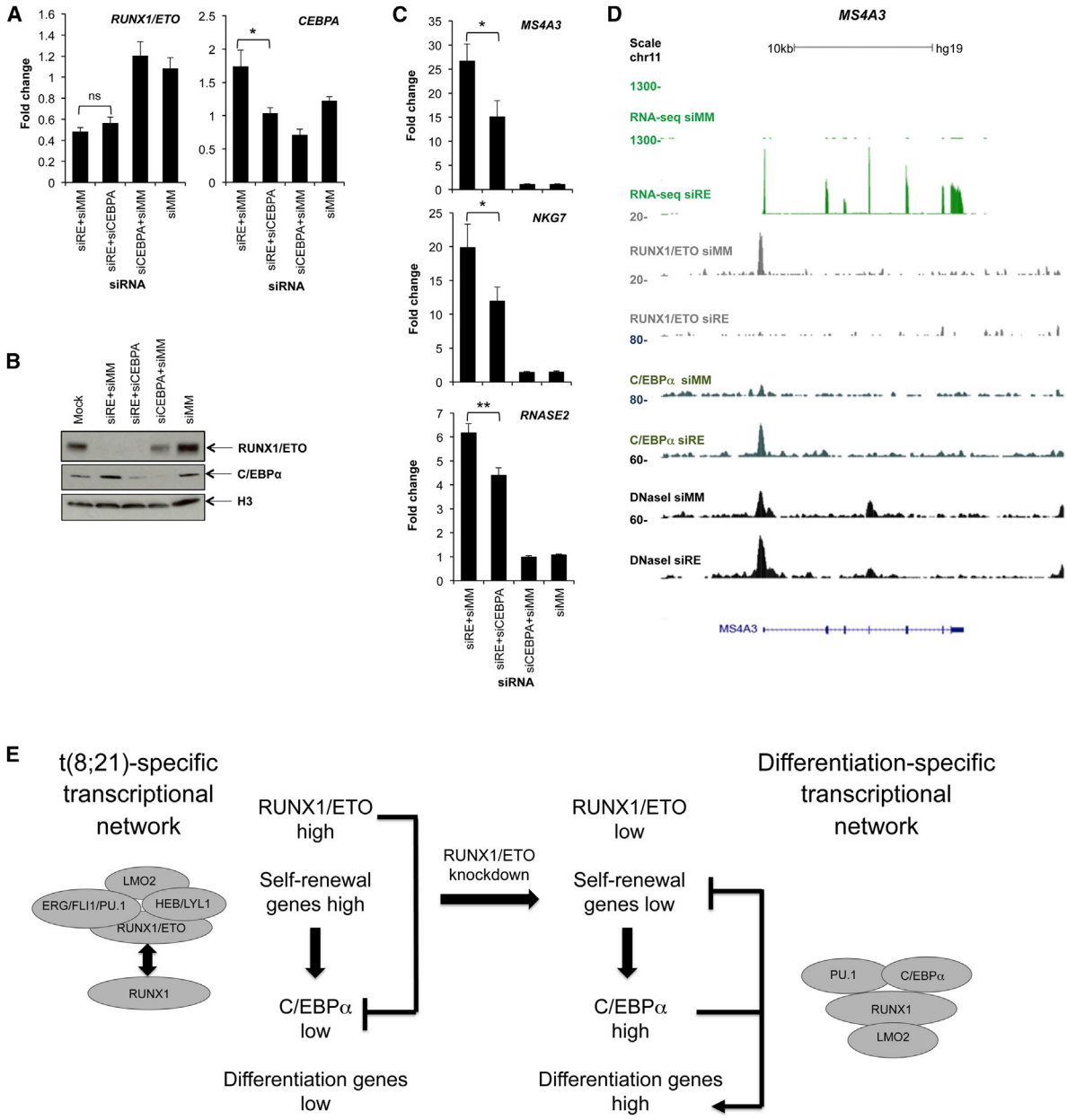


Figure 7. Loss of RUNX1/ETO Triggers C/EBPα-Driven Reorganization of the Leukemic Transcriptional Network

(A) *RUNX1/ETO* and *CEBPA* mRNA expression levels in Kasumi-1 cells 72 hr after electroporation with the indicated siRNAs. siRE, RUNX1/ETO siRNA; siCEBPA, C/EBPα siRNA; siMM, mismatch control siRNA. Results represent the mean ± SEM of five independent experiments. *$p < 0.05$; ns, not significant by paired Student's t test.

(B) Western blot indicating RUNX1/ETO and C/EBPα protein expression levels in single- and double-knockdown cells as indicated. An antibody against H3 was used as control. Mock, no siRNA.

(C) mRNA levels of *MS4A3*, *NKG7*, and *RNASE2* 72 hr after electroporation with the indicated siRNAs. Results represent the mean ± SEM of five independent experiments. *$p < 0.05$, **$p < 0.01$ by paired Student's t test.

(D) UCSC genome browser screenshot showing the binding pattern of RUNX1/ETO, C/EBPα, and DHSs at the *MS4A3* locus in Kasumi-1 cells treated for 48 hr with mismatch control siRNA (siMM) and with RUNX1/ETO siRNA (siRE).

(E) Model of RUNX1/ETO-mediated control of leukemic transcription. The competitive equilibrium in locus occupation between RUNX1/ETO and RUNX1 complexes drives leukemic self-renewal. Depletion of RUNX1/ETO increases the levels and DNA binding of its direct target gene, C/EBPα, which together with other differentiation genes reinstalls a transcriptional program that promotes myeloid differentiation.

See also Figure S7.

Consequently, perturbation of this equilibrium by depletion of RUNX1/ETO leads to loss of self-renewal, whereas knockdown of RUNX1 severely impairs viability (Ben-Ami et al., 2013; Dunne et al., 2006; Martinez et al., 2004; Martinez Soria et al., 2009). Currently, we do not know whether the different complexes exist independently or are in a rapid exchange. Evidence for both mechanisms exists; for example, in a previous study (Sun et al., 2013), neither p300 nor HDACs could be purified together with the RUNX1/ETO complex from t(8;21) cells using high stringency conditions. However, immunohistochemistry has demonstrated that RUNX1 and RUNX1/ETO are targeted to different subnuclear compartments (McNeil et al., 1999), a scenario that would be difficult to reconcile with a rapid exchange of factors binding to the same region of chromatin. Whatever the mechanism, it is likely that a mutually exclusive binding pattern can be found in other CBF leukemias. A similar colocalization with RUNX1 and its mutated counterpart has also been seen in AML with inversion 16 carrying the CBFB-MYH11 fusion protein (Mandoli et al., 2014), and furthermore, this type of AML is also dependent on the presence of an active copy of RUNX1 (Ben-Ami et al., 2013).

It was recently shown that aberrant RUNX1 expression is required for the maintenance of epithelial cancers (Scheitz et al., 2012). Moreover, RUNX1 plays a tumor-suppressive role by interacting with estrogen receptor α, and ERα-positive breast cancer patients carry mutations that disrupt these interactions (Chimge and Frenkel, 2013; Stender et al., 2010), highlighting increasing evidence that this factor and its deregulation or mutation are at the heart of multiple pathological processes. Moreover, alternative splicing of *RUNX1* leads to a C-terminally truncated isoform known as AML1a, which lacks the transactivation domain and promotes self-renewal of hematopoietic stem cells (Tsuzuki and Seto, 2012). We previously showed that during blood cell development, RUNX1 binding reshapes the epigenetic landscape by attracting other factors to its binding sites, and that this factor relocation is reversible (Lichtinger et al., 2012). Therefore, a dynamic equilibrium between different RUNX1 isoforms and other factors may also be relevant for cancers outside of the hematopoietic system.

A second important finding of our study is that the destruction of the RUNX1/ETO network establishes a transcription network dominated by the combinatorial binding of PU.1, RUNX1, and, in particular, C/EBPα (Figure 7E). Once RUNX1/ETO is depleted, C/EBPα expression levels increase and this factor then occupies a large number of binding sites, demonstrating at the genome-wide level that (1) C/EBPα is a major driver of myeloid differentiation and (2) the differentiation block in AML is partly caused by C/EBPα downregulation. The latter observation is consistent with the fact that a large number of AMLs involve mutations of C/EBPα (Preudhomme et al., 2002). However, the majority of binding sites are found in regions of previously accessible chromatin, indicating that (1) RUNX1/ETO targets binding sites that are destined for differentiation-driven factor exchange, and (2) shortly after its upregulation, C/EBPα resumes its original binding behavior and reorganizes existing transcription factor assemblies to drive myelopoiesis. These results tie in with the finding that PU.1 binding was largely invariant before and after RUNX1/ETO depletion. Although previous overexpression experiments indicated that RUNX1/ETO inactivated PU.1 (Vangala et al., 2003), our data indicate that, at least during the time window of 2 days, the PU.1 cistrome is largely unperturbed by the presence or absence of RUNX1/ETO and forms a platform upon which other factors dynamically assemble (Natoli et al., 2011).

In summary, our work sheds light on global mechanisms of the differentiation block in t(8;21) AML, which is of conceptual relevance for other types of AML and even other cancers. Many AML types are characterized by mutations in C/EBPα and RUNX1, which would impact many of the binding sites described here. The dynamic equilibrium between a mutated transcription factor and its wild-type counterpart allows a rapid reversion from a transcriptional program promoting malignant self-renewal to a differentiation program. Such dynamic behavior is likely to be the molecular cause of the good prognosis of t(8;21) AML and may also be a major angle for therapeutic intervention in other types of AML without mutations in other hematopoietic regulators.

EXPERIMENTAL PROCEDURES

More detailed descriptions of the materials and methods used in this work can be found in the Supplemental Experimental Procedures.

Human Patient Cells and Cell Lines

Patient material was obtained with approval from the NHS Research Ethics Committees (Leeds Teaching Hospitals NHS Trust and Newcastle upon Tyne Hospitals NHS Foundation Trust). Kasumi-1 cells were obtained from the DSMZ cell line repository (http://www.dsmz.de/) and were cultured in RPMI1640 containing 10% fetal calf serum (FCS). SKNO-1 cells were maintained in RPMI1640 supplemented with 20% FCS and 7 ng/ml granulocyte-macrophage colony-stimulating factor.

siRNA Transfections

Kasumi-1 and SKNO-1 cells were transfected with 200 nM siRNA using a Fischer EPI 3500 electroporator (Fischer) as described previously (Ptasinska et al., 2012). The following siRNAs were used: RUNX1/ETO siRNA (sense, CCUCGAAAUCGUACUGAGAAG; antisense, UCUCAGUACGAUUUCGAGG UU), mismatch control siRNA (sense, CCUCGAAUUCGUUCUGAGAAG; antisense, UC UCAGAACGAAUUCGAGGUU); and C/EBPα siRNA (sense, CCG GAGUUAUGACAAGCUUUC; antisense, AAGCUUGUCAUAACUCCGGUC).

Real-Time RT-PCR

RNA extraction and quantitative real-time RT-PCR were performed as described previously (Ptasinska et al., 2012). Primers are listed in Table S4.

Western Blotting

Kasumi-1 cells were lysed in RIPA buffer 2 days after electroporation. The following antibodies were used for western blot analysis: C/EBPα, ab15048 (Abcam); ETO, SC-9737 (Santa Cruz Biotechnology); GAPDH, ab8245 (Abcam); HDAC2, ab7029 (Abcam); HEB, SC-357 (Santa Cruz); LDB1, SC-11198 (Santa Cruz); LMO2, AF2726 (R&D Systems); LYL1, SC-374164 (Santa Cruz); PU.1, SC-352 (Santa Cruz); p300, SC-585 (Santa Cruz); and RUNX1, PC285 (Millipore).

ChIP

ChIP assays were performed as described previously (Ptasinska et al., 2012). Nuclei were essentially prepared as described previously (Lefevre et al., 2003). The following antibodies were used: C/EBPα, SC-61 (Santa Cruz Biotechnology); ETO (C terminus specific), SC-9737 (Santa Cruz); HDAC2, SC-6296 (Santa Cruz); HEB, SC-357 (Santa Cruz); LMO2, AF2726 (R&D Systems); LYL1, SC-374164 (Santa Cruz); PU.1, SC-352 (Santa Cruz); p300, SC-585 (Santa Cruz); RUNX1 (C terminus specific), ab23980 (Abcam) or IgG rabbit 12-370 (Millipore); IgG goat, SC-2346 (Santa Cruz); and IgG mouse,

SC-2025 (Santa Cruz). Precipitated material was subjected to library preparation and run on an Illumina Hiseq 2000 sequencer.

RNA-Seq

RNA samples from three independent biological replicates were processed using the Tru-seq RNA Sample Prep Kit v2 (Illumina) according to the manufacturer's protocol. Libraries were run in 4× multiplex on an Illumina Hiseq 2000 sequencer generating ~90 million paired-end reads per sample.

Re-ChIP

Re-ChIP was carried out as described above with minor modifications. Following the final ChIP wash, chromatin complexes were eluted twice in 50 μl of ChIP elution buffer (100 mM $NaHCO_3$, 1% SDS, PIC) for 15 min at room temperature with shaking. Eluates were combined and diluted 20 times with ChIP dilution buffer, followed by a 5 hr incubation with the second primary antibody or IgG. After elution with 100 mM $NaHCO_3$, 1% SDS for 30 min at room temperature, the re-ChIP products were analyzed by quantitative PCR (qPCR). Fold-enrichment values were calculated relative to a negative control region of the genome. Primers are listed in Table S4.

DHS Mapping

Genome-wide DHSs were mapped as described previously (Leddin et al., 2011).

Library Generation and Sequencing

Libraries of DNA fragments from ChIP or DNase I treatment were prepared from 10 ng of DNA according to standard procedures. ETO, RUNX1, C/EBPα, PU.1, LMO2 ChIP, and Kasumi-1 DNase I libraries were sequenced on an Illumina Genome Analyzer GAIIx using 36 bp single-end reads. For patients 1 and 2, DNase I (491 and 342 million reads, respectively) and control patient libraries (Table S1) were sequenced on an Illumina HiSeq using 50 bp single-end reads.

ACCESSION NUMBERS

The GEO accession numbers for the data reported in this paper are GSE29225 (Ptasinska et al., 2012) and GSE54478.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, and four tables and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2014.08.024.

AUTHOR CONTRIBUTIONS

A.P., M.R.I., N.M.-S., A.P., M.W., S.J., and M.H. performed experiments. S.A.A., P.C., J.P., S.O., D.R.W., and D.W. analyzed data. D.G.T. provided technical infrastructure and helped write the manuscript. P.N.C. supervised experiments and helped write the manuscript. C.B. and O.H. conceived the study, supervised experiments, and wrote the manuscript.

ACKNOWLEDGMENTS

The authors thank Simon Bomken, Luke Gaughan, and John Lunec for carefully reading and improving the manuscript. Research in the C.B. lab is supported by grants from Leukaemia & Lymphoma Research (7001 and 12007) and the Medical Research Council, UK. O.H. received support from Leukaemia & Lymphoma Research (10033 and 12055) and the North of England Children's Cancer Fund.

Received: January 31, 2014
Revised: June 19, 2014
Accepted: August 12, 2014
Published: September 18, 2014

REFERENCES

Amann, J.M., Nip, J., Strom, D.K., Lutterbach, B., Harada, H., Lenny, N., Downing, J.R., Meyers, S., and Hiebert, S.W. (2001). ETO, a target of t(8;21) in acute leukemia, makes distinct contacts with multiple histone deacetylases and binds mSin3A through its oligomerization domain. Mol. Cell. Biol. *21*, 6470–6483.

Ben-Ami, O., Friedman, D., Leshkowitz, D., Goldenberg, D., Orlovsky, K., Pencovich, N., Lotem, J., Tanay, A., and Groner, Y. (2013). Addiction of t(8;21) and inv(16) acute myeloid leukemia to native RUNX1. Cell Rep *4*, 1131–1143.

Chimge, N.O., and Frenkel, B. (2013). The RUNX family in breast cancer: relationships with estrogen signaling. Oncogene *32*, 2121–2130.

Davidson, E.H. (2010). Emerging properties of animal gene regulatory networks. Nature *468*, 911–920.

DeVilbiss, A.W., Sanalkumar, R., Johnson, K.D., Keles, S., and Bresnick, E.H. (2014). Hematopoietic transcriptional mechanisms: From locus-specific to genome-wide vantage points. Exp. Hematol. *42*, 618–629.

Diffner, E., Beck, D., Gudgin, E., Thoms, J.A., Knezevic, K., Pridans, C., Foster, S., Goode, D., Lim, W.K., Boelen, L., et al. (2013). Activity of a heptad of transcription factors is associated with stem cell programs and clinical outcome in acute myeloid leukemia. Blood *121*, 2289–2300.

Dunne, J., Cullmann, C., Ritter, M., Soria, N.M., Drescher, B., Debernardi, S., Skoulakis, S., Hartmann, O., Krause, M., Krauter, J., et al. (2006). siRNA-mediated AML1/MTG8 depletion affects differentiation and proliferation-associated gene expression in t(8;21)-positive cell lines and primary AML blasts. Oncogene *25*, 6067–6078.

Follows, G.A., Tagoh, H., Lefevre, P., Hodge, D., Morgan, G.J., and Bonifer, C. (2003). Epigenetic consequences of AML1-ETO action at the human c-FMS locus. EMBO J. *22*, 2798–2809.

Gaidzik, V.I., Bullinger, L., Schlenk, R.F., Zimmermann, A.S., Röck, J., Paschka, P., Corbacioglu, A., Krauter, J., Schlegelberger, B., Ganser, A., et al. (2011). RUNX1 mutations in acute myeloid leukemia: results from a comprehensive genetic and clinical analysis from the AML study group. J. Clin. Oncol. *29*, 1364–1372.

Goyama, S., Schibler, J., Cunningham, L., Zhang, Y., Rao, Y., Nishimoto, N., Nakagawa, M., Olsson, A., Wunderlich, M., Link, K.A., et al. (2013). Transcription factor RUNX1 promotes survival of acute myeloid leukemia cells. J. Clin. Invest. *123*, 3876–3888.

Heidenreich, O., Krauter, J., Riehle, H., Hadwiger, P., John, M., Heil, G., Vornlocher, H.P., and Nordheim, A. (2003). AML1/MTG8 oncogene suppression by small interfering RNAs supports myeloid differentiation of t(8;21)-positive leukemic cells. Blood *101*, 3157–3163.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors primecis-regulatory elements required for macrophage and B cell identities. Mol. Cell *38*, 576–589.

Kitabayashi, I., Yokoyama, A., Shimizu, K., and Ohki, M. (1998). Interaction and functional cooperation of the leukemia-associated factors AML1 and p300 in myeloid cell differentiation. EMBO J. *17*, 2994–3004.

Leddin, M., Perrod, C., Hoogenkamp, M., Ghani, S., Assi, S., Heinz, S., Wilson, N.K., Follows, G., Schönheit, J., Vockentanz, L., et al. (2011). Two distinct auto-regulatory loops operate at the PU.1 locus in B cells and myeloid cells. Blood *117*, 2827–2838.

Lefevre, P., Melnik, S., Wilson, N., Riggs, A.D., and Bonifer, C. (2003). Developmentally regulated recruitment of transcription factors and chromatin modification activities to chicken lysozymecis-regulatory elements in vivo. Mol. Cell. Biol. *23*, 4386–4400.

Levanon, D., Goldstein, R.E., Bernstein, Y., Tang, H., Goldenberg, D., Stifani, S., Paroush, Z., and Groner, Y. (1998). Transcriptional repression by AML1 and LEF-1 is mediated by the TLE/Groucho corepressors. Proc. Natl. Acad. Sci. USA *95*, 11590–11595.

Lichtinger, M., Ingram, R., Hannah, R., Müller, D., Clarke, D., Assi, S.A., Lie-A-Ling, M., Noailles, L., Vijayabaskar, M.S., Wu, M., et al. (2012). RUNX1

reshapes the epigenetic landscape at the onset of haematopoiesis. EMBO J. *31*, 4318–4333.

Liu, Y., Cheney, M.D., Gaudet, J.J., Chruszcz, M., Lukasik, S.M., Sugiyama, D., Lary, J., Cole, J., Dauter, Z., Minor, W., et al. (2006). The tetramer structure of the Nervy homology two domain, NHR2, is critical for AML1/ETO's activity. Cancer Cell *9*, 249–260.

Mandoli, A., Singh, A.A., Jansen, P.W., Wierenga, A.T., Riahi, H., Franci, G., Prange, K., Saeed, S., Vellenga, E., Vermeulen, M., et al. (2014). CBFB-MYH11/RUNX1 together with a compendium of hematopoietic regulators, chromatin modifiers and basal transcription factors occupies self-renewal genes in inv(16) acute myeloid leukemia. Leukemia *28*, 770–778.

Martens, J.H., Mandoli, A., Simmer, F., Wierenga, B.J., Saeed, S., Singh, A.A., Altucci, L., Vellenga, E., and Stunnenberg, H.G. (2012). ERG and FLI1 binding sites demarcate targets for aberrant epigenetic regulation by AML1-ETO in acute myeloid leukemia. Blood *120*, 4038–4048.

Martinez, N., Drescher, B., Riehle, H., Cullmann, C., Vornlocher, H.P., Ganser, A., Heil, G., Nordheim, A., Krauter, J., and Heidenreich, O. (2004). The oncogenic fusion protein RUNX1-CBFA2T1 supports proliferation and inhibits senescence in t(8;21)-positive leukaemic cells. BMC Cancer *4*, 44.

Martinez Soria, N., Tussiwand, R., Ziegler, P., Manz, M.G., and Heidenreich, O. (2009). Transient depletion of RUNX1/RUNX1T1 by RNA interference delays tumour formation in vivo. Leukemia *23*, 188–190.

McNeil, S., Zeng, C., Harrington, K.S., Hiebert, S., Lian, J.B., Stein, J.L., van Wijnen, A.J., and Stein, G.S. (1999). The t(8;21) chromosomal translocation in acute myelogenous leukemia modifies intranuclear targeting of the AML1/CBFalpha2 transcription factor. Proc. Natl. Acad. Sci. USA *96*, 14882–14887.

Michaud, J., Wu, F., Osato, M., Cottles, G.M., Yanagida, M., Asou, N., Shigesada, K., Ito, Y., Benson, K.F., Raskind, W.H., et al. (2002). In vitro analyses of known and novel RUNX1/AML1 mutations in dominant familial platelet disorder with predisposition to acute myelogenous leukemia: implications for mechanisms of pathogenesis. Blood *99*, 1364–1372.

Miyoshi, H., Kozu, T., Shimizu, K., Enomoto, K., Maseki, N., Kaneko, Y., Kamada, N., and Ohki, M. (1993). The t(8;21) translocation in acute myeloid leukemia results in production of an AML1-MTG8 fusion transcript. EMBO J. *12*, 2715–2721.

Natoli, G., Ghisletti, S., and Barozzi, I. (2011). The genomic landscapes of inflammation. Genes Dev. *25*, 101–106.

Pabst, T., Mueller, B.U., Harakawa, N., Schoch, C., Haferlach, T., Behre, G., Hiddemann, W., Zhang, D.E., and Tenen, D.G. (2001a). AML1-ETO downregulates the granulocytic differentiation factor C/EBPalpha in t(8;21) myeloid leukemia. Nat. Med. *7*, 444–451.

Pabst, T., Mueller, B.U., Zhang, P., Radomska, H.S., Narravula, S., Schnittger, S., Behre, G., Hiddemann, W., and Tenen, D.G. (2001b). Dominant-negative mutations of CEBPA, encoding CCAAT/enhancer binding protein-alpha (C/EBPalpha), in acute myeloid leukemia. Nat. Genet. *27*, 263–270.

Pimanda, J.E., and Göttgens, B. (2010). Gene regulatory networks governing haematopoietic stem cell development and identity. Int. J. Dev. Biol. *54*, 1201–1211.

Piper, J., Elze, M.C., Cauchy, P., Cockerill, P.N., Bonifer, C., and Ott, S. (2013). Wellington: a novel method for the accurate identification of digital genomic footprints from DNase-seq data. Nucleic Acids Res. *41*, e201.

Preudhomme, C., Sagot, C., Boissel, N., Cayuela, J.M., Tigaud, I., de Botton, S., Thomas, X., Raffoux, E., Lamandin, C., Castaigne, S., et al.; ALFA Group (2002). Favorable prognostic significance of CEBPA mutations in patients with de novo acute myeloid leukemia: a study from the Acute Leukemia French Association (ALFA). Blood *100*, 2717–2723.

Ptasinska, A., Assi, S.A., Mannari, D., James, S.R., Williamson, D., Dunne, J., Hoogenkamp, M., Wu, M., Care, M., McNeill, H., et al. (2012). Depletion of RUNX1/ETO in t(8;21) AML cells leads to genome-wide changes in chromatin structure and transcription factor binding. Leukemia *26*, 1829–1841.

Reed-Inderbitzin, E., Moreno-Miralles, I., Vanden-Eynden, S.K., Xie, J., Lutterbach, B., Durst-Goodwin, K.L., Luce, K.S., Irvin, B.J., Cleary, M.L., Brandt, S.J., and Hiebert, S.W. (2006). RUNX1 associates with histone deacetylases and SUV39H1 to repress transcription. Oncogene *25*, 5777–5786.

Saeed, S., Logie, C., Francoijs, K.J., Frigè, G., Romanenghi, M., Nielsen, F.G., Raats, L., Shahhoseini, M., Huynen, M., Altucci, L., et al. (2012). Chromatin accessibility, p300, and histone acetylation define PML-RARα and AML1-ETO binding sites in acute myeloid leukemia. Blood *120*, 3058–3068.

Scheitz, C.J., and Tumbar, T. (2013). New insights into the role of Runx1 in epithelial stem cell biology and pathology. J. Cell. Biochem. *114*, 985–993.

Scheitz, C.J., Lee, T.S., McDermitt, D.J., and Tumbar, T. (2012). Defining a tissue stem cell-driven Runx1/Stat3 signalling axis in epithelial cancer. EMBO J. *31*, 4124–4139.

Schwieger, M., Schüler, A., Forster, M., Engelmann, A., Arnold, M.A., Delwel, R., Valk, P.J., Löhler, J., Slany, R.K., Olson, E.N., and Stocking, C. (2009). Homing and invasiveness of MLL/ENL leukemic cells is regulated by MEF2C. Blood *114*, 2476–2488.

Scott, E.W., Simon, M.C., Anastasi, J., and Singh, H. (1994). Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages. Science *265*, 1573–1577.

Snaddon, J., Smith, M.L., Neat, M., Cambal-Parrales, M., Dixon-McIver, A., Arch, R., Amess, J.A., Rohatiner, A.Z., Lister, T.A., and Fitzgibbon, J. (2003). Mutations of CEBPA in acute myeloid leukemia FAB types M1 and M2. Genes Chromosomes Cancer *37*, 72–78.

Staber, P.B., Zhang, P., Ye, M., Welner, R.S., Nombela-Arrieta, C., Bach, C., Kerenyi, M., Bartholdy, B.A., Zhang, H., Alberich-Jordà, M., et al. (2013). Sustained PU.1 levels balance cell-cycle regulators to prevent exhaustion of adult hematopoietic stem cells. Mol. Cell *49*, 934–946.

Stender, J.D., Kim, K., Charn, T.H., Komm, B., Chang, K.C., Kraus, W.L., Benner, C., Glass, C.K., and Katzenellenbogen, B.S. (2010). Genome-wide analysis of estrogen receptor alpha DNA binding and tethering mechanisms identifies Runx1 as a novel tethering factor in receptor-mediated transcriptional activation. Mol. Cell. Biol. *30*, 3943–3955.

Sun, X.J., Wang, Z., Wang, L., Jiang, Y., Kost, N., Soong, T.D., Chen, W.Y., Tang, Z., Nakadai, T., Elemento, O., et al. (2013). A stable transcription factor complex nucleated by oligomeric AML1-ETO controls leukaemogenesis. Nature *500*, 93–97.

Taniuchi, I., Osato, M., Egawa, T., Sunshine, M.J., Bae, S.C., Komori, T., Ito, Y., and Littman, D.R. (2002). Differential requirements for Runx proteins in CD4 repression and epigenetic silencing during T lymphocyte development. Cell *111*, 621–633.

Taoudi, S., Bee, T., Hilton, A., Knezevic, K., Scott, J., Willson, T.A., Collin, C., Thomas, T., Voss, A.K., Kile, B.T., et al. (2011). ERG dependence distinguishes developmental control of hematopoietic stem cell maintenance from hematopoietic specification. Genes Dev. *25*, 251–262.

Tijssen, M.R., Cvejic, A., Joshi, A., Hannah, R.L., Ferreira, R., Forrai, A., Bellissimo, D.C., Oram, S.H., Smethurst, P.A., Wilson, N.K., et al. (2011). Genome-wide analysis of simultaneous GATA1/2, RUNX1, FLI1, and SCL binding in megakaryocytes identifies hematopoietic regulators. Dev. Cell *20*, 597–609.

Tsuzuki, S., and Seto, M. (2012). Expansion of functionally defined mouse hematopoietic stem and progenitor cells by a short isoform of RUNX1/AML1. Blood *119*, 727–735.

Valk, P.J., Verhaak, R.G., Beijen, M.A., Erpelinck, C.A., Barjesteh van Waalwijk van Doorn-Khosrovani, S., Boer, J.M., Beverloo, H.B., Moorhouse, M.J., van der Spek, P.J., Löwenberg, B., and Delwel, R. (2004). Prognostically useful gene-expression profiles in acute myeloid leukemia. N. Engl. J. Med. *350*, 1617–1628.

van Riel, B., Pakozdi, T., Brouwer, R., Monteiro, R., Tuladhar, K., Franke, V., Bryne, J.C., Jorna, R., Rijkers, E.J., van Ijcken, W., et al. (2012). A novel complex, RUNX1-MYEF2, represses hematopoietic genes in erythroid cells. Mol. Cell. Biol. *32*, 3814–3822.

Vangala, R.K., Heiss-Neumann, M.S., Rangatia, J.S., Singh, S.M., Schoch, C., Tenen, D.G., Hiddemann, W., and Behre, G. (2003). The myeloid master regulator transcription factor PU.1 is inactivated by AML1-ETO in t(8;21) myeloid leukemia. Blood *101*, 270–277.

Wang, J., Hoshino, T., Redner, R.L., Kajigaya, S., and Liu, J.M. (1998). ETO, fusion partner in t(8;21) acute myeloid leukemia, represses transcription by interaction with the human N-CoR/mSin3/HDAC1 complex. Proc. Natl. Acad. Sci. USA *95*, 10860–10865.

Wang, L., Gural, A., Sun, X.J., Zhao, X., Perna, F., Huang, G., Hatlen, M.A., Vu, L., Liu, F., Xu, H., et al. (2011). The leukemogenicity of AML1-ETO is dependent on site-specific lysine acetylation. Science *333*, 765–769.

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

Zhang, D.E., Zhang, P., Wang, N.D., Hetherington, C.J., Darlington, G.J., and Tenen, D.G. (1997). Absence of granulocyte colony-stimulating factor signaling and neutrophil development in CCAAT enhancer binding protein alpha-deficient mice. Proc. Natl. Acad. Sci. USA *94*, 569–574.

Zhang, P., Iwasaki-Arai, J., Iwasaki, H., Fenyus, M.L., Dayaram, T., Owens, B.M., Shigematsu, H., Levantini, E., Huettner, C.S., Lekstrom-Himes, J.A., et al. (2004). Enhancement of hematopoietic stem cell repopulating capacity and self-renewal in the absence of the transcription factor C/EBP alpha. Immunity *21*, 853–863.

Zhang, H., Alberich-Jorda, M., Amabile, G., Yang, H., Staber, P.B., Di Ruscio, A., Welner, R.S., Ebralidze, A., Zhang, J., Levantini, E., et al. (2013). Sox4 is a key oncogenic target in C/EBPα mutant acute myeloid leukemia. Cancer Cell *11*, 575–588.

Cell Reports, Volume *8*

Supplemental Information

Identification of a Dynamic Core Transcriptional Network in t(8;21) AML that Regulates Differentiation Block and Self-Renewal

Anetta Ptasinska, Salam A. Assi, Natalia Martinez-Soria, Maria Rosaria Imperato, Jason Piper, Pierre Cauchy, Anna Pickin, Sally R. James, Maarten Hoogenkamp, Dan Williamson, Mengchu Wu, Daniel G. Tenen, Sascha Ott, David R. Westhead, Peter N. Cockerill, Olaf Heidenreich, and Constanze Bonifer

SUPPLEMENTAL MATERIAL

1. Supplemental tables, figures and figure legends

Dataset	No of actual peaks	No. of high confidence peaks
DHS control	29379	NA
DHS siMM	31861	NA
DHS siRE	32814	NA
DHS CD34+	35773	NA
RUNX1/ETO control	7679	4821
RUNX1/ETO siMM	6816	4529
RUNX1/ETO siRE	563	103
RUNX1 siMM	8843	7191
RUNX1 siRE	11798	8665
CEBPa siMM	2389	1569
CEBPa siRE	7532	6077
LMO2 siMM	11470	8498
LMO2 siRE	6015	4730
PU.1 siMM	15229	9850
PU.1 siRE	13725	9789
RUNX1 CD34+	10909	N.A.

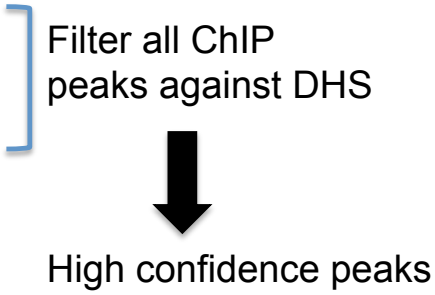


Table S1: The number of peaks of DHS, RUNX1/ETO, RUNX1, C/EBPα, LMO2 and PU.1 as determined by DNaseI-seq, ChIP-seq and as high confidence peaks associated within open chromatin regions

Table S2: Occupancy patterns correlated with genes that are up/down regulated (separate Excel file)

Table S 3: Occupancy patterns correlated with up/down regulated genes after RUNX1/ETO knock down (separate Excel file)

Table S4. Primers

(A) Primers used for ChIP-qPCR

Binding sites	Forward	Reverse
CSF1R (P)	AGAAGAGGTCAGCCCAAGGA	AGGGATCGGGACACTGGAC
CSF1R FIRE	GCCTGACGCCAACAATGTG	GGCAAAGGAGGGAAGTGAGAG
PU.1 (P)	CTGCCGCTGGGAGATAG	CGGCCAGAGACTTCCTGTA
PU.1 14 3H	AACAGGAAGCGCCCAGTCA	TGTGCGGTGCCTGTGGTAAT
MIR223	TTGGAAAGTTAGTGTCTGTTGAAGG	TGTTGTGAAAGGGTCTGCTACTG
TERT	AAATGGTCTCAGCCTCACCGTC	TTCCCTCCAATCACACCTTGC
C/EBPA	GCCAGTTTATGGAGGTGTGAGC	ATAGGTGGTGATGATGGTTGCC
LAT2	AAACCCAGAACAACCCAGGC	ATGAGGAAGGATGTGTGTGCGG
CTSG	TCAGTTGCTGCTGTGCTTC	TTCTCAATCCCCTGTCCCCAC
NFE2	*AATAGCGAGGCCCCTCTTAG*	ACCCAAACTGGAACACAAGG
JUN	TTGGGGTTACTGTAGCCATAAG	CGTGAAGTGACGGACTGTTC
CD34	TGTGGTTAGCCAAACTCCAGGTC	TGAGGAATGAAGCAGCAGTGG
TBP	CTGGCGGAAGTGACATTATCAA	GCCAGCGGAAGCGAAGTTA
CHR 18	ACTCCCCTTTCATGCTTCTG	AGGTCCCAGGACATATCCATT
DOK4	AGTAATGTTCCGTGCCCTTG	AGTAATGTTCCGTGCCCTTG
CDK6	AACAGAGCCCACAACATTCC	CCGTGAAAAATTGCATCCTT
KLF2	CTCCCACCGGGTCTACACTA	AATGCCGCAGACAGTACAAA
IGFBP7	GTCAAGCACTAAAAGGACAAACCG	TGAATGCCACTGGGAG
TMC3	*CACTCCCTGAGAAGGTCTGC*	*TACAGCTCTCAAAGGCAGCA*
GATA2	CACCGCACAGCAGTGATAGA	GCAGCCTGCTTTACCACATC
RUNX1	*GATACCGGAAAGGCCTGTGA*	*AGTGCCTGGAAATGAACGT*

(B) Primers used for qPCR

cDNA	**Forward**	**Reverse**
GAPDH	CCTGGCCAAGGTCATCCAT	AGGGGCCATCCACAGTCTT
RUNX1/ETO	TCAAAATCACAGTGGATGGGC	CAGCCTAGATTGCGTCTTCACA
RUNX1	CCCTCAGCCTCAGAGTCAGAT	AGGCAATGGATCCCAGGTAT
C/EBPα	GAGGGACCGGAGTTATGACA	AGACGCGCACATTCACATT
LMO2	TTCGGTTGAGAATGGAAACC	CTCCCCTCAAAATGAAGGTG
PU.1	TCTTGGCCACCAGGTCTCCTA	CGCCCTCCTCCTCATCTGA
LYL1	CATCTTCCCTAGCAGCCGGTTG	GTTGGTGAACACGCGCCG
LDB1	TGTTCCTCAAAGTCATTCAAGC	CCCACATCCCTATCCAGCAT
TAL1	AGCCCCCAGTCATCGAACT	CGGCCCTTTAAGTCTCTCG
HEB	TCATAGCTTGGGGATGAAGG	TTCCGTCAAATCCATCAACA
P300	GGGACTAACCAATGGTGGTG	ATGGCAGGCTGATTTACTGG
HDAC2	AATTCAAGGATGGCAAGCAC	GAGCTGTGAAGTTAAACCGACA
RNASE2	CCCCTGAACCCCAGAACAA	ACCATGTTTCCCAGTCTCCG
NKG7	CTGATTGCTTTGAGCACCGA	CCTGATATGATGTCCCCATGC
MS4A3	CCAAGCCATAAACAACCCCA	TTCTGGTCCCGTCTCACTGC

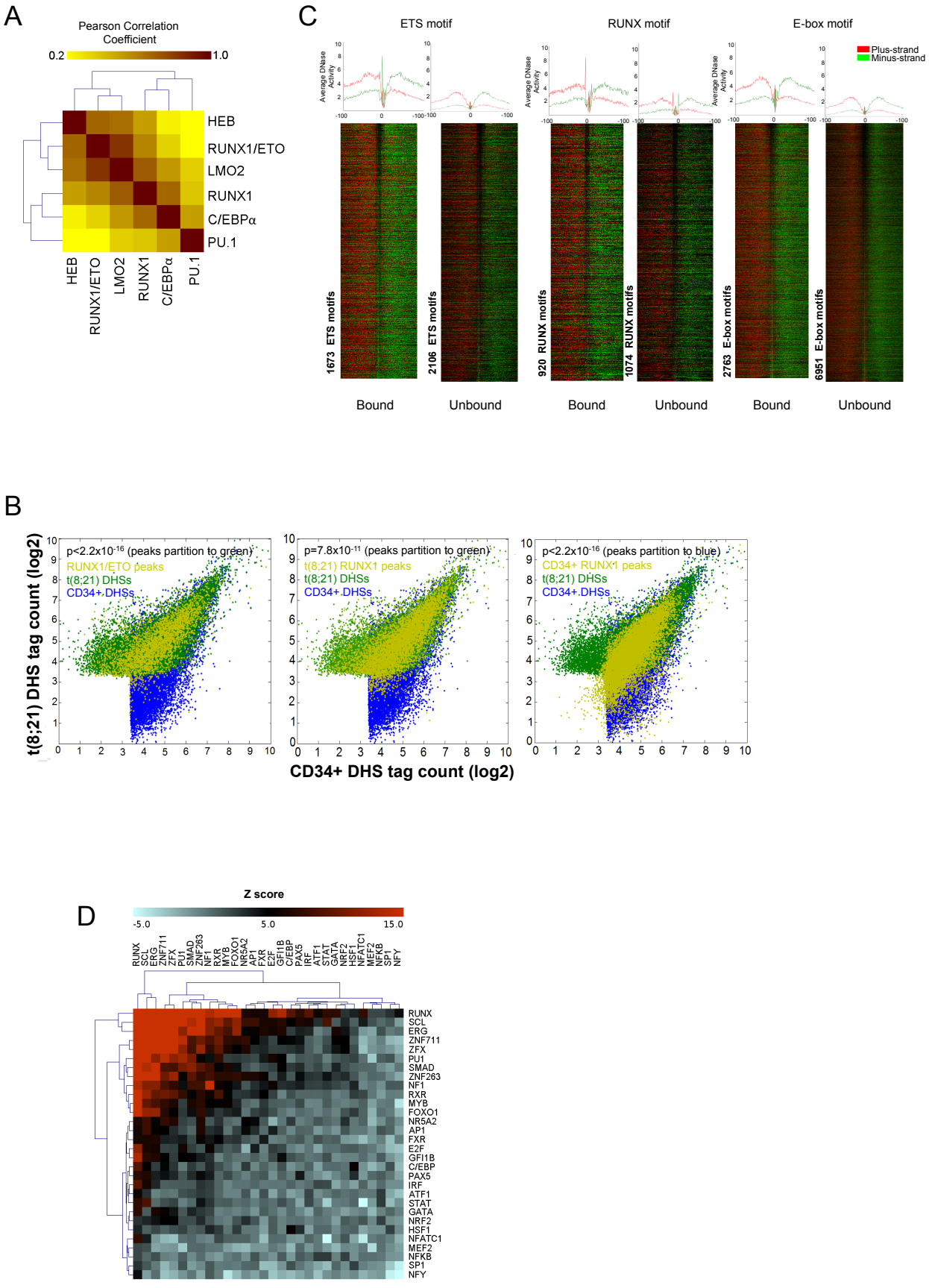


Figure S1 (related to Figure 1)

Similar transcription factor occupancy patterns in Kasumi-1 cells and primary t(8;21) cells. (A) Heatmap depicting hierarchical clustering of DNA-Sequences within ChIP-Seq peaks for the indicated transcription factors indicating similar or diverse binding patterns in Kasumi-1 cells treated for 48 h with mismatch control siRNA (siMM). A correlation matrix was generated and Pearson correlation coefficients are displayed after hierarchical clustering as a heatmap. Colors in the heatmap indicate the strength of association between each pair of transcription factors. **(B)** Scatter plot depicting DHS sequences from t(8;21) patients (union of two data-sets from two patients, green dots) on one axis and from normal CD34+ cells on the other axis (blue dots), both sorted by tag count (eliminating low tag counts). We then projected RUNX1/ETO bound sequences (left) and RUNX1 bound sequences (center) from Kasumi-1 cells and RUNX1 bound sequences from normal CD34+ cells (right) onto the plot, demonstrating that RUNX1 and RUNX1/ETO bound sequences in the Kasumi-1 cell line project on t(8;21)-specific DHS from patients. Statistical significance was tested with a Kolmogorov-Smirnov test.**(C)** ETS, RUNX and E-box motifs in sequences covering the union of LMO2, RUNX1, RUNX1/ETO and HEB peaks (13,584 peaks) identified in Kasumi-1 cells were subjected to footprinting analysis using DNAseI-seq data from t(8;21) patient J171 and the Wellington software (FDR=0.01). The DNAseI cuts, with the protected region in the middle and average profiles(upper panels) were plotted +/- 100 bp around ETS (left), RUNX (middle) and E-box (right) motifs and subdivided into motifs that were bound in Kasumi-1 cells (bound) and motifs that were not (unbound). Note the differences in average cutting intensity around bound and unbound regions depicted on top of the heatmap which illustrate how DNAse-seq footprinting is able to highlight preferentially bound motifs. Red lines in both average profiles and heatmap depict the positive DNA strand; green lines depict the negative strand. **(D)** Heatmap showing hierarchical clustering of transcription factor binding motif co-occurrences by Z-score within RUNX1/ETO ChIP peaks indicating possible cooperation between specific transcription factors binding to juxtapose cis-elements. The search for co-localizing motifs for the indicated transcription factor families was done within +/- 200bp from RUNX1/ETO peak centre. The distances between each motif pairs were calculated and motifs were counted as co-occuring when the first motif was within 50bp distance from the second motif. Z-scores were calculated from the mean and standard deviation of motif frequencies observed in random sets; a background data set was generated by analysing sequences equal to the number of RUNX1/ETO peaks from all Kasumi-1 DNAseI-seq data (i.e all regulatory sequences without enrichment). Red and blue colours indicate over- and under-represented motifs, respectively.

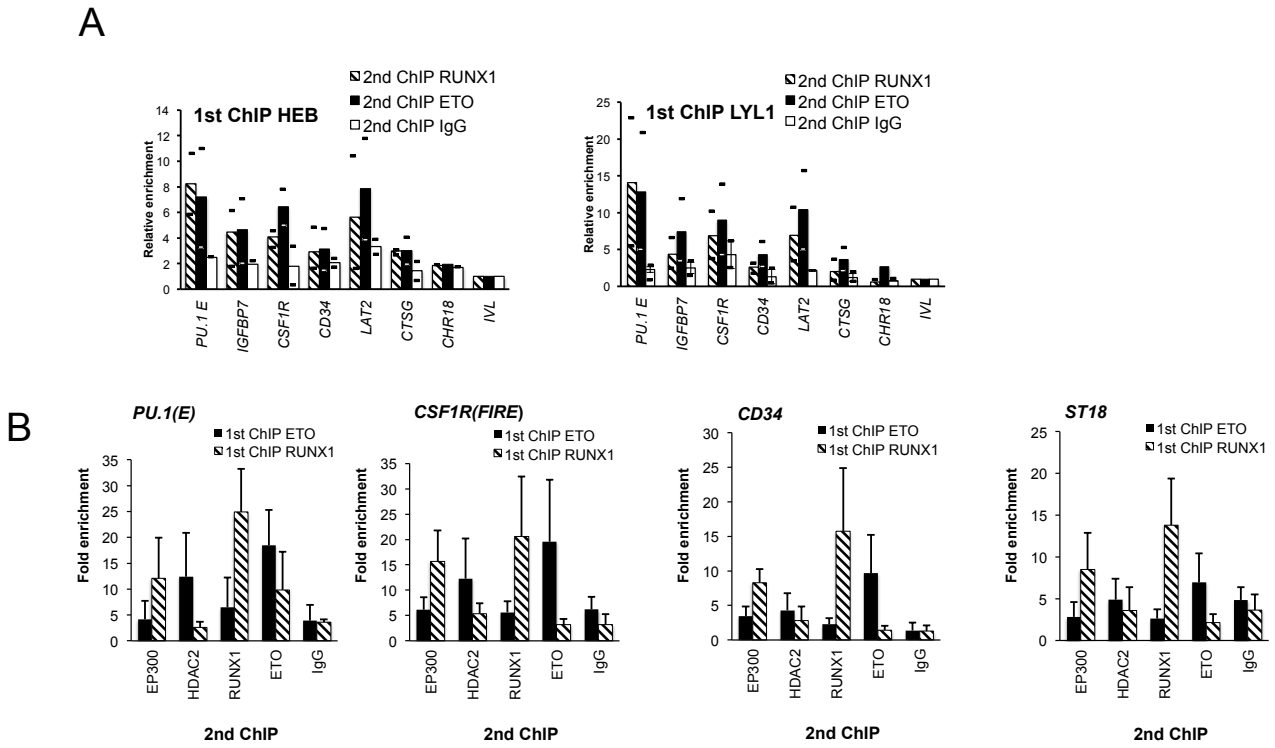


Figure S2 (related to Figure 2)

RUNX1 and RUNX1/ETO complexes differentially interact with co-activator and co-repressor complexes and binding to the same sites is mutually exclusive. Multiple RUNX1/ETO binding sites as well as control sequences (IVL, Chr18) were selected and validated for factor binding by a first round of ChIP followed by a second round with a different antibody or with just beads as indicated. **(A)** Both HEB and LYL1 are recruited to RUNX1 and RUNX1/ETO binding sites. Results represent the mean values of two biological duplicates yielding different absolute values and respective values are shown as upper bars and lower bars. **(B)** Re-ChIP experiments showing a preference for p300 recruitment to the RUNX1 complex and HDAC2 to the RUNX1/ETO complex to the same binding sites. Four binding sites (PU.1 (E), CSF1R (FIRE), CD34 and ST18) were selected and analyzed by re-ChIP as indicated. Note that CD34 and ST18 are active genes down-regulated by RUNX1/ETO knockdown.

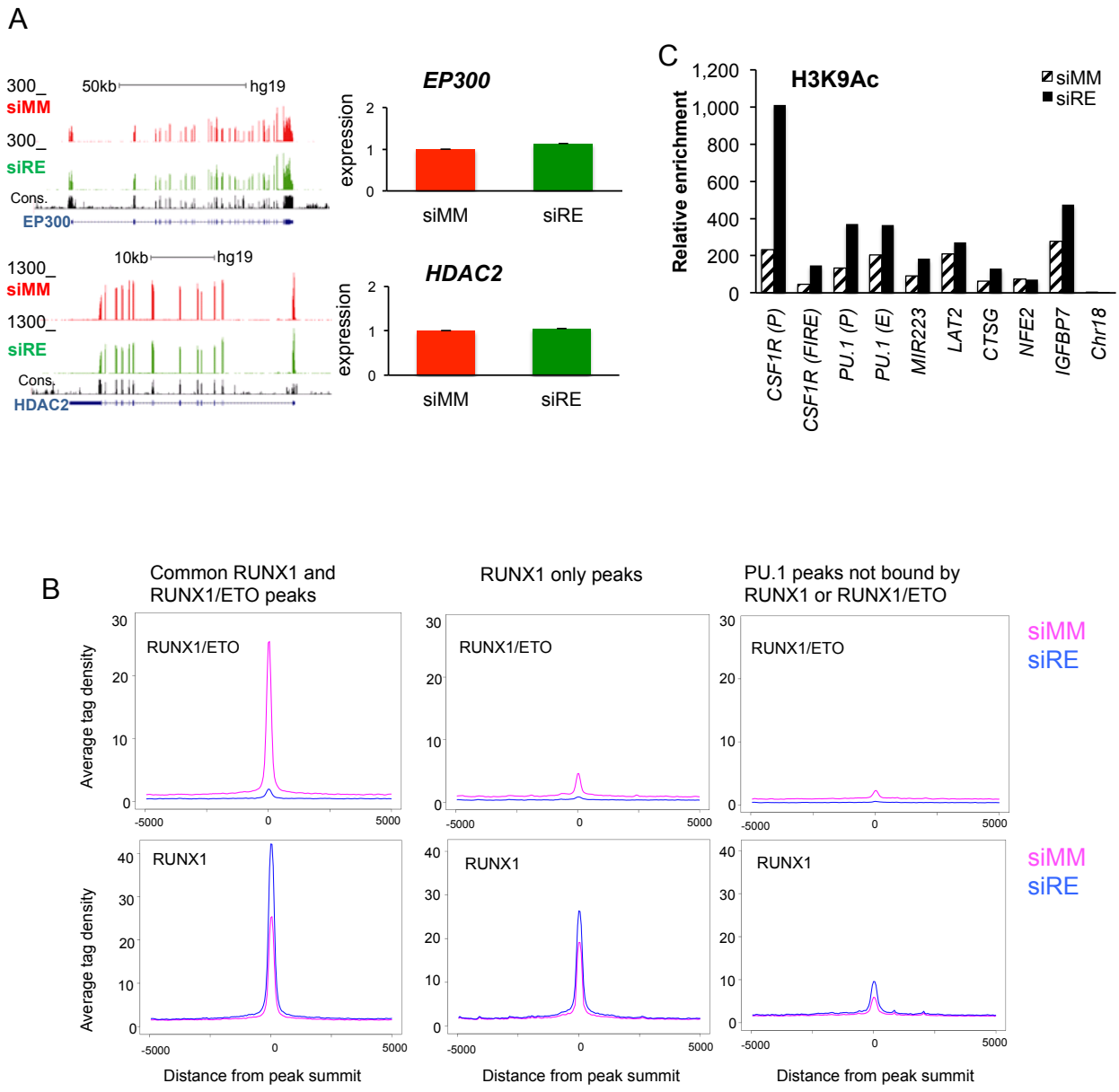


Figure S3 (related to Figure 3):

Knockdown of RUNX1/ETO results in increased histone acetylation and p300 binding at RUNX1/ETO target sites

(A) UCSC genome browser screenshots showing RNA-seq read coverage and real-time PCR analysis of mRNA expression of p300 and HDAC2 genes in Kasumi-1 cells treated for 48 h with mismatch control siRNA (siMM) and with RUNX1/ETO siRNA (siRE). **(B)** RUNX1/ETO knockdown leads to an increase in RUNX1 binding at RUNX1/ETO binding sites. Accumulated RUNX1/ETO (upper panel) and RUNX1 (lower panel) binding levels before and after RUNX1/ETO knockdown at distinct classes of binding sites as indicated. **(C)** H3K9 acetylation levels at selected RUNX1/ETO binding sites before and after RUNX1/ETO knockdown. For global acetylation data see (Ptasinska et al., 2012).

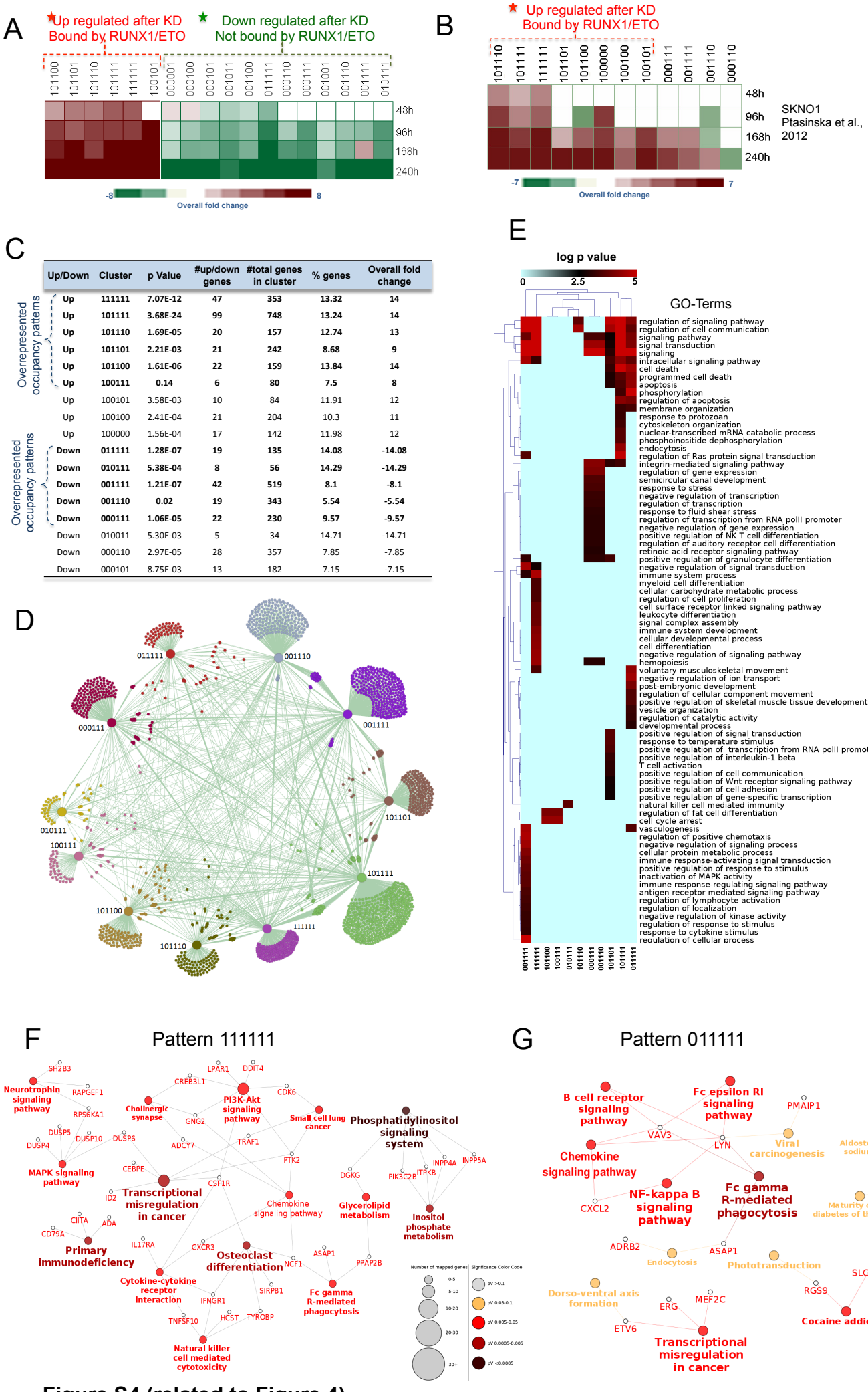


Figure S4 (related to Figure 4)

Combinatorial transcription factor binding patterns identify RUNX1/ETO responsive genes with distinct biological functions. (A) Heatmap showing overall fold change in gene expression in our previous study of Kasumi-1 cells during a 10-day time course after RUNX1/ETO knockdown (Ptasinska et al., 2012). Patterns that include RUNX1/ETO are enriched in up-regulated genes and those that do not bind RUNX1/ETO are enriched in down-regulated genes.**(B)** As in **(A)** but using data from our previous study of SKNO-1 cells (Ptasinska et al., 2012). **(C)** Table showing the number of up- or down-regulated genes related to the patterns (clusters) in Figure 4A whereby 1 indicates that a factor is bound and 0 indicates that it is not bound with the order of factors from the left to the right being described in Figure 4A. **(D)** A regulatory network model for genes regulated by eleven occupancy patterns for the transcription factors HEB, RUNX1/ETO, CEBPα, LMO2, PU.1 and RUNX1 before knockdown demonstrating that most genes are regulated by one specific pattern. Each small circle represents a single gene and each edge (collection of circles) shows genes regulated by the same factor combination. **(E)** Differential binding patterns are indicative of genes with different biological function. The heatmap shows a clustering of GO terms after gene ontology analysis of genes associated with the indicated binding patterns. The color reflects the enrichment significance of the terms. This analysis demonstrates, for example, a highly significant overrepresentation of myeloid differentiation, hematopoiesis and signal transduction for genes associated with the regions bound by RUNX1/ETO and all other transcription factors, whereas genes controlling apoptosis and negatively regulate signal transduction lack binding of RUNX1/ETO. Significant overrepresentation of GO terms is indicated in red, the actual GO terms are listed at the right side of the panel; patterns are depicted below the heat-map. **(F)** KEGG pathway analysis of the patterns 111111: (RUNX1/ETO|CEBPα|HEB|LMO2|PU.1|RUNX1) (upper panel) and 011111 (lower panel): (CEBPα|HEB|LMO2|PU.1|RUNX1) indicating different genes within different biological pathways being associated with differential transcription factor binding. Significantly grouped KEGG pathway network terms as defined by kappa statistics were implemented by ClueGO to link the terms in the network. Sizes of circles indicate the number of genes; the colour depicts significance, as indicated on the right. The right-sided enrichment (depletion) test based on the hyper-geometric distribution was used for terms and groups. Groups were created by iterative merging of initially defined groups based on the kappa score threshold. The colour reflects the enrichment significance of the terms. The sizes of the nodes represent the number of the genes related to the term. The network was automatically laid out using the layout algorithm supported by Cytoscape.

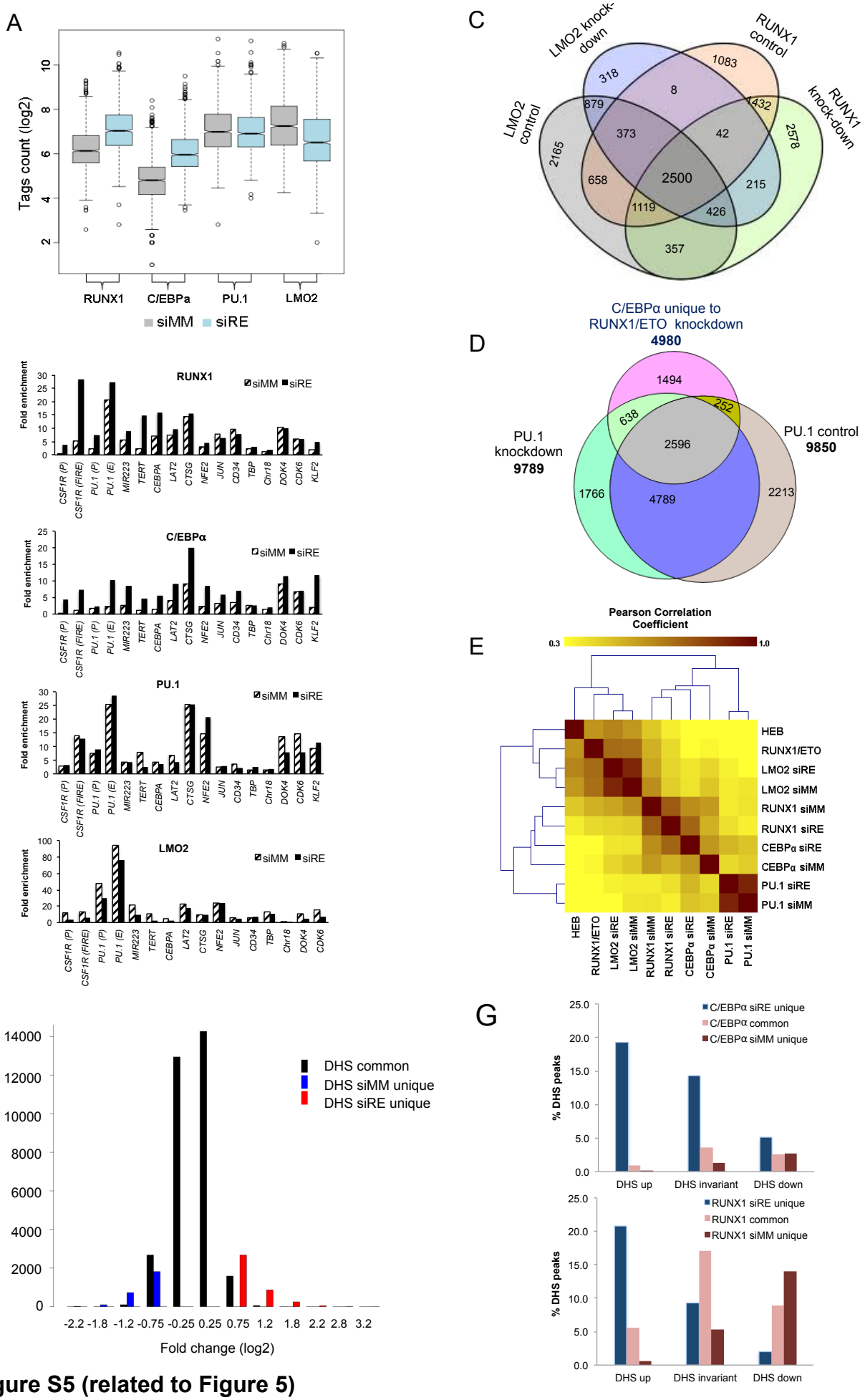


Figure S5 (related to Figure 5)

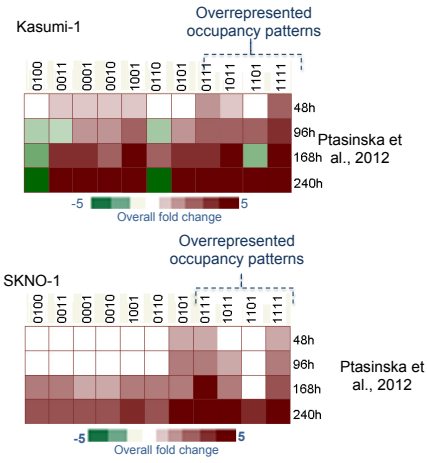
The knockdown of RUNX1/ETO leads to a reorganization of transcription factor assemblies mostly within pre-existing open chromatin regions. (A) Knockdown of RUNX1/ETO leads to a global redistribution of RUNX1, C/EBPα, PU.1 and LMO2 binding as shown by boxplots for the tag counts (+/- 200 bps) around peak centers before and after RUNX1/ETO knockdown. **(B)** Manual validation of ChIP-sequencing data for selected loci. RUNX1, C/EBPα, PU.1 and LMO2 binding was examined at the indicated loci by ChIP-qPCR. Values represent fold enrichment over non-specific binding at the negative control region (Involucrin (IVL)). **(C)** Four-way Venn diagram showing the overlap between RUNX1 and LMO2 peaks in Kasumi-1 before and after RUNX1/ETO knockdown, indicating that the loss of LMO2 binding after RUNX1/ETO knockdown occurs outside of RUNX1 binding sites. **(D)** Three-way Venn diagram showing the overlap between C/EBPα peaks unique to RUNX1/ETO knockdown cells and PU.1 peaks before and after RUNX1/ETO depletion, indicating that more than half of new C/EBPα binding occurs at pre-existing PU.1 binding sites. **(E)** Heatmap depicting hierarchical clustering of transcriptional factor bound sequences based on similar binding patterns of different ChIP-seq data, in Kasumi-1 cells before and after RUNX1/ETO knockdown demonstrating that a large number of PU.1 binding sites are distinct from those of C/EBPα and RUNX1. Pearson correlation coefficients are displayed after hierarchical clustering as a heatmap. Heatmap colours indicate the strength of association between each pair of transcription factor bound sequences. **(F)** Number of unique and common DNaseI peaks identified in Kasumi-1 cells with (siRE) and without (siMM) and RUNX1/ETO. **(G)** Percentage of DNAseI and C/EBPα or RUNX1 peaks in dependence on RUNX1/ETO knockdown. Panels show the distribution of all DHS peaks amongst the three groups. siRE unique, peaks found only in RUNX1/ETO knockdown cells; common, peaks found in both knockdown and control cells; siMM unique, peaks found only in control cells. This shows that DNaseI sensitivity follows transcription factor binding.

A

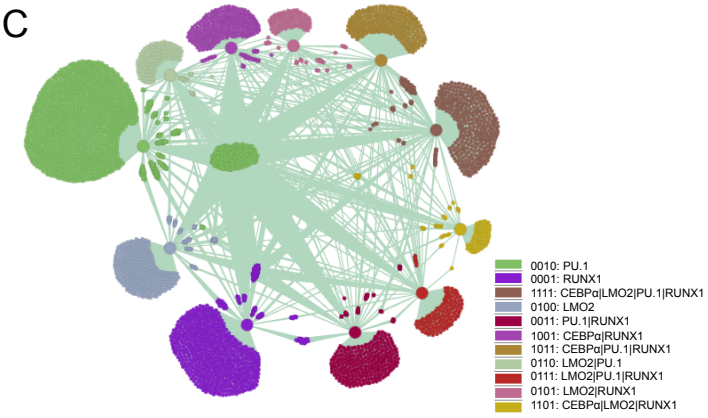
Overrepresented occupancy patterns

Up/Down	Cluster	p Value	#up/down genes	#total genes in cluster	% genes	Overall fold change
Up	1111	1.98E-26	129	1106	11.67	12
Up	0111	2.30E-11	49	391	12.54	13
Up	1011	6.74E-11	73	751	9.73	10
Up	1101	1.33E-04	16	127	12.6	13
Down	0100	4.57E-05	39	591	6.6	-6.6
Up	0101	8.38E-05	25	247	10.13	11
Up	0110	8.99E-04	30	371	8.09	9
Up	1001	3.01E-03	40	581	6.89	7
Up	0010	3.56E-03	118	2142	5.51	6
Up	0001	4.07E-03	73	1229	5.94	6
Up	0011	0.01	42	664	6.33	7

B

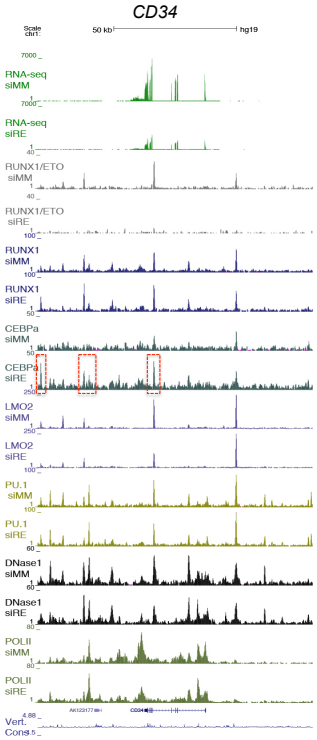


C

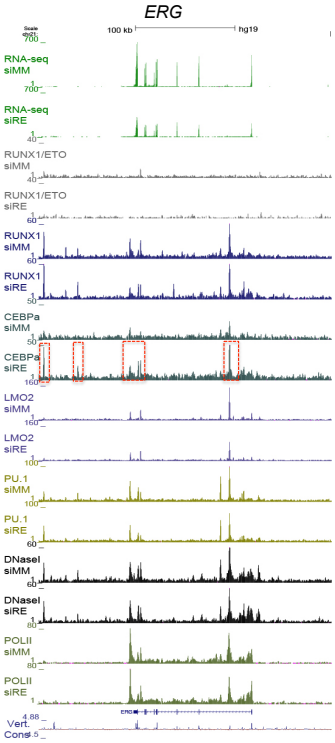


D

E



F



G

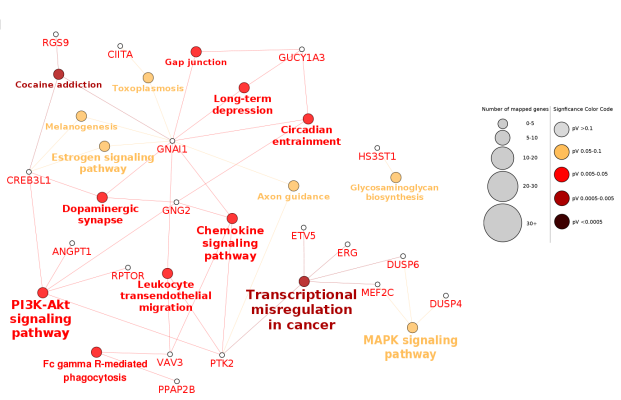


Figure S6 (related to Figure 6)

Knockdown of RUNX1/ETO drives the formation of a new transcriptional network dominated by the binding of C/EBPα. (A) Table showing the number of up- or down-regulated genes related to the patterns (clusters) in Figure 6D. **(B)** Heatmaps showing overall fold change in gene expression from our previous study of Kasumi-1 (top) and SKNO-1 (bottom) cells during a time course after RUNX1/ETO knockdown (Ptasinska et al., 2012). **(C**) A regulatory network model for genes regulated by the eleven occupancy patterns shown in B. **(D)** Clustering of GO terms after gene ontology analysis for genes that are up-regulated (left heatmap) and down-regulated genes (right heatmap) for the main dominant occupancy patterns (1111, 1011, and 0111). For other explanations see Figure S4E. **(E, F)** UCSC genome browser screenshot showing the binding patterns of transcription factors RUNX1/ETO, RUNX1, C/EBPα, LMO2, PU.1 as well as those of DHS, H3K9Ac and RNA-Pol II at the *CD34* (**E)** and *ERG* **(F)** loci in Kasumi-1 cells. Both genes are down-regulated after RUNX1/ETO depletion. Regions with changing C/EBPα binding sites are highlighted by red rectangles. **(G)** KEGG pathway analysis of the down-regulated genes that are bound by CEBPα after R/E KD. Significantly grouped KEGG pathway network terms using kappa statistics were implemented by ClueGO to link the terms in the network. The right-sided enrichment (depletion) test based on the hyper-geometric distribution was used for terms and groups. Groups were created by iterative merging of initially defined groups based on the kappa score threshold. The color reflects the enrichment significance of the terms. The sizes of the nodes represent the number of the genes related to the term. The network is automatically laid out using the layout algorithm supported by Cytoscape. For other explanations see Figure S4F.

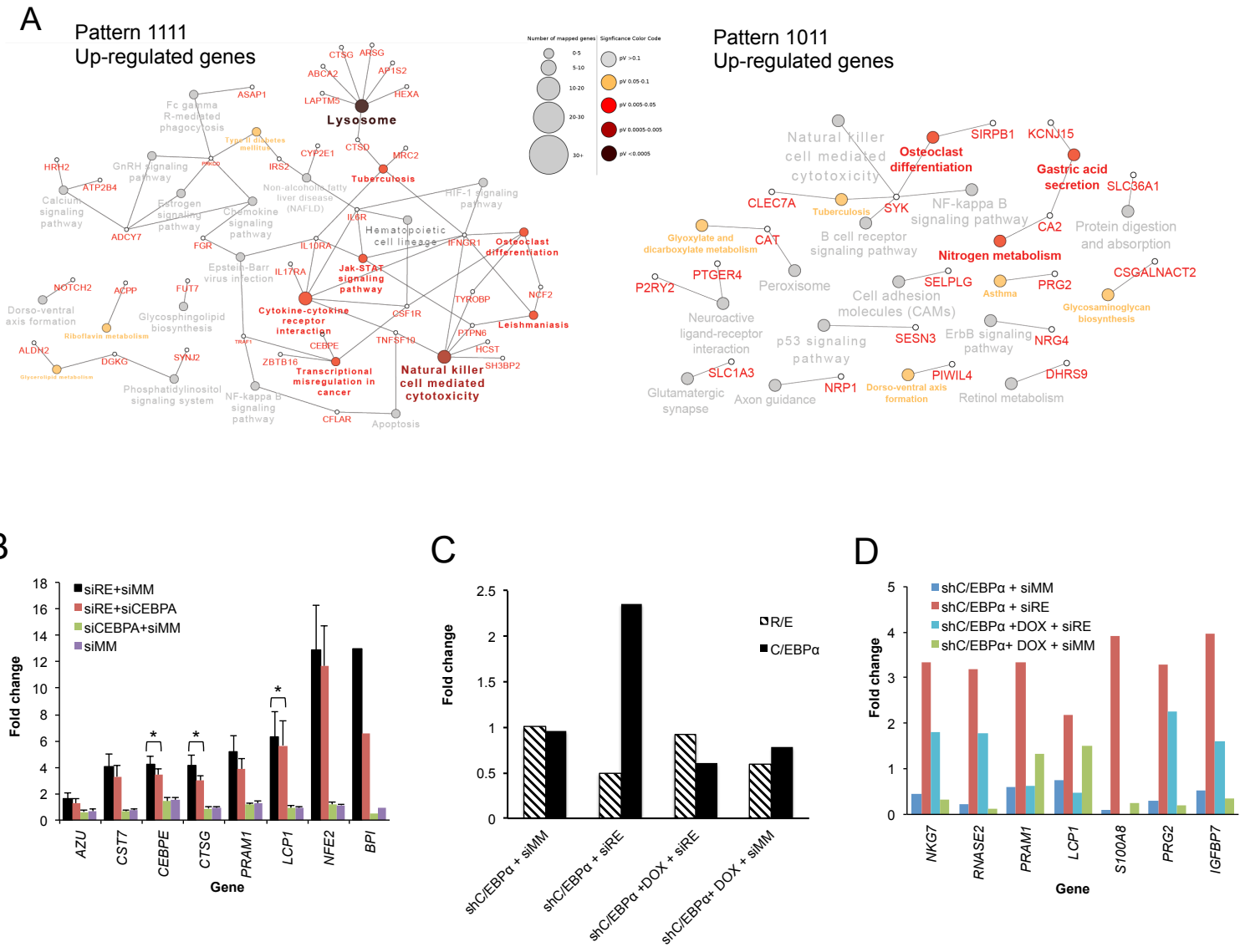


Figure S7 (related to Figure 7)

C/EBPα is required for the up-regulation of specific RUNX1/ETO target genes. (A) KEGG pathway analysis of the up-regulated genes after RUNX1/ETO knockdown that are bound by C/EBPα; pattern 1111: (CEBPα|LMO2|PU.1|RUNX1) and 0111 (CEBPα|PU.1|RUNX1). (**B**): mRNA levels of up-regulated target genes of RUNX/ETO and C/EBPα 72 h after electroporation of Kasumi-1 cells with the indicated siRNAs. Columns show the mean of 5 independent experiments with the exception of *BPI* with two independent experiments; error bars indicate SEM. *, p<0.05 by paired Student's t test **(C)** *RUNX1/ETO* and *C/EBPα* and **(D)** mRNA levels of several target genes in SKNO-1 cells carrying a doxycycline-inducible C/EBPα shRNA, 72 hours after electroporation/induction with the indicated si/shRNAs. siRE, RUNX1/ETO siRNA; shCEBPA, C/EBPα shRNA; siMM, mismatch control siRNA. Data from one experiment measured in duplicate are shown.

2. Supplementary Methods

Human primary cells and cell lines

All patient material was obtained with the required ethical approval from the NHS Research Ethics Committees (Leeds Teaching Hospitals NHS Trust and Newcastle upon Tyne Hospitals NHS Foundation Trust). Initial analysis of AML patient samples was carried out by the Haematological Malignancy Diagnostic Service (St James's Hospital, Leeds), where cytogenetic abnormalities and sample immunophenotype were determined at the time of disease diagnosis. Several sets of 6–colour FACS staining were performed on a FACSCanto or FACSCanto II flow cytometer using antibodies obtained from BD Bioscience (Oxford, UK) as follows: Anti-CD34 - APC / PerCP:Cy5.5 (clone 8G12), Anti-CD117 – PE / PE:Cy7 (clone 10402), Anti-CD45 – PerCP:Cy5.5 / APC:Cy7 (Clone 2D1), Anti-CD15 – FITC (clone MMA), Anti-CD13 – PE (clone L138), Anti-HLADR – APC:Cy7 (Clone L243 (G46-6)), Anti-CD33 – PE (clone P37.6), Anti-CD7 – FITC (clone M-T701), Anti-CD19 – PerCP:Cy5.5 (clone SJ25C1), Anti-CD56 – APC (clone NCAM16.2), Anti-CD14 – FITC (clone MφP9), Anti-CD38-PE:Cy7 (clone HB7) and Anti-CD64 – PE (clone MD22). The result of flow cytometry analysis is shown in Supplementary Table 2. The presence of the t(8;21) translocation and expression of mRNA encoding the RUNX1/ETO fusion protein were determined using RT-PCR, with the primers 5'-TCAAAATCACAGTGGATGGGC and 5'-CAGCCTAGATTGCGTCTTCACA. Mononuclear cells freshly obtained from patients #1 and #2 and from the control patient were prepared by differential centrifugation using Lymphoprep (Axis-Shield UK, Cambridgeshire, UK), and CD34+ blast cells were then isolated using MACS Micro Beads (MiltenyiBiotec GmbH **)** staining and separation on magnetic columns according to the manufacturer's guidelines. Purity of CD34+ cells was >85%. The Kasumi-1 cell line was obtained from the DSMZ cell line repository (http://www.dsmz.de/) and was cultured in RPMI1640 containing 10% fetal calf serum (FCS). SKNO-1 cells were maintained in RPMI1640 supplemented with 20% FCS and 7 ng/ml granulocyte-macrophage colony-stimulating factor.

Real-time RT-PCR

RNA was isolated 2 days after siRNA transfection using RNeasy extraction kit (Qiagen, Hilden, Germany). Reverse transcription was performed by using 1–2 μg of RNA, oligo(dT) primer and Moloney Murine Leukemia Virus reverse transcriptase (Promega, Madison, WI, USA) according to the manufacturer's protocol. Real time PCRs were performed using Sybr-Green Mix (Applied Biosystems, Warrington, UK), 100 nM primers and 50 times

diluted RT reaction mix using standard conditions on a 7500 Sequence detection system (Applied Biosystems, Foster City, CA, USA). The analysis was carried out at least on biological triplicates measured in duplicate. Primers are listed in Table S4.

Chromatin immunoprecipitation (ChIP)

The ChIP assay was performed as described previously (Ptasinska et al., 2011). Briefly, the cells were harvested and then resuspended to 2 x 10^7 in 10 ml of growth medium and cross-linked with 1% formaldehyde (Pierce, Rockford, IL, USA) for 10 min at RT. The cross-linking reaction was stopped by adding glycine to a final concentration of 0.4 M, followed by two washes with ice–cold PBS. Cells were resuspended in 10 ml of ice-cold ChIP buffer A (10 mM HEPES pH 8.0, 10 mM EDTA, 0.5 mM EGTA, 0.25% Triton X-100, proteinase inhibitor cocktail (Roche UK, Burgess Hill, UK) and 0.1 mM PMSF), incubated for 10 min at 4°C with rotation, and centrifuged 5 min at 500 x g at 4°C. The pellet was resuspended in 10 ml of ice–cold ChIP buffer B (10 mM HEPES pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.01% Triton X-100, protease inhibitor cocktail and 0.1 mM PMSF), incubated for 10 min at 4 °C with rotation and centrifuged for 5 min at 500 x g at 4 °C. Cells were resuspended in 600 µl of ice-cold ChIP lysis buffer (25 mM Tris-HCl pH 8.0, 150 mM NaCl, 2 mM EDTA, 1% Triton X-100, 0.25% SDS,protease inhibitor cocktail and 0.1 mM PMSF), incubated 10 min on ice and sonicated at 5 °C using a Bioruptor™ (Diagenode, Liege, Belgium) to generate fragments an average length of 200-500 bp (10 min with 30 s “ON” and “OFF” cycles, power setting high). The lysates were centrifuged for 5 min at 16,000 x g at 4 °C and the supernatants were diluted with two volumes of ice-cold ChIP dilution buffer (25 mM Tris-HCl pH 8.0, 150 mM NaCl, 2 mM EDTA, 1% Triton X-100, 7.5% glycerol, protease inhibitor cocktail and 0.1 mM PMSF). For each IP, 15 µl of Dynabeads® protein G were pre–incubated with 50 µg BSA and 2 µg antibody against C/EBPα Santa Cruz SC-61, ETO Santa Cruz SC-9737, HDAC2 Abcam ab7029, HEB Santa Cruz SC-357, LMO2 R&D Systems AF2726, LYL1 Santa Cruz SC-374164, PU.1 Santa Cruz SC-352, p300 Santa Cruz SC-585, RUNX1 Abcam ab23980 or IgG rabbit Milipore 12-370, IgG goat Santa Cruz SC-2346 and IgG mouse Santa Cruz SC-2025 for 2 h at 4 °C with rotation. The blocked antibody-bound protein G mix was added to 20–25 µg chromatin in a total volume of 500 µl diluted ChIP lysis buffer and incubated for 3 h at 4°C with rotation. After magnetic separation the beads were washed once with 1 ml wash buffer 1 (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 2 mM EDTA, 1% Triton X-100, 0.1% SDS), twice with 1 ml wash buffer 2 (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 2 mM EDTA, 1% Triton X-100, 0.1% SDS), once with 1 ml LiCl buffer (10 mM Tris-HCl pH 8.0, 250 mM

LiCl, 1 mM EDTA, 0.5% NP-40, 0.5% Na-deoxycholate) and twice with 1 ml TEN buffer (10 mM Tris-HCl pH 8.0, 50 mM NaCl, 1 mM EDTA). For each wash the beads were mixed with ice-cold washing buffers for 10 min at 4 °C. The immunoprecipitated DNA was eluted two times in 50 µl ChIP elution buffer (100 mM $NaHCO_3$, 1% SDS) for 15 min at RT with shaking. At this step the input control (1% of the starting material) was included in the experimental procedure after first adjusting the final volume to 100 µl with ChIP elution buffer. The eluted DNA was incubated overnight at 65 °C in the presence of 50µg proteinase K. The DNA was finally purified using Agencourt® AMPure® (Beckman Coulter, High Wycombe, UK) magnetic beads according to the manufacturer's instructions, eluted with 50 µl TE and analyzed by qPCR. Fold enrichment values were calculated relative to a negative control region of the genome. Primers are listed in Table S4.

Sequential chromatin immunoprecipitation: Re-ChIP

Re-ChIP was carried out as described above with minor modifications; following the final ChIP wash, chromatin complexes were eluted twice in 50 µl of ChIP elution buffer (100 mM $NaHCO_3$, 1% SDS, proteinase inhibitor cocktail) for 15 min at RT with shaking. Eluates were combined and diluted 20 times with ChIP incubation buffer (50 mM Tris pH 8.0, 100 mM NaCl, 2 mM EDTA, 1% NP40, 1mM DTT, protease inhibitor cocktail) followed by 5 h incubation with the second antibody or just beads for the second ChIP. After elution with 100 mM $NaHCO_3$, 1% SDS for 30 min at RT, Re-ChIP products were analysed by qPCR. Fold enrichment values were calculated relative to a negative control region of the genome. Primers are listed in Table S4.

DNase I hypersensitive site mapping

DNase I digestions were carried out on permeabilised CD34+ cells as previously described (2). Briefly, cells were suspended at a concentration of 3×10^7 cells/ml in nuclei digestion buffer (60 mM KCl, 15 mM NaCl, 5 mM MgCl2, 10 mM Tris pH7.4, 300 mM glucose). Digestions were then performed in 2 – 6 µg/ml DNase I (Worthington, DPPF grade) at 22ºC for 3 minutes, by adding the enzyme in an equal volume of digestion buffer containing 2 mM $CaCl_2$ and 0.4% Nonidet P-40. Genomic DNA was extracted from DNase I-treated cells using phenol/chloroform extraction, and then run out on 0.8% agarose gels. Levels of DNase I digestion were assessed using real time PCR, measuring the ratio of presence of known DNase I hypersensitive regions to more resistant gene free regions. Sequences of real time PCR primers are listed in Table S4. Fragments in the range of

100-600 bp in size, from samples with similar low levels of DNase I digestion, were excised and purified from gel slices in preparation for sequencing library generation.

Library preparation

Libraries of DNA fragments from chromatin immunoprecipitation or DNase I treatment were prepared from approximately 10 ng of DNA. Firstly, overhangs were repaired by treatment of sample material with T4 DNA polymerase, T4 PNK and Klenow DNA polymerase (all enzymes obtained from New England Biolabs UK) in a reaction also containing 50 mM Tris-HCl, 10 mM MgCl2, 10 mM Dithiothreitol, 0.4 mM dNTPs and 1 mM ATP. Samples were purified after each step using Qiagen MinElute columns (according to the manufacturer's guidelines). Adenosine bases were added to 3' ends of fragments using Klenow Fragment (3´- 5´ exo-minus), allowing for subsequent ligation of adapter oligonucleotides (Illumina part #1000521) using Quick T4 DNA ligase. After a further column clean up to remove excess adaptors, fragments were amplified in an 18 cycle PCR reaction using adapter-specific primers (sequences 5'-CAAGCAGAAGACGGCATACGAGCTCTTCCGATC*T and 5'-AATGATACGGCGACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATC*T). The libraries were purified and adapter dimers removed by running PCR products on 2% agarose gels and excising gel slices corresponding to fragments approximately 200-300 bp in size, which were then extracted using the Qiagen gel extraction kit. Libraries were validated using quantitative PCR for known targets, and quality assessed by running 1 µl each sample on an Agilent Technologies 2100 Bioanalyser. Once prepared, DNA libraries were subject to massively parallel DNA sequencing on an Illumina Genome Analyzer. ETO, RUNX1, C/EBPα, PU.1, LMO2 ChIP and Kasumi-1 DNase I libraries were sequenced employing the Illumina Genome Analyzer GAIIx, by using 36 base pair single end reads. For t(8;21) patient ♯1 and ♯2 DNase I libraries we used 50 bp single end reads on Illumina Hi-Seq.

Data analysis

Analysis of ChIP-sequencing data: The raw sequence reads in fast q format returned by the Illumina Pipeline was aligned to the hg19 human genome build using BWA (Li and Durbin, 2010). The reads in the resulting alignment files in sam format were used to generate density maps using bed-tools (Quinlan and Hall, 2010) and data was displayed using the UCSC Genome Browser (Kent et al., 2002). Regions of enrichment (peaks) of ChIP and DNAse1 sequencing data were identified using MACS (Zhang et al., 2008) and

cisGenome (Ji et al., 2008) software. The resulting peaks common for the two peak calling methods were considered for further analysis. Peaks overlap and gene annotations were performed using in-house scripts.

High confidence ChIP-Seq peaks were defined as peaks overlapping with the DNaseI-seq peaks. Overlaps between ChIP- and DNaseI- seq peaks were defined by requiring the summit of a peak in the ChIP dataset to lie between start and end coordinates of a peak in the DNaseI peaks. Peaks were allocated to genes if located in either their promoters or within the region of 500 bp downstream and 2000 bp upstream of the transcription start sites (TSS), as intragenic if not in promoter but within the gene body region, or if intergenic, to the nearest gene located within 100 kb.

Hierarchical clustering with Euclidean distance and complete linkage clustering was used for clustering of transcription factors (Figure 4D and S1A) based on similar binding patterns of different ChIP-seq data, in Kasumi-1 cells. The high confidence peaks for all transcriptional factors were intersected and merged when overlapping. The read counts for all union peaks were normalised with quantile normalisation and then Pearson's correlation coefficients were calculated between samples using log2 of the normalised read counts. A correlation matrix was generated and Pearson correlation coefficients are displayed after hierarchical clustering as a heatmap. Colors in the heatmap (Figure 4D and S1A) indicate the strength of association between each pair of transcription factors. Heatmaps were generated using Mev from TM4 microarray software suite (Saeed et al., 2006).

Analysis of p300 profiles were performed as follows: Common RUNX1 and RUNX1/ETO peaks, RUNX1 only peaks or PU.1 peaks that not bound by RUNX1 or RUNX1/ETO were used as reference coordinates against all aligned reads for p300 before and after RUNX1/ETO knockdown. Mean read density profiles were calculated for each 50bp-sized bins around peak summits up to +/-5000bp, these were normalised by the total p300 read counts. RUNX1 and RUNX1/ETO profiles were also performed in a similar way.

RUNX1 and RUNX1/ETO peaks were used as reference coordinates against all aligned readsfor the joint t(8;21) versus the CD34 positive cells DNaseI data to produce the scatter plots in supplementary Figure 1C. Normalised read counts were calculated for each peak +/- 200bp around peak summits.

Motif Analysis: HOMER (Heinz et al., 2010) was used for motif analysis. The Annotate Peaks function in HOMER was used to find occurrences of motifs in peaks. In this case we used HOMER database of known motif position weight matrices (PWM) with the most significant log p value and those that are expressed in RNA-seq of Kasumi-1 cells.

Digital footprinting of t(8;21) AML patients 1 and 2 from DNaseI high-depth sequencing data was performed using the Wellington algorithm (Piper et al., 2013) with FDR=0.01. The function dnase_average_profile.py function was used to generate the DNAse average profile plots and dnase_to_javatreeview.py to generate DNaseI cut where JavaTreeView was used to visualize the data. High read depth DNaseI-seq data from patient J171 were used. Footprinting analysis and motif searches were done within the union of PU.1, RUNX1 and HEB bound sequences in Kasumi- 1 cells (13584 peaks). Motif search was done within the footprint coordinates lying with the sequences that were bound by the above factors, but also in the footprints lying within sequences where factors did not bind. The DNaseI cuts and average profile were plotted +/- 100bp around ETS, RUNX and E-box motifs for both cases.

For the heat map that shows hierarchical clustering of motif occurrences within RUNX1/ETO peaks (Figure S1C), a motif positions search was done within +/- 200bp from RUNX1/ETO peaks centre. The distance between the centres of each motif pairs was calculated and the motif frequency was counted if the first motif is within 50bp distance from the second motif. Z-scores were calculated from the mean and standard deviation of motif frequencies observed in random sets using bootstrap analysis. For bootstrapping, peak sets of 400bps width and a population equal to that of RUNX1/ETO peaks were randomly obtained from the union of Kasumi-1 DNaseI-seq peaks. Motif search was repeated for each random set and then the mean and the standard deviation for the total motif frequencies of the random peak sets were calculated and compared with the actual motif frequencies to obtain the Z-scores. A matrix was generated and Z scores were displayed after hierarchical clustering as a heatmap. Red colour means that motifs are overrepresented and light blue indicates that motif is underrepresented. The same procedure was repeated with sequences containing RUNX1/ETO peaks that were only footprinted in t(8;21) patients. Motif search was done within the footprint coordinates and the random sets were generated from the total number of footprints in the patient sample.

Analysis of microarray data: The microarray gene expression data of Kasumi-1, SKNO-1 and t(8;21) primary cells were analysed as described in (Ptasinska et al., 2012) using Genome Studio software (Illumina, Little Chesterford, UK) with background subtraction. The raw data output by Genome Studio was analysed using the Lumi R package with quantile normalisation. The 10% threshold (p value <= 0.1) was applied to all data.

RNA-seq: RNA samples from three experimental replicates per condition were processed using the Tru-seq RNA Sample Prep Kit v2 from Illumina according to manufacturer's protocol. Libraries were run in 4x multiplex on an Illumina Hi-Seq 2000 sequencer generating ~90 million paired-end reads per sample. Quality of reads was verified using FASTQC analysis and reads aligned to HG19 Human Genome assembly using RNA-star(Dobin et al., 2013). Alignment quality was verified using RNA-SeQC(DeLuca et al., 2012) and read counts generated using HT-seq-count to align to the GENCODEv17 transcriptome. Finally differential expression analysis was performed using DEseq(Anders and Huber, 2010). Alternatively, RNA-Seq reads were aligned to the hg19 human genome build using TopHat. Fragments per Kilobase of transcript per Million mapped reads (FPKM) values for each gene were extracted using Cufflinks and differentially expressed genes were extracted using the limma R package. All genes with p-value ≤ 0.01 were considered. There are in total 12,121 genes that are expressed in Kasumi-1 cells with 1,030 changing expression at least twofold after RUNX1/ETO knockdown;586 genes were up- and 445 genes down-regulated. RNA-seq gene expressions were in good correlation with microarray gene expressions in Kasumi-1 cells before and after RUNX1/ETO knockdown with Pearson Correlation Coefficient ranging from 0.77 to 0.79.

Analysis of Combinatorial Binding: The high confidence peaks for all transcriptional factors before knockdown (RUNX1/ETO, CEBPα, HEB, LMO2, PU.1 and RUNX1) were intersected and merged when overlapping. This yielded to 61 different combinations involving binding of one or more factors. Z scores were calculated from the mean and standard deviation of random peak sets using bootstrap analysis; for bootstrapping, peak sets of 400bps width and populations equal to that of the six transcription factors peaks were randomly obtained from the union of all Kasumi-1 DNAseI-seq peaks. Peaks were intersected for each random peak sets and then the mean and the standard deviation for the total peaks overlap of the random peak sets were calculated and compared with the actual peak overlaps to obtain the z scores. Z scores were used here to indicate significance of deviation between observed and expected instances for all 61 binding patterns. The same procedure was also applied to the transcriptional factors after knockdown (CEBPα, LMO2, PU.1 and RUNX1)

The GSEA software (Subramanian et al., 2005) was used to perform gene set enrichment analysis on group of genes. In case a gene set had more than 500 genes, genes were selected such that peaks were ranked according to their position relative to the

transcription start site. The normalised enrichment score (NES), the p-value and the FDR q-value are displayed on the enrichment plot.

Module map (Segal et al., 2004) implemented by Genomic software is used to find which groups of genes were significantly up- or down-regulated using a statistical test based on the hyper-geometric distribution.

The regulatory network model for genes regulated by the occupancy patterns for the transcription factors was generated using BioLayout3D software (Theocharidis et al., 2009) using Markov Cluster Algorithm.

Gene ontology (GO) analysis was performed using Bingo (Maere et al., 2005) and David online tools at david.abcc.ncifcrf.gov (Huang da et al., 2009)using Hypergeometric for overrepresentation and Benjamini and Hochberg (FDR) correction for multiple testing corrections. Non redundant GO terms were filtered using REVIGO online tools at (http://revigo.irb.hr) with simRel as a similarity measure and a medium allowed similarity. KEGG Pathway network analysis was performed using clueGO tools (Bindea et al., 2009) with kappa score = 0.3. Kappa statistics was used to link the terms in the network. The right-sided enrichment (depletion) test based on the hyper-geometric distribution was used for terms and groups. Groups were created by iterative merging of initially defined groups based on the kappa score threshold. The color reflects the enrichment significance of the terms. The sizes of the nodes represent the number of the genes related to the term. The network was automatically laid out using the layout algorithm supported by Cytoscape

DNaseI-Seq profiles: A number of tools are designed for testing for differential DNaseI peaks; The MAnormR/Bioconductor package (Shao et al., 2012) was used to extract the DNaseI-seq differential peaks in Kasumi-1 cells.To determine the average DNaseI cutting frequency, start and end coordinates for positive and negative strand reads, respectively, were used as 5´ends of DNaseI digested fragments. Densities for DNaseI cut sites were thus generated for each base pair of the human genome. Average cut-site per bp profiles within sets of DHS were subsequently generated by retrieving densities [-50bp; +50bp] around the candidate motifs identified above using the Wellington BEDtoJTV function, and subdividing them into promoter and distal regions. In each case, DNaseI cut count normalization was performed based on average cut counts in the active space (promoter and distal regions) of each sample. Following normality testing, p-values of difference significance between the siRE and siMM datasets were subsequently computed via one-sided, paired t-test between total cut counts [+50bp; -50bp] around each candidate motif,

resulting in two lists containing a total number of cuts per region. Average DNaseI cut profiles were plotted using R software.

Supplementary References:

http://www-huber.embl.de/users/anders/HTSeq/doc/overview.html.

http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc.

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol *11*, R106.

Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., Fridman, W.H., Pages, F., Trajanoski, Z., and Galon, J. (2009). ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics *25*, 1091-1093.

DeLuca, D.S., Levin, J.Z., Sivachenko, A., Fennell, T., Nazaire, M.D., Williams, C., Reich, M., Winckler, W., and Getz, G. (2012). RNA-SeQC: RNA-seq metrics for quality control and process optimization. Bioinformatics *28*, 1530-1532.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15-21.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell *38*, 576-589.

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc *4*, 44-57.

Ji, H., Jiang, H., Ma, W., Johnson, D.S., Myers, R.M., and Wong, W.H. (2008). An integrated software system for analyzing ChIP-chip and ChIP-seq data. Nat Biotechnol *26*, 1293-1300.

Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., and Haussler, D. (2002). The human genome browser at UCSC. Genome Res *12*, 996-1006.

Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics *26*, 589-595.

Maere, S., Heymans, K., and Kuiper, M. (2005). BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. Bioinformatics *21*, 3448-3449.

Martinez, N., Drescher, B., Riehle, H., Cullmann, C., Vornlocher, H.P., Ganser, A., Heil, G., Nordheim, A., Krauter, J., and Heidenreich, O. (2004). The oncogenic fusion protein RUNX1-CBFA2T1 supports proliferation and inhibits senescence in t(8;21)-positive leukaemic cells. BMC Cancer *4*, 44.

Piper, J., Elze, M.C., Cauchy, P., Cockerill, P.N., Bonifer, C., and Ott, S. (2013). Wellington: a novel method for the accurate identification of digital genomic footprints from DNase-seq data. Nucleic Acids Res *41*, e201.

Ptasinska, A., Assi, S.A., Mannari, D., James, S.R., Williamson, D., Dunne, J., Hoogenkamp, M., Wu, M., Care, M., McNeill, H.*, et al.* (2012). Depletion of RUNX1/ETO in t(8;21) AML cells leads to genome-wide changes in chromatin structure and transcription factor binding. Leukemia *26*, 1829-1841.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841-842.

Saeed, A.I., Bhagabati, N.K., Braisted, J.C., Liang, W., Sharov, V., Howe, E.A., Li, J., Thiagarajan, M., White, J.A., and Quackenbush, J. (2006). TM4 microarray software suite. Methods Enzymol *411*, 134-193.

Segal, E., Friedman, N., Koller, D., and Regev, A. (2004). A module map showing conditional activity of expression modules in cancer. Nat Genet *36*, 1090-1098.

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

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S.*, et al.* (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A *102*, 15545-15550.

Theocharidis, A., van Dongen, S., Enright, A.J., and Freeman, T.C. (2009). Network visualization and analysis of gene expression data using BioLayout Express(3D). Nature protocols *4*, 1535-1550.

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