RESOURCE

nature immunology

A global network of transcription factors, involving E2A, EBF1 and Foxo1, that orchestrates B cell fate

Yin C Lin[1,9], Suchit Jhunjhunwala[1,9], Christopher Benner[2], Sven Heinz[2], Eva Welinder[1,3], Robert Mansson[1], Mikael Sigvardsson[4], James Hagman[5], Celso A Espinoza[6], Janusz Dutkowski[7,8], Trey Ideker[7,8], Christopher K Glass[2] & Cornelis Murre[1]

It is now established that the transcription factors E2A, EBF1 and Foxo1 have critical roles in B cell development. Here we show that E2A and EBF1 bound regulatory elements present in the *Foxo1* locus. E2A and EBF1, as well as E2A and Foxo1, in turn, were wired together by a vast spectrum of *cis*-regulatory sequences. These associations were dynamic during developmental progression. Occupancy by the E2A isoform E47 directly resulted in greater abundance, as well as a pattern of monomethylation of histone H3 at lysine 4 (H3K4) across putative enhancer regions. Finally, we divided the pro-B cell epigenome into clusters of loci with occupancy by E2A, EBF and Foxo1. From this analysis we constructed a global network consisting of transcriptional regulators, signaling and survival factors that we propose orchestrates B cell fate.

The basic unit of the chromatin fiber is the nucleosome. A nucleosome consists of a 146–base pair (bp) DNA segment folded around an octamer of histones composed of two copies each of H2A, H2B, H3 and H4. The amino-terminal tails of H2A, H2B, H3 and H4 are epigenetically marked and can be post-translationally modified by acetylation of lysine residues, methylation of arginine and lysine residues and ubiquitination of lysine residues, as well as phosphorylation of serine and threonine residues[1]. The acetylation of core histone residues generally correlates well with a chromatin fiber that is accessible to the transcriptional machinery[2]. Histone acetyltransferases catalyze the acetylation of histone tails. In contrast, deacetylation of histone core tails is mediated by histone deacetylases and is associated mainly with transcriptional repression. Methylation of histone residues promotes interactions of core histones with factors that have critical roles in chromatin remodeling. Methylation of core histone lysine residues is mediated by methyltransferases, whereas histone demethylases act to remove methyl groups from the tails of core histones. Nucleosome depletion is generally associated with active promoters in yeast as well as in *Drosophila*[2]. Active promoters are epigenetically marked by the acetylation of multiple H3 and H4 residues as well as trimethylation of histone H3 at lysine 4 (H3K4me3)[3]. In contrast, enhancer elements are characterized by monomethylation of H3K4 (H3K4me1)[4].

Hematopoiesis is initiated from a small pool of hematopoietic stem cells. These are pluripotent cells with the ability to differentiate into lymphoid-primed multipotent progenitors or, alternatively, into megakaryocyte-erythroid progenitors. Lymphoid-primed multipotent progenitors, in turn, differentiate into common lymphoid progenitors or granulocyte-macrophage progenitors. B cell development requires a hierarchy involving many transcription factors, including E2A (A000804). In common lymphoid progenitors, the helix-loop-helix protein E47 (A000806), an E2A isoform encoded by *Tcf3* (called '*E2A*' here) induces the expression of early B cell factor 1 (EBF1 (A000809))[5–8]. EBF1 in turn activates the expression of the transcription factor Pax5 (A000403)[9]. *Pax5* expression is also regulated by the concerted activities of the ETS-family transcription factor PU.1 and the interferon-regulatory factors IRF4 and IRF8 (refs. 9–11).

In addition to E2A, EBF1 and Pax5, other transcriptional regulators that modulate the developmental progression of B lineage cells have been identified. Prominent among these are Bcl-11a and Foxo1 (A000944). B cell development in Bcl-11a-deficient mice is arrested at a stage similar to that described for E2A- or EBF1-deficient mice[12]. Foxo1-deficient mice show a block at the pro-B cell stage, and Foxo1 directly activates expression of recombination-activating gene 1 (*Rag1*) and *Rag2* (refs. 13,14). Collectively, these factors form the core of the transcriptional machinery that promotes commitment to the B cell lineage, suppresses the expression of genes associated with alternative cell fates and coordinates cellular population expansion with developmental progression[15,16].

As a first approach to determine how those factors are linked to promote B cell fate, we did a genome-wide analysis for occupancy by E2A, EBF1 and Foxo1. We found that the DNA-binding activities of E2A, EBF1 and Foxo1 were linked to global patterns of

[1]Department of Molecular Biology, University of California, San Diego, La Jolla, California, USA. [2]Department of Cellular and Molecular Medicine, University of California, San Diego, La Jolla, California, USA. [3]Center for Stem Cell Biology and Cell Therapy, Lund University, Lund, Sweden. [4]Department for Biomedicine and Surgery, Linkoping University, Linkoping, Sweden. [5]Integrated Department of Immunology, National Jewish Health, Denver, Colorado, USA. [6]Ludwig Institute for Cancer Research, University of California, San Diego, La Jolla, California, USA. [7]Department of Bioengineering and Department of Cellular and Molecular Medicine, University of California, San Diego, La Jolla, California, USA. [8]The Institute for Genomic Medicine, University of California, San Diego, La Jolla, California, USA. [9]These authors contributed equally to this work. Correspondence should be addressed to C.K.G. (cglass@ucsd.edu) or C.M. (murre@biomail.ucsd.edu).

Received 8 March; accepted 19 May; published online 13 June 2010; doi:10.1038/ni.1891

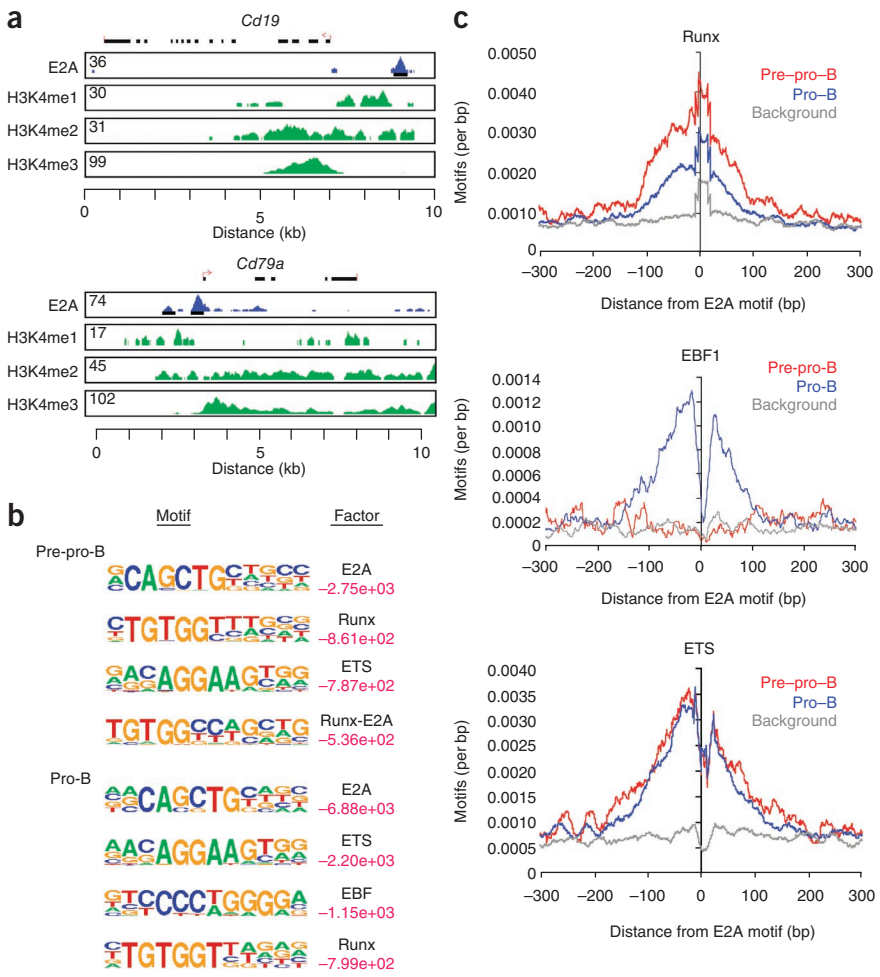

Figure 1 E2A occupancy and epigenetic marking in cultured EBF1-deficient pre-pro-B cells and RAG-1-deficient pro-B cells. (**a**) E2A occupancy (blue) and H3K4me1, H3K4me2 and H3K4me3 regions (green) in the *Cd19* and *Cd79a* loci, identified by ChIP-Seq. Red arrows indicate TSS; numbers in top left corners indicate tags observed. (**b**) *Cis*-regulatory sequences associated with E2A occupancy in pre-pro-B cells and pro-B cells, identified by comparison of enriched peaks with randomly selected genomic DNA sequences. Letter size indicates nucleotide frequency. Right margin (red), log *P* values; smaller values indicate greater enrichment for a given motif than for randomly selected regions. (**c**) Frequency of EBF, ETS and RUNX consensus binding sites located in close genomic proximity to E2A-bound regions. Data are from five independent experiments.

H3K4 methylation. The *cis*-acting targets ('cistromes') associated with E2A occupancy were mostly monomethylated at H3K4 but underwent further epigenic alteration during the transition from the pre-pro-B cell stage to the pro-B cell stage. These altered cistromes frequently associated with EBF1 and with Foxo1, albeit to a lower extent. We also found that in pro-B cells, E2A and EBF1 bound regulatory elements in the *Foxo1* locus and that E2A, EBF and Foxo1 in turn interacted with putative enhancer elements in the *Pax5* locus. The coordinated DNA-binding activities of E2A and Foxo1 had functional importance, as mice heterozygous for both *E2A* and *Foxo1* showed a nearly complete absence of B cells. These data link E2A and the phosphatidylinositol-3-OH kinase (PI(3)K) cascade in a common pathway. Finally, we describe the pro-B cell state in terms of a global network of transcription factors that involves E2A, EBF1 and Foxo1.

RESULTS

Cis-regulatory sequences associated with E2A occupancy

Published observations have established that the E2A proteins act to modulate lymphocyte differentiation, cell survival and cellular population expansion and to suppress the development of lymphoma[17]. To determine how the E2A proteins mechanistically coordinate developmental progression and cell growth, we assessed E2A occupancy in pre-pro-B cells (EBF1 deficient) and pro-B cells (RAG-1 deficient) by chromatin immunoprecipitation combined with deep DNA sequencing (ChIP-Seq). We identified more than 4,531 E2A-bound sites in pre-pro-B cells, whereas in pro-B cells, 11,846 binding sites showed E2A occupancy. Approximately 44% of the E2A-bound sites in pre-pro-B cells were the same sites occupied by E2A in pro-B cells (1,993 of 4,531 sites). However, the majority of E2A occupancy was associated exclusively with pro-B cells (9,853 of 11,846 sites), consistent with the critical roles of E2A proteins at this developmental stage. Although we observed binding of E2A to sites in close proximity to the transcription start sites (TSSs), E2A occupancy was widespread and was associated mainly with either introns or intergenic regions (**Supplementary Fig. 1a** and **Fig. 1a**).

To determine whether E2A interacts 'collaboratively' with alternative transcriptional regulators in EBF1-deficient pre-pro-B cells and RAG-1-deficient pro-B cells, we examined E2A-associated sites for enrichment in sequence elements. We used the motif-finder algorithm HOMER (Hypergeometric Optimization of Motif EnRichment) to identify motifs localized within 100 bp upstream or downstream of E2A occupancy (**Fig. 1b** and **Supplementary Table 1**). The HOMER algorithm ranks motifs in transcription factor–associating DNA sequences on the basis of their statistical enrichment. As expected, the E2A consensus DNA sequence (CAGCTG) was ranked as the top-scoring motif in both pre-pro-B cells and pro-B cells (**Fig. 1b**). In pre-pro-B cells, E2A-bound regions were enriched consensus motifs for both the transcription factor Runx1 (A000523) and ETS (**Fig. 1b**). In contrast to such regions in pre-pro-B cells, E2A-bound regions in pro-B cells were considerably enriched for EBF1 high-affinity binding sites (**Fig. 1b**,**c** and **Supplementary Table 1**).

To determine the relationship between E2A occupancy and its associated *cis*-regulatory elements, we plotted the genomic distances that separated EBF1, ETS and Runx consensus motifs from E2A-binding sites (**Fig. 1c**). This analysis showed that a 200-bp area centered on E2A-bound regions was highly enriched for EBF1, Runx and ETS motifs (**Fig. 1c**). Notably, the pattern of Runx occupancy was distinct from that observed for EBF1 and ETS DNA binding. Rather than showing a bimodal distribution, the Runx consensus motif seemed to overlap E2A occupancy. These data indicate that the spacing that separates E2A- and Runx-binding sites is under strict selective pressure and may reflect a special requirement for 'cooperative' rather than 'collaborative' DNA binding.

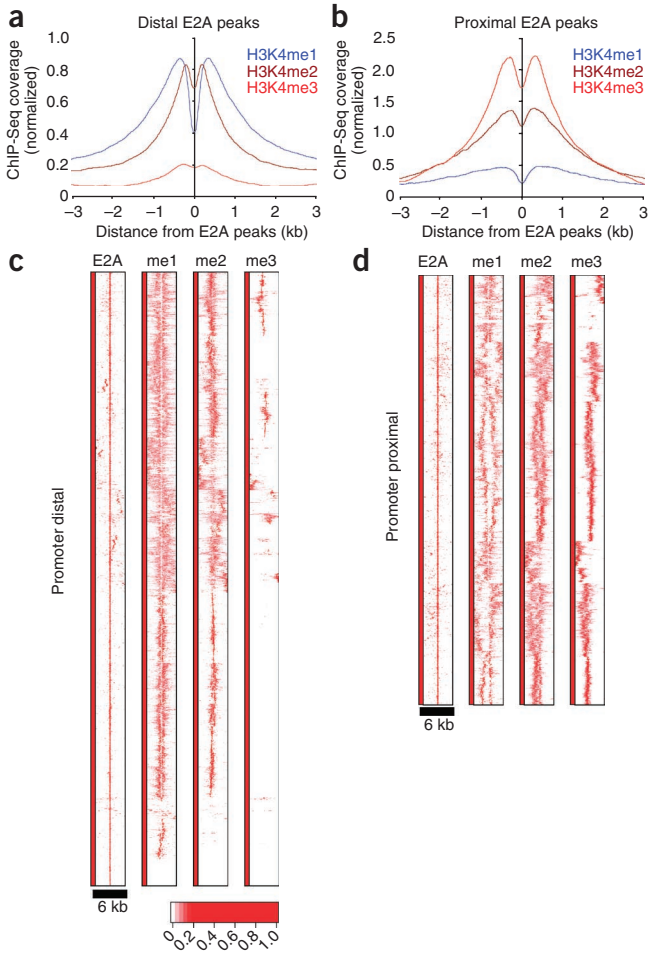

Figure 2 E2A occupancy and H3K4 methylation in cultured RAG-1-deficient pro-B cells. (**a**,**b**) H3K4 methylation (H3K4me1, H3K4me2 and H3K4me3) plotted as a function of genomic separation from promoter-distal E2A-bound sites (>3 kb from nearby TSSs; **a**) or promoter-proximal E2A-bound sites (<3 kb from nearby TSSs; **b**). (**c**,**d**) Heat map of E2A occupancy and H3K4me1, H3K4me2 and H3K4me3 patterns for promoter-distal DNA sequences (>3 kb from nearby TSSs; **c**) or promoter-proximal regions (<3 kb from nearby TSSs; **d**). Each row indicates an E2A-bound DNA fragment. Colors indicate enrichment for E2A binding and H3K4 methylation (scale below). Data are representative of four independent experiments.

Association of E2A occupancy with H3K4 methylation

Published studies have demonstrated tight correlation between transcriptionally active promoters and trimethylation of H3K4, whereas monomethylation of H3K4 has been associated with enhancer activity[2–4,18–20]. To determine whether E2A occupancy was allied with either promoter or enhancer regions, we immunoprecipitated DNA fragments from lysates of RAG-1-deficient pro-B cells with antibody to H3K4me1 (anti-H3K4me1), anti–dimethylated H3K4 (H3K4me2) and anti-H3K4me3 and analyzed them by ChIP-Seq. Consistent with previous observations[3], H3K4me3 was closely associated with the active promoter regions of the *Cd19* and *Cd79a* loci (**Fig. 1a**). We also analyzed RAG-1-deficient pro-B cells separately for global patterns of acetylated H3 (H3K9 and H3K14)[21]. Most E2A binding was associated with monomethylation of H3K4, whereas only a subset was associated with H3 acetylation (**Supplementary Fig. 1a**).

To determine whether epigenetic marks are globally aligned with E2A occupancy, we plotted cumulative sequence tag frequencies associated with E2A-bound sites for each of the histone marks (**Fig. 2** and **Supplementary Fig. 1**). To demonstrate the distribution of H3K4 methylation, we plotted the degree of H3K4 methylation as a function of genomic separation from the E2A-bound sites (**Fig. 2a**,**b**). This analysis showed a bimodal distribution with substantially less H3K4 methylation centered on the E2A-bound sites (**Fig. 2a**,**b**).

To examine the shape of the distributions centered on individual E2A-bound sites, we generated heat maps of 1×10^4 bound regions and analyzed these genome wide. We examined E2A occupancy in promoter-distal regions (separated from the TSS by >3 kilobases (kb)), as well as in promoter-proximal regions (separated from the TSS by <3 kb; **Fig. 2c**,**d**). We plotted normalized tag densities for promoter-distal or promoter-proximal regions for E2A occupancy as well as H3K4me1, H3K4me2 and H3K4me3. We aligned E2A-binding sites at the center and designated these 'position 0'. We identified and visualized genomic regions in a 3-kb window upstream and downstream of E2A occupancy (**Fig. 2c**,**d**). E2A occupancy was associated mainly with H3K4me1 and H3K4me2 in distal regions (**Fig. 2c**). Only a minor proportion of E2A occupancy was accompanied by H3K4me3 in regions distal to the promoter (**Fig. 2c**). As expected, however, in proximal elements, most binding of E2A was linked with trimethylation of H3K4 (**Fig. 2d**).

Two basic patterns of H3K4 monomethylation emerged: a focused and narrow distribution, as well as a diffuse and asymmetric pattern (**Fig. 2c**,**d**). The pattern of H3K4me1 in promoter regions was diffuse for most of the segments analyzed[22] (**Fig. 2d**). When we focused on H3K4me1 sites in pro-B cells, we found that 22% were associated with E2A occupancy (data not shown). In sum, this analysis indicates that a substantial proportion of putative enhancer elements in pro-B cells are bound by E2A. Furthermore, the data show that the patterns of H3K4 methylation linked with E2A occupancy are distinct in promoter-proximal regions versus promoter-distal regions.

E2A occupancy and H3K4 methylation

To determine whether E2A occupancy changes during early B lineage development, we examined binding of E2A to DNA in both EBF1-deficient pre-pro-B cells and RAG-1-deficient pro-B cells. We gated the binding sites on genomic regions located distal to the promoter. We aligned 1×10^4 randomly chosen E2A-binding sites in pre-pro-B cells as well as pro-B cells at the center. We plotted the promoter-distal subset of those sites (7,296) and directly compared them with the patterns of H3K4me1, H3K4me2 and H3K4me3.

As we observed many patterns of E2A occupancy and H3K4 methylation in pre-pro-B cells and pro-B cells, we used hierarchical clustering to categorize patterns of E2A occupancy and H3K4 methylation across the two developmental stages (**Fig. 3**). To identify such clusters, we used the Ward hierarchical clustering method[23]. This analysis showed distinct groups that were distinguished from each other on the basis of E2A occupancy and its associated patterns of H3K4me. A subset of DNA segments (clusters I–IV) showed E2A occupancy in pro-B cells but not in pre-pro-B cells. Cluster V showed binding of E2A at the pre-pro-B cell stage, whereas another subset (cluster VI) showed E2A occupancy at both the pre-pro-B cell stage and the pro-B cell stage.

The difference between pre-pro-B cells and pro-B cells in their patterns of H3K4me was distinct but overall correlated well with E2A occupancy across the entire spectrum of binding sites (**Fig. 3**). Most E2A occupancy correlated well with H3K4me1 and H3K4me2 in both pre-pro-B cells and pro-B cells. In pre-pro-B cells, H3K4me1 symmetrically flanked most E2A-binding sites (clusters V and VI).

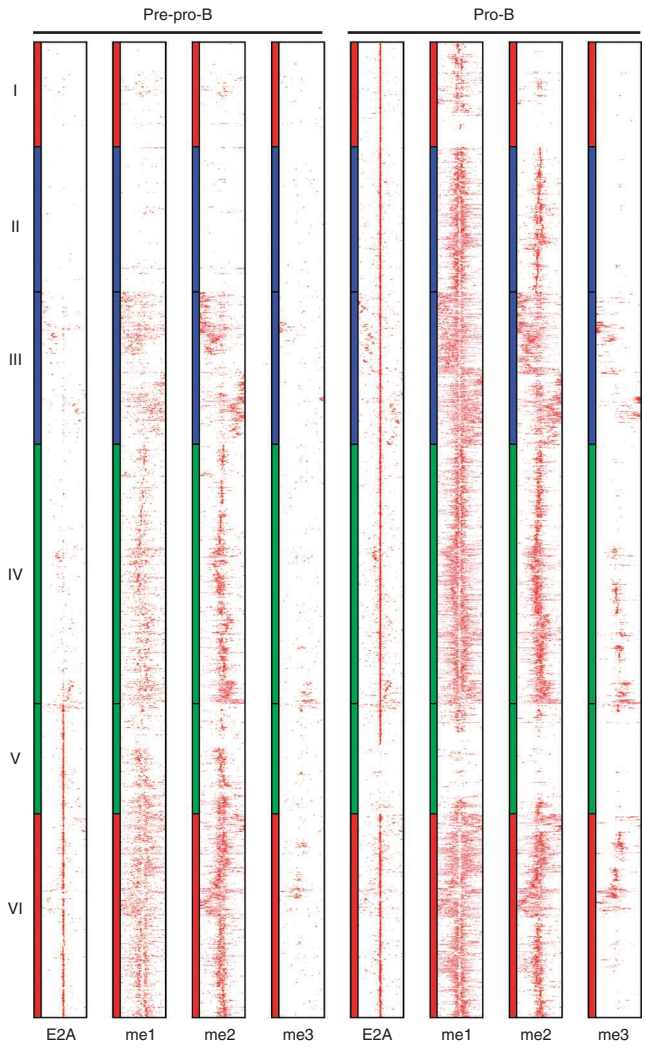

Figure 3 E2A occupancy and patterns of H3K4 methylation in cultured EBF1-deficient pre-pro-B cells and RAG-1-deficient pro-B cells. Heat map of E2A occupancy and distribution of H3K4me1, H3K4me2 and H3K4me3 for promoter-distal regions (>3 kb from nearby TSS; 7,296 sites) in pre-pro-B cells (left) and pro-B cells (right); cistronic elements were grouped by the Ward hierarchical clustering method. Each row represents an E2A-bound DNA fragment. Colors (far left) identify clusters. Clusters I, II, III, IV, V and VI (left margin) represent groups of 785, 1,083, 1,140, 1,940, 824 and 1,524 binding sites, respectively, more closely related to each other than those assigned to different clusters. Data are from eight independent experiments.

However, a subset of E2A-binding sites in pre-pro-B cells was not associated with H3K4me1. In pro-B cells, we observed many patterns of H3K4 methylation that correlated with E2A occupancy. Cluster I either lacked H3K4me1 or showed a diffuse and subtle pattern of H3K4me1. Clusters II and IV showed a focused pattern of H3K4me1 that was centered on the E2A-bound sites. In contrast, cluster III was associated with a diffuse and asymmetric distribution of H3K4me1. Cluster VI showed a mixed pattern of H3K4me1 that was mostly focused and symmetric but also included regions with asymmetric distribution.

To explore the possibility that binding patterns were associated with unique *cis*-regulatory sequences, we assessed the presence of enriched DNA sequences that were associated with E2A occupancy by both the motif-finding tool MEME (Multiple EM for Motif Elicitation) as well as HOMER (for statistical values). The E2A consensus motif was largely invariant across each of the clusters (**Fig. 4**). However, in each cluster, distinct *cis*-regulatory sequences were associated with E2A occupancy.

Notably, cluster I showed a substantial enrichment for *cis*-regulatory sequences closely associated with CTCF, a protein involved in modulating large-scale chromatin structure. The potential association of E2A with CTCF raised the possibility that E2A and CTCF act in concert to modulate long-range chromatin structure. Consistent with that idea, *cis*-regulatory elements containing both E2A- and CTCF-binding sites predominantly lacked the enhancer mark H3K4me1 (**Figs. 3** and **4**). Clusters II and III were associated mainly with a composite ETS-E2A consensus DNA sequence (AGGAAG-CAGCTG) as well as an EBF1 consensus DNA sequence (TCCCNNGGGA). Cluster IV showed substantial enrichment for a PU.1 consensus DNA sequence (AGGAAG). Cluster V, associated mainly with E2A occupancy in pre-pro-B cells but not in pro-B cells, was linked to ETS-E2A consensus DNA sequences. Consensus DNA sequences for Foxo1 (G/AGAAACAGCTG, where 'G/A' indicates guanosine or adenosine, and underlining indicates the consensus sequence) as well as Runx (TGTGGTT) seemed to be linked mainly with cluster VI, a group of DNA segments that showed E2A occupancy in both pre-pro-B cells and pro-B cells (**Figs. 3** and **4**). Together, the clustering analyses showed that in pro-B cells, E2A occupancy was associated with a wide spectrum of *cis*-regulatory elements.

Coordinated E2A, EBF1, Foxo1 and CTCF occupancy

The E2A-associated motif-identification analyses presented above suggested coordinated DNA-binding activity of E2A and EBF, E2A and Foxo1, as well as E2A and CTCF, to a subset of regulatory elements. To confirm the concerted DNA-binding activities of E2A and its putative partners, we immunoprecipitated proteins from lysates of RAG-1-deficient pro-B cells with anti-EBF1, anti-Foxo1 and anti-CTCF and analyzed them by ChIP-Seq. Occupancy by EBF1, Foxo1and CTCF was associated mainly with intergenic and intronic regions (data not shown).

The top-scoring motif associated with EBF1 occupancy was the EBF1 consensus-binding site (**Fig. 5a**, left, and **Supplementary Table 1**). Consistent with the data described above, the E2A-binding site was frequently associated with EBF1 occupancy (**Fig. 5a** and **Supplementary Fig. 2**). Using the clustering analysis, we predicted that E2A and Foxo1 also bind in a coordinated manner to a subset of *cis*-regulatory elements. To test our prediction directly, we immunoprecipitated proteins from lysates of RAG-1-deficient pro-B cells with anti-Foxo1 and analyzed them by ChIP-Seq. DNA fragments associated with Foxo1 occupancy were, as expected, enriched for the Foxo1 consensus binding site, as well as E2A and ETS-Foxo1 consensus binding sites (**Fig. 5a**, right). We plotted normalized tag densities for distally located regulatory elements for E2A, H3K4me1, H3K4me2 and H3K4me3 for a 6-kb window across Foxo1 DNA-binding sites (**Fig. 5b**). A sizable fraction of Foxo1-bound sites were associated with E2A occupancy (**Fig. 5b**).

As described above, the computational analysis showed that a small fraction of E2A-binding sites were enriched for CTCF consensus binding sequences (Cluster I; **Fig. 4**). These data raised the possibility that a tiny but important fraction of binding of E2A is associated with CTCF occupancy. To directly assess whether E2A and CTCF bind coordinately to a subset of *cis*-regulatory sequences, we immunoprecipitated CTCF from pro-B cell lysates and analyzed it by ChIP-Seq. As predicted from the computational analysis, a subset

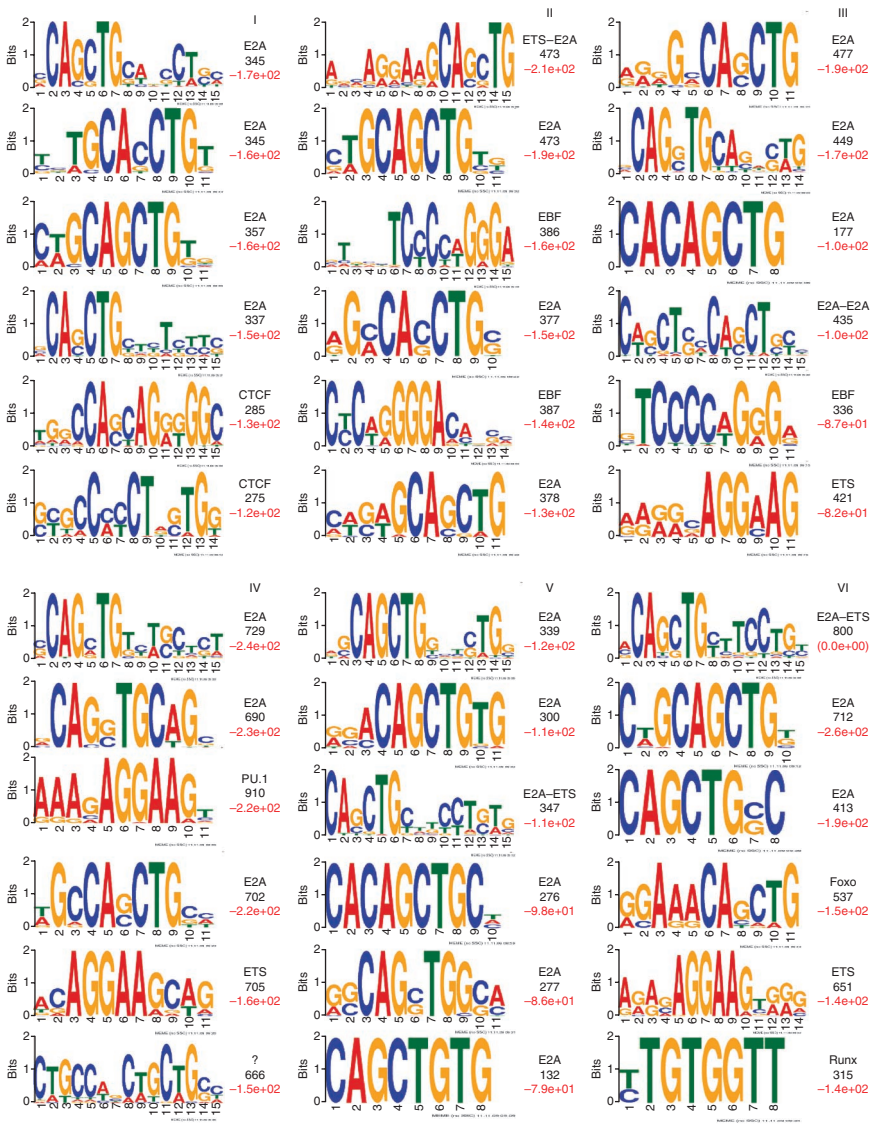

Figure 4 Distinct *cis*-regulatory DNA sequences associate with E2A occupancy in cultured RAG-1-deficient pro-B cells. *Cis*-regulatory sequences associated with E2A occupancy and patterns of H3K4me1 identified by comparison of enriched binding sites (clusters as described in **Fig. 3**) to randomly selected genomic DNA sequences, gated on DNA sequences located 100 bp upstream or downstream of the E2A-binding sites; sequences distinct from each cluster were pooled and analyzed by the MEME algorithm. Letter size indicates frequency. Right, ranking of the top six scoring motifs (by log *P* value (red; HOMER)), showing the transcription factors associated with *cis*-regulatory sequences (CTCF, ETS, EBF, Pu.1, Foxo and Runx); numbers in black indicate nonredundant motif occurrences. Data are from one experiment.

of CTCF occupancy showed coordinated binding of E2A (**Supplementary Figs. 3a**,**b** and **4**). Together these results showed that in pro-B cells, E2A binds coordinately with EBF1, Foxo1 and CTCF to a wide spectrum of *cis*-regulatory elements.

Global transcription factor occupancy and B cell development

To determine the extent of coordinated binding of E2A and EBF1, we measured genomic distances between the two closest binding sites. Nearly 50% of EBF1-binding sites were flanked by E2A-binding sites within 150 bp (**Supplementary Fig. 5a**). To determine whether coordinated binding of E2A and EBF1 directly relates to differences in gene expression, we correlated transcript abundance in pre-pro-B cells and pro-B cells with occupancy by E2A and EBF1. We assigned E2A- and EBF1-binding sites to a specific locus on the basis of the genomic distance that separated them from the nearest TSS. For this, we derived mRNA from EBF1-deficient pre-pro-B cells and RAG-1-deficient pro-B cells and assessed absolute gene expression by microarray analysis. We determined the distribution of transcript abundance in pre-pro-B cells and pro-B cells (**Fig. 6a**). A substantial fraction of genes enriched for coordinated binding of E2A and EBF1 were positively associated with a greater abundance of pro-B cell–specific transcripts than that of genes with binding sites for binding of either E2A or EBF1 alone (**Fig. 6a**). The *P* values were highly significant because of the large number of genes analyzed. Among the loci with coordinated binding of E2A and EBF1 DNA were a large set of loci known to have critical roles in early B cell development, including *Vpreb1, Vpreb3, Cd19, Bst1, Foxo1, Pou2f1, Cd79a* and *Pax5* (**Supplementary Table 2**). We found that pro-B cells had higher expression of 1,753 genes analyzed than did pre-pro-B cells (>1.5-fold change, *P*-value < 0.01). We found that 9% of those showed coordinated binding of E2A and EBF1 within a 150-bp DNA segment (**Supplementary Table 2**). Less than 2% of the genes contained one EBF1-binding site but lacked E2A occupancy (**Supplementary Table 3**). We found that 43% of loci were associated with at least one E2A-binding site but did not show coordinated EBF1 occupancy. Together these data indicate that E2A and EBF1 bind in a coordinated way to activate a B lineage program of gene expression. These data are consistent with genetic data indicating that E2A and EBF1 act synergistically to modulate early B cell development[11].

We also observed co-occupancy by Foxo1 and E2A (**Supplementary Fig. 5b**), but to a much lesser degree for Foxo1 and EBF1 (**Supplementary Fig. 5c**). The fraction of genes enriched for coordinated E2A- and Foxo1-binding activity was positively associated with a greater abundance of pro-B cell–specific transcripts than that of loci with binding of either E2A or Foxo1 (**Fig. 6a**). Among the loci with co-occupancy by E2A and Foxo1 were *Blk*, *Dntt* (encoding terminal deoxynucleotidyltransferase) and *Erg* (**Supplementary Tables 4** and **5**). The well known E2A and Foxo1 target gene *Bcl2l11* (encoding the proapoptotic molecule Bim) showed coordinated binding of E2A and Foxo1 (**Supplementary Table 4**). Notably, CTCF-binding sites were less frequently associated with loci showing pro-B cell lineage–specific transcription than were E2A-binding sites (**Fig. 6a**).

To further explore the coordinated activities of E2A and Foxo1 in B cell development, we adopted a genetic strategy. We analyzed abnormalities in B cell development in mice heterozygous for *E2A* (*E2A*$^{+/-}$) or heterozygous for a *lox*P-flanked *Foxo1* allele with expression of Cre recombinase driven by the estrogen receptor gene

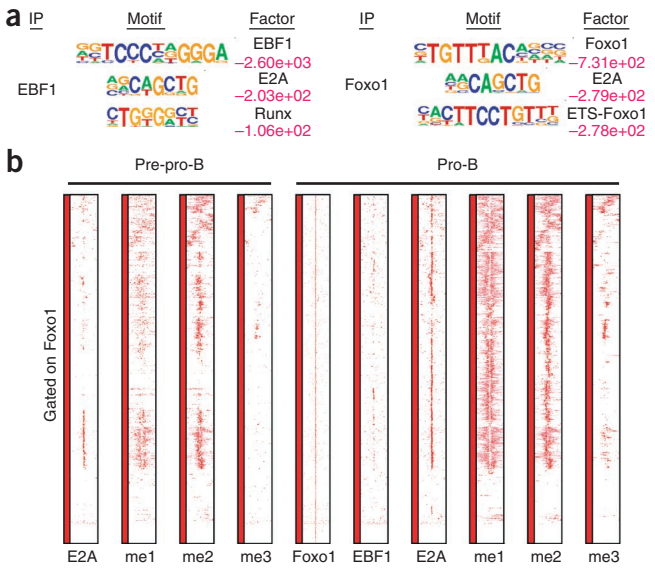

Figure 5 Distinct patterns of H3K4 monomethylation are associated with coordinated occupancy by E2A and EBF1 and Foxo1. (**a**) ChIP-Seq analysis of *cis*-regulatory sequences associated with occupancy by EBF1 or Foxo1 in lysates of RAG-1-deficient pro-B cells immunoprecipitated (IP) with anti-EBF1 (left) or anti-Foxo1 (right); motifs with E2A occupancy were identified by comparison of enriched peaks to randomly selected genomic DNA sequences. Letter size indicates frequency of detection. Numbers at right (red) indicate log *P* values; smaller values represent greater enrichment than that of randomly selected regions. (**b**) Heat map of Foxo1 occupancy and H3K4me1, H3K4me2 and H3K4me3 patterns centered on 1,118 promoter-distal Foxo1-binding sites, showing enrichment for binding of Foxo1 and E2A as well as H3K4 methylation. Each row represents a 6-kb genomic region centered on Foxo1 occupancy. Data are from ten independent experiments.

(*Foxo1*$^{f/+}$–ER-Cre), or mice heterozygous for both (*E2A*$^{+/-}$*Foxo1*$^{f/+}$–ER-Cre), after treatment with tamoxifen. The pro-B cell compartment in *E2A*$^{+/-}$*Foxo1*$^{f/+}$–ER-Cre mice was only modestly affected 10 d after the initial tamoxifen treatment (**Fig. 6b**). In contrast, the pre-B cell and immature B cell compartments were much smaller (both percentage and absolute numbers) in *E2A*$^{+/-}$*Foxo1*$^{f/+}$–ER-Cre mice than in *E2A*$^{+/-}$ or *Foxo1*$^{f/+}$–ER-Cre mice (**Fig. 6b**), which reflected a critical role for the coordinated activities of E2A and Foxo1 during the transition from the pro-B cell stage to the pre-B cell stage.

The coordinated DNA-binding activities of E2A and EBF1 to sites in H3K4me1 islands is consistent with their acting as enhancer-binding proteins. To determine directly whether the binding patterns indeed reflected lineage-specific enhancer activity, we assayed the enhancer activity of 24 segments that were H3K4 monomethylated and that showed coordinated binding of E2A and EBF1. We inserted putative enhancer elements upstream of a basal promoter (adenovirus major late promoter) in a luciferase reporter construct and transfected those constructs into a mouse pro-B cell line (22D6) and two T cell lines (166 and A12), then assessed luciferase activity 1 d after transfection. As expected, the majority of reporter constructs (20 of 24) that contained coordinated E2A- and EBF1-binding sites showed substantial enhancer activity in the pro-B cell line 22D6 (**Fig. 6c**). In contrast, only a minor proportion of these (9 of 24) showed enhancer activity in the T cell lines 166 and A12 (**Fig. 6c**). We found that 3 of 24 regulatory elements containing both E2A- and EBF1-binding sites suppressed reporter gene expression in the Pro-B cell line 22D6 (**Fig. 6c**). Additionally, several of the reporters showed activation in B lineage cells, whereas we observed repression in T cells (**Fig. 6c**). Thus, regulatory elements carrying E2A- and EBF1-binding sites can function as either transcriptional enhancers or silencers depending on the cell context.

Genome-wide monomethylation of H3K4 by E47

As described above, we found that the E2A proteins bound to a wide repertoire of putative enhancers. This left the question of how the E2A proteins modulate enhancer function on a global scale. As a first approach to answering this, we determined whether the E2A proteins were able to modulate chromatin structure. For this, we transduced E2A-deficient pre-pro-B cells with virus carrying the full-length coding DNA sequence of E47 fused to the estrogen receptor domain (E47-ER), as well as a human CD25 reporter. As a control, we transduced cells with virus carrying the gene encoding the E47 bHLH (basic helix-loop-helix) region but fused to the estrogen receptor domain (bHLH-ER). At 21 h after infection, we cultured cells for an additional 1 h or 6 h in the presence of 4-hydroxytamoxifen. We enriched samples for transduced cells through the use of magnetic beads, then immunoprecipitated E2A as well as H3K4me1 and analyzed these by ChIP-Seq. In the presence of E47 lacking the amino-terminal transactivation domains (bHLH-ER), the pattern of H3K4 monomethylation peaked across the E47-binding sites (**Fig. 7a**), which indicated that H3K4me1 residues were closely associated with the E47 bHLH domain. Notably, after forced expression of E47 containing the bHLH domain as well as the amino-terminal transactivation domains (E47-ER), we observe a bimodal distribution of H3K4me1 within 1 h of treatment with 4-hydroxytamoxifen (**Fig. 7a**), which indicated that the E2A proteins acted directly to alter the pattern of H3K4me1. We observed a similar redistribution of H3K4me1 after forced E47 expression in an E2A-deficient thymoma cell line (**Fig. 7b**). To exclude the possibility that the E2A-binding sites where histone was remodeled were different from the bHLH-binding sites where histone was not remodeled, we directly compared the H3K4me1 patterns for bHLH-ER as well as full-length E47 and found that the binding sites overlapped (**Supplementary Fig. 6**). These data show that E47 acts directly to increase the abundance as well as the pattern of H3K4me1 across a wide repertoire of putative enhancer regions.

Network construction

The data presented above raised the question of how the genome-wide DNA-binding analysis of individual transcription factors can be linked to a global network of interacting components that collectively promote the B cell fate. To answer this, we divided the genome into *cis*-coding elements occupied by E2A, EBF1 and Foxo1 using the computational approach Cytoscape[24,25]. We identified putative regulatory targets of E2A on the basis of two criteria. First, we selected genes with a significant change in transcript abundance in pro-B cells versus E2A-deficient pre-pro-B cells (≥1.5-fold change; $P \leq 0.05$). Second, we assessed this subset of genes for E2A occupancy. Similarly, to analyze EBF1 targets, we compared transcript abundance in pro-B cells and EBF1-deficient pre-pro-B cells, then assessed this subset of genes for sites with EBF1 occupancy. We did not directly analyze loci with Foxo1 occupancy for differences in transcript abundance in pro-B cells versus Foxo1-ablated cells. Differences in transcript abundance in E2A- and EBF1-deficient pre-pro-B cells versus wild-type pro-B cells may not necessarily reflect direct regulation by either E2A or EBF1 but instead reflect the developmental stages at which B cell maturation is arrested. Moreover, co-occupancy relates here

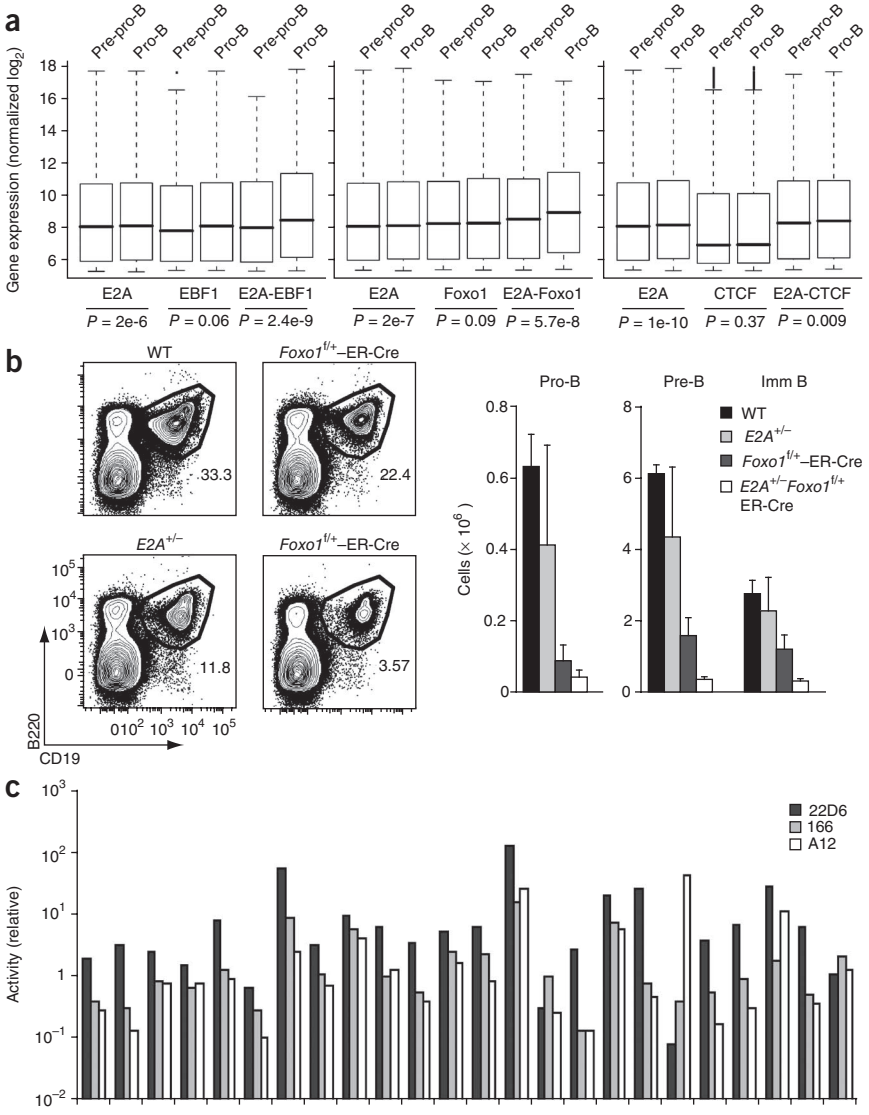

Figure 6 Coordinated binding of E2A, EBF and Foxo1 is associated with a B lineage–specific program of gene expression. (**a**) Distribution of RNA expression in pre-pro-B and pro-B cells. Genes closest to the binding sites (measured from the TSS) were categorized (left graph) as those with E2A- and EBF1-binding sites located within 150 bp of each other (right two bars); those in the vicinity of EBF1-binding sites with no E2A peak within 150 bp (middle two bars); and those in the vicinity of E2A-binding sites with no EBF1 peak within 150 bp (left two bars). Middle and right graphs, same analysis of E2A- and Foxo1-binding sites (middle) and E2A- and CTCF-binding sites (right). Bottom, *P* values (paired *t*-test). Data are from one experiment. (**b**) Flow cytometry analysis of the expression of B220 and CD19 (left) in wild-type (WT), *Foxo1*$^{f/+}$–ER-Cre, *E2A*$^{+/-}$ and *E2A*$^{+/-}$*Foxo1*$^{f/+}$–ER-Cre bone marrow ($n = 4$ mice per group), and absolute number of pro-B cells, pre-B cells and immature (Imm) B cells in the bone marrow (right). Numbers adjacent to outlined areas (left) indicate percent cells in each. Data are from one representative of two independent experiments with similar results (error bars, s.d.). (**c**) Enhancer activity of H3K4me1 DNA segments with coordinated binding of E2A and EBF1 in the pro-B cell line 22D6 and the T cell lines A12 and 166, assessed for genes closest to the respective enhancer elements (measured from the TSS; horizontal axis; $n = 24$); results were normalized to renilla activity and are presented relative to basal promoter activity (logarithmic scale). Data are representative of at least two independent experiments.

to binding in a genomic locus rather than coordinated DNA-binding activity in putative enhancer elements.

We identified 2,020 putative E2A, EBF1 and Foxo1 target sites that we categorized into eight groups on the basis of the criteria described above (**Fig. 8**). The overlap of E2A and EBF1 targets was highly significant, with over two thirds of EBF targets being potential regulatory targets of E2A ($P = 2.29 \times 10^{-253}$; **Supplementary Fig. 7a**). Many genes occupied by E2A and EBF1 also showed enrichment for Foxo1 binding ($P = 9.88 \times 10^{-324}$; **Supplementary Fig. 7b**). In terms of gene function, as expected, we found that the network containing all potential target genes was substantially enriched for genes encoding proteins involved in lymphocyte development (**Supplementary Fig. 7c**).

The analysis described above raised the question of to what extent the regions identified indeed showed enhancer activity. First we confirmed several of these binding sites by ChIP followed by quantitative PCR (**Supplementary Fig. 8**). Next, we inserted the putative regulatory elements identified by the network of several loci into a luciferase reporter and assayed transcriptional activity (**Supplementary Fig. 9a**). As expected, given the presence of H3K4me1 as well as occupancy by E2A, EBF1 and/or Foxo1, the regulatory elements showed high reporter activity in pro-B cells (**Supplementary Fig. 9b**). Moreover, most elements increased transcription to a much higher degree in pro-B cells than in T cell lines (**Supplementary Fig. 9b**). In sum, we have used a computational approach to separate the pro-B cell genome into clusters of loci characterized by occupancy by E2A, EBF and Foxo1. From this analysis, we have generated a global network that we propose orchestrates B cell fate (**Fig. 8**).

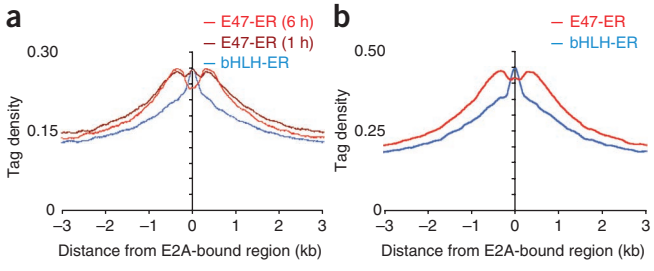

Figure 7 Binding of E47 to DNA alters the pattern of H3K4 monomethylation. (**a**) Distribution of H3K4me1 across E2A occupancy in E2A-deficient pre-pro-B cells transduced for 21 h with E47-ER or bHLH-ER (control), then incubated for 1 or 6 h with tamoxifen to induce E47 activity, crosslinked with formaldehyde, immunoprecipitated with anti-E2A or anti-H3K4me1 and analyzed by ChIP-Seq; results are presented as individual immunoprecipitated mapped reads (or tags) per base pair. (**b**) ChIP-Seq analysis of E2A-deficient A12 cells transduced as described in **a**. Data are from two independent experiments.

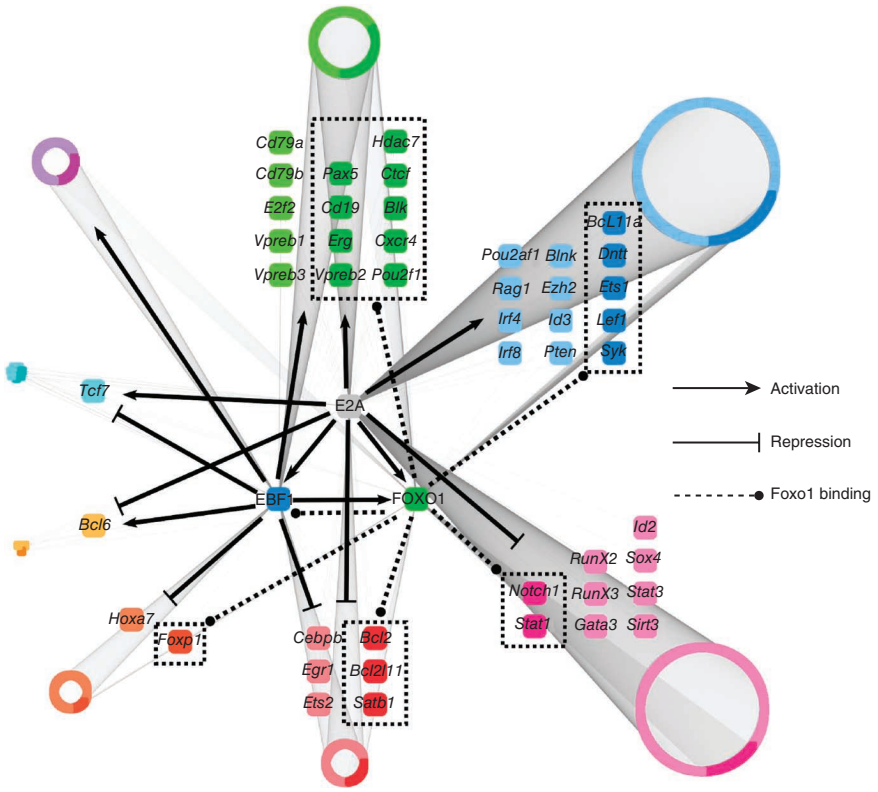

Figure 8 Regulatory network that links the activities of an ensemble of transcriptional regulators, signaling components and survival factors in developing B cells into a common pathway. Transcriptional regulatory targets of E2A and EBF1, as well as genes bound by Foxo1, identified by integrative genome-wide analysis of protein-DNA binding, are separated into eight groups on the basis of positive or negative changes change in transcript amount in pro-B cells relative to that in E2A-deficient pre-pro-B cells or EBF1-deficient pre-pro-B cells. Genes with critical roles in B cell commitment or potential regulators are indicated; all other genes are grouped into rings (ring size indicates the relative number of members). Darker colors indicate loci with Foxo1 occupancy.

DISCUSSION

Regulatory networks have been proposed that link the activities of E2A, EBF1, PU.1 and Pax5 (refs. 26,27). Here we have extended that analysis by genome-wide screens and have generated a global network in which we separated the genome into clusters of genes on the basis of distinct sets of *cis*-regulatory sequences. The network showed links among E2A, EBF1 and Pax5, consistent with published observations[17]. The network also included Foxo1, which until now was not directly connected to other factors in the context of early B cell development. On the basis of these as well as published observations[5,6,9,10,16], we propose that E2A activates expression of *EBF1* and *FOXO1*. E2A and EBF1 and possibly Foxo1, in conjunction with IRF4, IRF8 and PU.1, then act in concert to induce *Pax5* transcription and promote the B cell fate.

The linkage involving E2A and Foxo1 is of interest, as these molecules connect E2A and the PI(3)K cascade into a common module. Published studies have suggested that the PI(3)K–PI(3)K antagonist PTEN–kinase Akt–Foxo1 pathway and E-proteins are linked. Depletion of the long-term hematopoietic stem cell and common lymphoid progenitor compartments has been observed in both E2A-deficient mice and mice triply deficient in Foxo1, Foxo3 and Foxo4 (refs. 28–31). The roles of E2A and Foxo1 in early B cell development also overlap[13,32]. In the immature B cell compartment, both PI(3)K signaling and E2A have critical roles in receptor editing[33,34]. The marginal versus follicular zone B cell fate 'decision' is also modulated by the PI(3)K and E2A pathways[34,35]. Class-switch recombination is considerably perturbed in both E2A-deficient activated B cells and PTEN-deficient activated B cells[36,37]. In thymocyte development, both E2A and PTEN enforce the pre–T cell antigen receptor checkpoint[38,39]. Finally, PTEN, Foxo and E2A act to prevent the development of T cell lymphoma[39–42]. The data described here now provide a mechanism that underpins the linkage between the E2A and Foxo1 proteins: E2A acts immediately upstream of Foxo1, and E2A and Foxo1 are wired together by a wide spectrum of *cis*-regulatory sequences.

The analyses described here also showed that E2A, Foxo1 and EBF1 bound to a large subset of genes involved in pre–B cell antigen receptor signaling. Additionally, *LEF1*, *ETS1* and *Bcl11a* are potential targets of E2A and Foxo1, which links the key regulators encoded by these genes into a common framework. Of particular interest is LEF, which mediates Wnt signaling and modulates B cell proliferation[43]. Thus, these data raise the possibility that E2A, PI(3)K and Wnt signaling are connected by LEF to coordinate cell growth and developmental progression. As expected, many genes seemed to be repressed during the developmental transition from the pre-pro-B cell stage to the pro-B cell stage. Consistent with published observations, these included genes expressed in alternative lymphoid cell lineages, such as *Cebpb, Notch1, Runx2, Runx3* and *Gata3* (refs. 6,16).

The network included a small proportion of loci such as *Tcf7* with greater transcript abundance in pro-B cells than in E2A-deficient pre-pro-B cells, whereas we observed the reverse pattern in EBF1-deficient pre-pro-B cells. It remains to be determined whether E2A and/or EBF1 indeed modulate the expression of loci assigned to such clusters. Instead of comparing transcript abundance in wild-type pro-B cells with that in E2A- or EBF1-deficient pre-pro-B cells, future studies should ablate expression of E2A, EBF1 and Foxo1 in pro-B cells.

In sum, we have generated here a global network that involves the transcription factors E2A, EBF1 and Foxo1. We suggest that this network of transcription factors orchestrates B cell fate. The global network described here has confirmed previous connections[26] and has identified previously unknown links and participants. The network presented here is merely a starting point, as we did not include key regulators such as Pax5 and PU.1 in our analysis. Ultimately, this approach can be applied to other stages of B cell development and other developmental pathways. Global networks that reflect distinct stages of a developmental pathway provide insight into how enhancers and *cis*-regulatory sequences change in developing cells to induce lineage- and stage-specific programs of gene expression. For example, in the context of the B cell lineage, it will be useful to determine how global networks change in response to pre–B cell antigen receptor signaling, self-reactivity and invading pathogens. Ultimately, it will be possible to describe the entire B lineage developmental pathway in terms of global networks using the strategy described here.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/natureimmunology/.

Accession codes. UCSD-Nature Signaling Gateway (http://www.signaling-gateway.org): A000804, A000806, A000809, A000403, A000944 and A000523; GEO: microarray data, GSE21978 and GSE21512.

Note: Supplementary information is available on the Nature Immunology website.

ACKNOWLEDGMENTS

We thank G. Hardiman, C. Ludka, L. Edsall and Z. Ye for help with Solexa DNA sequencing; R. DePinho (Harvard Medical School) for Foxo1-deficient mice; J. Sprague and R. Sasik for microarray analysis; and members of the Murre laboratory for comments on the manuscript. Supported by the National Institutes of Health (1F32CA130276 to Y.C.L., P01DK074868 to C.B., F32HL083752 to S.H., CA52599 to C.K.G., AI05466 to J.H. and CA054198-20 to C.M.) and the National Science Foundation (IIS-0803937 to T.I. and BIOGEM DK063491 to the University of California, San Diego Core Facility).

AUTHOR CONTRIBUTIONS

Y.C.L. designed and did experiments, analyzed data and wrote the manuscript; S.J. and C.B. wrote programs and analyzed data; S.H. did CTCF ChIP-Seq and monomethylation of H3K4 in RAG-deficient pro-B cells; J.H. generated EBF-deficient pre-pro-B cells; M.S. provided anti-EBF; E.W. and R.M. analyzed E2A-Foxo1–deficient mice; C.A.E. did ChIP-Seq experiments during the initial phase of the study; J.D. and T.I. applied computational approaches to generate a global network; C.K.G. analyzed data and edited the manuscript; and C.M. designed experiments, analyzed data and wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/natureimmunology/.
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/.

1. Cosgrove, M.S. Histone proteomics and the epigenetic regulation of nucleosome mobility. *Expert Rev. Proteomics* **4**, 465–478 (2007).
2. Ruthenburg, A.J., Li, H., Patel, D.J. & Allis, C.D. Multivalent engagement of chromatin modifications by linked binding modules. *Nat. Rev. Mol. Cell Biol.* **8**, 983–994 (2007).
3. Pokholok, D.K. *et al.* Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell* **26**, 517–527 (2005).
4. Heintzman, N.D. *et al.* Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat. Genet.* **39**, 311–318 (2007).
5. Kee, B.L. & Murre, C. Induction of EBF and multiple B lineage genes by the helix-loop-helix protein E12. *J. Exp. Med.* **188**, 699–713 (1998).
6. Ikawa, T., Kawamoto, H., Wright, L.Y. & Murre, C. Long-term cultured E2A-deficient hematopoietic progenitor cells are pluripotent. *Immunity* **20**, 349–360 (2004).
7. Inlay, M.A. *et al.* Ly6d marks the earliest stage of B-cell specification and identifies the branchpoint between B-cell and T-cell development. *Genes Dev.* **23**, 2376–2381 (2009).
8. Beck, K., Peak, M.M., Ota, T., Nemazee, D. & Murre, C. Distinct roles for E12 and E47 in B cell specification and the sequential rearrangement of immunoglobulin light chain loci. *J. Exp. Med.* **206**, 2271–2284 (2009).
9. Decker, T. *et al.* Stepwise activation of enhancer and promoter regions of the B cell commitment gene *Pax5* in early lymphopoiesis. *Immunity* **30**, 508–520 (2009).
10. Roessler, S. *et al.* Distinct promoters mediate the regulation of Ebf1 gene expression by interleukin-7 and Pax5. *Mol. Cell. Biol.* **2**, 579–594 (2007).
11. O'Riordan, M. & Grosschedl, R. Coordinate regulation of B cell differentiation by the transcription factors E2A and EBF. *Immunity* **11**, 21–31 (1999).
12. Liu, P. *et al.* Bcl11a is essential for normal lymphoid development. *Nat. Immunol.* **4**, 525–531 (2003).
13. Dengler, H.S. *et al.* Distinct functions for the transcription factor Foxo1 at various stages of B cell differentiation. *Nat. Immunol.* **9**, 1388–1398 (2008).
14. Amin, R.H. & Schlissel, M.S. Foxo1 directly regulates the transcription of recombination-activating genes during B cell development. *Nat. Immunol.* **9**, 396–404 (2008).
15. Nutt, S.L., Heavey, B., Rolink, A.G. & Busslinger, M. Commitment to the B-lymphoid lineage depends on the transcription factor Pax5. *Nature* **401**, 556–572 (1999).
16. Pongubala, J.M. *et al.* Transcription factor EBF restricts alternative lineage options and promotes B cell fate commitment independently of Pax5. *Nat. Immunol.* **9**, 203–215 (2008).
17. Murre, C. Developmental trajectories in early hematopoiesis. *Genes Dev.* **15**, 2366–2370 (2009).
18. Cui, K. *et al.* Chromatin signatures in multipotent human hematopoietic stem cells indicate the fate of bivalent genes during differentiation. *Cell Stem Cell* **4**, 80–93 (2009).
19. Visel, A. *et al.* ChIP-Seq accurately predicts tissue-specific activity of enhancers. *Nature* **12**, 854–858 (2009).
20. Heintzman, N. *et al.* Histone modifications at human enhancers reflect global cell-type specific gene expression. *Nature* **459**, 108–112 (2009).
21. Agata, Y. *et al.* Regulation of T cell receptor beta gene rearrangements and allelic exclusion by the helix-loop-helix protein, E47. *Immunity* **27**, 871–884 (2007).
22. Barski, A. *et al.* High resolution profiling of histone methylations in the human genome. *Cell* **129**, 823–837 (2007).
23. Ward, J.H. Hierarchical grouping to optimize an objective function. *J. Am. Stat. Assoc.* **58**, 236–242 (1963).
24. Shannon, P. *et al.* Cytoscape a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
25. Maere, S., Heymans, K. & Kuiper, M. Bingo: a Cytoscape plugin to assess overrepresentation of Gene Ontology categories in biological networks. *Bioinformatics* **21**, 3448–3449 (2005).
26. Medina, K.L. *et al.* Assembling a gene regulatory network for specification of the B cell fate. *Dev. Cell* **7**, 607–617 (2004).
27. Ikawa, T., Kawamoto, H., Goldrath, A.W. & Murre, C. E proteins and Notch signaling cooperate T lineage specific progression and commitment. *J. Exp. Med.* **15**, 1329–1342 (2006).
28. Tothova, Z. *et al.* FoxOs are critical mediators of hematopoietic stem cell resistance to physiologic oxidative stress. *Cell* **128**, 325–339 (2007).
29. Yang, Q. *et al.* E47 controls the developmental integrity and cell cycle quiescence of multipotential hematopoietic progenitors. *J. Immunol.* **181**, 5885–5894 (2008).
30. Dias, S., Mansson, R., Gurbuxani, S., Sigvardsson, M. & Kee, B.L. E2A proteins promote development of lymphoid-primed multipotent progenitors. *Immunity* **29**, 217–227 (2008).
31. Semerad, C.L., Mercer, E.M., Inlay, M.A., Weissman, I.L. & Murre, C. E2A proteins maintain the hematopoietic stem cell pool and promote the maturation of myelolymphid and myeloerythorid progenitors. *Proc. Natl. Acad. USA* **106**, 1930–1935 (2009).
32. Bain, G. *et al.* Both E12 and E47 allow commitment to the B cell lineage. *Immunity* **6**, 145–154 (1997b).
33. Llorian, M., Stamataki, Z., Hill, S., Turner, M. & Martensson, I.L. The PI3K p110δ is required for down-regulation of RAG expression in immature B cells. *J. Immunol.* **178**, 1981–1985 (2007).
34. Quong, M.W. *et al.* Receptor editing and marginal zone B cell development are regulated by the helix-loop-helix protein, E2A. *J. Exp. Med.* **199**, 1101–1112 (2004).
35. Anzelon, A.N., Wu, H. & Rickert, R.C. *Pten* inactivation alters peripheral B lymphocyte fate and reconstitutes CD19 function. *Nat. Immunol.* **4**, 287–294 (2003).
36. Sayegh, C.E., Quong, M.W., Agata, Y. & Murre, C. E-proteins directly regulate expression of activation-induced deaminase in mature B cells. *Nat. Immunol.* **4**, 586–593 (2003).
37. Suzuki, A. *et al.* Critical roles of Pten in B cell homeostasis and immunoglobulin class switch recombination. *J. Exp. Med.* **197**, 657–667 (2003).
38. Engel, I. & Murre, C. E2A proteins enforce a proliferation checkpoint in developing thymocytes. *EMBO J.* **23**, 202–211 (2004).
39. Hagenbeek, T.J. *et al.* The loss of PTEN allows TCR αβ lineage thymocytes to bypass IL-7 and Pre-TCR-mediated signaling. *J. Exp. Med.* **200**, 883–894 (2004).
40. Engel, I. & Murre, C. Disruption of pre-TCR expression accelerates lymphomagenesis in E2A-deficient mice. *Proc. Natl. Acad. Sci. USA* **99**, 11322–11327 (2002).
41. Paik, J.H. *et al.* FoxOs are lineage-restricted redundant tumor suppressors and regulate endothelial cell homeostasis. *Cell* **28**, 309–323 (2007).
42. Bain, G. *et al.* E2A deficiency leads to abnormalities in α/β T-cell development and to rapid development of T-cell lymphomas. *Mol. Cell. Biol.* **17**, 4782–4791 (1997a).
43. Reya, T.O., Riordan, M., Okamura, R., Devaney, E., Willert, K., Nusse, R. & Grosschedl, R. Wnt signaling regulates B lymphocyte proliferation through a LEF-1 dependent mechanism. *Immunity* **13**, 15–24 (2000).

ONLINE METHODS

Mice. C57BL/6, Foxo1-deficient and RAG-1-deficient mice were maintained in ventilated cages. Animal studies were approved by the Institutional Animal Care and Use Committee of the University of California, San Diego.

Culture conditions, viral infection, and luciferase reporter assay. *Ebf1*$^{-/-}$ hematopoietic progenitors were isolated and grown in the presence of interleukin 7, stem cell factor and the hematopoietic growth factor Flt3L as described[6]. RAG-1-deficient pro-B cells were grown in the presence of Opti-MEM with 10% (vol/vol) FCS, 2% (vol/vol) penicillin-streptomycin and glutamine, 50 μM β-mercaptoethanol plus interleukin 7 and stem cell factor. E2A-deficient pre-pro-B cells were grown in conditions similar to those used for EBF1-deficient cells. Retroviral supernatants containing E47-ER or the bHLH-ER control were prepared and then transduced into E2A-deficient cell lines as described[6]. E47-ER and bHLH-ER were activated for 1 h or 6 h with 1 μM tamoxifen and transduced cells were isolated by magnetic selection (Miltenyi) for coexpressed human CD25. For transfection studies, reporter constructs were transiently transfected into the pro-B cell line 22D6, the E2A-deficient T cell line A12 or the p53-deficient T cell line 166 with the *Trans*IT-LT1 reagent (Mirus). Both 22D6 and A12 cells were grown in the presence of RPMI medium with 10% (vol/vol) FCS, 2% (vol/vol) penicillin-streptomycin-glutamine and 50 μM β-mercaptoethanol; 166 cells were grown in the presence of Opti-MEM with 10% (vol/vol) FCS, 2% (vol/vol) penicillin-streptomycin-glutamine and 50 μM β-mercaptoethanol. The cloning of luciferase constructs has been described[36]. Dual-luciferase reporter assays were done as described by the manufacturer (Promega). Primers for generating constructs, which yielded putative enhancers 500–600 bp in length, are in **Supplementary Table 6**. Each construct was verified by DNA sequencing.

Profiling of expression arrays. Total RNA derived from E2A-deficient or EBF-deficient pre-pro-B cells and RAG-1-deficient pro-B cells was purified with an RNeasy kit (Qiagen). Samples (250 ng RNA) were amplified and labeled with the Quick AMP Labeling kit (Agilent) and then hybridized to Whole Mouse Genome Oligo Microarray (G4122A) as described by the manufacturer (Agilent). Images were quantified with Agilent's Feature Extraction software.

ChIP-Seq. Chromatin was immunoprecipitated as described[21]. Cells were fixed for 5–20 min at 21 °C with 1% (wt/vol) formaldehyde, then were lysed and sonicated. Sonicated chromatin was immunoprecipitated with 10 μg anti-E2A (sc-349; Santa Cruz Biotechnology), anti-EBF1 (H00001879; Abnova), anti-Foxo1 (sc-11350; Santa Cruz Biotechnology) or anti-CTCF (07-0729; Millipore). Anti-H3K4me1 (ab8895), anti-H3K4me2 (ab7766) and anti-H3K4me3 (ab8580) were from Abcam. Anti-H3K27me3 (07-449) and antibody to acetylated H3 (06-599) were from Millipore. Samples were washed, then bound chromatin was eluted and incubated overnight at 65 °C for reversal of crosslinking. Samples were then treated with RNase A and proteinase K and were purified with a PCR Purification kit (Qiagen). Immunoprecipitated chromatin was ligated to an adaptor and amplified by PCR according to the manufacturer's protocol (Illumina). Amplified ChIP DNA was selected by size (150–250 bp) by 8% PAGE or by 2% agarose gel electrophoresis (200–300 bp) and was sequenced with a Genome Analyzer I or II (Illumina). The flow cell was sequenced for 36 cycles to generate 'reads' of 23–25 bp and was aligned to the mm8 assembly (National Center for Biotechnology Information build 36) with the ELAND alignment tool, which allows two-nucleotide mismatches. Tags that mapped to unique DNA sequences were analyzed further (**Supplementary Table 7**). Visualization was achieved by the generation of custom tracks for the University of California at Santa Cruz Genome Browser.

DNA sequence analysis. Data were analyzed with HOMER software. ChIP-Seq-bound elements were identified with HOMER. The position of each tag 3′ of its position was adjusted by 75 bp. One tag from each unique position was examined to eliminate peaks generated by clonal amplification. Binding sites were identified by searching for groups of tags located in a sliding 200-bp window. Adjacent clusters were required to be separated from each other by at least 500 bp. The threshold for the number of tags that generate a bound element was selected for a false-discovery rate of 0.001. Additionally, required peaks were to have at least fourfold more tags (normalized versus total number) than input control samples had. They were also required to have fourfold more tags relative to the local background region in a 10-kb region to avoid identification of DNA segments containing genomic duplications or nonlocalized occupancy. Peaks were associated with gene products by identification of the nearest National Center for Biotechnology Information Reference Sequence collection TSS. For the identification of methylated regions in H3K4, the peak-finding procedure was modified. The peak region was required to show fourfold more tags (relative to total tag count) than controls had. Additionally, DNA-bound regions were required to contain fourfold more tags in a 1-kb region than the flanking 10-kb DNA segments had. For enrichment for occupancy of H3K4me1 located in putative enhancer regions, coordinated H3K4me1 and H3K4me3 regions (in a 3-kb region) were eliminated. In addition, H3K4me1 regions located within 3 kb of an annotated TSS were removed from further analysis.

Flow cytometry. Cre-mediated deletion of *lox*P-flanked alleles was induced in mice 9–10 weeks old by intraperitoneal injection of 1 mg tamoxifen (Sigma) emulsified in 200 μl sunflower seed oil (Sigma) for 5 consecutive days. At 10 d after the start of injections, bone marrow was collected and Fc fragments were blocked with anti-CD16-CD32 (2.5G2), then cells were stained with fluorescein isothiocyanate–conjugated anti-CD11B (M1/70), anti-Gr-1 (RB6-8C5), anti-Ter119 (Ter119), anti-NK1.1 (PK136) and anti-CD3 (145-2c11); Alexa Fluor 700–conjugated anti-B220 (RA3-6B2); phycoerythrin–cyanine 5.5–conjugated anti-CD19 (ID3); allophycocyanin-conjugated anti-c-Kit (2B8); Pacific blue–conjugated immunoglobulin D (11/26); biotin-conjugated immunoglobulin M (11/41; visualized with streptavidin–quantum dot 655); and propidium iodine (all antibodies from eBioscience except QD655 (Invitrogen)). An LSR II (BD Bioscience) was used for analytical flow cytometry and data were analyzed with FlowJo software (TreeStar).

Network construction. Input gene expression and ChIP-Seq data were preprocessed with in-house scripts of the Python programming language. The network was prepared and analyzed with the Cytoscape platform[24]. Gene-ontology enrichment was analyzed with the BiNGO (Biological Network Gene Ontology) plugin[25]. The significance of overlap between target sets was calculated by Fisher's exact test. The R environment of the R Project for Statistical Computing was used for statistical analysis.

doi:10.1038/ni.1891