Cell

Interactome Maps of Mouse Gene Regulatory Domains Reveal Basic Principles of Transcriptional Regulation

Kyong-Rim Kieffer-Kwon,[1,12] Zhonghui Tang,[2,12] Ewy Mathe,[1,12] Jason Qian,[1,12] Myong-Hee Sung,[3,12] Guoliang Li,[2] Wolfgang Resch,[1] Songjoon Baek,[3] Nathanael Pruett,[1] Lars Grøntved,[3] Laura Vian,[1] Steevenson Nelson,[1] Hossein Zare,[4] Ofir Hakim,[5] Deepak Reyon,[6,7] Arito Yamane,[1] Hirotaka Nakahashi,[1] Alexander L. Kovalchuk,[8] Jizhong Zou,[9] J. Keith Joung,[6,7] Vittorio Sartorelli,[4] Chia-Lin Wei,[10] Xiaoan Ruan,[2] Gordon L. Hager,[3,12] Yijun Ruan,[2,12] and Rafael Casellas[1,11,12,*]

[1]Genomics and Immunity, NIAMS, National Institutes of Health, Bethesda, MD 20892, USA
[2]The Jackson Laboratory for Genomic Medicine, and Department of Genetic and Development Biology, University of Connecticut, 400 Farmington, CT 06030, USA
[3]Laboratory of Receptor Biology and Gene Expression, NCI, National Institutes of Health, Bethesda, MD 20892, USA
[4]Laboratory of Muscle Stem Cells and Gene Regulation, NIAMS, National Institutes of Health, Bethesda, MD 20892, USA
[5]Bar-Ilan University, Ramat-Gan 5290002, Israel
[6]Molecular Pathology Unit, Center for Computational and Integrative Biology, and Center for Cancer Research, Massachusetts General Hospital, Charlestown, MA 02129, USA
[7]Department of Pathology, Harvard Medical School, Boston, MA 02115 USA
[8]Laboratory of Immunogenetics, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Rockville, MD 20852, USA
[9]Laboratory of Stem Cell Biology, NIAMS, National Institutes of Health, Bethesda, MD 20892, USA
[10]DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
[11]Center of Cancer Research, NCI, National Institutes of Health, Bethesda, MD 20892, USA
[12]These authors contributed equally to this work
*Correspondence: rafael.casellas@nih.gov
http://dx.doi.org/10.1016/j.cell.2013.11.039

SUMMARY

A key finding of the ENCODE project is that the enhancer landscape of mammalian cells undergoes marked alterations during ontogeny. However, the nature and extent of these changes are unclear. As part of the NIH Mouse Regulome Project, we here combined DNaseI hypersensitivity, ChIP-seq, and ChIA-PET technologies to map the promoter-enhancer interactomes of pluripotent ES cells and differentiated B lymphocytes. We confirm that enhancer usage varies widely across tissues. Unexpectedly, we find that this feature extends to broadly transcribed genes, including *Myc* and *Pim1* cell-cycle regulators, which associate with an entirely different set of enhancers in ES and B cells. By means of high-resolution CpG methylomes, genome editing, and digital footprinting, we show that these enhancers recruit lineage-determining factors. Furthermore, we demonstrate that the turning on and off of enhancers during development correlates with promoter activity. We propose that organisms rely on a dynamic enhancer landscape to control basic cellular functions in a tissue-specific manner.

INTRODUCTION

Gene expression during development is orchestrated by promoter sequences and a variety of distal *cis*-regulatory elements. Key among these are enhancers, which associate with promoters to increase the transcriptional output of target genes in a tissue-specific manner (Visel et al., 2009). Enhancers are typically distinguished from nonregulatory DNA by their hypersensitivity to DNaseI digestion (Sabo et al., 2006) and binding of chromatin modifiers. The CBP/p300 acetyltransferase for instance mediates H3K27 acetylation of chromatin at active enhancers (Creyghton et al., 2010). In addition, enhancers display high levels of H3K4 monomethylation (H3K4me1; Buecker and Wysocka, 2012), and a relative depletion of H3K4me3 (Heintzman et al., 2007) and the histone variant H2AZ (Kouzine et al., 2013). Based on these parameters, ~400,000 genomic sites displaying enhancer-like features were recently discovered, spanning nearly 10% of the human genome (ENCODE Project Consortium et al., 2012).

Enhancers control lineage identity by recruiting transcription factors, cofactors, and RNA Polymerase II (PolII) to target genes. They physically interact with promoters resulting in looping out of intervening sequences (Krivega and Dean, 2012), which in some instances can span over 1 Mb of DNA (Nobrega et al., 2003). In contrast to promoters and insulators, which vary little across cell types, the enhancer landscape changes considerably during development (Thurman et al., 2012). This feature predicts that

functional connectivity in mammalian cells (1) must display a high degree of tissue specificity and (2) should closely reflect transcriptome changes during cell differentiation. However, these ideas have not been fully explored because of the difficulty of mapping promoter-enhancer connections during development.

In the absence of direct approaches, enhancers have been typically assigned to "cognate" promoters based on linear proximity or shared chromatin states. This strategy has limitations because enhancers do not always regulate nor share chromatin profiles with the nearest promoter. Alternatively, chromosome conformation capture techniques have been used to explore regulatory interactions at predefined genomic loci. However, the resolution of 3C-based techniques alone is insufficient to map promoter-enhancer connectivity in entire genomes (Xie and Ren, 2013). To overcome this challenge, the ChIA-PET protocol was recently developed (Fullwood et al., 2009). ChIA-PET is a ChIP-based method that captures long-range chromatin interactions involving or mediated by a protein of interest such as estrogen receptor α in adenocarcinoma cells (Fullwood et al., 2009) or RNA PolII in human cell lines (Li et al., 2012).

We here introduce the NIH Mouse Regulome Project, an initiative that seeks to define the 3D interplay of gene regulatory domains in developing mouse primary cells. In this first report we compare pluripotent embryonic stem (ES) cells and differentiated B lymphocytes. By combining ChIA-PET, CpG methylomes, DNaseI hypersensitivity, transcriptomes, digital footprinting, and TALEN-mediated genome editing, our studies reveal the dynamics of the mouse regulome during ontogeny.

RESULTS

A Comprehensive Map of Regulatory Domains and Their Interactions in Mouse Primary Cells

To characterize the mouse regulome in primary cells we first applied DNaseI hypersensitivity (DHS) followed by deep-sequencing to $CD43^-$ B lymphocytes activated in the presence of lipopolysaccharide and interleukin-4 (LPS+IL-4). From two independent experiments (268 million aligned reads), we identified 90,015 high-confidence DNaseI domains in B cells. As expected, DNaseI-seq profiles were highly reproducible between biological replicates (Spearman's $\rho = 0.99$; Figure S1A available online).

To identify DHS sites associated with gene regulatory domains, we next mapped Nipbl, Med12, and p300 by ChIP-seq (Figure 1A). We chose the cohesin-loading factor Nipbl and the Med12 component of mediator because they demarcate enhancers tethered with core promoters (Kagey et al., 2010). Likewise, the transcription regulator p300 occupies a subset of active promoters and enhancers (Chen et al., 2008). There was a considerable, although incomplete, overlap in the recruitment of these factors in B cells (Figures 1A and S1B). Active promoters were thus identified as $Nipbl^+$, $Med12^+$, or $p300^+$ DHS sites that overlapped with ENSEMBL-annotated transcription start sites (TSSs). Based on this strategy we identified 17,004 DHS promoter elements associated with 16,931 genes, 49,763 enhancer elements, and an additional 23,248 that did not overlap with either (Figure 1B). Promoters were in general H3K4me1-^{low}H3K4me3^{high} H2AZ^{high}, whereas the opposite signature demarcated enhancer domains (Figure 1A; Kouzine et al., 2013). DHS analysis of mouse ES cells uncovered a similar number of DHS promoters (16,771) and an increased number of enhancers (62,766, Figure 1B).

To directly map the promoter-enhancer interactome, we applied chromatin interaction analysis by paired-end-tag sequencing (ChIA-PET; Fullwood et al., 2009; Zhang et al., 2012), which combines PolII ChIP with 3C technology (Figures 1C, S2A). We generated two independent B cell ChIA-PET libraries, from which ~15 million reads were uniquely aligned and classified into two separate data sets: 5.7 million reads of PolII chromatin occupancy, and 9.2 million reads clustered into 14,247 high-confidence PolII long-range *cis* interactions or PETs (Figure 1D and Table S1). Both data sets were correlated between replicates (Spearman's $\rho > 0.83$, Figures S2B and S2C).

Attesting to the specificity of ChIA-PET, most PolII long-range interactions (13,070, 92%) were linked to at least one gene regulatory domain (Figure S1C). Furthermore, of 16,931 B cell promoters associated with DHS domains, 6,890 were involved in PolII long-range interactions. In general, these genes were transcribed 2-fold higher ($p < 2 \times 10^{16}$, Figure S1D) and recruited more PolII ($p < 2 \times 10^{16}$, Figure S1E) than nonanchored ones. We also detected 6,813 DHS enhancer domains involved in PolII interactions. Of these, 71% were active ($H3K27Ac^+$), whereas up to 60% of nonanchored ones were poised ($H3K27Ac^-$, Figure S1F). In general, the number of ChIA-PET interactions per regulatory site was proportional to the extent of DNaseI digestion (Figure S1G). Thus, ChIA-PET preferentially detects PolII long-range interactions involving $H3K27Ac^+$ enhancers and transcriptionally active promoters.

As previously shown (Li et al., 2012), PolII interactions fell into four distinct groups: (1) intragenic, connecting promoters to gene bodies; (2) extragenic, connecting promoters to distal regulatory elements; (3) intergenic, tethering promoters from different genes; and (4) enhancer-enhancer interactions (Figure S1H). Examples of these are provided in Figure 1D for the *Atp5j-Mir155* gene locus. Consistent with high expression of *Mir155* in activated B cells (Kuchen et al., 2010), its promoter was associated with 83 long-range interaction tags (Figure 1D, upper). Of these, 70 were extragenic, involving 5 upstream enhancer domains, while 13 were intragenic, connecting the promoter to downstream sequences. An additional 23 PolII long-range interactions interconnected the 5 enhancers upstream of *Mir155*. In contrast to B cells, ES cells actively transcribe *Mrpl39*, *Jam2*, and *Atp5j* but express little *Mir155* mRNA (Figure 1D, lower). Consistent with this, we identified 30 intergenic connections between *Mrpl39*, *Jam2*, and *Atp5j* promoters in ES cells, whereas few connections involved *Mir155* (Figure 1D). As in previous ChIA-PET studies, both direct and indirect interactions were considered in our analysis (Figure S1I).

TALEN-Mediated Validation of Promoter-Enhancer Connectivity

ChIA-PET confirmed established connections between gene regulatory domains. For instance, the pluripotent gene *Sox2* was associated in ES cells with a series of enhancers recently described by 5C studies (Figure S3A; Phillips-Cremins et al., 2013). Likewise, the immunoglobulin heavy chain (*Igh*) 3′Eα

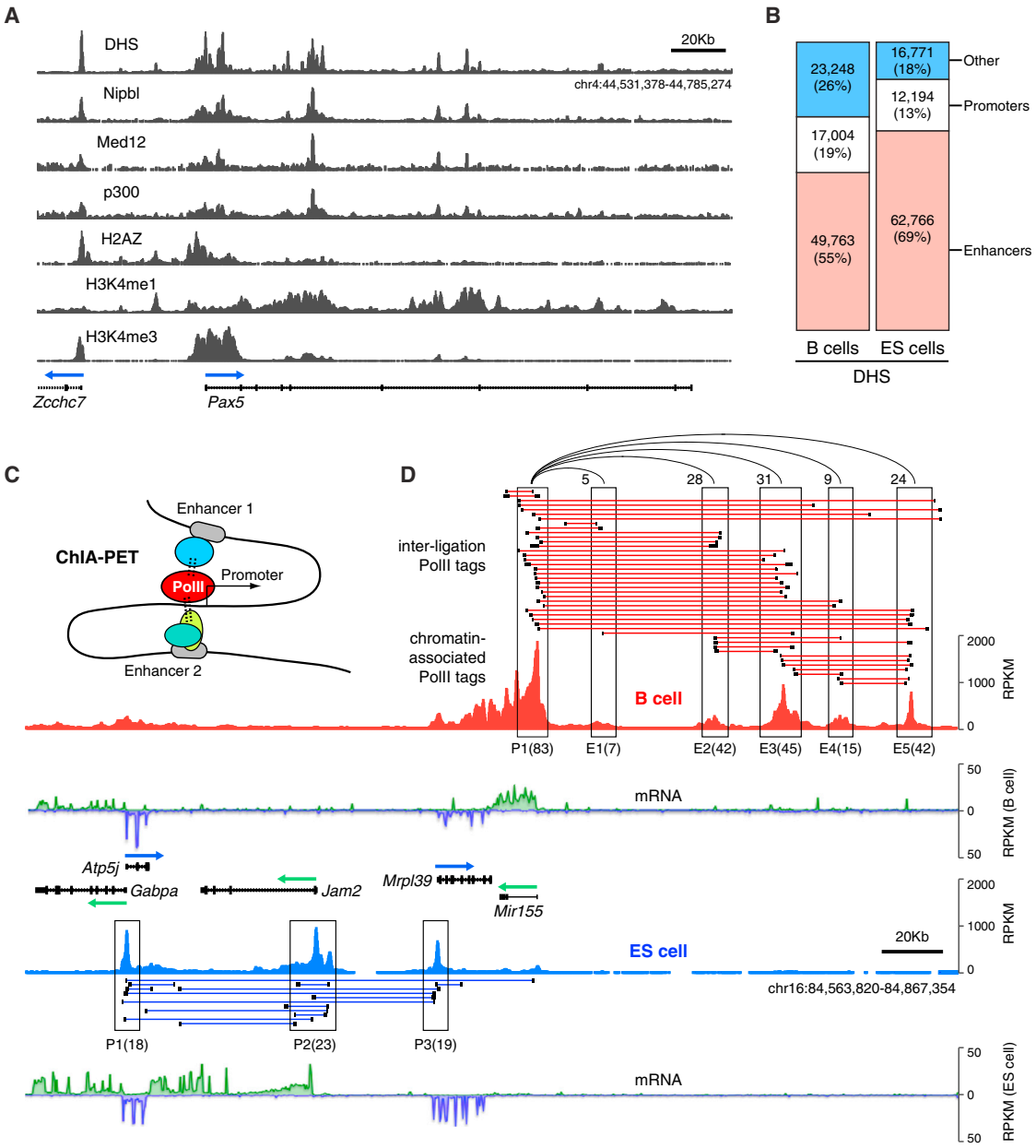

Figure 1. Characterization of Regulatory Domains and Their Interactions in Mouse Primary B and ES Cells

(A) *Pax5* locus in activated B cells displaying DNaseI hypersensitivity (DHS); recruitment of Nipbl, Med12, and p300; and chromatin marks H2AZ, H3K4me1, and H3K4me3.

(B) Bar graphs showing the number of DHS islands in B and ES cells overlapping with promoters (TSS+, white), enhancers (TSS-, Nipbl+, or Med12+, or p300+, red), or nonoverlapping (blue).

(C) The ChIA-PET protocol combines PolII ChIP with conformation capturing techniques to map the interaction of active promoters with gene regulatory domains.

(D) Examples of ChIA-PET clusters at the *Mir155* locus in activated B cells (red connectors) or ES cells (blue connectors). Each connector links two or more long-range interactions (PETs) separated by <500 bps (Figure S2A). ChIP-seq data are represented as reads per kb per million sequences (RPKM). Promoters (P) and enhancers (E) are boxed and the number of total PETs is provided in parenthesis. Interactions between enhancers and *Mir155* are represented by semi-circle connectors. mRNA expression is provided for B and ES cells as RPKM values (+, strand transcription in green; −, strand in blue).

See also Figures S1, S2, and S3, and Table S1.

enhancer was found in spatial proximity to transcribed *Igμ*, *Igγ1*, and *Igε* in B cells only (Figure S3B). We also found evidence of *Igμ-Igγ1* associations, representing either synapses between the recombining genes (Wuerffel et al., 2007) or fully recombined DNA. At the *Igκ* locus, the *VJκ*-5′Eκ domain was connected to the previously characterized 3′Eκ and Ed enhancers (Figure 2A; Liu

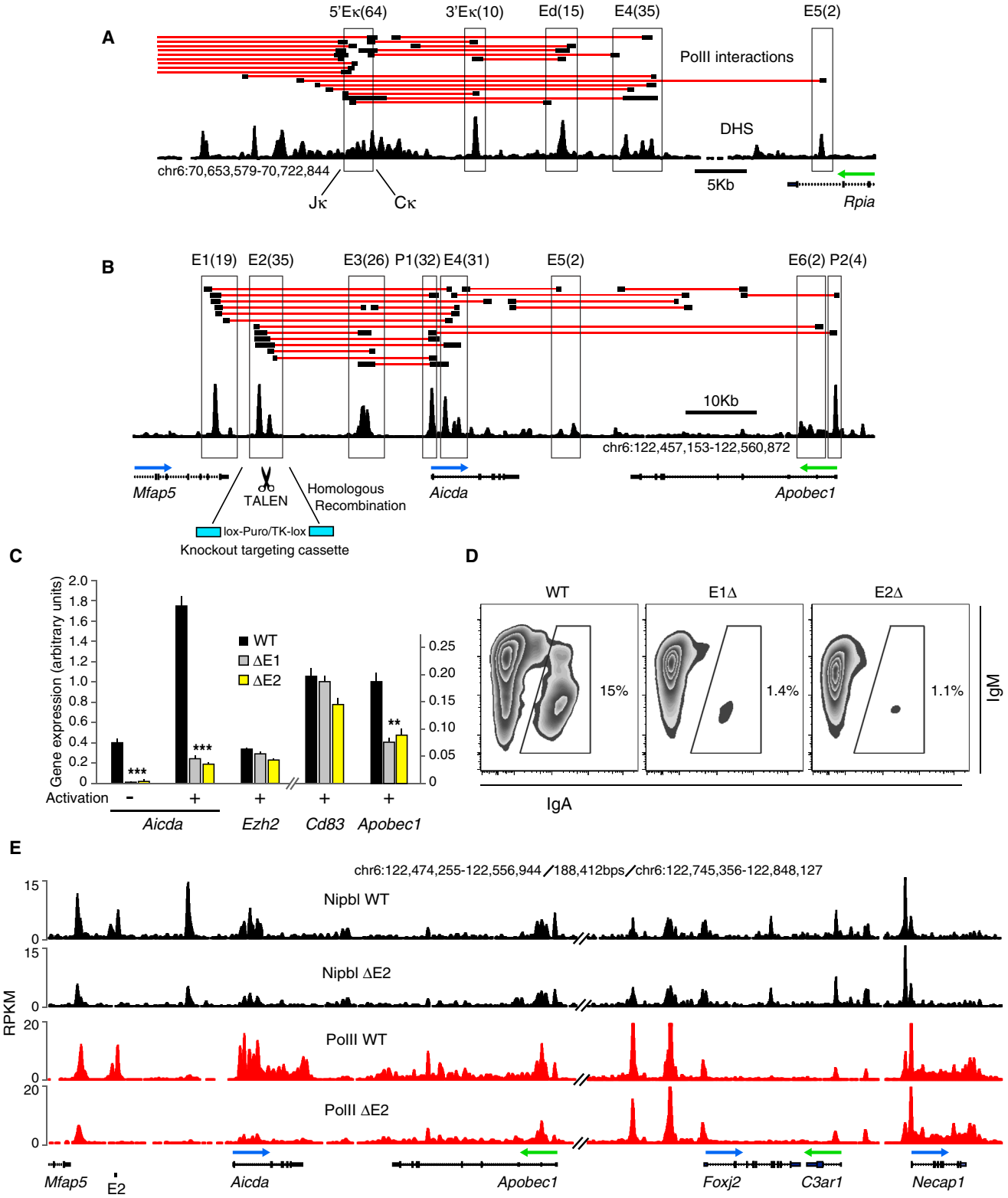

Figure 2. In Vivo Validation of ChIA-PET by Genome Editing

(A) ChIA-PET at the *Igκ* locus identifies previously characterized 5′Eκ, 3′Eκ, and Ed enhancers, as well as new enhancers E4 and E5. Number of PETs associated with each regulatory domain (boxed) are provided in parenthesis. The DHS activated B cell track is also provided (black).

(legend continued on next page)

et al., 2002; Meyer and Neuberger, 1989). Unexpectedly, the analysis uncovered two additional enhancers located 8 kb (E4) and 15 kb (E5) downstream of Ed (Figure 2A).

We also found additional enhancers (E1-E2) associated with the activation induced deaminase (AID) gene *Aicda* (Figure 2B). The three enhancers previously shown to regulate AID transcription in vivo were also linked by PolII long-range interactions in the analysis (E3-E5, Figure 2B; Crouch et al., 2007; Huong et al., 2013; Sayegh et al., 2003). The *Apobec1* promoter and a sixth enhancer located in *Apobec1* intron 2 were also clustered (Figure 2B). To validate ChIA-PET associations, we deleted E1 and E2 in CH12 mouse lymphoma cells. We chose this B cell line because upon activation it transcribes high levels of AID and undergoes efficient *Igμ-Igα* recombination (Nakamura et al., 1996). To facilitate homozygous gene targeting, knockout constructs were cotransfected with enhancer-specific transcription activator-like effector nucleases (TALENs), assembled via a solid-phase high-throughput system (Reyon et al., 2012; Figure S4A). Upon activation, wild-type CH12 cells increased AID mRNA expression ~5-fold and recombined to IgA (15%, Figures 2C and 2D). Deletion of E1 or E2 however markedly reduced AID transcription and IgA expression (Figures 2C and 2D), consistent with the notion that the extent of switching is proportional to AID expression (Takizawa et al., 2008). Transcription of *Apobec1* was also impaired in the mutant cells, whereas noninteracting *Ezh2* and *Cd83* genes were unaffected (Figure 2C). Importantly, $E1^{-/-}$ and $E2^{-/-}$ cells displayed an overall reduction in Nipbl and PolII occupancy at all regulatory domains within the *Aicda* locus, including the *Apobec1* promoter (Figure 2E and S4B). In contrast, this effect was not observed at the *Foxj2-Necap1* locus, ~190 kb downstream of *Apobec1* (Figure 2E and S4B). Thus, E1 and E2 regulate *AID* and *Apobec1* transcription by controlling local recruitment of PolII.

To further validate the ChIA-PET results, we targeted additional regulatory elements associated with the *Pou2af1* (OCA-B), and *Cd79a* genes. We uncovered an intronic enhancer (E3) ~15 kb downstream of OCA-B TSS required for transcriptional upregulation upon B cell activation but dispensable for basal transcription in nonstimulated cells (Figure S4C). This activity is consistent with the reported dynamics and signaling requirements of OCA-B expression during B cell differentiation (Casellas et al., 2002; Qin et al., 1998). A similar analysis confirmed the presence of enhancer elements that augment basal *Cd79a* transcription (Figure S4D). Additional gene targeting experiments within the *Pim1* oncogene locus are discussed below (Figure 5). Taken together, these results demonstrate that at least a fraction of PolII long-range interactions, as defined by ChIA-PET, represent functional promoter-enhancer connections in B lymphocytes.

Single- and Higher-Order Gene Clusters in Primary Mouse Cells

Up to 54% of genes recruiting PolII in activated B cells were associated with long-range interactions (6,890 of 12,652). Of these, 1,231 (18%) represented single promoters tethered to at least one enhancer (Figure 3A). These clusters created complex architectures and spanned an average of 78 kb of genomic DNA (Figure S5A). The most elaborate of this group was the *Gpr183* promoter, which was connected either directly or indirectly to 12 enhancers via 76 long-range interactions (Figure 3A). Another example was *Cd83*, which displayed 158 PolII intragenic and extragenic connections (Figure 3B).

Among single-promoter gene clusters, we found examples of the recently dubbed superenhancer domains (Whyte et al., 2013), which were defined based on clustering of gene regulatory domains: e.g., mir290/295 and Sox2 loci (Figures S5B and S3A). However, the vast majority of anchored genes (5,606, 81%) formed higher-order multigene complexes (1,481 B cell clusters, Figure 3C), which could not be easily deduced as interacting based on visual inspection of DHS island distribution. The average span of these clusters was 179 kb (Figure S5A). Its prime example in B cells was the *Rela* cluster in chromosome 19, which was composed of 66 genes and 398 long-range interactions (Figure 3C). Promoters linked by intergenic connections displayed higher PolII density and mRNA synthesis relative to genes from single-promoter clusters or not anchored to other domains (Figure S5C). Furthermore, families of genes coexpressed during ontogeny were overrepresented in the multiple-promoter gene group (see Experimental Procedures). Among these we found the major histocompatibility complex *H2-Mβ* cluster (Fisher exact test $p = 1.3 \times 10^{14}$), the *Hist1h* histone family ($p = 7.7 \times 10^{82}$), and the lymphoid signaling *Gimap* cluster ($p = 7.6 \times 10^{14}$, Figure S5D).

In contrast to promoters, which readily formed higher-order complexes, the vast majority of enhancers (~90%) were linked to a single promoter, and less than 2% of all enhancers were linked to more than two promoters (Table S1). One exception was an enhancer downstream of *Gimap6*, which was directly linked to seven promoters (Figure S5D).

Transcriptional Correlation between lncRNA and Associated Coding Genes

Long noncoding RNAs (lncRNA) are a new class of RNAs believed to play regulatory functions (Batista and Chang, 2013). ChIA-PET identified hundreds of associations between protein-coding and lncRNA genes in multiple-promoter clusters. For instance, lncRNA *E*(ENSMUSG)*85930* is extensively associated with *Clec2d* and to a lesser extent with *Cd69* (Figure 3D). Other examples involving genes key for B cell development included *Ptprcap-E90702*, *Cd81-E59277*, and *Bcl11a-E123592* pairs (Figure S5E). lncRNAs are believed to modulate transcription of

(B) Regulatory map of the *Aicda-Apobec1* locus in activated B cells. Deletion of selected enhancers (E1 and E2) was carried out in CH12 B cells using knockout targeting cassettes (cyan) and TALEN endonucleases.

(C) qPCR analysis of *Aicda*, *Apobec1*, *Ezh2*, and *Cd83* expression in wild-type (WT), and E1, or E2 deleted (Δ) CH12 cells. Data are represented as the mean ± SEM (n = 6). p values were < 0.0001 (*Aicda*), and = 0.008 (*Apobec1*).

(D) Flow cytometry analysis of recombination to IgA in activated WT, ΔE1, or ΔE2 cells.

(E) Nipbl (black) and PolII (red) occupancy at the *Aicda-Apobec1* and *Foxj2-Necap1* loci in WT or ΔE2 cells. The two loci are separated on chromosome 6 by 188 kb. See also Figure S4 and Table S3.

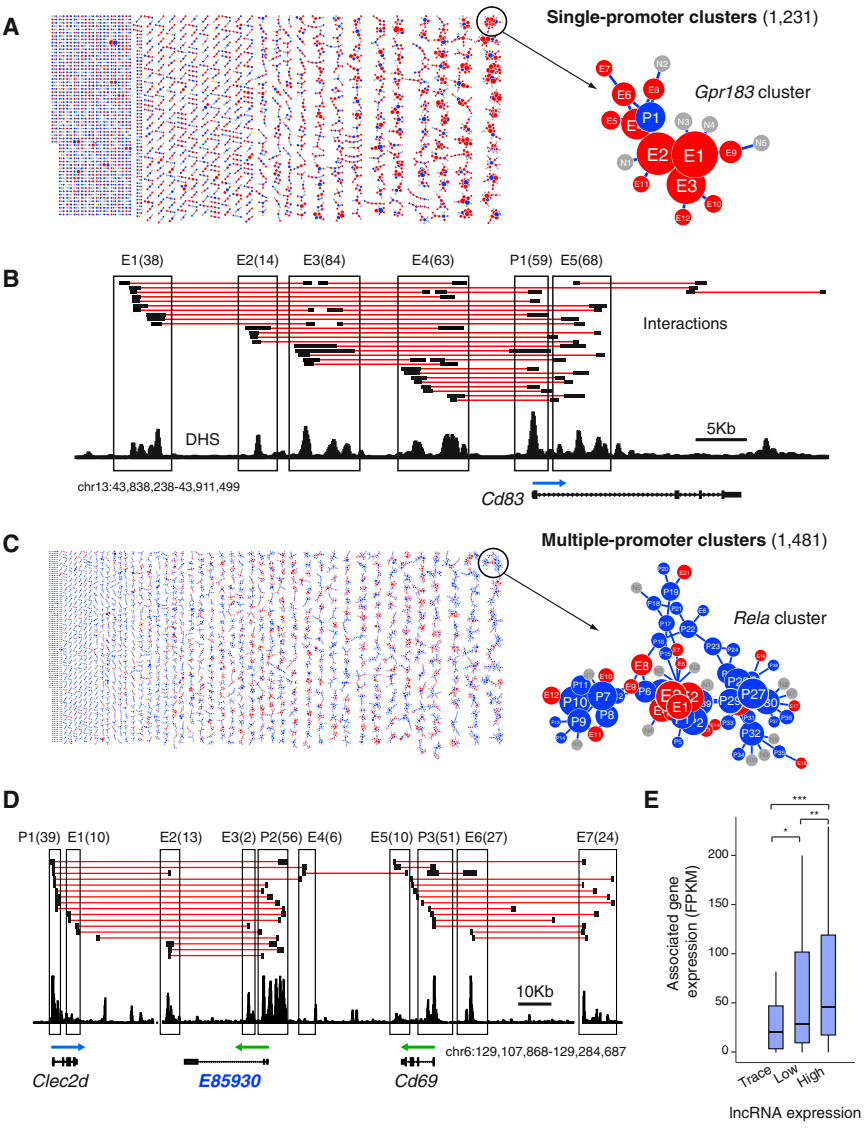

Figure 3. Gene Clusters Identified by ChIA-PET

(A) Single-promoter clusters in activated B cells connecting 1,231 gene promoters to at least one enhancer. Right: e.g., the *Gpr183* promoter (blue circle) is linked to 12 enhancers (red circles) via 76 interactions. PETs anchored outside enhancers are represented with gray circles. Circles are sized according to the absolute number of anchored PETs.

(B) Interactions at the *Cd83* single gene cluster.

(C) Multipromoter clusters (n = 1,481) identified in B cells. The *Rela* cluster display 398 interactions involving 66 genes.

(D) PolII connections between lncRNA *E85930*, *Clec2d*, and *Cd69*. Promoters and enhancers are boxed and number of PETs are provided in parenthesis.

(E) Transcription levels of genes associated with trace (no detectable FPKM), low (<0.9 FPKM), or highly transcribed lncRNAs (≥0.9 FPKM).

See also Figure S5.

neighboring genes by promoting local topological changes in chromatin (Ponting et al., 2009). Consequently, transcription of lncRNAs and their targeted genes is often coordinated (Guil and Esteller, 2012). To test this idea across the genome we measured expression of lncRNAs and their interacting protein-coding genes as defined by ChIA-PET. We found that genes associated with highly abundant lncRNAs were transcribed at higher levels than those associated with lncRNAs detected at low or trace levels ($p < 0.05$, Figure 3E). These findings are consistent with the proposal that transcription of lincRNAs and their targets can be coordinated. Whether the same scenario applies to promoter-promoter clusters not involving lncRNA remains to be determined.

Broadly Expressed Genes Are Linked to Cell-Type-Specific Enhancers

As expected, genes differentially expressed in B and ES cells were linked to tissue-specific regulatory elements. The pluripotent gene *Sox2* for instance was associated with ES-cell-specific enhancers (Figure S3A). Conversely, the B-cell-specific *Cd79b* gene was only anchored to enhancers in the B cell compartment (Figure S6A).

We next turned our attention to genes transcribed in both cell types. Of 6,890 promoters anchored by ChIA-PET in B cells, 4,854 (70%) were also anchored in ES cells (Figure 4A, Venn diagram). As an example, the *Heximl-2* genes were linked to the same downstream enhancer (E1) in B and ES cells (Figure 4B). Surprisingly, most anchored promoters in the two cell types (4,430, 94%) were associated with at least one additional tissue-specific enhancer (Figure 4A, lower pie chart). A striking example was the *Myc* proto-oncogene, which displayed a completely different enhancer landscape in ES and B cells (Figure 4C). In B cells, *Myc* was linked to ten enhancers (E5-E14) located near or downstream of exon 3 of the lncRNA *Pvt1*, whereas in ES cells all enhancers associated with *Myc* (E1-E4) were found upstream of this site (Figure 4C). Other examples included *Tgif1*, *Smad7*, and *Malat1*, which were preferentially linked to ES-cell-specific enhancers, whereas *Swap70*, *Etv5*, and *Pim1* were tethered to a greater number of enhancers in B lymphocytes (Figure 4D).

To explore whether changes in the enhancer landscape impacts transcription of this gene group, we measured their expression by calculating mRNA copy numbers per cell. Genes that turned on or off a single tissue-specific enhancer displayed little or no changes in transcription levels in the two cell types (Figure 4E). However, as genes interacted with two or more additional enhancers their expression was significantly different ($p < 9 \times 10^{10}$, Figure 4E). *Myc*, for instance, was transcribed ~4 times

higher in B cells than in ES cells (Table S2). This observation is consistent with the notion that, in general, transcription levels of a given promoter are commensurate with the number of regulatory domains it is regulated by (Li et al., 2012). On the basis of these findings we conclude that (1) broadly expressed genes can be regulated by cell-type-specific enhancers and (2) the turning on and off of enhancers during ontogeny impacts transcription levels.

Dynamic CpG Methylation of Cell-Type-Specific Regulatory Domains

Cellular differentiation is accompanied by changes in DNA methylation at promoters and distal regulatory domains (Shen et al., 2013; Song et al., 2013; Stadler et al., 2011; Ziller et al., 2013). To explore whether the dynamics of DNA methylation correlate with differential enhancer usage, we applied bisulphite sequencing (Bis-Seq) and generated methylome libraries at single-nucleotide resolution. Bis-Seq of activated B cells provided a total of 148 billion mappable methylome bases. We complemented this data set with Bis-Seq libraries from mouse ES cells (Stadler et al., 2011) and calculated the percentage of CpG methylation at gene regulatory domains. With few exceptions, activated B-cell-specific enhancers were highly methylated in ES cells (>80% of CpGs), whereas enhancers common to both cell types displayed a broad range of CpG methylation levels (Figure 5A). High methylation was also observed at ES cell-specific enhancers in activated lymphocytes, whereas B cell enhancers displayed low methylation levels (Figure 5A). Importantly, the level of CpG methylation was lower at active than at poised enhancers ($p < 3 \times 10^{16}$, Figure 5A). As expected, promoters of silent genes displayed on average higher CpG methylation than active ones ($p < 2.2 \times 10^{16}$, Figure 5A). Methylation levels were also inversely proportional to the extent of ChIA-PET signals (Figure S6B). Thus, enhancers are highly methylated when inactive, but become demethylated during development concomitant with the presence of tissue-specific PolII interactions.

We explore in Figure 5B transcriptional regulation of the *Pim1* oncogene, whose promoter is tethered to an entirely different set of enhancers in activated B cells and ES cells. The analysis shows a direct correlation between CpG demethylation and enhancer usage. For instance, B cell enhancers E2 and E5 and ES cell enhancers E6 and E7 display nearly complete CpG demethylation in a cell-type-specific manner (Figure 5B). To confirm that these enhancers truly promote *Pim1* transcription, we targeted E2 and E6 in CH12 B cells and ES cells, respectively. As measured by qPCR, we found a significant decrease in *Pim1* mRNA levels in the targeted cells, whereas expression of *Brd2*, *Mtch1*, and *Gapdh* was unaffected (Figure 5C). Attempts to delete E2 and E6 in the cell type where they are inactive were unsuccessful (not shown), likely due to the inability of TALENs to target methylated DNA (Bultmann et al., 2012).

In ES cells, B-cell-specific enhancers, including those linked to broadly expressed genes, were hypermethylated (Figure S6C). To examine at which stage during B lymphopoiesis these regulatory elements become demethylated, we generated methylome libraries from bone marrow hematopoietic stem cells (KSL), B lymphoid precursors (CLP), and peripheral G_0 resting B cells. Of 1,518 B-cell-specific enhancers that were linked by long-range interactions and hypomethylated during activation, only 82 (5%) were also hypomethylated in ES cells (Figure 5D). However, in KSL precursors nearly half (714, 47%) of activated B cell enhancers were already demethylated (Figure 5D). This group included *Pim1* enhancers E2, E3, and E5, which displayed nearly identical CpG methylation levels in KSL, CLP, G_0 resting, and cycling B cells (Figure S6D). As KSLs develop into CLP B cell precursors, demethylation was observed in 61% (926) of B cell enhancers. In resting G_0 B cells, this number increased to 80% (1,207, Figure 5D). At this stage of development, the overall mean methylation was comparable between activated B cell enhancers and those functional both in B and ES cells (Figure S6C). Thus, most activated B cell enhancers, including those associated with broadly expressed genes, are demethylated by the time naive lymphocytes migrate from the bone marrow to the periphery. This finding is consistent with the notion that the genome of G_0 lymphocytes is primed for activation and that most genes expressed during the humoral immune response are transcribed at basal levels in the naive compartment (Kouzine et al., 2013; Nie et al., 2012). At the same time, it is important to point out that ~20% of activated B cell enhancers do not become fully demethylated until activation occurs. Among these, we find *Pim1* E4 and *Aicda* E3 and E4 (Figures S6D and S6E).

Digital Genomic Footprinting Characterizes TF Binding in the Mouse Genome

The observation that cell-type-specific enhancers can promote transcription of broadly expressed genes implies that factors driving lineage specification are involved in this regulation. To explore this idea, we sought to comprehensively catalog transcription factor occupancy in mouse B lymphocytes and ES cells. To this end, we took advantage of the fact that transcription factors protect their binding sites from DNaseI cleavage, leaving nucleotide-resolution footprints within DHS islands (Neph et al., 2012). Figure 6A, for instance, shows four DNaseI footprints at the *Pold4* gene promoter in G_0 and cycling B cells. Importantly, these footprints overlap with recognizable binding motifs for transcription factors PU.1, Ebf1, Egr1, and Sp1 (Figure 6A). By applying an established algorithm (Baek et al., 2012), we detected 706,669 high-confidence (FDR < 5%) footprints within 75,917 B cell DHS domains (70% of total DHS). To link these footprints to known transcription factor recognition sequences, we examined all empirically defined DNA binding motifs, compiled by HOMER, UniPROBE, JASPAR, and similar databases. We found a significant enrichment in transcription factor binding motifs within DHS footprints ($p < 1 \times 10^{6}$, Figure S7A). Altogether, we linked 247 distinct transcription factor DNA motifs to 122,505 footprints in B cells (Table S1). In addition, de novo motif discovery yielded 18 new binding sites that did not match known recognition sequences (Table S1). A similar analysis linked 306 DNA motifs to 346,284 footprints in ES cells (Table S1).

Figure 6B shows examples of cleavage profiles for transcription factors Irf8, Sp1, Nrf1, PU.1, and CTCF (an extended view of footprints is provided in Figure S7B). Importantly, ChIP-seq analysis showed a correlation between PU.1 and CTCF occupancy, their DNA binding motifs, and cognate footprints

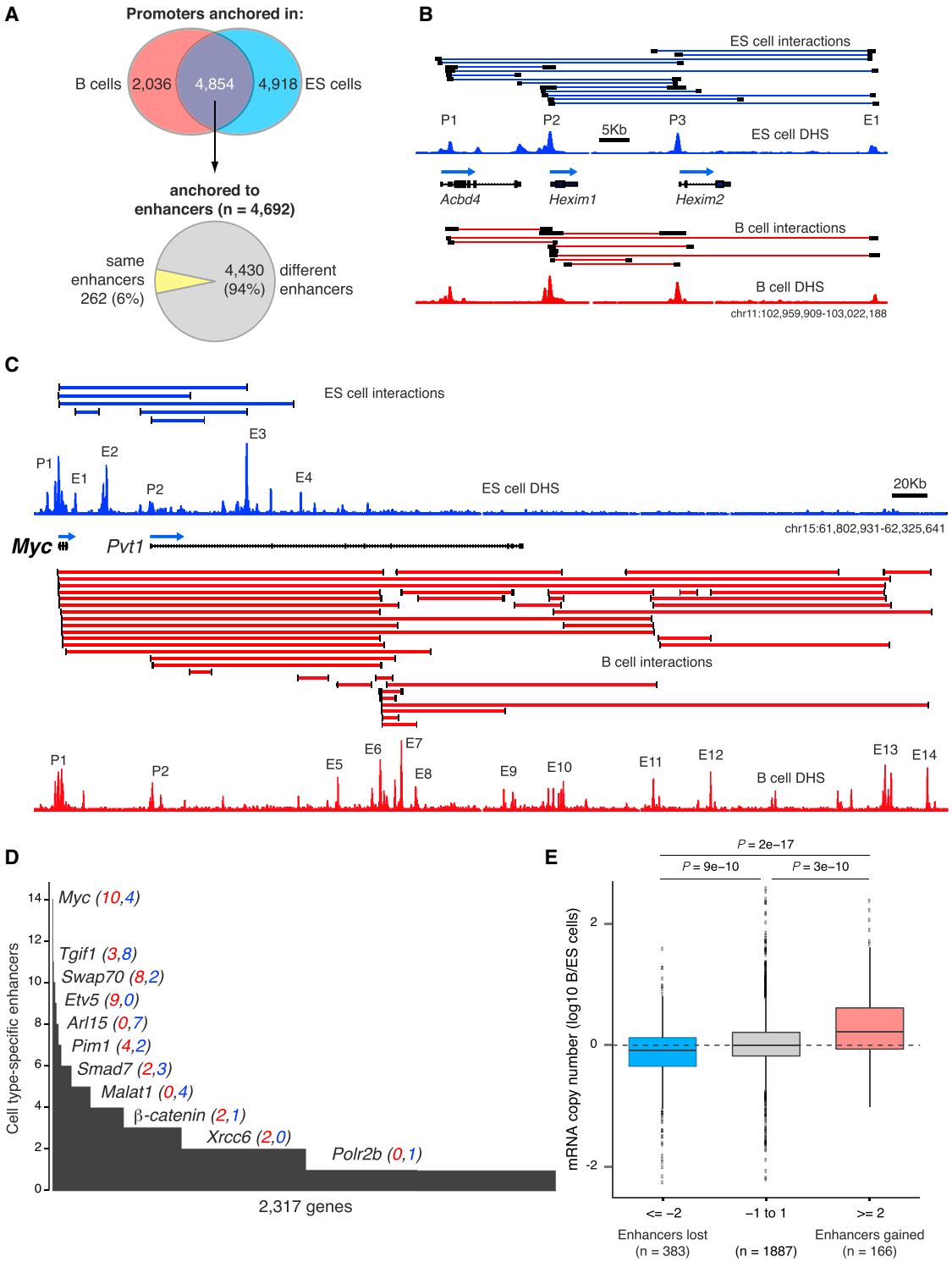

Figure 4. Tissue-Specific Enhancers Contribute to the Transcriptional Regulation of Broadly Expressed Genes

(A) Venn diagram showing the number of ChIA-PET anchored promoters in B cells (left), ES cells (right), or in both cell types (middle). For the latter group, the pie chart below shows the number of promoters linked to the same (yellow) or to at least one cell-type-specific enhancer (gray).

(B) *Acbd4* and *Hexim1/2* gene promoters associate with the same downstream enhancer (E1) in B and ES cells.

(C) The *Myc* oncogene is linked to an entirely different set of enhancers in ES (blue) and B (red) cells. Enhancers are numbered from 1–10 based on proximity to the *Myc* promoter (P1).

(legend continued on next page)

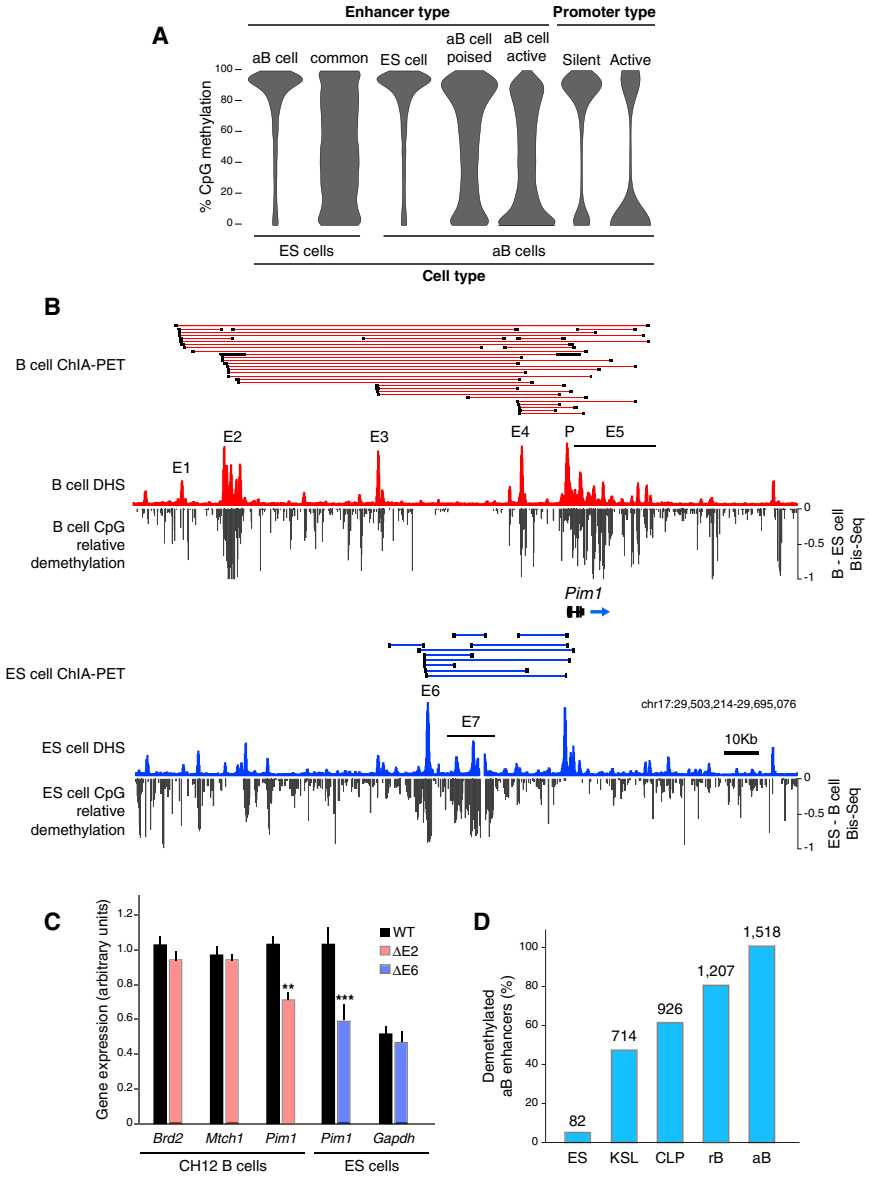

Figure 5. DNA Demethylation Demarcates Differential Enhancer Usage

(A) CpG methylation at regulatory elements in B and ES cells.

(B) DNA demethylation at the *Pim1* mouse locus in B (upper) and ES (lower) cells. Demethylation was calculated by subtracting normalized ES cell methylation values from the activated B cell Bis-Seq and vice versa.

(C) Bar graph showing qPCR analysis of *Brd2*, *Mtch1*, and *Pim1* expression in WT (black) or ΔE2 (red) CH12 cells. Same analysis for WT or ΔE6 (blue) ES cells. Data are represented as the mean ± SEM (n = 6). *Pim1* $p < 0.005$ in B cells and $p < 0.003$ in ES cells.

(D) Bar graph represents the fraction (%) of 1,518 activated B cell enhancers that are CpG demethylated in ES, KSL, CLP, and G_0 resting (r) B cells. Absolute numbers of demethylated enhancers are provided on top of each bar.

cleavage paralleled the topology of the protein-DNA interphase with a marked depression at DNA binding motif (Figure 6E; Huang et al., 2001). We conclude that DNaseI cleavage profiles can recapitulate the structural features of transcription factor:DNA interactions, and thus they reflect the occupancy of gene regulatory proteins across the genome.

Lineage-Determining Factors Regulate Transcription of Broadly Expressed Genes

Having validated the genomic footprinting approach, we next turned to the question of whether lineage-specific factors associate with enhancers controlling transcription of broadly expressed genes. To this end we classified B and ES cells enhancers into three groups: (1) those shared between the two cell types and linked to the same promoter as determined by ChIA-PET, (2) cell-type-specific enhancers bound to promoters active both in B and ES cells, and (3) cell-type-specific enhancers linked to cell-type-specific promoters (Figure 7A). Within the B cell compartment, all transcription factors analyzed were associated with the three enhancer groups. Ebf1, Oct2, and E2A footprints for instance showed no significant differences in their distribution regardless of enhancer specificity (Figure 7B). Other factors, such as E2f2 and Foxo1 displayed statistical significant biases for shared and cell-type-specific enhancers, respectively (Figure 7B). However, even in

(Figure 6C). Similar results were obtained with Irf8 and Ebf1 (not shown). To confirm that footprinting profiles truly reflect the morphology of DNA-protein interactions, we turned to available transcription factor:DNA cocrystal structures. For instance, consistent with the published Ebf1:DNA structure (Treiber et al., 2010), backbone phosphates in direct contact with Ebf1 amino acids via hydrogen bounds (in red in Figure 6D) were the most protected from DNaseI cleavage, whereas more exposed residues (in cyan) displayed by comparison greater susceptibility to digestion. Similarly, for the cRel (NF-κB) factor DNaseI

(D) Genes classified by the total number of tissue-specific enhancers they interact with. Individual examples are highlighted and the number of B-cell-specific (red) and ES-cell-specific (blue) enhancers is provided in parenthesis.

(E) Box plot providing changes in mRNA copy number for genes that are anchored both in B and ES cells and that associate with ≤2 ES-cell-specific enhancers (blue), ≥2 B-cell-specific enhancers (red), or genes in between (gray). See also Figure S6 and Table S2.

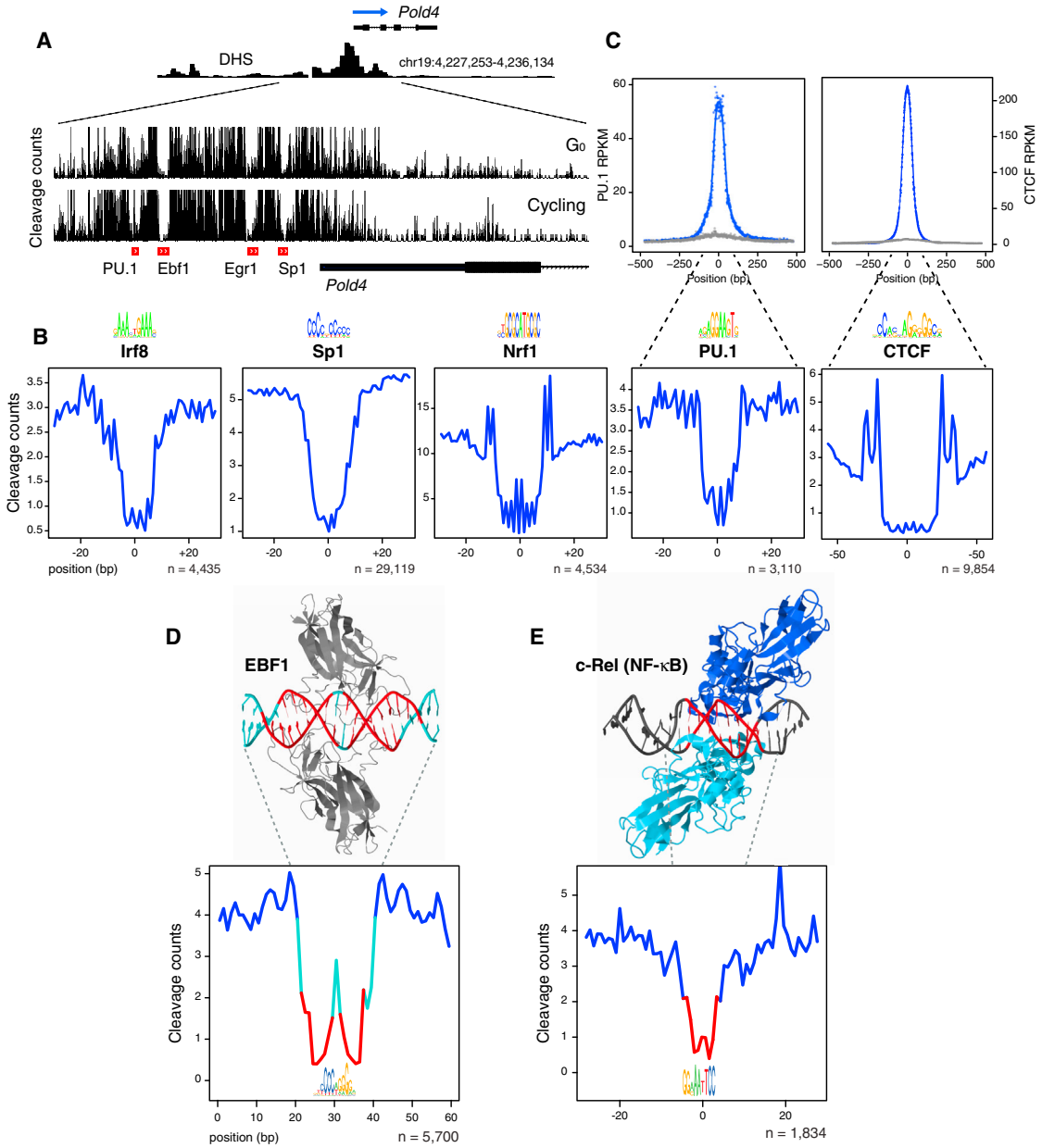

Figure 6. Digital Genomic Footprinting

(A) Characterization of footprints at the *Pold4* promoter in primary resting or cycling CH12 B cells. Transcription factor binding motifs overlapping with each footprint are shown below the graph (red rectangles). Tracks were configured to display the maximum (light gray) and one standard deviation above the mean (dark gray).

(B) Examples of footprints overlapping with Irf8, Sp1, Nrf1, PU.1, and CTCF DNA recognition motifs.

(C) Composite of PU.1 and CTCF ChIP-seq (blue data points, upper graphs) and cumulative footprinting (lower graphs) associated with cognate binding motifs (middle logos) in B cells. The absolute number of motif occurrences is provided. Grey data points represent ChIP-seq signals at footprints not associated with PU.1 or CTCF motifs.

(D) The cocrystal structure of Ebf1 bound to its DNA ligand is compared to its cognate footprint profile. Motif nucleotides least sensitive to cleavage are depicted in red; most sensitive residues are depicted in cyan.

(E) Similar analysis as in (D) for the NFκB-cRel dimer.

these cases, footprints were not excluded from any enhancer group. A similar distribution was observed in ES cells (Figure S7C). Thus, lineage specification factors associate both with tissue-specific and broadly active enhancers. Confirming this finding, ChIP-seq analysis of the pluripotent factors Nanog, Oct4, and Sox2 showed occupancy of B and ES cell shared enhancers at the *Ube2g1* locus as well as binding to ES-cell-specific enhancers at the *Myc* locus (Figure 7C). A global analysis of

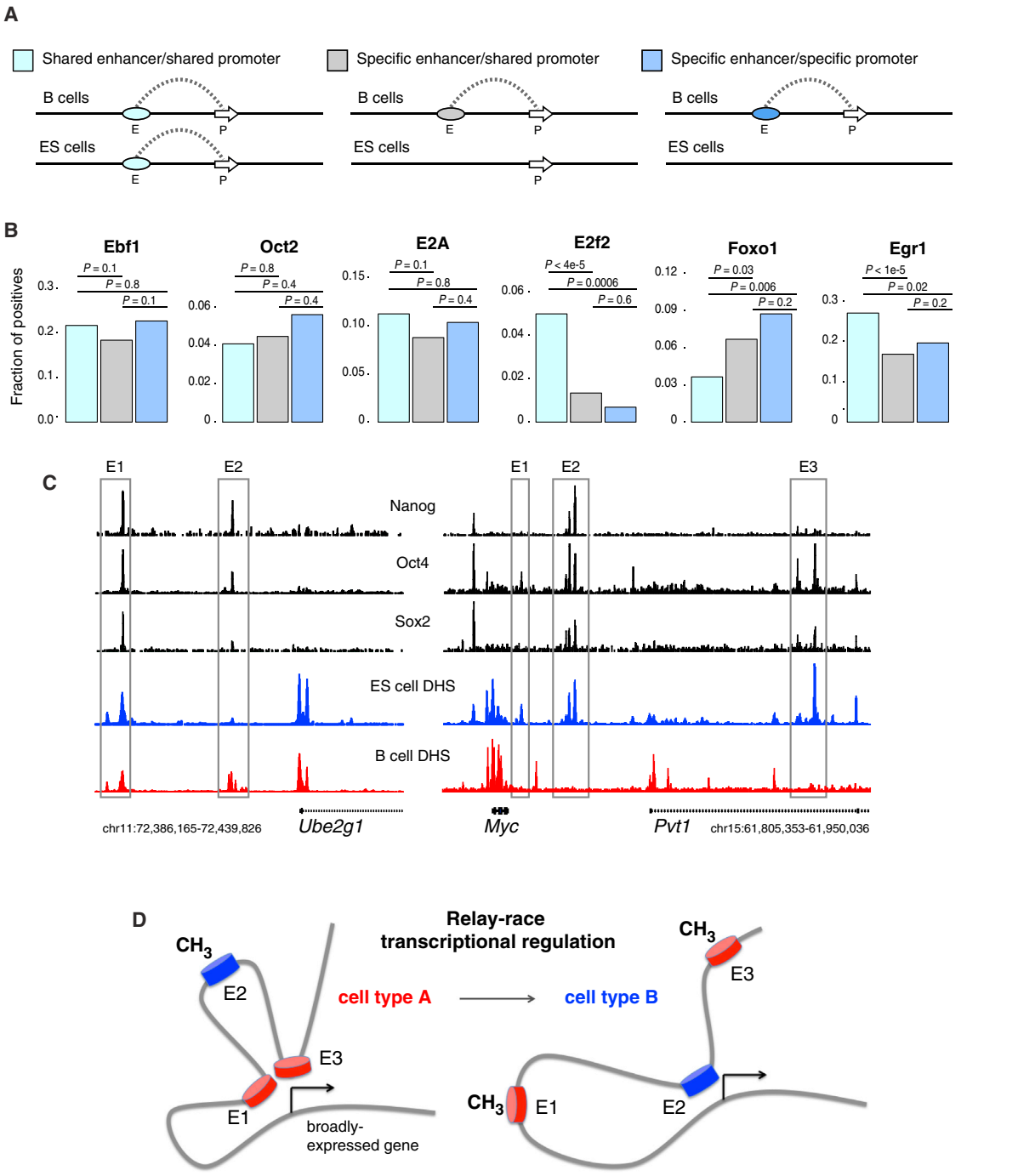

Figure 7. Distribution of Transcription Factor Binding across the B and ES Cell Genome

(A) Classification of B and ES cell enhancers into those present in both cell types and associated with shared promoters (light blue), cell-type-specific enhancers associated with broadly active promoters (gray), and cell-specific enhancers associated with cell-specific promoters (blue).

(B) Fraction of each enhancer type harboring a particular motif in a footprint.

(C) ChIP-seq analysis of pluripotent factors Nanog, Oct4, and Sox2 at the *Ube2g1* and *Myc* gene loci.

(D) Transcription of broadly expressed genes is driven during development by a relay race type of regulation, where the landscape of active enhancers varies in different cell types concomitant with the dynamics of CpG methylation. See also Figure S7.

Irf4, Irf8, Pu.1, Ebf1, and Stat6 ChIP-seq from B cells also corroborated the results (data not shown). We conclude that transcription factors driving lineage specification associate with regulatory elements of broadly expressed genes. We point out that the functional significance of TF recruitment as defined by digital footprinting remains to be empirically determined. At the same time, our findings are consistent with the notion that broadly expressed genes can be regulated in a tissue-specific manner.

DISCUSSION

We have here characterized the first interactomes of gene regulatory domains in primary cells. The data provide a wealth of information from the mouse genome with thousands of promoter-enhancer pairs. Even for loci that have been extensively studied, the analysis uncovered new connectivities. At the *Igκ* locus, for instance, we identified novel enhancers linked to 5′Eκ, the regulatory domain from which NF-κB was originally isolated 27 years ago (Sen and Baltimore, 1986). ChIA-PET also uncovered three additional enhancers tethered at a long distance (up to 50 kb) to the AID gene promoter; the previously characterized AID enhancers are all located within 15 kb of the TSS. Thus, one clear advantage of ChIA-PET lies in its ability to identify long-range interactions, even when they leapfrog noninteracting genes, as is the case for 65% of all enhancers (Table S1). A striking example is a pair of giant enhancers linked to *Pax5* by skipping over ~250 kb of DNA containing the *Zcchc7* gene (Figure S7D).

Like other conformation capturing techniques, ChIA-PET does not directly address functionality of chromatin interactions. This can only be determined empirically by other means. Typically, enhancer activity is defined by luciferase-based plasmids or LacZ transgenes. However, these experimental approaches only provide an incomplete view of transcriptional regulation because they either lack proper chromatin structure or the influence of neighboring enhancers, insulators, and silencers is not taken into account. Conversely, genome editing provides a means to measure the impact of enhancer deletion in the physiological context. In most cases, we found that cognate promoter activity was partially reduced following enhancer ablation, supporting the model that the contribution of individual enhancers to gene expression is additive in nature. Examples of this category were enhancers linked to the *Cd79a*, *Pou2af1*, and *Pim1*. On the other hand, the targeting of AID 5′ enhancers E1 or E2 nearly entirely abolished AID expression and activity, a result that is reminiscent of those obtained upon deletion of AID enhancers E3 and E5 in BAC transgenic mice (Crouch et al., 2007; Huong et al., 2013), or by interference with E4 activity in primary B cells (Sayegh et al., 2003). Thus, rather than working in additive fashion, AID gene regulatory elements seem to synergize or act as a cooperative unit. Conceivably, the local topology of the *Aicda-Apobec1* locus requires activation of all enhancers for optimal transcription to occur. This strategy may be useful for genes that require tight regulation as AID, whose expression is strictly limited to activated B cells to minimize its well known tumorigenic activity (Casellas et al., 2009). Consistent with this model, our methylome analysis indicates that the AID locus is not completely demethylated until B cells are activated.

A direct comparison between the B and ES cell interactomes revealed that up to 95% of genes anchored in both cell types are associated with at least one tissue-specific enhancer. Transcription modularity, the mechanism whereby genes accumulate regulatory elements during ontogeny, is a well-described phenomenon that controls the spatiotemporal expression of developmental genes (Davidson, 2001; Visel et al., 2007). For instance, the cardiac homeobox gene *Nkx2-5* is targeted to specific subregions of the developing heart by turning on additional *cis*-regulatory domains over time (Schwartz and Olson, 1999). Similarly, expression of the human apolipoprotein E gene is triggered in hepatocytes and astrocytes by enhancers only active in those tissues (Grehan et al., 2001). Our studies demonstrate that the turning on and off of enhancers is not a singularity of developmental gene loci. Instead, it is a widespread mechanism that regulates broadly expressed genes involved in basic cellular functions, such as cell-cycle initiation (*Myc*), signal transduction (*β-catenin*), and cellular motility (*Malat1*).

Mechanistically, we show that the changing enhancer landscape in mammalian development results from the unbiased recruitment of lineage-determining factors, which associate with enhancers anchored not only to tissue-specific promoters but also to constitutively active ones. Based on these observations we propose a "relay race" model of transcriptional regulation, whereby broadly active genes make use of tissue-specific cis-regulatory elements and transcription factors as cells progress through development (Figure 7D). For genes that only replace a subset of their regulatory domains, transcription is roughly maintained at comparable levels in different cell types. However, as the number of connected enhancers fluctuates considerably, promoter activity can be significantly altered.

We can think of at least two reasons why higher organisms modulate the enhancer landscape of broadly expressed genes. First, as aforementioned, it enables fine-tuning of protein output, which in turns controls protein activity. Second, it places basic cellular functions under the control of tissue-specific factors. In the B cell compartment, these strategies are perhaps best illustrated during the immune response to invading pathogens. In this microenvironment B lymphocytes move rapidly from a G_0, quiescent state to one of the fastest proliferative rates among eukaryotic cells (Liu et al., 1991). Key in this process is *Myc* which, along with TFIIH, triggers a ~10-fold amplification of the lymphocyte transcriptome concomitant with cell-cycle entry (Kouzine et al., 2013; Nie et al., 2012). Thus, vis-à-vis proliferation B cells differ substantially from continuously dividing ES cells, in that they must rapidly engage a burst of *Myc* expression and activity during the immune response to cope with fast dividing pathogens. Our studies imply that this unique response is mediated, at least in part, by the large number of *cis*- and *trans*-responsive elements that associate with the *Myc* promoter in the B cell compartment.

EXPERIMENTAL PROCEDURES

Cell Isolation and Culture

Hematopoietic stem cells were isolated from the bone marrow of 25 6-week-old C57BL6 mice (Jackson Laboratory). Cells were purified following Ema et al.'s protocol (PMID: 17406558). KSL (KIT+, SCA1+, Lin-, IL-7R-) and CLP (KIT+, SCA1+, Lin-, IL-7R+) cells were sorted using MoFlo Legacy (Beckman Coulter) and BD FACSAria III. Resting splenic B cells were isolated from

6- to 8-week-old wild-type C57BL6/J mice with anti-CD43 Microbeads (anti-Ly48; Miltenyi Biotech) and were activated for 48-60 hr with LPS (50 μg/ml; Sigma), IL-4 (5 ng/ml; Sigma) and 0.5 μg/ml of anti-CD180 (RP105) antibody (RP/14, BD PharMingen). E14 tg2A mouse embryonic stem cells were maintained as described in (PMID:18555785). Switchable $IgM^+ \rightarrow IgA^+$ murine CH12-F3 Ly-1^+ B cell lymphoma line was maintained and passaged every 2 days in RPMI 1640 supplemented with 10% FBS (ATCC), 1% penicillin/streptomycin (Invitrogen), 55 μM 2-β mercaptoethanol (Invitrogen). All cells were maintained at 37°C and 5% CO_2 in a humidified incubator.

ACCESSION NUMBERS

The Short Read Archive Project Number for the ChIP-seq data in activated B cells (other than CTCF) and CH12 cells, DHS-seq in activated B cells, RNA-seq, whole genome methylation (other than ES cells), and ChIA-PET data reported in this paper is SRP029721.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, seven figures, and three tables and can be found with this article online at http://dx.doi.org/10.1016/j.cell.2013.11.039.

ACKNOWLEDGMENTS

We thank Kefei Yu for CH12 cells and protocols; J. Simone and J. Lay for cell sorting; G. Gutierrez for technical assistance with Illumina analyzer. This work was supported by the Intramural Research Program of NIAMS and NCI, and internal Jackson Laboratory fund JAX19020120 to Y.R. J.K.J. was supported by NIH grants DP1 GM105378 and P50 HG005550, the Defense Advanced Research Projects Agency grant W911NF-11-2-0056 and The Jim and Ann Orr Massachusetts General Hospital Research Scholar Award. J.Q. was supported by the NIH UGSP program. All animal experiments were performed according to NIH guidelines. High-performance computation was performed using NIH Helix Systems (http://helix.nih.gov). J.K.J. has a financial interest in Transposagen Biopharmaceuticals. J.K.J.'s interests were reviewed and are managed by Massachusetts General Hospital and Partners HealthCare in accordance with their conflict of interest policies.

Received: September 11, 2013
Revised: November 1, 2013
Accepted: November 25, 2013
Published: December 19, 2013

REFERENCES

Baek, S., Sung, M.H., and Hager, G.L. (2012). Quantitative analysis of genome-wide chromatin remodeling. Methods Mol. Biol. *833*, 433–441.

Batista, P.J., and Chang, H.Y. (2013). Long noncoding RNAs: cellular address codes in development and disease. Cell *152*, 1298–1307.

Buecker, C., and Wysocka, J. (2012). Enhancers as information integration hubs in development: lessons from genomics. Trends Genet. *28*, 276–284.

Bultmann, S., Morbitzer, R., Schmidt, C.S., Thanisch, K., Spada, F., Elsaesser, J., Lahaye, T., and Leonhardt, H. (2012). Targeted transcriptional activation of silent oct4 pluripotency gene by combining designer TALEs and inhibition of epigenetic modifiers. Nucleic Acids Res. *40*, 5368–5377.

Casellas, R., Jankovic, M., Meyer, G., Gazumyan, A., Luo, Y., Roeder, R., and Nussenzweig, M. (2002). OcaB is required for normal transcription and V(D)J recombination of a subset of immunoglobulin kappa genes. Cell *110*, 575–585.

Casellas, R., Yamane, A., Kovalchuk, A.L., and Potter, M. (2009). Restricting activation-induced cytidine deaminase tumorigenic activity in B lymphocytes. Immunology *126*, 316–328.

Chen, X., Xu, H., Yuan, P., Fang, F., Huss, M., Vega, V.B., Wong, E., Orlov, Y.L., Zhang, W., Jiang, J., et al. (2008). Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. Cell *133*, 1106–1117.

Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A., et al. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc. Natl. Acad. Sci. USA *107*, 21931–21936.

Crouch, E.E., Li, Z., Takizawa, M., Fichtner-Feigl, S., Gourzi, P., Montaño, C., Feigenbaum, L., Wilson, P., Janz, S., Papavasiliou, F.N., and Casellas, R. (2007). Regulation of AID expression in the immune response. J. Exp. Med. *204*, 1145–1156.

Davidson, E.H. (2001). Genomic regulatory systems: development and evolution (San Diego: Academic Press).

ENCODE Project Consortium, Bernstein, B.E., Birney, E., Dunham, I., Green, E.D., Gunter, C., and Snyder, M. (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57–74.

Fullwood, M.J., Liu, M.H., Pan, Y.F., Liu, J., Xu, H., Mohamed, Y.B., Orlov, Y.L., Velkov, S., Ho, A., Mei, P.H., et al. (2009). An oestrogen-receptor-alpha-bound human chromatin interactome. Nature *462*, 58–64.

Grehan, S., Tse, E., and Taylor, J.M. (2001). Two distal downstream enhancers direct expression of the human apolipoprotein E gene to astrocytes in the brain. J. Neurosci. *21*, 812–822.

Guil, S., and Esteller, M. (2012). Cis-acting noncoding RNAs: friends and foes. Nat. Struct. Mol. Biol. *19*, 1068–1075.

Heintzman, N.D., Stuart, R.K., Hon, G., Fu, Y., Ching, C.W., Hawkins, R.D., Barrera, L.O., Van Calcar, S., Qu, C., Ching, K.A., et al. (2007). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat. Genet. *39*, 311–318.

Huang, D.B., Chen, Y.Q., Ruetsche, M., Phelps, C.B., and Ghosh, G. (2001). X-ray crystal structure of proto-oncogene product c-Rel bound to the CD28 response element of IL-2. Structure *9*, 669–678.

Huong, T., Kobayashi, M., Nakata, M., Shioi, G., Miyachi, H., Honjo, T., and Nagaoka, H. (2013). In vivo analysis of Aicda gene regulation: a critical balance between upstream enhancers and intronic silencers governs appropriate expression. PLoS ONE *8*, e61433.

Kagey, M.H., Newman, J.J., Bilodeau, S., Zhan, Y., Orlando, D.A., van Berkum, N.L., Ebmeier, C.C., Goossens, J., Rahl, P.B., Levine, S.S., et al. (2010). Mediator and cohesin connect gene expression and chromatin architecture. Nature *467*, 430–435.

Kouzine, F., Wojtowicz, D., Yamane, A., Resch, W., Kieffer-Kwon, K.-R., Bandle, R., Nelson, S., Nakahashi, H., Awasthi, P., Feigenbaum, L., et al. (2013). Global regulation of promoter melting in naive lymphocytes. Cell *153*, 988–999.

Krivega, I., and Dean, A. (2012). Enhancer and promoter interactions-long distance calls. Curr. Opin. Genet. Dev. *22*, 79–85.

Kuchen, S., Resch, W., Yamane, A., Kuo, N., Li, Z., Chakraborty, T., Wei, L., Laurence, A., Yasuda, T., Peng, S., et al. (2010). Regulation of microRNA expression and abundance during lymphopoiesis. Immunity *32*, 828–839.

Li, G., Ruan, X., Auerbach, R.K., Sandhu, K.S., Zheng, M., Wang, P., Poh, H.M., Goh, Y., Lim, J., Zhang, J., et al. (2012). Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. Cell *148*, 84–98.

Liu, Y.J., Zhang, J., Lane, P.J., Chan, E.Y., and MacLennan, I.C. (1991). Sites of specific B cell activation in primary and secondary responses to T cell-dependent and T cell-independent antigens. Eur. J. Immunol. *21*, 2951–2962.

Liu, Z.M., George-Raizen, J.B., Li, S., Meyers, K.C., Chang, M.Y., and Garrard, W.T. (2002). Chromatin structural analyses of the mouse Igkappa gene locus reveal new hypersensitive sites specifying a transcriptional silencer and enhancer. J. Biol. Chem. *277*, 32640–32649.

Meyer, K.B., and Neuberger, M.S. (1989). The immunoglobulin kappa locus contains a second, stronger B-cell-specific enhancer which is located downstream of the constant region. EMBO J. *8*, 1959–1964.

Nakamura, M., Kondo, S., Sugai, M., Nazarea, M., Imamura, S., and Honjo, T. (1996). High frequency class switching of an IgM+ B lymphoma clone CH12F3 to IgA+ cells. Int. Immunol. *8*, 193–201.

Neph, S., Vierstra, J., Stergachis, A.B., Reynolds, A.P., Haugen, E., Vernot, B., Thurman, R.E., John, S., Sandstrom, R., Johnson, A.K., et al. (2012). An expansive human regulatory lexicon encoded in transcription factor footprints. Nature *489*, 83–90.

Nie, Z., Hu, G., Wei, G., Cui, K., Yamane, A., Resch, W., Wang, R., Green, D.R., Tessarollo, L., Casellas, R., et al. (2012). c-Myc is a universal amplifier of expressed genes in lymphocytes and embryonic stem cells. Cell *151*, 68–79.

Nobrega, M.A., Ovcharenko, I., Afzal, V., and Rubin, E.M. (2003). Scanning human gene deserts for long-range enhancers. Science *302*, 413.

Phillips-Cremins, J.E., Sauria, M.E., Sanyal, A., Gerasimova, T.I., Lajoie, B.R., Bell, J.S., Ong, C.T., Hookway, T.A., Guo, C., Sun, Y., et al. (2013). Architectural protein subclasses shape 3D organization of genomes during lineage commitment. Cell *153*, 1281–1295.

Ponting, C.P., Oliver, P.L., and Reik, W. (2009). Evolution and functions of long noncoding RNAs. Cell *136*, 629–641.

Qin, X.F., Reichlin, A., Luo, Y., Roeder, R.G., and Nussenzweig, M.C. (1998). OCA-B integrates B cell antigen receptor-, CD40L- and IL 4-mediated signals for the germinal center pathway of B cell development. EMBO J. *17*, 5066–5075.

Reyon, D., Tsai, S.Q., Khayter, C., Foden, J.A., Sander, J.D., and Joung, J.K. (2012). FLASH assembly of TALENs for high-throughput genome editing. Nat. Biotechnol. *30*, 460–465.

Sabo, P.J., Kuehn, M.S., Thurman, R., Johnson, B.E., Johnson, E.M., Cao, H., Yu, M., Rosenzweig, E., Goldy, J., Haydock, A., et al. (2006). Genome-scale mapping of DNase I sensitivity in vivo using tiling DNA microarrays. Nat. Methods *3*, 511–518.

Sayegh, C.E., Quong, M.W., Agata, Y., and Murre, C. (2003). E-proteins directly regulate expression of activation-induced deaminase in mature B cells. Nat. Immunol. *4*, 586–593.

Schwartz, R.J., and Olson, E.N. (1999). Building the heart piece by piece: modularity of cis-elements regulating Nkx2-5 transcription. Development *126*, 4187–4192.

Sen, R., and Baltimore, D. (1986). Multiple nuclear factors interact with the immunoglobulin enhancer sequences. Cell *46*, 705–716.

Shen, L., Wu, H., Diep, D., Yamaguchi, S., D'Alessio, A.C., Fung, H.L., Zhang, K., and Zhang, Y. (2013). Genome-wide analysis reveals TET- and TDG-dependent 5-methylcytosine oxidation dynamics. Cell *153*, 692–706.

Song, C.X., Szulwach, K.E., Dai, Q., Fu, Y., Mao, S.Q., Lin, L., Street, C., Li, Y., Poidevin, M., Wu, H., et al. (2013). Genome-wide profiling of 5-formylcytosine reveals its roles in epigenetic priming. Cell *153*, 678–691.

Stadler, M.B., Murr, R., Burger, L., Ivanek, R., Lienert, F., Schöler, A., van Nimwegen, E., Wirbelauer, C., Oakeley, E.J., Gaidatzis, D., et al. (2011). DNA-binding factors shape the mouse methylome at distal regulatory regions. Nature *480*, 490–495.

Takizawa, M., Tolarová, H., Li, Z., Dubois, W., Lim, S., Callen, E., Franco, S., Mosaico, M., Feigenbaum, L., Alt, F.W., et al. (2008). AID expression levels determine the extent of cMyc oncogenic translocations and the incidence of B cell tumor development. J. Exp. Med. *205*, 1949–1957.

Thurman, R.E., Rynes, E., Humbert, R., Vierstra, J., Maurano, M.T., Haugen, E., Sheffield, N.C., Stergachis, A.B., Wang, H., Vernot, B., et al. (2012). The accessible chromatin landscape of the human genome. Nature *489*, 75–82.

Treiber, N., Treiber, T., Zocher, G., and Grosschedl, R. (2010). Structure of an Ebf1:DNA complex reveals unusual DNA recognition and structural homology with Rel proteins. Genes Dev. *24*, 2270–2275.

Visel, A., Bristow, J., and Pennacchio, L.A. (2007). Enhancer identification through comparative genomics. Semin. Cell Dev. Biol. *18*, 140–152.

Visel, A., Rubin, E.M., and Pennacchio, L.A. (2009). Genomic views of distant-acting enhancers. Nature *461*, 199–205.

Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl, P.B., Lee, T.I., and Young, R.A. (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell *153*, 307–319.

Wuerffel, R., Wang, L., Grigera, F., Manis, J., Selsing, E., Perlot, T., Alt, F.W., Cogne, M., Pinaud, E., and Kenter, A.L. (2007). S-S synapsis during class switch recombination is promoted by distantly located transcriptional elements and activation-induced deaminase. Immunity *27*, 711–722.

Xie, W., and Ren, B. (2013). Developmental biology. Enhancing pluripotency and lineage specification. Science *341*, 245–247.

Zhang, J., Poh, H.M., Peh, S.Q., Sia, Y.Y., Li, G., Mulawadi, F.H., Goh, Y., Fullwood, M.J., Sung, W.K., Ruan, X., and Ruan, Y. (2012). ChIA-PET analysis of transcriptional chromatin interactions. Methods *58*, 289–299.

Ziller, M.J., Gu, H., Müller, F., Donaghey, J., Tsai, L.T., Kohlbacher, O., De Jager, P.L., Rosen, E.D., Bennett, D.A., Bernstein, B.E., et al. (2013). Charting a dynamic DNA methylation landscape of the human genome. Nature *500*, 477–481.