11070–11087 *Nucleic Acids Research, 2017, Vol. 45, No. 19* *Published online 17 August 2017*
doi: 10.1093/nar/gkx722

Developmentally regulated higher-order chromatin interactions orchestrate B cell fate commitment

Ravi Boya[1,†], Anurupa Devi Yadavalli[1,†], Sameena Nikhat[1], Sreenivasulu Kurukuti[1], Dasaradhi Palakodeti[2] and Jagan M. R. Pongubala[1,*]

[1]Department of Animal Biology, School of Life Sciences, University of Hyderabad, Hyderabad 500046, India and [2]Institute for Stem Cell Biology and Regenerative Medicine, National Centre for Biological Sciences, Bangalore 560065, India

Received May 25, 2017; Revised July 21, 2017; Editorial Decision August 07, 2017; Accepted August 09, 2017

ABSTRACT

Genome organization in 3D nuclear-space is important for regulation of gene expression. However, the alterations of chromatin architecture that impinge on the B cell-fate choice of multi-potent progenitors are still unclear. By integrating *in situ* Hi-C analyses with epigenetic landscapes and genome-wide expression profiles, we tracked the changes in genome architecture as the cells transit from a progenitor to a committed state. We identified the genomic loci that undergo developmental switch between A and B compartments during B-cell fate determination. Furthermore, although, topologically associating domains (TADs) are stable, a significant number of TADs display structural alterations that are associated with changes in *cis*-regulatory interaction landscape. Finally, we demonstrate the potential roles for Ebf1 and its downstream factor, Pax5, in chromatin reorganization and transcription regulation. Collectively, our studies provide a general paradigm of the dynamic relationship between chromatin reorganization and lineage-specific gene expression pattern that dictates cell-fate determination.

INTRODUCTION

It is increasingly evident that the assembly of higher-order genome structures and their associated sub-nuclear compartments are intimately linked with transcriptional activity (1,2). Recent advances in high-throughput Chromosome Conformation Capture (3C)-derived methods have enabled quantitative measurement of physical interactions of chromatin in 3D nuclear space (2–6). These studies have demonstrated that chromatin is organized into transcriptionally permissive (A) and repressive (B) compartments indicating that chromatin positioning in 3D nuclear space may be associated with gene activity. For instance, B cell specification is associated with relocalization of *Igh* alleles from the nuclear periphery (a repressive compartment) towards center of the nucleus (an active compartment), where they undergo long-range interactions and subsequent rearrangements (7–9). These findings provide a functional link between sub-nuclear localization of the chromatin and gene activity. Recent studies indicate that chromatin compartments are further organized into varying sizes of dense and highly self-interacting regions, known as Topologically Associating Domains (TADs). These chromatin domains have been found to be stable and conserved across various cell types (10). In mammalian cells, insulator binding protein, CTCF, is found to be enriched in TAD boundaries (10). The deletion of boundary regions results in an increase in inter-domain interactions indicating the structural and functional role of insulators in maintenance of discrete, functional chromatin domains (11,12). Further it was demonstrated that loss of CTCF results in dose dependent insulation defects at most of the TAD boundaries (13). However, recent studies suggest that depletion of cohesin-loading factor *Nipbl*, but not CTCF, results in genome-wide disappearance of TADs, reinforcing the critical role of cohesin in the formation of TADs by loop extrusion mechanism (Schwarzer et al., 2016; Kubo et al., 2017, Unpublished). Although TADs are invariant, the intrinsic interactions within these TADs were found to be varying (10,14). Moreover, several studies show that the cell type-specific gene expression is regulated through interactions between promoters and distantly located *cis*-regulatory elements, particularly enhancers, by looping out of intervening DNA sequences (15–18). These long-range interactions were found to be associated with changes in histone modifications and DNA methylation (19–21). Furthermore, the transcriptional output is controlled by a combinatorial binding of transcription factors at *cis*-regulatory elements (22–25). Thus, a number of molecular mechanisms contribute to the precise regulation of gene expression pat-

*To whom correspondence should be addressed. Tel: +91 40 23134580; Fax: +91 40 23010145; Email: jpsl@uohyd.ernet.in
†These authors contributed equally to this work as first authors.

tern that drives lineage differentiation to maintain cell identity.

Gene ablation studies demonstrate that the B cell developmental program is orchestrated by unique sets of transcription factors including PU.1 (*Sfpi1*), Ikaros (*Ikzf1*), E2A (*Tcf3*), Ebf1 (*Ebf1*) and Pax5 (*Pax5*) (26–31). PU.1, Ikaros and E2A promote the lymphoid developmental competence of multipotent progenitors (MPPs) (30,32,33). Loss of any of these factors results in a severe block to the development of B cells prior to *Igh* V to DJ recombination (26,28,34). These mutant progenitors fail to express both Ebf1 and Pax5. While PU.1, Ikaros and E2A are required for B cell development *in vivo*, their functions can be bypassed *in vitro* by complementing with Ebf1 but not with Pax5 (33,35,36). Thus, PU.1, Ikaros and E2A are necessary for the development of early lymphoid progenitors, whereas Ebf1 and Pax5 function as primary and secondary regulators of B cell fate determination (37–40). Correspondingly, *Ebf1*−/− hematopoietic progenitors display multilineage developmental potential, both *in vitro* and *in vivo*. Restoration of Ebf1 expression in *Ebf1*−/− progenitors inhibits their alternative lineage choice and induces B cell development, independently of Pax5 (41). Although much is known about the concerted interplay of transcription factors that are important for B cell determination, far less is known about the genome-wide composition of *cis*-regulatory interactions controlling B-lineage-specific gene expression program and relevance of these interactions on hierarchical organization of the chromatin during early B cell developmental transition.

To obtain a comprehensive view of the relationship between higher-order chromatin reorganization and induction of developmentally regulated B lineage-specific gene expression program, we carried out *in situ* Hi-C, in pre-pro-B cells (*Ebf1*−/− progenitors) and pro-B cells (*Rag2*−/−). Our comparative analysis of chromatin interactomes between pre-pro-B and pro-B cells revealed a distinct set of genomic loci that switch between A and B compartments. In addition, we show that TADs serve as coordinated sub-units of chromatin and undergo structural alterations during developmental transition from pre-pro-B to pro-B stage. Correspondingly, we demonstrate that the *cis*-regulatory interaction landscape displays extensive rewiring, thereby, modulating the transcriptional activity during pre-pro-B to pro-B cell transition. Finally, we show that Ebf1 regulates differential gene expression pattern, at least in part, through chromatin relocalization and establishment of long-range promoter-*cis*-regulatory interactions. Collectively, our results demonstrate that, B lineage-specific gene expression pattern is closely associated with dynamic reorganization of chromatin in a developmental stage-specific manner.

MATERIALS AND METHODS

Cell culture

Pre-pro-B cells (*Ebf1*−/− progenitors) were maintained on stromal layer (OP9 cells) in the presence of Opti-MEM (Gibco) containing 4% (v/v) fetal calf serum, β-mercaptoethanol (50 μM), penicillin (10 U/ml) and streptomycin (10 μg/ml) and supplemented with SCF (10 ng/ml), Flt3L (10 ng/ml) and IL-7 (5 ng/ml). Pro-B cells (*Rag2*−/− cells) were maintained under similar conditions except that the media was supplemented with only IL-7 (5 ng/ml). Both pre-pro-B cells and pro-B cells were used for preparation of RNA for RT-PCR and chromatin for the 3C and Hi-C assays.

In situ Hi-C and 3C experiments

In situ Hi-C as well as 3C experiments were carried out using pre-pro-B and pro-B cells as described previously (2,3). During *in situ* Hi-C, chromatin cross-linking, restriction enzyme (HindIII) digestion, biotin fill-in and ligation reactions were performed in intact nuclei (42,43). In case of 3C experiments, chromatin ligation following restriction digestion were performed in intact nuclei and the interaction frequencies between pre-pro-B and pro-B cells were normalized using a control region in *Ercc3* gene.

Identification of topologically associated domains

Iteratively corrected relative contact probability matrices at 40 kb resolution, generated by implementing HiResHiC module of hiclib were converted into the format specified by Domain Caller (10), where the first three columns represent the chromosome number followed by start and end of the bin. Domain Caller is a simple and straightforward approach with greater flexibility to identify biologically relevant domain structures.

Generation of 3D structures of TADs

We have generated 3D structures of TADs in both pre-pro-B and pro-B cells by implementing AutoChrom3D (44), which uses a novel sequencing-bias-relaxed parameter to normalize chromatin interactions.

Determination of statistically significant *cis*-regulatory interactions

To discriminate between random polymer loops and specific chromatin loops, we have used Fit-Hi-C (45), a tool for assigning statistical confidence estimates to mid-range contacts. We have prepared 'FRAGSFILE' containing midpoints (or start indices) of the fragments and 'INTERSFILE' containing interactions between fragment pairs from the dict-file obtained through fragment level filtering. The BIASFILE is prepared by using the python code that implements the iterative correction in sparse mode by filtering out loci that are less mappable than the threshold (cut off $\geq$ 0.5). The significant interactions obtained by implementing Fit-Hi-C, were further integrated with various epigenetic modifications (H3K4me1, H3K4me3, H3K4me2 and H3K9/14ac) to identify potential *cis*-regulatory interactions.

Microarray analysis

Pre-pro-B cells were transduced by spin-infection with retrovirus encoding GFP or Ebf1-GFP or Pax5-GFP and maintained for 2 days in lymphoid culture conditions as

previously described (41). After two days, GFP+ transductants were FACS-sorted, total RNA was isolated with TRIzol reagent (Invitrogen) and further purified on RNeasy columns (Qiagen). RNA quality control analysis was performed as previously described (41). Biotin-labeled cRNA was generated and hybridized to the Mouse Genome 430 2.0 Array (Affymetrix, Santa Clara, USA) according to the manufacturer's instructions. Data were analyzed as previously described (41).

Statistical analysis

All statistical analyses except for identifying significant promoter-*Cis*-interactions, were performed using R package. Statistical significance was evaluated by unpaired two sample *t*-test. For all the tests performed, statistical significance was assessed as $*P < 0.05$, $**P < 0.01$, $***P < 0.001$.

RESULTS

Differential chromatin compartmentalization promotes the B lineage gene expression program

To determine programmatic changes in chromatin organization during B cell development, we performed *in situ* Hi-C (Supplementary materials and methods), a high-throughput molecular approach (42,43) that captures genome-wide chromatin interactions, using *Ebf1*−/− progenitors that represent the pre-pro-B cell stage (41) and *Rag2*−/− cells that represent the pro-B cell stage (Supplementary Figures S1 and S2A–D). The *in situ* Hi-C approach is similar to the previously described dilution Hi-C method (2), except that the reactions: chromatin crosslinking, restriction enzyme digestion (HindIII), fill-in of 5′ overhangs and ligation of chromatin ends present in close proximity, were performed in intact nuclei (42). The Hi-C libraries were generated from both pre-pro-B and pro-B cells and then subjected to paired-end sequencing. Following high-throughput sequencing, the uniquely aligned (reference genome mm10) raw-reads were extensively filtered to eliminate various systemic biases originating from experimental procedures and intrinsic properties of the genome (fragment length, GC content and mappability). For this, we employed hiclib that implements filtering at multiple levels to determine the corrected contact counts (46) (Supplementary materials and methods). This approach has been known to selectively highlight the specific contacts and to facilitate the generation of corrected relative contact probability matrices, which are critical for determination of changes in chromatin architecture between the two different cell types. Thus, in comparison with similar studies (47), our strategy has two major advantages. First, *in situ* Hi-C captures specific DNA–DNA proximity ligations compared to dilution Hi-C (42,43). Second, the ICE (Iterative Correction and Eigen vector decomposition implemented by hiclib) approach significantly reduces the frequency of spurious contacts and permits fair comparison of chromatin interactome data between pre-pro-B and pro-B cells.

To gain a comprehensive understanding of progressive changes occurring in intra-chromosomal (*cis*) interactions between pre-pro-B and pro-B cells, iteratively corrected contact maps for each chromosome were generated at 1 Mb resolution (Supplementary Figure S3). Our analyses captured many of the previously identified long-range chromatin interactions (48,49), indicating that the *in situ* Hi-C approach was performed under optimal conditions and the captured interactions are valid *in vivo* (Supplementary Figure S4A and B). In line with previous studies (50), relative contact probability maps showed an ordered, dense pattern of varying sized blocks spanning across the diagonal (Figure 1A; Supplementary Figure S3). The majority of the interactions (60.0%) were limited to a range of 1–3 Mb and the frequency of such interactions decreased gradually with increasing linear genomic distance. In order to understand the differences in chromatin interaction patterns between the two cell types, we implemented Principal Component Analysis (PCA) at 1 Mb resolution (Supplementary materials and methods). As expected, these analyses revealed that chromatin is segregated into A or B compartments, which are defined by enriched or minimal interactions respectively (Figure 1A; Supplementary Figure S3). A compartments were found to contain a higher number of genes (Figure 1B) with an increase (4 fold) in CpG islands than the B compartments. Accordingly, the A compartments were substantially enriched for active histone modifications (H3K4me3, H3K4me1 and H3K9/14ac) (47) (Figure 1C) and displayed higher transcript levels when compared to the B compartments (Figure 1D), indicating that chromatin compartmentalization mirrors gene activity in both cell-types.

To investigate the possibility that selective changes in chromatin compartmentalization provide a structural framework for B-lineage gene expression (2,8), we performed PCA analysis at a higher resolution (100 kb). From these analyses, we were able to define the chromatin state of a total number of 22,360 common genes that were captured by *in situ* Hi-C in both the cell types. Of these, 16,045 genes in pre-pro-B cells and 16,643 genes in pro-B cells were found to be present in A compartments, whereas 6,315 genes in pre-pro-B cells and 5,717 genes in pro-B cells were found to be present in B compartments. Further examination of these common genes between pre-pro-B cells and pro-B cells revealed three distinct classes, including a common set of genes that are localized in either A (Group I; 68.44%) or B (Group II; 22.25%) compartments in both cell types. Consistent with previous observations (47), although a major fraction (90.69%) of genes remained in the same compartment (Group I or II) in both cell types, a distinct set of genes (Group III; 9.31%) switched between A and B compartments. Of these, 1,339 (5.98%) genes transitioned from B to A compartment, while, 741 (3.31%) genes relocalized from A to B compartment during differentiation of pre-pro-B cells to pro-B cells (Supplementary Tables S1 and S2). These observations demonstrate that B cell developmental progression from a multipotential progenitor to a specified state encompasses notable changes in chromatin compartmentalization.

In order to test whether the differential chromatin compartmentalization is associated with B lineage-specific gene expression pattern, we compared the abundance of nascent transcript levels as determined by RNA-Seq (GSE52450) of Group III genes in pre-pro-B cells and pro-B cells. We observed that the genes, which switch from the B compartment to the A compartment during differentiation, dis-

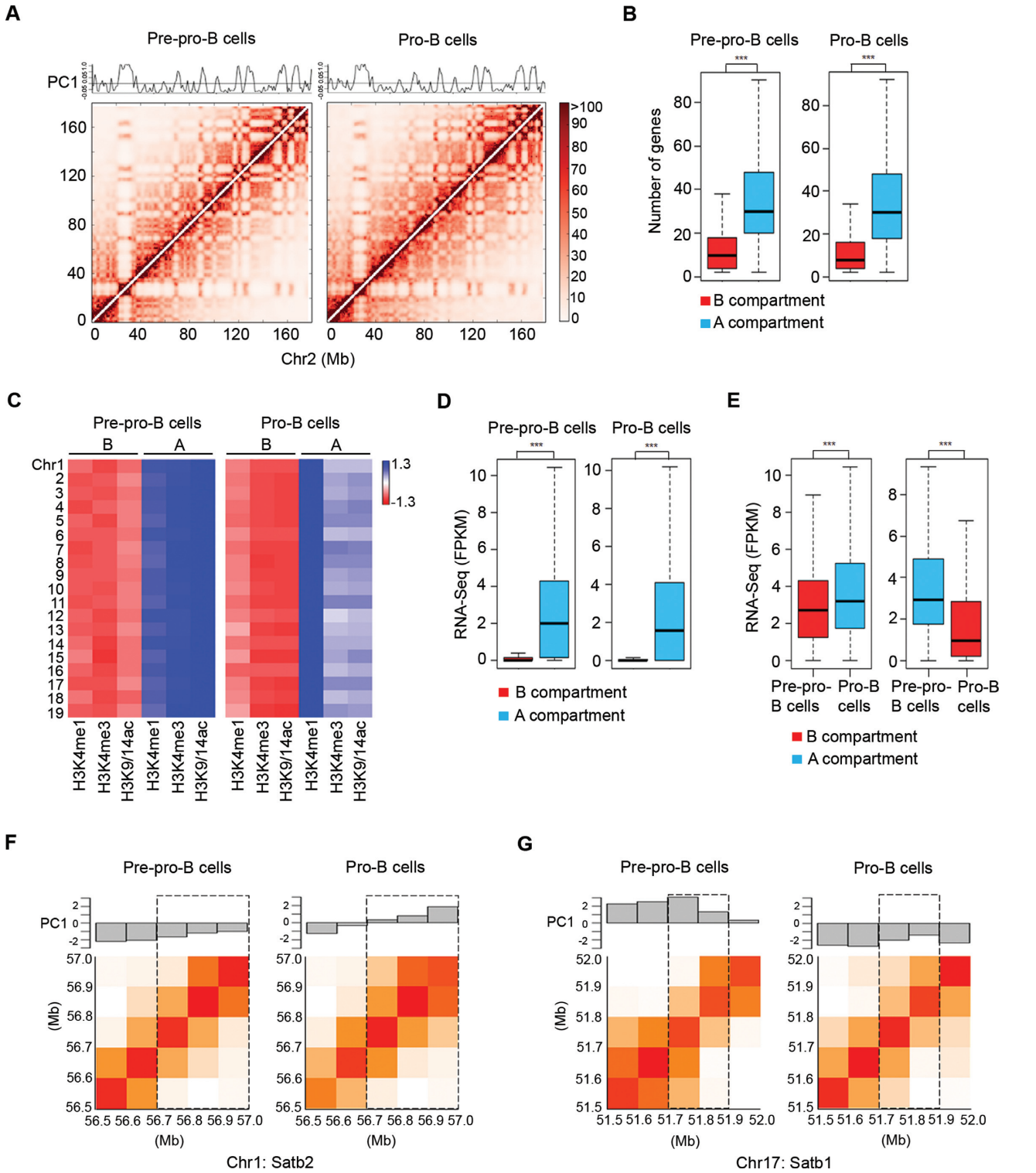

Figure 1. Chromatin compartmentalization is closely associated with gene activity. (**A**) Iteratively corrected intra-chromosomal contact count matrix of chromosome 2, representing the frequency of interactions at 1 Mb resolution. The first principal components (PC1) indicate the chromatin state on a linear genomic scale. (**B**) Distribution of genes in A and B compartments for both pre-pro-B and pro-B cell types (***$P < 0.001$). (**C**) A and B compartments that are defined by PC1 were integrated with active methylation marks (H3K4me1, H3K4me3 and H3K9/14ac). Normalized heat maps were generated by employing Matrix2png. Rows represent individual chromosomes, whereas the columns represent normalized count of respective methylation mark. (**D**) Comparative analysis of transcript levels of genes, based on RNA-Seq, present in A and B compartments (***$P < 0.001$) for both pre-pro-B and pro-B cells. (**E**) Comparative analysis of transcript abundance of genes that relocate from B to A compartment (left panel) and A to B compartment (right panel) during differentiation of pre-pro-B cells into pro-B cells (***$P < 0.001$). (**F**, **G**) Iteratively corrected contact count matrices derived from genomic regions comprising *Satb2* (chr1) and *Satb1* (chr17) for both pre-pro-B and pro-B cells. The PC1 values indicate the chromatin state of respective genomic loci. Dotted boxes represent genomic regions of *Satb2* and *Satb1*.

played higher transcript levels in pro-B cells (Figure 1E, left panel). For instance, *Satb2* (Figure 1F), *Tead1*, *Pou2af1* and *Tlr4* that are essential for B cell development (51,52) are re-localized from the B compartment to the A compartment during pre-pro-B to pro-B cell transition (Supplementary Table S1). Likewise, genes that relocate from the A compartment to the B compartment displayed lower transcript levels in pro-B cells (Figure 1E, right panel). Notably, genes that are associated with multipotent progenitors such as *Satb1* (Figure 1G), *cKit* and *Cd34* as well as key alternate lineage determinants such as *Gata3*, *Zbtb16*, *Klf4*, *Vav3* and *Sox6* are found to be relocated to the B compartment in pro-B cells (Supplementary Table S2). In comparison with pre-pro-B cells, a significant number of genes within the chromosomes 10, 11 and 16 switch from the B compartment to the A compartment. Similarly, genes that are located in chromosomes 6 and 7, switch from the A compartment to the B compartment in pro-B cells (Supplementary Figure S5). Interestingly, our studies reveal that majority of functionally important B-lineage-specific genes (*Ebf1*, *Pax5*, *Foxo1*, *IRF4*, *IRF8*, *Cd79a*, *Cd79b* and *Cd19*) are localized in A compartments in both cell types. However, some of the key alternate lineage genes (*Gata3*, *Zbtb16*, *Klf4*, *Vav3* and *Sox6*) switch to the B compartment in pro-B cells. Thus, these observations indicate that relocalization of alternate lineage genes into B compartments is closely associated with their transcriptional repression. Collectively, our studies demonstrate that switching of selective genomic loci between A and B compartments is closely associated with the B lineage-specific gene expression pattern. However, these studies cannot rule out the possibility that chromatin relocalization and its associated changes may be a result of alteration of transcription.

Global analysis of topologically associating domains (TADs) during B cell specification

At sub-megabase level, A and B compartments of chromatin are organized into dense and contiguous self-interacting regions termed topologically associating domains, TADs (10). These chromatin domains have been proposed to be stable and conserved across cell types, yet their intrinsic chromatin interactions were found to be varying (14). This raised a possibility that changes in the interaction patterns within TADs may serve as a framework for differential gene activity and contribute to the developmental progression of the cell. In order to capture the changes in chromatin structure within these domains in pre-pro-B and pro-B cells, we employed domain caller software (10) to identify TADs from iteratively corrected relative contact probability matrices generated at 40 kb resolution. Our analyses revealed that the genome of pre-pro-B cells is partitioned into a total of 2,008 TADs, whereas the genome of pro-B cells comprised of 1,810 TADs with a total genomic occupancy of 90.74% and 89.50%, respectively. Strikingly, we found that the median TAD size is higher in pro-B cells (920 kb) as compared to pre-pro-B cells (800 kb). Collectively, these studies provide the first indication that the structural organization of TADs may be subjected to alterations during developmental transition from pre-pro-B to pro-B cell stage.

To gain further insights into changes in structural organization of the chromatin at the sub-megabase level, we cross-compared the TADs between two cell types (pre-pro-B and pro-B cells) based on their linear genomic position. In line with the previous reports (10), a substantial number of TADs, 1,023 (pre-pro-B cells: 50.9%, pro-B cells: 56.5%) were found to be stable, in both cell types (Supplementary Figure S6). The remaining TADs (pre-pro-B: 985, pro-B: 787) exhibited re-organization in terms of their genomic positions and were categorized as 'dynamic'. It is possible that the stable TADs may maintain persistent chromatin interactions and thus account for uniform gene activity between two cell types. Alternatively, chromatin regions within these stable TADs may be subjected to epigenetic modifications and concomitant changes in intra-molecular interactions, resulting in cell-type specific gene expression pattern. To investigate these possibilities, we have compared the transcript levels of genes present within the stable TADs between pre-pro-B and pro-B cells and observed significant differences in their activities. This differential gene expression pattern may possibly be due to alterations in intrinsic chromatin interaction landscape. To examine this, we have calculated Aggregation Preference (AP), a parameter that quantitatively measures interaction patterns of TADs (53). During this analysis, the local high-frequency chromatin interactions, violating the distance-dependence decay principle, were measured and segregated according to their spatial aggregation by employing DBSCAN. The weighted density of clustered groups, defined as Aggregation Preference (AP), was used to quantitatively measure interaction patterns within each TAD. As expected, we found that TADs with higher AP values were comprised of high density chromatin interaction blocks in both cell types. Accordingly, TADs with higher AP values were enriched with active methylation marks (H3K4me1, H3K4me2 and H3K4me3) and displayed higher nascent transcript levels. On the other hand, TADs with low AP values displayed sparse chromatin interactions and were found to be depleted with active epigenetic marks (Supplementary Figure S7A). Furthermore, permissive TADs displayed higher AP values as compared to repressive TADs (Supplementary Figure S7B). Thus, AP values define transcriptional status and may serve as an appropriate measure of functional activity of TADs.

Interestingly, of the 1,023 stable TADs, a majority (867, 85%) of them displayed similar AP values (<0.2) between pre-pro-B and pro-B cells, indicating that the cumulative number of chromatin interactions with these TADs are comparable. Thus the differential gene expression pattern observed within stable TADs between the two cell types may possibly be attributed to the combinatorial changes in their promoter and *cis*-regulatory interactions. To test this, we first identified statistically significant ($P < 0.05$) chromatin interactions in both cell types by implementing the spline-fit model (45) (Supplementary materials and methods). Next, these significant interactions were integrated with genome-wide epigenetic marks (H3K4me3, H3K4me1 and H3K4me2) to identify potential promoters and *cis*-regulatory interactions (32). Only those promoters located within close proximity (± 2.5 kb) of transcription start sites (TSS) (Supplementary Figure S8) and the *cis*-regulatory elements located ≥ 1 kb away from putative promoters, were

considered for further analysis. In total, we have identified 30,150 and 46,263 potential interactions involving promoters and *cis*-regulatory elements in pre-pro-B and pro-B cells, respectively. We have mapped these promoter-*cis*-regulatory interactions to the stable TADs (867, 85%) with similar AP values. From these analyses, we found that 6,678 and 9,468 *cis*-regulatory interactions were associated with stable TADs in pre-pro-B and pro-B cells, respectively. Of these interactions, 1,715 promoters were found to be common in pre-pro-B and pro-B cells, whereas 1,147 (48.8%) and 952 (57.1%) promoters were found to be unique in pre-pro-B and pro-B cells, respectively. The majority of these promoters (common and unique) were found to interact with developmental stage-specific enhancers (Figure 2A). For instance, *Polg2*, which is highly expressed in pre-pro-B cells and *Cd79b*, which is induced at pro-B cell stage, are both located in a stable TAD with similar AP values. Interestingly, we observed that *Polg2* interacts with multiple enhancers (3) in pre-pro-B cells, whereas no such promoter and *cis*-regulatory interactions were captured in pro-B cells. Conversely, *Cd79b* promoter interacts with multiple enhancers (4) in pro-B cells, while only one such interaction was observed in pre-pro-B cells (Figure 2B). These results suggest that although a substantial number of TADs are stable with respect to the genomic position, their intrinsic chromatin interactions involving *cis*-regulatory elements are dynamic. These intrinsic changes in chromatin interactions may have a limited effect on the structural maintenance of the TAD, but may be critical for sustaining the cell type-specific gene expression pattern.

Structural reorganization of TADs corresponds to changes in *cis*-regulatory interaction landscape

While a substantial number (1,023) of TADs are constant in both pre-pro-B and pro-B cell types, a considerable number of TADs (pre-pro-B:985, pro-B:787) were found to be altered as indicated by changes in their genomic positions. We classified these altered TADs into two groups: unique and merged. Unique TADs are defined as those present in pre-pro-B cells but not in pro-B cells and *vice versa*. We propose that unique TADs (pre-proB:100, pro-B:65) may have been generated as a result of increased local genomic interactions to facilitate cell type-specific gene expression pattern (Supplementary Tables S3 and S4). Consistent with this assumption, key alternate lineage genes like *Ccl3*, *Serpini1* and *Vav3* that are highly expressed in multipotent progenitors, were found to be associated with TADs in pre-pro-B cells. On the contrary, in pro-B cells, these genes are located in the boundary regions (Supplementary Figure S9A). Besides the unique TADs, we observed that few larger TADs in pre-pro-B cells (110) partitioned into two or more minor TADs in pro-B cells. Conversely, two or more minor TADs in pre-pro-B cells coalesce into a larger 'merged' TAD in pro-B cells (183). We propose that merged TADs may have been formed as a result of increased inter-TAD interactions in pro-B cells (Figure 2C). Accordingly, the normalized contact frequency of inter-TAD regions of minor TADs in pre-pro-B cells is significantly lower as compared to the counter regions of merged TADs in pro-B cells (Supplementary Figure S9B). These observations are in line with increased median TAD size (920 kb) of pro-B cells, as compared to the size of pre-pro-B cells (800 kb). Correspondingly, we observed a significant increase in the number (pre-pro-B cells: 30,150, pro-B cells: 46,263) as well as in the median distance (pre-pro-B cells: 298 kb, pro-B cells: 330 kb) between promoter-*cis*-regulatory interactions in pro-B cells. This raises the possibility that inter-TAD promoter-*cis*-regulatory interactions may contribute for re-organization of TADs. To test this, we have mapped promoter-*cis*-regulatory interactions of merged TADs in pro-B cells and compared to their counter TADs in pre-pro-B cells. We found a significant increase in inter-TAD promoter-*cis*-regulatory interactions (2,600) in pro-B cells as compared to those (1,570) in pre-pro-B cells (Figure 2C). Taken together, these analyses provide insight into the dynamic re-organization of TADs which is closely associated with changes in the *cis*-regulatory interaction landscape during developmental transition from pre-pro-B to pro-B cell stage.

To rigorously demonstrate the dynamic organization of TADs observed between pre-pro-B and pro-B cells, we employed two distinct strategies. First, we used Directionality Index (DI), which quantitatively measures the 'interaction bias' of a given genomic region (10), as a parameter to detect the structural variations of TADs between the two cell types. Comparative DI analyses revealed that stable TADs display significantly higher correlation as compared to the dynamic TADs (merged and unique) (Figure 3A). These observations suggest that, unlike stable TADs, dynamic TADs display dramatic structural alterations. Second, we built 3D models of merged and stable TADs to determine the changes in position order of chromatin using AutoChrom3D (44). Compared to conventional 3D modeling methods (54–57), AutoChrom3D employs a novel sequencing-bias-relaxed parameter to derive 3D chromatin models. Next, we compared the spatial distance between start and end regions of merged TADs in pro-B cells with their counter regions in pre-pro-B cells. We reason that, in pro-B cells, the ends of a merged TAD should be in close spatial proximity compared to their counter regions in pre-pro-B cells. Consistent with this supposition, we found that the spatial distance was significantly lower in pro-B cells as compared to pre-pro-B cells (Figure 3B, left panel; Supplementary Figure S10A–D). In contrast, no significant difference in the spatial distance was observed for stable TADs (Figure 3B, right panel; Supplementary Figure S10E). Collectively, these analyses demonstrate that 3D models reflect the changes in 2D interaction maps. To validate these results, we performed 3C-qPCR (Chromosome Conformation Capture) for a merged TAD (Chr12:69720000–71160000) found in pro-B cells formed as a result of coalescence of three minor TADs in pre-pro-B cells. The spatial distance between start and end of this TAD is found to be lower in pro-B cells (2.41 Å) compared to pre-pro-B cells (8.00 Å) as shown by AutoChrom3D. Correspondingly, our 3C experiment using primers close to start (+1615 bp) and end regions (–7004 bp) of merged TAD revealed higher cross-linking frequency in pro-B cells than pre-pro-B cells (Figure 4) (Supplementary Table S5). These results support our hypothesis that TADs undergo dynamic structural alterations as a result of changes in chromatin in-

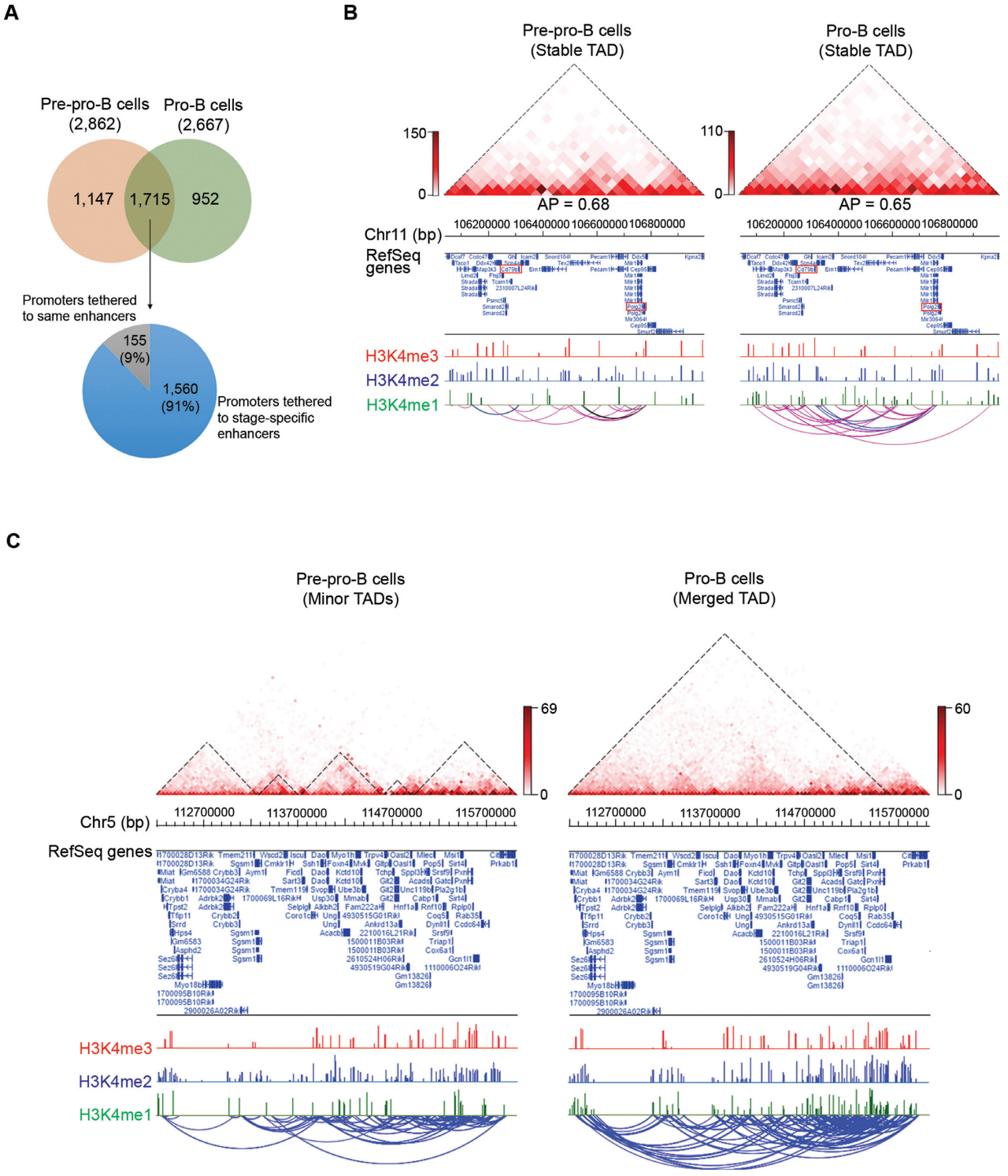

Figure 2. TADs are dynamic and undergo structural reorganization during early B cell development. (**A**) Venn diagram indicating the number of promoters interacting with *cis*-regulatory elements that are present in the stable TADs with similar AP values in both pre-pro-B and pro-B cells. The pie chart represents the number of common promoters tethered to same (grey) or cell type-specific (blue) enhancers. (**B**) Comparative analysis of promoter-*cis*-regulatory interactions between stable TADs with similar AP values spanning a genomic region (106.04–107 Mb) of chromosome 11. TADs are mapped with active epigenetic marks, H3K4me3 (enriched at promoter regions), H3K4me1 and H3K4me2 (enriched at enhancers) as determined by ChIP-Seq in both pre-pro-B and pro-B cells. TADs were defined by domain calling approach and are highlighted by dotted lines. The genomic positions of promoter-*cis*-regulatory interactions within the TADs are represented by arcs. *Polg2* (black) and *Cd79b* (blue) interactions are highlighted (**C**) Comparative analysis of promoter-*cis*-regulatory interactions between merged TAD (pro-B cells) and its counter TADs (pre-pro-B cells) spanning the genomic region (112.20–115.64 Mb) of chromosome 5. TADs are demarcated by domain calling approach and highlighted by dotted lines. TADs were mapped with active epigenetic marks: H3K4me3 (for promoters) and H3K4me1 and H3K4me2 (for enhancers) as determined by ChIP-Seq in both pre-pro-B and pro-B cells. Promoter-*cis*-regulatory interactions are represented by blue arcs.

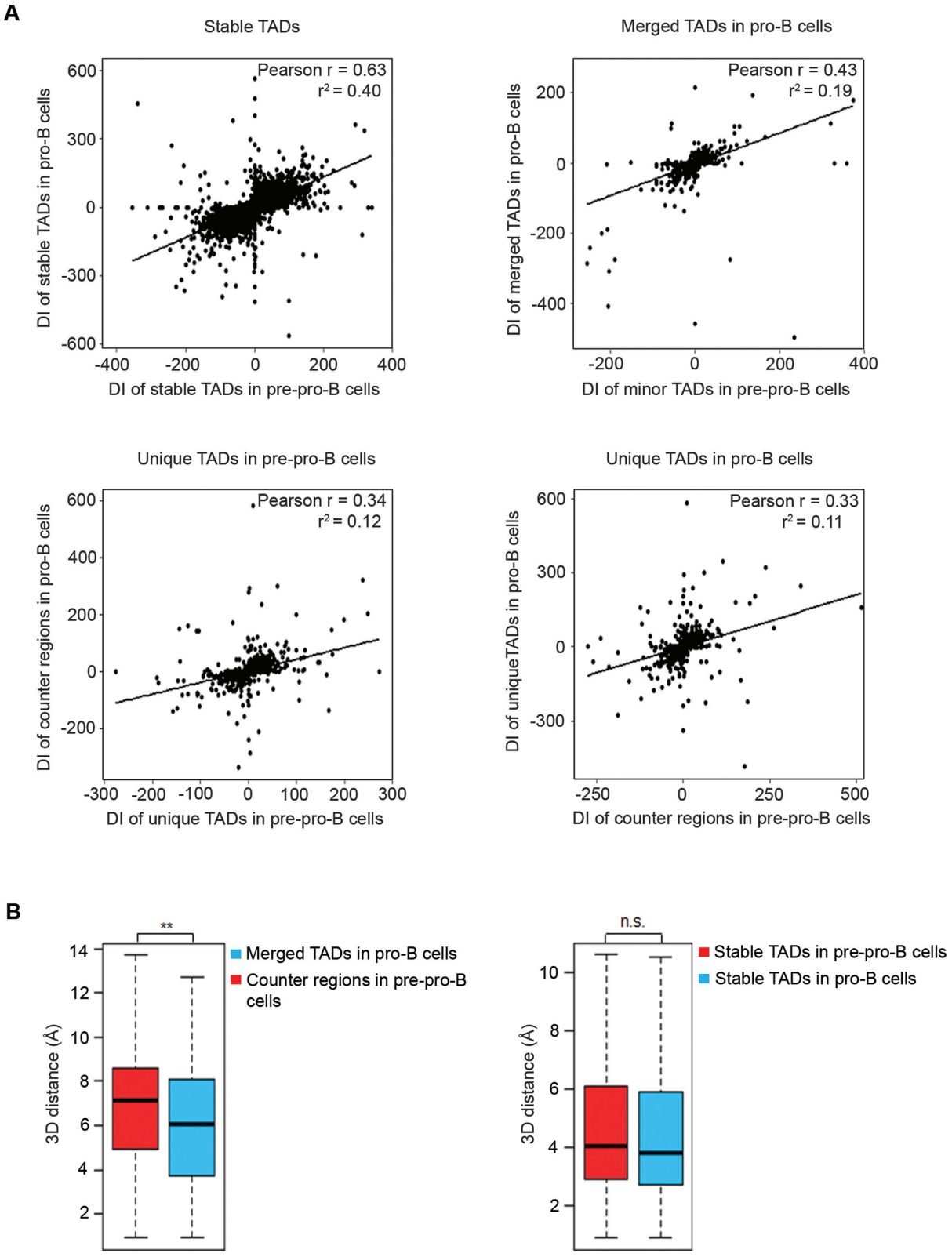

Figure 3. Comparative analysis of structural organization of TADs between pre-pro-B and pro-B cells. (**A**) Pearson's Correlation Coefficient for directionality index (DI) calculated for stable as well as dynamic TADs (merged and unique) between pre-pro-B and pro-B cells. (**B**) Genome-wide comparative analysis of 3D spatial distances between start and end regions of merged TADs in pro-B cells and their counter regions in pre-pro-B cells (**$P < 0.01$) (left panel). Similar analysis of 3D spatial distances for stable TADs in both pre-pro-B and pro-B cells (n.s. = not significant) (right panel).

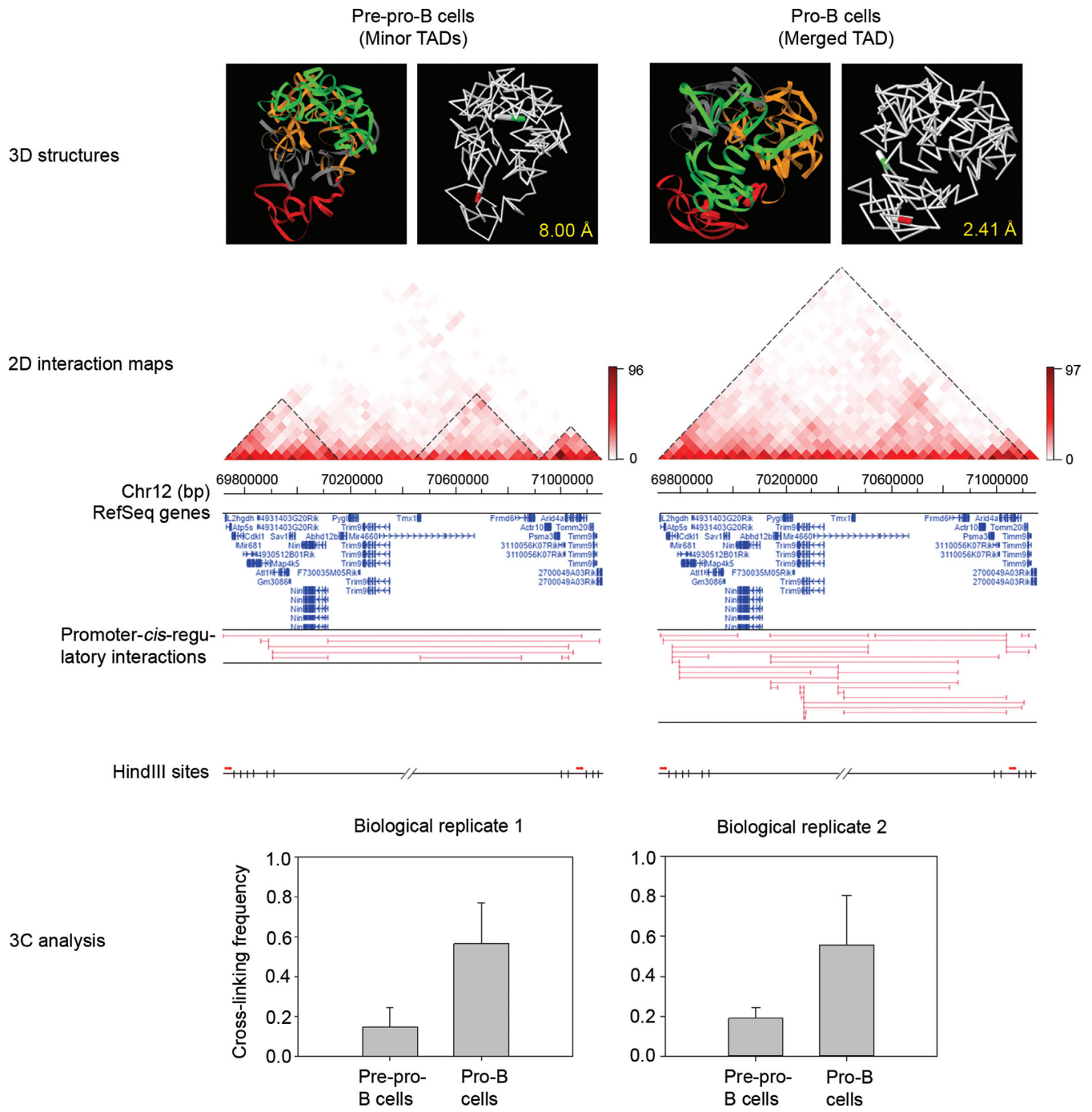

Figure 4. Validation of TADs reorganization by 3C analysis. Comparative analysis of 3D spatial distances and promoter-*cis*-regulatory interactions between merged TAD (pro-B cells) and its counter TADs (pre-pro-B cells) spanning the genomic region (69.72–71.16 Mb) of chromosome 12. 3D models generated by AutoChrom3D were colored distinctly based on minor TADs in pre-pro-B cells and the same color code is given for corresponding genomic regions of merged TAD in pro-B cells. The start and end regions of merged TAD in pro-B cells and its counter regions in pre-pro-B cells are highlighted by green and red respectively in the back bone 3D structure and the spatial distance between these regions is indicated in Å units. 3D models were generated at 8 kb resolution (upper panel). TADs were demarcated by domain calling approach and highlighted by dotted lines. TADs were mapped with active epigenetic marks H3K4me3 (for promoters), and H3K4me1 and H3K4me2 (for enhancers) as determined by ChIP-Seq in both pre-pro-B and pro-B cells. Promoter-*cis*-regulatory interactions are represented by horizontal lines (middle panel). 3C analysis of interaction frequency between ends of merged TAD (Chr12: 69.72–71.16 Mb) in pro-B cells and its counter regions in pre-pro-B cells. HindIII restriction sites are shown above the 3C plots. The location of primers used for measuring cross-linking frequency is indicated by red arrows (lower panel).

teraction patterns that may be important for transcription regulation.

To investigate if our findings could be extrapolated to other cell types, we compared the structural organization of TADs in mESC and cortex cells using publicly available data (10). Similar to what we have observed with our cell types, the total number of TADs (mESC-2085, cortex-1519) as well as their median size (mESC-880 kb, cortex-1.3 Mb) differed between mESC and cortex cells. Moreover, the comparative analysis of DI and relative contact probabilities suggests that TADs undergo structural reorganization between mESC and cortex (Supplementary Figure S11A and B). In concordance with our data for pro-B cells, we noticed that cortex cells have more number of merged TADs (269) than the mESCs (95). In comparison with the pluripotent cells (pre-pro-B and mESC), the observed increase in number of merged TADs and the associated increase in the average size of TADs in differentiated cells (pro-B and cortex) can be attributed to their compact chromatin organization (58,59). These findings are further supported by an increase in long-range interactions in pro-B cells compared to those in pre-pro-B cells. Consistently, a recent study suggests that the increase in TAD size as well as long-range interactions in sperm cells may be due to the dense packaging of its genome (60). Collectively, these findings demonstrate that the differences in the TAD organization between various cell types are dependent on the differences in their long-range interactions and chromatin compaction.

TADs constitute structural frameworks for coordinated gene expression

Next, we sought to determine the relationship between the structural organization of TADs and the differential gene expression pattern in pre-pro-B and pro-B cells. For this, we integrated TAD regions with PC1 values at 100 kb resolution to assess their chromatin state. Interestingly, we found that majority of the TADs are either transcriptionally permissive or repressive. However, a small percentage of TADs are comprised of both permissive and repressive chromatin regions and are referred as uncharacterized (Figure 5A). As expected, genes present in permissive TADs, displayed higher nascent transcript levels compared to those in repressive TADs (Supplementary Figure S12). These observations suggest the possibility that TADs serve as structural frameworks for coordinated regulation of genes. To rigorously demonstrate this, we calculated Pearson's correlation coefficient (PCC) for all possible gene pairs, using publicly available genome-wide expression data sets for pre-pro-B and pro-B cell types (61). PCC for pre-pro-B cells was calculated by comparing microarray measurements of hematopoietic stem cells (HSCs) and CLPs as they mimic pre-pro-B cells. Likewise, for pro-B cells, PCC was calculated using microarray measurements of CLPs and pro-B cells (pro-B.FrBC.BM). These analyses revealed that in both pre-pro-B and pro-B cells, genes within a given TAD exhibit significantly higher correlation values ($P < 0.001$) in relation to the genes that are present in other TADs, indicating that TADs represent co-regulated sub-units of the genome (Figure 5B–D). We note that such coordinated regulation of genes within the TADs facilitate activation/repression of gene clusters in a cell type-specific manner. For instance, the HOXA gene cluster (Chromosome 6), which is localized in a single stable TAD is transcriptionally active in pre-pro-B cells, whereas the same cluster is found to be transcriptionally inactive in pro-B cells; suggesting that TADs not only serve as fundamental sub-units for coordinate regulation of genes, but they also provide a framework to sustain lineage-specific gene expression pattern.

The *cis*-regulatory interaction landscape undergoes rewiring during B cell fate commitment

Although, it is well established that promoter-*cis*-regulatory interaction landscape determines a lineage-specific gene expression pattern (15,62), much less is known about the genome-wide composition of these interactions during B cell development. From *in situ* Hi-C analyses, we identified a total of 31,190 and 47,711 potential promoter-tethered interactions in pre-pro-B and pro-B cells, respectively. As expected, genes whose promoters are involved in *cis*-regulatory interactions showed significantly higher expression levels than the genes that are not involved in any such interactions (Supplementary Figure S13A). The majority of the promoter-*cis*-regulatory interactions (83.5% in pre-pro-B cells and 79% in pro-B cells) are within the range of 1 Mb with a median value of 298 kb and 330 kb in pre-pro-B and pro-B cells, respectively (Supplementary Figure S13B). We have classified these promoter-tethered interactions into three basic groups: intergenic (promoter–promoter), extragenic (promoter–enhancer) and intragenic (promoter–gene body). We observed 8,410 and 10,556 intergenic interactions in pre-pro-B and pro-B cells, respectively. Similarly, we have captured 1,040 and 1,448 intragenic interactions in pre-pro-B cells and pro-B cells, respectively. Strikingly, we observed a significant increase in promoter–enhancer interactions in pro-B cells compared pre-pro-B cells. We found a total number of 21,740 promoter–enhancer (extragenic) interactions involving 8,096 promoters and 10,637 enhancers in pre-pro-B cells, wherein each promoter on average interacts with 2.68 enhancers. In the case of pro-B cells, 35,707 promoter–enhancer (extragenic) interactions involve about 9,424 promoters and 14,904 enhancers, wherein each promoter on average interacts with 3.79 enhancers. Among 9,424 promoters captured in pro-B cells, 6,331 (67.2%) promoters were also captured in pre-pro-B cells. Interestingly, 5,101 (80.5%) of common promoters were found to interact with cell type-specific enhancers and only 1,230 (19.5%) promoters share common enhancers (Supplementary Figure S13C and D; Supplementary Table S6). These results reveal that during B lineage-specification, the promoter–enhancer interaction landscape undergoes extensive rewiring. We also note that in pre-pro-B cells, nearly 13.3% of enhancers interact with more than three promoters. Likewise, in pro-B cells, ∼19% of enhancers were found to interact with more than three promoters (Supplementary Figure S13E and F). These observations support the assertion that multiple genes interacting with the same enhancer may be co-expressed (63).

Furthermore, our data revealed that in pre-pro-B cells, nearly 25.4% of promoters interact with only one enhancer,

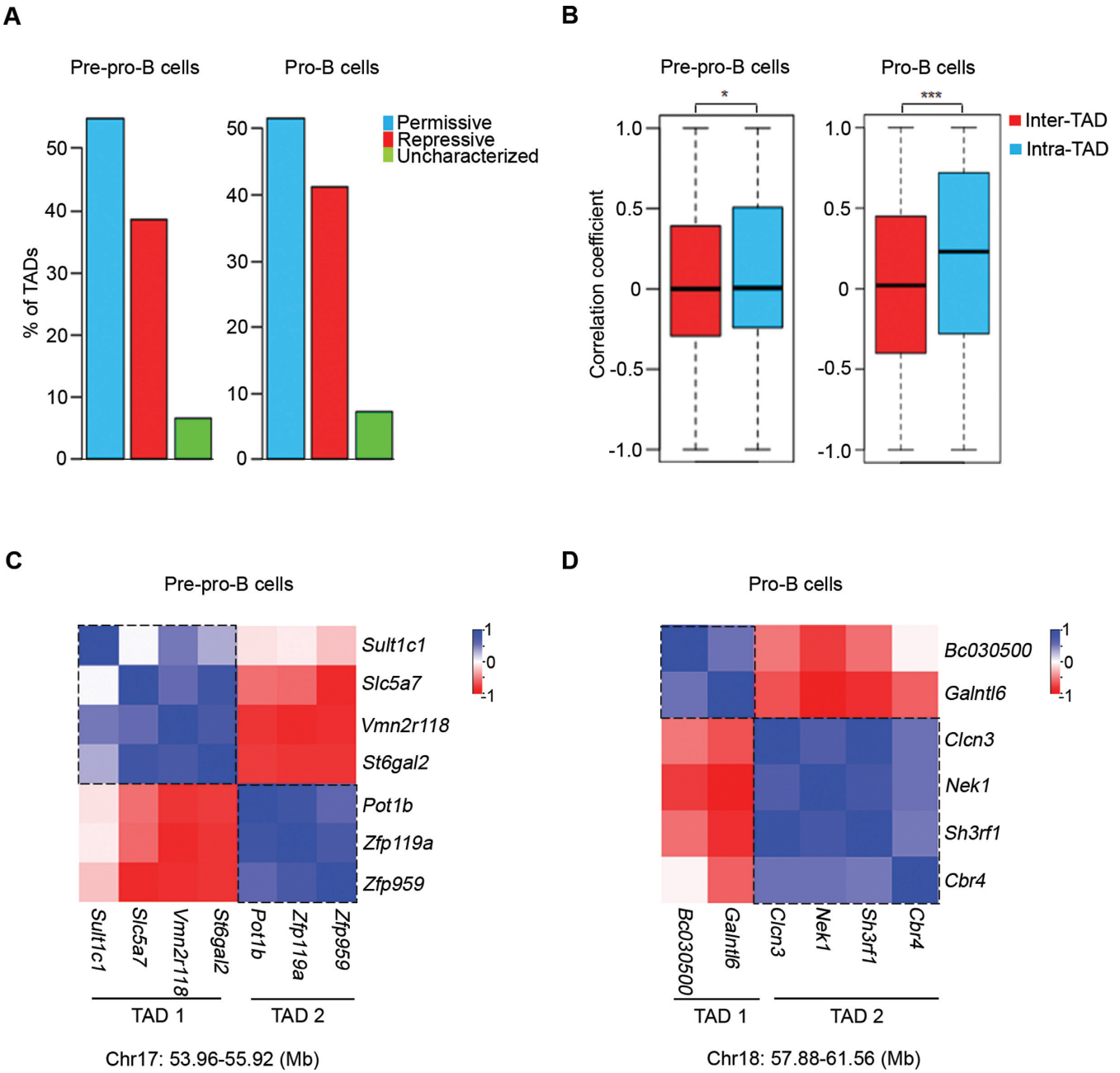

Figure 5. TADs represent chromatin subunits of coordinate gene expression. (**A**) Histogram representing chromatin state of TADs as defined by Principal Component Analysis (PCA). (**B**) Comparative analysis of Pearson's correlation coefficient (PCC) for gene-pairs present in the same TAD against gene-pairs present in other TADs (*$P < 0.05$, ***$P < 0.001$) in pre-pro-B and pro-B cells. PCC was calculated by considering microarray measurements of hematopoietic stem cells (HSCs), common Lymphoid Progenitors (CLPs) and pro-B cells (pro-B.FrBC.BM). (**C**, **D**) Representation of Pearson's correlation coefficient for gene-pairs present in two different TADs spanning genomic region (53.96–55.92 Mb) of chromosome 17 and (57.88–61.56 Mb) of chromosome 18 for pre-pro-B and pro-B cells respectively. Blue represents positive correlation whereas red represents negative correlation. Each dotted box represents an individual TAD.

while remaining promoters interact with two or more enhancers. Likewise, in pro-B cells, the majority of the promoters (81.1%) were found to interact with two or more enhancers (Supplementary Figure S13G and H). To test whether transcriptional activity of a promoter depends on the number of its *cis*-regulatory interactions, we assessed the transcript levels of the corresponding genes. Interestingly, these studies depicted a positive correlation between gene expression and a number of *cis*-regulatory interactions in both cell types (Figure 6A). Next, we sought to determine if loss or gain of these interactions induce differential gene expression patterns. For this, we compared the *cis*-regulatory interaction landscape of genes that show +10-fold differential expression between pre-pro-B and pro-B cells. The analysis showed that the expression pattern is closely associated with an increase in the number of *cis*-regulatory

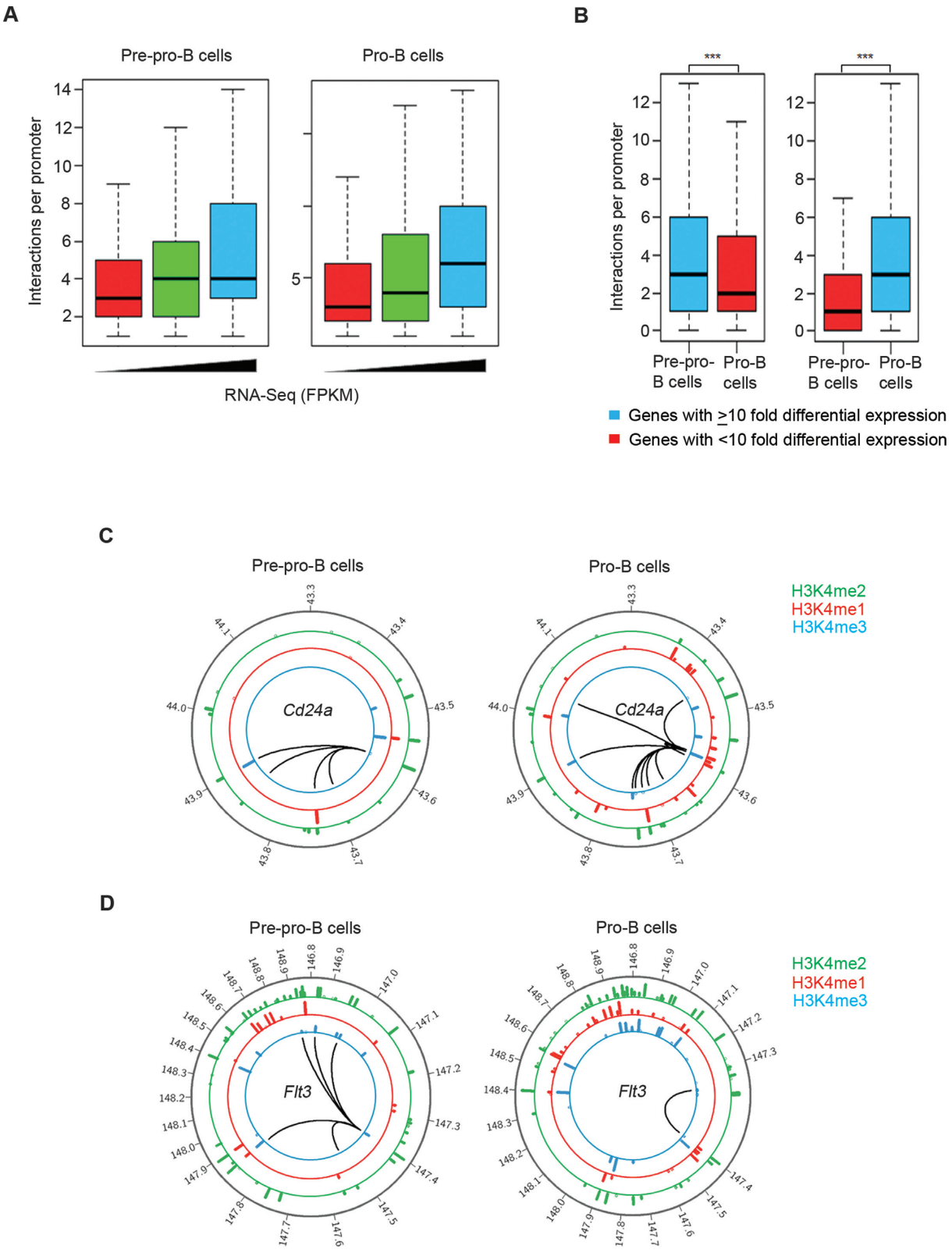

Figure 6. *cis*-regulatory interaction landscape determines differential gene expression pattern. (**A**) Box plots showing the relation between the number of *cis*-regulatory elements that are interacting with promoters and their expression levels, in pre-pro-B and pro-B cells. Transcript levels of genes were measured by RNA-Seq. (**B**) Box plots representing comparative analysis of promoter-*cis*-regulatory interactions for a set of genes with ≥10-fold differential expression in pre-pro-B cells (right panel) and in pro-B cells (left panel) (***$P<0.001$). (**C**, **D**) Circos plots showing promoter-*cis*-regulatory interactome of *Cd24a* (Chr11:43.3–44.1 Mb) and *Flt3* (Chr5:14.68–14.89 Mb) in pre-pro-B (left panel) and pro-B cells (right panel). Black arcs represent promoter-*cis*-regulatory interactions.

interactions (Figure 6B; Supplementary Tables S7 and S8). For instance, *Cd24a*, which is highly induced in pro-B cells, interacts with 13 *cis*-regulatory elements, whereas, it is involved in only four such interactions at pre-pro-B cell stage, where its expression is considerably low (Figure 6C). Correspondingly, the genes: *Flt3* and *Ccl3*, that are important for maintenance of MPPs and differentiation of T-cells, respectively, are transcriptionally active at the pre-pro-B stage. These genes were found to be involved in more number of *cis*-regulatory interactions (*Flt3*:6, *Ccl3*:8) in pre-pro-B cells as compared to pro-B cells (*Flt3*:1, *Ccl3*:0) (Figure 6D). The examples depicted here demonstrate the prevalence of dynamic promoter-*cis*-regulatory interactions across B cell developmental stages. To rigorously validate these findings, we carried out 3C analysis of promoter–enhancer interactions of *Ccl3* locus in pre-pro-B and pro-B cells (Supplementary materials and methods) (Supplementary Table S5). We observed that the interaction frequency between the *Ccl3* promoter and with its upstream enhancer (located 64 kb away) was higher in pre-pro-B cells compared to that in pro-B cells (Figure 7A). Correspondingly, the quantitative RT-PCR analysis revealed thirty-fold higher *Ccl3* transcript levels in pre-pro-B cells as compared to the levels in pro-B cells (Figure 7B). These analyses confirm that reinforcement of lineage-specific gene expression is contingent upon specificity and frequency of interactions between promoters and their *cis*-regulatory elements.

Ebf1 coordinates B cell specific *cis*-regulatory interaction landscape

A major goal of this study was to understand the molecular relationship between chromatin architecture and differential transcriptional cascade. It has been shown that Ebf1 is essential for induction of early B lineage gene expression program and targeted inactivation of Ebf1 results in a complete block prior to B cell commitment (28,41). This raises a possible role for Ebf1 in chromatin relocalization and establishment of B lineage-specific *cis*-regulatory interaction landscape. To test this, we scanned for highly specific and significant Ebf1 binding sites in *cis*-regulatory regions of genes that switched to A or B compartments in pro-B cells, using publicly available databases (Jasper, Homer and Uniprobe). From these analyses, we observed that Ebf1 and/or Pax5 bind to *cis*-regulatory sequences of differentially switched genes (65.3%). Although a subset of genes undergoes differential compartmentalization, this may not solely account for the induction of B lineage expression program. We propose that activation of B lineage-specific genes may be regulated at multiple levels including binding of lineage-specific transcription factors (E2A, Ebf1, Foxo1 and Pax5) to their target promoter-*cis*-regulatory interacting elements. To examine this, we integrated promoter-*cis*-regulatory interactions that are captured in pro-B cells with binding events of these factors. Importantly, we found that in pro-B cells, Ebf1 binds either alone or in combination with Pax5 to 5390 (57.2%) promoters and 7629 (51.2%) *cis*-regulatory elements that are involved in long-range interactions as defined by *in situ* Hi-C (Figure 8A). Consistent with these observations, binding of Ebf1/Pax5 at *cis*-regulatory elements of their target genes is positively correlated with increased expression levels (Figure 8B and C).

To rigorously demonstrate the induction of B lineage genes in response to Ebf1 and/or Pax5, we carried out genome-wide expression analysis following restoration of Ebf1 or Pax5 in *Ebf1*−/− progenitor cells (Supplementary materials and methods). As expected, Ebf1 and/or Pax5 induced a spectrum of genes associated with B cell identity, including those that are involved in pre-B and B cell receptor signaling, antigen presentation, DNA recombination, and repair. Conversely, Ebf1 and/or Pax5 repressed a subset of genes that are involved in the development and function of natural killer (NK), dendritic and T cells (Figure 8D and E). Integration of *in situ* Hi-C interactome (promoter-*cis*-regulatory interactions) with microarray data sets revealed that the genes that are upregulated (>2-fold) in response to Ebf1 (124, 39.7%) or Pax5 (231, 44%) or both (37, 72.5%) are involved in long-range interactions (Figure 8F). We note that activation of these genes could be due to direct binding of Ebf1 and/or Pax5 to their respective promoter or distant regulatory elements that are brought in close proximity by looping-out of intervening DNA sequences. To determine this, we scanned the promoter and their corresponding *cis*-regulatory elements that are upregulated in response to Ebf1 (161) and/or Pax5 (268) for their binding. *De novo* motif analyses revealed that Ebf1 binds to either promoter regions (17.6%) or distant *cis*-regulatory elements (45.3%) or both (36.9%). Similarly, Pax5 binds either at promoter regions (19.2%) or distant *cis*-regulatory elements (54.1%) or both (26%) (Figure 8G). We note that both Ebf1 and Pax5 co-bind to a number of key B lineage genes such as *Cd19*, *Cd24a*, *Socs3* and *Dtx1*. Taken together, *in situ* Hi-C analyses in combination with genome-wide expression analysis and DNA occupancy studies, we demonstrate that activation of B lineage-specific genes is associated with changes in long-range interactions and many of these genes are potentially regulated by lineage-specific transcription factors, Ebf1 and Pax5.

DISCUSSION

Precise and coordinated control of gene expression is important for the cell fate determination of multipotent progenitors (1,58,64,65). Recent studies indicate that structural organization of the genome in 3D nuclear space is closely associated with modulation of transcriptional activity and establishment of cell type-specific gene expression program, indicating a potential relationship between nuclear architecture and mechanistic control of transcription. 3C-based studies indicate that genome is organized in a hierarchical manner: folding of chromatin loops, TADs, and large-scale compartments (1,2,10,14,47). However, the comprehensive understanding of how multilayer organization of chromatin regulates cell-type-specific transcriptional activity remains unclear. Specifically, the following questions arise: Does chromatin relocalization precedes lineage commitment? Do chromatin domains undergo structural reorganization? What are the roles of lineage determinants during chromatin reorganization? We have attempted to address these questions by integrating genome-wide chromatin interaction data with epigenetic landscape and tran-

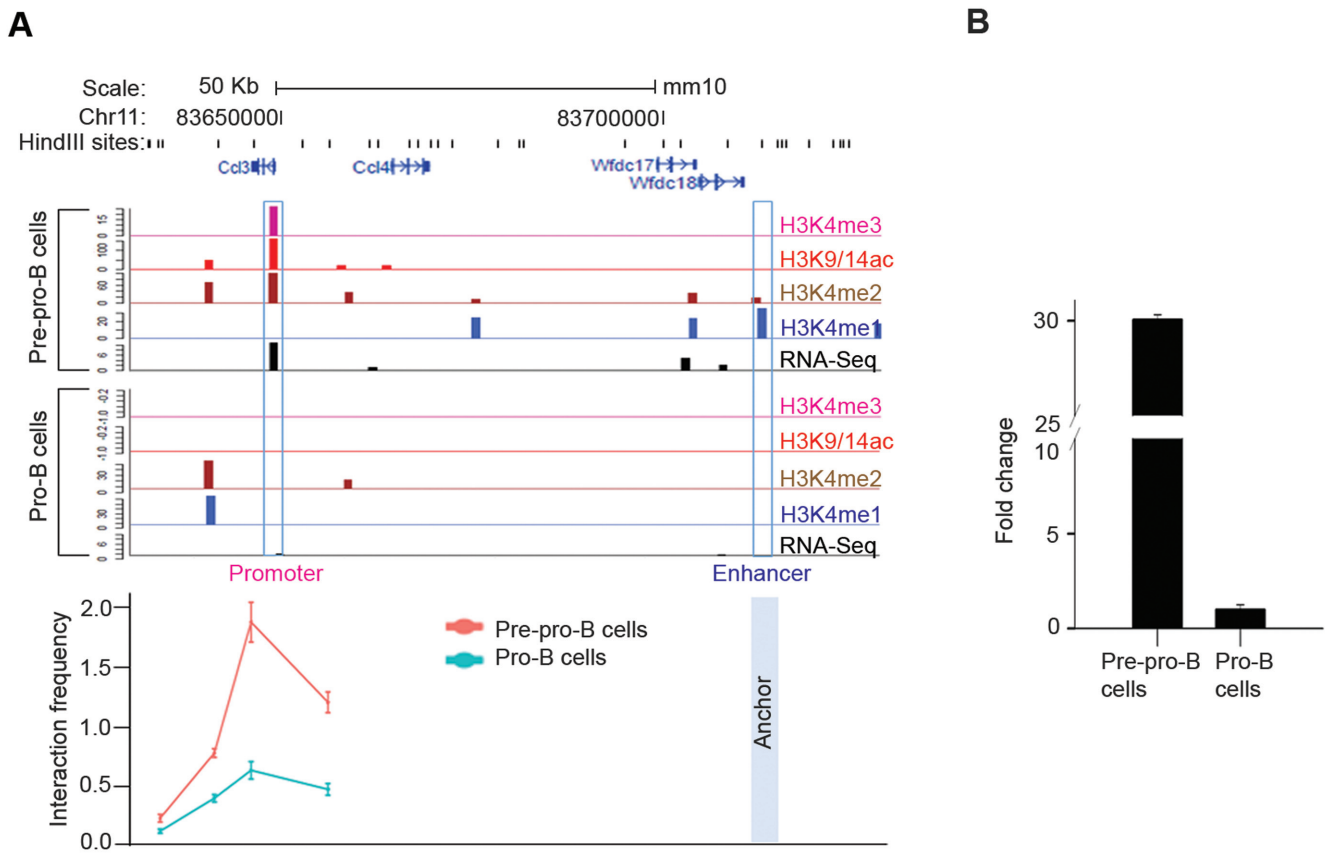

Figure 7. Validation of promoter–enhancer interactions by 3C analysis. (**A**) *Ccl3* locus, overlaid with various epigenetic marks as determined by ChIP-Seq in both pre-pro-B and pro-B cells (upper panel). Interaction frequency between promoter region of *Ccl3* and its distant putative enhancer located at 64 kb upstream of promoter in both pre-pro-B and pro-B cells (lower panel) was measured by 3C-qPCR analyses. Data are representative of two independent biological experiments (error bars, S.E.). (**B**) Relative transcript levels of *Ccl3* as measured by quantitative RT-PCR in pre-pro-B cells and pro-B cells. *Hprt* was used as endogenous control and values were normalized against pro-B cells as a reference control. Data are representative of two independent biological experiments (error bars, S.E.).

scription profiling of cells that represent two distinct stages (pre-pro-B and pro-B) of B cell development. In line with previous reports, our *in situ* Hi-C analyses revealed that chromatin is non-randomly organized into A and B compartments (1,2). We identified that a distinct set of genes switch between A and B compartments during the developmental transition from pre-pro-B to pro-B cell stage. For instance, genes that are important for B cell development including *Satb2*, *Tead1*, *Pou2af1* and *Tlr4* switch from the B to A compartment during the pre-pro-B to pro-B transition. Likewise, genes such as *Gata3*, *Klf4*, *Satb1* and *Zbtb16* that are important for disparate lineage differentiation programs localize to B compartment in pro-B cells, where they are silenced. In contrast, the majority of the downstream targets of these master regulators were found to be in A compartments in both cell types. These studies suggest that sequestering master regulators of alternate lineages into B compartments may ensure lineage-specification. These observations are further supported by a previous study (66) wherein Th2-specific regulator, GATA3, was found to be localized in the nuclear periphery in a transcriptionally inactive state in Th1 cells. However, the downstream targets of GATA3 (IL-2, IL-3, and IL-4) were retained in the permissive compartment in both Th1 and Th2 cells. Thus selective relocalization of lineage determinants appears to play an important role during developmental transition from a progenitor to lineage-committed state. However, the mechanistic details of how the genomic loci switch between A/B compartments is unclear. We propose that binding of transcription factors along with chromatin activation complexes enable the genomic loci to relocate from B to A compartment, whereas binding of polycomb group (PcG) proteins enable the loci to relocate from A to B compartments.

Our findings show that, although, a substantial number of TADs are stable between pre-pro-B and pro-B cells, chromatin interaction patterns particularly promoter-*cis* regulatory interactions within these TADs remodel to facilitate cell type-specific gene expression pattern. Strikingly, in addition to the stable TADs, we found a number of TADs that are dynamic and display structural alterations during B cell developmental transition. Specifically, we uncovered a set of unique TADs that are exclusively present either in pre-pro-B or pro-B cells, which contain genes that are selectively expressed at these stages. Additionally, we found a distinct set of merged TADs in pro-B cells, which were generated by coalesce of contiguous TADs present in pre-pro-B cells. This may account for the presence of fewer TADs with an increase in average TAD size and the corresponding

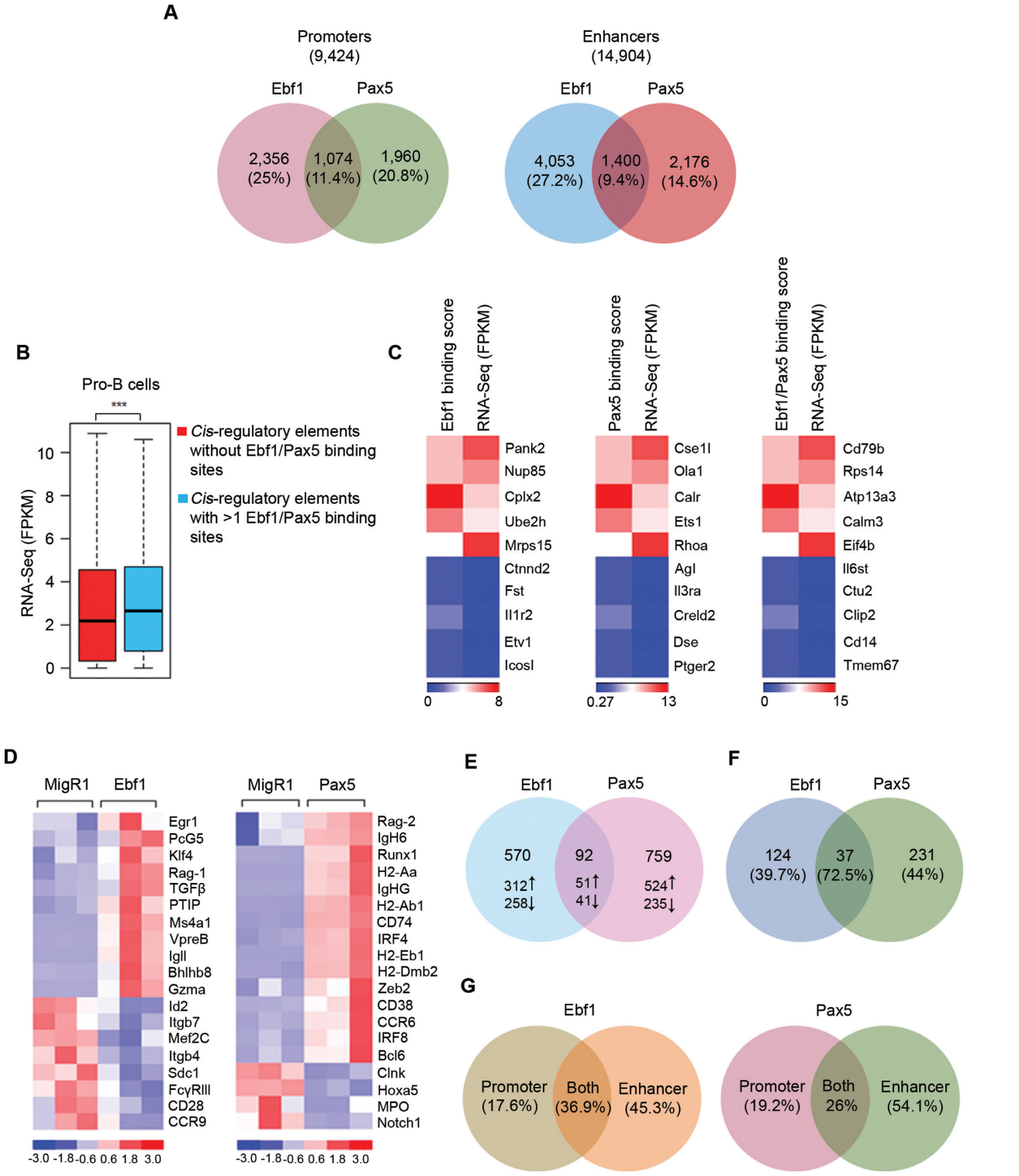

Figure 8. Ebf1 regulates B lineage-specific gene expression pattern in part by binding at *cis*-regulatory interacting elements. (**A**) Venn diagram representing the motifs of the TFs (Ebf1 and Pax5) binding at the promoters and respective *cis*-regulatory elements. (**B**) Box plot representing genome-wide comparative analysis of transcript levels of genes with or without Ebf1/Pax5 binding sites in the promoter-*cis*-regulatory interacting elements (***$P < 0.001$). (**C**) Heat maps showing correlation between transcript levels of genes and Ebf1 and/or Pax5 binding events in the promoter-*cis*-regulatory interacting elements. (**D**) Heat maps showing the genome-wide expression patterns of B lineage-specific genes (fold change ≥ 2; P-value < 0.05) obtained by microarray analysis of pre-pro-B cells (*Ebf1*−/− progenitors) transduced with Ebf1 or Pax5. (**E**) Venn diagram indicating the number of genes that are regulated by Ebf1 and/or Pax5. Up headed arrow represents activated, down headed arrow represents repressed genes. (**F**) Venn diagram representing the percentage of upregulated targets of Ebf1 and/or Pax5 that are involved in *cis*-regulatory interactions in pro-B cells. (**G**) Venn diagrams representing the percentage of Ebf1 or Pax5 target genes containing Ebf1 and/or Pax5 binding sites within the *cis*-regulatory sequences that are involved in long-range interactions.

gain of promoter-*cis*-regulatory interactions in pro-B cells as compared to pre-pro-B cells. We propose that the merging of TADs is due to an increase of inter-TAD interactions that are associated with epigenetic modifications and regulation of lineage-specific transcription factors. However, the precise molecular mechanisms that regulate the merging of TADs remains to be elucidated. A recent study (67), in drosophila cells demonstrates a dramatic increase in inter-TAD promoter-*cis*-regulatory interactions upon heat shock treatment, which was shown to be associated with a corresponding redistribution of architectural proteins from borders to inside of TADs. Moreover, we show that genes that are present in a given TAD exhibit higher correlation of expression compared to genes present in other TADs, supporting the notion that TADs provide a structural framework for coordinated gene regulation (12). Thus the studies presented here, provide new insights into the structural organization of TADs and their propensity to undergo alterations during developmental progression.

In this study, we provide a comprehensive map of long-range interactions between promoters and their corresponding *cis*-regulatory elements in pre-pro-B and pro-B cells. We found a significant expansion in promoter-*cis*-regulatory interaction landscape during the developmental transition from pre-pro-B to pro-B stage is evidenced by not only an increase in a total number of promoter–enhancer interactions, but also by the average increase in the ratio of promoters to enhancers. Additionally, we show that the transcript levels are positively correlated with the number of *cis*-regulatory interactions in both cell types, indicating that enhancer usage dictates transcriptional output. These studies corroborate the observations that modulation of gene activity is regulated by the enhancer landscape (48,68). Furthermore, our studies reveal that a significant number of promoter and enhancer elements that are engaged in long-range interactions contain Ebf1 and/or Pax5 binding sites. Many of these *cis*-regulatory elements are important for expression of developmentally regulated genes during B cell fate commitment. Accordingly, Ebf1 targeted genes displayed high levels of gene expression in pro-B cells. These results are further strengthened by the fact that, a subset of Ebf1 targeted genes, activated upon complementation of *Ebf1*−/− progenitors with Ebf1 or Pax5 were found to be involved in long-range interactions. Our analyses revealed novel molecular functions of Ebf1 and its potential role in the establishment of *cis*-regulatory interactions and activation of B lineage-specific genes. Nevertheless, the molecular mechanisms by which Ebf1 regulates these interactions remains to be understood. Binding of Ebf1 has been shown to recruit chromatin remodeling complexes like SWI-SNF to the *Cd79a* promoter there by increasing local chromatin accessibility for subsequent activation (69). Similar mechanisms may also operate for the establishment of long-range interactions, wherein binding of key lineage-determining transcription factors like Ebf1 to cell type-specific enhancers recruits a distinct combinatorial set of factors, thereby positioning the enhancers in close proximity to their target promoters. This raises the exciting possibility that Ebf1 may mediate lineage-specific long-range interactions crucial for B lineage gene expression program. Collectively, our studies demonstrate that dynamic alterations of chromatin organization associated with changes in *cis*-regulatory interactions that are regulated by lineage determinants impinge on the induction of lineage-specific gene expression. As our understanding of the detailed molecular mechanisms that govern the dynamics of higher-order chromatin organization continues to expand, the relationship between the 3D organization of the genome and lineage-specific gene expression will be better understood.

ACCESSION NUMBERS

In situ Hi-C and microarray data were deposited in the Gene Expression Omnibus database under accession number GSE85858.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank C-CAMP, Bangalore, for providing technical assistance with Illumina sequencing. High-performance computational analyses were carried out at the Centre for Modelling Simulation and Design (CMSD), University of Hyderabad and Centre for DNA Fingerprinting and Diagnostics (CDFD), Hyderabad.

FUNDING

Department of Biotechnology (DBT) [BT/PR13507/BRB/10/763/2010, BT/PR1825/AGR/36/679/2011]; Department of Science and Technology (DST) [SR/SO/BB-0071/2013]. R.B., A.D.Y., and S.N. are recipients of CSIR/UGC fellowships; D.P. is a DBT India Alliance fellow. Funding for open access charge: University of Hyderabad.
Conflict of interest statement. None declared.

REFERENCES

1. Gibcus,J.H. and Dekker,J. (2013) The hierarchy of the 3D genome. *Mol. Cell*, **49**, 773–782.
2. Lieberman-Aiden,E., van Berkum,N.L., Williams,L., Imakaev,M., Ragoczy,T., Telling,A., Amit,I., Lajoie,B.R., Sabo,P.J., Dorschner,M.O. *et al.* (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*, **326**, 289–293.
3. Dekker,J., Rippe,K., Dekker,M. and Kleckner,N. (2002) Capturing chromosome conformation. *Science*, **295**, 1306–1311.
4. Dostie,J., Richmond,T.A., Arnaout,R.A., Selzer,R.R., Lee,W.L., Honan,T.A., Rubio,E.D., Krumm,A., Lamb,J., Nusbaum,C. *et al.* (2006) Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome Res.*, **16**, 1299–1309.
5. 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.
6. Simonis,M., Klous,P., Splinter,E., Moshkin,Y., Willemsen,R., de Wit,E., van Steensel,B. and de Laat,W. (2006) Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nat. Genet.*, **38**, 1348–1354.
7. Fuxa,M., Skok,J., Souabni,A., Salvagiotto,G., Roldan,E. and Busslinger,M. (2004) Pax5 induces V-to-DJ rearrangements and locus contraction of the immunoglobulin heavy-chain gene. *Genes Dev.*, **18**, 411–422.

8. Kosak,S.T., Skok,J.A., Medina,K.L., Riblet,R., Le Beau,M.M., Fisher,A.G. and Singh,H. (2002) Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development. *Science*, **296**, 158–162.
9. Schneider,R. and Grosschedl,R. (2007) Dynamics and interplay of nuclear architecture, genome organization, and gene expression. *Genes Dev.*, **21**, 3027–3043.
10. Dixon,J.R., Selvaraj,S., Yue,F., Kim,A., Li,Y., Shen,Y., Hu,M., Liu,J.S. and Ren,B. (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*, **485**, 376–380.
11. Ibn-Salem,J., Kohler,S., Love,M.I., Chung,H.R., Huang,N., Hurles,M.E., Haendel,M., Washington,N.L., Smedley,D., Mungall,C.J. *et al.* (2014) Deletions of chromosomal regulatory boundaries are associated with congenital disease. *Genome Biol.*, **15**, 423.
12. Nora,E.P., Lajoie,B.R., Schulz,E.G., Giorgetti,L., Okamoto,I., Servant,N., Piolot,T., van Berkum,N.L., Meisig,J., Sedat,J. *et al.* (2012) Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature*, **485**, 381–385.
13. Nora,E.P., Goloborodko,A., Valton,A.L., Gibcus,J.H., Uebersohn,A., Abdennur,N., Dekker,J., Mirny,L.A. and Bruneau,B.G. (2017) Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization. *Cell*, **169**, 930–944.
14. Dixon,J.R., Jung,I., Selvaraj,S., Shen,Y., Antosiewicz-Bourget,J.E., Lee,A.Y., Ye,Z., Kim,A., Rajagopal,N., Xie,W. *et al.* (2015) Chromatin architecture reorganization during stem cell differentiation. *Nature*, **518**, 331–336.
15. Carter,D., Chakalova,L., Osborne,C.S., Dai,Y.F. and Fraser,P. (2002) Long-range chromatin regulatory interactions in vivo. *Nat. Genet.*, **32**, 623–626.
16. Marinic,M., Aktas,T., Ruf,S. and Spitz,F. (2013) An integrated holo-enhancer unit defines tissue and gene specificity of the Fgf8 regulatory landscape. *Dev. Cell*, **24**, 530–542.
17. Ruf,S., Symmons,O., Uslu,V.V., Dolle,D., Hot,C., Ettwiller,L. and Spitz,F. (2011) Large-scale analysis of the regulatory architecture of the mouse genome with a transposon-associated sensor. *Nat. Genet.*, **43**, 379–386.
18. Spitz,F., Gonzalez,F. and Duboule,D. (2003) A global control region defines a chromosomal regulatory landscape containing the HoxD cluster. *Cell*, **113**, 405–417.
19. Jaenisch,R. and Bird,A. (2003) Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat. Genet.*, **33**(Suppl), 245–254.
20. Jenuwein,T. and Allis,C.D. (2001) Translating the histone code. *Science*, **293**, 1074–1080.
21. Kouzarides,T. (2007) Chromatin modifications and their function. *Cell*, **128**, 693–705.
22. Drissen,R., Palstra,R.J., Gillemans,N., Splinter,E., Grosveld,F., Philipsen,S. and de Laat,W. (2004) The active spatial organization of the beta-globin locus requires the transcription factor EKLF. *Genes Dev.*, **18**, 2485–2490.
23. Sahlen,P., Abdullayev,I., Ramskold,D., Matskova,L., Rilakovic,N., Lotstedt,B., Albert,T.J., Lundeberg,J. and Sandberg,R. (2015) Genome-wide mapping of promoter-anchored interactions with close to single-enhancer resolution. *Genome Biol.*, **16**, 156.
24. Song,S.H., Hou,C. and Dean,A. (2007) A positive role for NLI/Ldb1 in long-range beta-globin locus control region function. *Mol. Cell*, **28**, 810–822.
25. Vakoc,C.R., Letting,D.L., Gheldof,N., Sawado,T., Bender,M.A., Groudine,M., Weiss,M.J., Dekker,J. and Blobel,G.A. (2005) Proximity among distant regulatory elements at the beta-globin locus requires GATA-1 and FOG-1. *Mol. Cell*, **17**, 453–462.
26. Bain,G., Maandag,E.C., Izon,D.J., Amsen,D., Kruisbeek,A.M., Weintraub,B.C., Krop,I., Schlissel,M.S., Feeney,A.J., van Roon,M. *et al.* (1994) E2A proteins are required for proper B cell development and initiation of immunoglobulin gene rearrangements. *Cell*, **79**, 885–892.
27. Georgopoulos,K., Bigby,M., Wang,J.H., Molnar,A., Wu,P., Winandy,S. and Sharpe,A. (1994) The Ikaros gene is required for the development of all lymphoid lineages. *Cell*, **79**, 143–156.
28. Lin,H. and Grosschedl,R. (1995) Failure of B-cell differentiation in mice lacking the transcription factor EBF. *Nature*, **376**, 263–267.
29. Peschon,J.J., Morrissey,P.J., Grabstein,K.H., Ramsdell,F.J., Maraskovsky,E., Gliniak,B.C., Park,L.S., Ziegler,S.F., Williams,D.E., Ware,C.B. *et al.* (1994) Early lymphocyte expansion is severely impaired in interleukin 7 receptor-deficient mice. *J. Exp. Med.*, **180**, 1955–1960.
30. Scott,E.W., Simon,M.C., Anastasi,J. and Singh,H. (1994) Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages. *Science*, **265**, 1573–1577.
31. Urbanek,P., Wang,Z.Q., Fetka,I., Wagner,E.F. and Busslinger,M. (1994) Complete block of early B cell differentiation and altered patterning of the posterior midbrain in mice lacking Pax5/BSAP. *Cell*, **79**, 901–912.
32. Lin,Y.C., Jhunjhunwala,S., Benner,C., Heinz,S., Welinder,E., Mansson,R., Sigvardsson,M., Hagman,J., Espinoza,C.A., Dutkowski,J. *et al.* (2010) A global network of transcription factors, involving E2A, EBF1 and Foxo1, that orchestrates B cell fate. *Nat. Immunol.*, **11**, 635–643.
33. Reynaud,D., Demarco,I.A., Reddy,K.L., Schjerven,H., Bertolino,E., Chen,Z., Smale,S.T., Winandy,S. and Singh,H. (2008) Regulation of B cell fate commitment and immunoglobulin heavy-chain gene rearrangements by Ikaros. *Nat. Immunol.*, **9**, 927–936.
34. Zhuang,Y., Soriano,P. and Weintraub,H. (1994) The helix-loop-helix gene E2A is required for B cell formation. *Cell*, **79**, 875–884.
35. Medina,K.L., Pongubala,J.M., Reddy,K.L., Lancki,D.W., Dekoter,R., Kieslinger,M., Grosschedl,R. and Singh,H. (2004) Assembling a gene regulatory network for specification of the B cell fate. *Dev. Cell*, **7**, 607–617.
36. Seet,C.S., Brumbaugh,R.L. and Kee,B.L. (2004) Early B cell factor promotes B lymphopoiesis with reduced interleukin 7 responsiveness in the absence of E2A. *J. Exp. Med.*, **199**, 1689–1700.
37. Singh,H., Medina,K.L. and Pongubala,J.M. (2005) Contingent gene regulatory networks and B cell fate specification. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 4949–4953.
38. Boller,S. and Grosschedl,R. (2014) The regulatory network of B-cell differentiation: a focused view of early B-cell factor 1 function. *Immunol. Rev.*, **261**, 102–115.
39. Rothenberg,E.V. (2014) Transcriptional control of early T and B cell developmental choices. *Annu. Rev. Immunol.*, **32**, 283–321.
40. Miyazaki,K., Miyazaki,M. and Murre,C. (2014) The establishment of B versus T cell identity. *Trends Immunol.*, **35**, 205–210.
41. Pongubala,J.M., Northrup,D.L., Lancki,D.W., Medina,K.L., Treiber,T., Bertolino,E., Thomas,M., Grosschedl,R., Allman,D. and Singh,H. (2008) Transcription factor EBF restricts alternative lineage options and promotes B cell fate commitment independently of Pax5. *Nat. Immunol.*, **9**, 203–215.
42. Nagano,T., Lubling,Y., Stevens,T.J., Schoenfelder,S., Yaffe,E., Dean,W., Laue,E.D., Tanay,A. and Fraser,P. (2013) Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature*, **502**, 59–64.
43. Rao,S.S., Huntley,M.H., Durand,N.C., Stamenova,E.K., Bochkov,I.D., Robinson,J.T., Sanborn,A.L., Machol,I., Omer,A.D., Lander,E.S. *et al.* (2014) A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*, **159**, 1665–1680.
44. Peng,C., Fu,L.Y., Dong,P.F., Deng,Z.L., Li,J.X., Wang,X.T. and Zhang,H.Y. (2013) The sequencing bias relaxed characteristics of Hi-C derived data and implications for chromatin 3D modeling. *Nucleic Acids Res.*, **41**, e183.
45. Ay,F., Bailey,T.L. and Noble,W.S. (2014) Statistical confidence estimation for Hi-C data reveals regulatory chromatin contacts. *Genome Res.*, **24**, 999–1011.
46. Imakaev,M., Fudenberg,G., McCord,R.P., Naumova,N., Goloborodko,A., Lajoie,B.R., Dekker,J. and Mirny,L.A. (2012) Iterative correction of Hi-C data reveals hallmarks of chromosome organization. *Nat. Methods*, **9**, 999–1003.
47. Lin,Y.C., Benner,C., Mansson,R., Heinz,S., Miyazaki,K., Miyazaki,M., Chandra,V., Bossen,C., Glass,C.K. and Murre,C. (2012) Global changes in the nuclear positioning of genes and intra- and interdomain genomic interactions that orchestrate B cell fate. *Nat. Immunol.*, **13**, 1196–1204.
48. Kieffer-Kwon,K.R., Tang,Z., Mathe,E., Qian,J., Sung,M.H., Li,G., Resch,W., Baek,S., Pruett,N., Grontved,L. *et al.* (2013) Interactome maps of mouse gene regulatory domains reveal basic principles of transcriptional regulation. *Cell*, **155**, 1507–1520.

49. Palstra,R.J., Tolhuis,B., Splinter,E., Nijmeijer,R., Grosveld,F. and de Laat,W. (2003) The beta-globin nuclear compartment in development and erythroid differentiation. *Nat. Genet.*, **35**, 190–194.
50. Naumova,N., Imakaev,M., Fudenberg,G., Zhan,Y., Lajoie,B.R., Mirny,L.A. and Dekker,J. (2013) Organization of the mitotic chromosome. *Science*, **342**, 948–953.
51. Dobreva,G., Dambacher,J. and Grosschedl,R. (2003) SUMO modification of a novel MAR-binding protein, SATB2, modulates immunoglobulin mu gene expression. *Genes Dev.*, **17**, 3048–3061.
52. Laurenti,E., Doulatov,S., Zandi,S., Plumb,I., Chen,J., April,C., Fan,J.B. and Dick,J.E. (2013) The transcriptional architecture of early human hematopoiesis identifies multilevel control of lymphoid commitment. *Nat. Immunol.*, **14**, 756–763.
53. Wang,X.T., Dong,P.F., Zhang,H.Y. and Peng,C. (2015) Structural heterogeneity and functional diversity of topologically associating domains in mammalian genomes. *Nucleic Acids Res.*, **43**, 7237–7246.
54. Barbieri,M., Chotalia,M., Fraser,J., Lavitas,L.M., Dostie,J., Pombo,A. and Nicodemi,M. (2012) Complexity of chromatin folding is captured by the strings and binders switch model. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, 16173–16178.
55. Kalhor,R., Tjong,H., Jayathilaka,N., Alber,F. and Chen,L. (2012) Genome architectures revealed by tethered chromosome conformation capture and population-based modeling. *Nat. Biotechnol.*, **30**, 90–98.
56. Duan,Z., Andronescu,M., Schutz,K., McIlwain,S., Kim,Y.J., Lee,C., Shendure,J., Fields,S., Blau,C.A. and Noble,W.S. (2010) A three-dimensional model of the yeast genome. *Nature*, **465**, 363–367.
57. Hu,M., Deng,K., Qin,Z., Dixon,J., Selvaraj,S., Fang,J., Ren,B. and Liu,J.S. (2013) Bayesian inference of spatial organizations of chromosomes. *PLoS Comput. Biol*, **9**, e1002893.
58. Gorkin,D.U., Leung,D. and Ren,B. (2014) The 3D genome in transcriptional regulation and pluripotency. *Cell Stem Cell*, **14**, 762–775.
59. Meshorer,E., Yellajoshula,D., George,E., Scambler,P.J., Brown,D.T. and Misteli,T. (2006) Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. *Dev. Cell*, **10**, 105–116.
60. Battulin,N., Fishman,V.S., Mazur,A.M., Pomaznoy,M., Khabarova,A.A., Afonnikov,D.A., Prokhortchouk,E.B. and Serov,O.L. (2015) Comparison of the three-dimensional organization of sperm and fibroblast genomes using the Hi-C approach. *Genome Biol.*, **16**, 77.
61. Heng,T.S. and Painter,M.W. (2008) The Immunological Genome Project: networks of gene expression in immune cells. *Nat. Immunol.*, **9**, 1091–1094.
62. Hatzis,P. and Talianidis,I. (2002) Dynamics of enhancer-promoter communication during differentiation-induced gene activation. *Mol. Cell*, **10**, 1467–1477.
63. Chepelev,I., Wei,G., Wangsa,D., Tang,Q. and Zhao,K. (2012) Characterization of genome-wide enhancer-promoter interactions reveals co-expression of interacting genes and modes of higher order chromatin organization. *Cell Res.*, **22**, 490–503.
64. Bickmore,W.A. and van Steensel,B. (2013) Genome architecture: domain organization of interphase chromosomes. *Cell*, **152**, 1270–1284.
65. Misteli,T. (2007) Beyond the sequence: cellular organization of genome function. *Cell*, **128**, 787–800.
66. Hewitt,S.L., High,F.A., Reiner,S.L., Fisher,A.G. and Merkenschlager,M. (2004) Nuclear repositioning marks the selective exclusion of lineage-inappropriate transcription factor loci during T helper cell differentiation. *Eur. J. Immunol.*, **34**, 3604–3613.
67. Li,L., Lyu,X., Hou,C., Takenaka,N., Nguyen,H.Q., Ong,C.T., Cubenas-Potts,C., Hu,M., Lei,E.P., Bosco,G. *et al.* (2015) Widespread rearrangement of 3D chromatin organization underlies polycomb-mediated stress-induced silencing. *Mol. Cell*, **58**, 216–231.
68. 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.
69. Hagman,J., Ramirez,J. and Lukin,K. (2012) B lymphocyte lineage specification, commitment and epigenetic control of transcription by early B cell factor 1. *Curr. Top. Microbiol. Immunol.*, **356**, 17–38.