Supplementary Figures, Tables, and Methods

Global changes in nuclear positioning of genes and intra- and inter-domain genomic interactions that orchestrate B cell fate

Yin C Lin¹,⁴, Christopher Benner²,³,⁴,⁶, Robert Mansson¹,⁵, Sven Heinz²,⁵, Kazuko Miyazaki¹, Masaki Miyazaki¹, Vivek Chandra¹, Claudia Bossen¹, Christopher K Glass²,⁶ & Cornelis Murre¹,⁶

¹Department of Molecular Biology
²Department of Cellular and Molecular Medicine
University of California, San Diego
La Jolla, CA 92093
³Integrative Genomics Core
Salk Institute for Biological Studies
10010 North Torrey Pines Road
La Jolla, CA92037

⁴These authors contributed equally to this work
⁵These authors contributed equally to this work

⁶Correspondence should be addressed to: cbenner@salk.edu; cglass@ucsd.edu; murre@biomail.ucsd.edu
SUPPLEMENTARY METHODS

**Hi-C.** The Hi-C experiment was performed as previously described with the following exceptions\(^7\). Forty million of pro-B or pre-pro-B cells were cross-linked with 1.5 mM of ethylene glycol bis(succinimidylsuccinate) (EGS) in 20 ml of phosphate buffer saline (PBS) for 45 minutes at 21°C under rotation. The cells were further cross-linked with 1% formaldehyde for an additional 20 minutes while rotating. The reaction was quenched with 0.2 M glycine for 20 minutes while rotating. The cells were washed twice with 20 ml ice-cold PBS. The pro-B cell replicate was cross-linked according to the original Hi-C protocol\(^7\). Ten million nuclei of each cell type were used for chromatin digestion. The nuclei were spun down at 1400 rpm (394 xg) instead of 5000 rpm. For chromatin restriction digestion with HindIII, 400U of fresh HindIII was added to the overnight HindIII digestion for an additional 2-hour incubation in the following morning. Instead of using 50 \(\mu l\) of Invitrogen T4 DNA ligase for the proximity ligation, 10 \(\mu l\) of Fermentas T4 DNA ligase (30 Weiss U/\(\mu l\)) was used per tube.

**ChIP-Seq.** Chromatin was immunoprecipitated as described\(^3^0\). Most of the cells were fixed for 5–10 min at room temperature in the presence of 1% (wt/vol) formaldehyde, followed by lysis and sonication. For the p300 ChIP, the cells were fixed with 1.5 mM of EGS for 15 minutes, followed by an 8-minute fixation in the presence of 1% formaldehyde. Sonicated chromatin was immunoprecipitated with 6 \(\mu l\) of anti-dimethyl-Histone H3 (Lys4) (07-030, Millipore), 6 \(\mu l\) of anti-trimethyl-Histone H3 (K27) (07-449, Millipore), 15 \(\mu l\) of anti-CTCF (07-729, Millipore), 8 \(\mu l\) of anti-Rad21 (ab992, Abcam), 3 \(\mu l\) of anti-H3K36me3 (ab9050, Abcam), 5 \(\mu l\) of anti-c-Myc (sc-764x, Santa Cruz Biotechnology) or anti-p300 (sc-585, Santa Cruz Biotechnology) antibody. Samples were washed, followed by elution and incubated overnight at 65 °C for reversal of cross-linking. Samples were next treated with
RNase A and proteinase K and purified using a PCR Purification kit (Qiagen). Immunoprecipitated chromatin was ligated to adaptors and multiplex amplified by PCR according to the manufacturer's protocol (Illumina). Amplified ChIP DNA was selected by size (150–300 bp) by 8% PAGE and was sequenced with a HiSeq 2000 sequencer (Illumina) except for the c-Myc ChIP-seq, which was sequenced on a GAII sequencer (Illumina). ChIP-Seq and GRO-Seq libraries were sequenced for 50 cycles and was aligned to the mm9 assembly (National Center for Biotechnology Information build 37) with the Bowtie alignment tool. Tags that mapped to unique DNA sequences were analyzed further. Visualization files and peaks were generated using HOMER (http://biowhat.ucsd.edu/homer/) and read pile-ups visualized using the UCSC Genome Browser. Gene expression was calculated by quantifying GRO-Seq reads across gene-bodies using HOMER. Differential expression was determined using EdgeR (http://www.bioconductor.org/packages/release/bioc/html/edgeR.html). Differential genes were defined as genes exhibiting greater than 2-fold change and an FDR of 1e-5.

3D-FISH. Fluorescent in situ hybridization (FISH) and image acquisition were done as described previously. The BAC probes used for this study were RP23-118P17, RP23-286D14, and RP24-74E7 and were obtained from the BACPAC Resource Center (BPRC) at Children’s Hospital Oakland Research Institute. 3D fluorescent images were acquired on a deconvolution microscope (Deltavision) using a 100x objective. Optical sections (z stacks) of 0.2 µm apart were obtained in the DAPI, FITC, Red, and Cy5 channels. The distances separating the probes were measured using the center of mass as focal points. The center of mass was determined by SoftWorx software package. Chromatic aberration was corrected using Tetraspeck measurements as described previously.
**Nuclear lamina and Ebf1 domain association with immuno-FISH.** Fluorescent in situ hybridization (FISH) and image acquisition were done as described previously\textsuperscript{47} with the following changes. The BAC probe used for Ebf1 domain was RP23-118P17 as listed above. It was labeled with Alexa Fluor 488-5-dUTP. The nuclear lamina was stained first with the primary anti-Lamin A and B antibodies (sc-6214 and sc-6217) from Santa Cruz Biotechnology, followed by the secondary staining of donkey anti-goat IgG antibody conjugated to Alexa Flour 594 (A11058) from Invitrogen instead with biotinylated goat-specific donkey antibodies. Fluorescent images were acquired on a deconvolution microscope (Deltavision) using a 100x objective. Optical sections (z stacks) of 0.2 µm apart were obtained in the DAPI, FITC, and Red channels. The distances between the nuclear lamina and the Ebf1 domain were measured in 2D with SoftWorx software package.

**GRO-Seq.** GRO-Seq was performed as previously described\textsuperscript{48} with minor modifications. Briefly, nuclei from 10 million cells isolated by hypotonic lysis, and RNA polymerases were run-on for 5 minutes at 30°C in the presence of sarkosyl, BrUTP, ATP, GTP and limiting concentrations of CTP. Total RNA was purified with Trizol/isopropanol precipitiation, DNase-treated (TURBO-DNase, Ambion), fragmented with fragmentation buffer (Ambion), and re-buffered by size exclusion chromatography. RNA fragments were 3’ dephosphorylated with polynucleotide kinase (Enzymatics), and BrUTP-labeled run-on RNA was immunopurified with anti-BrdUTP-coated agarose beads, washed, eluted and EtOH-precipitated. Run-on RNA was then de-capped with tobacco acid pyrophosphatase (Epicentre), 5’ phosphorylated with polynucleotide kinase (Enzymatics) and purified with Trizol LS/isopropanol precipitiation. Sequencing libraries were prepared by ligating a single-stranded, 3’-blocked, 5’-adenylated 3’ oligonucleotide with mutant (K227Q) truncated RNA ligase 2 (NEB) to the 3’ end of the RNA fragments, followed by annealing a
reverse transcription primer complementary to the 3’ adapter to suppress adapter dimer formation, and ligating a hybrid 5’ DNA-RNA oligonucleotide using RNA ligase I and then reverse-transcribing with SuperScript III reverse transcriptase (Invitrogen). The cDNA was purified with AMPure XL beads and PCR-amplified with primers bearing primer landing sites compatible with Illumina indexed sequencing. The library was size-selected on a PAGE gel to 60-110 bp insert size and sequenced on a HiSeq 2000 (Illumina).

**Hi-C analysis.** Computational data analysis was performed with the software developed in-house. The routines for Hi-C analysis were integrated into HOMER, a general Next-gen sequencing analysis suite, and are freely available ([http://biowhat.ucsd.edu/homer/](http://biowhat.ucsd.edu/homer/)). Specialized visualization of HOMER analysis files was accomplished using Java Tree View, Circos and Cytoscape software.

**Read filtering.** Hi-C libraries were sequenced on an Illumina HiSeq 2000 using 50x2 paired-end sequencing. The ends from each read were mapped separately to the mm9 genome using bowtie, and only paired-end reads with both ends mapping to single, unique genomic locations were retained. Reads were further filtered based on their proximity to HindIII restriction sites, retaining only those containing at least one HindIII site located within 500 bp 3’ from the beginning of the read. 500 bp was the approximate size of fragments used for sequencing. Independent analysis using reads lacking a proximal HindIII site showed considerable more noise. All read-pairs with ends that mapped to locations separated by less than 1 kb were removed from the data. Since Hi-C does not involve any selection and should yield relatively even coverage across the genome, 10 kb regions with high coverage (>5x the average number of reads per region) were also removed to eliminate artifacts or genomic duplications that might confound the analysis.
Sequence reads were also checked for GC and other nucleotide biases. We noted that all experiments were composed of fragments closely resembling the expected genomic frequency (42.8% and 42.6% for pro-B and pre-pro-B cells, respectively; expected frequency was 41.4%, Figure S1B).

**Normalization.** For Hi-C analyses we examined genomic regions at windows of 50 kb and 10 kb. The description of analysis methods below applies broadly to regions of any window size. Since Hi-C analysis is an unbiased assay of nuclear topology, the expectation is that roughly equal numbers of Hi-C reads should originate from each region of equal size in the genome. If different numbers of Hi-C reads are observed, this is likely due to bias in mapping (e.g. repeats or duplicated genomic sequences in the region where no reads can be mapped), a variable number of restriction enzyme recognition sites (e.g. HindIII), or a technical artifact (i.e. inaccessibility of HindIII to DNA). For this reason, Hi-C reads between any two regions of a given size must be normalized for sequencing depth by dividing the total number of interaction reads that each region participates in. To calculate the expected Hi-C reads between two given regions, we use the following equation:

$$ e_{ij} = \frac{n_i n_j}{N} $$

Where N is the total number of reads in the Hi-C experiment and n the total number of reads at each region i and j. This formulation assumes that each region has a uniform probability of interacting with any other region in the genome.

In addition to sequencing depth, it is also useful to normalize data based on the distance between interacting regions. The constrained proximity of regions along linear DNA is by far the strongest signal in Hi-C data, with regions found at close linear distances.
much more likely to generate Hi-C reads than regions located at large linear distances along the chromosome or located on separate chromosomes. By computing the average number of Hi-C reads as a function of distance and sequencing depth, read frequencies between specific regions can be reported relative to the average Hi-C read density for their linear distances to help reliably identify proximity relationships that are not simply a result of the general linear compaction along the DNA.

The general idea was to modify our calculation of the expected number of Hi-C reads between any two loci to account for both their linear distance and sequencing depth. However, this calculation requires that we know the true number of interaction reads originating from a given locus, and not the measured number of interaction reads. For example, if region A (mapped) is adjacent to a region B (repeat region, unmappable), the total number of Hi-C reads mapping to A will likely be much less than a scenario in which B can be mapped as well. This is particularly important of regions adjacent to regions that cannot be mapped where their linear proximity would predict a significant numbers of Hi-C reads to connect the regions. This phenomenon causes a problem when trying to estimate the expected number of reads connecting region A to another region C. Since the total number of reads mapped to A is less because of region B adjacent to A that cannot be mapped, we likely underestimate the number of Hi-C reads mapping A to C.

To account for this, we find the expected number of Hi-C reads can be found with the following equation:

\[ e_{ij} = f(i - j) \frac{n_i^* n_j^*}{N^*} \]

Where \( f \) is the expected frequency of Hi-C reads as a function of distance, \( N^* \) is estimated total number of reads, and \( n^* \) is the estimated total number of interaction reads at each region. The goal was to identify \( n^*_i \) such that the total number of expected reads at each
region i as a function of sequencing depth and distance ($S_i = \text{sum}(e_{ij})$) is equal to the observed number of reads at each region $n_i$. Because the expected number of reads in any given region depends on the number of reads in all other regions, the resulting nonlinear system is difficult to solve directly. Instead, a simple hill climbing optimization was used to estimate inferred total reads. To calculate the inferred number of reads at each region, the model for expected interactions above is used to compute the expected number of read totals for each region, using the actual numbers of interaction reads as the initial values. The difference between the observed number of reads and the expected number of reads is then used to scale the values for the estimated numbers of reads. This is computed for each region iteratively and repeated until the error between expected and observed Hi-C read totals per region is near zero.

**Calculating Hi-C correlation matrices and PCA.** The correlation between the interaction profiles of any two regions can be calculated to assess the similarity between their shared interaction partners, increasing the sensitivity of the analysis. The Pearson’s correlation coefficient is calculated using vectors of the ratio of observed to expected Hi-C reads at every region relative to the two regions of interest. In cases were the expected reads from distal regions become very small, particularly at small window size, such that it is unlikely to observe even a single interaction, adjacent regions are pooled such that the total expected reads for each data point is at least 3. Regions with fewer than 15% of the average number of reads were excluded from the analysis to avoid data points with sparse coverage.

Principal component analysis (PCA) was carried out on each chromosome individually in a similar manner to the method described\textsuperscript{15}. Correlation matrices for each chromosome were analyzed using R (princomp) and the coordinate of each region relative
to the primary principal component was assigned (PC1 values). The values were multiplied by -1 if negative values were significantly associated with gene-rich regions.

To directly compare correlation matrices between Hi-C experiments, correlation matrices were first computed separately for each experiment. Next, the vector of correlation values from the first experiment relative to a given region was compared to the vector of correlation values from the 2nd experiment for the same region. The correlation between these vectors is calculated, revealing the similarity between interaction partners for the region between the two experiments.

**Finding significant interactions.** True proximal positioning between genomic loci should reflect a significant deviation in the number of Hi-C reads from the expected value. To assess the significance of this deviation, we use the binomial distribution to score the probability that x number of Hi-C reads would occur between two given regions by chance, given their expected number of interaction reads between the two regions and the total number of reads for each locus. Since an interaction involves the association between two regions, the region with the fewest total interaction reads is used as the total number of possible interacting Hi-C reads to keep the significance calculation conservative.

Due to the large number of possible interactions in the genome, particularly at smaller window, a false discovery rate (FDR) is calculated to avoid analyzing interactions that may appear simply by chance. To calculate the FDR, we generated random Hi-C experiments with the same sequencing depth per region and linear distance expectation and repeated the interaction calculations to serve as an empirical distribution of expected P-values. Individual interactions were visualized using HOMER to format input files for Circos. Circos was subsequently used to generate the images.
Interaction endpoint feature enrichment. To assess the significance of overlap between significant interaction end-points and other features, end-points were compared to ChIP-Seq peaks and various genome annotations. For each feature, the co-localization of interaction end-points and features were found at 10 kb window. The observed number of overlapping features was compared to the expected number of overlapping features assuming a conservative effective genome size of $2 \times 10^9$ and significance calculated using the Fisher Exact Test.

Interaction end-point co-localization enrichment was calculated by first considering the total number of end-points co-localizing with each genomic feature (10 kb window). The expected number of interactions linking any two features was calculated by assuming each feature was independently distributed among all interaction end-points. The significance of the difference between observed and expected links between any two features was calculated using the binomial distribution based on the total number of interactions. Cytoscape was used to visualize the results for all pair-wise comparisons between the different features.

Structure Interaction Matrix Analysis (SIMA). To test if a given set of features mediates a greater than expected share of interactions between any two large domains, the total number of interactions connecting specific features from one domain to specific features in the other domain was determined at a window of 10 kb ($+/-5$ kb from feature position). This value was normalized to the expected number of interactions at those locations. This normalized interaction value was compared to the values found when randomizing the feature positions 10,000 times within the domain itself to assess the feature connectivity enrichment and the significance of the enrichment. If a domain was compared to itself, only interactions spanning greater than 20 kb were considered to avoid interactions along the diagonal. The results were reported as the ratio of observed normalized interactions
between specific features divided by the average normalized interactions between randomized positions. Only ratios with a corresponding P-value less than 0.05 were reported, otherwise the comparison was assigned a ratio of 1.
Supplementary Figure 1. The folding pattern of the pro-B cell genome. (a) The compartment structures of the murine pro-B and human mature-B cell genomes are highly conserved. Normalized genome-wide contact matrix revealing intrachromosomal interactions involving chromosome 11. Indicated are the ratios of observed versus expected reads. PC1 values corresponding to the coordinates of each genomic region relative to the first principal component from principal component analysis (PCA) are shown along the length of chromosome 11. PC1 values derived from human Hi-C data in lymphoblasts were mapped to orthologous regions and displayed in conjunction with mouse PC1 values. Regions lacking obvious orthologs were removed from the analysis. H3K4me2 ChIP-Seq read density from pro-B cells and RefSeq genes are also shown. (b) Histogram of GC% of pro-B cell Hi-C fragments was shown relative to fragments from random genomic positions. (c) Distribution of Hi-C reads with respect to HindIII sites in the genome. (d) Reproducibility of pro-B Hi-C replicates. PC1 values were calculated for each Hi-C replicate and compared at 50 kb windows across the pro-B cell genome. Comparison is shown for formaldehyde versus EGS/formaldehyde fixed cells. (e) Comparisons between PC1 values of ESC and transformed pro-B Hi-C data. PC1 values were calculated for each Hi-C and compared at 50 kb windows. (f) Comparisons between PC1 values of ESC and pro-B Hi-C data. PC1 values were calculated for each Hi-C and compared at 50 kb windows.
Supplementary Figure 2. The folding pattern of individual chromosomes of the pro-B cell genome. Normalized genome-wide contact matrix revealing intrachromosomal interactions involving all chromosomes are shown. Indicated are the ratios of observed versus expected reads. PC1 values corresponding to the coordinates of each genomic region relative to the first principal component from principal component analysis (PCA) are shown along the length of the chromosomes. Indicated are observed versus expected frequencies. Red pixels indicate higher than expected frequencies of interaction whereas blue reflects lower observed versus expected frequencies.
a  E2A

Structured Interaction Matrix Analysis (SIMA) shows distinct anchors acting at different length scales. Intra- and inter-domain interactions across the transcriptionally permissive compartment were examined for enrichment involving for (a) E2A and (b) CTCF occupancy for each chromosome.

Supplementary Figure 3. Identification of distinct classes of anchors across the pro-B cell genome using Structured Interaction Matrix Analysis (SIMA).
**Supplementary Figure 4.** (a) Normalized genome-wide contact matrix of pre-pro-B cell and pro-B cell for chromosome 11 at 100 kb window. Indicated are the log2 ratios of observed versus expected interaction densities. Red pixels indicate higher than expected interactions. Blue pixels reflect less than expected interactions. (b) Partial list of domains that relocate between the transcriptionally inert and permissive compartments between pre-pro-B and pro-B cells are shown. (c) PC1 values, GRO-Seq, H3K4me2, and E2A ChIP-Seq read densities from pre-pro-B and pro-B cells for the *Igκ* and *Hmga2* loci are visualized.
Supplementary Figure 5. Identification of epigenetic marks and putative anchor-associated chromatin domains in pre-pro-B and pro-B cells. (a) The genome was segregated into four classes of compartments defined by regions inert in both cell types, permissive in only pre-pro-B (pre-pro-B switched), permissive in only pro-B (pro-B switched), or permissive in both. The distinct classes of compartments were examined for association with transcription factor occupancy and deposition of epigenetic marks. Red boxes represent enrichment for epigenetic deposition and factor occupancy higher than expected frequencies whereas blue indicates lower than expected frequencies. (b) Repositioning of chromatin domains during developmental progression from a transcriptionally repressive to a permissive environment is not exclusively associated with activation of gene expression. Logarithm-transformed ratio of GRO-Seq and H3K27me3 ChIP-Seq reads over genomic input reads at the promoters of 703 genes that switch from the inert to permissive compartment. Genes are sorted by their fold induction in pro-B relative to pre-pro-B cells. Yellow indicates higher than expected frequencies whereas blue shows less than expected frequencies. (c) Structured Interaction Matrix Analysis (SIMA) identifies distinct Pax5, PU.1, Foxo1 and p300 acting at different length scales across the pro-B cell interactome. Binding sites within domains were pooled and examined for interactions across the domains to binding sites in other domains. These values were compared to peak positions that were randomized and P-values associated with enrichment for interactions were calculated.
Supplementary Figure 6. Distribution of SINE repeats near genes expressed in embryonic stem cells.

(a) LINE and SINE density, PC1 values, and H3K4me2 read densities from embryonic stem cells (ES) and pro-B cells are visualized. The ES cell-specific gene Zfp42 (Rex-1) is indicated. (b) SINE and LINE repeat elements are localized in regions showing high PC1 versus low PC1 values in pro-B cells. (c) Whisker-box plot of H3K4me2 enrichment across the pro-B, embryonic stem cells and 3T3-L1 genomes in inert/repressive domains that contain more than 30 copies of SINE repeats. (d) Changes in nascent transcription are strongly associated with switched versus constitutive domains.
Supplementary Figure 7. Switching nuclear compartments is closely associated with chromatin remodeling across the *Ebf1* locus. Circos diagrams of the *Ebf1* locus in pre-pro-B and pro-B cells are shown. Top diagram shows the looping (black lines) across the *Ebf1* locus in pre-pro-B cells. Bottom diagram indicates chromatin interactions (black lines) across the *Ebf1* locus in pro-B cells. Purple tiles indicate annotated genes. Orange indicates GRO-Seq signal. Blue reflects E2A occupancy in pro-B cells. Green indicates CTCF binding. Red reflects deposition of H3K4me2.
Hi-C experiments:

| Experiment          | Uniquely mapped reads | Filtered reads (near HindIII) |
|---------------------|-----------------------|------------------------------|
| pro-B (EGS)         | 230,344,220           | 105,936,217                  |
| pro-B (Formaldehyde)| 71,807,415            | 32,569,008                   |
| pre-pro-B (EGS)     | 180,100,803           | 50,746,548                   |

ChIP-Seq experiments:

| Cell type             | Ab         | Uniquely mapped reads | Peaks |
|-----------------------|------------|-----------------------|-------|
| pre-pro-B (E2A-/-)    | H3K4me2    | 16,295,392            | 25,153|
|                       | H3K36me3   | 15,218,329            | 40,989|
|                       | CTCF       | 15,923,805            | 30,311|
|                       | Rad21      | 14,269,546            | 7,404 |
|                       | GRO-Seq    | 26,411,116            |       |
|                       | Input      | 9,104,978             |       |
|                       | PU.1       | 5,058,325             | 39,451|
|                       | Sox2       | 9908245               | 8740  |
|                       | Foxo1      | 39704074              | 11850 |
|                       | Pax5       | 19673836              | 22718 |

1,2,3,4 Previously published as part of GSE21978, GSE21512, GSE11431, and GSE35024, respectively

**Supplementary Table 1.** Summary of reads and peaks obtained from Hi-C, GRO-Seq and ChIP-Seq analysis for the pre-pro-B and pro-B cell genomes are indicated.
| chr | start | end   | Symbol  | F.C.  | chr | start | end   | Symbol  | F.C.  | chr | start | end   | Symbol  | F.C.  |
|-----|-------|-------|---------|-------|-----|-------|-------|---------|-------|-----|-------|-------|---------|-------|
| 1   | 55462789 | 55611549 | 4bla | 84.50 | 5 | 58649865 | 58742445 | Accbg2 | 62.10 | chr16 | 10257906 | 10364760 | Lzic | 5.77 |
| 1   | 10257906 | 10364760 | Lzic | 5.77 | chr9 | 8750419 | 88039686 | Aspd1 | 9.74 | chr10 | 86510999 | 86599918 | Kifap3 | 3.35 |
| 1   | 8750419 | 88039686 | Aspd1 | 9.74 | chr18 | 21706857 | 21790680 | Fdsc | 6.47 | chr17 | 75957091 | 76016870 | Ms4a1 | 101.10 |
| 1   | 21706857 | 21790680 | Fdsc | 6.47 | chr8 | 8593981 | 8595253 | Tusn | 6.37 | chr17 | 75957091 | 76016870 | Ms4a1 | 101.10 |
| 1   | 8593981 | 8595253 | Tusn | 6.37 | chr15 | 10645325 | 10692630 | Timp1 | 15.25 | chr1 | 16193708 | 16207517 | Timp1 | 15.25 |
| 1   | 16193708 | 16207517 | Timp1 | 15.25 | chr10 | 10086658 | 10090632 | Timp1 | 15.25 | chr15 | 10645325 | 10692630 | Timp1 | 15.25 |
| 1   | 10086658 | 10090632 | Timp1 | 15.25 | chr15 | 10645325 | 10692630 | Timp1 | 15.25 | chr1 | 16193708 | 16207517 | Timp1 | 15.25 |
| 1   | 16193708 | 16207517 | Timp1 | 15.25 | chr15 | 10645325 | 10692630 | Timp1 | 15.25 | chr1 | 16193708 | 16207517 | Timp1 | 15.25 |
| 1   | 10086658 | 10090632 | Timp1 | 15.25 | chr15 | 10645325 | 10692630 | Timp1 | 15.25 | chr1 | 16193708 | 16207517 | Timp1 | 15.25 |
| 1   | 16193708 | 16207517 | Timp1 | 15.25 | chr15 | 10645325 | 10692630 | Timp1 | 15.25 | chr1 | 16193708 | 16207517 | Timp1 | 15.25 |
| 1   | 10086658 | 10090632 | Timp1 | 15.25 | chr15 | 10645325 | 10692630 | Timp1 | 15.25 | chr1 | 16193708 | 16207517 | Timp1 | 15.25 |
| 1   | 16193708 | 16207517 | Timp1 | 15.25 | chr15 | 10645325 | 10692630 | Timp1 | 15.25 | chr1 | 16193708 | 16207517 | Timp1 | 15.25 |

Supplementary Table 2. List of genes whose expression is modulated during the developmental progression from the pro-pro-B to the pro-B cell stage and switch nuclear location from the transcriptionally inert to the permissive compartments during the pro-pro-B to the pro-B cell transition is shown. Induction values are included for genes exhibiting significant differences in expression (>2x induction, FDR < 1e-5).
Supplementary Table 3. List of genes whose expression is modulated during the developmental progression from the pre-pro-B to the pro-B cell stage and switch nuclear location from the transcriptionally permissive to inert compartment during the pre-pro-B to pro-B cell transition is shown. Fold changes inductions are included for genes exhibiting significant differences in expression (>2x induction, FDR < 1e-5).