Supplementary Information: Network analysis identifies chromosome intermingling regions as regulatory hotspots for transcription

Anastasiya Belyaeva Saradha Venkatachalapathy Mallika Nagarajan
G.V. Shivashankar Caroline Uhler

November 30, 2017

SI Appendix

Obtaining Hi-C matrices

The Hi-C matrices were obtained from [12] at 250kb resolution. Matrices were corrected for bias using interchromosomal matrix balancing based on the Knight-Ruiz algorithm using software in [3]. Centromeric regions, as well as peri-centromeric regions within 2Mb of the centromere were filtered out. Repeat regions, outliers based on row and column sums (outside of 1.5×interquartile range interval) in the Hi-C contact matrix, and regions already masked in [12], were removed from the analysis. The final Hi-C matrix was $\log(1+x)$ transformed and normalized by mean contact frequency and standard deviation, computed over all interchromosomal contact pairs that were not filtered out.

LAS Algorithm for identifying highly interacting regions

The LAS algorithm [13] takes a real-valued data matrix $X(m \times n)$ as input and outputs contiguous submatrices $U(k \times l)$ that have a high average, τ. This is done via the following iterative algorithm:

Repeat until $\tau\sqrt{kl} < threshold$:

1. Search: greedily, by updating one row and column set at a time, find a submatrix U that maximizes the submatrix score
$$S(U) = -\log\left[(m-k+1)(n-l+1)\Phi(-\tau\sqrt{kl})\right] \quad (1)$$
2. Remove: identify rows and columns corresponding to U in X and subtract the submatrix average τ from this set of rows and columns.

The LAS algorithm search space was limited to contiguous submatrices of at most 10Mb × 10Mb in size, i.e. 40×40 submatrices (at 250kb resolution). For each chromosome pair each iteration of the search procedure was initialized at a random contiguous $k \times l$ submatrix in the interchromosomal Hi-C map. The threshold for the algorithm was chosen based on a Gaussian approximation such that $P(\tau\sqrt{kl} > threshold) = 1E-15$. This stringent cutoff guarantees that highly interacting regions in the whole-genome Hi-C map are identified with FDR controlled at 4.16×10^{-8} (see the following paragraph). Returning submatrices U, as determined by LAS, for each interchromosomal contact matrix results in a list of highly interacting pairs of 250kb regions.

FWER and FDR computation for Large Average Submatrix (LAS) algorithm

The LAS algorithm takes a real-valued matrix $X(m \times n)$ as input and outputs contiguous submatrices $U(k \times l)$ of high average [13]. The null hypothesis is that the interchromosomal Hi-C matrix is a standard Gaussian random matrix, and the alternative hypothesis is that the interchromosomal Hi-C matrix is a sum of K constant (> 0) submatrices plus a standard Gaussian random matrix, i.e., that the Hi-C contact matrix contains substructure [9]. More precisely, each entry in the alternative model can be expressed as
$$x_{i,j} = \sum_{k=1}^{K} \alpha_k I(i \in A_k, j \in B_k) + \epsilon_{ij}, \quad (2)$$
where $A_k \subseteq [m]$ and $B_k \subseteq [n]$ are the row and column sets of the kth submatrix, α_k is the constant corresponding to the kth submatrix, ϵ_{ij} are independent noise variables sampled from $\mathcal{N}(0,1)$, and $I(\cdot)$ is the indicator function. Note that $K = 0$ corresponds to the null model. The individual entries in the $\log(1+x)$ transformed Hi-C matrices have an R^2 of 0.972 with the standard normal distribution, which justifies using a standard Gaussian random matrix as null hypothesis.

Let $\tau := Avg(U)$, i.e., the average of the submatrix U. Under the null hypothesis, $\tau\sqrt{kl} \sim N(0,1)$ and thus the probability of observing a $k \times l$ submatrix V with an average of τ or greater is $P(Avg(V) \geq \tau) = \Phi(-\tau\sqrt{kl})$, where Φ is the standard normal cdf. Let A denote the event that there exists a $k \times l$ submatrix V with average greater than or equal to τ in an $m \times n$ matrix. Note that this event is bounded as follows: $P(A) \leq \sum P(Avg(V) \geq \tau)$, where the sum is over all $k \times l$ submatrices in the $m \times n$ matrix. Hence, under the null hypothesis, $P(A) \leq N\Phi(-\tau\sqrt{kl})$, where $N = (m-k+1) \times (n-l+1)$, the total number of contiguous submatrices of size $k \times l$ in an $m \times n$ matrix.

The search space of the LAS algorithm was limited to contiguous submatrices of at most 10Mb × 10Mb in size, which corresponds to 40×40 submatrices (at 250kb resolution). In order to calculate

the total number of hypotheses for each interchromosomal matrix, we summed the number of possible contiguous submatrices for all combinations of k and l within the [1,40] range. Considering all pairs of interchromosomal matrices, the total number of hypotheses was 9.33×10^{10}. In our procedure, we applied a p-value threshold, namely $P(\tau\sqrt{kl} > threshold) = 1 \times 10^{-15}$, for the discovery of significant submatrices. Using a formulation based on the Bonferroni correction, we can estimate the familywise error rate (FWER), which is the probability of making at least one type I error. Let p be the p-value threshold, b the number of hypotheses and α the FWER level. The Bonferroni correction rejects the null hypothesis when p-value $\leq \frac{\alpha}{b}$, thereby controlling the FWER at $\leq \alpha$. With our p-value threshold of 1×10^{-15}, the FWER is ≤ 0.0000933.

We can also calculate the false discovery rate (FDR), i.e. the fraction of false discoveries among all discoveries, using the Benjamini-Hochberg procedure. Let a be the number of discoveries, b the number of hypotheses, p the p-value threshold, and α the FDR level. The Benjamini-Hochberg procedure rejects the null hypothesis when p-value $\leq \frac{a}{b}\alpha$, thereby controlling the FDR at $\leq \alpha$. With our p-value threshold of 1×10^{-15} and the resulting number of discoveries $a = 2244$, the FDR is $\leq 4.16 \times 10^{-8}$.

Genomic features

Pre-processed data for 48 features including histone modifications, transcription factor ChIP-seq, DNase-seq and RNA-seq, were retrieved from ENCODE (32), Roadmap Epigenomics (33), the GEO database, and previous studies (49) (SI Appendix, Table S2) for the IMR90 cell line. In order to obtain the genomic profile for a 250kb region, matching the resolution of Hi-C data, the number of peaks overlapping the 250kb region was calculated for each feature. For each feature, the feature matrix was $\log(1 + x)$ transformed and z-scored by computing the mean and standard deviation of the feature across all regions in the genome that were not removed by the Hi-C filtering step.

Weighted Correlation Clustering

The weighted network of 250kb regions was partitioned into clusters using weighted correlation clustering on networks [4]. This method determines clusters by drawing cluster boundaries across edges with low weights but not across edges with high weights by solving a non-convex minimization problem. Weighted correlation clustering was run in 25 replicates and the clustering with the lowest value of the objective function was chosen for further analysis. The resulting clusterings were robust across replicate runs, as evidenced by the high adjusted mutual information between cluster labels across runs (Fig. S12).

Classification into Intermingling and Non-intermingling Domains

In order to identify features that may be important for chromosome intermingling, a binary classification task was performed. The training and test data consisted of genomic feature profiles (SI Appendix, Table S2) for intermingling versus non-intermingling regions, weighted by the number of samples in each class. Classification was done using eXtreme gradient boosting trees with $n\_estimators = 1000$, $learning\_rate = 0.1$, $max\_depth = 5$, $min\_child\_weight = 1$ with 10-fold cross-validation. Feature importances were computed by the relative rank of a feature in the decision tree, calculated via $feature\_importances\_$ function in scikit-learn [11] in python. Additionally, features were evaluated using iterative feature elimination by removing one feature at a time and optimizing the AUC.

Fold Enrichment of Genomic Features

Fold enrichment for the intermingling regions as well as for specific clusters was calculated as follows:

$$\frac{\frac{\#\text{ bases in cluster and having feature}}{\#\text{ bases in genome}}}{\left(\frac{\#\text{ bases in cluster}}{\#\text{ bases in genome}}\right)\left(\frac{\#\text{ bases having feature}}{\#\text{ bases in genome}}\right)} \tag{3}$$

A fold enrichment of 1 indicates that the two events - belonging to the intermingling regions or a particular cluster and belonging to a particular feature - are independent events.

Comparison to a random network - Stochastic Block Model

To analyze the importance of the spatial interactions for the function and properties of the determined clusters, we performed a comparison based on a "similar" network in which the spatial interactions have been randomized. To be more precise, we generated a network from a stochastic block model, where each chromosome is a community and the edge probabilities within and between communities are computed from the number of interactions in the Hi-C matrix as determined by the LAS algorithm. In order to obtain similar cluster sizes as in the original network, we sampled the edge weights from the observed distribution of edge weights.

Using a 2-sided χ^2-test we tested whether the proportion of intermingling regions in the observed network was equal (H_A: not equal) to the proportion of intermingling regions in a random network. Generating 50 networks from the stochastic block model and using the average proportion of intermingling regions as test statistic, the null hypothesis was rejected with a p-value $< 2.2 \times 10^{-16}$.

Further, in order to test the functional relevance of the determined clusters, we tested using a 1-sided χ^2-test whether the proportion of clusters enriched for all five active marks (RNAPII, H3K9ac, H3K36me3, H3K4me3, H3K4me1) in the observed network was equal (H_A: larger than) in a random network. As in the previous test, generating 50 networks from the stochastic block model and using the average proportion as test statistic, the null hypothesis was rejected with a p-value of 1.398×10^{-5}.

Finally, we tested the regulatory event that active and inactive clusters are spatially separated. To do this, we tested using a 1-sided χ^2-test whether the proportion of clusters enriched for all five active marks and the inactive mark (H3K9me3) in the observed network was equal (H_A: smaller than) in a random network. Following the same procedure as in the previous test, the null hypothesis was rejected with a p-value of 4.699×10^{-4}.

Venn diagrams

Venn diagrams were constructed using the software provided by [1].

Gene ontology

Expressed genes for IMR90 with reads per kilobase of transcript per million mapped reads (RPKM) > 0 were obtained from ENCODE [5]. For each cluster, we identified the genes that resided in the cluster and were expressed. For each cluster, we then performed gene ontology (GO) term analysis on these genes using DAVID [6, 7].

Cluster ranking

In order to select clusters for experimental validation, we ranked each cluster based on the number of TFBS present in each region in the cluster. Several methods and databases were used for ranking the clusters in order to choose a robust set of clusters for experimental validation. The whole genome was scanned for TFBS using position frequency matrices (PFMs) from the JASPAR2016 database for humans. MOODS software [8] was used to identify motif matches. TFBS were further filtered by ChIP-seq from ENCODE (32), resulting in 52 TFs or robust CAGE peaks, resulting in 386 TFs. The CAGE peaks, which indicate transcription start sites, were obtained from the FANTOM5 project [2], which pooled CAGE analysis over 573 human cell samples. These peaks were flanked by 400bp upstream and 50bp downstream as suggested by [10] and overlayed with the TFBS data to obtain the final set of TFBS.

In the following, we explain how we used the determined TFBS to rank the clusters based on a permutation test. First, we constructed a score function to compare observed and randomized matrices. For each cluster we construct a matrix Z of size $m \times n$ consisting of the TFBS counts for each of the m 250kb domains that are clustered together for each of the n transcription factors that were analyzed. The number m may change from cluster to cluster, while the number of considered transcription factors n is the same for all clusters. Let A be the set of TFs that have TFBS on multiple chromosomes in the considered cluster. Then the score function for each cluster is computed as follows:

$$Score(Z) = \sum_{j \in A} \sum_{i=1}^{m} Z_{ij} \tag{4}$$

For each matrix Z, a set of 1000 random matrices is generated to compute the background score distribution. Assuming that the number of TFBS for a specific transcription factor is independent of the other

transcription factors, a random matrix for a particular cluster with corresponding matrix Z is generated by the following procedure:

1. Let k denote the number of nonzero entries in Z. The probability of having a nonzero entry for each of the n transcription factors is defined by p_j, where $p_j = \frac{\text{\# nonzero entries for TF}_\text{j}}{\text{total \# of nonzero entries}}$. The number of nonzero entries for each transcription factor, x_j, is drawn from a multinomial distribution, $(x_1, \cdots, x_n) \sim Mult(k, p_1, \cdots, p_n)$.

2. After determining the number of nonzero entries for each transcription factor, these nonzero entries must be distributed across the m clustered 250kb regions. Let q_i be the probability of assigning a nonzero entry to that specific region, where $q_i = \frac{\text{\# nonzero entries in region}_\text{i}}{k}$. For each transcription factor j, the number of nonzero entries for each region, $(y_{1j}, \cdots, y_{mj})$ is drawn from a multinomial, $(y_{1j} \cdots y_{mj}) \sim Mult(x_j, q_1, \cdots, q_m)$.

3. By now the positions of nonzero entries within the randomly generated matrix have been chosen, and only the number of TFBS (counts) remain to be assigned to each of the k nonzero entries. For each transcription factor, samples are drawn from the observed count distribution for that transcription factor over active clusters.

Finally, the p-value of the observed score was computed using the background score distribution that we obtained by calculating a score for each of the 1000 randomly generated matrices described above. In order to ensure stability of the ranking procedure, the background distribution was computed in 10 replicates, resulting in 10 different p-values. We observed that the p-values across different runs were consistent. For each replicate, we obtained the cluster rankings based on their p-values. The final rank of each cluster was computed from the median rank across the 10 replicate runs.

Negative controls - chromosomes that do not intermingle

As negative controls, we identified by a whole-genome analysis analysis pairs of chromosomes that do not intermingle (in Hi-C) and are anti-correlated in terms of genomic features (Fig. S8). First, we determined chromosome pairs for which the LAS analysis did not result in any intermingling regions. These chromosomes formed the nodes of a network with edges drawn between pairs of chromosomes with no intermingling regions. The weight of the edges was calculated as $1 - |\rho|$, where ρ is the correlation between the genomic features averaged over the whole chromosome. Chromosomes 3 and 20 were chosen as a negative control pair, since they were representative of the network of non-intermingling chromosomes.

Cell culture and chromosome FISH

BJ fibroblast cells were cultured in Low Glucose DMEM (Life Technologies, USA) supplemented with 10% (vol/vol) FBS (GIBCO, Life Technologies, USA) at 37°C in 5% CO2. BJ fibroblast cells were cultured overnight on fibronectin-coated cleaned glass slides. Cells were then washed with 1× PBS to remove cell culture medium followed by incubation on ice for 5-8 minutes, with 0.25% Triton in CSK buffer (100 mM NaCl, 300 mM Sucrose, 3 mM MgCl2, 10 mM PIPES with pH 6.8). Cultured cells were fixed with 4% PFA (Paraformaldehyde) for 10 minutes, briefly rinsed with 0.1 M Tris-HCl and washed with 1× PBS wash. This was followed by permeabilization with 0.5% Triton for 10-15 minutes. The cells were then incubated overnight in 20% glycerol at 4°C and subjected to 5-6 freeze-thaw cycles in liquid nitrogen. After this, cells were washed with 1× PBS a few times, before and after treatment with 0.01% HCl for 5-10 minutes, followed by protein digestion with 0.002% porcine pepsin (Sigma Aldrich, USA) in 0.01N HCl at 37°C for 4 minutes. Cells were then fixed with 1% PFA for 4 minutes, briefly rinsed in 1× PBS before being treated with RNAse (Promega, USA, 200 microgram/ml made in 2× SSC-0.3M sodium chloride and 30mM trisodium citrate) at 37°C for 15-20 minutes to digest RNA. The cells were then washed with 2× SSC and equilibrated in 50% Formamide / 2× SSC (pH 7.4) overnight at 4°C. Hybridization was set up the following day. Chromosome fish probes (Chrombios, Germany) tagged with different fluorophores were thawed to room temperature and mixed with hybridization buffer provided by the supplier. The DNA was denatured in 50% Formamide / 2× SSC at 85°C for 2-3 minutes and then incubated with the fluorescently labeled human chromosome FISH probe mix. The slides were then sealed with a Sigmacote (Sigma Aldrich, USA) coated hydrophobic coverslip and rubber cement to incubate for 18-48 hours in a moist chamber at 37°C with shaking. At the end of the incubation period, slides were washed three times in 50% Formamide / 2× SSC at 45°C and 0.1× SSC at 60°C. After the last stringent wash with 50% Formamide made in 0.1× SSC at 45°C, the nuclei were blocked in 5% BSA

solution made in 2× SSC and then subjected to primary and secondary antibodies diluted in 5% BSA solution made in 2× SSC. In case indirect labels such as chromosome probes conjugated with biotin and digoxigenin (DIG), were used during hybridization detection step, the procedure also involved the use of fluorophore labeled streptavidin/avidin and anti-DIG. The primary antibodies used here were: RNAPII CTD repeat YSPTSPS (phospho S5) (Abcam - ab5131, 1:500 dilution), mouse monoclonal (21H8) to DIG (Abcam-ab420; 1:500 dilution). Finally, the nuclei were stained with Hoechst 33342 (Sigma Aldrich, USA) for 10 minutes and then mounted with Prolong Gold antifade mounting medium (Life Technologies, USA), sealed with a coverslip, and imaged.

Confocal Imaging and Image Analysis

Slides for chromosome FISH were scanned using a Nikon A1 Confocal microscope (Nikon, USA) with a 100×, 1.4 NA oil objective. Stacks of 12-bit gray scale two-dimensional images were obtained with a pixel size of 130 nm in XY direction and 500 nm in the Z direction and used for the quantitative evaluation. The image analysis was performed using a custom code in ImageJ2. The code first identified the nuclear boundary using Otsu 3D thresholding method. This was followed by the identification of chromosome territories in the nuclear region using ReyniEntropy 3D thresholding. The threshold for identifying signal and background in each image was determined using the intensity histogram from the 3D image stack. The thresholded image was binarized. The overlapping region between two chromosomes, i.e. the intermingling region (IMR), was identified by performing the AND function over the 3D binary stacks of both chromosomes. The chromosome and IMR volumes were computed by summing up the volumes of the non-zero voxels in the respective binary images. The intermingling degree was calculated by dividing the volume of the IMR between two chromosomes by the total volume of the two chromosomes. Similarly, the amount of active RNAPII in the nucleus and the IMR was obtained by passing the RNAPII image and the binary images of the nucleus and IMR through the AND filter, respectively. The enrichment of active RNAPII in the intermingling regions was obtained by dividing the mean intensity of active RNAPII in the IMR by the mean intensity of the active RNAPII in the entire nucleus. R was used for testing statistical significance and for data visualization.

References

[1] P. Bardou, J. Mariette, F. Escudié, C. Djemiel, and C. Klopp. jvenn: an interactive Venn diagram viewer. *BMC Bioinformatics*, 15(1):293, 2014.

[2] T. F. Consortium, the RIKEN PMI, and C. (DGT). A promoter-level mammalian expression atlas. *Nature*, 507(7493):462–470, 2014.

[3] N. C. Durand, J. T. Robinson, M. S. Shamim, I. Machol, J. P. Mesirov, E. S. Lander, and E. L. Aiden. Juicebox Provides a Visualization System for Hi-C Contact Maps with Unlimited Zoom. *Cell Systems*, 3(1):99–101, 2016.

[4] M. Elsner and W. Schudy. Bounding and comparing methods for correlation clustering beyond ILP. *Proceedings of the Workshop on Integer Linear Programming for Natural Language Processing*, (June):19–27, 2009.

[5] ENCODE Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature*, 489(7414):57–74, 2012.

[6] D. W. Huang, R. a. Lempicki, and B. T. Sherman. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, 4(1):44–57, 2009.

[7] D. W. Huang, B. T. Sherman, and R. A. Lempicki. Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Research*, 37(1):1–13, 2009.

[8] J. H. Korhonen, K. Palin, J. Taipale, and E. Ukkonen. Fast motif matching revisited: high-order PWMs, SNPs and indels. *Bioinformatics*, 33(4):514–521, 2016.

[9] E. Lieberman-aiden, N. L. V. Berkum, L. Williams, M. Imakaev, T. Ragoczy, A. Telling, I. Amit, B. R. Lajoie, P. J. Sabo, M. O. Dorschner, R. Sandstrom, B. Bernstein, M. A. Bender, M. Groudine, A. Gnirke, J. Stamatoyannopoulos, and L. A. Mirny. Comprehensive Mapping of Long-Range

Interactions Reveals Folding Principles of the Human Genome. *Science*, 326(October):289–293, 2009.

[10] D. Marbach, D. Lamparter, G. Quon, M. Kellis, Z. Kutalik, and S. Bergmann. Tissue-specific regulatory circuits reveal variable modular perturbations across complex diseases. *Nature Methods*, 13(4):366–370, 2016.

[11] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and É. Duchesnay. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2012.

[12] S. S. P. Rao, M. H. Huntley, N. C. Durand, E. K. Stamenova, I. D. Bochkov, J. T. Robinson, A. L. Sanborn, I. Machol, A. D. Omer, E. S. Lander, and E. L. Aiden. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*, 159(7):1665–1680, 2014.

[13] A. A. Shabalin, V. J. Weigman, C. M. Perou, and A. B. Nobel. Finding large average submatrices in high dimensional data. *Annals of Applied Statistics*, 3(3):985–1012, 2009.

SI Tables

Table S1: Total sizes of intermingling regions

Region	Size
Intermingling domains from LAS	903.25 Mb
Intermingling regions after clustering	459.5 Mb
Intermingling regions in active clusters	179.75 Mb

Table S2: Datasets and accessions used to obtain the genomic features. The symbol * indicates that peaks were retrieved from a previous study (49) that had already pre-processed the data from the GEO database.

Name	Accession	Category
RNA-seq	GSE24565	active
RNAPII	GSE31477	active
H3K4me1	GSE16256	active
H3K4me2	GSE16256	active
H3K4me3	GSE16256	active
H3K36me3	GSE16256	active
H3K9ac	GSE16256	active
H3K27me3	GSE16256	repressive
H3K9me3	GSE16256	repressive
YAP1*	GSE61852	other
RFX5	GSE31477	other
RELA*	GSE43070	other
RCOR1	GSE31477	other
RBL2*	GSE19899	other
RB1*	GSE19899	other
RAD21	GSE31477	other
MXI1	GSE31477	other
MECP2*	GSE47678	other
MAZ	GSE31477	other
MAFK	GSE31477	other
LMNB1*	GSE53332	other
H4K91ac	GSE16256	other
H4K8ac	GSE16256	other
H4K5ac	GSE16256	other
H4K20me1	GSE16256	other
H3K9me1	GSE16256	other
H3K79me2	GSE16256	other
H3K79me1	GSE16256	other
H3K56ac	GSE16256	other
H3K4ac	GSE16256	other
H3K27ac	GSE16256	other
H3K23ac	GSE16256	other
H3K18ac	GSE16256	other
H3K14ac	GSE16256	other
H2BK5ac	GSE16256	other
H2BK20ac	GSE16256	other
H2BK15ac	GSE16256	other
H2BK12ac	GSE16256	other
H2BK120ac	GSE16256	other
H2A.Z	GSE16256	other
MacroH2A1.1	GSE54847	other
H2AK9ac	GSE16256	other
H2AK5ac	GSE16256	other

EP300*	GSE43070	other
DNase-seq	GSE18927	other
CTCF	GSE31477	other
CHD1	GSE31477	other
CEBPB	GSE31477	other

Table S3: Clusters identified via network analysis, enrichment in genomic features, and cluster type. Clusters are given by chromosome number and start position in kb; each region in the cluster is 250kb in length. Only active and inactive clusters with less than seven 250kb regions in the cluster are shown. For a full list refer to Dataset S1.

Clusters (kb)	H3K36me3	RNAPII	H3K9ac	H3K4me3	H3K4me1	H3K9me3	Label
chr2:233250, chr5:170750	1.423	1.348	1.841	1.942	1.399	0.910	active
chr10:30750, chr16:87000, chr21:37250	0.560	0.863	1.371	1.081	1.578	1.369	inactive
chr3:124750, chr3:124500, chr11:73000, chr1:25000	2.870	2.212	3.228	1.998	3.507	0.323	active
chr17:48250, chr13:114500	2.251	9.898	4.081	4.902	3.152	0.489	active
chr11:74750, chr15:69750, chr2:26750, chr1:203000	1.019	0.635	1.719	0.826	1.360	1.386	inactive
chr9:128000, chr1:10500, chr3:197500, chr10:103500, chr2:26250	3.739	1.842	2.066	2.413	2.111	0.297	active
chr15:70000, chr1:25250, chr16:86000, chr10:71250, chr11:69500	0.518	0.936	1.786	0.600	1.470	1.835	inactive
chr19:17500, chr16:29500, chr19:17750, chr17:48500, chr19:18000, chr22:29500	1.945	2.285	2.834	3.075	1.676	0.658	active
chr5:28000, chr21:31250	0.010	0.000	0.048	0.101	0.027	2.068	inactive
chr15:65250, chr2:218750, chr9:132750, chr11:75250, chr15:65000, chr1:201750	3.197	3.558	3.583	2.963	2.724	0.440	active
chr19:32000, chr20:18750	0.032	0.431	0.507	0.159	0.305	2.107	inactive
chr3:97000, chr1:142750	0.026	0.000	0.013	0.024	0.000	1.575	inactive
chr22:42750, chr1:24000, chr10:71750, chr17:40000, chr11:72250, chr15:43500	2.859	2.759	2.354	2.835	1.566	0.749	active
chr15:40750, chr10:75000, chr1:28000, chr6:35750	2.898	2.655	1.621	2.327	1.089	0.521	active
chr22:19250, chr20:3000, chr20:2750	4.106	1.151	2.284	2.950	1.697	0.457	active
chr9:40750, chr15:54500	0.051	0.000	0.055	0.111	0.003	1.182	inactive
chr4:123500, chr6:134250	0.614	6.399	1.826	2.480	2.036	0.331	active
chr19:50250, chr22:20250, chr21:48000	1.874	2.491	1.840	2.709	1.249	0.382	active
chr21:44000, chr11:1000, chr7:500, chr4:500	1.916	0.234	1.659	1.234	1.415	1.587	inactive
chr15:72000, chr6:34750, chr1:202250, chr11:72500	2.680	1.368	1.354	1.090	1.692	0.301	active
chr15:67250, chr1:201500, chr17:75250	4.084	12.517	6.157	3.522	5.362	0.241	active
chr8:141000, chr8:141250, chr7:156500	1.791	1.557	1.354	1.194	1.214	0.549	active
chr2:210500, chr4:179500	0.005	0.000	0.079	0.081	0.037	1.854	inactive
chr12:49250, chr17:38250, chr12:49500	2.858	6.080	4.032	6.611	2.574	0.253	active
chr17:73250, chr22:31500, chr15:76000	2.944	3.592	3.124	3.613	2.582	0.210	active
chr15:42750, chr11:65750, chr10:70250, chr17:37500, chr15:43000	3.788	1.665	2.142	2.485	1.680	0.234	active
chr18:78000, chr19:59000	0.592	3.355	1.156	2.669	0.567	0.136	active
chr1:34750, chr1:34000, chr1:34250, chr8:142500, chr1:35000, chr1:34500	0.060	0.000	0.371	0.123	0.130	3.054	inactive
chr21:34000, chr20:2500	1.858	2.082	1.627	2.315	0.942	0.770	active
chr14:29000, chr21:22250, chr14:29250	0.042	0.000	0.201	0.639	0.221	1.273	inactive
chr5:179500, chr1:6000, chr5:180000, chr1:6250, chr6:170750	2.057	1.357	1.962	2.489	1.267	0.990	active
chr4:187000, chr2:237250, chr2:237750, chr5:174250, chr2:237000	0.428	0.339	1.007	0.990	0.649	1.696	inactive
chr20:57750, chr21:40750, chr22:33250, chr22:33500	0.539	0.823	0.729	0.553	0.319	1.421	inactive
chr6:170250, chr1:13750	0.265	0.000	0.686	0.218	0.298	3.152	inactive
chr1:232750, chr5:173750, chr5:174000, chr4:185000	0.235	0.285	1.547	1.026	1.480	1.193	inactive
chr15:30500, chr14:47000	0.007	0.000	0.013	0.022	0.002	1.210	inactive
chr20:48750, chr20:48500, chr21:47250, chr22:30500	2.066	13.599	4.770	3.083	3.801	0.419	active
chr12:53500, chr4:1000, chr12:53250	2.535	5.004	4.081	5.002	2.787	0.544	active
chr3:47000, chr12:123500, chr1:21000	4.802	1.555	2.516	2.085	2.701	0.119	active
chr5:137500, chr1:6500, chr5:137750	3.557	3.871	3.037	4.253	1.852	0.341	active
chr19:48250, chr17:43750, chr19:48500	1.459	0.202	1.187	0.987	0.431	1.104	inactive
chr3:5250, chr12:128000, chr12:127750	0.179	0.000	0.443	0.211	0.825	2.736	inactive
chr1:23500, chr12:122750, chr3:50000	4.037	3.956	2.656	2.446	2.539	0.250	active
chr9:137500, chr17:75000	3.986	1.556	3.424	1.442	3.416	0.679	active
chr15:32500, chr2:5000	0.000	0.000	0.005	0.000	0.000	1.068	inactive
chr17:81000, chr7:0	0.996	2.609	2.496	2.603	2.174	0.691	active
chr11:73500, chr2:26000, chr15:65750, chr10:104500, chr17:41000	3.520	2.012	2.080	2.076	1.940	0.272	active
chr5:164000, chr5:164750, chr5:165000, chr5:164500, chr2:229500, chr5:164250	0.034	0.000	0.317	0.085	0.288	1.948	inactive
chr7:158500, chr2:3250, chr5:179250, chr6:167250, chr5:179000	4.010	2.967	2.302	2.818	1.571	0.574	active
chr15:69500, chr22:41000	2.384	3.361	2.303	1.928	2.475	0.305	active
chr5:167500, chr1:228000	0.286	0.431	1.353	1.198	0.719	1.367	inactive
chr3:49750, chr3:49500, chr1:24250, chr17:46750	1.997	3.171	2.259	3.072	1.747	0.871	active
chr22:43250, chr20:32500	3.093	2.639	2.669	1.674	3.034	0.409	active
chr19:34250, chr20:56250, chr19:33750, chr19:34000, chr20:56500, chr20:56000	0.849	0.879	1.382	0.913	0.711	1.104	inactive
chr17:44000, chr7:101500, chr22:40750, chr1:23250	3.297	1.651	2.278	1.107	3.477	0.480	active
chr2:750, chr2:1250, chr2:500, chr2:1000, chr10:135250	0.197	0.095	0.474	0.513	0.266	2.477	inactive
chr21:37500, chr18:9500, chr18:9750	3.222	2.515	2.108	2.129	2.767	0.425	active
chr4:12250, chr15:28750, chr4:12500	0.091	0.000	0.099	0.250	0.012	1.874	inactive
chr15:66750, chr6:36750, chr1:202000, chr11:66750	2.146	2.426	2.612	2.046	2.223	0.455	active
chr4:1750, chr16:1500, chr13:114000, chr13:113750, chr7:2250	4.343	2.571	3.261	3.426	2.147	0.356	active
chr5:175000, chr1:233500, chr1:234000, chr5:175250, chr4:187250, chr1:233750	0.086	0.144	0.337	0.384	0.231	2.217	inactive
chr10:125000, chr1:5750, chr10:125250	0.584	0.000	0.442	0.044	0.137	3.377	inactive
chr22:26500, chr20:55750, chr20:55250, chr21:42750, chr20:55500, chr20:55000	0.746	0.461	1.064	0.663	0.573	1.890	inactive
chr2:237500, chr4:188750	0.206	0.000	0.963	0.192	1.112	1.906	inactive
chr9:137250, chr17:78750, chr1:9250, chr16:85000	3.272	6.178	4.438	2.449	3.959	0.446	active
chr14:74000, chr17:40500	3.647	4.099	3.848	5.281	3.036	0.204	active
chr12:52250, chr15:73750	2.234	14.967	3.360	3.774	2.636	0.373	active
chr19:49250, chr17:46500, chr22:36500	2.830	8.133	3.922	7.375	2.628	0.647	active
chr12:2500, chr9:130000, chr15:42250, chr12:2750	3.392	2.228	1.599	1.484	1.169	0.698	active
chr15:42000, chr11:66500, chr1:22000	4.266	2.732	2.940	1.626	2.606	0.230	active
chr6:162500, chr1:237000, chr1:237250, chr1:237500	0.415	0.000	0.160	0.123	0.122	2.830	inactive
chr18:12250, chr21:33000	2.656	5.633	1.964	3.184	1.808	0.516	active
chr20:23250, chr19:32750	2.224	3.350	1.316	1.882	1.051	0.454	active
chr16:68000, chr17:45000, chr9:132500, chr16:68250	3.519	2.843	2.394	2.560	1.636	0.374	active
chr17:17750, chr22:21000	3.784	1.273	2.893	2.013	2.485	0.205	active
chr19:56000, chr19:55500, chr22:50250, chr19:55750	2.654	8.154	3.389	5.962	2.105	0.921	active
chr21:25500, chr20:12000	0.033	0.000	0.179	0.040	0.336	1.065	inactive
chr15:44000, chr22:41250	3.009	2.925	2.309	3.444	1.705	0.171	active
chr20:33250, chr16:3750, chr12:1750, chr10:74750, chr4:2500, chr20:33000	4.149	2.123	2.010	1.689	2.323	0.362	active
chr16:30000, chr19:10750, chr16:29750, chr19:10500	3.381	4.911	3.608	5.989	2.148	0.357	active
chr3:127750, chr3:128000, chr9:129500	2.495	1.965	2.381	2.126	2.127	0.666	active
chr17:45250, chr16:29250	2.322	3.078	2.735	1.269	2.271	0.449	active
chr5:172250, chr2:238000	2.995	13.301	6.091	2.195	4.571	0.356	active
chr14:44500, chr10:47250	0.006	0.000	0.066	0.116	0.039	1.057	inactive
chr1:15500, chr6:169750	1.486	0.681	2.208	1.049	2.357	1.029	inactive
chr4:186500, chr6:169250, chr5:167250, chr4:186750	0.087	0.431	0.792	0.506	0.792	1.211	inactive
chr16:84250, chr4:2000, chr4:2250, chr13:114250	1.938	0.540	2.065	1.811	1.603	1.275	inactive
chr7:129500, chr3:195000	3.296	1.596	1.623	1.711	2.162	0.196	active
chr22:30250, chr12:1500, chr15:40000, chr17:78500, chr16:4000	2.484	1.680	2.494	2.361	3.330	0.401	active
chr16:67000, chr17:49000, chr16:67500	3.557	4.573	3.451	4.481	2.734	0.152	active
chr2:220250, chr1:16750, chr2:220000	3.158	5.466	3.607	4.629	2.036	0.410	active
chr17:38500, chr22:36750, chr16:86500, chr1:27000, chr10:75500, chr11:65000	3.009	18.696	4.696	6.521	3.370	0.373	active
chr13:82750, chr7:143250	0.056	0.000	0.061	0.029	0.008	1.114	inactive

Table S3 : Clusters identified via network analysis, enrichment in genomic features, and cluster type.

Clusters (kb)	H3K36me3	RNAPII	H3K9ac	H3K4me3	H3K4me1	H3K9me3	Label
chr11:72000, chr6:37250	3.090	1.311	1.605	1.052	1.545	0.532	active
chr19:31000, chr20:19000	0.027	0.000	0.220	0.313	0.067	2.469	inactive
chr2:229750, chr5:166250, chr5:166750	0.027	0.000	0.075	0.117	0.013	2.378	inactive
chr19:18250, chr17:1750, chr22:19750	3.458	12.328	4.399	4.620	2.914	0.500	active
chr21:44250, chr17:17000	4.258	2.441	3.226	2.723	2.309	0.302	active
chr16:69000, chr22:36250, chr22:36000	2.334	1.400	2.394	2.296	3.074	0.349	active
chr14:52000, chr15:63250, chr14:51750, chr14:52250	1.487	4.368	2.244	2.423	3.152	0.314	active
chr4:123750, chr8:121500, chr6:136250, chr6:136000	2.371	1.273	1.429	1.249	2.284	0.408	active
chr6:103750, chr9:68000	0.000	0.000	0.003	0.000	0.000	1.153	inactive
chr2:90000, chr16:32250	0.008	0.000	0.017	0.105	0.000	1.277	inactive
chr17:51000, chr18:70500	0.026	0.000	0.085	0.320	0.064	2.408	inactive
chr22:46250, chr9:130750, chr11:2750	2.343	3.706	3.933	4.750	2.968	0.736	active
chr16:28500, chr19:10250	2.722	2.962	2.572	3.376	1.503	0.266	active
chr1:203250, chr17:74500	1.615	7.969	4.034	3.857	3.157	0.574	active
chr22:50500, chr17:7750, chr17:8000	3.033	9.854	4.030	7.875	2.236	0.558	active
chr20:32250, chr22:46500, chr20:32000	3.607	1.580	2.798	2.834	2.204	0.411	active
chr17:44250, chr16:4250, chr7:100000, chr7:99500	2.402	5.604	3.392	5.203	2.140	0.359	active
chr4:32500, chr10:46250	0.232	0.000	0.018	0.000	0.021	1.095	inactive
chr5:172750, chr2:235750, chr2:235250, chr5:169000, chr2:235500, chr4:189750	0.832	2.242	2.200	1.104	2.355	0.870	active
chr8:7000, chr5:26000	0.006	0.000	0.028	0.094	0.015	1.354	inactive
chr15:41500, chr21:47750, chr17:500, chr22:29750	4.841	1.438	2.278	2.285	1.448	0.328	active
chr17:73000, chr8:144250	3.095	3.572	3.374	4.836	1.779	0.566	active
chr6:165500, chr8:136750	0.100	0.000	0.115	0.107	0.024	2.881	inactive
chr7:88750, chr9:43250	0.002	0.000	0.077	0.000	0.054	1.361	inactive
chr5:25750, chr5:24250, chr21:31750, chr5:28750, chr21:32000, chr21:31500	0.012	0.000	0.036	0.004	0.001	2.701	inactive
chr19:19250, chr16:30250, chr19:19500, chr19:19000, chr16:30500	3.413	4.223	3.442	4.317	1.867	0.352	active
chr6:160500, chr1:236500	1.486	0.326	0.907	0.461	0.524	1.501	inactive
chr22:33000, chr20:19750	3.212	3.549	3.452	2.324	4.253	0.642	active
chr9:68500, chr8:114750	0.032	0.000	0.025	0.123	0.000	1.268	inactive
chr14:26750, chr14:26500, chr14:27000, chr14:27250, chr21:26000	0.012	0.000	0.170	0.200	0.144	1.619	inactive
chr8:137000, chr6:164500	0.022	0.000	0.069	0.000	0.011	3.353	inactive
chr22:44750, chr16:84500, chr17:750	3.241	1.953	3.647	2.710	3.580	0.723	active
chr4:140250, chr1:10000	4.023	1.444	1.476	1.444	1.663	0.310	active
chr4:181000, chr2:212250	0.014	0.000	0.119	0.111	0.059	1.988	inactive
chr12:122000, chr5:139000	2.344	3.149	3.928	5.351	2.539	0.536	active
chr22:21250, chr19:48750, chr19:49000, chr17:37750	1.827	3.699	2.985	4.730	2.123	0.591	active
chr2:155250, chr17:44500	0.035	0.000	0.013	0.065	0.002	1.409	inactive
chr17:11500, chr17:10750, chr5:6250, chr17:11000, chr17:11250	0.113	0.079	0.172	0.131	0.035	3.264	inactive
chr1:1000, chr12:123250, chr1:750, chr9:133500, chr8:145000, chr13:114750	3.042	9.057	4.072	6.169	2.975	0.512	active
chr15:75750, chr17:46000, chr1:23750, chr19:46000	3.265	7.361	3.753	4.392	2.487	0.307	active
chr12:57000, chr12:56750, chr12:56500, chr17:47750	3.511	4.664	3.052	4.413	1.854	0.436	active
chr2:219000, chr1:1500	3.456	4.024	3.255	4.225	2.055	0.528	active
chr11:73250, chr10:32500, chr15:72750	2.392	2.497	1.302	1.975	1.861	0.258	active
chr9:138500, chr11:70750	2.725	0.000	1.296	1.055	1.212	2.219	inactive
chr17:4750, chr17:4500, chr22:19000	3.594	3.942	3.115	4.055	1.736	0.609	active
chr2:240750, chr5:173500, chr7:154750, chr2:240500	0.893	0.157	0.684	0.528	0.269	2.573	inactive
chr12:125500, chr8:140750	1.633	0.863	1.199	0.684	0.765	1.627	inactive
chr6:36500, chr10:74000, chr17:42500	1.894	5.294	3.049	3.359	2.395	0.316	active
chr21:38500, chr20:47500	4.092	3.421	1.757	1.943	1.683	0.489	active
chr14:104500, chr14:104750, chr13:112000	0.096	0.000	0.358	0.195	0.078	2.994	inactive
chr7:85500, chr2:88000	0.012	0.000	0.013	0.086	0.002	1.000	inactive
chr21:32250, chr20:15250	0.036	0.431	0.184	0.050	0.126	3.001	inactive
chr4:180000, chr2:212000	0.024	0.000	0.044	0.000	0.033	2.117	inactive
chr5:53250, chr7:81250, chr5:52500	0.923	1.151	1.618	0.780	2.005	0.546	active
chr10:51250, chr1:73750	0.057	0.000	0.142	0.261	0.053	1.199	inactive
chr19:20500, chr9:68750	0.409	0.000	0.108	0.386	0.029	1.230	inactive
chr1:239750, chr6:163250	0.033	0.000	0.107	0.104	0.107	2.523	inactive
chr17:19250, chr22:46750	1.003	1.133	2.438	2.517	1.840	0.972	active
chr16:84750, chr6:43750	2.250	3.458	4.078	0.984	3.821	0.581	active
chr2:89750, chr16:46250	0.005	0.000	0.003	0.000	0.000	1.181	inactive
chr5:139250, chr1:7000	0.256	0.662	1.214	1.338	0.578	1.521	inactive
chr6:165250, chr8:138500	0.025	0.000	0.012	0.000	0.000	3.019	inactive
chr2:230250, chr5:166500	0.014	0.000	0.520	0.177	0.394	2.267	inactive
chr6:164000, chr6:164250, chr5:174500	0.102	0.288	0.699	0.253	1.071	1.782	inactive
chr21:35250, chr12:25000, chr12:25250	1.470	1.391	1.234	1.645	2.054	0.566	active
chr15:62750, chr13:45250, chr13:45500, chr15:63000	2.138	1.255	1.881	0.717	2.577	0.331	active
chr21:26500, chr20:15500	0.018	0.000	0.398	0.053	0.584	1.051	inactive
chr22:45250, chr20:31750	0.871	0.431	1.223	0.790	0.543	1.268	inactive
chr4:181750, chr3:192250	0.002	0.431	0.189	0.084	0.083	1.799	inactive
chr2:0, chr5:178250	1.406	0.431	0.514	1.117	0.161	3.225	inactive
chr22:47000, chr20:59750, chr20:60000, chr22:47500	1.967	0.677	1.636	1.144	1.605	1.828	inactive
chr7:128250, chr1:202500	3.514	2.811	2.314	2.513	2.007	0.522	active
chr19:30750, chr20:19250	0.031	0.000	0.205	0.000	0.032	3.048	inactive
chr14:105000, chr12:132500, chr14:105250, chr12:132250, chr14:105500	2.922	1.970	3.126	3.626	1.957	0.672	active
chr1:26000, chr3:48500	4.793	3.654	4.920	4.844	3.350	0.191	active
chr22:27750, chr21:39000, chr22:27500	0.184	0.288	1.045	0.348	1.517	1.456	inactive
chr14:74250, chr17:41250, chr15:72500	3.161	6.297	2.513	3.337	1.498	0.382	active
chr2:230000, chr3:193000, chr4:181250	0.091	0.000	0.230	0.232	0.112	2.292	inactive
chr21:35750, chr20:4000	1.365	4.249	2.536	1.895	2.464	0.845	active
chr17:250, chr22:41500, chr13:115000	3.364	2.010	2.073	1.969	1.656	0.261	active
chr1:22250, chr19:45250	3.044	4.325	3.363	4.009	2.605	0.282	active
chr2:242500, chr9:140500	5.277	2.551	2.854	2.910	2.378	0.738	active
chr1:27250, chr11:76250	1.625	3.649	3.604	2.211	2.840	0.601	active
chr1:203500, chr2:28750	1.761	2.422	3.377	1.720	3.440	0.298	active
chr20:58250, chr20:58000, chr22:32500	0.028	0.000	0.337	0.173	0.063	1.841	inactive
chr15:72250, chr12:51500	2.259	2.756	1.405	1.906	1.654	0.224	active
chr12:124500, chr2:242750	0.480	0.000	0.975	0.589	0.891	1.905	inactive
chr22:40250, chr21:35500, chr17:44750, chr15:69250	0.703	0.658	0.978	0.998	1.252	1.310	inactive
chr4:172000, chr1:143750	0.023	0.000	0.164	0.430	0.221	1.208	inactive
chr6:42750, chr16:84000, chr6:43000	3.616	4.106	3.462	4.169	1.970	0.422	active
chr17:17500, chr21:33750	2.590	4.865	4.207	3.336	3.252	0.517	active
chr4:189500, chr1:234750	0.112	2.033	1.362	0.571	1.740	0.684	active
chr11:76000, chr8:124250, chr11:75750, chr15:66500, chr3:196000, chr1:24750	2.473	1.571	1.521	2.227	1.447	0.588	active
chr8:125250, chr6:129750	0.918	2.158	1.365	1.715	1.959	0.292	active
chr17:62000, chr1:24500	1.627	3.359	2.185	1.579	2.042	0.777	active
chr1:8750, chr2:202500	2.404	4.061	2.698	3.316	2.729	0.358	active
chr1:11000, chr3:195500, chr2:219750, chr1:10750, chr3:195250	1.777	0.739	1.608	2.123	1.121	1.040	inactive
chr11:72750, chr10:104750, chr1:21250	2.931	1.596	2.194	1.318	3.258	0.380	active
chr19:3250, chr9:137000, chr17:43000, chr16:67250	2.709	4.559	4.219	4.096	2.997	0.532	active
chr1:2250, chr5:175750, chr1:2000	2.764	3.389	3.953	4.985	2.833	0.542	active
chr6:123750, chr4:132000	0.012	0.000	0.033	0.080	0.000	2.334	inactive
chr18:10000, chr21:40250, chr18:10250	0.373	0.000	1.907	0.479	2.498	1.041	inactive
chr14:74750, chr17:47500	1.938	1.478	1.868	1.836	1.581	0.951	active
chr12:133250, chr2:242250, chr2:242000	4.862	3.385	2.725	3.590	1.787	0.514	active
chr3:46500, chr9:129750	2.089	1.608	3.240	1.433	3.290	0.705	active
chr1:7250, chr3:170250, chr2:217250	0.907	1.640	2.063	0.882	2.443	0.626	active
chr2:1500, chr6:169500	2.103	3.478	3.079	1.628	2.936	0.759	active
chr17:40250, chr10:76750	3.593	3.626	3.204	2.686	3.092	0.202	active
chr1:242000, chr6:160750	0.409	0.616	0.706	0.751	0.432	1.909	inactive
chr4:183750, chr2:236000, chr5:174750	1.070	0.687	1.096	0.640	1.385	1.163	inactive
chr2:78000, chr16:32000	0.006	0.000	0.028	0.177	0.000	1.297	inactive
chr1:23000, chr9:132000, chr10:71500, chr10:71000, chr6:35000, chr15:68750	2.260	1.235	3.158	1.360	3.054	0.568	active
chr20:20750, chr19:32250	0.009	0.000	0.158	0.000	0.011	2.669	inactive
chr1:238250, chr6:162250	0.022	0.000	0.097	0.000	0.079	3.127	inactive

Table S3 : Clusters identified via network analysis, enrichment in genomic features, and cluster type.

Clusters (kb)	H3K36me3	RNAPII	H3K9ac	H3K4me3	H3K4me1	H3K9me3	Label
chr15:64000, chr10:32250	1.699	1.183	1.174	0.914	1.443	0.499	active
chr2:211750, chr4:179750	0.014	0.000	0.056	0.000	0.004	2.129	inactive
chr21:43250, chr5:250	4.284	7.366	3.647	3.850	3.270	0.512	active
chr1:26750, chr7:99750, chr22:39250, chr17:76000, chr1:26500, chr11:66250	2.601	2.151	2.782	3.094	1.654	0.497	active
chr7:100250, chr7:100750, chr11:65250, chr17:42750	2.942	10.116	4.162	6.824	2.254	0.542	active
chr12:125000, chr2:217500, chr3:197000	1.204	1.938	3.385	2.155	2.996	0.851	active
chr2:224250, chr1:241750, chr2:224500, chr2:223750	0.875	1.432	1.254	0.830	1.844	0.410	active
chr21:34750, chr20:3750	4.801	4.828	3.022	4.542	2.644	0.161	active
chr16:4750, chr19:11000	4.777	3.499	3.601	3.363	1.879	0.266	active
chr14:102250, chr12:123750	4.778	1.552	2.580	2.851	1.959	0.127	active
chr15:47000, chr9:69250	0.002	0.000	0.005	0.103	0.011	1.181	inactive
chr21:26250, chr20:7000, chr20:7250, chr20:6750	0.034	0.000	0.129	0.081	0.048	1.327	inactive
chr6:163500, chr1:241000, chr1:241250	0.020	0.000	0.180	0.078	0.071	2.818	inactive
chr22:19500, chr17:4250, chr22:18750	0.716	1.405	1.946	1.725	1.378	0.748	active
chr12:50000, chr6:35250	3.305	5.211	4.071	3.695	2.959	0.311	active
chr22:40000, chr10:72250, chr15:73500	0.890	0.575	1.361	0.327	1.460	1.203	inactive
chr20:33500, chr14:75250, chr15:68500	3.449	1.319	2.876	2.120	2.818	0.319	active
chr4:96750, chr9:44250	0.000	0.000	0.009	0.051	0.007	1.133	inactive
chr5:172000, chr2:239750	1.693	4.227	2.929	2.526	2.330	0.867	active
chr20:12500, chr21:27750, chr20:12250	0.008	0.000	0.083	0.098	0.029	2.230	inactive
chr7:101250, chr7:101750, chr17:37250, chr9:134500	2.647	1.083	2.300	1.976	2.017	0.560	active
chr9:24750, chr15:23250	0.155	0.000	0.037	0.142	0.000	1.112	inactive
chr14:74500, chr15:64250, chr17:43500	2.590	1.404	1.752	1.954	1.235	0.435	active
chr2:226500, chr3:190000	0.076	1.441	0.690	0.922	0.923	0.321	active
chr17:80750, chr14:103250, chr14:103000	4.074	1.744	1.953	1.522	2.308	0.583	active
chr6:159500, chr1:236750, chr6:159750	0.512	0.466	0.649	0.562	0.344	1.650	inactive
chr4:167250, chr11:89500	0.004	0.000	0.013	0.000	0.009	1.149	inactive
chr21:43500, chr22:28000	2.108	0.431	2.572	1.525	2.187	1.250	inactive
chr11:2250, chr4:3500, chr11:2000	0.902	0.648	1.810	1.867	1.079	1.698	inactive
chr20:59500, chr20:58750, chr20:59000, chr20:59250, chr19:32500	0.048	0.173	0.346	0.117	0.240	2.953	inactive
chr6:167000, chr1:15250	2.255	3.144	2.708	1.945	2.622	0.928	active
chr14:75000, chr20:33750	3.305	3.009	3.358	2.495	3.306	0.230	active
chr3:126500, chr1:26250, chr3:126250	2.018	1.263	2.060	1.802	1.826	0.652	active
chr17:18000, chr22:43500, chr21:34500	2.844	2.199	3.090	2.922	2.016	0.643	active
chr5:171000, chr2:239500	0.442	0.000	0.953	0.167	0.972	1.879	inactive
chr22:26250, chr20:22000, chr20:21750	0.011	0.000	0.101	0.000	0.000	3.180	inactive
chr22:31000, chr17:41750	1.296	3.118	2.121	2.411	1.755	0.510	active
chr16:18500, chr7:85750	0.153	0.000	0.089	0.255	0.033	1.084	inactive
chr7:77250, chr4:129000	3.458	1.922	1.859	1.990	2.510	0.510	active
chr9:67250, chr3:96000	0.038	0.000	0.147	0.226	0.123	1.328	inactive
chr22:50750, chr17:7250	4.009	8.212	4.114	7.026	2.330	0.379	active
chr1:8000, chr1:8250, chr2:201500	1.912	8.022	3.473	2.939	3.169	0.337	active
chr21:25750, chr18:2000, chr21:25250, chr14:26000	0.004	0.000	0.063	0.101	0.083	2.095	inactive
chr1:13000, chr2:156500, chr2:156750	0.001	0.000	0.028	0.016	0.001	1.749	inactive
chr1:148750, chr7:122500	0.258	0.431	0.275	0.816	0.223	1.363	inactive
chr11:74500, chr14:50500	3.703	2.628	1.569	1.047	1.424	0.485	active
chr17:48750, chr12:56250	2.341	3.686	3.359	4.556	2.131	0.355	active
chr9:43500, chr4:35000	0.006	0.000	0.005	0.050	0.000	1.248	inactive
chr17:74250, chr22:45500	3.395	9.272	3.875	3.819	3.048	0.237	active
chr18:50250, chr6:94750	0.008	0.000	0.013	0.000	0.000	2.185	inactive
chr19:46500, chr15:74000	1.264	0.885	1.669	1.725	1.270	1.580	inactive
chr5:138750, chr1:9750	3.071	2.643	2.611	1.631	2.488	0.278	active
chr4:132250, chr6:123500	0.004	0.000	0.071	0.000	0.003	2.125	inactive
chr13:42000, chr15:67000	2.075	1.369	3.628	1.317	3.524	0.611	active
chr12:750, chr15:40250, chr11:73750	3.137	1.712	1.928	1.926	2.219	0.364	active
chr2:236500, chr4:186000	2.826	1.503	1.840	1.479	2.496	0.233	active
chr3:192500, chr2:227500	1.484	1.191	2.374	3.144	2.752	0.285	active
chr7:49000, chr16:32750	0.000	0.000	0.045	0.111	0.002	1.675	inactive
chr12:27000, chr21:38250	2.675	2.585	1.401	2.292	1.615	0.464	active
chr22:24500, chr21:40500	3.886	2.481	2.730	2.319	2.185	0.331	active
chr18:51250, chr9:44000	0.011	0.000	0.036	0.102	0.007	1.661	inactive
chr6:36000, chr12:49750	3.422	1.485	1.614	1.645	2.239	0.529	active
chr1:1250, chr12:124750	4.612	5.699	4.692	5.677	3.030	0.619	active
chr16:33000, chr14:48500	0.008	0.000	0.037	0.013	0.000	1.104	inactive
chr17:62250, chr15:63750	2.392	3.266	1.613	1.707	2.339	0.282	active
chr10:2750, chr8:4750	0.029	0.000	0.179	0.034	0.130	2.943	inactive
chr21:43750, chr4:6000, chr4:6250	0.451	0.288	1.150	0.919	0.448	1.762	inactive
chr1:20750, chr12:122500	2.355	2.797	2.540	3.067	1.713	0.587	active
chr1:25750, chr8:124000, chr15:70750, chr11:74000	2.119	1.210	2.020	1.927	2.134	0.727	active
chr9:68250, chr1:143000	0.135	0.000	0.123	0.592	0.133	1.173	inactive
chr9:134750, chr21:38000, chr22:29250, chr12:2000	2.299	1.597	1.721	2.063	2.350	0.737	active
chr2:211500, chr4:179250	0.055	0.000	0.032	0.000	0.002	2.138	inactive
chr4:65000, chr9:69000	0.082	0.000	0.174	0.538	0.097	1.085	inactive
chr4:97500, chr1:143250	0.027	0.000	0.045	0.113	0.000	2.069	inactive
chr11:63750, chr16:67750	4.255	4.689	4.485	4.824	2.781	0.098	active
chr2:13750, chr7:74500	0.069	0.000	0.086	0.125	0.056	1.168	inactive
chr12:53750, chr20:34250	4.557	5.981	3.342	3.659	3.046	0.175	active
chr3:47250, chr12:122250	3.474	2.642	2.998	2.097	2.140	0.265	active
chr17:45500, chr16:70000	3.437	1.180	1.407	1.455	1.193	0.082	active
chr15:75500, chr11:0	2.093	3.742	1.518	2.394	0.810	0.197	active
chr1:239250, chr6:163000	0.008	0.000	0.082	0.150	0.026	2.992	inactive
chr4:178750, chr2:214250, chr2:214500	0.004	0.000	0.128	0.049	0.178	1.321	inactive
chr1:27750, chr3:46750	2.212	1.008	3.330	2.640	2.765	0.589	active
chr16:86250, chr12:2250	1.691	4.227	4.145	1.433	4.180	0.598	active
chr7:102000, chr12:121750	3.615	3.167	2.451	2.288	1.243	0.105	active
chr10:75250, chr15:43250	3.421	1.987	1.505	1.929	1.410	0.260	active
chr16:28750, chr19:11250	3.369	4.231	3.245	4.878	1.669	0.301	active
chr16:30750, chr19:17250, chr16:31000	3.641	5.382	3.843	6.132	1.980	0.325	active
chr6:161000, chr1:239500	0.017	0.000	0.091	0.144	0.012	2.839	inactive
chr10:1500, chr8:2750	0.025	0.000	0.192	0.000	0.076	3.523	inactive
chr6:164750, chr5:166000	0.013	0.000	0.037	0.000	0.000	3.116	inactive

Table S4: Fold enrichment of high-occupancy target (HOT) regions as compared to low-occupancy target regions (LOT) in active clusters.

Feature	Enrichment
HOT:LOT	2.94

Table S5: Top 15 inactive clusters, ranked by enrichment for H3K9me3. Clusters are given by chromosome number and start position in kb; each region in the cluster is 250kb in length. Only clusters with less than seven 250kb regions in the cluster are included.

Cluster #	**Clusters (kb)**	**H3K9me3**	**RNAPII**
443	chr10:1500, chr8:2750	3.523	0.000
114	chr10:125000, chr1:5750, chr10:125250	3.377	0.000
204	chr8:137000, chr6:164500	3.353	0.000
213	chr17:11500, chr17:10750, chr5:6250, chr17:11000, chr17:11250	3.264	0.079
260	chr2:0, chr5:178250	3.225	0.431
370	chr22:26250, chr20:22000, chr20:21750	3.180	0.000
63	chr6:170250, chr1:13750	3.152	0.000
316	chr1:238250, chr6:162250	3.127	0.000
445	chr6:164750, chr5:166000	3.116	0.000
49	chr1:34750, chr1:34000, chr1:34250, chr8:142500, chr1:35000, chr1:34500	3.054	0.000
265	chr19:30750, chr20:19250	3.048	0.000
251	chr6:165250, chr8:138500	3.019	0.000
235	chr21:32250, chr20:15250	3.001	0.431
230	chr14:104500, chr14:104750, chr13:112000	2.994	0.000
434	chr1:239250, chr6:163000	2.992	0.000

Table S6: Top 15 active clusters, ranked by p-value of permutation test based on TFBS (JASPAR 2016, threshold = 0.000001, CAGE). Clusters are given by chromosome number and start position in kb; each region in the cluster is 250kb in length.

Cluster #	**Clusters (kb)**	**P-value**	**RNAPII**	**H3K9me3**
137	chr7:100250, chr7:100750, chr11:65250, chr17:42750	0.0000	10.1162	0.5417
111	chr1:22250, chr19:45250	0.0000	4.3250	0.2820
61	chr16:30000, chr19:10750, chr16:29750, chr19:10500	0.0002	4.9110	0.3573
57	chr19:56000, chr19:55500, chr22:50250, chr19:55750	0.0001	8.1535	0.9214
157	chr22:50750, chr17:7250	0.0005	8.2122	0.3790
27	chr20:48750, chr20:48500, chr21:47250, chr22:30500	0.0021	13.5989	0.4189
29	chr12:53500, chr4:1000, chr12:53250	0.0024	5.0042	0.5440
144	chr12:50000, chr6:35250	0.0031	5.2105	0.3111
92	chr1:1000, chr12:123250, chr1:750, chr9:133500, chr8:145000, chr13:114750	0.0043	9.0570	0.5125
180	chr1:27750, chr3:46750	0.0068	1.0079	0.5889
84	chr17:73000, chr8:144250	0.0068	3.5725	0.5660
68	chr2:220250, chr1:16750, chr2:220000	0.0070	5.4663	0.4099
185	chr16:30750, chr19:17250, chr16:31000	0.0073	5.3819	0.3246
17	chr12:49250, chr17:38250, chr12:49500	0.0075	6.0798	0.2533
78	chr1:203250, chr17:74500	0.0103	7.9689	0.5743

SI Figures

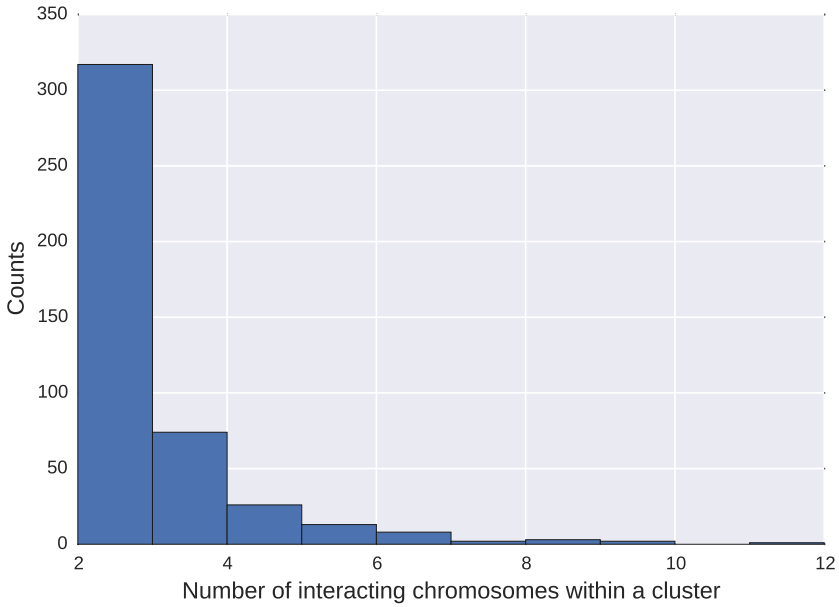


Figure S1: Number of chromosomes within an intermingling cluster.

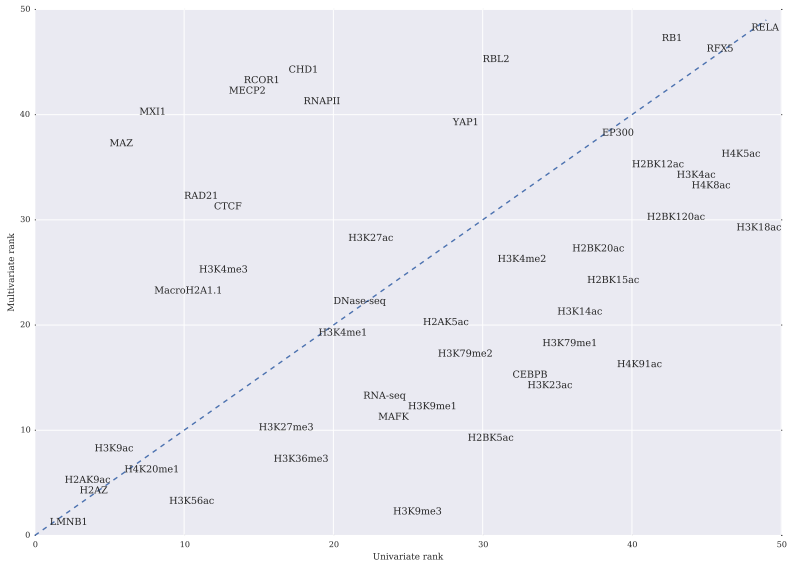
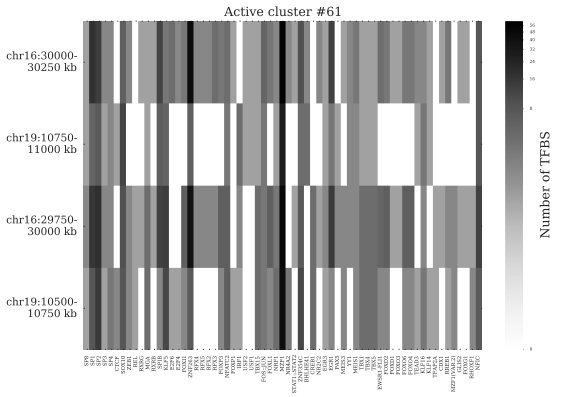
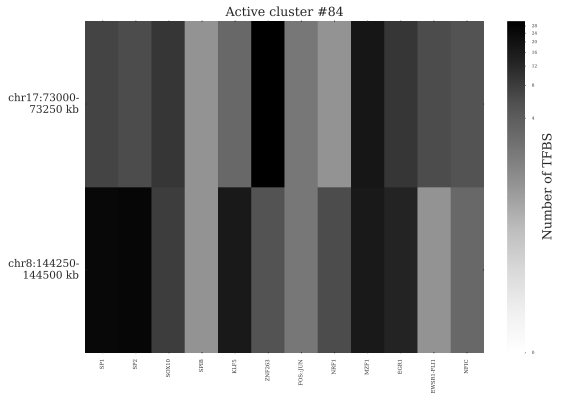
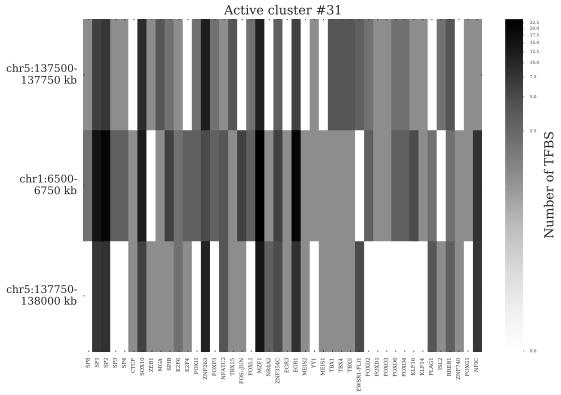


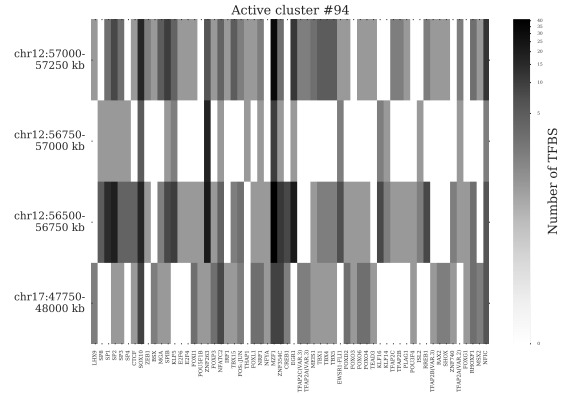
Figure S2: Scatterplot of univariate rank of a feature (two-sample Kolmogorov–Smirnov test) versus predictive rank of a feature when it is combined with all other features (relative importance in decision trees) for classification of intermingling versus non-intermingling regions. A similar analysis has been performed in (49) to explore context-dependency of features for classification.

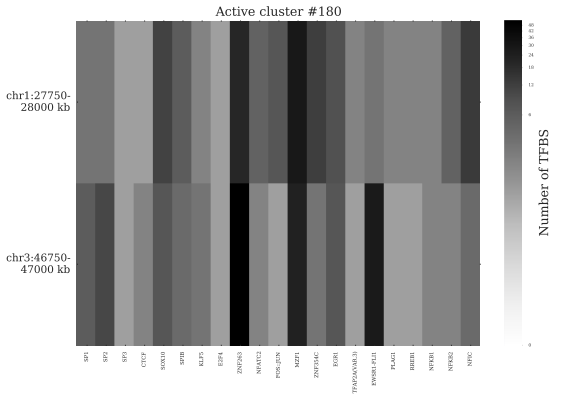
Figure S3: Comparison of gene expression in reads per kilobase of transcript per million mapped reads (RPKM) on log scale between active and inactive clusters. Active clusters show significantly higher gene expression (p-value = 0.004 under t-test).

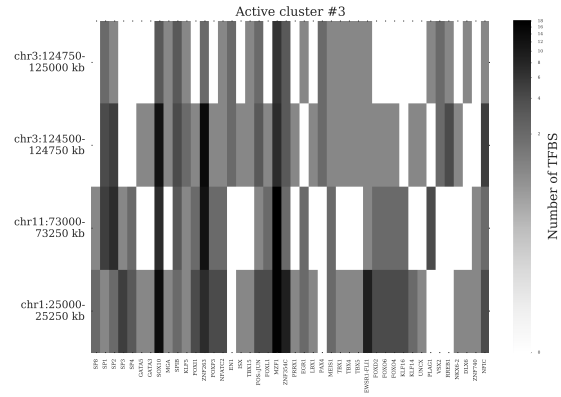

Active cluster #61
chr16:30000-30250 kb
chr19:10750-11000 kb
chr16:29750-30000 kb
chr19:10500-10750 kb
Number of TFBS

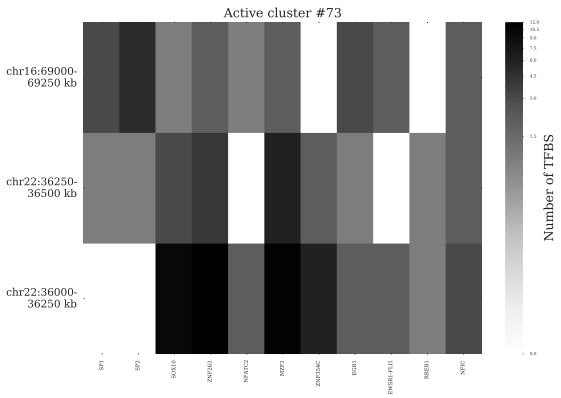

Active cluster #84
chr17:73000-73250 kb
chr8:144250-144500 kb
Number of TFBS

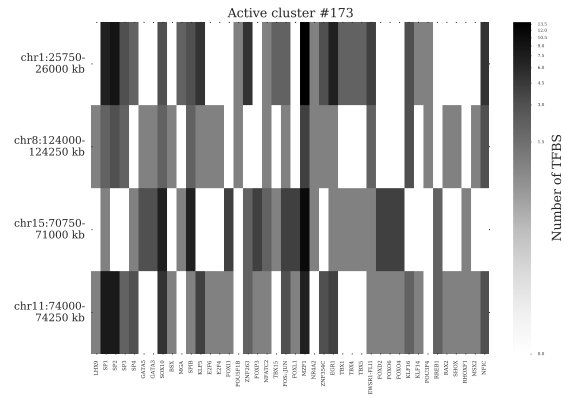

Active cluster #31
chr5:137500-137750 kb
chr1:6500-6750 kb
chr5:137750-138000 kb
Number of TFBS

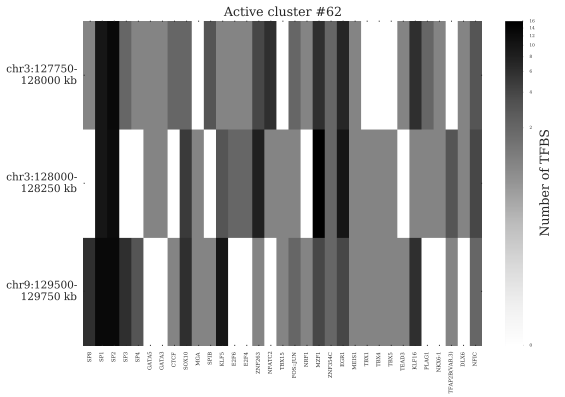

Active cluster #94
chr12:57000-57250 kb
chr12:56750-57000 kb
chr12:56500-56750 kb
chr17:47750-48000 kb
Number of TFBS

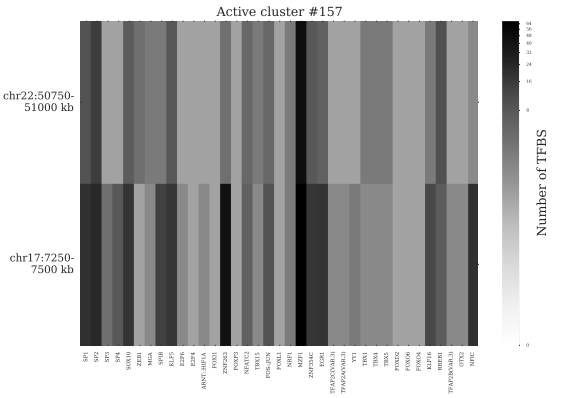

Active cluster #180
chr1:27750-28000 kb
chr3:46750-47000 kb
Number of TFBS

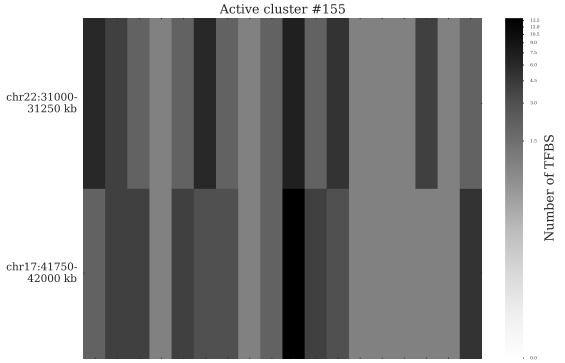

Active cluster #3
chr3:124750-125000 kb
chr3:124500-124750 kb
chr11:73000-73250 kb
chr1:25000-25250 kb
Number of TFBS

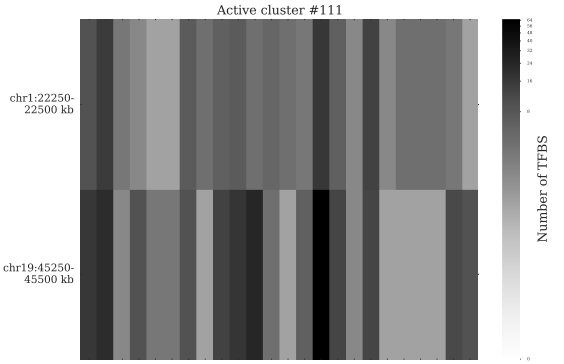

Active cluster #73
chr16:69000-69250 kb
chr22:36250-36500 kb
chr22:36000-36250 kb
Number of TFBS

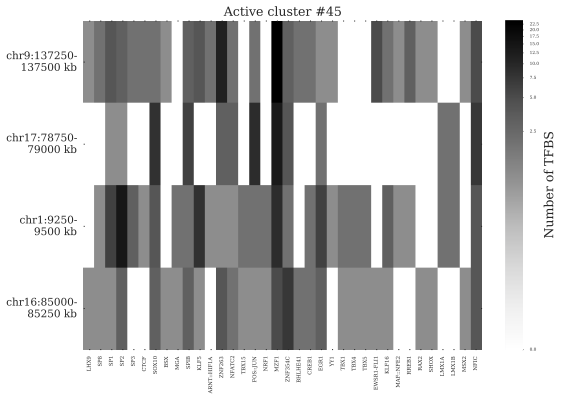

Active cluster #173
chr1:25750-26000 kb
chr8:124000-124250 kb
chr15:70750-71000 kb
chr11:74000-74250 kb
Number of TFBS

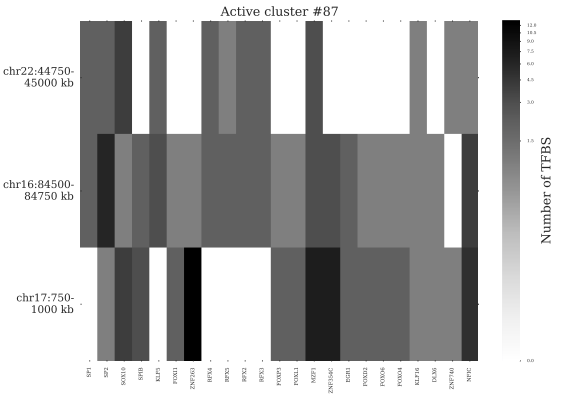

Active cluster #62
chr3:127750-128000 kb
chr3:128000-128250 kb
chr9:129500-129750 kb
Number of TFBS

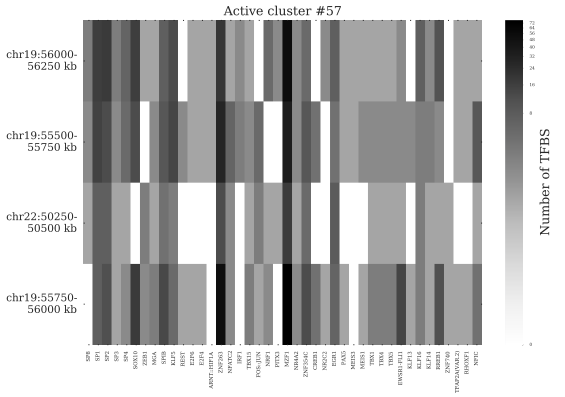

Active cluster #157
chr22:50750-51000 kb
chr17:7250-7500 kb
Number of TFBS

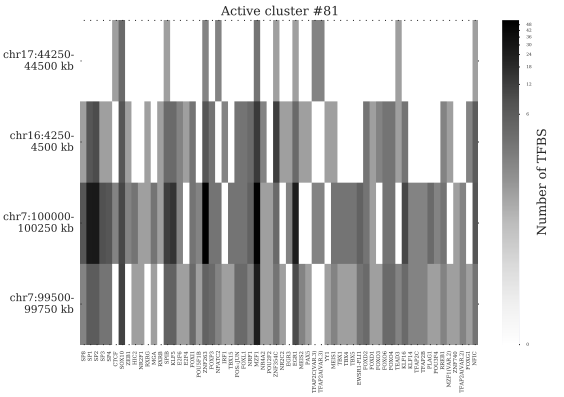

Active cluster #155
chr22:31000-31250 kb
chr17:41750-42000 kb
Number of TFBS

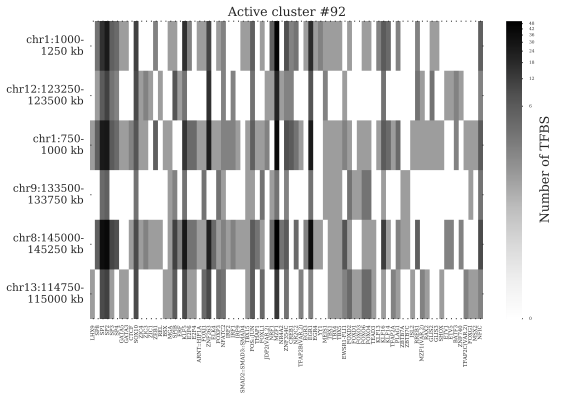

Active cluster #111
chr1:22250-22500 kb
chr19:45250-45500 kb
Number of TFBS

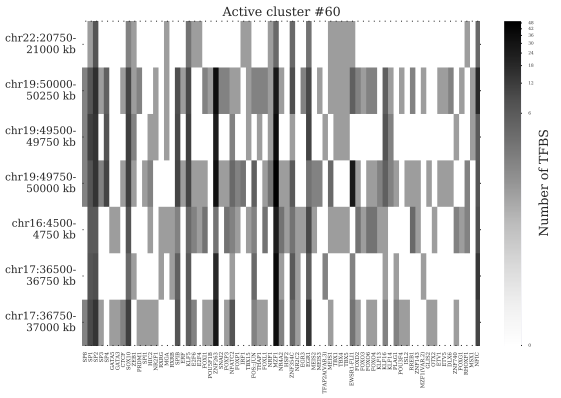

Active cluster #45
chr9:137250-137500 kb
chr17:78750-79000 kb
chr1:9250-9500 kb
chr16:85000-85250 kb
Number of TFBS

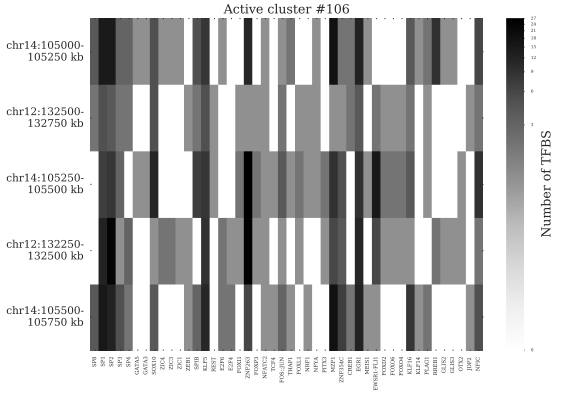

Active cluster #87
chr22:44750-45000 kb
chr16:84500-84750 kb
chr17:750-1000 kb
Number of TFBS

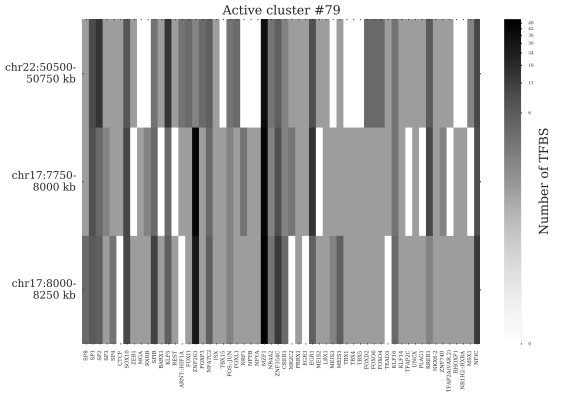

Active cluster #57
chr19:56000-56250 kb
chr19:55500-55750 kb
chr22:50250-50500 kb
chr19:55750-56000 kb
Number of TFBS

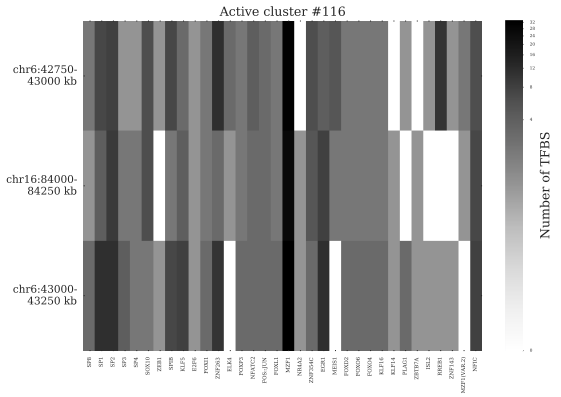

Active cluster #81
chr17:44250-44500 kb
chr16:4250-4500 kb
chr7:100000-100250 kb
chr7:99500-99750 kb
Number of TFBS

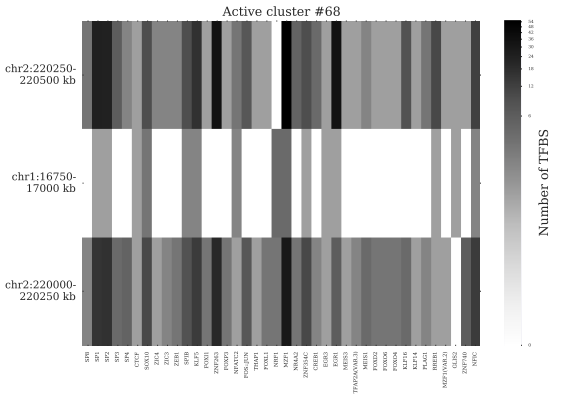

Active cluster #92
chr1:1000-1250 kb
chr12:123250-123500 kb
chr1:750-1000 kb
chr9:133500-133750 kb
chr8:145000-145250 kb
chr13:114750-115000 kb
Number of TFBS

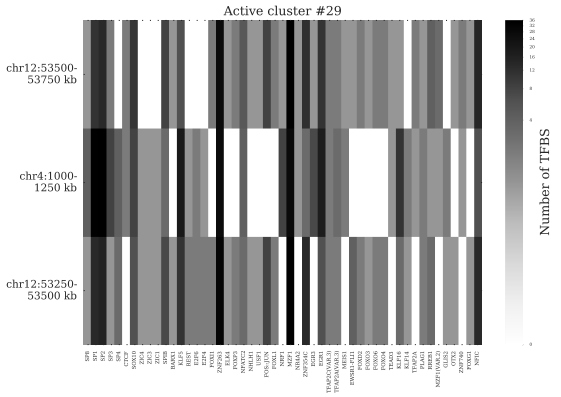

Active cluster #60
chr22:20750-21000 kb
chr19:50000-50250 kb
chr19:49500-49750 kb
chr19:49750-50000 kb
chr16:4500-4750 kb
chr17:36500-36750 kb
chr17:36750-37000 kb
Number of TFBS

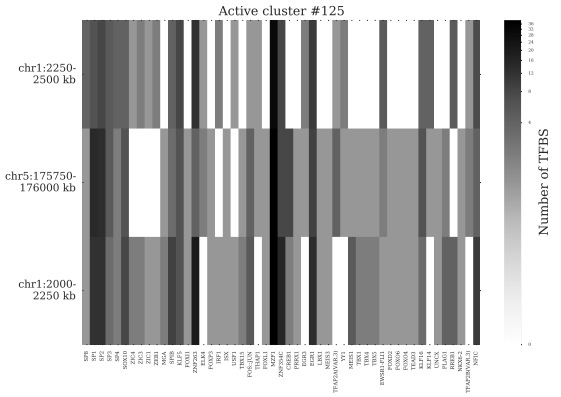

Active cluster #106
chr14:105000-105250 kb
chr12:132500-132750 kb
chr14:105250-105500 kb
chr12:132250-132500 kb
chr14:105500-105750 kb
Number of TFBS

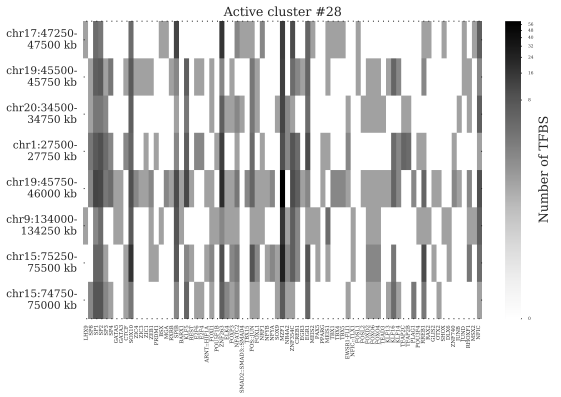

Active cluster #79
chr22:50500-50750 kb
chr17:7750-8000 kb
chr17:8000-8250 kb
Number of TFBS

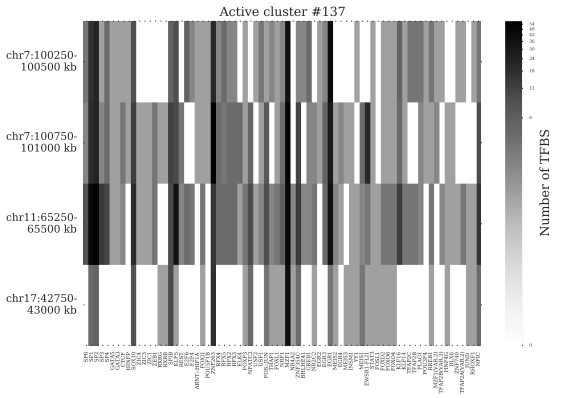

Active cluster #116
chr6:42750-43000 kb
chr16:84000-84250 kb
chr6:43000-43250 kb
Number of TFBS

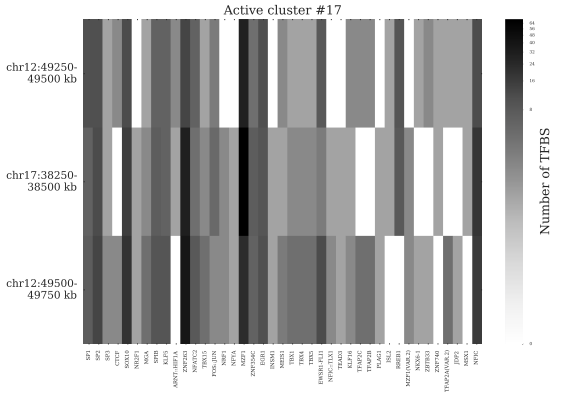

Active cluster #68
chr2:220250-220500 kb
chr1:16750-17000 kb
chr2:220000-220250 kb
Number of TFBS

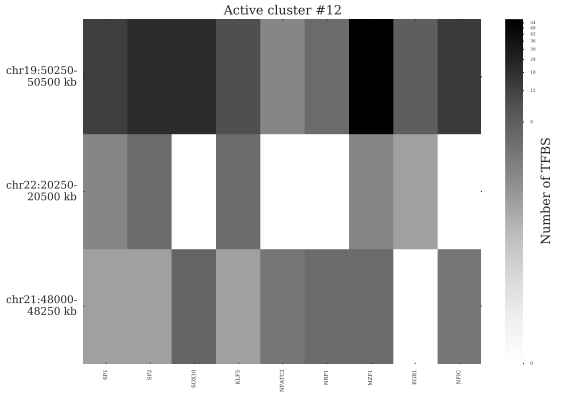

Active cluster #29
chr12:53500-53750 kb
chr4:1000-1250 kb
chr12:53250-53500 kb
Number of TFBS

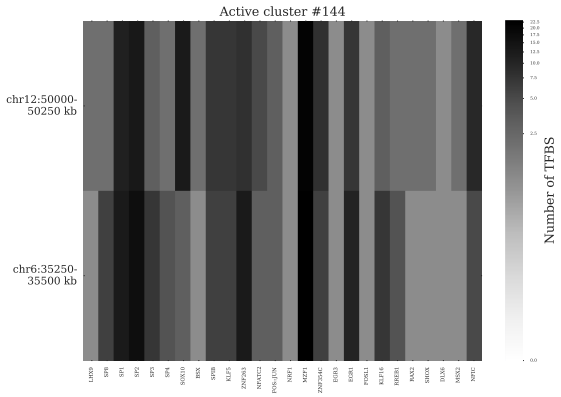

Active cluster #125
chr1:2250-2500 kb
chr5:175750-176000 kb
chr1:2000-2250 kb
Number of TFBS

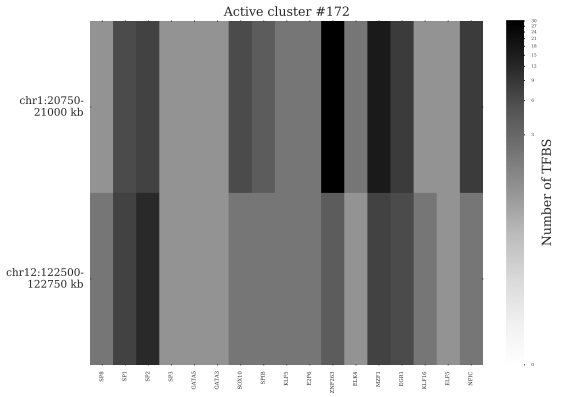

Active cluster #28
chr17:47250-47500 kb
chr19:45500-45750 kb
chr20:34500-34750 kb
chr1:27500-27750 kb
chr19:45750-46000 kb
chr9:134000-134250 kb
chr15:75250-75500 kb
chr15:74750-75000 kb
Number of TFBS

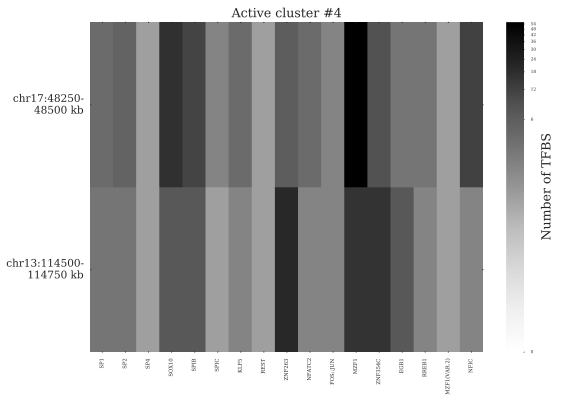

Active cluster #137
chr7:100250-100500 kb
chr7:100750-101000 kb
chr11:65250-65500 kb
chr17:42750-43000 kb
Number of TFBS

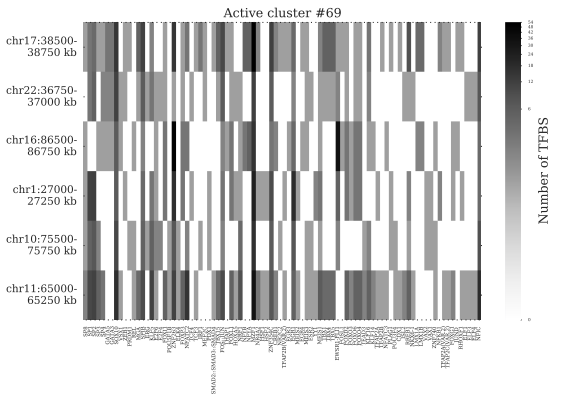

Active cluster #17
chr12:49250-49500 kb
chr17:38250-38500 kb
chr12:49500-49750 kb
Number of TFBS


Active cluster #12
chr19:50250-50500 kb
chr22:20250-20500 kb
chr21:48000-48250 kb
Number of TFBS


Active cluster #144
chr12:50000-50250 kb
chr6:35250-35500 kb
Number of TFBS


Active cluster #172
chr1:20750-21000 kb
chr12:122500-122750 kb
Number of TFBS


Active cluster #4
chr17:48250-48500 kb
chr13:114500-114750 kb
Number of TFBS


Active cluster #69
chr17:38500-38750 kb
chr22:36750-37000 kb
chr16:86500-86750 kb
chr1:27000-27250 kb
chr10:75500-75750 kb
chr11:65000-65250 kb
Number of TFBS

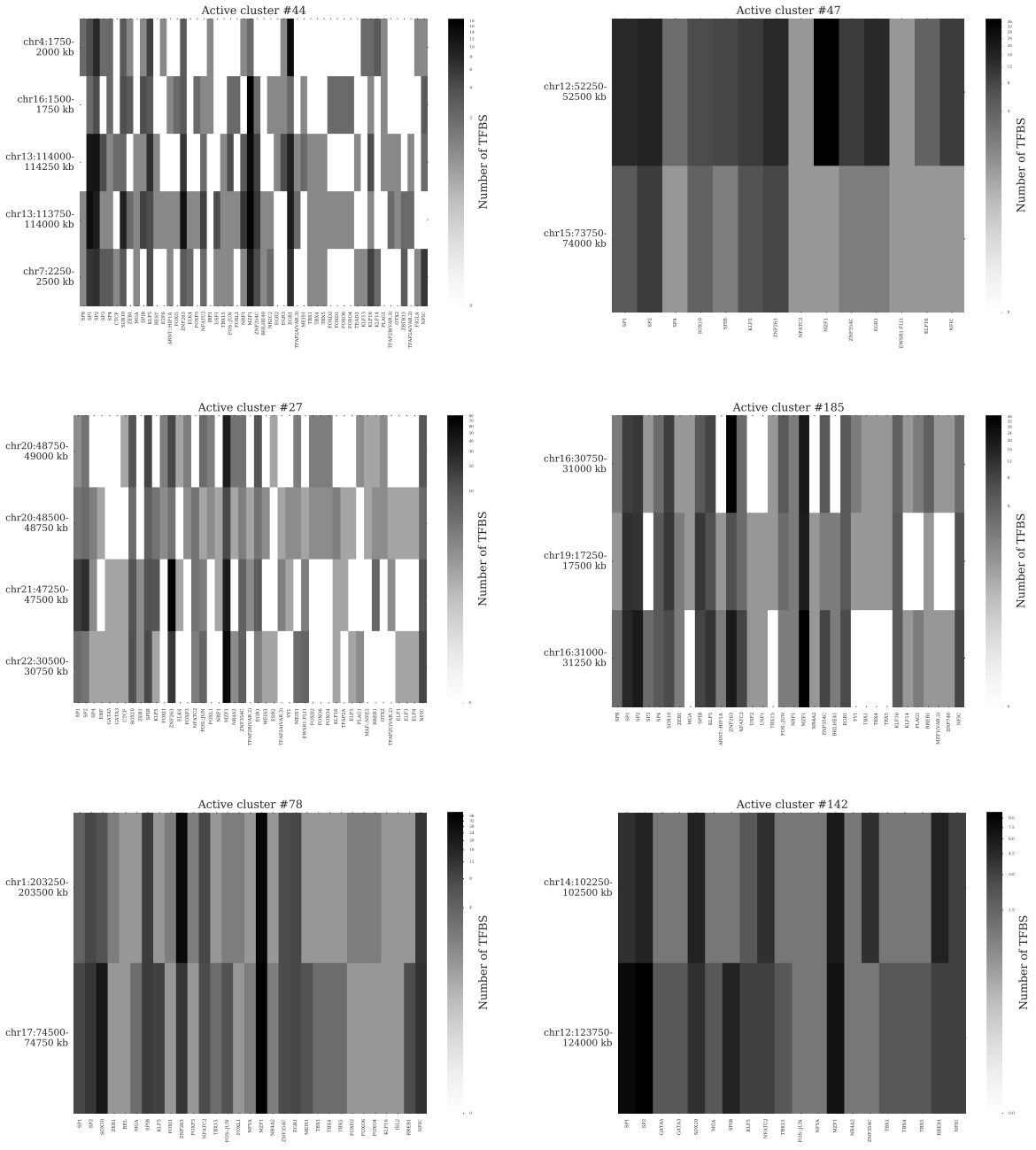
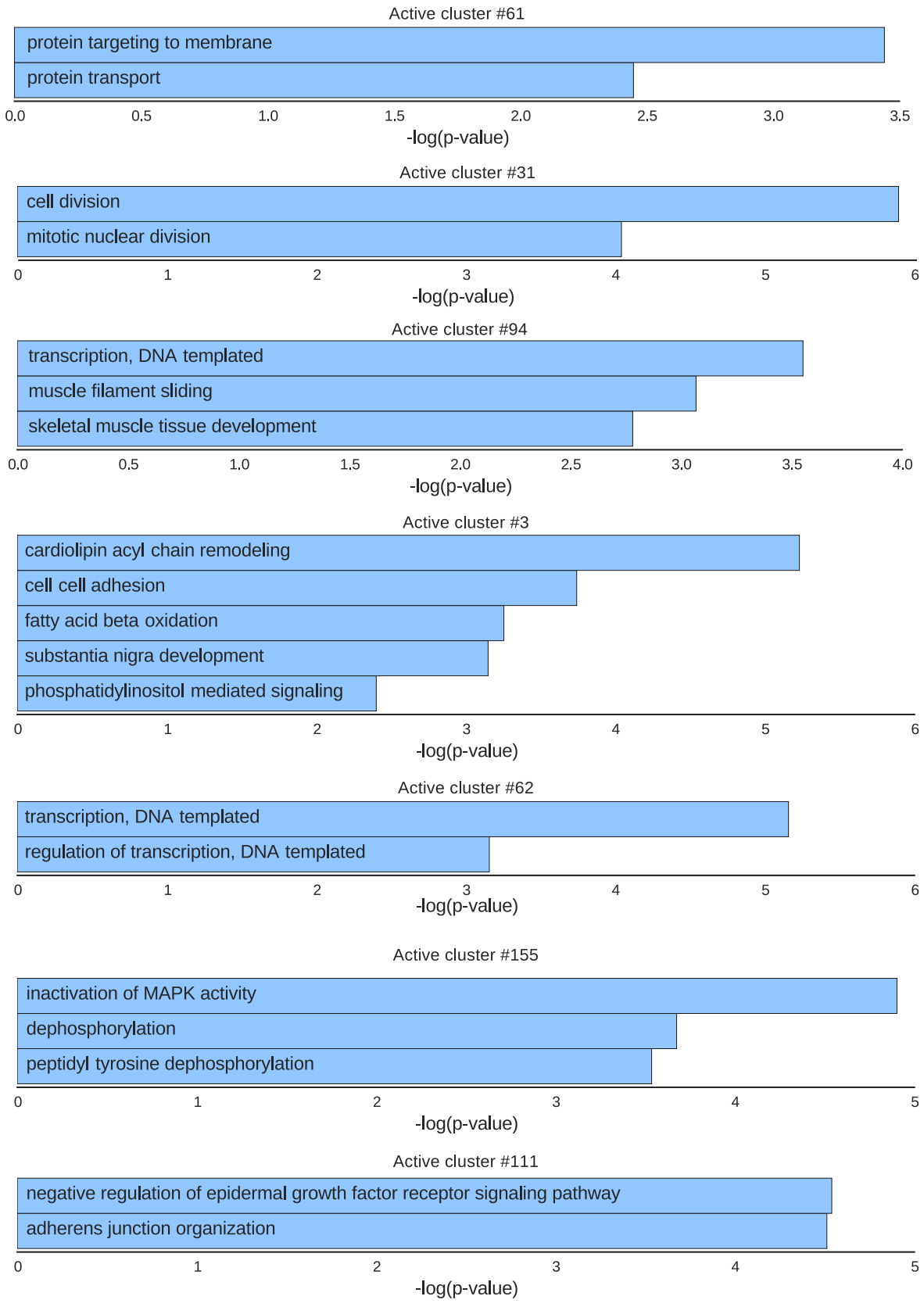
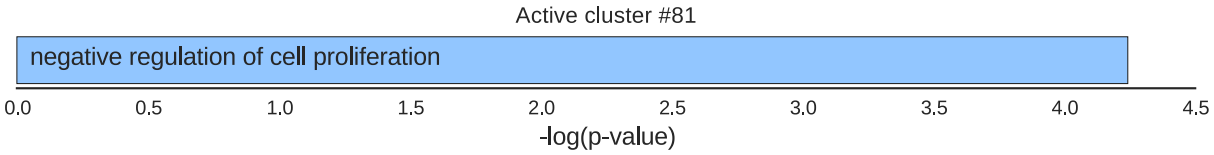
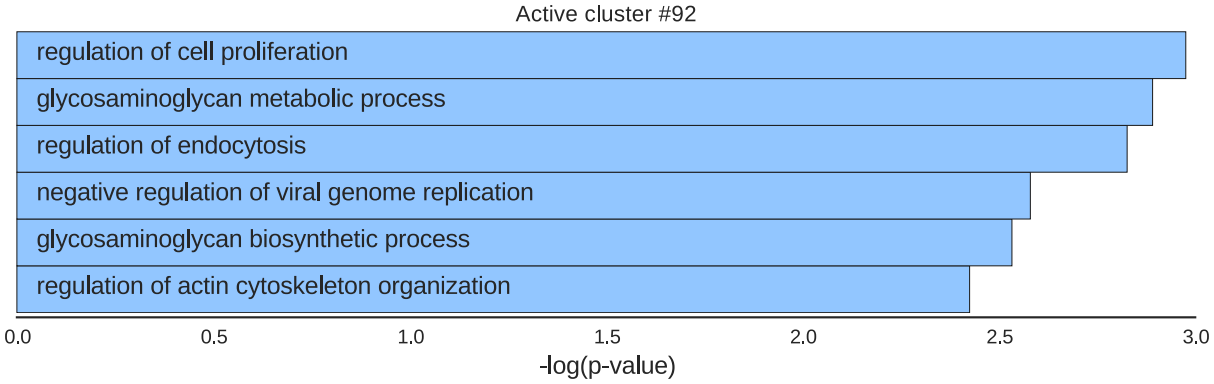
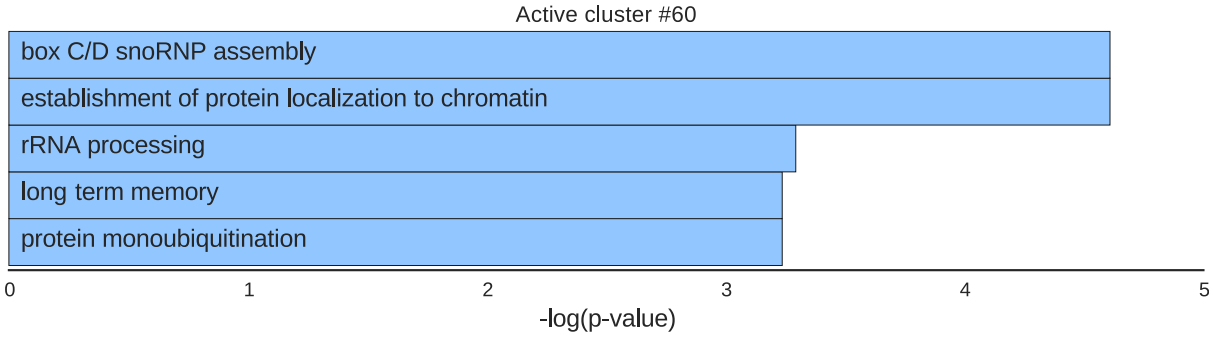


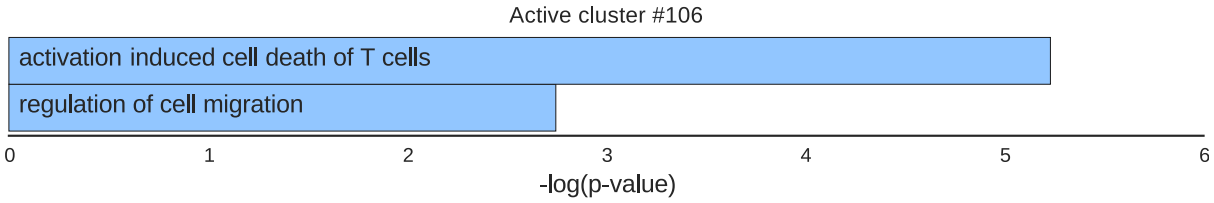
Figure S4: TFBS for top 15 active clusters given in Fig. S7 (JASPAR 2016, threshold = 0.000001, CAGE). Only TFs that span at least 2 regions in the cluster are shown.

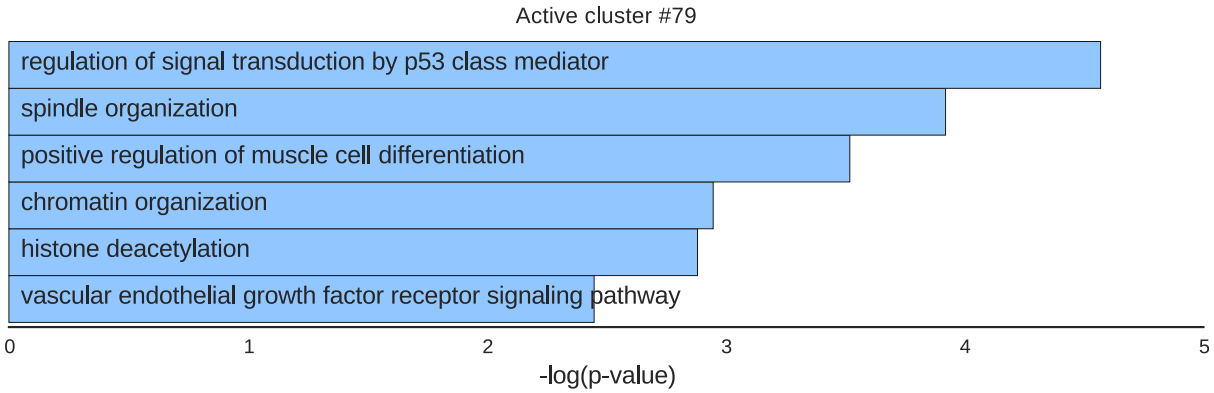

Active cluster #61
protein targeting to membrane
protein transport
0.0 0.5 1.0 1.5 2.0 2.5 3.0 3.5
-log(p-value)
Active cluster #31
cell division
mitotic nuclear division
0 1 2 3 4 5 6
-log(p-value)
Active cluster #94
transcription, DNA templated
muscle filament sliding
skeletal muscle tissue development
0.0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0
-log(p-value)
Active cluster #3
cardiolipin acyl chain remodeling
cell cell adhesion
fatty acid beta oxidation
substantia nigra development
phosphatidylinositol mediated signaling
0 1 2 3 4 5 6
-log(p-value)
Active cluster #62
transcription, DNA templated
regulation of transcription, DNA templated
0 1 2 3 4 5 6
-log(p-value)
Active cluster #155
inactivation of MAPK activity
dephosphorylation
peptidyl tyrosine dephosphorylation
0 1 2 3 4 5
-log(p-value)
Active cluster #111
negative regulation of epidermal growth factor receptor signaling pathway
adherens junction organization
0 1 2 3 4 5
-log(p-value)

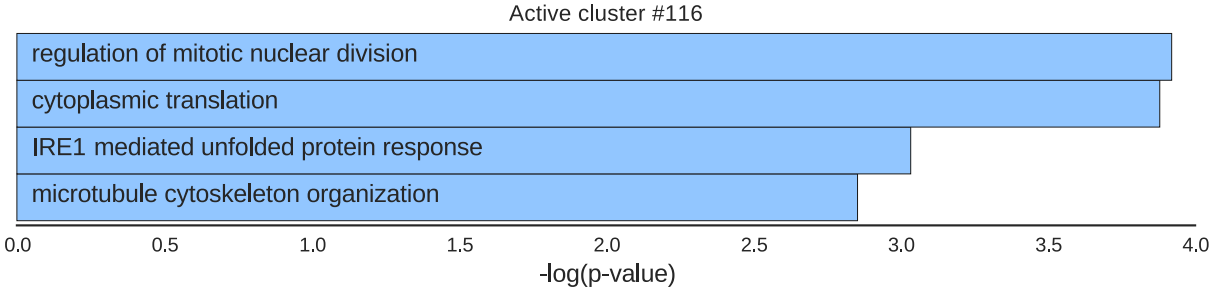

Active cluster #81
negative regulation of cell proliferation
0.0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5
-log(p-value)

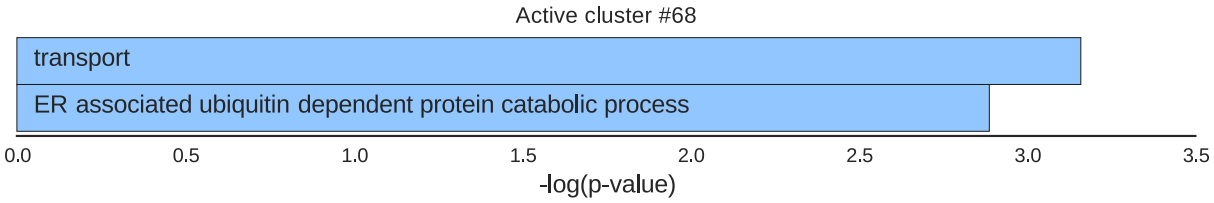

Active cluster #92
regulation of cell proliferation
glycosaminoglycan metabolic process
regulation of endocytosis
negative regulation of viral genome replication
glycosaminoglycan biosynthetic process
regulation of actin cytoskeleton organization
0.0 0.5 1.0 1.5 2.0 2.5 3.0
-log(p-value)

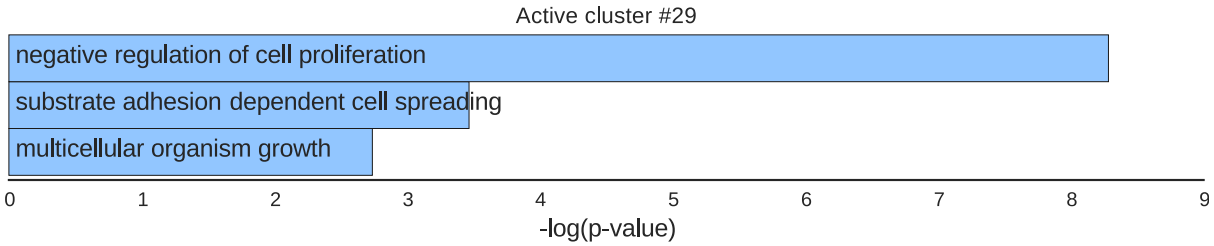

Active cluster #60
box C/D snoRNP assembly
establishment of protein localization to chromatin
rRNA processing
long term memory
protein monoubiquitination
0 1 2 3 4 5
-log(p-value)

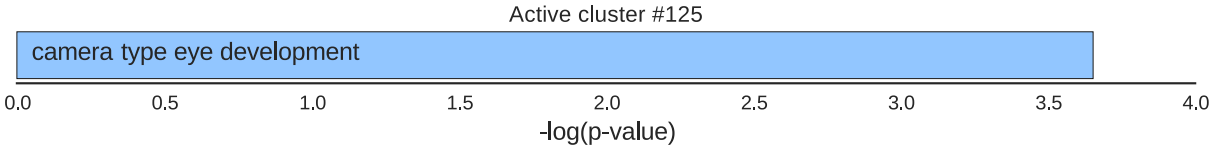

Active cluster #106
activation induced cell death of T cells
regulation of cell migration
0 1 2 3 4 5 6
-log(p-value)

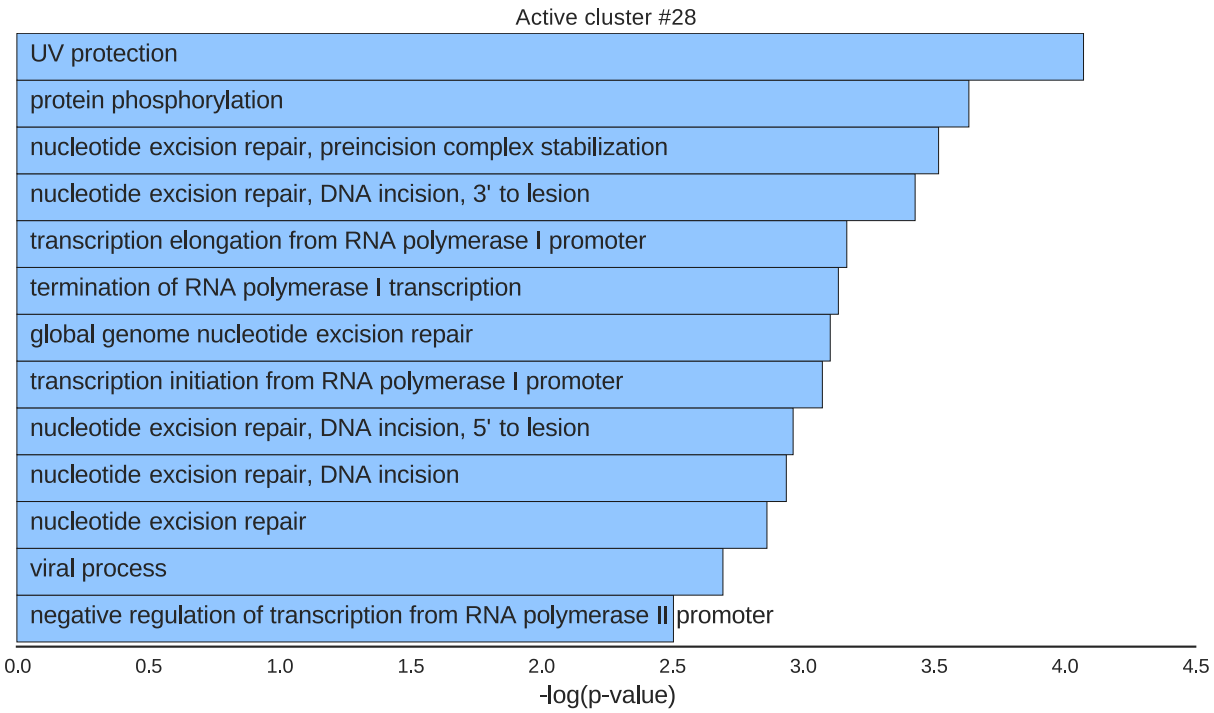

Active cluster #79
regulation of signal transduction by p53 class mediator
spindle organization
positive regulation of muscle cell differentiation
chromatin organization
histone deacetylation
vascular endothelial growth factor receptor signaling pathway
0 1 2 3 4 5
-log(p-value)

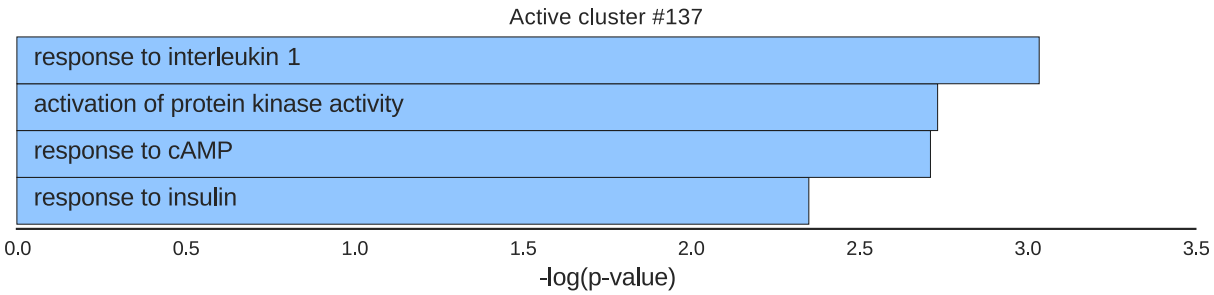

Active cluster #116
regulation of mitotic nuclear division
cytoplasmic translation
IRE1 mediated unfolded protein response
microtubule cytoskeleton organization
0.0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0
-log(p-value)

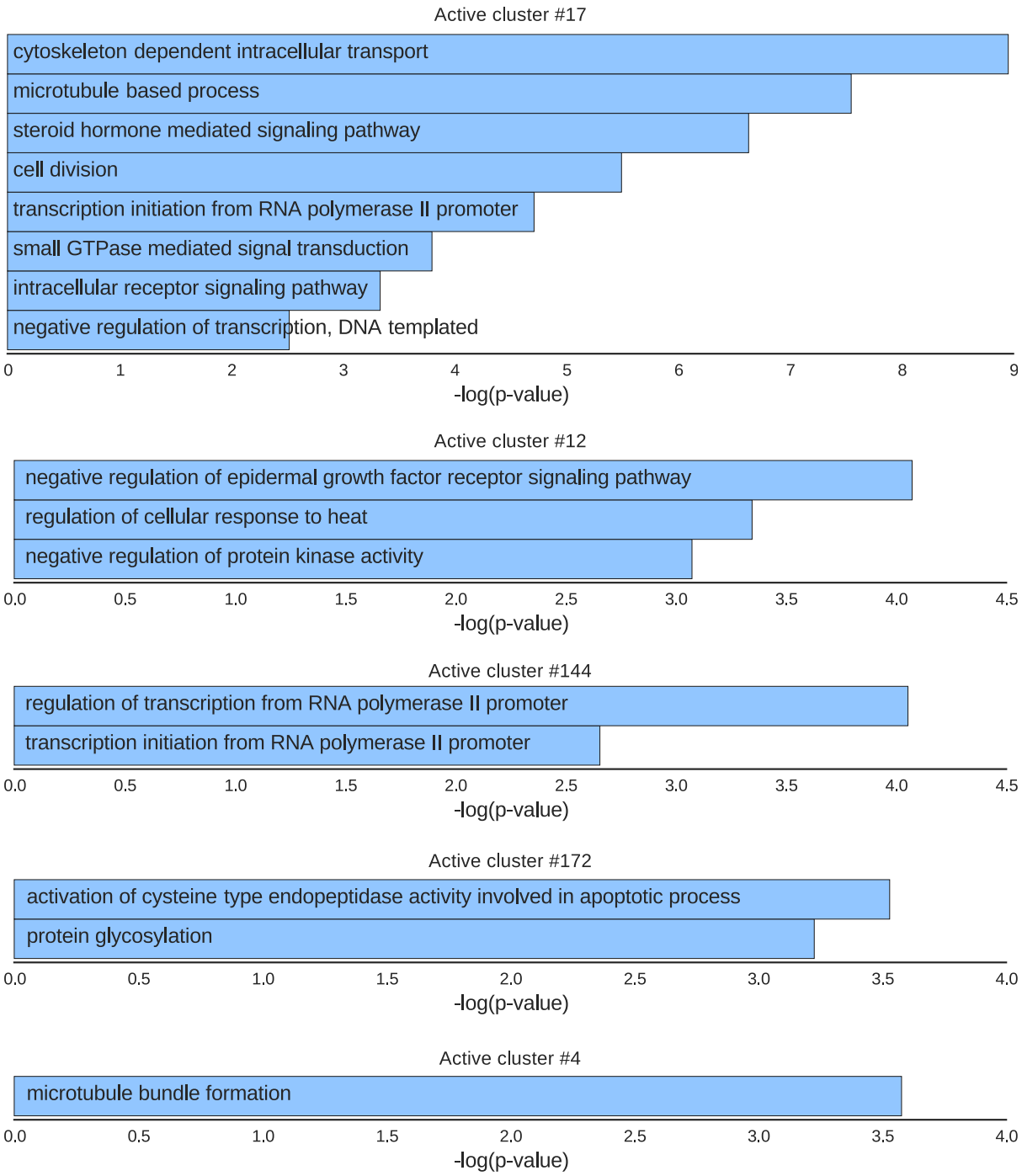

Active cluster #68
transport
ER associated ubiquitin dependent protein catabolic process
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
-log(p-value)

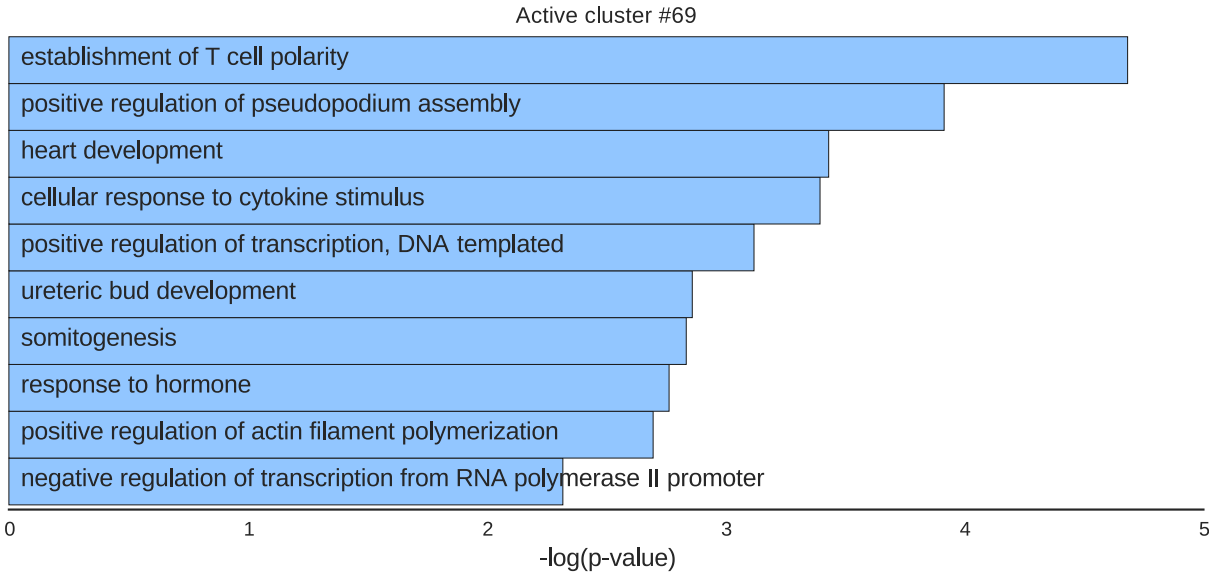

Active cluster #29
negative regulation of cell proliferation
substrate adhesion dependent cell spreading
multicellular organism growth
0
1
2
3
4
5
6
7
8
9
-log(p-value)

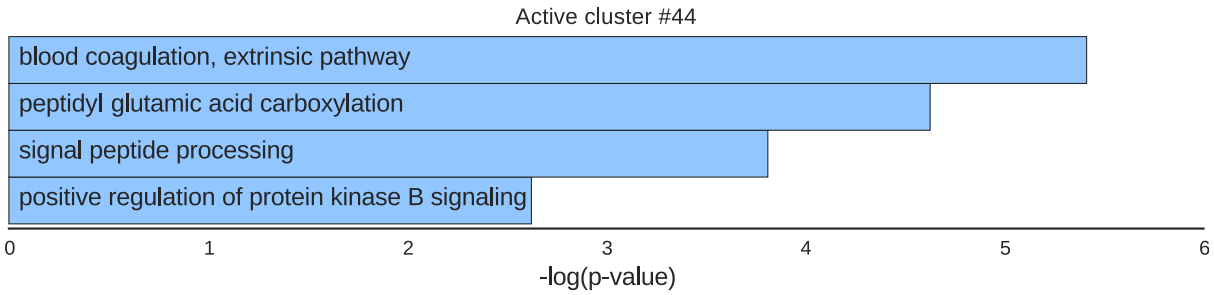

Active cluster #125
camera type eye development
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
-log(p-value)

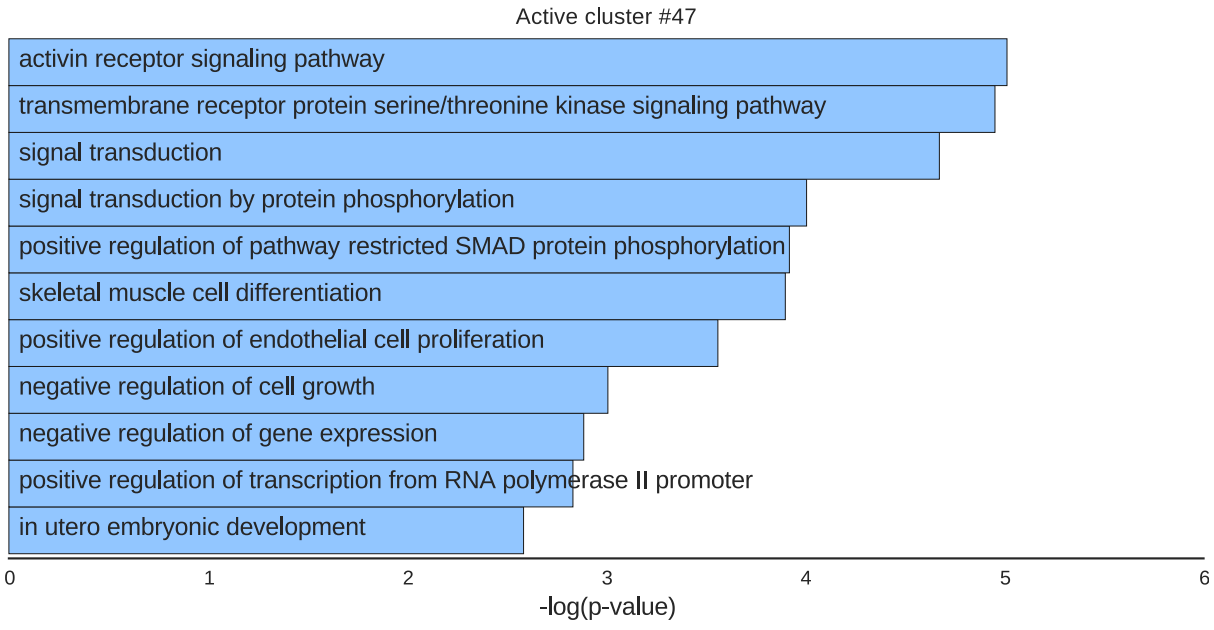

Active cluster #28
UV protection
protein phosphorylation
nucleotide excision repair, preincision complex stabilization
nucleotide excision repair, DNA incision, 3' to lesion
transcription elongation from RNA polymerase I promoter
termination of RNA polymerase I transcription
global genome nucleotide excision repair
transcription initiation from RNA polymerase I promoter
nucleotide excision repair, DNA incision, 5' to lesion
nucleotide excision repair, DNA incision
nucleotide excision repair
viral process
negative regulation of transcription from RNA polymerase II promoter
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
-log(p-value)


Active cluster #137
response to interleukin 1
activation of protein kinase activity
response to cAMP
response to insulin
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
-log(p-value)


Active cluster #17
cytoskeleton dependent intracellular transport
microtubule based process
steroid hormone mediated signaling pathway
cell division
transcription initiation from RNA polymerase II promoter
small GTPase mediated signal transduction
intracellular receptor signaling pathway
negative regulation of transcription, DNA templated
0 1 2 3 4 5 6 7 8 9
-log(p-value)
Active cluster #12
negative regulation of epidermal growth factor receptor signaling pathway
regulation of cellular response to heat
negative regulation of protein kinase activity
0.0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5
-log(p-value)
Active cluster #144
regulation of transcription from RNA polymerase II promoter
transcription initiation from RNA polymerase II promoter
0.0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5
-log(p-value)
Active cluster #172
activation of cysteine type endopeptidase activity involved in apoptotic process
protein glycosylation
0.0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0
-log(p-value)
Active cluster #4
microtubule bundle formation
0.0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0
-log(p-value)


Active cluster #69
establishment of T cell polarity
positive regulation of pseudopodium assembly
heart development
cellular response to cytokine stimulus
positive regulation of transcription, DNA templated
ureteric bud development
somitogenesis
response to hormone
positive regulation of actin filament polymerization
negative regulation of transcription from RNA polymerase II promoter
0
1
2
3
4
5
-log(p-value)


Active cluster #44
blood coagulation, extrinsic pathway
peptidyl glutamic acid carboxylation
signal peptide processing
positive regulation of protein kinase B signaling
0
1
2
3
4
5
6
-log(p-value)


Active cluster #47
activin receptor signaling pathway
transmembrane receptor protein serine/threonine kinase signaling pathway
signal transduction
signal transduction by protein phosphorylation
positive regulation of pathway restricted SMAD protein phosphorylation
skeletal muscle cell differentiation
positive regulation of endothelial cell proliferation
negative regulation of cell growth
negative regulation of gene expression
positive regulation of transcription from RNA polymerase II promoter
in utero embryonic development
0
1
2
3
4
5
6
-log(p-value)

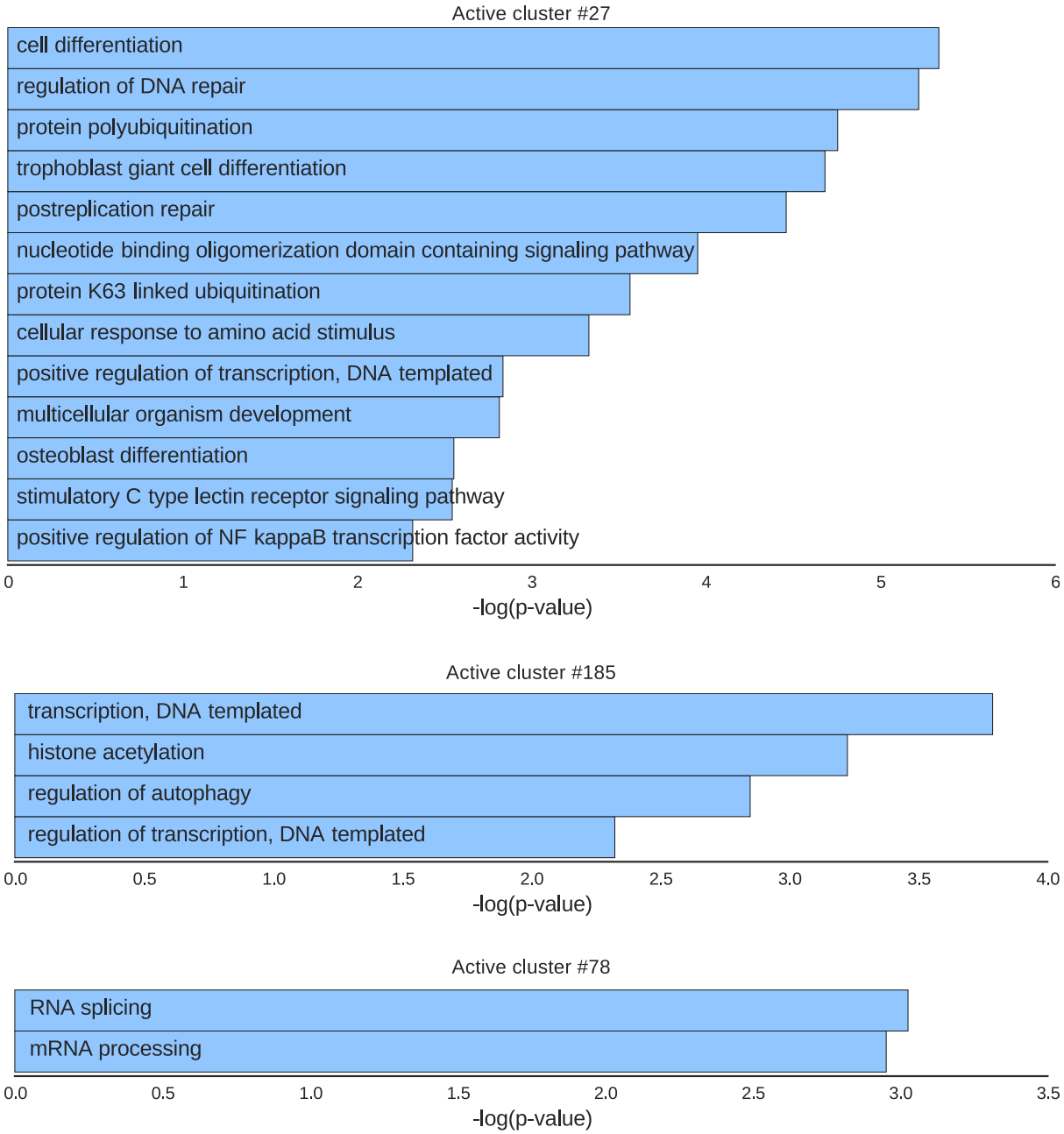
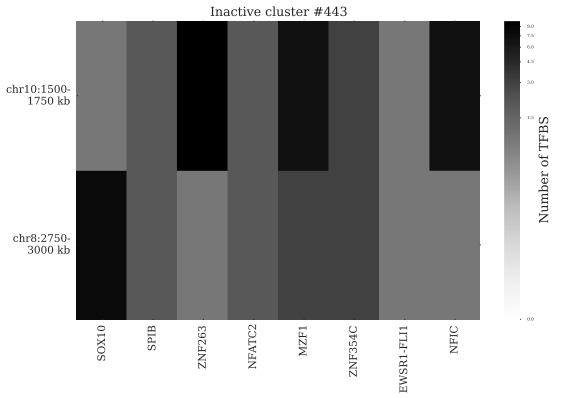
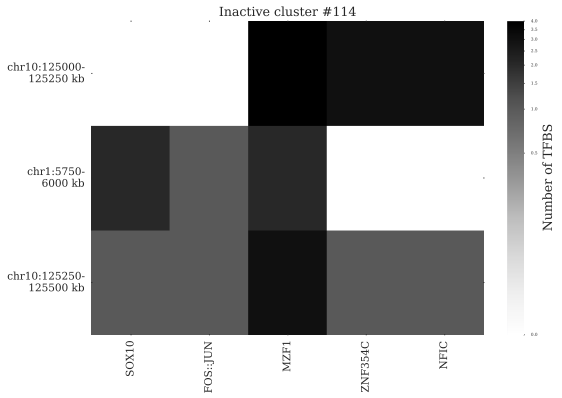
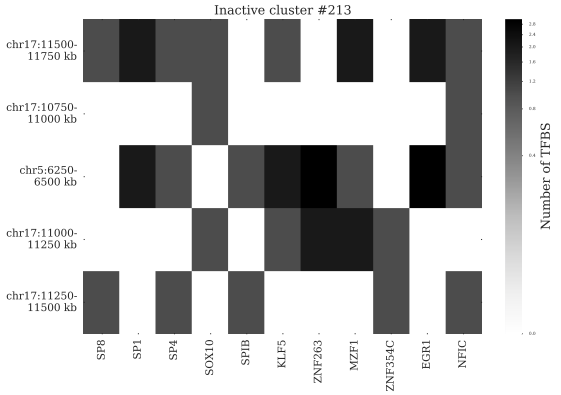
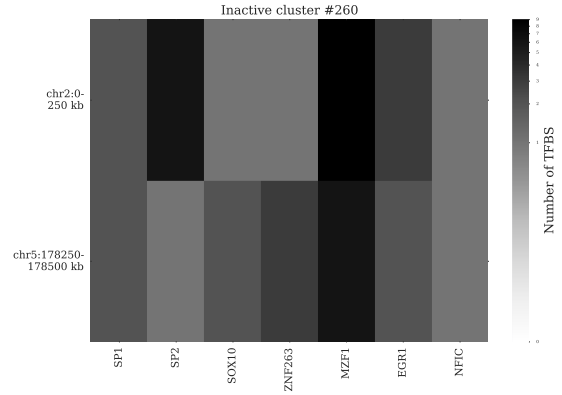


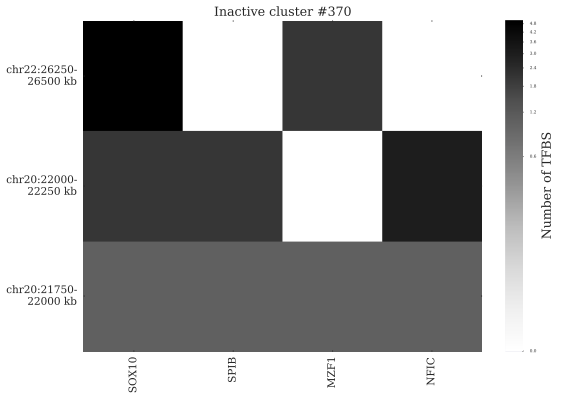
Figure S5: GO terms associated with top active clusters in Fig. S7. Only clusters that have less than 10 chromosomes in them and that had GO terms associated with the cluster are displayed.

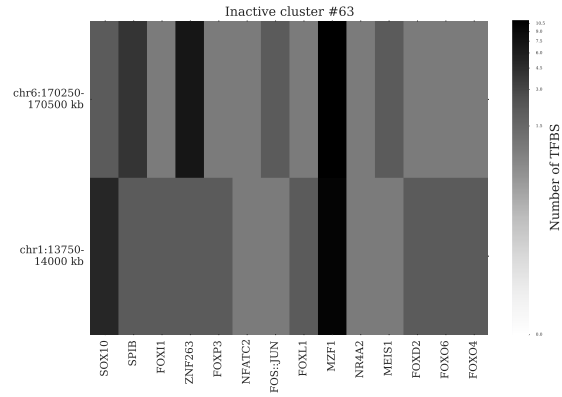

Inactive cluster #443
chr10:1500-1750 kb
chr8:2750-3000 kb
SOX10
SPIB
ZNF263
NFATC2
MZF1
ZNF354C
EWSR1-FLI1
NFIC
Number of TFBS

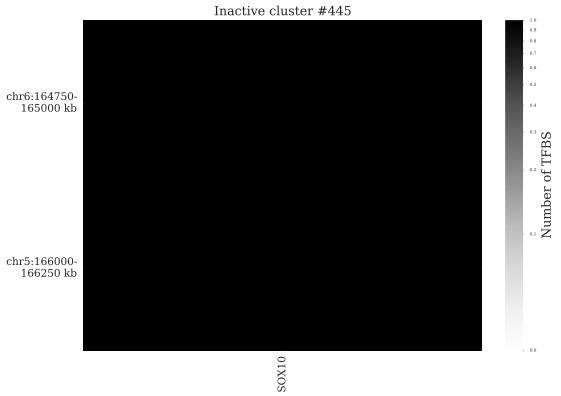

Inactive cluster #114
chr10:125000-125250 kb
chr1:5750-6000 kb
chr10:125250-125500 kb
SOX10
FOS::JUN
MZF1
ZNF354C
NFIC
Number of TFBS

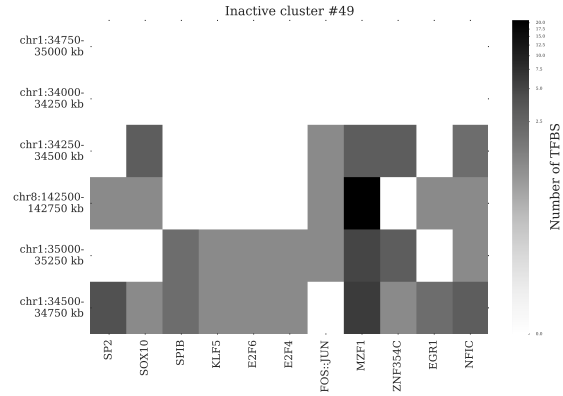

Inactive cluster #213
chr17:11500-11750 kb
chr17:10750-11000 kb
chr5:6250-6500 kb
chr17:11000-11250 kb
chr17:11250-11500 kb
SP8
SP1
SP4
SOX10
SPIB
KLF5
ZNF263
MZF1
ZNF354C
EGR1
NFIC
Number of TFBS


Inactive cluster #260
chr2:0-250 kb
chr5:178250-178500 kb
SP1
SP2
SOX10
ZNF263
MZF1
EGR1
NFIC
Number of TFBS


Inactive cluster #370
chr22:26250-26500 kb
chr20:22000-22250 kb
chr20:21750-22000 kb
SOX10
SPIB
MZF1
NFIC
Number of TFBS


Inactive cluster #63
chr6:170250-170500 kb
chr1:13750-14000 kb
SOX10
SPIB
FOXI1
ZNF263
FOXP3
NFATC2
FOS::JUN
FOXL1
MZF1
NR4A2
MEIS1
FOXD2
FOXO6
FOXO4
Number of TFBS


Inactive cluster #445
chr6:164750-165000 kb
chr5:166000-166250 kb
SOX10
Number of TFBS


Inactive cluster #49
chr1:34750-35000 kb
chr1:34000-34250 kb
chr1:34250-34500 kb
chr8:142500-142750 kb
chr1:35000-35250 kb
chr1:34500-34750 kb
SP2
SOX10
SPIB
KLF5
E2F6
E2F4
FOS::JUN
MZF1
ZNF354C
EGR1
NFIC
Number of TFBS

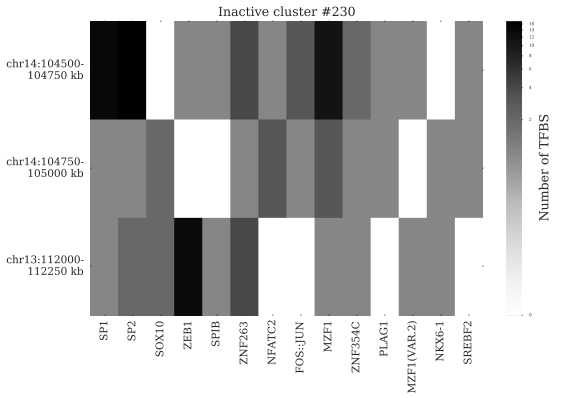


Figure S6: TFBS for top 15 inactive clusters given in Table S5 (with TFBS based on JASPAR 2016, threshold = 0.000001, CAGE). Only TFs that span at least 2 regions in the cluster are shown. Clusters with zero TFs spanning at least 2 regions are not shown.

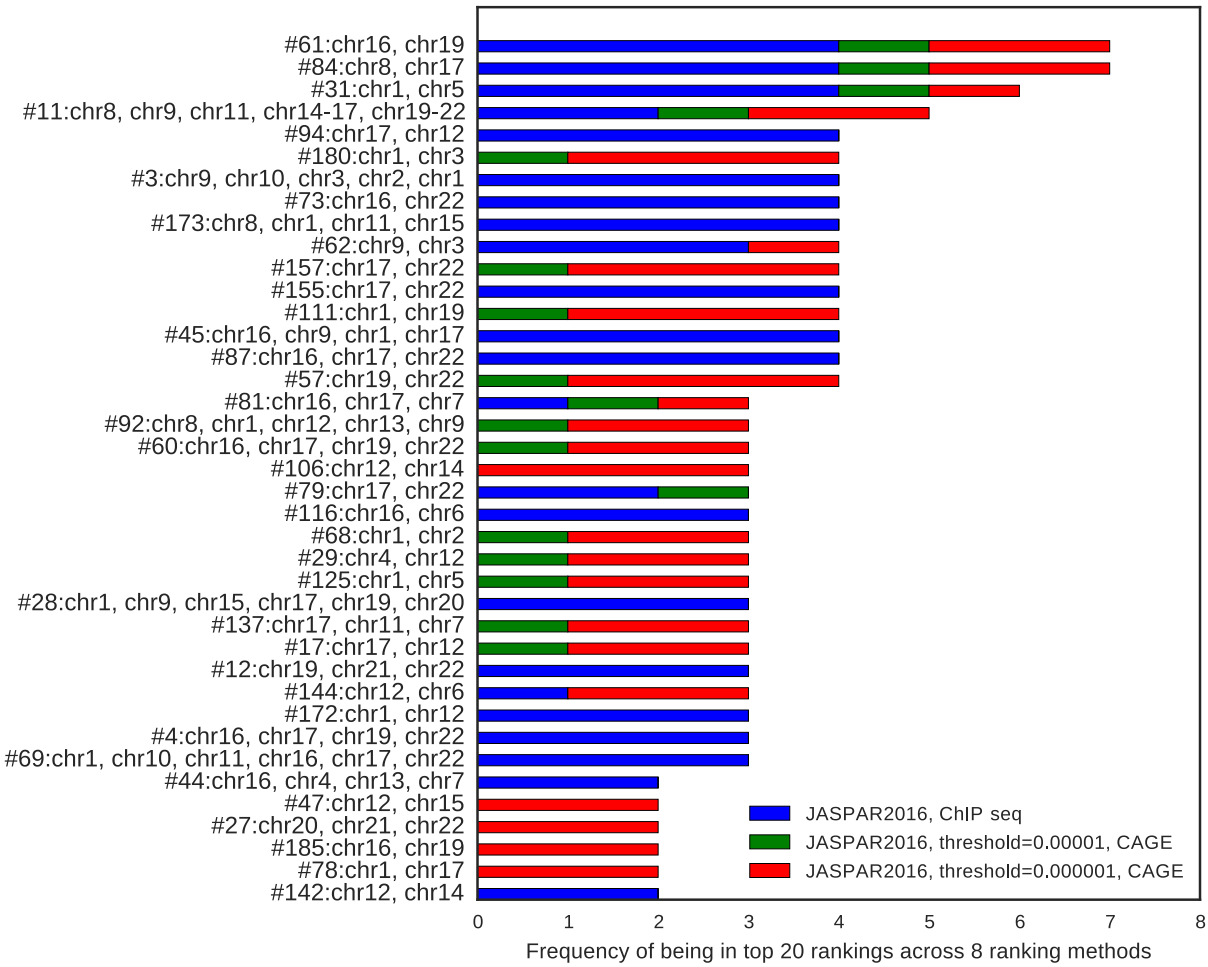


Figure S7: Top ranked clusters that appeared in the top 20 clusters across 8 different methods of evaluation. The eight rankings are grouped by the different choices for filtering the JASPAR 2016 database: The JASPAR2016 database was filtered by ChIP-seq (blue) and the rankings were obtained i) using active clusters as background and by generating random matrices from the observed counts for the permutation test, ii) using active clusters as background and by generating random matrices based on the Dirichlet distribution for the permutation test, iii) using all intermingling regions as background and by generating random matrices from the observed counts for the permutation test, iv) using the whole genome as background and by generating random matrices from the observed counts for the permutation test. The JASPAR2016 database TFBS were obtained with a threshold of 0.00001 and filtered by CAGE (green) and the rankings were obtained v) using active clusters as background and by generating random matrices from the observed counts for the permutation test. The JASPAR2016 database TFBS were obtained with a threshold of 0.000001 and filtered by CAGE (red) and the rankings were obtained vi) using active clusters as background and by generating random matrices from the observed counts for the permutation test, vii) using active clusters as background and by generating random matrices from the Dirichlet distribution for the permutation test, and viii) using intermingling regions as background and by generating random matrices from the observed counts for the permutation test.

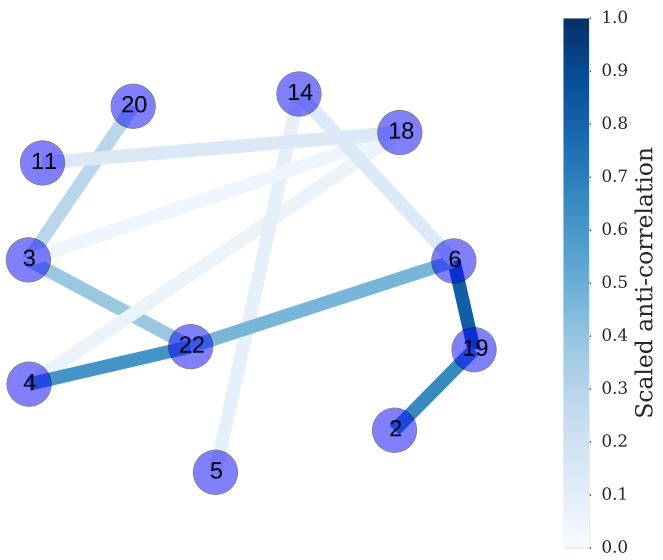


Figure S8: Determining chromosomes that do not intermingle for predicting negative controls. Edges link nodes, i.e., pairs of chromosomes, that showed no intermingling regions in the LAS analysis. The edge weights are given by the absolute value of the anti-correlation between the genomic features of the adjacent chromosomes at a whole chromosome level.

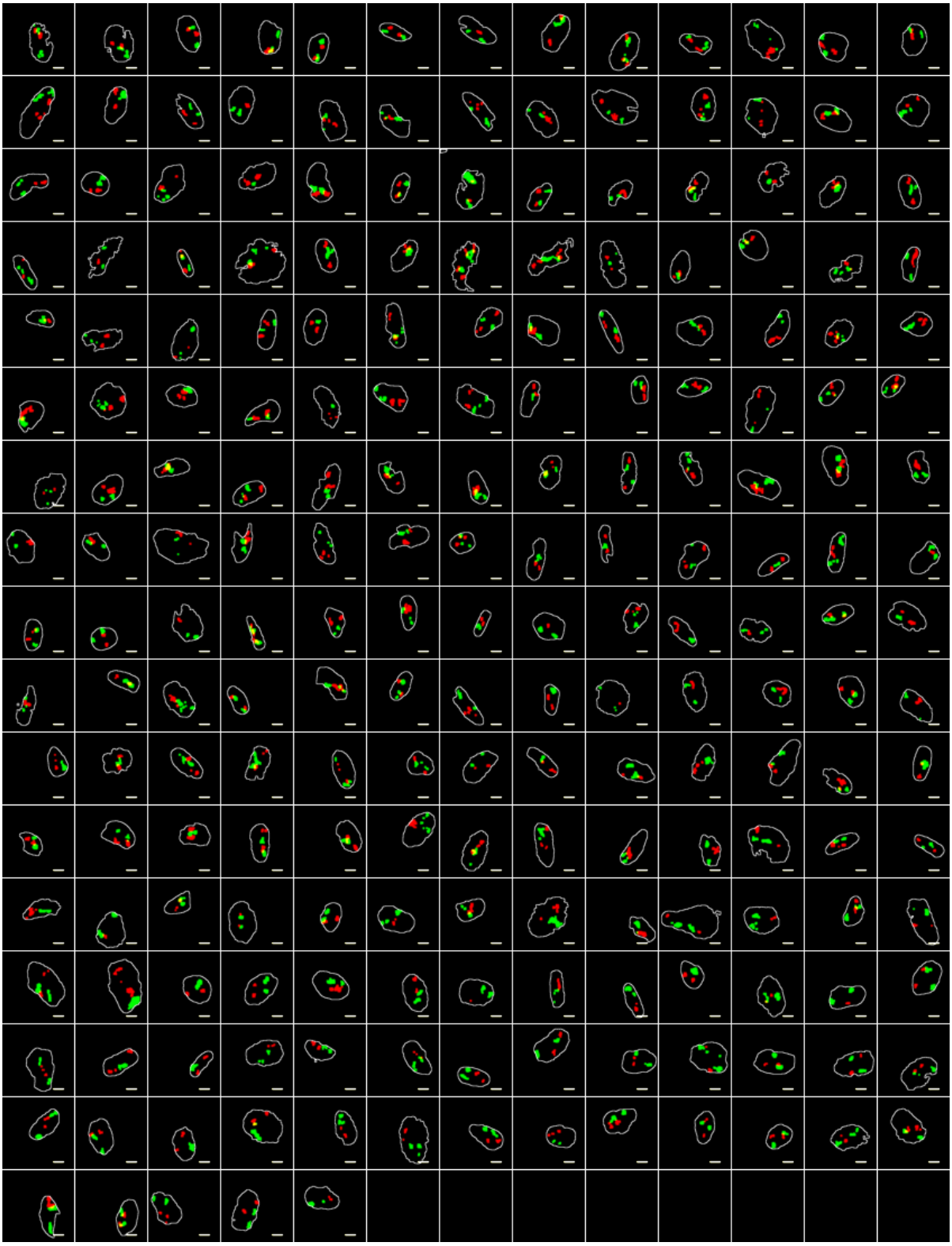


Figure S9: Experimental validation using FISH for predicted active cluster on chromosomes 12 and 17. Segmented images of nuclei from a population of cells with nuclear boundary shown in white, chromosome 17 in red, and chromosome 12 in green. The scale bar has a length of 5 µm.

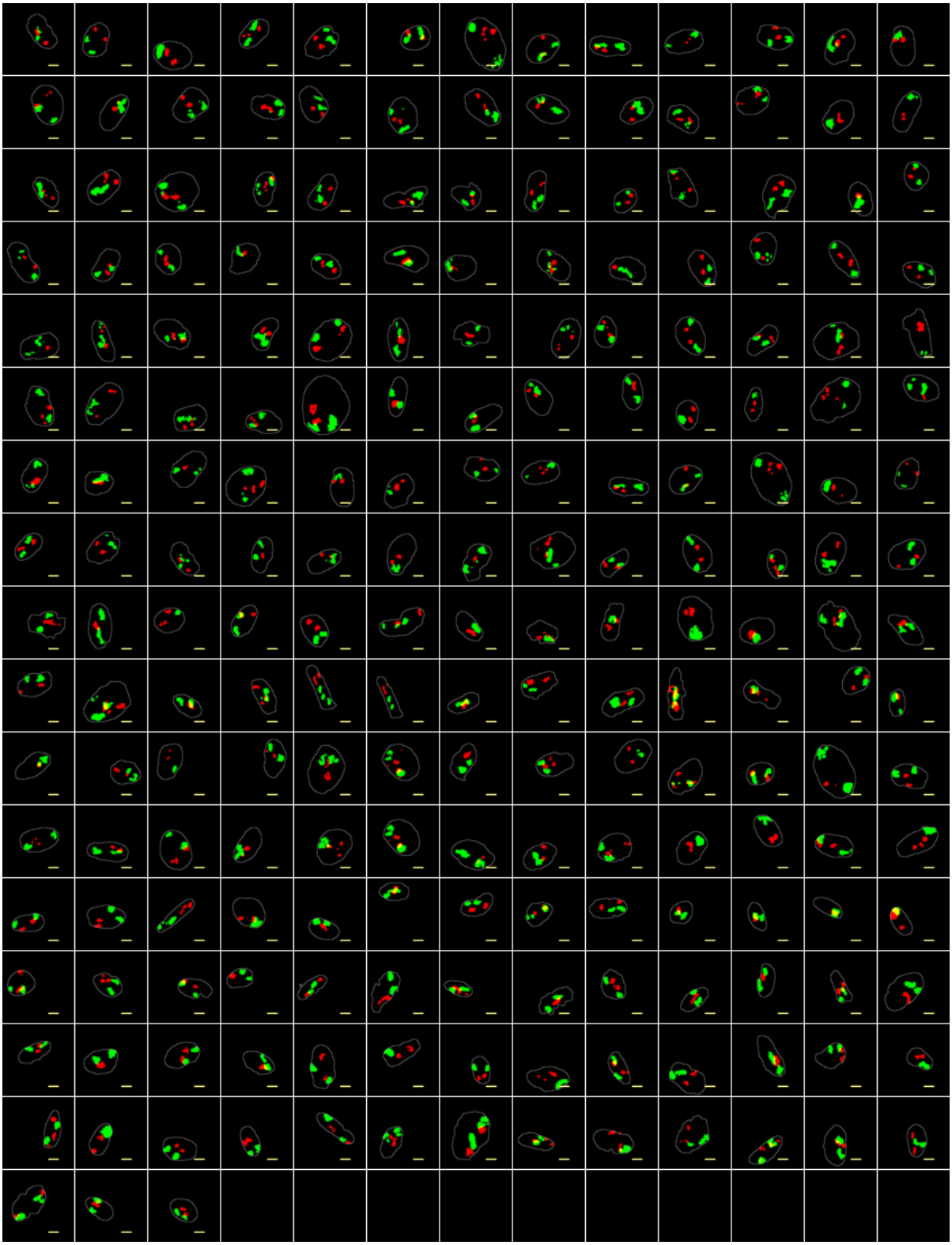


Figure S10: Experimental validation using FISH for predicted negative control, i.e. chromosomes 3 and 20 that are predicted to not intermingle. Segmented images of nuclei from a population of cells with nuclear boundary shown in white, chromosome 20 in red, and chromosome 3 in green. The scale bar has a length of 5 µm.

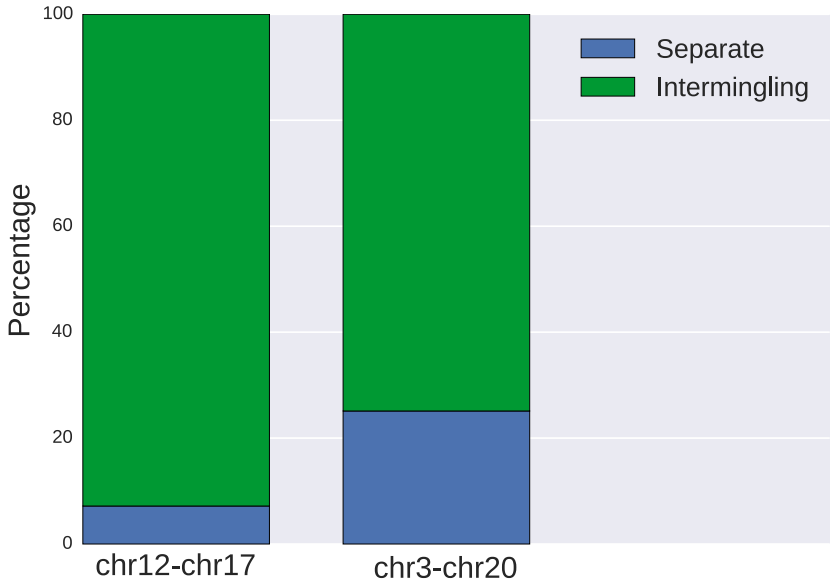


Figure S11: Breakdown of the percentage of nuclei that intermingle (intermingling degree > 0) versus don't intermingle for the chromosome pairs 12-17 and 3-20, as found by FISH experiments.

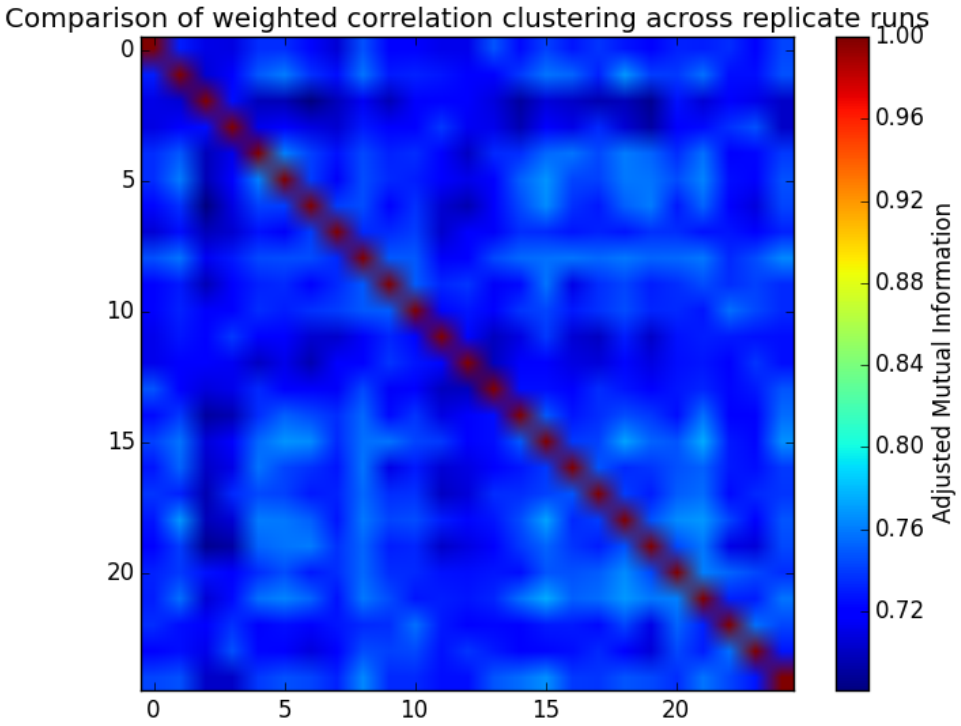


Figure S12: Adjusted mutual information between replicate clusterings from weighted correlation clustering.