

RESEARCH

Open Access



Dynamic epigenetic mode analysis using spatial temporal clustering

YangLan Gan^{1†}, Han Tao^{2†}, Guobing Zou³, Cairong Yan¹ and Jihong Guan^{2*}

From The 27th International Conference on Genome Informatics
Shanghai, China. 3-5 October 2016

Abstract

Background: Differentiation of human embryonic stem cells requires precise control of gene expression that depends on specific spatial and temporal epigenetic regulation. Recently available temporal epigenomic data derived from cellular differentiation processes provides an unprecedented opportunity for characterizing fundamental properties of epigenomic dynamics and revealing regulatory roles of epigenetic modifications.

Results: This paper presents a spatial temporal clustering approach, named STCluster, which exploits the temporal variation information of epigenomes to characterize dynamic epigenetic mode during cellular differentiation. This approach identifies significant spatial temporal patterns of epigenetic modifications along human embryonic stem cell differentiation and cluster regulatory sequences by their spatial temporal epigenetic patterns.

Conclusions: The results show that this approach is effective in capturing epigenetic modification patterns associated with specific cell types. In addition, STCluster allows straightforward identification of coherent epigenetic modes in multiple cell types, indicating the ability in the establishment of the most conserved epigenetic signatures during cellular differentiation process.

Keywords: Dynamic epigenetic mode, Spatial temporal clustering, Epigenetic modification, Cellular differentiation

Background

An epigenome consists of chemical modifications and variations to histones, DNA methylation and other proteins that package the genome [1, 2]. These epigenetic modifications crucially contribute to epigenetic maintenance of chromatin structures and gene expression regulation [3]. There are various interactions among these modifications, which act combinatorially to orchestrate gene expression in different cell types [4]. When heritable from one cell generation to the next, the epigenetic information can bring about lasting changes in gene expression [5, 6].

Embryonic development is a complex process that requires precise gene regulation to govern developmental decisions during cellular differentiation [7, 8]. However,

how gene expression is regulated and maintained along developmental transitions remains to be understood. Currently, it is well accepted that transcription factors binding to cis-regulatory sequences coordinately regulate gene expression in response to various environmental cues [9]. On the contrary, the regulatory functions of epigenetic modifications that accompany embryogenesis are largely unexplored. To fully investigate the mechanisms of epigenetic regulation in the cellular differentiation process, extensive research efforts provide genome-wide maps of epigenetic modifications at multiple developmental time points. Human embryonic stem cells were differentiated into a variety of precursor cell types [3, 10], including mesendoderm [11], trophoblast-like cells [12], mesenchymal stem cells [13] and neural progenitor cells [14]. Mouse embryonic stem cells were also differentiated into mesendoderm cells [15]. The availability of these temporal epigenomic data provides a unique opportunity for characterizing fundamental properties of epigenome

*Correspondence: jhguan@tongji.edu.cn

[†]Equal contributors

²Department of Computer Science and Technology, Tongji University, Shanghai, China

Full list of author information is available at the end of the article

dynamics and revealing regulatory roles of epigenetic modifications.

To establish combinatorial patterns of epigenetic modification, previous computational methods primarily utilize spatial information of epigenetic marks. For example, Chromasig was designed to study histone modification patterns using correlations of histone signals [16]. CoSBI also used the correlations within 5Kbp genomic segments to exhaustively searched for histone code [17]. ChromHMM applied a HMM model to annotate genomic sequences by co-occurrence of multiple epigenetic marks [18]. RFECS was developed based on random forests [19]. AWNFR explored epigenomic landscapes using the wavelet transforms [20]. Although these methods successfully identify the combinatorial epigenetic mode based on spatial epigenomic information, there is still an urgent need to exploit newly produced temporal information to study the dynamic patterns and functions of epigenetic modifications.

In this study, we developed a spatial temporal clustering approach that exploits the temporal variation information of epigenomes along the differentiation process, aiming to characterize dynamic properties of epigenetic modifications. This approach identifies significant spatial temporal patterns of epigenetic modifications during embryogenesis and cluster regulatory sequences by their spatial temporal epigenetic patterns. The results might shed a light on how epigenetic modifications evolve temporally and how the spatial temporal patterns of epigenetic modifications regulate gene expression during the process of cellular differentiation.

Methods

Datasets

In mammals, studying the epigenetic mechanisms of early embryonic development often requires access to embryonic cell types. In recent studies, to analyze early human developmental decisions, human embryonic stem cells (hESCs) were differentiated into trophoblast-like cells, mesendoderm, mesenchymal stem cells, and neural progenitor cells [3, 10]. The first three states represent developmental events that mirror critical developmental decisions in the embryo. Mesenchymal stem cells are fibroblastoid cells that are capable of expansion and multi-lineage differentiation to bone, cartilage, adipose, muscle, and connective tissues. In these cell types, genome-wide maps of main epigenetic marks have been generated using ChIP-seq [21]. In detail, the investigated epigenetic modifications were profiled, including H3K4me1/2/3, H3K36me3, H3K9me3, H3K27me3, H3K79me1, H2AK5ac, H2bK120ac, H2BK5ac, H3K18ac, H3K23ac, H3K27ac, H3K4ac, H3K9ac and H4K8ac. RNA expression profiles of these five cell types were also generated using Affymetrix GeneChip-arrays. Here, we downloaded these

datasets from the website of NIH Roadmap Epigenome Project (<http://www.epigenomebrowser.org/>) [22].

Methods

General scheme of the STCluster algorithm

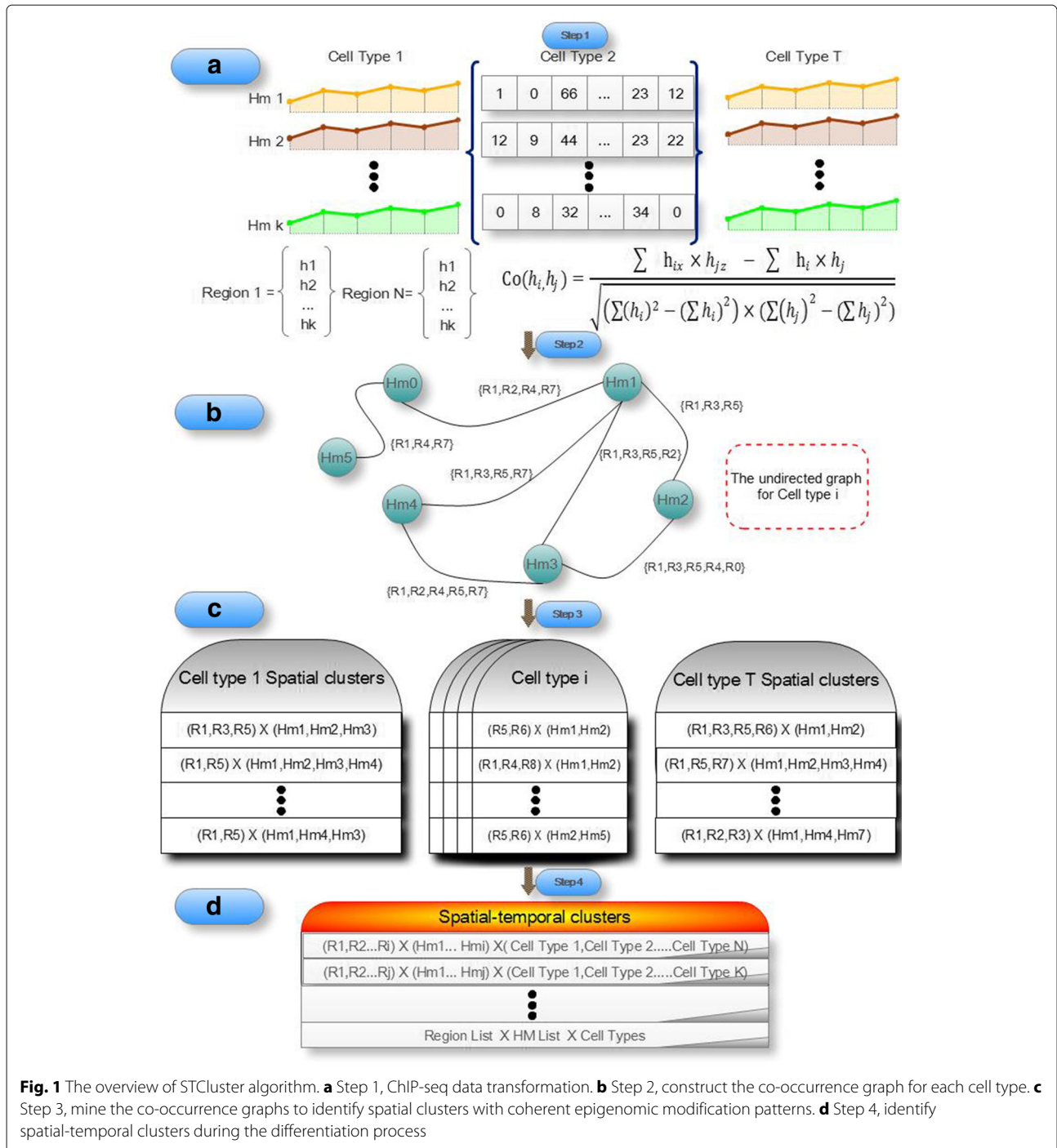
The STCluster algorithm analyzes genome-wide maps of epigenetic modifications to characterize dynamic epigenetic signatures during embryonic stem cell differentiation. There are four major steps in the STCluster algorithm: (i) ChIP-seq data transformation, (ii) construct the co-occurrence graph for each cell type, (iii) mine the co-occurrence graphs to identify spatial clusters of genomic segments with coherent epigenetic patterns, and (iv) mine the resulted spatial clusters of each cell type to identify spatial temporal clusters and discover conserved epigenetic signatures during the differentiation process. Figure 1 illustrates the scheme of the STCluster algorithm. In the following, we elaborate the process of epigenetic mode analysis step by step.

Step 1. Data transformation

The whole human genome were divided into non-overlapping 200bp bins. For each epigenetic modification map, we first computed the summary tag count of every bin. Then, in each cell type, raw sequence read counts of each epigenetic modification were normalized by the total number of reads followed by arcsine transformation [23], to remove noises resulting from spurious tag counts in the ChIP-seq experiments. Further, we divided the whole genome into 5Kbp genomic regions. In this way, for each cell type, the profiles of these epigenetic modifications are represented as a matrix R_i , where i is the index of the genomic regions ranging from 1 to N (assuming there are totally N genomic regions under consideration), as shown in Fig. 1a. In each region, the number of columns is denoted as B and the number of epigenetic modifications is denoted as K . The column vectors correspond to combinatorial epigenetic modification tag counts within individual genomic bins and the row vectors correspond to the contiguous genomic landscape of individual epigenetic modifications.

Step 2. Construct co-occurrence graph for each cell type

In this step, we computed the correlation coefficients of epigenetic modification pairs in each region, and then we constructed the corresponding co-occurrence graph for each cell type. Given the processed and organized epigenetic modification data of each cell type, correlation coefficients of any two histone modifications at every region were calculated to obtain a coefficient matrix. If the coefficients are higher than a given threshold, the two epigenetic modifications are regarded as coherent in this region. Subsequently, this region was added to the corresponding region set. Based on the coefficient matrix,



we further constructed the co-occurrence graph, which is modeled as an undirected graph $G = (V, E)$, where V is the set of all histone modifications. For any two epigenetic modification types h_i and h_j ($i \neq j$), if they are correlated at any region, there exists an edge $e \in E$ between vertices h_i and h_j . In addition, each edge in the co-occurrence graph is associated with the region set. Figure 1b shows an example of co-occurrence graph. Here, we set the correlation

coefficient threshold as 0.9 to achieve a high quality of spatial clusters.

Step 3. Mine spatial clusters from co-occurrence graph

The co-occurrence graph represents in a compact way all the correlated epigenetic modifications in different regions. It can be used to mine potential spatial clusters corresponding to each developmental stage, and filter

out most of the unrelated data. The STCluster algorithm applies a depth first search (DFS) strategy on the co-occurrence graph to mine all the spatial clusters. A typical spatial cluster represents a group of genomic regions that share spatial epigenetic patterns. To gain the significant epigenetic states, we set the minimum number of histone modifications as 5 and the minimum percent of regions as 0.1%. For each cell type, we identified a set of spatial clusters, as shown in Fig. 1c.

Step 4. Identify spatial-temporal clusters from spatial clusters

On obtaining the maximal spatial cluster set for all cell types, we utilized them to mine the maximal spatial-temporal clusters. This is accomplished by enumerating the subsets of the time points (Fig. 1d), using a process similar to the spatial cluster clique mining. The regions in each spatial temporal clusters exhibit similar changes of epigenetic modifications during the cellular differentiation process. Spatial-temporal clusters indicate specific conserved chromatin signatures that are shared by multiple time points along the embryonic stem cell differentiation.

Results and Discussion

Identifying combinatorial epigenetic states during differentiation

To investigate combinatorial epigenetic states during the differentiation of embryonic stem cells, we applied STCluster to the genome-wide epigenetic modification maps of five cell types, including H1, Mesendoderm, Trophoblast-like cells, Mesenchymal stem cells and Neuronal progenitor cells. STCluster first grouped genomic regions based on spatial patterns of epigenetic modifications to identify spatial clusters. For each cell type, we set the minimum number of histone modifications as

5 and the minimum percent of regions as 0.1%, which allow us to capture patterns that involve at least five epigenetic modifications and re-occur across at least 0.1% of the human genome. With this parameter setting, we respectively identified 3344, 667, 4726, 1422, 1984 spatial clusters in H1, Mesendoderm, Trophoblast-like cells, Mesenchymal stem cells and Neuronal progenitor cells.

Next, we evaluated the occurrence frequencies of all investigated epigenetic modifications in the identified spatial clusters. Specifically, the occurrence frequency of an epigenetic modification is computed as the ratio of the occurrence in these spatial clusters and the total number of 5Kbp non-overlapping regions in the genome. Their occurrence frequencies are depicted in Fig. 2. We found that epigenetic modification H3K18ac has a high frequency in all cell types, which indicates that these regions share the variation pattern of this epigenetic modification. Epigenetic modifications (H3K4me1/3) seldom occur in the spatial clusters of H1, whereas epigenetic modification (H3K4me2) frequently occur in the spatial clusters of cell types except H1. In the spatial clusters of Neuronal progenitor cells, most epigenetic modifications have median occurrence frequencies.

Expanding this research, we studied co-occurred epigenetic modifications at each developmental stage of the differentiation process. For each cell type, we ranked the identified spatial clusters according to the number of regions co-occupied by epigenetic modification set. We discovered groups of epigenetic modifications that frequently co-occur in each cell type. Table 1 summarizes the top 10 frequently co-occurred epigenetic modifications. Overall, the clustering results show that different cell types exhibit diverse cell type specific patterns of epigenetic modifications. The overlaps of epigenetic modifications among these cell types are small. However,

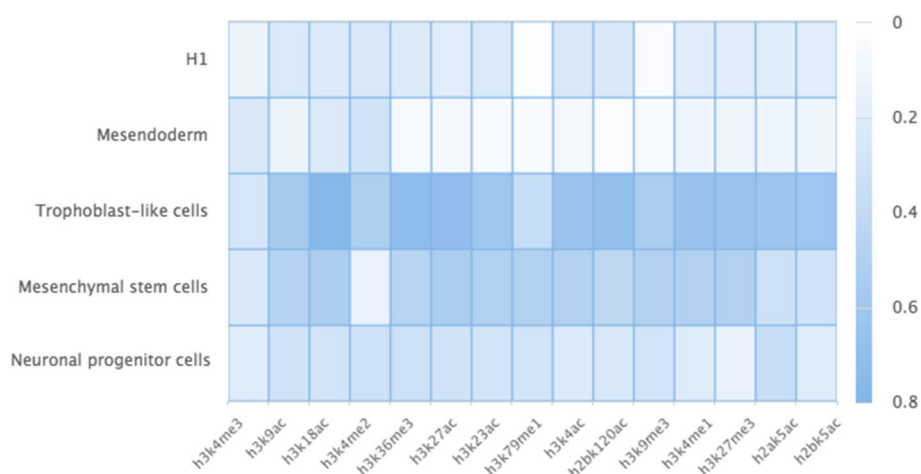


Fig. 2 The occurrence frequencies of all investigated epigenetic modifications in the identified spatial clusters

Table 1 The top 10 co-occurred epigenetic modifications in different cell types during the differentiation process

	H1	Mesendoderm	Trophoblast-like cells	Mesenchymal stem cells	Neuronal progenitor cells
1	H3K4me2,H3K4me1, H2aK5ac,H3K27ac, H2bK5ac,H3K23ac, H3K9me3,H3K79me1, H3K4ac,H3K9ac, H2bK120ac	H3K4me2,H3K4me1, H3K27ac,H3K23ac, H3K9ac,H3K18ac	H2aK5ac,H2bK5ac, H3K27me3,H3K23ac, H3K9me3,H3K79me1, H3K4ac,H3K9ac, H2bK120ac,H3K18ac	H3K4me2,H3K36me3, H3K27ac,H2bK5ac, H3K23ac,H3K9me3, H3K4ac,H3K9ac, H2bK120ac	H3K4me1,H3K36me3, H2aK5ac,H3K27ac, H2bK5ac,H3K27me3, H3K23ac,H3K79me1, H3K9ac,H2bK120ac
2	H2aK5ac,H2bK5ac, H3K23ac,H3K79me1, H3K4ac,H2bK120ac	H3K4me2,H3K4me1, H3K27ac,H3K23ac, H3K9ac,H2bK120ac	H3K4me1,H2bK5ac, H3K27me3,H3K23ac, H3K9me3,H3K79me1, H3K4ac,H3K9ac, H2bK120ac,H3K18ac	H3K4me2,H3K36me3, H2bK5ac,H3K23ac, H3K4ac,H2bK120ac	H3K4me1,H3K36me3, H3K27ac,H2bK5ac, H3K27me3,H3K9me3, H3K79me1,H3K9ac, H2bK120ac
3	H2bK5ac,H3K27me3, H3K23ac,H3K79me1, H3K4ac,H3K9ac	H2aK5ac,H2bK5ac, H3K79me1,H3K4ac, H2bK120ac,H3K18ac	H3K27ac,H2bK5ac, H3K27me3,H3K23ac, H3K9me3,H3K18ac	H3K36me3,H2aK5ac, H2bK5ac,H3K23ac, H3K9me3,H3K4ac H3K9ac	H3K4me2,H2aK5ac, H3K23ac,H3K9me3, H3K4ac,H2bK120ac
4	H3K4me2,H3K4me1, H2aK5ac,H3K27ac, H2bK5ac,H3K27me3, H3K23ac,H3K79me1, H3K4ac,H3K9ac,H3K18ac	H3K4me1,H3K36me3, H3K27ac,H3K23ac, H3K79me1,H3K4ac, H3K9ac,H2bK120ac, H3K18ac	H3K27ac,H2bK5ac, H3K27me3,H3K23ac, H3K9me3,H2bK120ac	H3K36me3,H2aK5ac, H2bK5ac,H3K23ac, H3K4ac,H3K9ac	H3K4me2,H2aK5ac, H3K23ac,H3K9me3, H3K9ac,H2bK120ac
5	H2bK5ac,H3K27me3, H3K23ac,H3K79me1, H3K4ac,H3K18ac	H3K27ac,H3K27me3, H3K9me3,H3K4ac, H2bK120ac,H3K18ac	H3K27ac,H2bK5ac, H3K27me3,H3K23ac, H3K9me3,H3K9ac	H3K4me2,H3K36me3, H2bK5ac,H3K23ac, H3K9me3,H3K9ac	H3K27ac,H2bK5ac, H3K27me3,H3K9me3, H3K9ac,H2bK120ac
6	H2aK5ac,H2bK5ac, H3K9me3,H3K79me1, H3K4ac,H2bK120ac	H3K4me2,H3K27ac, H3K9me3,H3K4ac, H2bK120ac,H3K18ac	H3K27ac,H2bK5ac, H3K27me3,H3K23ac, H3K9me3,H3K79me1	H3K36me3,H2aK5ac, H2bK5ac,H3K9me3, H3K4ac,H2bK120ac	H3K27ac,H2bK5ac, H3K27me3,H3K9me3, H3K4ac,H2bK120ac
7	H2aK5ac,H2bK5ac, H3K23ac,H3K79me1, H3K4ac,H3K9ac	H3K4me2,H3K4me1, H3K27ac,H3K27me3, H3K9me3,H3K18ac	H3K27ac,H2bK5ac, H3K27me3,H3K23ac, H3K9me3,H3K4ac	H3K4me2,H3K36me3, H2bK5ac,H3K23ac, H3K9me3,H3K4ac	H3K4me2,H3K4me1, H3K23ac,H3K9me3, H3K9ac,H2bK120ac
8	H3K4me2,H2aK5ac, H3K27ac,H2bK5ac, H3K27me3,H3K23ac, H3K9me3,H3K79me1, H3K4ac,H3K9ac,H3K18ac	H3K4me2,H3K4me1, H3K27ac,H3K27me3, H3K4ac,H3K18ac	H3K4me2,H3K27ac, H3K27me3,H3K23ac, H3K9me3,H3K9ac	H3K4me2,H3K36me3, H2bK5ac,H3K23ac, H3K4ac,H3K9ac, H2bK120ac	H3K4me2,H3K4me1, H3K23ac,H3K9me3, H3K4ac,H2bK120ac
9	H3K4me2,H3K4me1, H2aK5ac,H3K27ac, H2bK5ac,H3K23ac, H3K9me3,H3K79me1, H3K4ac,H3K9ac H3K18ac	H3K4me2,H3K4me1, H3K27ac,H3K27me3, H3K4ac,H2bK120ac	H3K4me1,H3K27ac, H3K27me3,H3K23ac, H3K9me3,H3K79me1	H3K4me2,H3K4me1, H3K9me3,H3K4ac, H3K9ac,H2bK120ac	H3K4me2,H3K4me1, H3K36me3,H2aK5ac, H3K27me3,H3K23ac, H3K79me1,H3K9ac, H2bK120ac
10	H2aK5ac,H2bK5ac, H3K9me3,H3K79me1, H3K4ac,H3K9ac	H3K4me2,H3K27ac, H3K27me3,H3K4ac, H2bK120ac,H3K18ac	H3K4me1,H3K27ac, H3K27me3,H3K23ac, H3K9me3,H3K79me1	H3K4me2,H3K36me3, H2bK5ac,H3K23ac, H3K9me3,H2bK120ac	H3K4me2,H3K4me1, H2aK5ac,H3K27ac, H3K9me3,H3K9ac, H2bK120ac

We found that epigenetic marks <H3K4me1, H3K4me2> and <H3K4me1, H3K4ac> frequently co-occur in the spatial clusters of different cell types. Part of these cell types, such as H1 cell line and Mesendoderm, Mesenchymal stem cells and Neuronal progenitor cells, share more epigenetic patterns than other groups of cell types.

Identifying conserved epigenetic states during the differentiation of ES cells

There are large differences among the investigated epigenetic modifications regarding their temporal variations. To identify conserved epigenetic states and explore the temporal patterns of these epigenetic modifications, we

applied STCluster to further group genomic regions based on the spatial clusters. The identified spatial temporal clusters are represented as triples (<genomic regions>, <epigenetic modifications>, <cell types>). Each cluster lists the genomic regions with the co-occupied epigenetic modifications, which exhibit little variation at different cell types during the differentiation process. Taking a typical spatial temporal cluster as an example, Fig. 3 displays the profiles of co-occurred epigenetic modifications in different regions during the differentiation process. In this cluster, eight epigenetic modifications (H3K4me2, H3K23ac, H3K27ac, H2BK120ac, H3K27me3, H3K79me1, H3K4ac and H3K9ac) co-occur the clustered

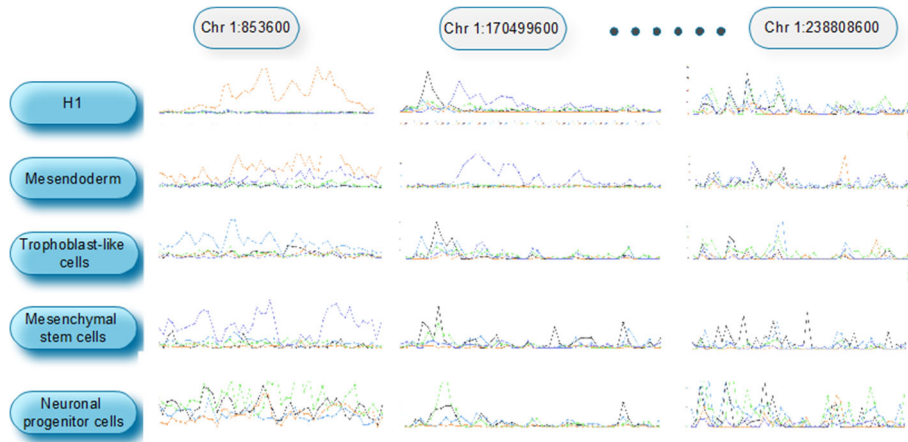


Fig. 3 Profiles of co-occurred epigenetic modifications display an overall conserved pattern at five different stages during the differentiation process

regions in five different cell types. For each region, these epigenetic marks display conserved modification patterns at all these five stages.

The detailed information of all identified spatial temporal clusters are listed in Additional file 1. The results indicate that there exist conserved epigenetic states during the differentiation process. We observed a high co-occurrence and stable patterns of H3K4me2 with

H3K23ac, H3K18ac, H3K27ac and H3K9ac at five different stages along the differentiation process. Our observation is consistent with the previous finding that H3K4me2 is one of the backbone epigenetic modifications along with H3K27ac and H3K9ac [18, 24]. On the contrary, some epigenetic modification patterns are only coherent in certain cell types. For example, the variation pattern of epigenetic modifications <H3K27me3, H3K9me3,

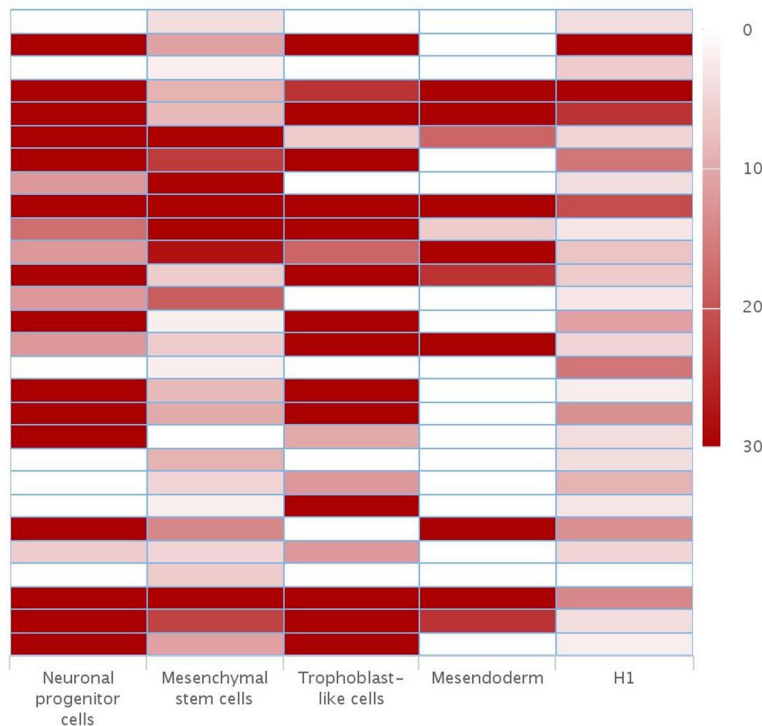


Fig. 4 The RNA expression levels of the identified spatial temporal clusters

H3K79me1, H3K4ac, H3K9ac> is shared in four cell types except Mesendoderm. <H3K4me2, H3K23ac, H3K27ac, H3K27me3, H3K79me1, H3K18ac> are only shared in H1, Mesendoderm, and Trophoblast-like cells except Mesenchymal stem cells and Neuronal progenitor cells. Notably, the identified spatial temporal clusters reveal more details of the differentiation process.

Analyzing the regulatory roles of epigenetic modifications during differentiation

As epigenetic marks were thought to be predictive of gene expression levels in a context-independent manner [25], we further analyzed the RNA expression levels of the identified spatial temporal clusters to see if this theory holds during embryonic stem cell differentiation. Specifically, we extracted the RNA expression data of the genomic regions included in the spatial temporal clusters, and compared the expression level in the corresponding cell types. Consistent with previous study, some epigenetic marks, such as H3K27ac and H3K36me3, are correlated with RNA expression level of different genomic regions at different developmental stages. However, several epigenetic modifications show cell type specific regulation on mRNA expression. As shown in Fig. 4, the variations of temporal epigenetic modifications are not correlated with gene expression changes in several spatial-temporal clusters, which are primarily located in promoter regions. These results imply that epigenetic patterns mediate gene regulation during cell differentiation in a complex way, rather than in a linear manner.

Conclusions

Identifying epigenomic dynamics is important to understand mechanisms for gene regulation. Our knowledge about the temporal patterns of epigenetic modifications and the consequence of them are still limited. There is a urgent need to develop new computational approach that exploits the complex epigenomic landscapes and discovers significant signatures out of them. In this study, we developed a spatial temporal clustering algorithm to explore the epigenomic landscapes of five cell types during embryonic stem cell differentiation. Using this approach, we identified spatial temporal patterns of epigenetic modifications in early embryogenesis. Different from previous computational methods, our approach is designed to investigate the dynamic epigenetic landscapes as well as the combinational epigenetic modes. The experimental results demonstrate that the proposed STCluster algorithm could successfully capture dynamic epigenetic modification patterns associated with specific cell types. In addition, STCluster allows straightforward identification of epigenetic conservation at multiple developmental stages during cell differentiation process.

Additional file

Additional file 1: The detailed information of identified spatial temporal clusters. Details for the spatial temporal clusters, represented as triples (<genomic regions>, <epigenetic modifications>, <cell types>). Each cluster lists the genomic regions with the co-occupied epigenetic modifications, which exhibit little variation at different cell types during the differentiation process. (TXT 3962 kb)

Acknowledgements

Authors are grateful to NIH Roadmap Epigenome Project for providing the genome-wide data to carry out this work. We thank anonymous reviewers for their useful comments on the manuscript.

Declarations

This article has been published as part of *BMC Bioinformatics* Volume 17 Supplement 17, 2016: Proceedings of the 27th International Conference on Genome Informatics: bioinformatics. The full contents of the supplement are available online at <http://bmcbioinformatics.biomedcentral.com/articles/supplements/volume-17-supplement-17>.

Funding

This work and the publication costs were supported in part by the National Natural Science Foundation of China (61272380, 61300100, 61303096), Shanghai Natural Science Foundation (13ZR1451000, 13ZR1454600, 15ZR1400900), Chen Guang project supported by Shanghai Municipal Education Commission and Shanghai Education Development Foundation, and the Fundamental Research Funds for the Central Universities (16D111208).

Availability of data and materials

Datasets used in this study can be accessed at the following URL <http://www.tongjidmb.com/human/index.html>.

Authors' contributions

YLG and JHG have designed the clustering algorithm. YLG and HT have implemented and tested the method. GBZ and CRY have coordinated data preprocessing. All authors have read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Consent for publication

Not applicable.

Ethics approval and consent to participate

No ethics approval was required for this work, as all results reported have already been published elsewhere.

Author details

¹School of Computer Science and Technology, Donghua University, Shanghai, China. ²Department of Computer Science and Technology, Tongji University, Shanghai, China. ³School of Computer Engineering and Science, Shanghai University, Shanghai, China.

Published: 23 December 2016

References

- Bernstein BE, Meissner A, Lander ES. The mammalian epigenome. *Cell*. 2007;128(4):669–81.
- Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K. High-resolution profiling of histone methylations in the human genome. *Cell*. 2007;129(4):823–37.
- Gifford CA, Ziller MJ, Gu H, Trapnell C, Donaghey J, Tsankov A, Shalek AK, Kelley DR, Shishkin AA, Issner R, et al. Transcriptional and epigenetic dynamics during specification of human embryonic stem cells. *Cell*. 2013;153(5):1149–1163.
- Chen T, Dent SY. Chromatin modifiers and remodellers: regulators of cellular differentiation. *Nat Rev Genet*. 2014;15(2):93–106.

5. Kouzarides T. Chromatin modifications and their function. *Cell*. 2007;128(4):693–705.
6. Zhou WW, Goren A, Bernstein BE. Charting histone modifications and the functional organization of mammalian genomes. *Nat Rev Genet*. 2011;12(1):7–18.
7. Vastenhouw NL, Schier AF. Bivalent histone modifications in early embryogenesis. *Curr Opin Cell Biol*. 2012;24(3):374–86.
8. Young RA. Control of the embryonic stem cell state. *Cell*. 2011;144(6):940–54.
9. Tsankov AM, Gu H, Akopian V, Ziller MJ, Donaghey J, Amit I, Gnirke A, Meissner A. Transcription factor binding dynamics during human es cell differentiation. *Nature*. 2015;518(7539):344–9.
10. Xie W, Schultz MD, Lister R, Hou Z, Rajagopal N, Ray P, Whitaker JW, Tian S, Hawkins RD, Leung D, et al. Epigenomic analysis of multilineage differentiation of human embryonic stem cells. *Cell*. 2013;153(5):1134–1148.
11. Yu P, Pan G, Yu J, Thomson JA. Fgf2 sustains nanog and switches the outcome of bmp4-induced human embryonic stem cell differentiation. *Cell Stem Cell*. 2011;8(3):326–34.
12. Xu RH, Chen X, Li DS, Li R, Addicks GC, Glennon C, Zwaka TP, Thomson JA. Bmp4 initiates human embryonic stem cell differentiation to trophoblast. *Nat Biotechnol*. 2002;20(12):1261–1264.
13. Vodyanik MA, Yu J, Zhang X, Tian S, Stewart R, Thomson JA, Slukvin II. A mesoderm-derived precursor for mesenchymal stem and endothelial cells. *Cell Stem Cell*. 2010;7(6):718–29.
14. Chen G, Gulbranson DR, Hou Z, Bolin JM, Ruotti V, Probasco MD, Smuga-Otto K, Howden SE, Diol NR, Propson NE, et al. Chemically defined conditions for human ipsc derivation and culture. *Nat Methods*. 2011;8(5):424–9.
15. Yu P, Xiao S, Xin X, Song CX, Huang W, McDee D, Tanaka T, Wang T, He C, Zhong S. Spatiotemporal clustering of the epigenome reveals rules of dynamic gene regulation. *Genome Res*. 2013;23(2):352–64.
16. Hon G, Ren B, Wang W. Chromasig: a probabilistic approach to finding common chromatin signatures in the human genome. *PLoS Comput Biol*. 2008;4(10):1000201.
17. Ucar D, Hu Q, Tan K. Combinatorial chromatin modification patterns in the human genome revealed by subspace clustering. *Nucleic Acids Res*. 2011;39(10):4063–75.
18. Ernst J, Kheradpour P, Mikkelsen TS, Shoresh N, Ward LD, Epstein CB, Zhang X, Wang L, Issner R, Coyne M, et al. Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature*. 2011;473(7345):43–9.
19. Rajagopal N, Xie W, Li Y, Wagner U, Wang W, Stamatoyannopoulos J, Ernst J, Kellis M, Ren B. Rfecs: a random-forest based algorithm for enhancer identification from chromatin state. *PLoS Comput Biol*. 2013;9(3):1002968.
20. Nguyen N, Vo A, Won KJ. A wavelet-based method to exploit epigenomic language in the regulatory region. *Bioinformatics*. 2014;30(7):908–14.
21. Furey TS. Chip-seq and beyond: new and improved methodologies to detect and characterize protein–dna interactions. *Nat Rev Genet*. 2012;13(12):840–52.
22. Bernstein BE, Stamatoyannopoulos JA, Costello JF, Ren B, Milosavljevic A, Meissner A, Kellis M, Marra MA, Beaudet AL, Ecker JR, et al. The nih roadmap epigenomics mapping consortium. *Nat Biotechnol*. 2010;28(10):1045–1048.
23. Pinello L, Xu J, Orkin SH, Yuan GC. Analysis of chromatin-state plasticity identifies cell-type-specific regulators of h3k27me3 patterns. *Proc Natl Acad Sci*. 2014;111(3):344–53.
24. Wang Z, Zang C, Rosenfeld JA, Schones DE, Barski A, Cuddapah S, Cui K, Roh TY, Peng W, Zhang MQ, et al. Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat Genet*. 2008;40(7):897–903.
25. Karlič R, Chung HR, Lasserre J, Vlahoviček K, Vingron M. Histone modification levels are predictive for gene expression. *Proc Natl Acad Sci*. 2010;107(7):2926–931.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

