

Chromosomal Architecture Changes Upon Cell Differentiation

Maxim Imakaev¹, Geoffrey Fudenberg², Leonid Mirny^{3,1,2}

¹*Massachusetts Institute in Technology, 77 Mass. ave, Boston, MA, USA, imakaev@mit.edu*

²*Harvard University, Program in Biophysics, Boston, MA, USA, geoff.fudenberg@gmail.com*

³*Institute for Medical Engineering and Science, MIT, Boston, MA, USA, leonid@mit.edu*

The recently developed Hi-C method provides a comprehensive whole-genome picture of physical contacts between distal loci¹. Analysis of these data has begun to reveal determinants of 3D genomic organization. However, the similarities and differences in chromosomal organization between cell-types remain unexplored. To analyze chromosomal architecture between cell types, it is crucial to have a consistent way of analyzing Hi-C data and removal of experimental biases. To this end we developed a comprehensive method of Iterative Correction and Eigenvector decomposition (ICE)². ICE maps Hi-C reads to the genome, filters mapped reads and obtains a Hi-C map of relative contact probabilities free of experimental biases. It then decomposes the maps into a set of genomic tracks characterizing high-order chromatin organization.

Using ICE, we analyze Hi-C data³ from human embryonic stem cells, and IMR90 lung fibroblast cells. We focus our analysis on the compartment profile, which has been shown to partition the genome into transcribed gene-rich regions, enriched in active chromatin marks (“active” regions), and “inactive” gene-poor regions. First, we show that in ES cells have a gradual transition between “active” and “inactive” chromatin interaction preferences, as demonstrated by a broad uni-modal distribution of values of the compartment profile. In contrast, differentiated IMR90 cells show one inactive chromatin state and a range of states at the active end. Second, we find that chromatin interactions in embryonic cells are best described by GC content of a genomic region. Conversely, for the differentiated cell line IMR90, transcription data (CAGE) is a much better predictor of chromatin interaction preferences than sequence-derived features. Lastly, we analyze changes in chromatin

interactions upon differentiation, and find that regions which belonged to an active compartment in ES cells often switch to inactive compartment in IMR90, while the opposite rarely happens.

Taken together, our results show that genome-wide chromatin interactions change upon differentiation of ES cells into IMR90, and suggest that sequence-dependent chromatin interactions in embryonic stem cells get overridden in a cell-type-specific manner. We show that upon differentiation regions change from an active to an inactive compartment, suggesting that change in chromatin interactions reflects cell-type-specific silencing of genomic regions.

1. E. Lieberman-Aiden et al. (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome, *Science*, **326(5950)**:289-293.
2. M. Imakaev et al. (2012) Iterative correction of Hi-C data reveals hallmarks of chromosome organization, *Nature Methods*, **9(10)**: 999-1003.
3. J.R. Dixon et al. (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions, *Nature*, **485(7398)**, 376-380