Spatial configuration of the alpha-globin gene domain in three cell types of G.gallus

Aleksandra A. Galitsyna
Moscow State University, Faculty of Bioengineering and Bioinformatics, Leninskie Gory 1-73, Moscow 119991, Russia
agalitzina@gmail.com

Ekaterina E. Khrameeva
Skolkovo Institute of Science and Technology, Novaya st. 100, Skolkovo 143025, Russia,
Institute for Information Transmission Problems, RAS, Bolshoy Karetny per. 19/1, Moscow 127051, Russia

Sergey V. Ulyanov
Institute of Gene Biology, RAS, Vavilova st. 34/5, Moscow 119334, Russia
sergey.v.ulyanov@gmail.com

Chromatin spatial configuration and its role in the regulation of gene expression can be analyzed with chromosome conformation capture technologies at high resolution and throughput [1]. The alpha-globin gene domain is one of the most popular models to study the mechanisms of long-range and tissue-specific gene regulation [2]. 3C analysis of the alpha-globin gene domain in chicken lymphoid and erythroid cells has revealed the dependence of its complex spatial configuration on the alpha-globin gene expression [3]. However, the role of its interactions with the neighboring region remains unclear. Chromosome Conformation Capture Carbon Copy (5C) technology is one of the possible ways to study this region, but it has never been used for the chicken chromatin analysis [4].

Here, we applied 5C technology to analyze the alpha-globin gene region in chicken cells.

We used the results of 5C sequencing for three cell types of G.gallus (lymphoid cells DT40, proerythroblasts HD3 and induced erythroblasts HD3). We also used the results of RNA-Seq and ChIP-Seq for chromatin architecture protein CTCF for these cells. These data were provided by Laboratory of Structural and Functional Organization of Chromosomes, Institute of Gene Biology of the Russian Academy of Sciences.

5C data were mapped, filtered, binned, balanced and smoothed, and contact probability heatmaps were produced. We tested two approaches for mapping (standard mapping tools and simple search algorithm), different binning windows sizes, several balancing methods (including BAC-normalization and iterative correction [5]), and different smoothing methods.

We used Armatus algorithm [6] to test the heatmaps for the presence of topological domains (TADs) [7], and have shown that chicken chromatin is organized in TADs.

In addition, we have identified two large compartments of chromatin in our data with PCA analysis [8]. The boundaries of compartments are concordant with TADs boundaries. Analysis of the RNA-Seq and ChIP-Seq for CTCF data has shown that one of the compartments (A) includes regions with high expression and the most CTCF peaks. Another compartment (B) includes region with inactive chromatin. Compartments and their boundaries are preserved between cell types and upon induction of proerythroblasts differentiation. The alpha-globin gene domain is located in the compartment of active chromatin.

The alpha-globin expression is absent in lymphoid cells, low in proerythroblasts and high in induced erythroblasts. To detect chromatin decompaction in induced erythroid cells, we have applied the test for changes in intra-domain interaction frequency [9]. Also, Jaccard coefficient has been used to compare heatmaps of the chicken alpha-globin gene region between cell types. The results support the hypothesis of chromatin decompaction upon gene activation for alpha-globin region.

This work was supported by RFBR grant number 13-04-40089. This is joint work with Mikhail S. Gelfand, Institute for Information Transmission Problems, RAS.

1. E. Wit, W. Laat (2012) A decade of 3C technologies: insights into nuclear organization,

Genes&Development, **26**:11-24

2. D. Bau et al. (2011) The three-dimensional folding of the alpha-globin gene domain reveals formation of chromatin globules, *Nature Structural&Molecular Biology,* **18(1):** 107-114.

3. A. Gavrilov A. (2008) Spatial configuration of the chicken alpha-globin gene domain: immature and active chromatin hubs, *Nucleic Acid Research,* **14**: 4629-4640

4. J. Dostie et al. (2006) Chromosome Conformation Capture Carbon Copy (5C): A massively parallel solution for mapping interactions between genomic elements, *Genome Research*, **16**: 1299–1309

5. M. Imakaev et al. (2012) Iterative correction of hi-c data reveals hallmarks of chromosome organization, *Nature methods*, **9(10)**: 999-1003

6. D. Fillipova et al. (2014) Identification of alternative topological domains in chromatin, *Algorithms for Molecular Biology,* 9:14

7. J. R. Dixon et al. (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions, *Nature*, **485**: 376-380

8. E. Lieberman-Aiden et al. (2009) Comprehensive mapping of long range interactions reveals folding principles of the human genome, *Science*, **326(5950)**: 289–293

9. S. S. Rao et al. (2014) A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping, *Cell*, **159**: 1–16