Role of DNA sequence in establishing histone modifications in humans

Anatoliy Zubritskiy[1], Yulia A. Medvedeva[1,2,3]

1 Institute of Bioengineering, Research Center of Biotechnology, Russian Academy of Sciences, Moscow 119071, Russian Federation

2 Moscow Institute of Physics and Technology, Dolgoprudny, Moscow Region 141701, Russian Federation

3 Vavilov Institute of General Genetics, Russian Academy of Sciences, Moscow 119991, Russian Federation

Epigenetic modifications, such as DNA methylation and histone marks are distributed unevenly along the genome being highly stable across cell types in some loci and highly variable in the others. There are some insights that DNA sequence can determine the likelihood of the CpG methylation. Insertion of a human chromosome into a mouse cell revealed that promoter-associated CGIs were refractory to DNA methylation regardless of the host species, demonstrating that DNA sequence plays a central role in specifying the hypomethylated state through evolutionarily conserved mechanisms [1]. Successful efforts are made in predicting of DNA methylation state based on its sequence only [2]. Moreover, DNA methylation can be restored for the DNA sequence of specific GC-content [3-4].

The role of DNA sequence for establishing histone modifications is less clear. It is widely accepted that the majority of histone modifications occur in the promoter regions (H3K4me3, H3K27ac, H3K27me3, etc), exons (H3K36me3), and enhancers (H3K4me1, H3K27ac), regions enriched for specific DNA motifs absent elsewhere in the genome. It has been shown that CpG island promoters demonstrate very distinct histone modification profile [5]. It is not entirely clear whether the effect of DNA sequence on histone modifications is direct or is mediated by an act of active transcription. Yet, DNA sequence enriched for CpG dinucleotides is able to recover H3K4me3 profile even being far away from active promoters [4]. Most likely the sequence pattern that determines establishing and maintenance of H3K4me3 is CpG-containing motif bound by the transcription factor Cfp1 (CXXC1), which helps precise localization of histone methylases in the genome [6]. Also, attraction of PRC2 complex and consequent H3K27me3 modification is correlated with local density of CpG dinucleotides [7]. Not only transcription factors, that recognize DNA motifs, can serve as a proxy between DNA sequence and histone modifications. Long non-coding RNAs could also direct chromatin modifying proteins towards specific locations in DNA (reviewed in [8]). Several mechanisms potentially connecting DNA sequence with epigenetic modifications are already known, yet even more has to be discovered.

In this work we perform a direct experiment to detect sequence patterns that determine acetylation of the histones (H3K27ac). To do so we test whether various sequences from genomic regions with high level of histone acetylation are capable of re-establishing acetylation after being inserted into human β-globin locus in CAKI-1 cell line, the region where level of histone acetylation is known to be low. The exact region of insertion was selected to be far from promoters to eliminate the effect of transcription initiation. To perform the insertion of the DNA fragments into the same locus high-throughput recombinase-mediated cassette exchange method is adapted from [4]. In brief, cells are transfected with floxed (flanked with loxP sites) HSVtk-NeoR selection cassette targeted to intergenic space of human β-globin locus. Library of genomic fragments are obtained by chromatin immunoprecipitation (ChIP) in CAKI-1 cell line using H3K27ac antibody. We clone obtained DNA fragments into floxed libraries and perform cassette exchange using Cre recombinase. Then we again perform chromatin immunoprecipitation with the same H3K27ac antibody on the libraries carrying foreign fragments inserted to β-globin locus. We amplify the inserted foreign fragments to produce two libraries for Illumina sequencing: (1) containing all fragments inserted into β-globin locus; (2) containing DNA fragments that re-establish H3K27ac profile. Bioinformatic analysis of the fragments from both libraries is performed to obtain sequence patterns responsible for re-establishing of H3K27ac.

The work is supported by RSF grant 15-14-30002 to YAM.

1. Long HK, King HW, Patient RK, Odom DT, Klose RJ. Protection of CpG islands from DNA methylation is DNA-encoded and evolutionarily conserved. Nucleic Acids Res. 2016 Aug 19; 44(14):6693-706.
2. Huska M, Vingron M. Improved Prediction of Non-methylated Islands in Vertebrates Highlights Different Characteristic Sequence Patterns. PLoS Comput Biol. 2016 Dec 16;12(12):e1005249.
3. Wachter E, Quante T, Merusi C, Arczewska A, Stewart F, Webb S, Bird A. Synthetic CpG islands reveal DNA sequence determinants of chromatin structure. Elife 3: e03397.
4. Krebs AR, Dessus-Babus S, Burger L, Schübeler D. High-throughput engineering of a mammalian genome reveals building principles of methylation states at CG rich regions. Elife. 2014 Sep 26; 3:e04094
5. Vavouri T., Lehner B. Human genes with CpG island promoters have a distinct transcription associated chromatin organization. Genome Biol. 2012 Nov 27;13(11):R110.

6. Clouaire T, Webb S, Skene P, Illingworth R, Kerr A, Andrews R, Lee JH, Skalnik D, Bird A. Cfp1 integrates both CpG content and gene activity for accurate H3K4me3 deposition in embryonic stem cells. Genes Dev. 2012 Aug 1;26(15):1714-28.
7. Jermann P, Hoerner L, Burger L, Schübeler D. Short sequences can efficiently recruit histone H3 lysine 27 trimethylation in the absence of enhancer activity and DNA methylation. Proc Natl Acad Sci USA. 2014 Aug 19;111(33):E3415-21.
8. Backofen, R., Vogel T. Biological and bioinformatical approaches to study crosstalk of long non-coding RNAs and chromatin-modifying proteins. Cell Tissue Res. 2014 Jun; 356(3):507-26.