LncRNAs and their role in chromatin modifications

S.V. Vinogradova

Moscow State University, Faculty of Bioengineering and Bioinformatics, 119991, Moscow, GSP-1, Leninskiye Gory, MSU, 1-73 kintany@gmail.com

M.A. Mironov

Moscow State University, Faculty of Bioengineering and Bioinformatics, 119991, Moscow, GSP-1, Leninskiye Gory, MSU, 1-73 mironov@bioinf.fbb.msu.ru

Long non-coding RNAs (lncRNAs) play important role in chromatin modifications and usually act as scaffolds between different chromatin-modifying complexes. One example is HOTAIR that binds polycomb complex PRC2 that methylates histone H3 on K27 and second complex that demethylates histone H3 on K4.

The ability of double-stranded DNA to form a triple-helix structure by hydrogen bonding with a third strand of RNA was shown 50 years ago [1, 2], but biological function of such structures is still unclear. Recently it has been shown that 'triplex-forming oligonucleotides' (TFOs) take part in regulation of the dehydrofolate reductase gene: a non-protein-coding transcript (ncRNA) interacts in a sequence specific way to form a triplex within the downstream major promoter [3]. It was also suggested that formation of triplex complexes provides contact points that may participate in chromatin organization possible through ncRNAs, which in turn may interact with different protein complexes [4]. One well-studied example of lncRNA that mediates chromatin modifications is HOTAIR. HOTAIR binds the polycomb complex PRC2 which methylates histone H3 on K27 to promote gene repression [5]. In addition, the 3' domain of HOTAIR was also found to interact with a second complex containing LSD1, CoREST, and REST that demethylates histone H3 on K4 to antagonize gene activation. So it was suggested that HOTAIR acts as a scaffold and bridges between two complexes. However it is unclear whether domains in HOTAIR may bind yet unidentified protein partners or to genomic DNA.

We searched for TFOs of length 4 in exon 6 of human HOTAIR and found 3 motifs of length

13. The same procedure was repeated for other mammalian genomes. The first motif (65-81 bp in human HOTAIR) was conservative in genomes of human, cow, horse and dog and was selected as the main motif for further analysis. We searched for TTSs to match this motif in the genomes. We found 2693 matches for the human, 1917 for cow, 2266 for dog and 3307 for horse. To verify that these TTS are not random, we mapped them on the genes in corresponding genomes and lifted over the matches between orthologous genes. We found that there are more interceptions than are expected by chance. Next we counted intersections between DNAse open sites and TTSs for human genome. To estimate importance of these intersections we shuffled positions of TTS sites and estimated numbers of intersections with DNAse track (z-score=18). So we propose that TTSs tend to be in open chromatin sites. Next we downloaded and analyzed histone modifications data for cell lines in which HOTAIR is expressed: HSMM (skeletal muscle myoblasts) and NHEK (epidermal keratinocytes). We've shown that some chromatin features are correlated with TTSs across the human genome.

1. Felsenfeld, G., D.R. Davies, and R. A. (1957) *Formation of a three-stranded polynucleotide molecule* J. Am. Chem. Soc., 1957. 79(8): p. 2023–2024.
2. Morgan, A.R. and R.D. Wells (1968) *Specificity of the three-stranded complex formation between double-stranded DNA and single-stranded RNA containing repeating nucleotide sequences*. J Mol Biol. 37(1): p. 63-80.
3. Martianov, I., et al. (2007) *Repression of the human dihydrofolate reductase gene by a non-coding interfering transcript*. Nature 445(7128): p. 666-70.
4. Zheng, R., et al. (2010) *Polypurine-repeat-containing RNAs: a novel class of long non-coding RNA in mammalian cells*. J Cell Sci, 123(Pt 21): p. 3734-44.
5. Tsai, M.C., et al. (2010) *Long noncoding RNA as modular scaffold of histone modification complexes.* Science. 329(5992): p. 689-93.
6. Buske, F.A., et al. (2012) Triplexator: detecting nucleic acid triple helices in genomic and transcriptomic data. Genome Res. 22(7):1372-81
7. Quinlan, A.R. and Hal,l I.M. (2010). *BEDTools: a flexible suite of utilities for comparing genomic features.* Bioinformatics. 26, 6: pp. 841–842.