Analysis of 3 dimensional interactions in DNA and chromatin

Riina Kaukonen

Human genome is composed of approximately 22 000 protein coding genes. All the somatic cells in the human body contain exactly the same genes, so why do we have various different cell types and tissues? The answer lies on strictly regulated gene expression. During the differentiation some genes are activated while other genes are silenced. Correct expression of the genes is crucial for the homeostasis of the cell. Deregulation of gene expression may lead to the development of cancer or other diseases. To ensure that gene is correctly expressed, each gene contains specific control regions called promoters and distal regulatory elements. Promoter is located upstream from the protein coding region and it binds the enzyme called RNA-polymerase II (RNA-pol II). RNA-pol II converts the information coded to the DNA into the messenger RNA (mRNA) in a process called transcription.

RNA-pol II recognizes and binds to the promoter and attracts other proteins called general transcription factors (GTFs) that are needed for gene to be expressed. RNA-pol II together with GTFs promotes transcription at a low rate that is called basal level. Additional regulatory proteins called silencers, activators and insulators binds to the distal regulatory elements and increases or decreases the rate of transcription. The main function of the regulatory proteins is to ensure that the gene expression is strictly controlled in time and place.

Various different methods have been developed to reveal the mystery behind the extremely complex regulation of gene expression. Chromatin immunoprecipitation (ChIP) is a method that can be used to analyze the binding sites of the regulatory proteins. ChIP is a method, where protein-chromatin interactions are fixed and protein of interest is precipitated with specific antibody. DNA from the protein-chromatin complexes can be analyzed via sequencing (ChIP-seq) or by the microarray (ChIP-chip). Sequences can be mapped to the reference genome in order to create genome wide map of protein binding sites and to identify regulatory elements.

There is a need for new methods to study gene expression due to the complexity of regulatory networks and interactions between the regulatory proteins. Latest ChIP-based method called Chromatin Interaction Analysis using Paired End Tag sequencing (ChIA-PET) can be used to study interactions between the promoters and distal regulatory elements. ChIA-PET combines immunoprecipitation to proximity ligation. Proximity ligation means that DNA-molecules that are close to each other are ligated together. This means that if DNA fragments are brought together due to the regulatory protein interactions, they will be proximity ligated and ligation products can be analyzed via sequencing.

ChIA-PET can provide genome wide information not only about the locations of regulatory elements but also about the interactions between the regulatory elements and the promoters. ChIA-PET method can be used to reveal the biological function of the regulatory elements. It may help to understand for example why some mutations lead to development of the disease and why are different versions of the same gene, called alleles, expressed differentially.