DOI 10.18699/SBB-2019-07

Computer tools for spatial chromosome contacts analysis by ChIA-PET and Hi-C data

Dergilev A.I.[1]*, Luzin A.N.[1], Kovalev S.S.[2], Babenko R.O.[1], Orlov Y.L.[1, 2]
[1] *Novosibirsk State University, Novosibirsk, Russia*
[2] *Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia*
** e-mail: arturd1993@yandex.ru*

Key words: genome, transcription factor, medical informatics, ChIP-seq

Motivation and Aim: 3D contacts and interactions in interphase nuclei of eukaryotic cell play critical role for gene expression regulation. Series of post-genome technologies have been developed to study the binding of transcription factors for transcription regulation, such as chromatin immunoprecipitation arrays (ChIP-Seq) [1]. Correspondingly, set of software tool for processing of such data has been developed. Identification of genome-wide distal chromatin interactions that lead the regulatory elements to their target genes may provide novel insights into the study of transcription regulation. Chromatin Interaction Analysis with Paired-End-Tag sequencing (ChIA-PET) method for such analysis requires development of specialized software [2]. The aim of the work was to review existing tools for 3D genome structure and spatial topological domains analysis.
Methods and Algorithms: The data have been obtained via available data sources containing experimental information from ChIP-seq, Hi-C, ChIA-PET tests using different sequencing platforms. Gene annotation was obtained from UCSC Genome Browser. We reviewed existing software and created a database prototype of bioinformatics tools for 3D genome structure analysis.
Results: We tested program for analysis of ChIA-PET experimental data. The result of the program is a distribution of CTCF transcription factor binding sites on domains on the human chromosomes. The distributions of human genes relative CTCF binding sites and a randomly generated list of such sites as the program output were used to estimate statistical significance of the associations found.
Conclusion: With the rapidly increasing resolution of Hi-C datasets, the size of the chromatin contact map will soon exceed the memory capacity of general computers. ChIA-PET and Hi-C technologies provide huge volume of data demanding development of new computer tools in different applications.
Acknowledgements: The work was supported by the RFBR and ICG SB RAS budget project (0324-0019-0040).

References

1. Li G. et al. Chromatin Interaction Analysis with Paired-End Tag (ChIA-PET) sequencing technology and application. *BMC Genomics*. 2014;15(Suppl. 12):S11.
2. Fullwood M.J. et al. An oestrogen-receptor-alpha-bound human chromatin interactome. *Nature.* 2009;462(7269):58-64.