Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 119 (2017) 182–189

Procedia Computer Science

www.elsevier.com/locate/procedia

6th International Young Scientists Conference in HPC and Simulation, YSC 2017, 1-3 November 2017, Kotka, Finland

Multiscale dynamic visualization of signal transduction processes with detailing of target-genes activation in three-dimensional genome structure

Kseniia Bureiko[a]*, Maria Semashko[a], Ksenia D. Mukhina[a] and Andrey Karsakov[a]

[a]*ITMO University, Saint Petersburg, Russia*

Abstract

Given the complexity of modern biological data it is essentially crucial to accord a consistent expounding. Interpreting such data into complex networks and visualizing them can reveal understanding of various processes in a cell. A consequence mapping of signal transduction processes to the spatial genome structure can benefit new insights in interaction detection in the spatial arrangement of genes. We present an approach for multiscale dynamic visualization of signal transduction processes with detailing of target-genes activation in spatial genome structure. The usage of this approach is demonstrated for the WNT signaling pathway in a human cell. We conclude with suggesting future research questions to improve our approach by considering new available data.

Peer-review under responsibility of the scientific committee of the 6th International Young Scientist conference in HPC and Simulation

Keywords: Bioinformatics visualization; signaling pathways; spatial genome structure

1. Introduction

In today's world, living cell simulation is a great challenge which can be solved using modern theoretical and experimental methods. There were some efforts to create a static model based on structure of cells and intracellular organelles, their biochemical composition, and location [1]. However, such complex problem as a dynamic model is

* Corresponding author. Tel.: +358-44-919-26-10
E-mail address: kseniia.bureiko@aalto.fi

1877-0509
Peer-review under responsibility of the scientific committee of the 6th International Young Scientist conference in HPC and Simulation
10.1016/j.procs.2017.11.175

required to consider changes of chemical composition and special aspects of intracellular processes in time and space [2].

Networks are often used in modeling and analyzing variety of biological systems. Molecular networks have primary importance among all biological networks. They include genes, proteins, metabolic and signaling networks as key components. Molecular networks analysis plays a vital role in identification of functional modules and determination of components' roles in cell functioning [3]. Thus, signaling networks is one of the most effective way to study signal transduction processes.

Signal transduction processes are often described as a composition of interactions between participants of signal pathways. However, cell scale does not reflect processes which occur after signal transmission in nucleus. The additional genome scale can yield understanding about relations between the signal pathway and the location of activated genes in chromatin.

Such combination of signaling transduction and gene regulation is commonly used for differential gene expression analysis. The main compounds of signal pathways in a cell are proteins encoded by genes. Due to that fact, it is important to consider signaling pathways with the current expression level of their encoding genes. Some of the pathways, Hedgehog pathway for example, has self-regulation mechanisms. They affect the way of signal transduction process through initialization of changes in genes transcription of encoding proteins.

However, in this work the main emphasis is on the positions of activated genes in chromatin regardless of whether they are or not a part of the signaling pathway. Our main contribution is to propose an approach to visualization of signal pathways which can reveal connections between two scales and to give new insights into the phenomenon of signal transduction.

2. Related works

Biological pathways are a natural representation of the coordinated reactions and actions occurring in a cell, thus they became one of the leading methods in -omics data analysis and visualization. In this part we review some tools for visual exploration of biological networks [4].

Most tools allow user to explore static and simple representation of pathways based on information from such popular databases as KEGG, Reactome, WikiPathways, BioCyc. The main goal of these tools is providing user as much detail as possible, but these tools do not reproduce processes step-by-step. Thus, these tools have some limitations and inconveniences in analysis of experimental data [4], [5].

Another type of visualization tools, VisANT[6] and Cytoscape[7], allows users to integrate their own data by rendering it on top of the pathway maps.

VisANT is a web-based platform for signal pathways analysis using metagraphs which allow users to represent nodes, edges and subnetworks in a nested structure. Moreover, metagraphs provide an opportunity to reveal roles of nodes in different compounds automatically. Cytoscape is a tool for representing molecular interactions as a network diagram. This tool mainly provides 2D representations, but it is suitable for large-scale network analysis with hundreds of thousands of nodes and edges. Cytoscape is an efficient tool for comparing networks between each other.

Despite the special aspects of each tool relating to analysis of signal pathways, visualization solutions are identical. These tools do not focus on visualization of processes' flow. The information is organized in hierarchical levels and is shown as a 2D static graph. Such structure allows users to analyze a composition of signal pathways, but unfortunately it does not show process which is dynamic in nature.

Tools for spatial visualization of the genome should be considered separately, since no tool for visualization of signaling pathways allows us to consider this aspect two together.

Modern biological methods provide a large quantity of heterogeneous and unnormalized datasets. Being a combination of chromatin contact experiments and next-generation sequencing (NGS), such techniques as Hi-C[8], ChiA-PET[9] etc. yield different representation of three-dimensional genome structure. Nowadays, it becomes possible to achieve chromosome reconstruction from chromosomal contacts [10]–[13].

3D-GNOME [14] provides a web-based, interactive 3D viewer to visualize and analyze the resulting 3D structure from ChiA-PET data, and includes options for the user to upload genomic annotation data to overlay on the structure. At each level, the chromatin is represented as a beads-and-springs polymer, where beads represent different genomic regions. GMOL [15] was developed based upon multi-scale approach that allows users to zoom in and out between