The System-Biological GLOBE 3D Genome Platform

Michael Lesnussa 1,2, Frank N. Kepper 1,2,3, Hubert B. Eussen 1,3 and Tobias A. Knoch 1,2,5
in cooperation with A. de Klein 1, K. Rippe 1, F. G. Grosvold 1

The GLOBE-Consortium
headed by B. Eussen and Tobias A. Knoch

The combination of genome sequence and structure, its annotation and experimental data in an accessible and comprehensible way is a major challenge. Increasingly there is a large number of extremely divergent data sets: the sequence itself, genes, regulatory regions, various forms of measuring sequence features and clone sets etc. Currently, one possibility to represent this information in a visual form - and thus to reveal its scientific meaning - is to use genome browsers such as "Ensembl" or "The UCSC Genome Browser". These browsers have been beneficial in the understanding of the complex organization of genomes. However, there are also limitations concerning their focus to linear presentation, standardized input and data accessibility. Also constructibility by a remote user with special requirements is difficult. The GLOBE-Consortium is developing ways to visualize multi-dimensional data sets from various sources in an easily accessible way. This allows the integration of these data sets into a single holistic display system giving a biological oriented view of genomes and advancing basic research, diagnostics and new treatments.

Introduction

The genome viewer presented here enables navigation to visualize and analyse the multi-dimensional aspects of genomes in a new intuitive way in combination with a data-warehouse and a computing grid also being set-up by the GLOBE-Consortium at the Erasmus Medical Center. In an environment with active new interactive possibilities, innovation has been created. This allows new possibilities for future research, leading to a better understanding of the holistic properties of genomes, which is necessary for advanced diagnostic services and perhaps ultimate treatments.

Fig. 1: Complete merged clone set (UCSC: hg18 assembly) of 15 chromosomes to individual bases. This new environment is given a biological oriented view of genomes and advancing basic research, diagnostics and new treatments.

Conclusion

The genome viewer presented here enables navigation to visualize and analyse the multi-dimensional aspects of genomes in a new intuitive way in combination with a data-warehouse and a computing grid also being set-up by the GLOBE-Consortium at the Erasmus Medical Center. In an environment with active new interactive possibilities, innovation has been created. This allows new possibilities for future research, leading to a better understanding of the holistic properties of genomes, which is necessary for advanced diagnostic services and perhaps ultimate treatments.

Fig. 2: Multi-chromosomal relation view between duplicon regions in and between chromosomes 15 & 21. Colours: duplicon spreading degree.

Fig. 3: Intra-chromosomal duplicons (Eichler et al.) compared to syndromes (blue/green), literature hot-spots (orange), and our defined bigrams (pink) in the chr 21q11 region.

Fig. 4: Dynamic zoom into the level of the DNA.

Fig. 5: Background image: Multi-chromosomal relation view (the duplicons of 15 & 15 to all other chromosomes. Coloured as in Fig. 3.)

Fig. 6: Correlation of a simulated 3D chromatin/chromosome topology, combined with the - in principle - linear information content in the DNA sequence and multi-dimensional mapping of Fig. 15.

Features

Flexible Customizable Intuitive Navigation
Real-Time Interaction & Analysis
Dynamical Resolution & Arrangement
Extremely Large & Multi-Dimensional Data

Bridge ALL Scales from Sequence to Morphology

Data Tracks
Syndrome Chromosome Break Points Ideogram Bands Duplication Chromatin Loops Repeat Regions Epigenetics Syndrome Chromosome Break Points Ideogram Bands Duplication Chromatin Loops Repeat Regions Epigenetics

Intra-Relations

Using the dynamic scaling range of the Intra-chromosomal relationship, it can be visualized and analyzed on a large range of data sources - from whole genome to single bases. This makes it possible to integrate diagnostics and treatments. Always can be projectable, related measured and labeled, so allowing us to various genome levels to scale-free insights.

Resolution Scale

The viewer has a large dynamic range in the size and resolution of the features it can display, from whole chromosomes to the individual bases. The user interface creates entire new possibilities for understanding genome organization.

Structure

There are several physical levels of genetic information storage, e.g. DNA, chromatin and chromosomes. The visualization of these levels of critical importance for genome function. The viewer allows the visualization of 3D genomic structures and to project and link these to a classical linear representation.

Inter-Relations

In addition to the simultaneous mapping on one chromosome, the viewer allows the analysis of inter-chromosomal relationships either on an external input (Fig. 2) or internal simulation analysis (Figs. 1, 4, 6). Every genome level can be generated for the visualization of diagnostics or gene function to investigate etc.

Multi-Mapping

The viewer allows the mapping of classical and experimental data tracks projected into metaphase chromosomes simultaneously (Fig. 1). The general tracks as well as every single track element layout is customizable e.g. colors, size and shape of a given component can be altered. It is possible to generate an unlimited number of elements.

Fig. 1: Complete merged clone set (UCSC: hg18 assembly) of 15 chromosome. Colours represent association with duplicon regions.
A New Holistic Genome Viewer for Molecular Genetics

–

The System-Biological GLOBE 3D Genome Platform

Lesnussa, M., Kepper, F. N., Eussen, H. G. & Knoch, T. A.


**Abstract**

Genomes are tremendous co-evolutionary holistic systems for molecular storage, processing and fabrication of information. Their system-biological complexity remains, however, still largely mysterious, despite immense sequencing achievements and huge advances in the understanding of the general sequential, three-dimensional and regulatory organization. Here, we present the GLOBE 3D Genome Platform a completely novel grid based virtual “paper” tool and in fact the first system-biological genome browser integrating the holistic complexity of genomes in a single easy comprehensible platform: Based on a detailed study of biophysical and IT requirements, every architectural level from sequence to morphology of one or several genomes can be approached in a real and in a symbolic representation simultaneously and navigated by continuous scale-free zooming within a unique three-dimensional OpenGL and grid driven environment. In principle an unlimited number of multi-dimensional data sets can be visualized, customized in terms of arrangement, shape, colour, and texture etc. as well as accessed and annotated individually or in groups using internal or external data bases/facilities. Any information can be searched and correlated by importing or calculating simple relations in real-time using grid resources. A general correlation and application platform for more complex correlative analysis and a front-end for system-biological simulations both using again the huge capabilities of grid infrastructures is currently under development. Hence, the GLOBE 3D Genome Platform is an example of a grid based approach towards a virtual desktop for genomic work combining the three fundamental distributed resources: i) visual data representation, ii) data access and management, and iii) data analysis and creation. Thus, the GLOBE 3D Genome Platform is the novel system-biology oriented information system urgently needed to access, present, annotate, and to simulate the holistic genome complexity in a unique gateway towards a real understanding, educative presentation and curative manipulation planning of this tremendous evolutionary information grail – genomes.

*Corresponding author email contact:* TA.Knoch@taknoch.org

**Keywords:**

Genome, genomics, genome organization, genome architecture, structural sequencing, architectural sequencing, systems genomics, coevolution, holistic genetics, genome mechanics, genome function, genetics, gene regulation, replication, transcription, repair, homologous recombination, simultaneous co-transfection, cell division, mitosis, metaphase, interphase, cell nucleus, nuclear structure, nuclear organization, chromatin density distribution, nuclear morphology, chromosome territories, subchromosomal domains, chromatin loop aggregates, chromatin rosettes, chromatin loops, chromatin fibre, chromatin density, persistence length, spatial

**Literature References**


