3DGV: Immersive Exploration of 3D Genome Structures using Virtual Reality

Éric Zhang^{1,*}, Chrisostomos Drogaris^{1,*}, Antoine Gédon¹, Aaron Sossin¹, Rajae Faraj¹, Haifen Chen², Yan Cyr³, Jacek Majewski², Mathieu Blanchette¹, and Jérôme Waldispühl^{1,✉}

¹School of Computer Science, McGill University, Montréal, QC, H3A 0E9, Canada

²Department of Human Genetics, McGill University, Montréal, QC, H3A 1B1, Canada

³Beam Me Up Labs Inc., Montréal, QC, H2T 1R2, Canada

The analysis of 3D genomic data is expected to revolutionize our understanding of genome organization and regulatory mechanisms. Yet, the complex spatial organization of this information can be difficult to interpret with 2D viewers. Virtual Reality (VR) technologies offer an opportunity to rethink our methods to visualize and navigate 3D objects. In this paper, we introduce the Virtual Reality 3D Genome Viewer (3DGV), an open platform to experiment and develop VR solutions to explore 3D genome structures.

Availability: <http://3dgv.cs.mcgill.ca/>

Key words: 3D genome, visualization, genome browser, virtual reality.

Correspondence: jeromew@cs.mcgill.ca

* *The first two authors should be regarded as joint first authors.*

Introduction

As chromosome conformation capture (3C) technologies continue to improve, the need for applications and software to process and review 3D genomic data is growing (1). In particular, the release of the first complete 3D models (2) suggests the development of customized tools to exploit this novel information.

Currently, software designed to explore 3D genome structures is primarily desktop-based and distributed as stand-alone applications (3–5) or integrated in a Web browser (6–8). But the browsing experience is limited by the representation of 3D objects as 2D projections on a screen and the manipulation of the data through keyboards and mice. Immersive technologies present an opportunity to address this challenge (9, 10).

In this paper, we introduce the Virtual Reality 3D Genome Viewer (3DGV) – a virtual reality (VR) version of our previous 3D genome browser (6). Notably, 3DGV introduces multiple concepts and techniques to facilitate the manipulation of the data in immersive environment, as well as visual representations allowing to overlay multiple levels of genomic annotations in 3D space. This software aims to lower the barrier to entry to immersive technologies for 3D genomics and serve as a base for further experimentations.

Methods

Input data. The Virtual Reality 3D Genome Viewer (3DGV) displays *pre-computed* 3D models of genome structures. In addition, users can also upload genomic annotations they wish to analyze in this structural context. Sequence data is stored in a basic text file. For structure coordinates and other annotations, 3DGV uses a simple tab-separated values (TSV) file format that enables us to easily convert annotations from any file format. We represent 3D models as a trajectory where each row of the file contains a node identifier, its genome index and spatial coordinates. A gene annotation file stores in each row the chromosome number, start and end indices, name and orientation of a gene. Other annotations use a generic format in which each row store a chromosome number, start and end indices, label and value associated to this region. The current distribution features 3D models of the 23 Human chromosomes from (11) with gene annotations stored in separate files. We also include CpG islands and Methylation data to illustrate the versatility of the file format and test 3DGV functionalities.

Visualization and annotations. 3D structures of chromosomes are presented as 3D splines on the top of which we overlay annotations (See Fig. 1a). Genes possess their own level of representation. We display the genomic region they occupy with a cylinder wrapped around the spline, where the color gradient show the orientation of this gene. Other annotations are represented at a third level with larger cylinders whose color encode the value associated to this segment. For instance, the relative amount of methylation of regions of a chromosome. Multiple annotations can be combined and mapped on the same 3D model, but overlaps can occur at the third level. We do not display the sequence data.

Navigation and Controls. By default, users navigate the 3D space and manipulate data with two controllers. The main menu allows the users to activate the 3D models and annotations to display. The speed of motion and resolution of the model (i.e zoom) can be adjusted. We also implemented a search function allowing the user to access immediately specific coordinates. During the exploration of the data, a pointer enables the users to query specific regions of the genome and display the information associated within a panel (See Fig. 1b). These panels feature a graphical summary and can

Fig. 1. (Left) 3D structures with annotated with gene (in green). The darker green ring shows the start of the gene. (Right) Information panel showing gene positions, CpG and Methylation data of the region highlighted in Magenta.

be conveniently organized to compare annotations in distinct regions.

In order to improve accessibility, keyboard and mouse functionality was added to the application in order to allow users who do not have access to virtual reality to use 3DGV.

Implementation and Distribution. The Virtual Reality Genome Viewer (3DGV) is distributed as a standalone application built using the Unity game engine that works with most popular commercial VR headsets and controllers. The current release has been tested on HTC Vive and Oculus Go. It can also be used on desktops using a non-VR mode under OSX and Windows operating systems.

Funding

This work has been supported by Genome Canada, Genome Québec and the Canada Institute of Health Research (Bioinformatics & Computational Biology competitions 2015 & 2018).

Bibliography

1. Jérôme Waldispühl, Eric Zhang, Alexander Butyaev, Elena Nazarova, and Yan Cyr. Storage, visualization, and navigation of 3D genomics data. *Methods*, 142:74–80, 06 2018. doi: 10.1016/j.ymeth.2018.05.008.
2. Erez Lieberman-Aiden, Nynke L van Berkum, Louise Williams, Maxim Imakaev, Tobias Ragoczy, Agnes Telling, Ido Amit, Bryan R Lajoie, Peter J Sabo, Michael O Dorschner, Richard Sandstrom, Bradley Bernstein, M A Bender, Mark Groudine, Andreas Gnirke, John Stamatoiyannopoulos, Leonid A Mirny, Eric S Lander, and Job Dekker. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*, 326(5950):289–93, Oct 2009. doi: 10.1126/science.1181369.
3. Thomas M Asbury, Matt Mitman, Jijun Tang, and W Jim Zheng. Genome3D: a viewer-model framework for integrating and visualizing multi-scale epigenomic information within a three-dimensional genome. *BMC Bioinformatics*, 11:444, Sep 2010. doi: 10.1186/1471-2105-11-444.
4. Jackson Nowotny, Avery Wells, Oluwatosin Oluwadare, Lingfei Xu, Renzhi Cao, Tuan Trieu, Chenfeng He, and Jianlin Cheng. GMOL: an interactive tool for 3D genome structure visualization. *Sci Rep*, 6:20802, Feb 2016. doi: 10.1038/srep20802.
5. Tuan Trieu, Oluwatosin Oluwadare, Julia Wopata, and Jianlin Cheng. GenomeFlow: a comprehensive graphical tool for modeling and analyzing 3D genome structure. *Bioinformatics*, Sep 2018. doi: 10.1093/bioinformatics/bty802.
6. Alexander Butyaev, Ruslan Mavlyutov, Mathieu Blanchette, Philippe Cudré-Mauroux, and Jérôme Waldispühl. A low-latency, big database system and browser for storage, querying and visualization of 3D genomic data. *Nucleic Acids Res*, 43(16):e103, Sep 2015. doi: 10.1093/nar/gkv476.
7. Ruifeng Li, Yifang Liu, Tingting Li, and Cheng Li. 3Disease browser: A web server for integrating 3D genome and disease-associated chromosome rearrangement data. *Sci Rep*, 6:34651, 10 2016. doi: 10.1038/srep34651.
8. François Serra, Davide Baù, Mike Goodstadt, David Castillo, Guillaume J Filion, and Marc A Marti-Renom. Automatic analysis and 3D-modelling of Hi-C data using TADbit reveals structural features of the fly chromatin colors. *PLoS Comput Biol*, 13(7):e1005665, Jul 2017. doi: 10.1371/journal.pcbi.1005665.
9. Jimmy F Zhang, Alex R Paciorkowski, Paul A Craig, and Feng Cui. BioVR: a platform for virtual reality assisted biological data integration and visualization. *BMC Bioinformatics*, 20(1):78, Feb 2019. doi: 10.1186/s12859-019-2666-z.
10. Stephen Todd, Peter Todd, Simon J McGowan, James R Hughes, Yasutaka Kakui, Frederic Fol Leymarie, William Latham, and Stephen Taylor. CSynth: a dynamic modelling and visualisation tool for 3D chromatin structure. *bioRxiv*, 2019. doi: 10.1101/499806.
11. Tuan Trieu and Jianlin Cheng. Large-scale reconstruction of 3D structures of human chromosomes from chromosomal contact data. *Nucleic Acids Res*, 42(7):e52, Apr 2014. doi: 10.1093/nar/gkt1411.