

# Exploring genomic data coupled with 3D chromatin structures using the WashU Epigenome Browser

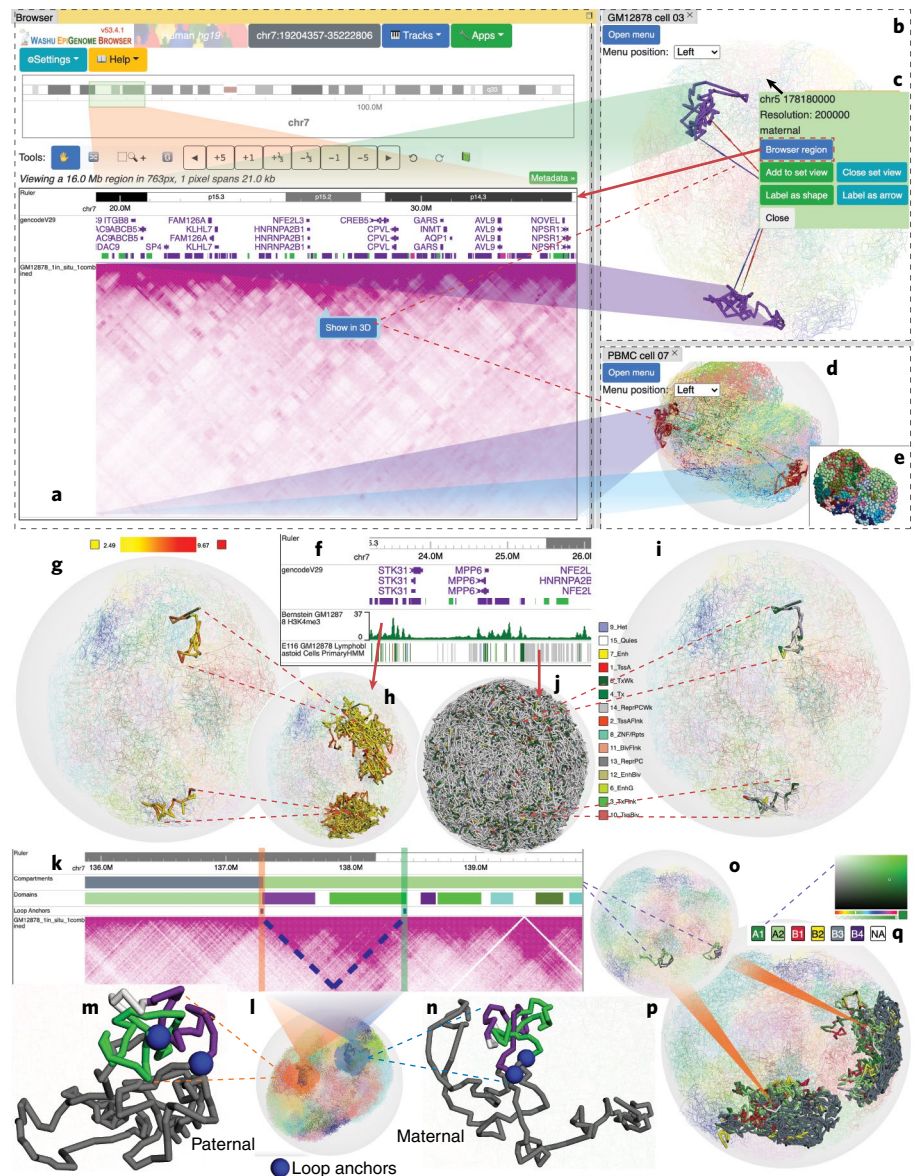
To the Editor — Three-dimensional (3D) genomic structures are vital for gene regulation and cell function<sup>1</sup>. High-throughput technologies based on chromosome conformation capture have been used to study genome-wide physical chromosome interactions<sup>2</sup>. These interactions can be visualized as 2D heatmaps<sup>3</sup> or as interaction networks decorated with genomic features<sup>4</sup>. In addition, computational approaches using interaction data based on models, such as constrained physical models, polymer models and population-based analysis, have been developed to predict the physical 3D structures of chromosomes<sup>5</sup>. Large consortia, such as ENCODE<sup>6</sup>, Roadmap<sup>7</sup>, and 4D Nucleome<sup>8</sup>, have generated tens of thousands of genome-wide datasets of transcription factor binding sites and epigenetic marks across numerous cell types and tissues. Biologists wish to visually explore the connections between these genome-wide profiles and 3D genome structures, which will facilitate the generation and testing of diverse hypotheses. This presents a challenge to conventional genome browsers<sup>9</sup>, where most genomic data is visualized in linear genomic coordinates. The WashU Epigenome Browser was invented in 2011 as an interactive tool for exploring genomic data in a web browser<sup>9,10</sup>. We have now expanded the browser functions to allow investigators to visually explore 1D, 2D and 3D genomic data on a single webpage. The key innovation is to thread the linear genomic coordinates onto a multi-resolution 3D model of the chromosome.

We created a data format called g3d to encode the 3D model and have provided tools to convert various 3D modeling approaches. Given a 3D model of the chromosome, any genomic information anchored on the linear coordinates can be displayed in a 3D model. The 3D browser contains the general WashU Epigenome Browser functions, as well as 3D-specific functions, which enable users to intuitively examine biological information that is difficult to display on a typical browser. To facilitate a direct comparison between the linear and 3D browser views, we designed a panel system for integrated data

visualization. The linear browser panel displays 1D and 2D data, while the 3D browser panel displays multiple 3D models (Fig. 1). The 3D highlighted region on the 3D browser and the linear browser are synchronized (Fig. 1b,d). Users can choose multiple non-adjacent genome regions from the 3D model and display them on the linear browser using the region-set-view mode or label the regions on the 3D model (Fig. 1c).

The 3D browser also contains a thumbnail viewer that provides a synchronized global display of the main viewer (Fig. 1e).

The 3D browser allows users to paint the 3D model with genomic data and features using two painting styles: annotation and numerical. In annotation painting, users can define regions on the 3D model using any genomic annotation in a segmentation format (Fig. 1i,j). In numerical painting,



**Fig. 1 | 3D visualization integration with panels, decoration in 3D structure, viewing chromatin loops and compartmental annotations. a–e.** The linear genome browser shows sequence, gene annotation, epigenetic marks and Hi-C heat map (a); two 3D structures are displayed in b,d. From the 3D structure, click a 3D segment to reveal a popup menu with options to highlight the linear browser region or label this segment (c). By default, the 3D viewer also comes with a thumbnail, which is synchronized with the main 3D viewer (e). **f–i.** Browser view shows cytobands, gene annotations, H3K4me3 signal and chromHMM segmental annotation for GM12878 (f). The H3K4me3 signal is used to paint the 3D model in the current region (g) and chromosome (h). Highlighting colors and scale can be customized from the legend and configuration menu. ChromHMM annotation is used to paint the 3D model in the current region (i) and chromosome. Other genomic annotation can be used to annotate the 3D in annotation painting (cytoband, Supplementary Fig. 1b,c; gene annotation, Supplementary Fig. 1d,e). **j.** Typical genome browser view with a ruler, compartment annotation with subtypes, bed track with loop anchor locations and Hi-C track for GM12878. Users can also customize category colors. **k–q.** Spatial compartment definition using five subtypes: A1, A2, B1, B2 and B3 (k). Blue dashed line indicates the domain or loop, which is highlighted in l,m,n. Orange and green rectangles indicate loop anchors. Two subdomains or loops colored purple and green in the browser view are also shown in the 3D view with the same colors. Blue spheres in the 3D model indicate the same loop anchors as in the browser view. Compartmental annotation is used to paint the 3D model in the current region (o) and chromosome (p). The color of each subtype can be customized by clicking on the subtype and selecting a color (q).

users can apply numerical data, such as GC content or epigenetic signals, to the 3D model. Users can configure paint thickness, background opacity, scales and color gradient (Fig. 1g,h). The 3D browser allows visualization of the multiple spatial compartment definitions (Fig. 1o,p and Supplementary Fig. 2). The spatial compartments can be customized in the 3D browser and viewed as an annotation track in the linear browser (Fig. 1q,k and Supplementary Fig. 2d,e). The 3D browser makes viewing chromatin loops and topologically associated domains (TADs) intuitive. Users can view Hi-C loop anchors on the 3D model and the linear browser, where genomic segments can be custom labeled corresponding to the ‘Domain’ track (Fig. 1k–n). Users can also load genomic locations of TADs to paint the 3D structure (Supplementary Fig. 3).

We collected 11,045 3D models from published studies, converted them to g3d format and built corresponding data hubs, including 3D models generated from single-cell studies (Supplementary Fig. 4 and Supplementary Table 1). Investigators can upload their own models for visualization, annotation and comparison. By displaying different models in multiple browser panels, investigators can compare differences in 3D modeling tools and compare single-cell and bulk Hi-C. Recent advances in genomic technologies and computational algorithms have provided unprecedented opportunities to probe chromatin interactions and

generate 3D models. These models not only facilitate investigation into the formation and function of the 3D genome, but also provide a different paradigm to display and interact with genomic data. Typical genome browsers anchor genomic data on a 1D, linear axis, emulating a process of untangling and straightening genomic DNA. This process makes it convenient and straightforward to overlay genomic data; however, it destroys the spatial configuration of the chromatin. We strived to maintain the 3D structure of the underlying genomic DNA by threading the linear coordinates through a 3D model. Our approach implements an updated coordinate system from the linear genome and the 3D spatial locations to build visualization tools into a seamless format. This enables investigators to intuitively examine 3D features, such as loops and TADs; to visualize all genomic data on 3D genome coordinates; and to explore the dynamics of the 3D genome structure.

#### Data availability

All data used in this study are published or public from open databases, which are listed in Supplementary Table 2.

#### Code availability

g3d-related source code is freely available at <https://github.com/lidaof/g3d>, and g3dtools documentation is available at <https://g3d.readthedocs.io/en/latest/>. The browser codebase is available at GitHub (<https://github.com/lidaof/eg-react>) and Zenodo (<https://doi.org/10.5281/zenodo.6353838>), and documentation can be found at <https://eg.readthedocs.io/en/latest/>. The repository <https://github.com/lidaof/eg-3d-demo> contains demo files for 3D visualization and brief instructions. Video tutorials are at <https://bit.ly/eg3dtutorial>. All code is open source. □

Daofeng Li<sup>1</sup>✉, Jessica K. Harrison<sup>1</sup>, Deepak Purushotham<sup>1</sup> and Ting Wang<sup>1,2</sup>✉

<sup>1</sup>Department of Genetics, The Edison Family Center for Genome Sciences & Systems Biology, Washington University School of Medicine, St. Louis, MO, USA. <sup>2</sup>McDonnell Genome Institute, Washington University School of Medicine, St. Louis, MO, USA. ✉e-mail: [dli23@wustl.edu](mailto:dli23@wustl.edu); [twang@wustl.edu](mailto:twang@wustl.edu)

Published online: 21 July 2022  
<https://doi.org/10.1038/s41592-022-01550-y>

#### References

- Cremer, T. & Cremer, C. *Nat. Rev. Genet.* **2**, 292–301 (2001).
- Kempfer, R. & Pombo, A. *Nat. Rev. Genet.* **21**, 207–226 (2020).
- Zhou, X. et al. *Nat. Methods* **10**, 375–376 (2013).
- Madrid-Mencia, M., Rainieri, E., Cao, T. B. N. & Pancaldi, V. *Nucleic Acids Res.* **48**, 4066–4080 (2020).
- Jerkovic, I. & Cavalli, G. *Nat. Rev. Mol. Cell Biol.* **22**, 511–528 (2021).
- ENCODE Project Consortium. *Nature* **489**, 57–74 (2012).
- Roadmap Epigenomics Consortium et al. *Nature* **518**, 317–330 (2015).
- Dekker, J. et al. *Nature* **549**, 219–226 (2017).
- Zhou, X. et al. *Nat. Methods* **8**, 989–990 (2011).
- Li, D. et al. *Nucleic Acids Res.* **50**, W774–W781 (2022).

#### Acknowledgements

We thank the 3Dmol.js team (<https://3dmol.csb.pitt.edu/>) for the fantastic 3D graphics library. We thank D. Sehnal for suggestions on the data structure of the g3d format and L. Tan and L. Fiorillo for data discussion. We also thank all the members of the Wang laboratory for testing and feedback. This work was supported by NIH R01HG007175, U01CA200060, U24ES026699, U01HG009391, UM1HG011585, U41HG010972 and U24HG012070.

#### Author contributions

T.W. conceived and oversaw all aspects of the study, supervised research and provided constructive feedback. D.L. developed and is maintaining the 3D visualization module in the Browser. J.K.H. and D.P. provided key contributions and useful feedback. The manuscript was written by D.L. and T.W. with additional contributions and edits from J.K.H. and D.P.

#### Competing interests

The authors declare no competing interests.

#### Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41592-022-01550-y>.

**Peer review information** *Nature Methods* thanks Vera Pancaldi and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.