Wang et al. Genome Biology (2018) 19:151
https://doi.org/10.1186/s13059-018-1519-9

Genome Biology

METHOD Open Access

The 3D Genome Browser: a web-based browser for visualizing 3D genome organization and long-range chromatin interactions

Yanli Wang[1], Fan Song[1], Bo Zhang[1], Lijun Zhang[2], Jie Xu[2], Da Kuang[3], Daofeng Li[4], Mayank N. K. Choudhary[4], Yun Li[5,6,7], Ming Hu[8], Ross Hardison[9], Ting Wang[4*] and Feng Yue[1,2*]

Abstract

Here, we introduce the 3D Genome Browser, http://3dgenome.org, which allows users to conveniently explore both their own and over 300 publicly available chromatin interaction data of different types. We design a new binary data format for Hi-C data that reduces the file size by at least a magnitude and allows users to visualize chromatin interactions over millions of base pairs within seconds. Our browser provides multiple methods linking distal *cis*-regulatory elements with their potential target genes. Users can seamlessly integrate thousands of other omics data to gain a comprehensive view of both regulatory landscape and 3D genome structure.

Background

The three-dimensional (3D) organization of mammalian genomes plays an essential role in gene regulation [1–4]. At the DNA level, distal regulatory elements such as enhancers have been shown to be in spatial proximity to their target genes. At a larger scale, topologically associating domains (TADs) have been suggested to be the basic unit of mammalian genome organization [5, 6]. Several recent high-throughput technologies based on chromatin conformation capture (3C) [7] have emerged (such as Hi-C [8], ChIA-PET [9], Capture-C [10], Capture Hi-C [11], PLAC-Seq [12], and HiChIP [13]) and have provided an unprecedented opportunity to study the genome spatial organization in a genome-wide fashion.

As the volume of chromatin interaction data keeps increasing, efficient visualization and navigation of these data become a major bottleneck for their biological interpretation. Due to the size and complexity of these interactome data, it is challenging for an individual lab to store and explore them efficiently. To tackle this challenge, several visualization tools have been developed, and each of them has its unique features and limitations. The Hi-C Data Browser [8] was the first web-based query tool that visualizes Hi-C data as heatmaps. Currently, it does not support zoom functionalities and only hosts limited number of datasets. The WashU Epigenome Browser [14, 15] can display both Hi-C and ChIA-PET data, and it also provides access to thousands of epigenomic datasets from the ENCODE and Roadmap Epigenome projects. Due to the large file size of Hi-C matrices, which could reach hundreds of gigabytes, its speed for uploading and exploring Hi-C data is still not optimal. Furthermore, it does not offer an option to display inter-chromosomal interaction data as heatmaps. Users can also explore Hi-C data in Juicebox [16] and Hi-Glass [17] with great speed, but currently, neither of them provide other types of chromatin interaction data, such as Capture Hi-C or ChIA-PET. Delta browser [18] is another visualization tool with many features and can display both physical view of 3D genome modeling and Hi-C data. However, all the aforementioned tools except for the WashU Epigenome Browser only display Hi-C as a heatmap, which is convenient for visualizing large

* Correspondence: twang@genetics.wustl.edu; fyue@hmc.psu.edu
[4]Department of Genetics, The Edison Family Center for Genome Sciences and Systems Biology, Washington University School of Medicine, St. Louis, MO 63108, USA
[1]Bioinformatics and Genomics Program, The Pennsylvania State University, University Park, State College, PA 16802, USA
Full list of author information is available at the end of the article

domain structures such as TADs, but may not be the most informative way for visualizing enhancer-promoter interactions.

Here, we present the 3D Genome Browser (www.3dgenome.org), which is a fast web-based browser that allows users to smoothly explore both published and their own chromatin interaction data. Our 3D Genome Browser features six distinct modes that allow users to explore interactome data tailored toward their own needs, from exploring organization of higher-order chromatin structures at domain level to investigating high-resolution enhancer-promoter interactions. Our browser provides convenient zoom and traverse functions in real time and supports queries by gene name, genomic loci, or SNP rsid. In addition, users can easily incorporate their UCSC Genome Browser and the WashU Epigenome Browser sessions and therefore can simultaneously query and supplement chromatin interaction data with thousands of genetic, epigenetic, and phenotypic datasets, including ChIP-Seq and RNA-Seq data from the ENCODE and Roadmap Epigenomics projects. So far, it has been visited by more than 60,000 unique users from 120 countries surpassing over 600,000 page views. In summary, the 3D Genome Browser represents an invaluable resource and ecosystem for the study of chromosomal organization and gene regulation in mammalian genomes.

Results and Discussion

Overall design and implementation of the system

The overall structure of the 3D Genome Browser is summarized in Fig. 1. Currently, our browser hosts more than 300 chromatin interaction datasets of a variety of different types (Table 1), including Hi-C, ChIA-PET, Capture Hi-C, PLAC-Seq, HiChIP, GAM [19], and SPRITE [20], in both human and mouse across multiple genome assemblies, making it one of the most comprehensive and up-to-date high-quality chromatin

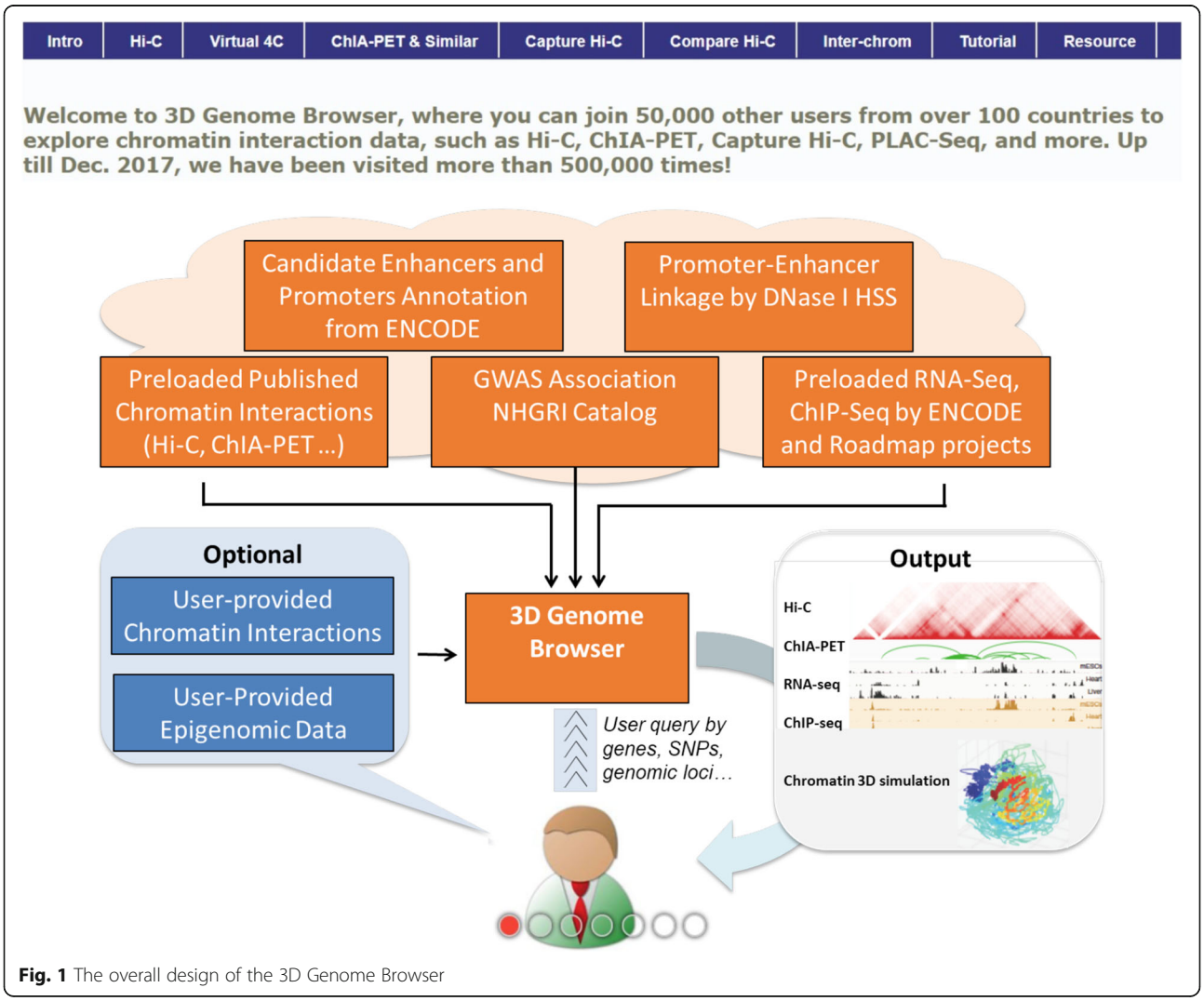

Fig. 1 The overall design of the 3D Genome Browser

Table 1 Summary of number of datasets available on the 3D Genome Browser

Data type	Samples and conditions	Total datasets
Hi-C	70	288
Virtual 4C, derived from Hi-C	Same as above	Same as above
ChIA-pet	14	14
Capture Hi-C	19	19
HiChIP	2	2
PLAC-Seq	3	3
GAM	1	1
DNase Hi-C	2	2
SPRITE	2	2
Total number	113	331

interaction data collection (details in Table S1, S2, S3). To increase their impacts and usability, we systematically re-mapped and generated interaction matrices for over 100 Hi-C datasets to the most current genome assembly (GRCh38 and mm10), using the same in-house data processing pipeline.

One of the important discoveries based on Hi-C data analysis is that the mammalian genomes are organized in mega-base pair chromatin domains, termed topologically associating domains (TADs). Therefore, we adopted the same pipeline from Dixon et al. [5] and systematically predicted TADs in all cell/tissue types (Fig. 2a, orange/blue bars) in our browser. Hi-C data has been shown to contain systematic noises [21]; therefore, we performed ICE (iterative correction and eigenvector decomposition) normalization to all the Hi-C datasets in our browser as well. To further assist users to explore 3D genome organization and gene regulation events simultaneously, we also collected the open chromatin data from the same cell type and display them in the same window (Fig. 2a, red bars). Finally, when users query the chromatin interaction information for a gene, we can also display the expression profiles of this gene across 109 cell/tissue types (Additional file 1: Figure S1), which was uniformly processed by the ENCODE consortium. In summary, for a given genomic loci, our browser can display TADs, chromatin interaction, RNA-Seq, and open chromatin region simultaneously and therefore give our users a comprehensive view of these regions.

To facilitate a user's unique interest, our 3D Genome Browser features six distinct modes that allow users to explore interactome data, including (1) intra-chromosomal Hi-C contact matrices as heatmaps, coupled with TADs and available genome annotation in the same cell type; (2) inter-chromosomal Hi-C heatmaps: this mode is particularly helpful for visualizing inter-chromosomal interactions and translocations; (3) compare Hi-C matrices: stacked Hi-C heatmaps from different tissues or even different species; (4) virtual 4C: Hi-C data is plotted as an arc for a queried gene or loci (bait), where the center is the bait region. This mode is particularly helpful for revealing chromatin interactions between two individual loci; (5) ChIA-PET or other ChIP-based chromatin interaction data such as PLAC-Seq and HiChIP; (6) Capture Hi-C or other capture-based chromatin interaction data. Below, we will use several examples to demonstrate these options and also illustrate how the 3D Genome Browser can be used to make novel biological discoveries.

Exploring chromatin interactions using Hi-C data

First, we demonstrate an example of exploring Hi-C data with the 3D Genome Browser for a large genomic region in Fig. 2a. It only takes ~ 5 s to show a 10-Mb region of GM12878 Hi-C interaction map on chr12 (~ 15–25 Mb) at a 25-kb resolution. The alternating yellow and blue bars are predicted TADs using the same in-house pipeline as in Dixon et al. [5]. The dark red vertical bars are DNase I hypersensitive sites (DHS) in the same cell type. Users can also adjust the color scale to reduce the background signals and make the TAD structure more visible.

Identifying cell/tissue-specific chromatin interactions is important, as it has been shown that chromatin structure plays an important role in determining cellular identity [22, 23]. In Fig. 2b, we notice a chromatin interaction in the 5-kb resolution Hi-C contact map in K562 cell line [24] (marked by the black arrow). To interpret biological meaning of this chromatin interaction, we integrated the WashU Epigenome Browser with gene annotation; histone modification H3K4me1, H3K4me3, and H3K27ac; and chromHMM [25] in K562 cells. We found that the two interacting loci are the promoter of *SLC25A37* and a putative enhancer predicted by histone modification patterns and chromHMM (Fig. 2b, vertical gray bar). This putative enhancer has been confirmed to exhibit enhancer activities that regulate *SLC25A37* expression during late-phase erythropoiesis [26]. Further, we checked the expression patterns profiled by the ENCODE consortium for *SLC25A37* on our browser and it showed high tissue specificity to K562 cells (Additional file 1: Figure S1).

Discovering high-resolution promoter-enhancer interactions using Capture Hi-C and DHS-linkage

While Hi-C data provides a viable way to suggest promoter-enhancer pairing, most of the current published Hi-C maps are at 10–40-kb resolution and therefore are not optimal for uncovering enhancer-promoter interactions. Sequence capture- or pull-down-based methods, such as Capture Hi-C or ChIA-PET, generally have higher resolution and therefore are more effective in identifying chromatin interactions between gene and

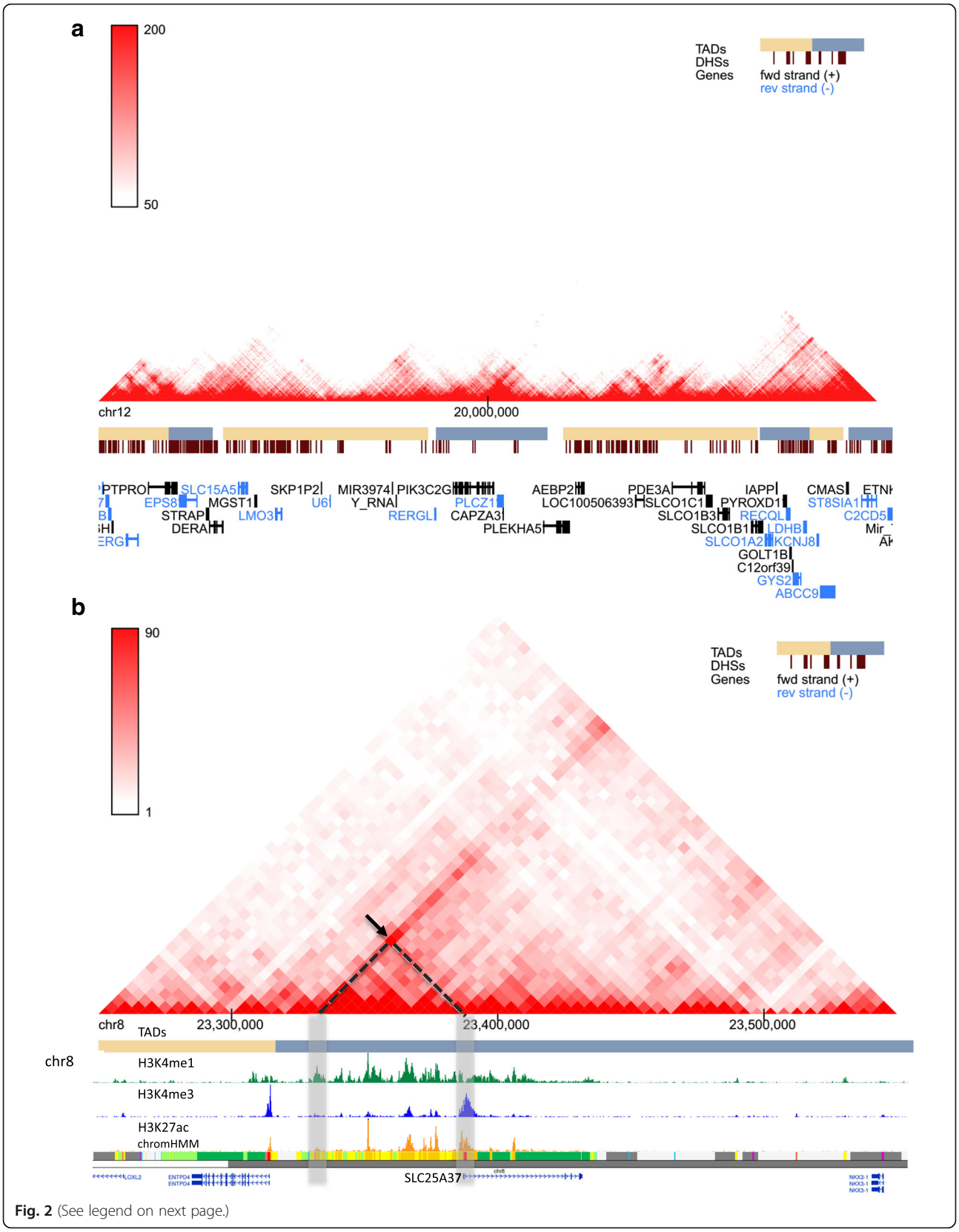

Fig. 2 (See legend on next page.)

(See figure on previous page.)
Fig. 2 Examples of using the 3D Genome Browser to explore Hi-C data. **a** A 10-Mb region of GM12878 Hi-C interaction map on chr12 (~ 15–25 Mb) at 25-kb resolution. The alternating yellow and blue bars are predicted TADs. The dark red bars are DHS in the same cell type. **b** Hi-C interaction map in K562 cells at 5-kb resolution. The black arrow points to a potential tissue-specific interaction between the *SLC25A37* promoter and a candidate enhancer region (marked by H3K4me1). The ChIP-Seq tracks for histone modifications, and chromHMM are visualized using the WashU Epigenome Browser

their *cis*-regulatory elements. In Fig. 3a, we give an example of Capture Hi-C [27], which seeks long-range interactions that involve selected elements of interests captured with pre-determined sequences (in this case, promoters). Capture Hi-C identified chromatin loops are presented as the green arcs (top track in Fig. 3a). The center of the track is the capture sequence in this region, which is the *PAX-5* gene promoter. We observed that the promoter interacts highly with the nearby regions and most of the interacting regions are enriched for strong enhancer marks (H3K4me1 and H3K27ac).

To further examine the predicted promoter-enhancer linkages, we also explored the linkage data by DNase I hypersensitive sites (DHS) in this region (blue curve line, second track in Fig. 3a), which represents another method of linking distal regulatory element with their target genes. It works by computing Pearson correlation coefficients between the gene proximal and distal DHS pairs across more than 100 ENCODE cell types, and only the pairs with PCC > 0.7 and within 500 kb are kept as the linked pairs [28]. In the example shown in Fig. 3a, we observed several interactions involving the promoter of the *PAX-5* gene and a potential enhancer (marked by both H3K4me1 and H3K27ac signals) downstream of the *ZCCHC7* gene in the naïve B cell Capture Hi-C dataset [27]. One region marked by enhancer-associated histone modifications has indeed been previously determined to be an enhancer for *PAX5*, and its disruption leads to leukemogenesis [29]. By integrating multiple lines of evidence, our browser provides a valuable resource for investigators to generate hypotheses connecting distal non-coding regulatory elements and their target genes.

Investigating potential target genes for non-coding genetic variants

Resolutions at loci-specific levels also hold significance in the discovery of the functions of non-coding genetic variants, such as single nucleotide polymorphisms (SNPs), which may disrupt transcription factor (TF) binding sites of *cis*-regulatory elements. In this section, we will first demonstrate how to use virtual 4C mode for such analyses. The 4C (circular chromosomal conformation capture [30, 31]) experiment is a chromatin ligation-based method that measures *one*-versus-*many* interactions in the genome, that is, the interaction frequencies between a "bait" locus and any other loci. Its data is plotted as a line histogram, where the center is the "bait" region and any peak signals in distal regions indicate the frequency of chromatin interaction events. In our browser, we use the queried region (gene name or SNP) as the bait and extract Hi-C data centered on the bait region, hence, *virtual* 4C. To bolster the power of the virtual 4C plot, our browser also supplements ChIA-PET and DHS-linkage data. In Fig. 3b, we queried the SNP rs12740374 in the virtual 4C mode. This SNP has been associated with high plasma low-density lipoprotein cholesterol (LDL-C) [32], which could lead to coronary artery disease and myocardial infarction. We plotted virtual 4C and ChIA-PET data from K562 in this region, as high-resolution Hi-C and ChIA-PET data are only available for K562, but not for hepatic cell lines. Since LDLs are processed by the liver, we examined the histone modifications in the Hep2G cell line and found rs12740374 is located within a candidate enhancer region as marked by H3K27ac signals. Hence, virtual 4C, ChIA-PET, and DHS-linkage all support a putative interaction between the enhancer harboring this SNP and the promoter region of *SORT1*. Further, it has been shown that the rs12740374 minor allele creates a C/EBPα-binding site which enhances *SORT1* expression leading to decreased LDL-C levels, thus suggesting that the minor allele confers a gain-of-function effect [33]. Still, despite the unusual conclusions reached by the study—as most minor alleles are usually loss-of-function—the virtual 4C mode of our 3D Genome Browser could aid in the hypothesis generation of not only the *cis*-regulatory elements and their putative target genes but also the effects of non-coding variants.

Exploring conservation of chromatin structure across species

Studying the evolutionary conservation of TADs could lead to a deeper understanding of their functional significance. The compare Hi-C mode of the 3D Genome Browser facilitates this endeavor by stacking Hi-C heatmaps from homologous regions of different species for visual contrast. In this mode, we observed the conservation of TADs and the genes near or at the TAD boundaries between human and mouse in their homologous region surrounding the *BCL-6*/Bcl-6 genes (Fig. 4), suggesting the chromatin structure may play a conserved role in the regulation of this proto-oncogene. This mode could be helpful for users to observe conserved or dynamic Hi-C interactions from different tissue/cell types.

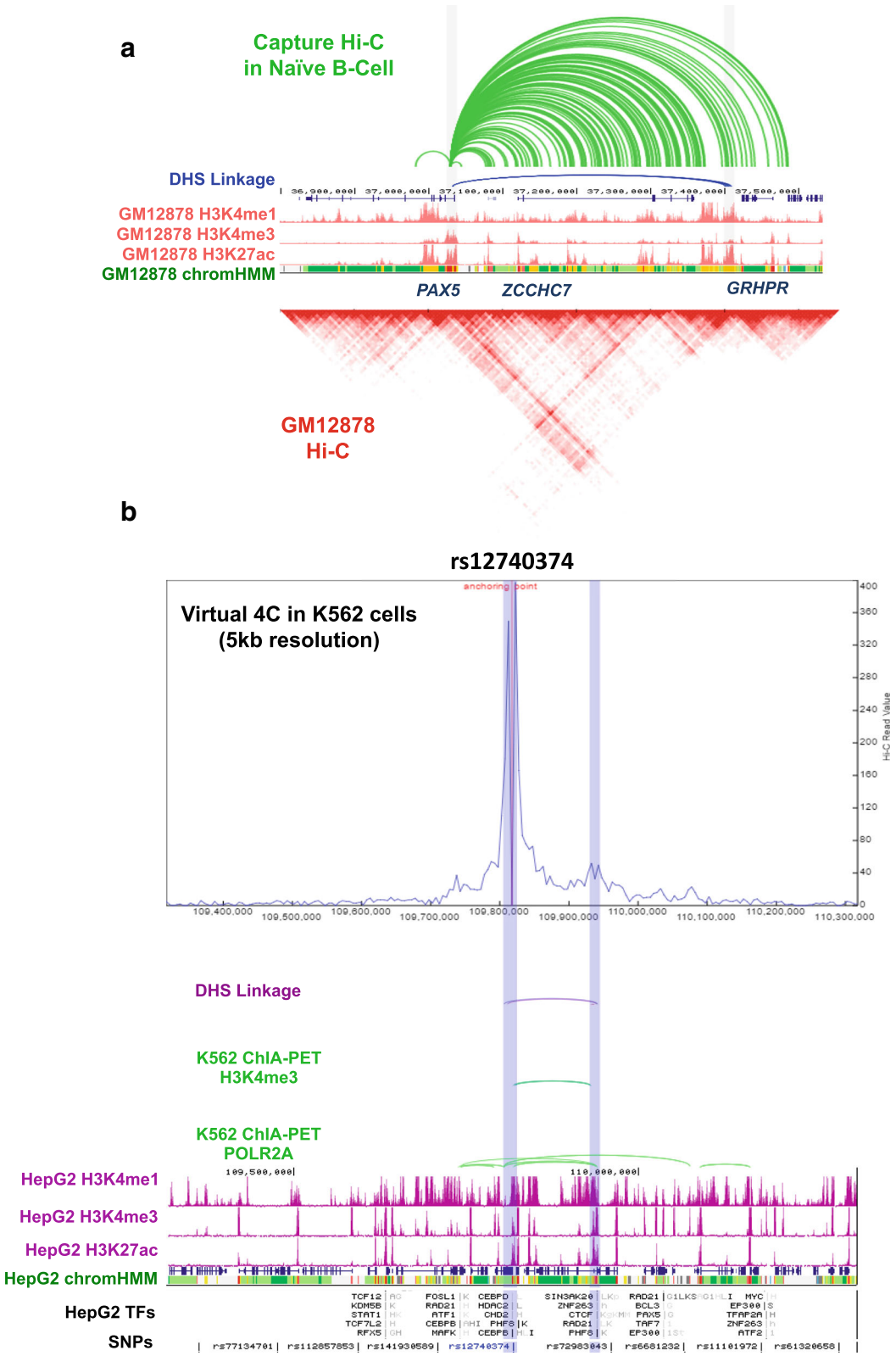

Fig. 3 (See legend on next page.)

(See figure on previous page.)
Fig. 3 Linking distal regulatory elements and SNPs with their target genes with the 3D Genome browser. **a** Capture Hi-C data in naïve B cells showing potential interactions (green curve lines) with PAX5 promoter region. The Capture Hi-C interactions are consistent with patterns from the 5-kb resolution Hi-C data in GM12878 cells. **b** Using virtual 4C, DHS-linkage, and ChIA-PET data to hypothesize the target gene for non-coding variant rs12740374. Based on the annotation by chromHMM in HepG2, this SNP is located at a putative enhancer region (orange). According to virtual 4C data, there is a potential interaction between this enhancer and the SORT1 promoter. This linkage is also supported by DHS-linkage, as well as by the H3K4me3 and POL2A ChIA-PET data in K562 cell line

Uncovering structural variations in cancer genomes

It has been shown recently that Hi-C data cannot only be used to detect chromatin interactions, but also may be used to denote structural variations [34–39]. Certain structural variations, such as deletions, insertions, inversions and translocations, establish signature patterns have been observed in Hi-C heatmaps. A striking structural variation is shown in Fig. 5 through the inter-chromosomal heatmap mode: we confirmed the oncogenic BCR-ABL inter-chromosomal translocations

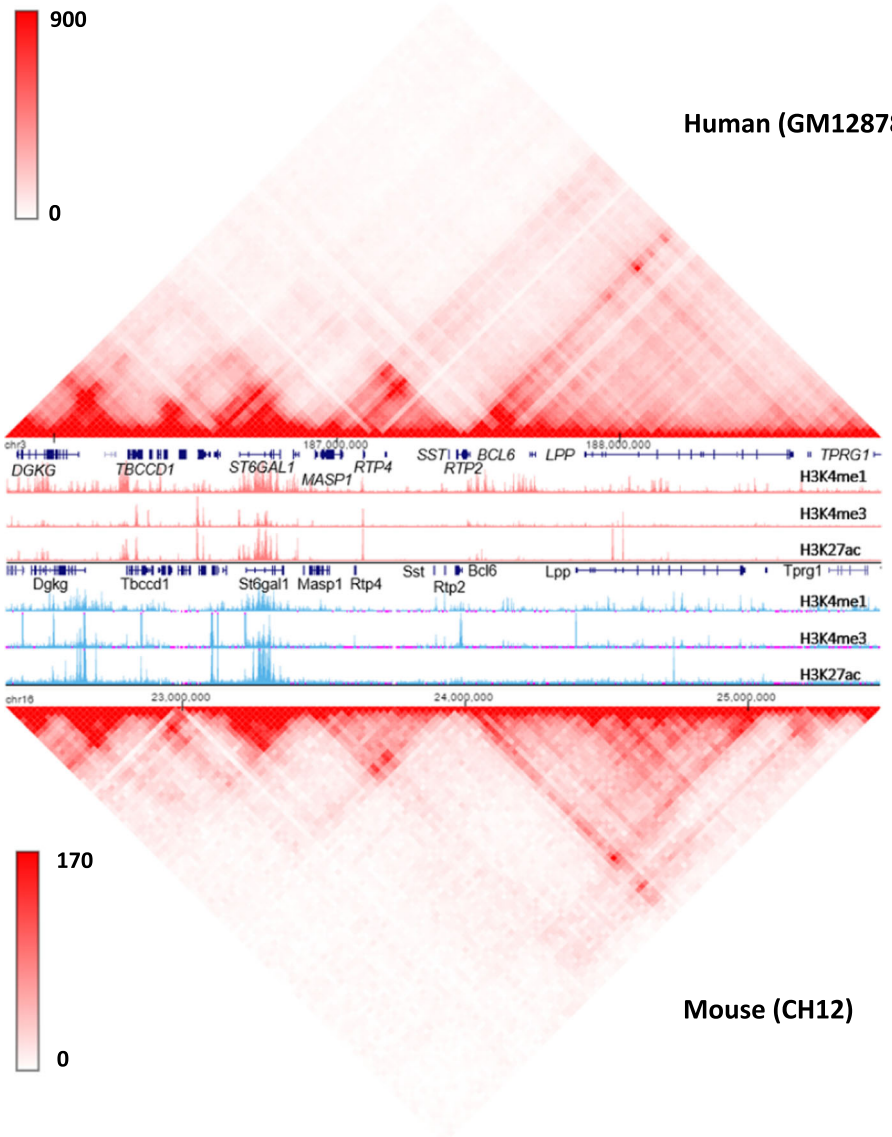

Fig. 4 Using the 3D Genome Browser to explore conserved chromatin structure across human and mouse. The similarity between human GM12878 Hi-C data and mouse CH12 Hi-C data at the region surrounding the *BCL6*/Bcl6 gene indicates an evolutionary conservation event of the chromatin structure between the two species

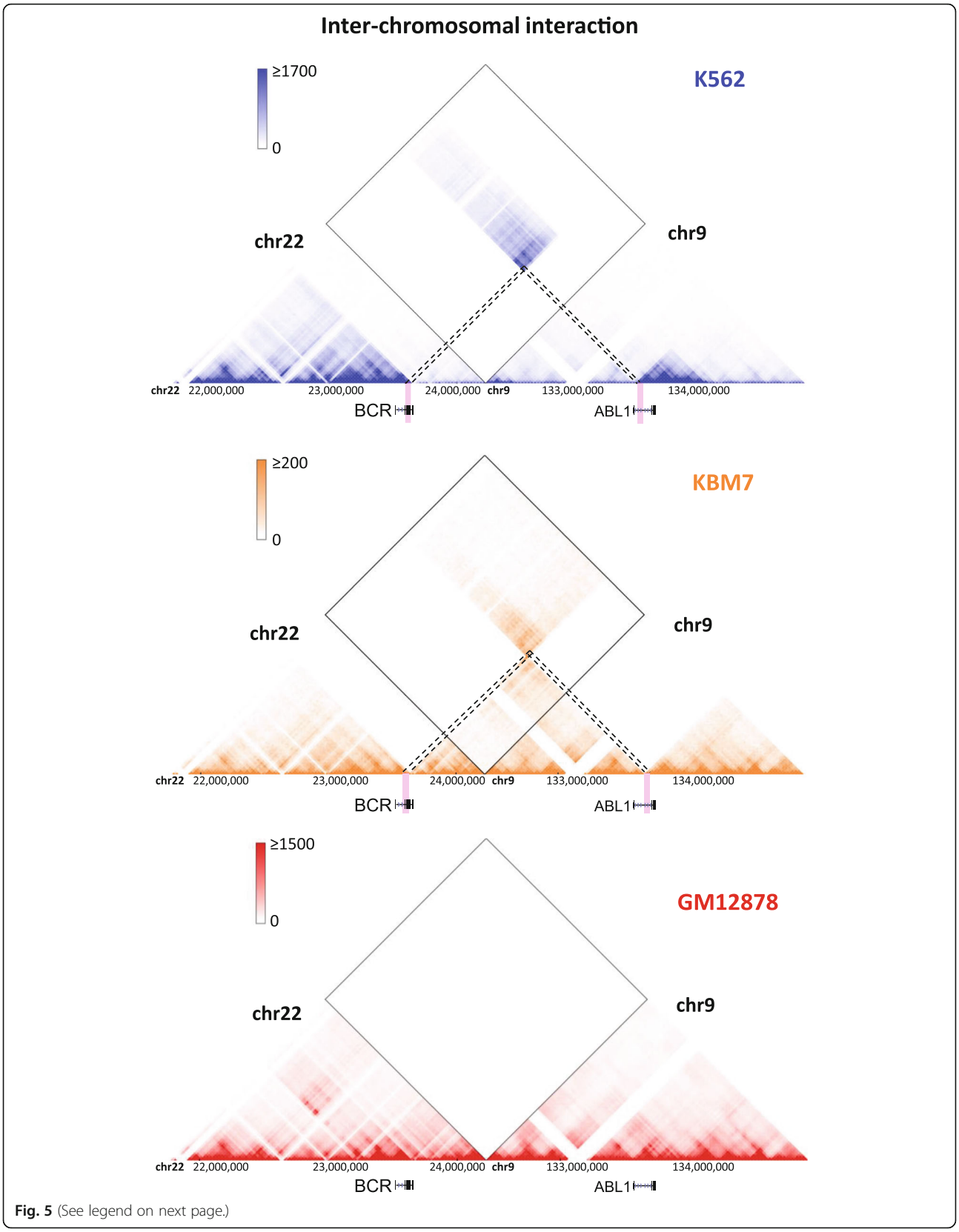

Fig. 5 (See legend on next page.)

(See figure on previous page.)
Fig. 5 Using the inter-chromosomal interaction mode of the 3D Genome Browser to discover structural variations in cancer cells. An inter-chromosomal translocation event (*BCR-ABL* fusion) in K562 and KBM7 CML cell lines appears as "inter-chromosomal interactions" on Hi-C maps. Such aberrant patterns are frequently observed in Hi-C maps in cancer cells, because the cancer genome is not available and Hi-C reads are mapped to the reference genome. We also noted that this translocation is reciprocal in KBM7 but not in K562 cells and that the breakpoint in ABL is different in the two cell lines. Such inter-chromosomal interactions are not observed in the karyotypically normal GM12878 cells

in two chronic myelogenous leukemia (CML) cell lines, K562 and KBM7. Such inter-chromosomal interactions are not observed in the karyotypically normal GM12878 cell line. We also noted that this translocation is reciprocal in KBM7 but not in K562 cells and that the breakpoint in ABL is different in the two cell lines. In addition, with the browser's compare Hi-C mode, the users could contrast the similarities and differences of chromosomal structure between distinct cells/tissues or even different species. Comparing the cancer cell line K562 to the normal cell line KBM7, we noted deletions specific to K562, one of which encompasses the tumor suppressor genes *CDKN2A* and *CDKN2B* (Additional file 1: Figure S2), as previously confirmed [40].

New binary Hi-C data format allows faster data retrieval and visualizing users' own Hi-C datasets

The 3D Browser supports a variety of features that allow users to browse unpublished data. First, our browser encourages integration with customized UCSC or WashU Epigenome browser sessions, wherein the users could add or modify existing tracks or upload their own genomic/epigenomic data. For example, to view a customized UCSC session, a user would only be required to enter the UCSC session URL. More importantly, the users could view their own Hi-C data by converting the contact matrices into a novel, indexed binary file format called *B*inary *U*pper *T*riangu*L*ar *M*at*R*ix (BUTLR file) developed by us. By hosting the BUTLR file on any HTTP-supported server and providing the URL to the 3D Genome Browser, a user can take full advantage of the features of our browser, without having to upload their Hi-C data since the browser would only query the selected region through binary indexing, rather than searching through the entire matrix. This capability is similar to the bigWig/bigBed mechanism invented by us and UCSC [41].

Additionally, BUTLR format dramatically reduces the file size of high-resolution Hi-C data not only through the binarization but also through the omission of redundant values (Additional file 1: Figure S3a; Additional file 2). The BUTLR file encodes an entire genome-wide chromatin interactions data into a binary, indexed format. While 1-kb resolution hg19 intra-chromosomal Hi-C contact matrices in the tab-delimited format require almost 1 TB, the BUTLR format of those same matrices would only take 11 GB (Additional file 1: Figure S3b). More importantly, the binary file format also greatly improves the query speed: using pre-loaded Hi-C datasets, the 3D browser generally returns the query results as a heatmap in a matter of seconds. We also want to note that our browser is designed as query-based to maximize its usability, and as a result, it excels at exploring locus of interest and gene-element relationship, but can be a little less dynamic than other tools when navigating Hi-C matrix for larger genomic regions.

Conclusion

In summary, we developed an interactive 3D Genome Browser that is defined by simple and easy-to-navigate graphical user interface, fast query-response time, and a comprehensive collection of publicly available chromatin interaction datasets. As our browser simultaneously displays the 3D chromatin interactions, functional (epi)genomic annotations, and disease/trait-associated SNPs, we provide an invaluable online tool for investigators from all over the world for the study of 3D genome organization and its functional implications in mammalian gene regulation.

Methods

Backend and user interface

The 3D Genome Browser is supported by the LAMP (Linux, Apache, MySQL, PHP) stack web service on the backend. At the user-interface level, the browser depends on HTML5 and JavaScript and its libraries JQuery and D3.js. All displays are rendered on HTML5 Canvas or Inline SVG.

In-house Hi-C data processing pipeline

We followed the pipeline in Dixon et al. [22] for Hi-C data processing. Briefly, raw fastq files were aligned to human reference genome GRCh38 with BWA aligner (0.7.15-r1140). Only uniquely mapped reads and properly paired reads on the same chromosome are retained. The genome is binned at different resolution (e.g., 40 kb and 10 kb) to generate Hi-C matrix. Paired reads were considered to be chromatin interactions connecting two bins. ICE (iterative correction and eigenvector decomposition) normalization was done using the "iced" Python package.

User query submission

The user may provide genomic coordinates or genome features such as gene symbols, RefSeq ID, Ensembl ID, or SNP rsid as queries for all modes of the 3D Genome Browser.

External genome browser integration and alignment

For the UCSC Genome Browser, we embed its sessions with the iframe and we align its content with our tracks by manipulating the scroll bars of the div HTML element containing the iframe. The WashU Epigenome Browser provides a JavaScript function for seamless integration into our browser. For both external browsers, it is possible for the user to embed a user-defined session consisting of user-selected tracks and options by providing the session URL to the 3D Genome Browser.

Determining homologous regions

For the compare Hi-C mode, we determine the homologous regions between two species by querying for homologous genes from the NCBI's HomoloGene database [42] as well as utilizing known inter-species chains [43].

BUTLR format

The BUTLR file encodes an entire genome-wide chromatin interactions data into a binary, indexed format. To compress the original contact matrices, BUTLR only stores the nonzero values of the upper triangular matrices of the intra-chromosomal data and the $n \times m$, where n and m are the number of interrogated loci and where $n < m$ of the inter-chromosomal data. The locations of each chromosome or chromosome-pair matrix, row indices of each matrix, and column indices of nonzero values along with nonzero values are binarized and indexed within the BUTLR file structure. Perl scripts that encode and decode BUTLR files are available at http://github.com/yuelab/BUTLRTools. All the Hi-C matrices in this manuscript are converted to BUTLR file format for visualization [5, 8, 19, 20, 22–24, 44-54].

Additional files

Additional file 1: Figure S1. Gene expression of *SCL25A37* across 109 issues. **Figure S2.** Using the 3D Genome Browser to determine intra-chromosomal structural variations. **Figure S3.** Design and performance of the BUTLR file format. **Table S1.** List of Hi-C datasets hosted by the 3D Genome Browser. **Table S2.** List of ChIA-PET, Capture Hi-C, PLAC-Seq and HiChIP datasets. **Table S3.** List of GAM, DNase Hi-C, and SPRITE datasets. (PDF 1295 kb)

Additional file 2: Review history. (DOCX 1354 kb)

Acknowledgements

We thank Dr. Jesse R. Dixon for the help with TADs and compartment calling. We are grateful to the members of Wang lab and Yue lab for useful discussions.

Funding

This work was supported by NIH grants R35GM124820, R01HG009906, and U01CA200060 (F.Y.). F.Y. is also supported by Leukemia Research Foundation, PhRMA Foundation, and Penn State CTSI. T.W. is also supported by NIH grants R01HG007175, R01HG007354, R01ES024992, U24ES026699, and U01HG009391. M.H. is partially supported by NIH U54DK107977. Y.L. is partially supported by NIH R01HG006292 and R01HL129132.

Availability of data and materials

All data are available at http://3dgenome.org. The source code of the website is deposited at https://github.com/yuelab/3dgenome [55] and Zenodo (DOI: 10.5281/zenodo.1402785) [56]. The code for the 3D Genome Browser is freely available under an MIT license.
No new experimental datasets were generated within this study. Publicly available datasets included in the browser are listed in the supplementary tables.

Review history

The review history for this manuscript is available as Additional file 2.

Authors' contributions

YW, TW, and FY conceived the project. YW, FS, and FY designed and implemented the project. YW, TW, and FY wrote the manuscript with input from MC, YL, MH, and RCH. BZ, LZ, JX, DK, DL, and MC helped with the data processing and integration. All authors read and approved the final manuscript.

Ethics approval

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Author details

[1]Bioinformatics and Genomics Program, The Pennsylvania State University, University Park, State College, PA 16802, USA. [2]Department of Biochemistry and Molecular Biology, College of Medicine, The Pennsylvania State Hershey, Hershey, PA 17033, USA. [3]Department of Computer and Information Science, University of Pennsylvania, Philadelphia, PA 19104, USA. [4]Department of Genetics, The Edison Family Center for Genome Sciences and Systems Biology, Washington University School of Medicine, St. Louis, MO 63108, USA. [5]Department of Genetics, University of North Carolina, Chapel Hill, NC 27599, USA. [6]Department of Biostatistics, University of North Carolina, Chapel Hill, NC 27599, USA. [7]Department of Computer Science, University of North Carolina, Chapel Hill, NC 27599, USA. [8]Department of Quantitative Health Sciences, Lerner Research Institute, Cleveland Clinic Foundation, Cleveland, OH 44195, USA. [9]Center for Computational Biology and Bioinformatics, Huck Institutes of the Life Sciences, The Pennsylvania State University, University Park, State College, PA 16802, USA.

Received: 23 December 2017 Accepted: 29 August 2018
Published online: 04 October 2018

References

1. Dekker J, Marti-Renom MA, Mirny LA. Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. Nat Rev Genet. 2013;14:390–403. https://doi.org/10.1038/nrg3454
2. Sexton T, Cavalli G. The role of chromosome domains in shaping the functional genome. Cell. 2015;160:1049–59. https://doi.org/10.1016/j.cell.2015.02.040
3. Bickmore WA. The spatial organization of the human genome. Annu Rev Genomics Hum Genet. 2013;14:67–84. https://doi.org/10.1146/annurev-genom-091212-153515
4. Gorkin DU, Leung D, Ren B. The 3D genome in transcriptional regulation and pluripotency. Cell Stem Cell. 2014;14:762–75. https://doi.org/10.1016/j.stem.2014.05.017
5. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B. Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature. 2012;485:376–80. https://doi.org/10.1038/nature11082
6. Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N, Piolot T, van Berkum NL, Meisig J, Sedat J, et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature. 2012;485:381–5. https://doi.org/10.1038/nature11049

7. Dekker J, Rippe K, Dekker M, Kleckner N. Capturing chromosome conformation. Science. 2002;295:1306–11. https://doi.org/10.1126/science.1067799
8. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science. 2009;326:289–93. https://doi.org/10.1126/science.1181369
9. Li G, Fullwood MJ, Xu H, Mulawadi FH, Velkov S, Vega V, Ariyaratne PN, Mohamed YB, Ooi HS, Tennakoon C, et al. ChIA-PET tool for comprehensive chromatin interaction analysis with paired-end tag sequencing. Genome Biol. 2010;11:R22. https://doi.org/10.1186/gb-2010-11-2-r22
10. Hughes JR, Roberts N, McGowan S, Hay D, Giannoulatou E, Lynch M, De Gobbi M, Taylor S, Gibbons R, Higgs DR. Analysis of hundreds of cis-regulatory landscapes at high resolution in a single, high-throughput experiment. Nat Genet. 2014;46:205–12. https://doi.org/10.1038/ng.2871
11. Mifsud B, Tavares-Cadete F, Young AN, Sugar R, Schoenfelder S, Ferreira L, Wingett SW, Andrews S, Grey W, Ewels PA, et al. Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C. Nat Genet. 2015;47:598–606. https://doi.org/10.1038/ng.3286
12. Fang R, Yu M, Li G, Chee S, Liu T, Schmitt AD, Ren B. Mapping of long-range chromatin interactions by proximity ligation-assisted ChIP-seq. Cell Res. 2016;26:1345–8. https://doi.org/10.1038/cr.2016.137
13. Mumbach MR, Rubin AJ, Flynn RA, Dai C, Khavari PA, Greenleaf WJ, Chang HY. HiChIP: efficient and sensitive analysis of protein-directed genome architecture. Nat Methods. 2016;13:919–22. https://doi.org/10.1038/nmeth.3999
14. Zhou X, Maricque B, Xie M, Li D, Sundaram V, Martin EA, Koebbe BC, Nielsen C, Hirst M, Farnham P, et al. The human epigenome browser at Washington University. Nat Methods. 2011;8:989–90. https://doi.org/10.1038/nmeth.1772
15. Zhou X, Lowdon RF, Li D, Lawson HA, Madden PA, Costello JF, Wang T. Exploring long-range genome interactions using the WashU Epigenome Browser. Nat Methods. 2013;10:375–6. https://doi.org/10.1038/nmeth.2440
16. Durand NC, Robinson JT, Shamim MS, Machol I, Mesirov JP, Lander ES, Aiden EL. Juicebox provides a visualization system for Hi-C contact maps with unlimited zoom. Cell Syst. 2016;3:99–101. https://doi.org/10.1016/j.cels.2015.07.012
17. Kerpedjiev P, Abdennur N, Lekschas F, McCallum C, Dinkla K, Strobelt H, Luber JM, Ouellette SB, Ahzir A, Kumar N, et al. HiGlass: web-based visual comparison and exploration of genome interaction maps. bioRxiv. 2017; https://doi.org/10.1101/121889
18. Tang B, Li F, Li J, Zhao W, Zhang Z. Delta: a new web-based 3D genome visualization and analysis platform. Bioinformatics. 2018;34:1409–10. https://doi.org/10.1093/bioinformatics/btx805
19. Beagrie RA, Scialdone A, Schueler M, Kraemer DC, Chotalia M, Xie SQ, Barbieri M, de Santiago I, Lavitas LM, Branco MR, et al. Complex multi-enhancer contacts captured by genome architecture mapping. Nature. 2017;543:519–24. https://doi.org/10.1038/nature21411
20. Quinodoz SA, Ollikainen N, Tabak B, Palla A, Schmidt JM, Detmar E, Lai MM, Shishkin AA, Bhat P, Takei Y, et al. Higher-order inter-chromosomal hubs shape 3D genome organization in the nucleus. Cell. 2018; https://doi.org/10.1016/j.cell.2018.05.024.
21. Imakaev M, Fudenberg G, McCord RP, Naumova N, Goloborodko A, Lajoie BR, Dekker J, Mirny LA. Iterative correction of Hi-C data reveals hallmarks of chromosome organization. Nat Methods. 2012;9:999–1003. https://doi.org/10.1038/nmeth.2148
22. Dixon JR, Jung I, Selvaraj S, Shen Y, Antosiewicz-Bourget JE, Lee AY, Ye Z, Kim A, Rajagopal N, Xie W, et al. Chromatin architecture reorganization during stem cell differentiation. Nature. 2015;518:331–6. https://doi.org/10.1038/nature14222
23. Schmitt AD, Hu M, Jung I, Xu Z, Qiu Y, Tan CL, Li Y, Lin S, Lin Y, Barr CL, Ren B. A compendium of chromatin contact maps reveals spatially active regions in the human genome. Cell Rep. 2016;17:2042–59. https://doi.org/10.1016/j.celrep.2016.10.061
24. Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES, Aiden EL. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell. 2014;159:1665–80. https://doi.org/10.1016/j.cell.2014.11.021
25. Ernst J, Kellis M. ChromHMM: automating chromatin-state discovery and characterization. Nat Methods. 2012;9:215–6. https://doi.org/10.1038/nmeth.1906
26. Huang J, Liu X, Li D, Shao Z, Cao H, Zhang Y, Trompouki E, Bowman TV, Zon LI, Yuan GC, et al. Dynamic control of enhancer repertoires drives lineage and stage-specific transcription during hematopoiesis. Dev Cell. 2016;36:9–23. https://doi.org/10.1016/j.devcel.2015.12.014
27. Javierre BM, Burren OS, Wilder SP, Kreuzhuber R, Hill SM, Sewitz S, Cairns J, Wingett SW, Varnai C, Thiecke MJ, et al. Lineage-specific genome architecture links enhancers and non-coding disease variants to target gene promoters. Cell. 2016;167:1369–84. e1319. https://doi.org/10.1016/j.cell.2016.09.037
28. Thurman RE, Rynes E, Humbert R, Vierstra J, Maurano MT, Haugen E, Sheffield NC, Stergachis AB, Wang H, Vernot B, et al. The accessible chromatin landscape of the human genome. Nature. 2012;489:75–82. https://doi.org/10.1038/nature11232
29. Puente XS, Bea S, Valdes-Mas R, Villamor N, Gutierrez-Abril J, Martin-Subero JI, Munar M, Rubio-Perez C, Jares P, Aymerich M, et al. Non-coding recurrent mutations in chronic lymphocytic leukaemia. Nature. 2015;526:519–24. https://doi.org/10.1038/nature14666
30. Simonis M, Klous P, Splinter E, Moshkin Y, Willemsen R, de Wit E, van Steensel B, de Laat W. Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). Nat Genet. 2006;38:1348–54. https://doi.org/10.1038/ng1896
31. Zhao Z, Tavoosidana G, Sjolinder M, Gondor A, Mariano P, Wang S, Kanduri C, Lezcano M, Sandhu KS, Singh U, et al. Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. Nat Genet. 2006;38:1341–7. https://doi.org/10.1038/ng1891
32. Sandhu MS, Waterworth DM, Debenham SL, Wheeler E, Papadakis K, Zhao JH, Song K, Yuan X, Johnson T, Ashford S, et al. LDL-cholesterol concentrations: a genome-wide association study. Lancet. 2008;371:483–91. https://doi.org/10.1016/S0140-6736(08)60208-1
33. Musunuru K, Strong A, Frank-Kamenetsky M, Lee NE, Ahfeldt T, Sachs KV, Li X, Li H, Kuperwasser N, Ruda VM, et al. From noncoding variant to phenotype via SORT1 at the 1p13 cholesterol locus. Nature. 2010;466:714–9. https://doi.org/10.1038/nature09266
34. Dixon J, Xu J, Dileep V, Zhan Y, Song F, Le VT, Yardimci GG, Chakraborty A, Bann DV, Wang Y, et al: An integrative framework for detecting structural variations in cancer genomes. *bioRxiv* 2017. https://doi.org/10.1101/119651.
35. Burton JN, Adey A, Patwardhan RP, Qiu R, Kitzman JO, Shendure J. Chromosome-scale scaffolding of de novo genome assemblies based on chromatin interactions. Nat Biotechnol. 2013;31:1119–25. https://doi.org/10.1038/nbt.2727
36. Engreitz JM, Agarwala V, Mirny LA. Three-dimensional genome architecture influences partner selection for chromosomal translocations in human disease. PLoS One. 2012;7:e44196. https://doi.org/10.1371/journal.pone.0044196
37. Naumova N, Imakaev M, Fudenberg G, Zhan Y, Lajoie BR, Mirny LA, Dekker J. Organization of the mitotic chromosome. Science. 2013;342:948–53. https://doi.org/10.1126/science.1236083
38. Chakraborty A, Ay F. Identification of copy number variations and translocations in cancer cells from Hi-C data. Bioinformatics. 2017; https://doi.org/10.1093/bioinformatics/btx664.
39. Seaman L, Chen H, Brown M, Wangsa D, Patterson G, Camps J, Omenn GS, Ried T, Rajapakse I. Nucleome analysis reveals structure-function relationships for colon cancer. Mol Cancer Res. 2017;15:821–30. https://doi.org/10.1158/1541-7786.MCR-16-0374
40. Sherborne AL, Hosking FJ, Prasad RB, Kumar R, Koehler R, Vijayakrishnan J, Papaemmanuil E, Bartram CR, Stanulla M, Schrappe M, et al. Variation in CDKN2A at 9p21.3 influences childhood acute lymphoblastic leukemia risk. Nat Genet. 2010;42:492–4. https://doi.org/10.1038/ng.585
41. Kent WJ, Zweig AS, Barber G, Hinrichs AS, Karolchik D. BigWig and BigBed: enabling browsing of large distributed datasets. Bioinformatics. 2010;26:2204–7. https://doi.org/10.1093/bioinformatics/btq351
42. Geer LY, Marchler-Bauer A, Geer RC, Han L, He J, He S, Liu C, Shi W, Bryant SH. The NCBI BioSystems database. Nucleic Acids Res. 2010;38:D492–6. https://doi.org/10.1093/nar/gkp858
43. Kent WJ, Baertsch R, Hinrichs A, Miller W, Haussler D. Evolution's cauldron: duplication, deletion, and rearrangement in the mouse and human genomes. Proc Natl Acad Sci U S A. 2003;100:11484–9. https://doi.org/10.1073/pnas.1932072100
44. Jin F, Li Y, Dixon JR, Selvaraj S, Ye Z, Lee AY, et al. A high-resolution map of the three-dimensional chromatin interactome in human cells. Nature. 2013;503:290–4. https://doi.org/10.1038/nature12644
45. Leung D, Jung I, Rajagopal N, Schmitt A, Selvaraj S, Lee AY, Yen CA, Lin S, Lin Y, Qiu Y, et al. Integrative analysis of haplotype-resolved epigenomes across human tissues. Nature. 518:350–4. https://doi.org/10.1038/nature14217

46. Won H, De La Torre-Ubieta L, Stein JL, Parikshak NN, Huang J, Opland CK, Gandal MJ, Sutton GJ, Hormozdiari F, Lu D, et al. Chromosome conformation elucidates regulatory relationships in developing human brain. Nature. 538:523–7. https://doi.org/10.1038/nature19847
47. Nagano T, Lubling Y, Várnai C, Dudley C, Leung W, Baran Y, Mendelson Cohen N, Wingett S, Fraser P, Tanay A. Cell-cycle dynamics of chromosomal organization at single-cell resolution. Nature. 547:61–7. https://doi.org/10.1186/s13059-015-0753-7
48. Rubin A, Barajas B, Furlan-Magaril M, Lopez-Pajares V, Mumbach M, Howard I, Chang H, Fraser P, Khavari P. Lineage-specific dynamic and pre-established enhancer-promoter contacts cooperate in terminal differentiation. Nat Genet. 2017;49(10):1522–8. https://doi.org/10.1038/ng.3935
49. Rao SSP, Huang SC, Glenn St Hilaire B, Engreitz JM, Perez EM, Kieffer-Kwon KR, Sanborn AL, Johnstone SE, Bascom GD, Bochkov ID, et al. Cohesin loss eliminates all loop domains. Cell. 171:305–20. e24. https://doi.org/10.1016/j.cell.2017.09.026
50. Phanstiel DH, Van Bortle K, Spacek D, Hess GT, Shamim MS, Machol I, Love MI, Aiden EL, Bassik MC, Snyder MP. Static and dynamic DNA loops form AP-1-bound activation hubs during macrophage development. Mol Cell. 2017;67(6):1037–48. https://doi.org/10.1016/j.molcel.2017.08.006
51. Jiang Y, Loh YE, Rajarajan P, Hirayama T, Liao W, Akbarian S, et al. The methyltransferase SETDB1 regulates a large neuron-specific topological chromatin domain. Nat Genet. 2017;49(8):1239–50. https://doi.org/10.1038/ng.3906
52. Bonev B, Mendelson Cohen N, Szabo Q, Fritsch L, Papadopoulos GL, Lubling Y, Xu X, Lv X, Hugnot JP, Tanay A, Cavalli G. Multiscale 3D genome rewiring during mouse neural development. Cell. 2017;171(3):557–72. e24. https://doi.org/10.1016/j.cell.2017.09.043
53. Doynova MD, Markworth JF, Cameron-Smith D, Vickers MH, O'Sullivan JM. Linkages between changes in the 3D organization of the genome and transcription during myotube differentiation in vitro. Skelet. Muscle. 7:1–14. https://doi.org/10.1186/s13395-017-0122-1
54. Ma W, Ay F, Lee C, Gulsoy G, Deng X, Cook S, et al. Fine-scale chromatin interaction maps reveal the cis-regulatory landscape of lincRNA genes in human cell. Nature Methods. 12(1):71–8. https://doi.org/10.1038/nmeth.3205
55. Wang Y, Song F, Zhang B, Zhang L, Xu J, et. al. The 3D Genome Browser: a web-based browser for visualizing 3D genome organization and long-range chromatin interactions. Github repository. https://github.com/yuelab/3dgenome. (2018).
56. Wang Y, Song F, Zhang B, Zhang L, Xu J, et al. The 3D Genome Browser: a web-based browser for visualizing 3D genome organization and long-range chromatin interactions. Zenodo: Source code; 2018. https://doi.org/10.5281/zenodo.1402785.