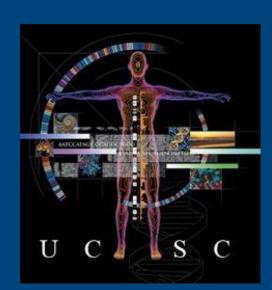
UCSC Genome Browser: HubSpace Track Storage



Christopher Lee, Jairo Navarro Gonzalez, Luis Nassar, Jonathan Casper, Max Haeussler

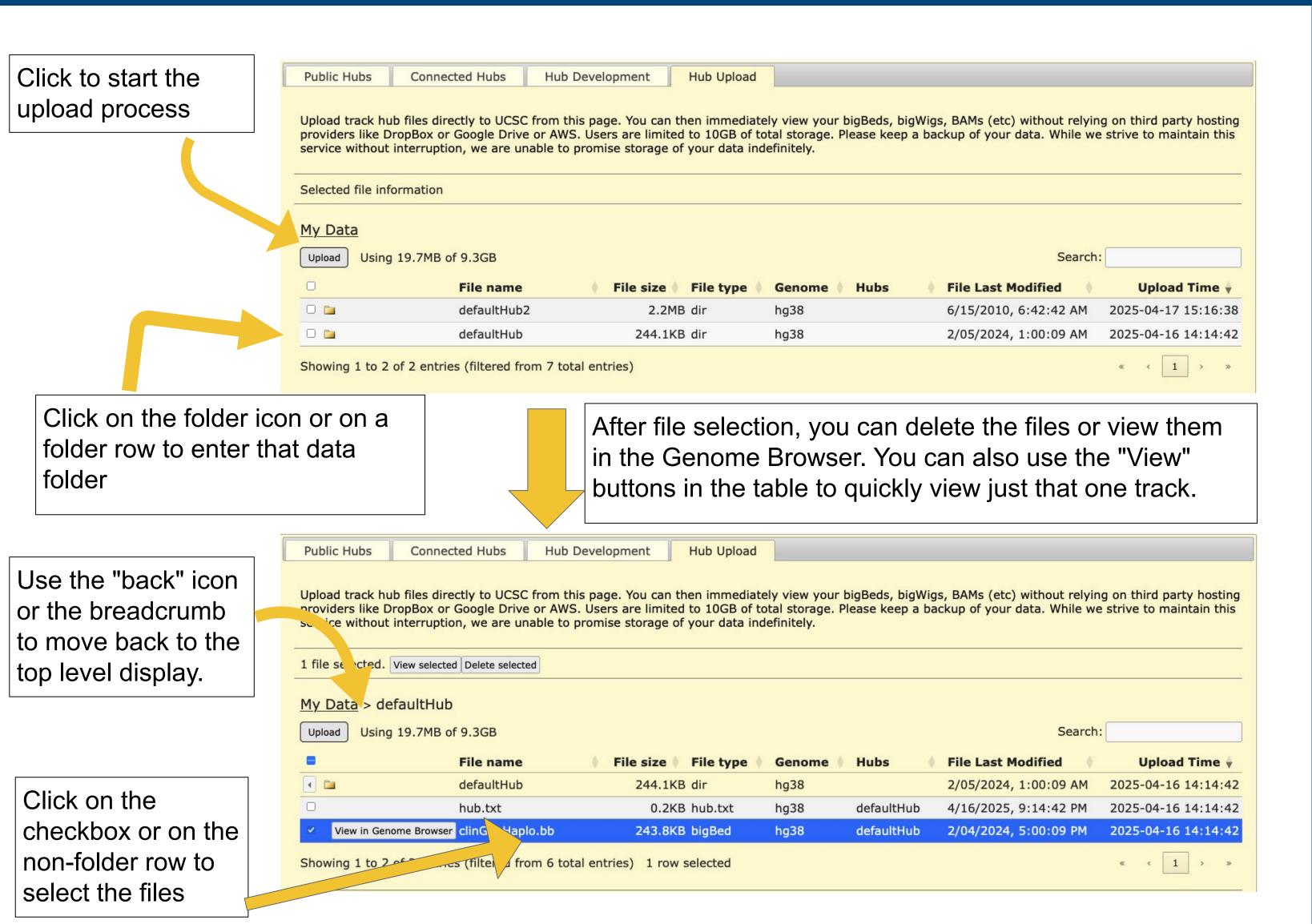
Directly upload track hub data without needing web storage

- Track hub data can now be directly uploaded to UCSC, no longer do you need an institutional web server or DropBox, Github, etc accounts
- Each user gets 10Gb of storage, with more available upon request
- Files can be uploaded on the track hub connect page, or via the command line program hubtools
- Files are stored directly on UCSC servers, optimizing both speed and robustness

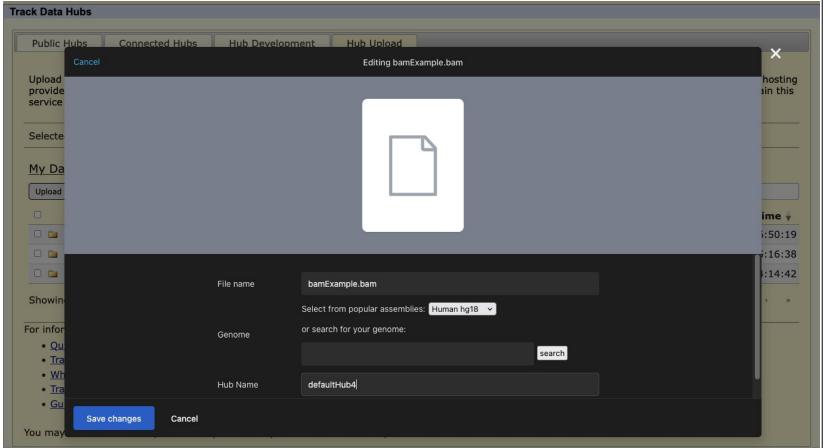
What are track hubs?

Track hubs are collections of remotely accessible data files like bigBed, bigWig, BAM, and VCF, that are described by a text file that defines how the track data is to be visualized in the Genome Browser. Traditionally, track hubs are hosted at your own institution or via Cloud Providers like DropBox and Github, and the Genome Browser only downloads the data needed to build the view for the particular genomic range being viewed. Over time, more visualization features exclusive to track hubs have been added, making them preferable to custom tracks. To facilitate ease of use, we are committed to providing dedicated storage for track hubs, up to 10Gb per user, perhaps even more in the future.

Using the HubSpace Interface

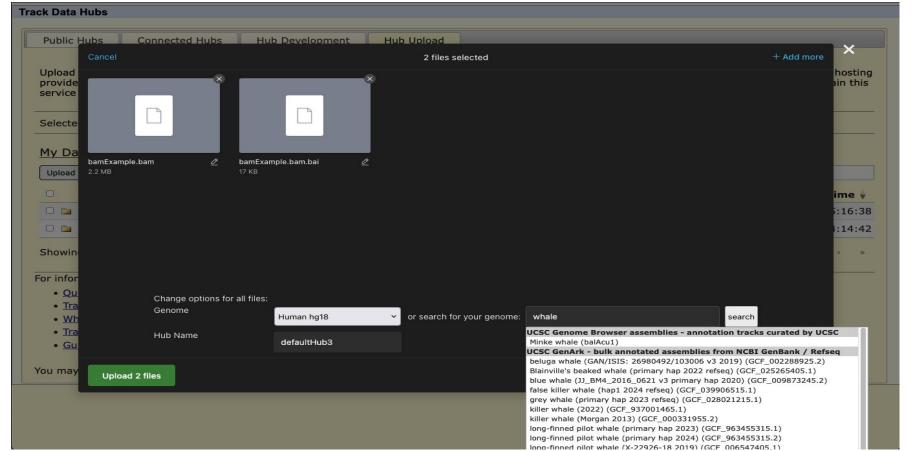


Example BAM file upload



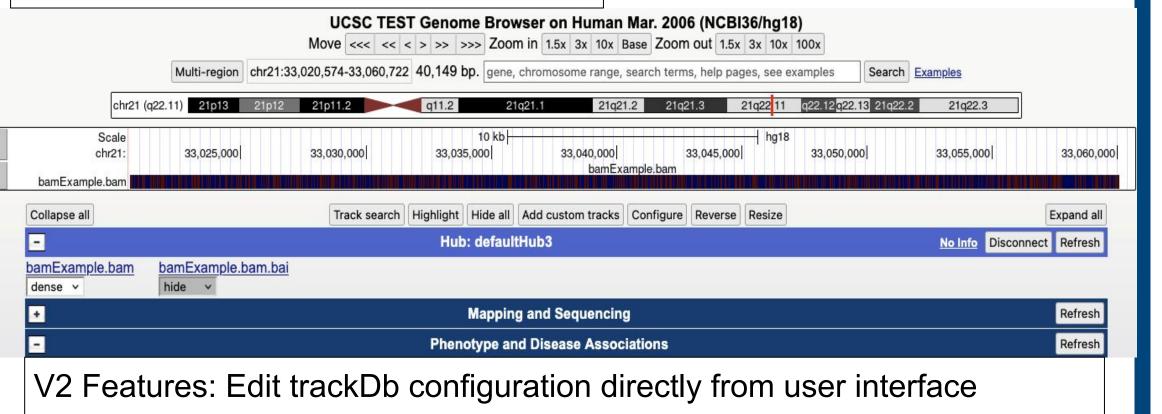
Track type and genome are automatically inferred, but can be manually configured with the dropdowns or by changing the file name extension. A track hub name is automatically generated for you.

When uploading multiple files at once, a batch editor is available



A track hub is automatically created for you with your uploaded tracks filled in. Click the "View in Genome Browser" link to visualize the data

Choose a popular genome from the dropdown list, or search from thousands of GenArk genomes available from NCBI/GenBank



References

The UCSC Genome Browser database: 2024 update Raney BJ, Barber GP, Benet-Pagès A, Casper J, Clawson H, Cline MS, Diekhans M, Fischer C, Navarro Gonzalez J, Hickey G et al. Nucleic Acids Res. 2024 Jan 5;52(D1):D1082-D1088.

DOI: 10.1093/nar/gkad987; PMID: 37953330;

PMC: PMC10767968.

GenArk: towards a million UCSC genome browsers.

Clawson H, Lee BT, Raney BJ, Barber GP, Casper J, Diekhans M, Fischer C, Gonzalez JN, Hinrichs AS, Lee CM et al.

Genome Biol. 2023 Oct 2;24(1):217.

DOI: <u>10.1186/s13059-023-03057-x</u>; PMID: <u>37784172</u>;

PMC: PMC10544498

tus: https://tus.io/ uppy: https://uppy.io/

Acknowledgements

This work was funded by NHGRI award 2U24HG002371 to the UCSC Genomics Institute. We would like to acknowledge the work of the UCSC Genome Browser technical staff (https://genome.ucsc.edu/staff.html), our many collaborators, and our users for their feedback and support.

More Information

Have questions?

Send us a question on our public mailing list: genome@soe.ucsc.edu

Genome Browser tutorials and videos: https://genome.ucsc.edu/training/index.html

Digital copy of this poster available at:

