



Integrating data sources for visualization in the Integrated Genome Browser



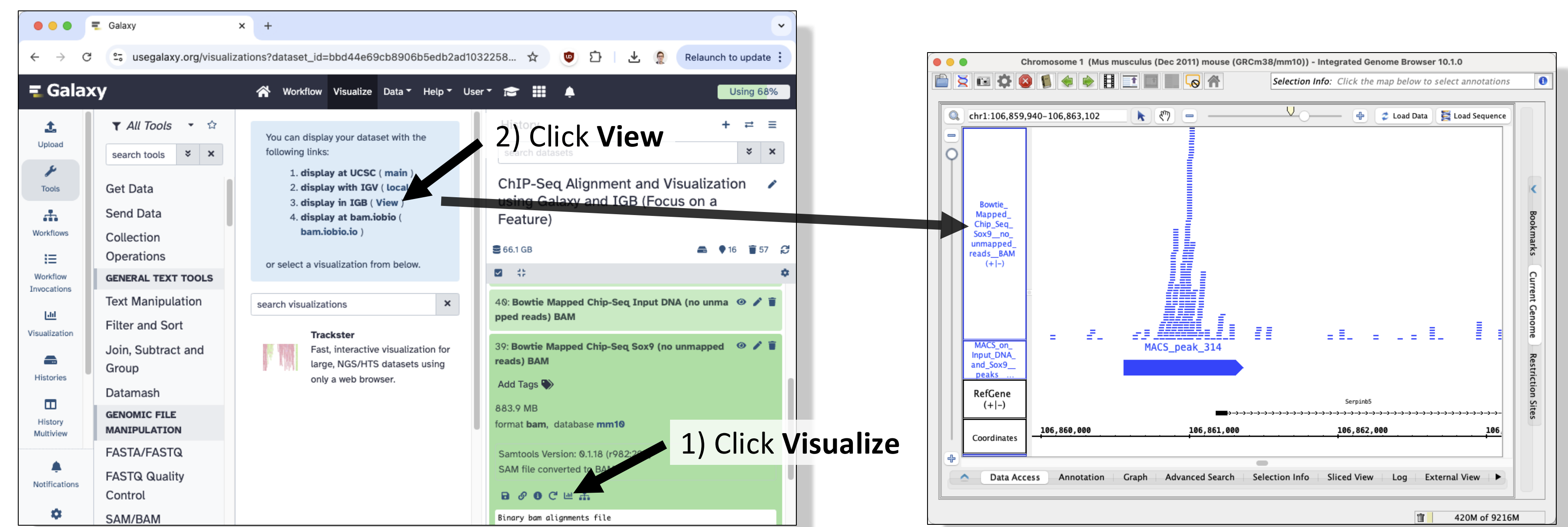
Nowlan H. Freese, Jaya Sravani Sirigineedi, Karthik Raveendran, Paige J. Kulzer, Ann E. Loraine

Abstract

We develop the Integrated Genome Browser (IGB), an open-source desktop genome browser that supports interactive visualization of genomic datasets such as aligned sequences (BAM/CRAM/SAM from "Seq" experiments), gene annotations (BED/GFF), variation (VCF/BCF), and coverage (wiggle/bedGraph). With the rapid output of new genomes, it has become increasingly difficult for our small team to keep up with curating and storing new genomes. Therefore, we have pushed to make IGB even more "integrated" by connecting to resources that provide genome sequences and annotations (UCSC, track hubs, Quickload), data storage and analytics capability (Galaxy, CyVerse), and interactive web-based "Seq" data exploration interfaces (BAR eFP-Seq). For example, by utilizing the UCSC Genome Browser REST API we can provide most of the assemblies and tracks from UCSC for visualization within IGB. As IGB runs locally, a user can now view data stored on their computer in IGB alongside UCSC tracks without the need to upload their data to the web. Future work will add support for viewing genomes from Ensembl via their API. Our goal is to make it as easy as possible for users to view data from many sources within the Integrated Genome Browser.

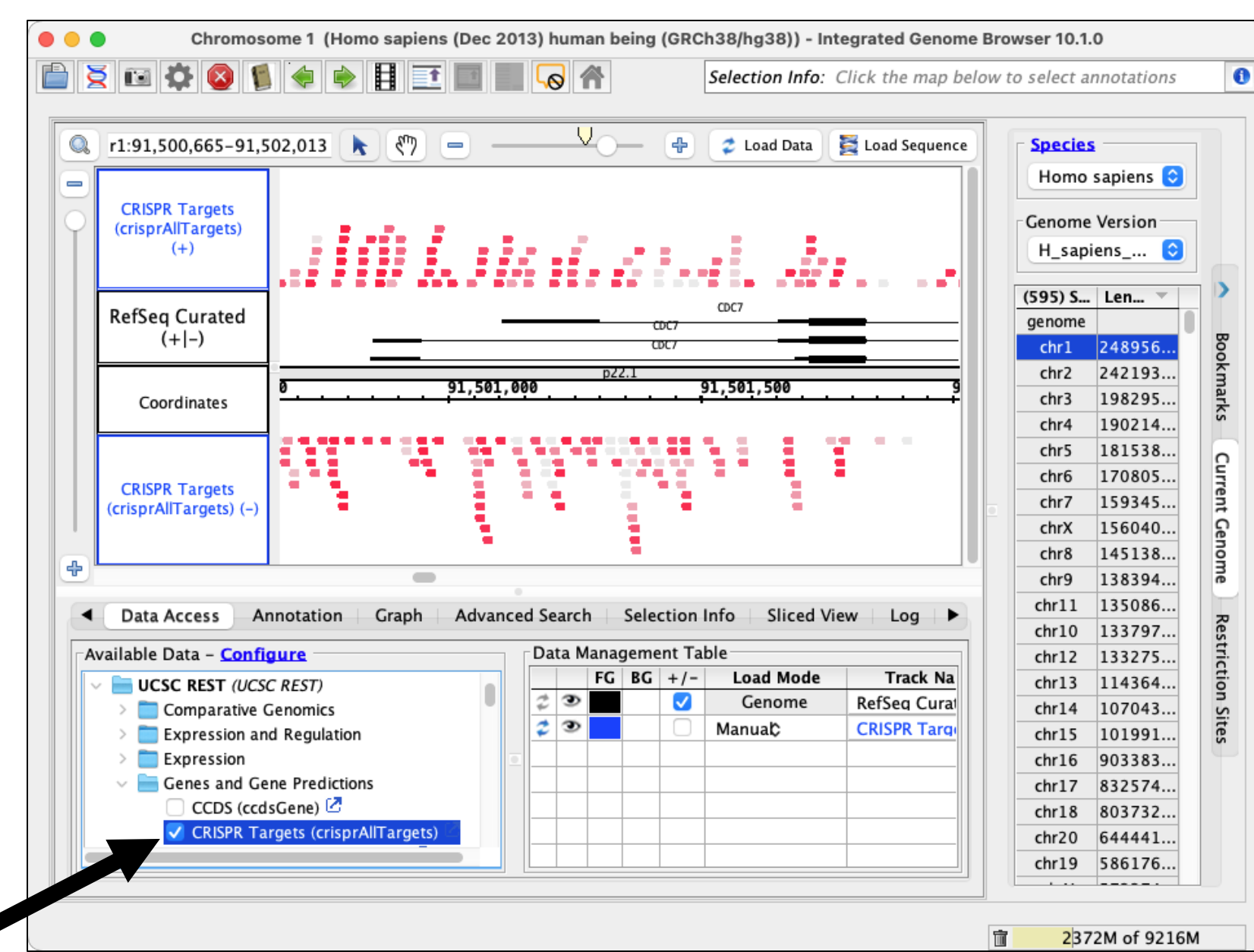


Galaxy is an open source, web-based platform for data intensive biomedical research. Data stored in Galaxy, such as BAM files, can be viewed in IGB by clicking on the **Visualize** button and then clicking **View** next to **display in IGB**.

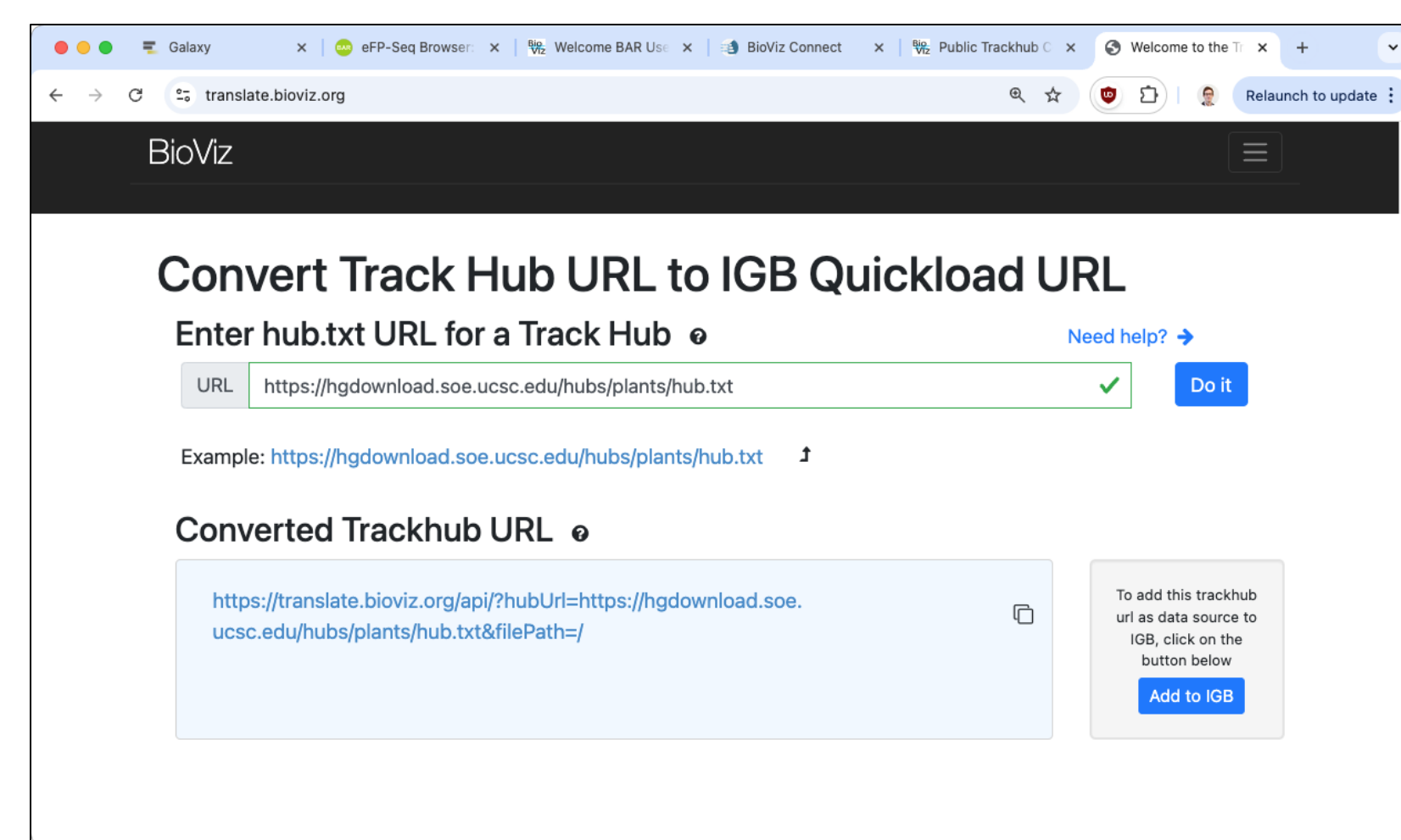


The **UCSC Genome Browser** is a widely utilized web-based tool for visualization and analysis of genomic data, encompassing over 4000 assemblies from diverse organisms. By using the UCSC REST API most of the tracks and assemblies provided by the UCSC genome browser are now available to view in IGB. Assemblies can be selected in the IGB **Species** menu and UCSC Tracks can be found under **Available Data** in the **UCSC REST** folder. UCSC tracks can be viewed alongside a user's own data in IGB.

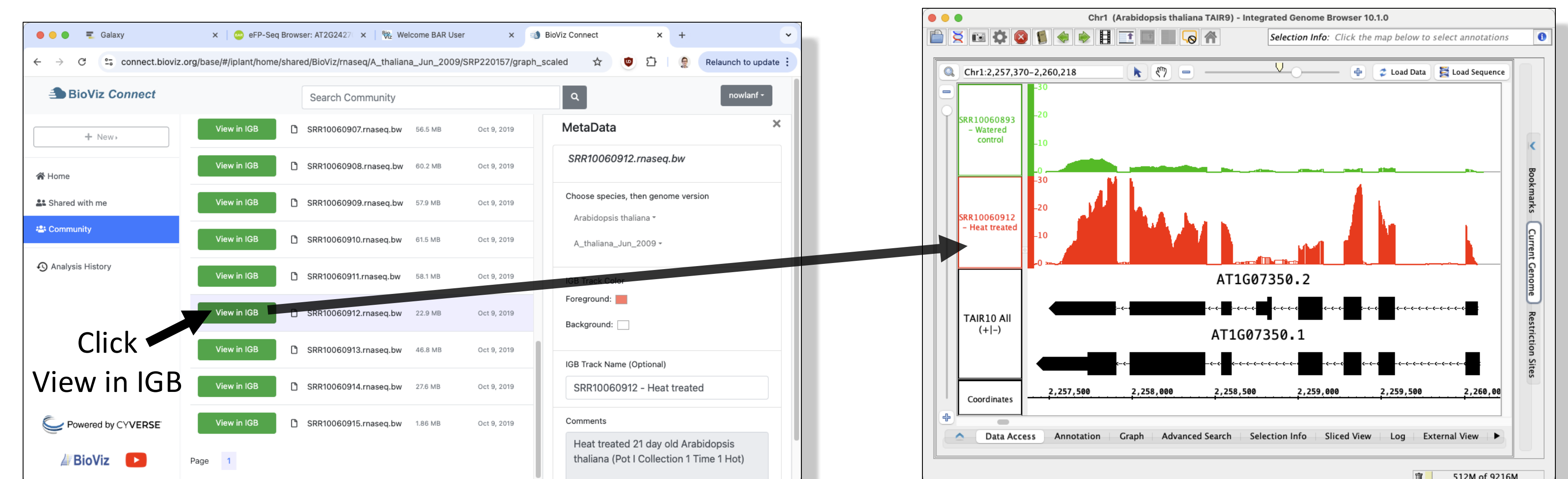
Select a UCSC track to view



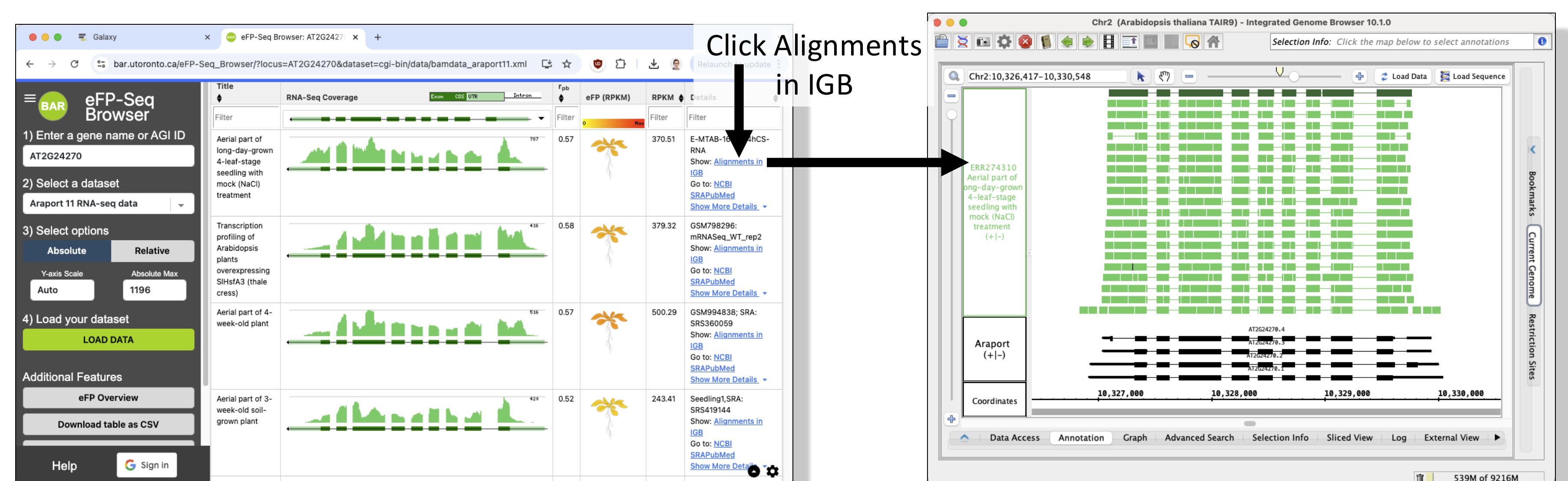
Track hubs are web-accessible directories of genomic data that can be viewed on the UCSC Genome Browser. We have created the website translate.bioviz.org that converts track hubs to IGB Quickloads so that they can be added to IGB and the data visualized. This includes both user created custom track hubs as well as the approximately 100 public track hubs that have been registered with UCSC.



CyVerse provides scientists with powerful platforms to handle huge datasets and complex analyses. CyVerse data can be visualized in IGB through our web application connect.bioviz.org by clicking on the **View in IGB** button.



The **eFP-Seq Browser** allows you to explore RNA-seq-based gene expression levels for your gene of interest in expression datasets. Data from the eFP-Seq browser can be viewed in IGB by clicking on **Alignments in IGB**.



Download IGB at BioViz.org

Research reported in this publication was supported by the NIH NIGMS (R35GM139609)