

JBrowse Genome Browser Basics

Learning objectives

- Navigate to the genome browser
- Use the menu and navigation bars
- Run searches
- Add pre-loaded data tracks
- Upload your own data tracks
- Configure tracks
- Download track data

Introduction

JBrowse is an interactive, web-based genome browser built with JavaScript and HTML5 that allows users to explore and visualize genomic data directly in their browser. It supports a wide range of data types, including sequences, annotations, and high-throughput sequencing tracks, all displayed in a fast, responsive interface. Users can zoom, search, and navigate genomic regions easily, making it useful for both research and teaching. JBrowse is widely used because it is lightweight, customizable, and designed to handle large genomic datasets efficiently. This fast and full-featured genome browser is integrated into VEuPathDB.

You can read more about JBrowse and its features here:

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4830012/>

1. Navigating to the Genome Browser (JBrowse). Links to the genome browser are available from multiple locations:

- a. The tools menu in the banner of any page¹.

The screenshot shows the homepage of TriTrypDB (Release 66, 7 May 2024). At the top, there's a search bar and a navigation bar with links like 'My Strategies', 'Searches', 'Tools', 'My Workspace', 'Data', 'About', 'Help', and 'Contact Us'. On the right, there are social media icons and a 'Guest' link. Below the header, there's a sidebar titled 'Search for...' with categories like 'Genes', 'Organisms', 'Popset Isolate Sequences', etc. The main content area has a section titled 'Overview of Resources available' with links to 'Apollo', 'BLAST (multi-query capable)', 'Companion', 'CRISPR guide design tool', 'Genome browser', 'LeishGEdit', 'NCBI Primer3', 'PubMed and Entrez', and 'Sequence retrieval'. A prominent orange box highlights 'Sustainability and Future' with a sub-section about VEuPathDB's mission and a 'Take the Survey' button. To the right, there's a sidebar for 'My Organism Preferences (88 of 88)' and a 'News and Events' section.

- b. From record pages such as gene, SNP or genomic sequence pages – these links are usually to a specific JBrowse configuration that includes data relevant to the section on that record page. For example, a JBrowse link from an RNAseq dataset on the gene page would display the gene of interest along with the RNAseq data as density or coverage plots. These links are usually indicated by the “View in JBrowse genome browser” button.

View in JBrowse genome browser

¹ Note: this exercise uses TriTrypDB (<https://TriTrypdb.org>) as an example database, but the same functionality is available on all VEuPathDB resources.

2. Getting around JBrowse. Use any of the above described JBrowse linking strategies to get to the genome browser. Once in JBrowse examine the following features:

- The menu bar: located at the top of the JBrowse frame. This includes the Genome menu, Track menu, View menu, Help menu and the Sharing link on the right side of the screen. [What do each of these do/provide?](#)
- The navigation bar: located below the menu bar. This contains
 - zooming (magnifying glass icons)
 - panning (left/right arrows) and highlighting (yellow highlighter) buttons
 - sequence selector (dropdown that allows you to select the contig, scaffold, or chromosome sequence) located between the zoom buttons and the text box
 - a text box to search for features such as gene IDs
 - an overview bar which shows the location of the region in view
- The genome view: this is where the data tracks are displayed.
- Selecting tracks: click on the “select track” button (top left). You can search/filter for tracks and then select them for display by checking the check box next to the track name.

JBrowse Screenshot with annotations:

- Menu:** Points to the top navigation bar (Genome, Track, View, Help) and the genome identifier (Trypanosoma brucei brucei TREU927).
- Navigation:** Points to the zoom and panning controls at the top of the genome view.
- Genome:** Points to the main genome view area showing mRNA transcripts.
- Select tracks:** Points to the sidebar on the left containing the "Select Tracks" interface.

Select Tracks Sidebar:

My Tracks: Currently Active, Recently Used

Category:

- Comparative Genomics
- Epigenetics
- Gene Models
- Genetic Variation
- Proteomics
- Sequence Analysis
- Transcriptomics

Subcategory:

- (no data)
- Array Probes
- BLAT and Blast Alignments
- ChIP-Seq
- DNA polymorphisms
- Gene Expression
- Origins of Replication
- Orthology and Synteny
- Protein Expression
- RNA-Seq
- Sequence assembly
- Sequence conservation, complexity and repeats
- Sentence sites, features
- Splice Sites
- Transcripts

Dataset:

- (no data)
- 1 Nucleic-acid proteome, prokaryotic form (Tb927)
- 8 Adipose tissue vs muscle tissue transcriptomes
- 1 Aligned BAC End Sequences (RPCI93 library)
- 1 Aligned cell line cell for Lister strain 427
- 2 Aligned genomic sequence reads - strain TREU927

Search and Filter:

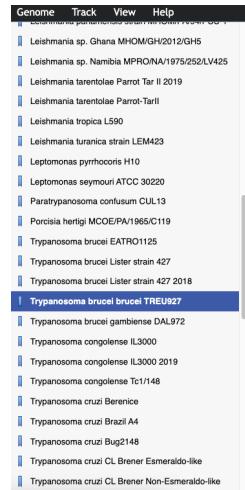
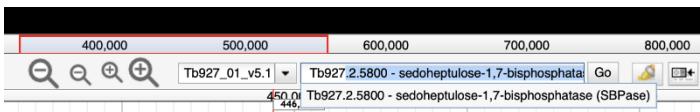
- Back to browser
- Clear All Filters
- Contains text: 659 tracks

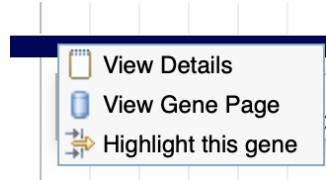
Name	Category	Subcategory	Dataset	Track Type	RNA-Seq Alignment	RNA-Seq Strand
Influence of glucose on transcription of SS, LS, PC - 001.1 - SS + gluc (unique forward) Coverage	Transcriptomics	RNA-Seq	Influence of glucose on transcription of SS, LS, PC	Coverage	unique	forward
Influence of glucose on transcription of SS, LS, PC - 001.2 - SS + gluc (non-unique forward) Coverage	Transcriptomics	RNA-Seq	Influence of glucose on transcription of SS, LS, PC	Coverage	non-unique	forward
Influence of glucose on transcription of SS, LS, PC - 002.1 - SS + gluc + Pro, Thr (unique forward) Coverage	Transcriptomics	RNA-Seq	Influence of glucose on transcription of SS, LS, PC	Coverage	unique	forward
Influence of glucose on transcription of SS, LS, PC - 002.2 - SS + gluc + Pro, Thr (non-unique forward) Coverage	Transcriptomics	RNA-Seq	Influence of glucose on transcription of SS, LS, PC	Coverage	non-unique	forward
Influence of glucose on transcription of SS, LS, PC - 003.1 - SS + gluc (unique forward) Coverage	Transcriptomics	RNA-Seq	Influence of glucose on transcription of SS, LS, PC	Coverage	unique	forward
Influence of glucose on transcription of SS, LS, PC - 003.2 - SS + gluc (non-unique forward) Coverage	Transcriptomics	RNA-Seq	Influence of glucose on transcription of SS, LS, PC	Coverage	non-unique	forward
Influence of glucose on transcription of SS, LS, PC - 004.1 - SS + gluc + Pro, Thr (unique forward)	Transcriptomics	RNA-Seq	Influence of glucose on transcription of SS, LS, PC	Coverage	unique	forward

3. Navigate to a specific gene in JBrowse.

The goal of this step is to navigate to the sedoheptulose-1,7-bisphosphatase (SBPase) gene of *T. brucei* 927.

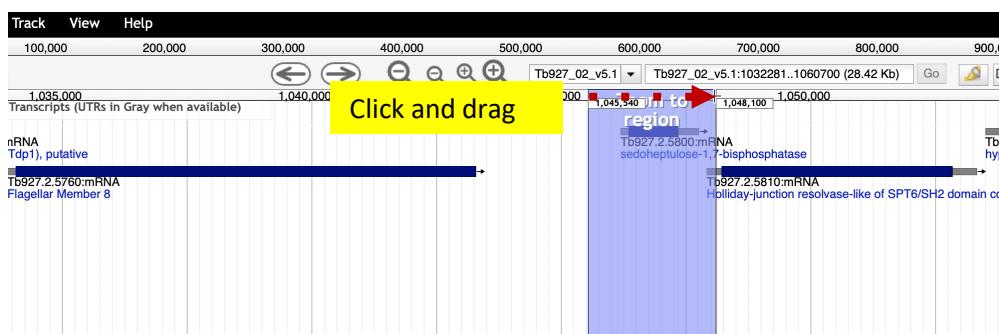
- Make sure the *Trypanosoma brucei brucei* TREU927 genome is selected from the genome menu.
- Start typing the word sedoheptulose in the search box. After a few seconds you should see the result of the search (do not hit enter). Select the gene from the search dropdown and press GO. This will take you to Tb927.2.5800.



- You can get information about any feature in the genome view window by clicking on it. Click on the gene. [What information is in the popup?](#)
- You can also right click (or control click) on a feature to display a menu which provides quick links to highlight a feature, go to the feature page (like the gene page) or get the info popup (the same one you get when you click on the feature).
- 
- [What genes are immediately upstream and downstream of SBP?](#) (Hint: use the zoom out button in the navigation bar).

What is the difference between the small and large zoom buttons? (Tip1: another way to zoom in and out is by clicking on shift and the up or down arrows. Tip2: you can also zoom in by clicking and dragging your cursor in the location ruler in the navigation bar.)

[What happens if you click shift and left or right arrows?](#)



4. Explore transcription start sites.

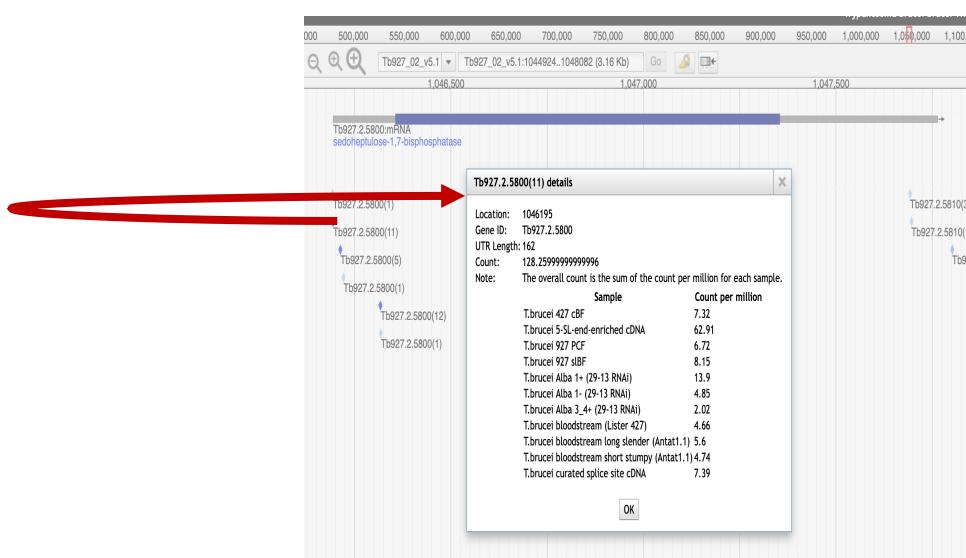
Examine the features of the sedoheptulose-1,7-bisphosphatase gene. Gene models in JBrowse use color to indicate the direction of transcription: blue features run from left to right (forward strand), and red features run from right to left (reverse strand). The thicker colored blocks represent exons, while the white blocks represent untranslated regions (UTRs), which are transcribed but not translated into protein. Introns are represented as thin lines connecting the exon blocks within a gene model.

Are you confident about the gene transcription start site of this gene? What additional data track would be useful for you to assess the transcription start site? (hint: Click on the “Select Tracks” button to reveal all available tracks. Now type the word “splice” in the “contains text” box. This will filter all tracks that contain the word splice. Find the one called “Unified Splice Leader Addition Sites” and select it. Click on the “Back to browser” button).

The screenshot shows the 'Select Tracks' interface. At the top, there's a search bar with 'Contains text: splice' and a result count of '7 matching tracks'. Below the search bar is a table with columns: Name, Category, Subcategory, Dataset, Track Type, RNASeq Alignment, and RNASeq Strand. The 'Unified Spliced Leader Addition Sites' track is listed and checked. On the left, there's a sidebar with categories like 'My Tracks', 'Gene Models', 'Poly A Sites', 'Splice Sites', etc., with some items expanded to show sub-options. A 'Back to browser' button is at the top left of the sidebar.

Examine the Unified Splice Leader Addition Sites track. What do the different diamond colors mean? Each track title has a drop-down menu that contains a link to an ABOUT popup. Click on this to see what the colors mean. Click on the diamonds to see information about that specific feature.

Which color provides the most evidence for a splice junction?



5. Explore synteny between genomes. Synteny helps define conservation of homologous genes and gene order between genomes.

- Go to the “Select Tracks” tab on the left of the page and turn on the track called “Syntenic Sequences and Genes”. How did you find this track? One option is to click on the “Comparative Genomics” category on the left side to filter the tracks.

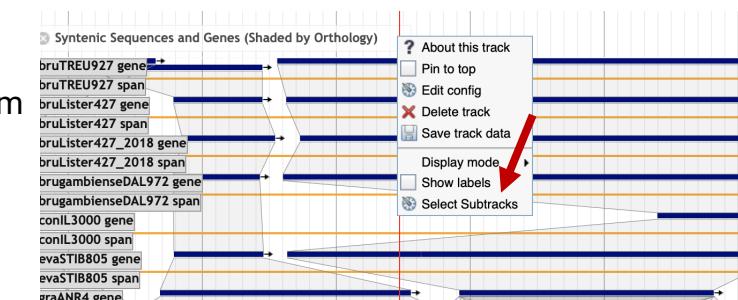
The screenshot shows the 'Select Tracks' interface. On the left, there's a sidebar with 'My Tracks' (Currently Active, Recently Used), 'Category' (Comparative Genomics, Gene Models, Genetic Variation, Proteomics, Sequence Analysis, Transcriptomics), and a search bar ('Contains text'). The main area lists tracks: 'Syntenic Sequences and Genes (Shaded by Orthology)' is checked and highlighted in blue. Other tracks listed include 'Comparative Genomics', 'Gene Models', 'Genetic Variation', 'Proteomics', 'Sequence Analysis', and 'Transcriptomics'. The header says '1 matching track'.

- Return to the browser by clicking “Back to Browser” and zoom out so you can see a couple of genes on either side of SBP (does not have to be exact).
- Configure the synteny track to include the following species subtracks: *Trypanosoma brucei* 927, *T. brucei* 427, *T. brucei gambiense*, *T. congolense*, *T. evansi*, *T. theileri* and *T. grayi*, *Crithidia fasciculata*, *Leishmania amazonensis*, *L. braziliensis*, *L. donovani*, *L. infantum*.

To configure the subtracks, Click on the down arrow in the track name



Select the option called “Select Subtracks” from the menu

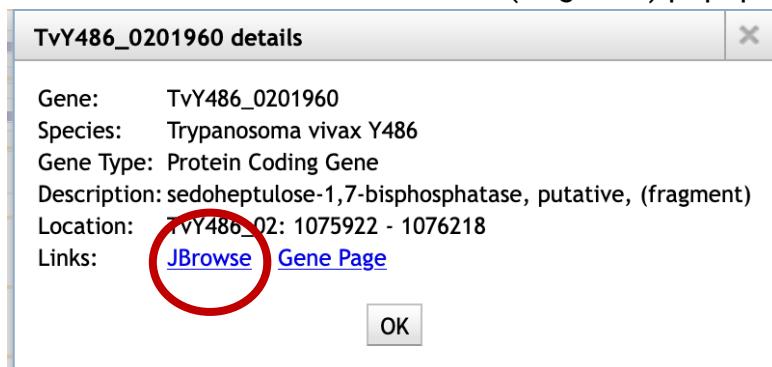


In the next popup,

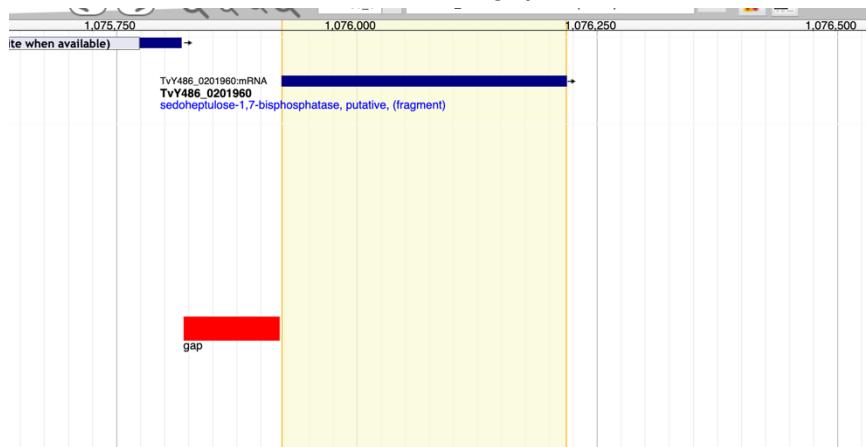
- first uncheck all organisms
- second use the filters on the left to select Trypanosoma
- third select the species of interest (note that you should select both the gene and span subtracks for each species and that there may be two genome versions for an organism)
- fourth click on the save button at the bottom of the popup, this takes you back to the genome view

The screenshot shows the 'Select Subtracks' dialog box. It lists 12 matching tracks under 'My Tracks' (Currently Selected). The tracks are categorized by species: *Trypanosoma brucei* (8 tracks), *Trypanosoma congolense* (2 tracks), *Trypanosoma evansi* (2 tracks), *Trypanosoma grayi* (2 tracks), *Trypanosoma rangeli* (2 tracks), *Trypanosoma theileri* (2 tracks), and *Trypanosoma vivax* (2 tracks). The dialog also includes filters for 'Species', 'Kingdom', 'Phylum', 'Class', 'syntype', and 'taxon'.

- d. Observe the synteny track
- What does the synteny track in this region look like? Feel free to zoom out some more. Are genes (in general) similarly organized between these species? What does the shading between genes mean?
 - What direction is the SBPase gene relative to the chromosome?
 - What genes are upstream and downstream of the SBPase? Are these genes syntenic?
 - What does synteny look like if you add more distantly related species? Does SBPase appear to have orthologs in *Leishmania* species or *Crithidia*?
- e. Add a synteny subtrack for *T. vivax*. Examine the gene corresponding to the *T. vivax* SBPase in the synteny track. Hover over the gene image to find the gene name in the popup. Does this gene appear to be a fragment? What could be some possible reasons for this?
- f. Click on the Jbrowse link in the *T. vivax* SBPase (fragment) popup.



- g. Add the track called “scaffolds and gaps” .



What does this track tell you about the sequence in that area? Based on this do you expect *T. vivax* to contain a full SBPase sequence?

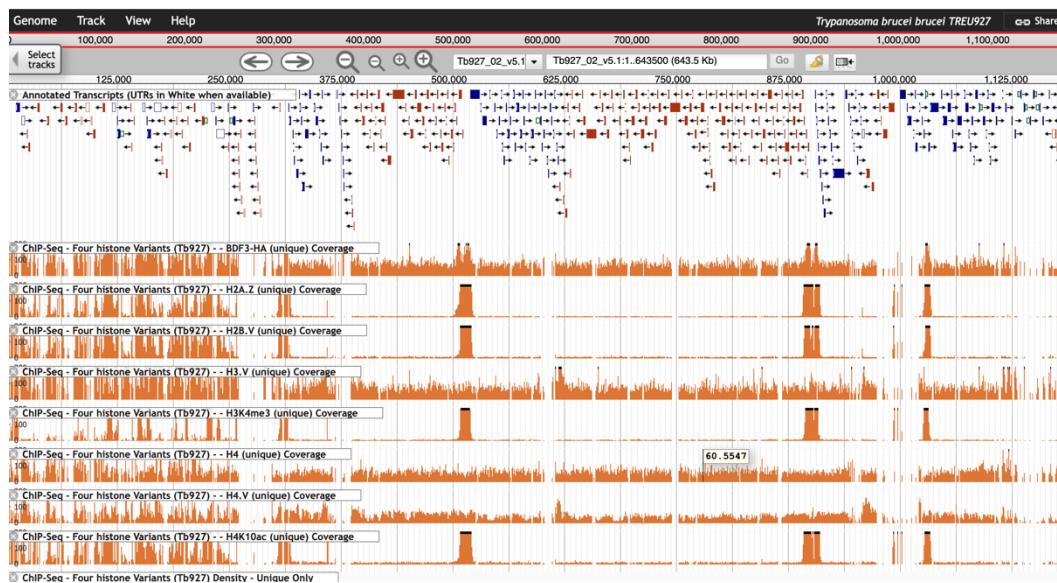
6. Explore other data tracks in JBrowse. For this example, we will view *T. brucei* TREU927 data, so the data tracks you turn on will display data only if the data is aligned to the *T. brucei* genome.

- Return to the SBPase gene in *T. brucei* by searching for the gene ID (Tb927.2.5800) in 'Landmark or Region' to redirect the browser.
- Turn on the ChIP-seq coverage plots and turn off the syntenic gene and region tracks. The data tracks are from an experiment called: ChIP-Seq - Four histone Variants ChIP-Seq Coverage aligned to T brucei TREU927 (Cross) (linear plot). For this experiment, chromatin was immunoprecipitated using several different histone antibodies. The DNA that precipitated with the histone was sequenced and aligned to the *T. brucei* TREU927 genome. Peaks in the sequence coverage plots represent areas of histone binding. Different histone variants can be associated with start and termination sites for transcription

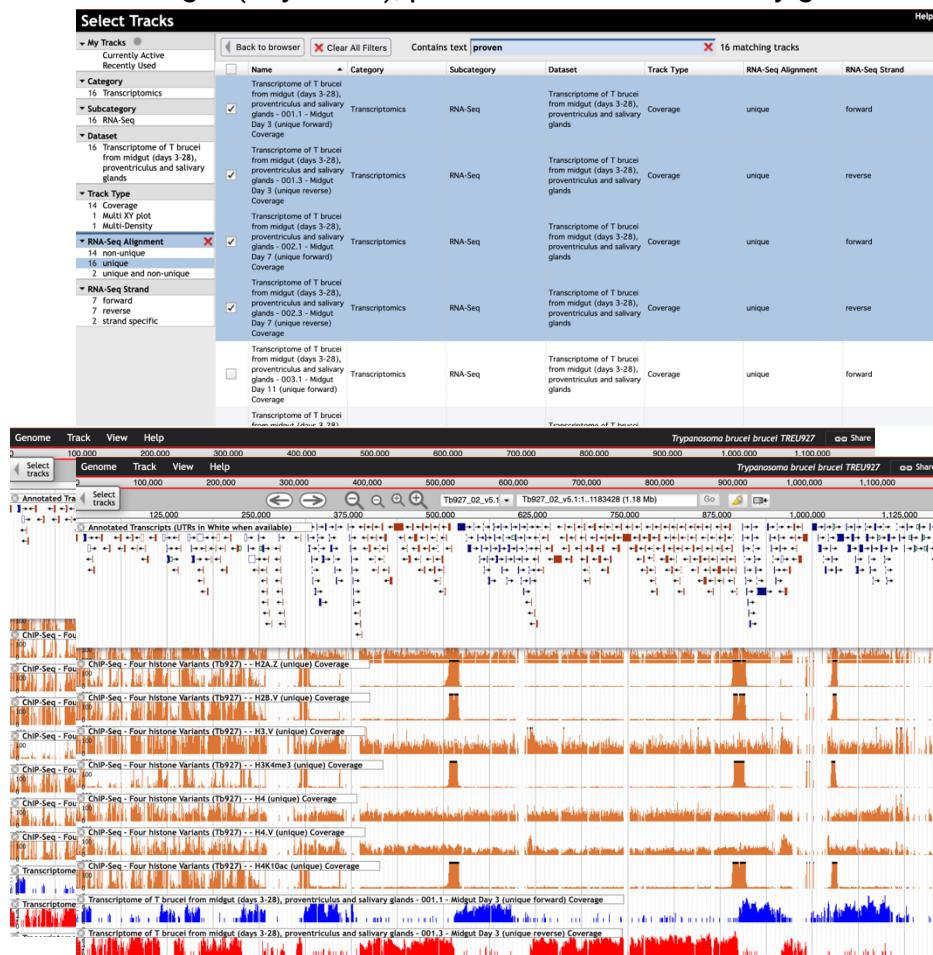
(<http://www.ncbi.nlm.nih.gov/pubmed/19369410>).

Name	Category	Subcategory	Dataset	Track Type	RNASeq Alignment	RNASeq Strand
BDf3-HA (unique) Coverage	Epigenomics	ChIPSeq	ChIP-Seq - Four histone Variants	Coverage
ChIP-Seq - Four histone Variants Density - Unique And Non-Unique	Epigenomics	ChIPSeq	ChIP-Seq - Four histone Variants	Multidensity
H2Az (unique) Coverage	Epigenomics	ChIPSeq	ChIP-Seq - Four histone Variants	Coverage
H2Bv (unique) Coverage	Epigenomics	ChIPSeq	ChIP-Seq - Four histone Variants	Coverage
H3K4me3 (unique) Coverage	Epigenomics	ChIPSeq	ChIP-Seq - Four histone Variants	Coverage
H3v (unique) Coverage	Epigenomics	ChIPSeq	ChIP-Seq - Four histone Variants	Coverage
H4 (unique) Coverage	Epigenomics	ChIPSeq	ChIP-Seq - Four histone Variants	Coverage
H4K10ac (unique) Coverage	Epigenomics	ChIPSeq	ChIP-Seq - Four histone Variants	Coverage
H4v (unique) Coverage	Epigenomics	ChIPSeq	ChIP-Seq - Four histone Variants	Coverage

- Zoom out to display the entire chromosome. What does this data show you?
- Roughly how many polycistronic units does this chromosome have? Zoom out to the entire chromosome.

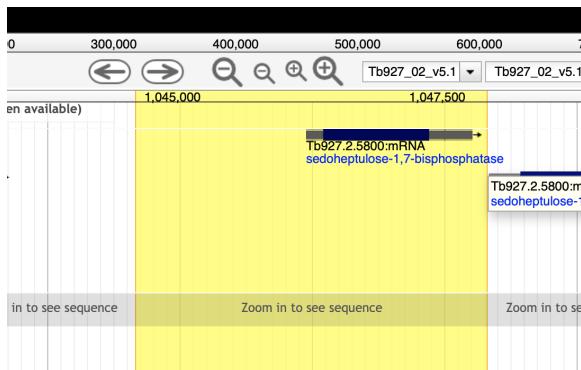
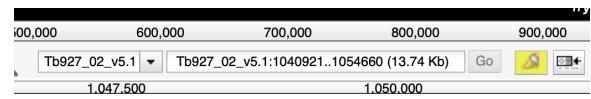


- e. Do the ChIP-seq peaks correlate with the direction of gene transcription (blue vs. red)?
- f. What if you turn on a strand-specific RNAseq dataset – for example, turn on some of the uniquely mapped forward and reverse coverage tracks from the experiment: “Transcriptome of *T brucei* from midgut (days 3-28), proventriculus and salivary glands”.



7. Retrieve data from JBrowse. Download sequence in FASTA format from the SBPase gene (your region of interest):

- Make sure the “annotated transcripts” and the “reference sequence” tracks are turned on.
- Click on the “highlight a region” button in the navigation bar. The button should turn yellow when activated.
- Click and drag in the genome view region (in the Annotated Transcripts track) and select the area you would like to highlight.



- Click on the down arrow on the reference sequence track and select “Save track data”.
- In the next popup window you can keep everything as the default and either save or view the sequence.



8. Upload your own tracks to JBrowse: JBrowse can accept several standard-format data files by direct upload or through a URL if the data is stored remotely. Some file formats like BAM and VCF require indexing before uploading.

In this exercise we will download a bigwig file from GEO and then upload it to JBrowse:

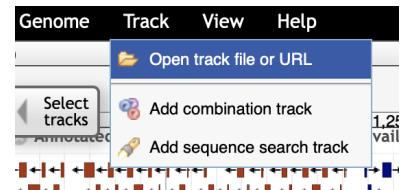
- Go to this GEO sample record:

<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM2407365>

- Scroll down to the bottom of the page and download the bigwig file with the http link.

Supplementary file	Size	Download	File type/resource
GSM2407365_BF_WT_HNI_VO2_rep2-T_brucei_427.bigwig	12.4 Mb	(ftp)(http)	BIGWIG

- Once the file is downloaded go to JBrowse and select *Trypanosoma brucei brucei* Lister 427 as the reference genome (hint: use the Genome link in the menu panel, top left).
- Turn on the track for annotated transcripts if it is not on already.
- Click on the Tracks menu item and select “Open track file or URL”.
- In the popup click on select file then select the file you just downloaded. JBrowse should automatically recognize that the file is in bigwig format.
- Click on the Open button.



- The bigwig output should appear very quickly in your browser.

