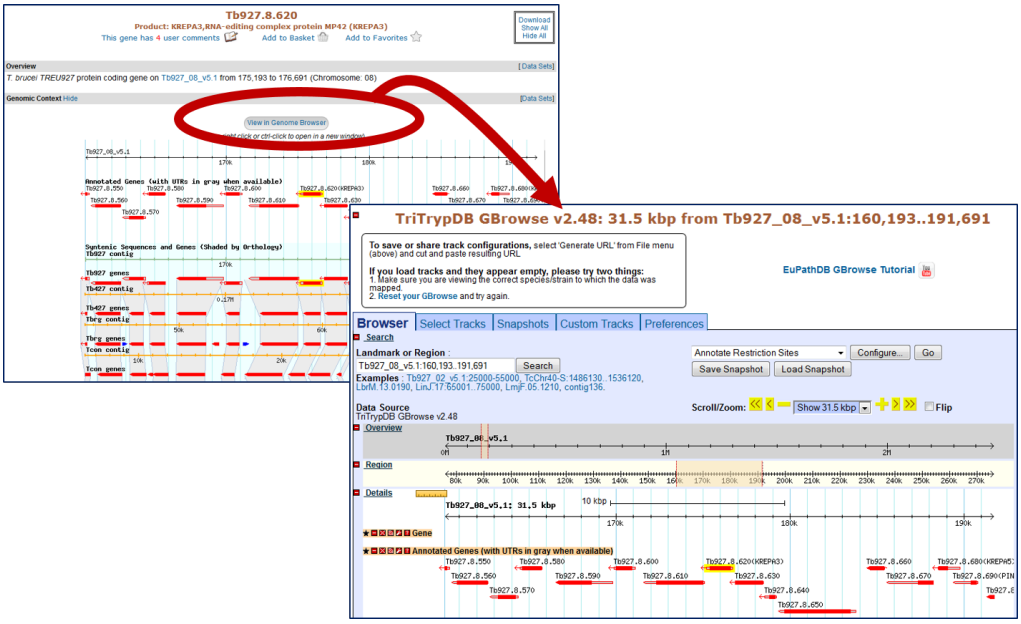
Browser Exercises
Alignments and Comparative genomics

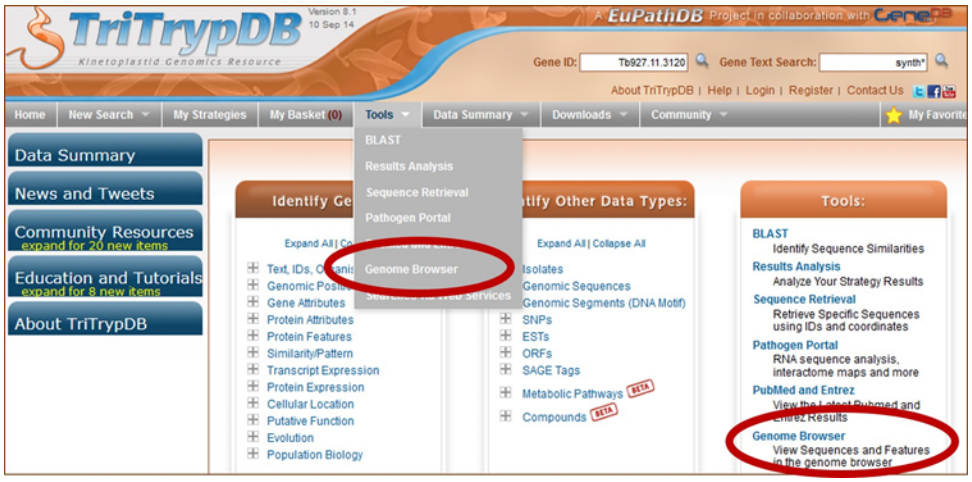
1. Navigating to the Genome Browser (GBrowse)

Note: For this exercise use http://www.tritrypdb.org

1a. There are two ways to navigate to GBrowse from TriTrypDB.

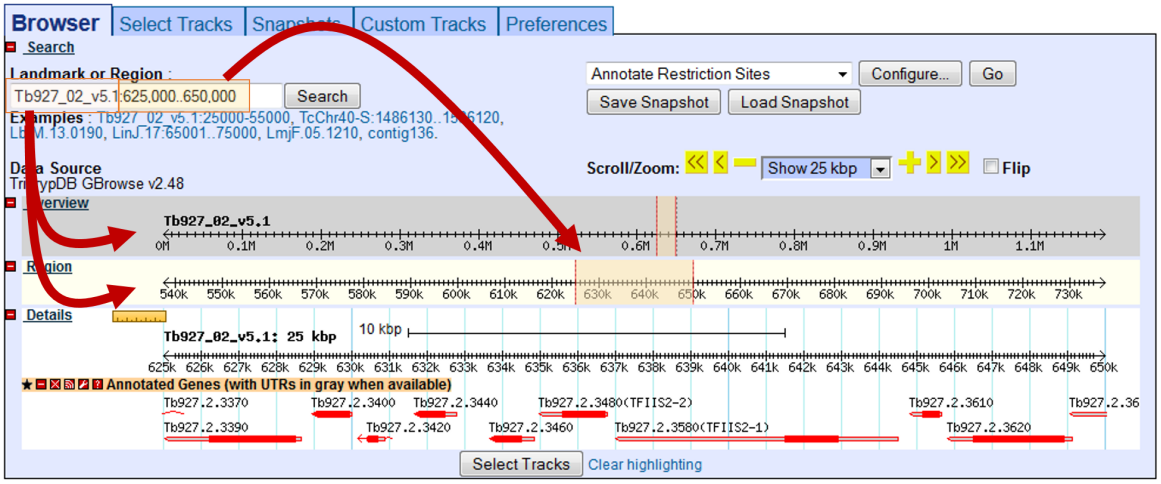
- From record pages, like a gene page, genomic sequence or EST page, click on the "View in Genome Browser" link. You can also use the Tools section on the homepage or the grey toolbar in the header section



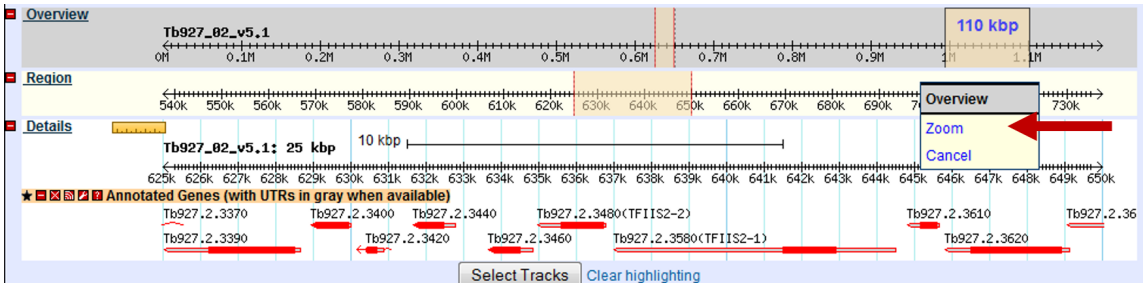


1b. **Go to GBrowse from the TriTrypDB home page.** Explore this page – take note of the different sections: Instructions, Search, Overview, Region, Details, Tracks, etc…

1c. Look at the "Landmark or Region" box.

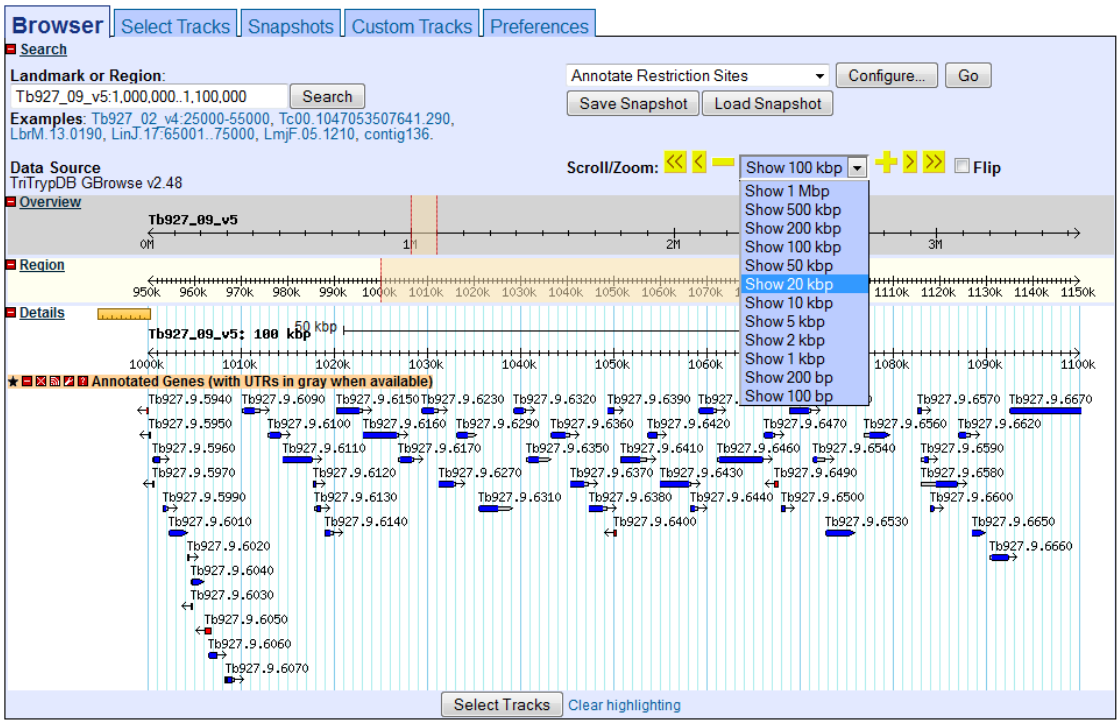


- What information does the "Landmark or Region" box contain? The Landmark or Region box should read – Tb927_02_v5.1:625,000..650,000.
- What chromosome is displayed?
- What location of the chromosome is displayed?
- Move to a different genomic region on this chromosome – for example, visit the right arm of this chromosome.
    - Hint: change the coordinate numbers in the "landmark or region" box to correspond to an area in that region. Look at the overview to give you an indication of the total size of this chromosome, ie. 1000000..1100000).
    - OR highlight the area representing approx. 1000000-1100000 on the scale in the Overview section and then choose zoom from the popup.

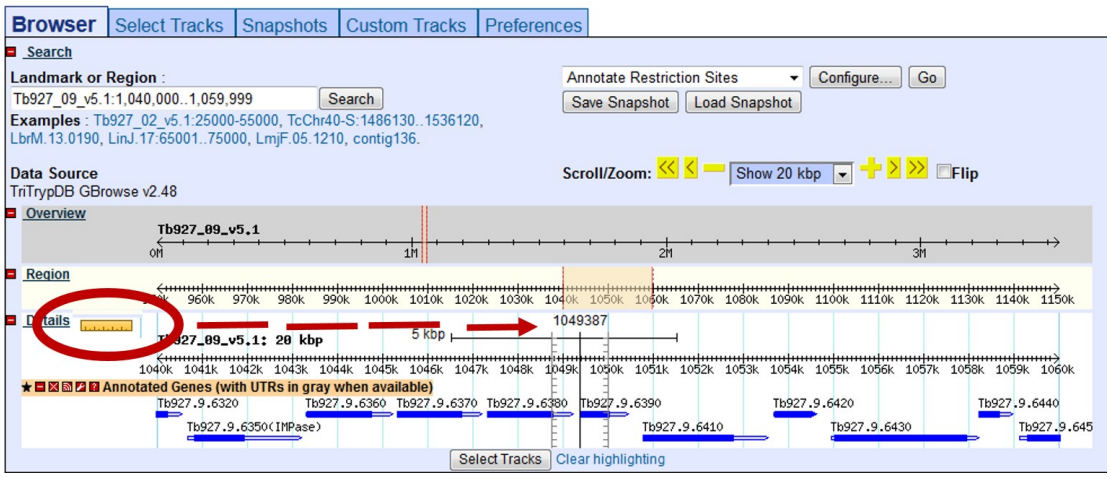


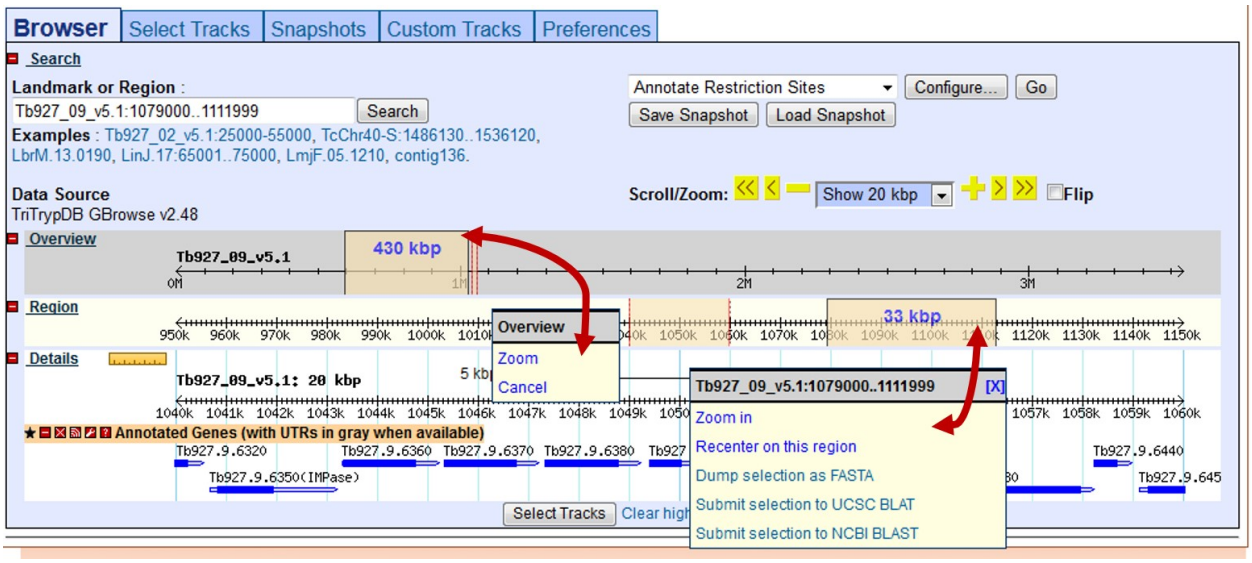
- Move to chromosome 9. How did you do this?
    - Hint: Change the chromosome number in the "landmark or region" box – it should look like this: Tb927_**09**_v5.1:1,000,000..1,100,000.

- Zoom in to a 20Kb region. Select 20Kb from the Scroll/zoom drop down menu.



- What genes are in this region? Mouse over the gene graphics and look at the popups.
- Explore the ruler tool. Click on the ruler to engage then drag it across the window. The ruler tool displays the nucleotide coordinates of the ruler's solid center line. This is very useful for comparing between the annotation data track and others that we will add later.
- There are other ways to move and zoom. Try highlighting an area along the scale in the overview, region or details sections of GBrowse.





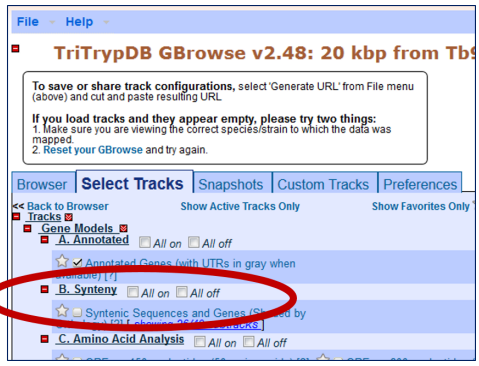
1d. What if you want to go to a specific gene in Gbrowse? Try to figure out how to go to this gene: Tb927.2.5800

- Type the ID in the "landmark or region" box. The landmark box has a search function that supports gene IDs. What else does it support?
- What is this gene?

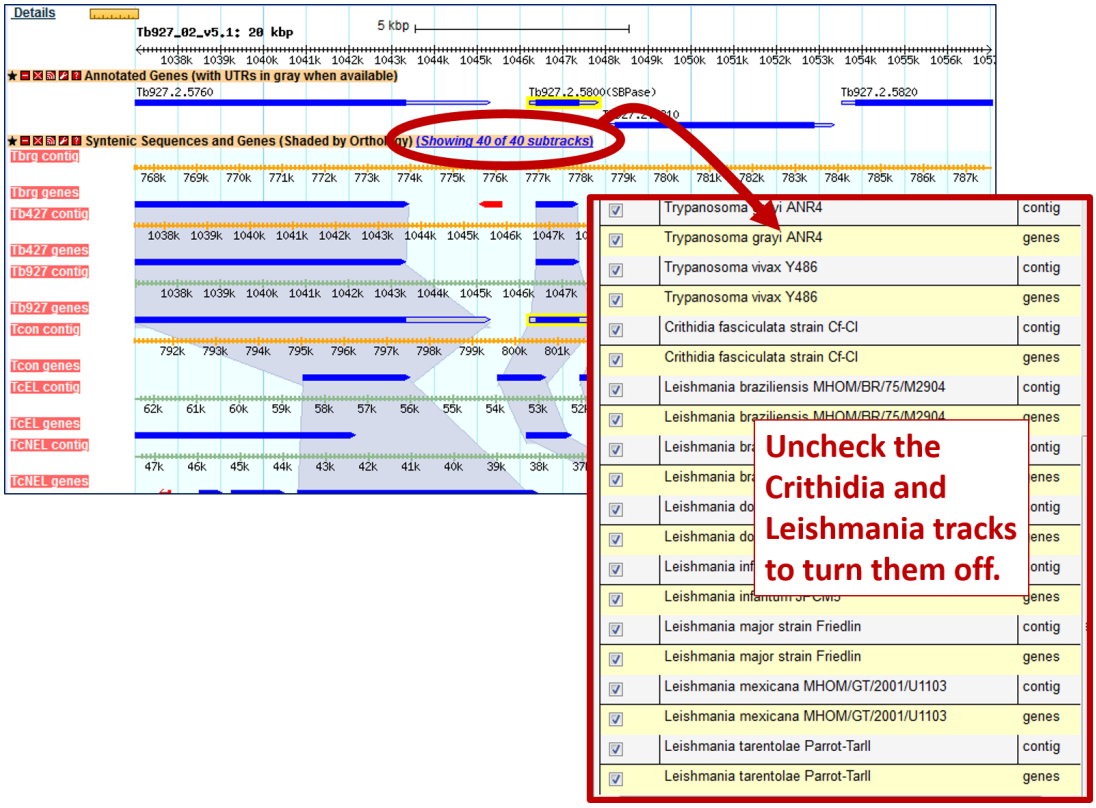
2. Exploring Synteny data tracks in GBrowse

2a. Is the region containing the sedoheptulose-1,7-bisphosphatase (SBPase) gene syntenic in all kinetoplastids?

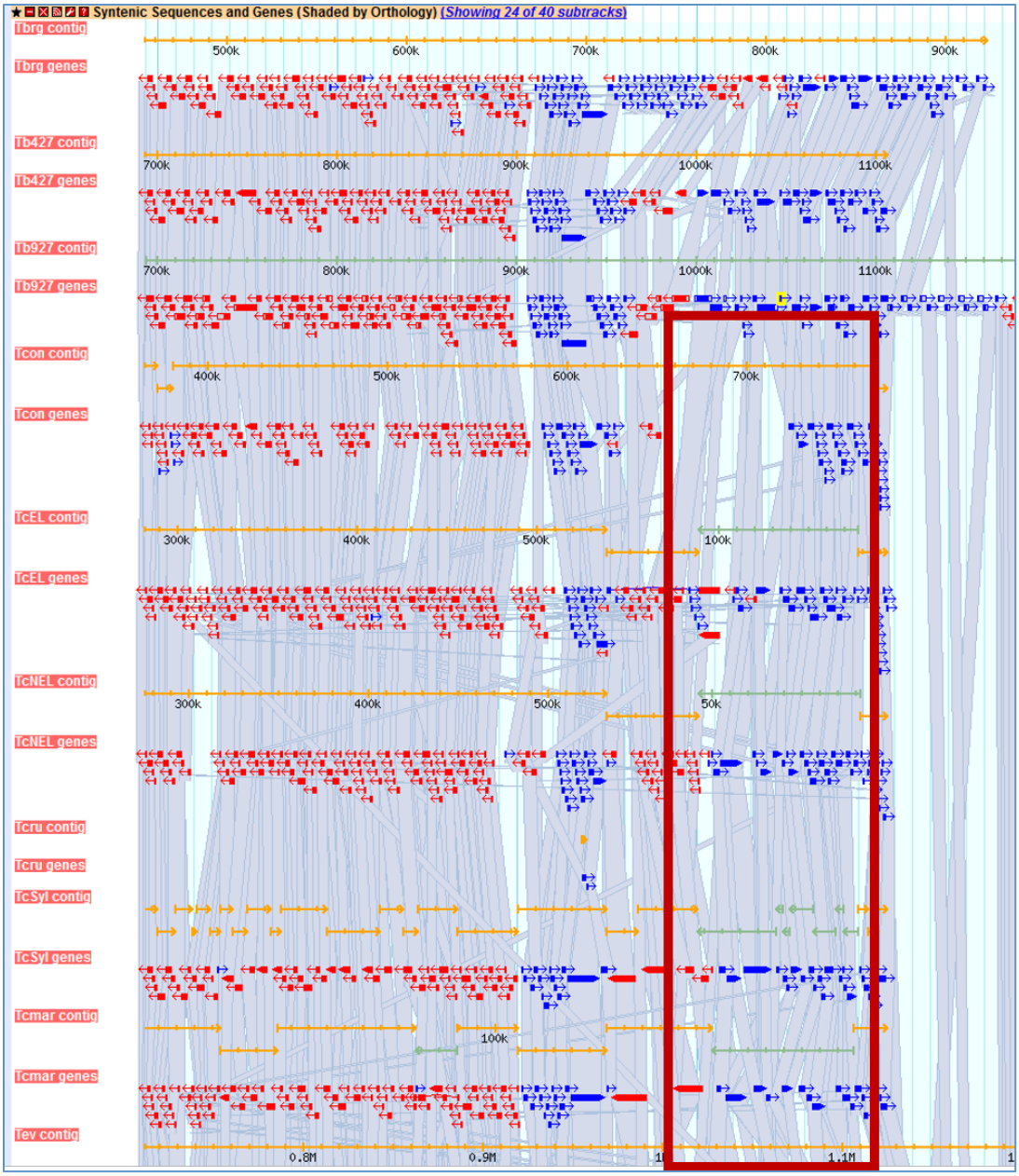
Hint: Go to the "Select Tracks" section and turn on the track called "Syntenic Sequences and Genes". The browser is automatically updated with tracks you select. Note that this track contains multiple subtracks.



- Return to the browser by clicking the “Browser” tab and **zoom out to 20Kb**.
- What does this region look like?
- What direction is the gene relative to the chromosome?
- What genes are upstream and downstream of the SBPase?
- Modify the subtracks to remove *Leishmania* species from the view. Click on the link ‘showing 40 of 40 subtracks’, wait for the popup and uncheck all *Leishmania* and *Crithidia* species. **Then click Change**.



- Zoom out to 500KB – what could you conclude about this region in *T. congolense*? (See image on next page if needed).
- You will also notice that some of the genomes have contigs that are not contiguous. Why is that?
- Mouse over the two contigs and look at the information in the popups – do these pieces belong to the same chromosome? What does this mean?


Syntenic Sequences and Genes (Shaded by Orthology) (Showing 24 of 40 subtracks)
Tbrg contig
Tbrg genes
Tb427 contig
Tb427 genes
Tb927 contig
Tb927 genes
Tcon contig
Tcon genes
TcEL contig
TcEL genes
TcNEL contig
TcNEL genes
Tcru contig
Tcru genes
TcSyl contig
TcSyl genes
Tcmar contig
Tcmar genes
Tev contig

2b. Exploring other data tracks in Gbrowse.

In this example we are viewing *T. brucei* SBPase, so the data tracks you turn on will display data only if the data is aligned to the *T. brucei* genome.

Turn on the ChIP-seq coverage plots and turn off the syntenic gene and region tracks. The data track is called: **ChIP-Seq - Four histone Variants ChIP-Seq Coverage aligned to T brucei TREU927 (Cross) (log plot)**.

- What does this data show you?
  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 and transcription start sites.
  http://www.ncbi.nlm.nih.gov/pubmed/19369410
- Roughly how many polycistronic units does this chromosome have? Zoom out to the entire chromosome.
- Do the ChIP-seq peaks correlate with the direction of gene transcription (blue vs. red)?

3. Uploading your own tracks to GBrowse

- **Uploading your own tracks is also possible.** One reason to upload your own tracks is if you have data that you would like to display on a chromosome or genomic segment and view it in the context of gene models and other data. To do this you have to follow some rules to ensure that the file you are uploading can be understood by GBrowse.

  - Now let's load a complex graphic, a bigwig file of some RNA Sequencing data. For this we posted the file to a public site and are using the URL to direct GBrowse to the file location. In the field "Fetch track file from this URL", enter the following and click Import:
    http://loquat.rcc.uga.edu/swfeltz/bigwig/TREU927_Cross_RNASeq.bw