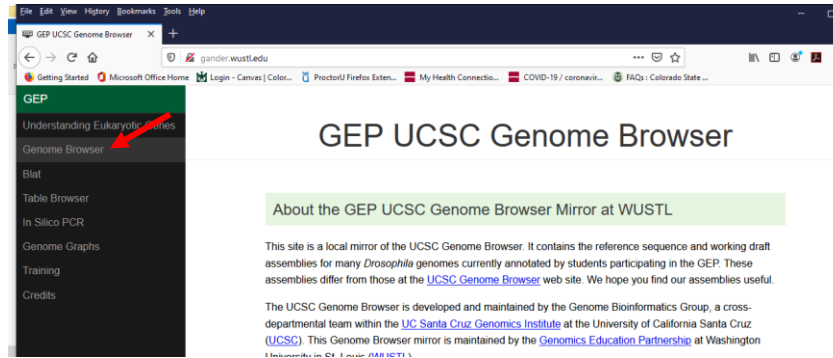


Addendum

Retrieving your Contig and Navigating the gene Browser

Retrieving you Contig through the GEP Genome Browser

1. go to the Genome Browser in <http://gander.wustl.edu>,



2. This takes you to the Gateway Page. On the left side of the resulting page, click “*D. ananassae*” on the phylogenetic tree, then using the drop-down menus, choose “Sep. 2019 (GEP/Muller F Element)”, and type your contig # in position /search term box. Click GO

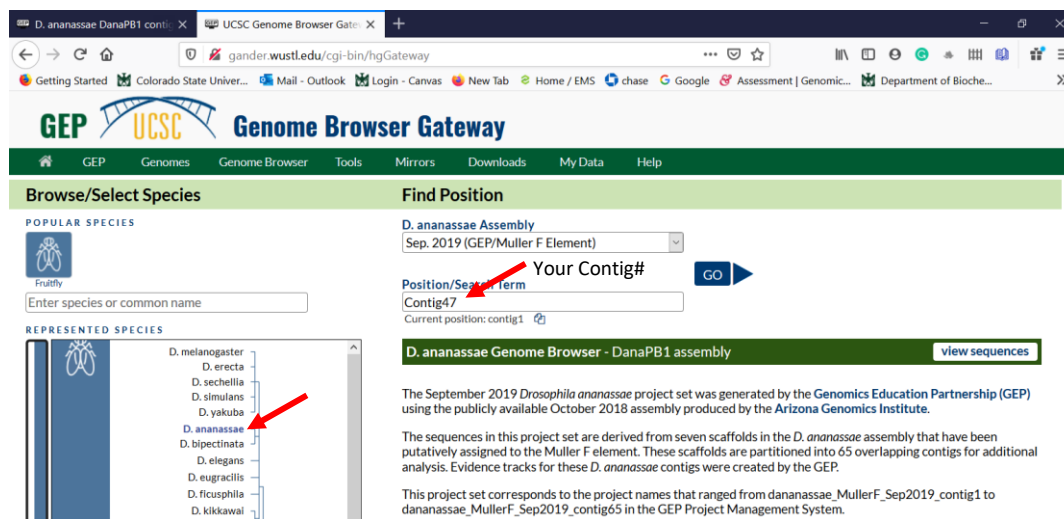


Fig 1:GEP Genome Browser page

You will see the BlastX page where your Contig DNA was translated into proteins and was Compared to *Drosophila melanogaster* proteins. The page will show several genes aligned to your contig with many “Evidence tracks” un the bottom that were used to find gene alignment with you contig. To simplify the page click on “hide all” (red box in figure below) (Figure 2)

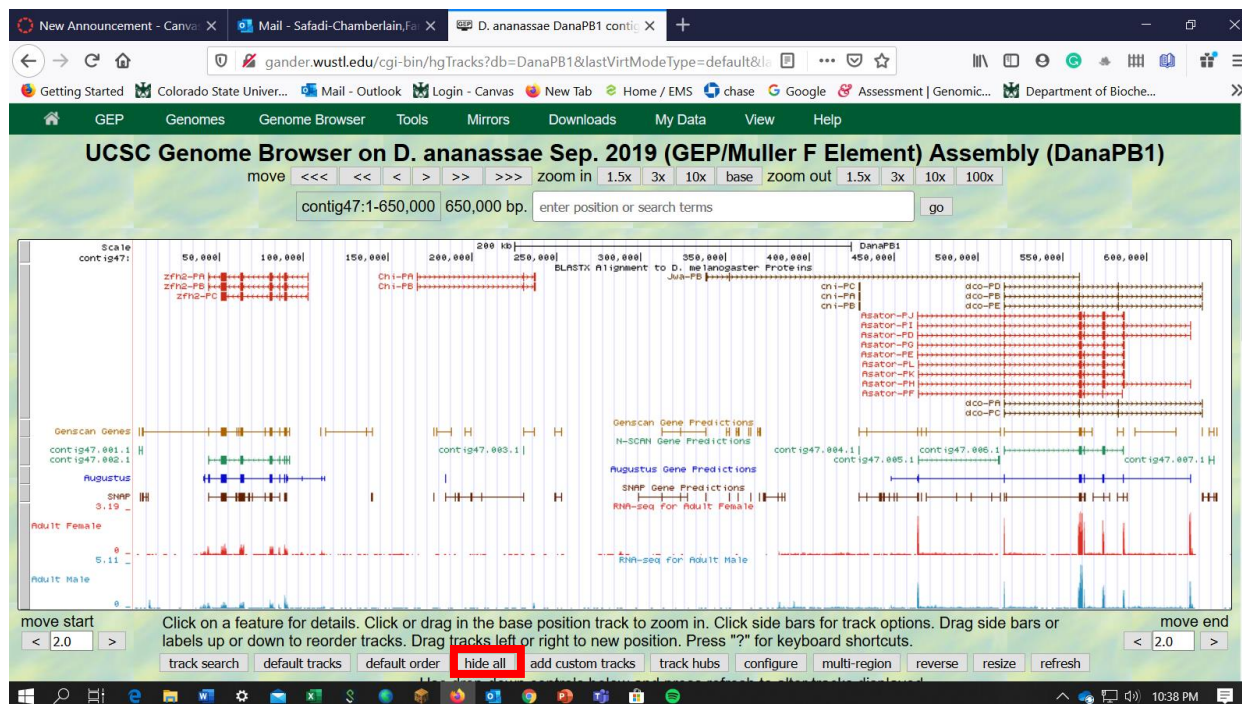
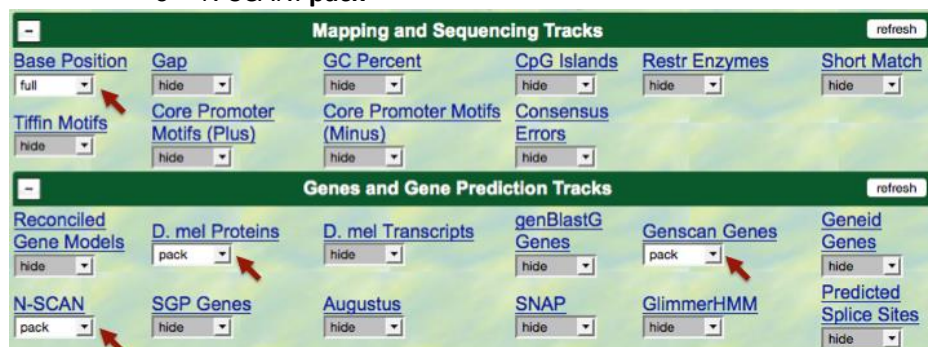


Figure 1 Click on the “hide all” button to hide all the evidence tracks.

and then adjust the following tracks by Using the drop-down boxes in the track configuration sections to change the display options for the following evidence tracks and then click on the “refresh” button (**Error! Reference source not found.**)

- Under “Mapping and Sequencing Tracks”
 - Base Position: **full**
- Under “Genes and Gene Prediction Tracks”
 - D. mel Proteins: **pack**
 - Genscan Genes: **pack**
 - N-SCAN: **pack**



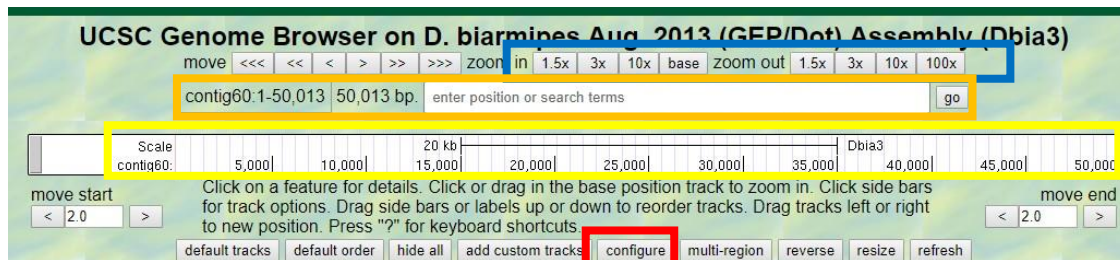
3. Now the white box should be much smaller. The white box with the vertical hatch marks represents the entire nucleotide sequence of our contig, with marker points every 5000 bases (base 0 is on the left, and it goes all the way to a bit over 50,000 on the right (yellow box below). In the position/search field above, it lists exactly which bases are “displayed” in the window—it says contig#: 1-50,013 (the total number of bases in this contig). **The actual nucleotide sequence is not shown at this point, since the computer can’t fit that many characters on one line to display in your browser.** Only when we zoom in will we be able to see the actual sequence.

Note: To fit the white hatched box to your screen, click “configure” (red box below) and set the “image width” to 1200 pixels (or whatever width seems appropriate for your device). Increase the “text size” to 10, and click “submit”.

NOTE: replace “#” with your contig number

There are multiple ways to zoom in and out on this long sequence:

- Click on the “zoom in” or “zoom out” buttons at the top of the browser (blue box below).
- Highlight the numbers in the hatched box, in the region you would like to zoom in to (yellow box below).
- In the position/search box, type in a range of numbers that you would like to go to after “contig60:”, then click “go” (orange box below).



- Practice zooming in and out in the three ways listed above. When you would like to zoom back out completely, the easiest way is to clear all the numbers in the position/search box, just leaving “contig#”.
- Next, let’s zoom in enough to see sequence.
- Enter “contig#:1-100” in the position/search box, and click “go”.
- Scroll down, and set the “Base Position” track to “full”, then click “refresh”
- Now you can see the nucleotide sequence of the contig, starting from base #1

For the STEP 4: Fine Mapping of your gene today you will do the following adjustments

- Add RNA-seq data. Under “RNA Seq Tracks”, go to “RNA-Seq Alignment Summary” and choose “show” from the drop-down menu, then click “refresh”. You will see the sources of *D. ananassae* from which the RNA which has been sequenced (such as for *D. ananassae* sources were: mixed embryos, adult females, and adult males). The higher the peak, the greater the amount of mRNA similar to the sequence in this region was found.
- Also under “Genes and Gene Prediction Tracks”, set “Predicted Splice Sites” on dense. This now highlights high, medium, and low quality splice donor and splice acceptor sites in the DNA sequence. This track will be useful in providing additional evidence for the location of splice junctions.
- Under “Genes and Gene Prediction Tracks”, set all of the following to “dense”: Genscan Genes, Geneid Genes, N-SCAN, SGP Genes, Augustus, SNAP, and GlimmerHMM. These are the *ab initio* gene prediction programs which have been run on this sequence, and the predicted genes are now displayed as rectangles (possible exons) connected by lines (possible introns).
- Click “refresh

Gene annotation Step 4

(50 points)

In the future when annotating, it will be important to display all of the key evidence tracks. Use the screenshot below as a shortcut to get things set up correctly. Some different species have slightly different evidence tracks available. If the evidence on the species that you will annotate differ significantly, consult your instructor about which tracks you should display.

The screenshot displays a genome browser interface with the following sections and tracks:

- Genes and Gene Prediction Tracks** (refresh button):
 - D. mel Proteins (pack dropdown)
 - D. mel Transcripts (hide dropdown)
 - genBlastG Genes (hide dropdown)
 - GeMoMa Genes (hide dropdown)
 - Genscan Genes (dense dropdown)
 - Geneid Genes (hide dropdown)
 - N-SCAN (pack dropdown)
 - N-SCAN PASA-EST (hide dropdown)
 - SGP Genes (hide dropdown)
 - Augustus (dense dropdown)
 - SNAP (dense dropdown)
 - GlimmerHMM (hide dropdown)
 - Gnomon Genes (hide dropdown)
 - Predicted Splice Sites (hide dropdown)
- ChIP Seq Tracks** (refresh button):
 - Histone Modifications (hide dropdown)
- Expression and Regulation** (refresh button):
 - Combined RAMPAGE TSS (hide dropdown)
 - RAMPAGE TSS Read Density (hide dropdown)
 - Eye Discs ATAC-Seq (hide dropdown)
- RNA Seq Tracks** (refresh button):
 - RNA-Seq Coverage (Stages) (full dropdown)
 - RNA-Seq Coverage (Tissues) (hide dropdown)
 - Combined Splice Junctions (hide dropdown)
 - StringTie Transcripts (hide dropdown)
 - Trinity Transcripts (hide dropdown)
- Comparative Genomics** (refresh button):
 - Drosophila Gnomon Transcripts (hide dropdown)
 - D. mel. Chain (hide dropdown)
 - D. mel. Net (hide dropdown)
- Variation and Repeats** (refresh button):
 - Simple Repeats
 - RepeatMasker