

JBrowse Basics

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

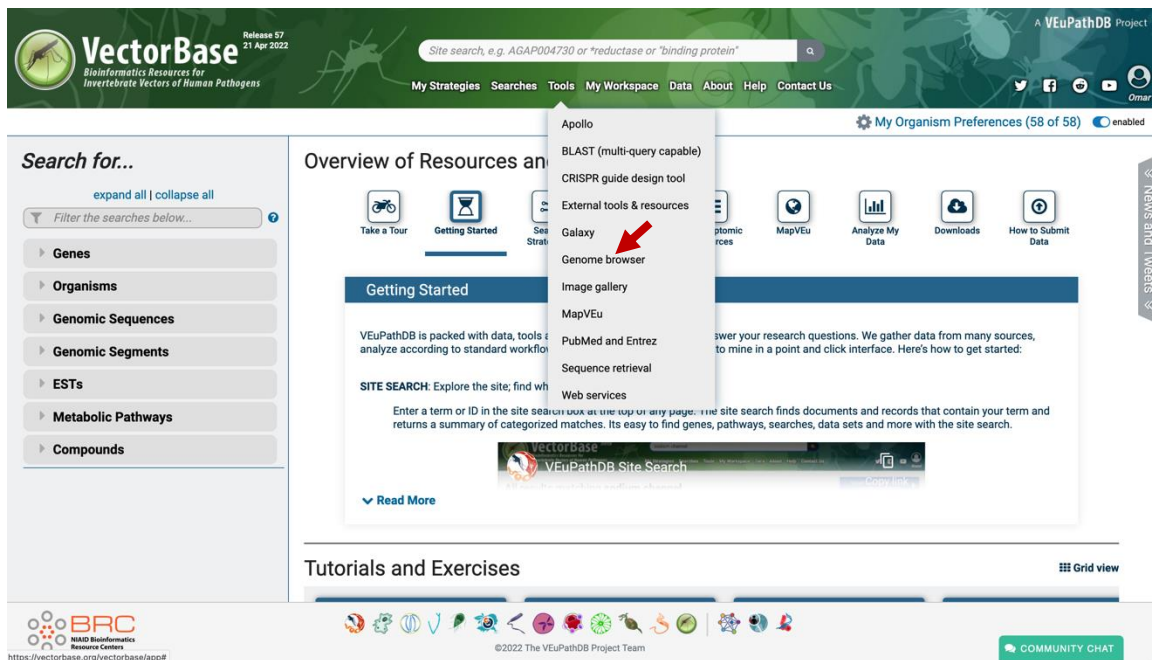
Learning objectives:

- 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

1. **Navigate to the Genome Browser (JBrowse)** JBrowse is a fast and full-featured genome browser built with JavaScript and HTML5. You can read more about JBrowse and its features here: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4830012/>

Links to the genome browser are available from multiple locations:

- a. **The tools menu** in the banner of any page.



- b. **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 “View in JBrowse genome browser” button.

View in JBrowse genome browser

2. **Getting around JBrowse.** Use the tools menu or the View in JBrowse button to open JBrowse. Once in JBrowse examine the following features:
 - a. 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. What do each of these do/provide?
 - b. The **navigation bar**: located below the menu bar. This contains zooming (magnifying glass icons), panning (left/right arrows) and highlighting (yellow highlighter) buttons, reference sequence selector (drop down with sequences from the selected genome sorted by length), a text box to search for features such as gene IDs and overview bar which shows the location of the region in view. Zoom features are also built into the scale on the top of the navigation panel. Select an area to zoom to that location
 - c. The **genome view**: this is where the data tracks are displayed. When viewing the annotation track (top most track), you can move upstream and downstream by dragging the track features left or right.
 - d. **Select tracks**: Click on the “select track” button (top left). This menu contains all the data tracks that are aligned to the genome that you are viewing. The list of tracks can be filtered using the ‘clickable’ left panel categories, or with the search/filter function above the detailed right panel.

Menu

Navigation

Genome Or Data View

Select Tracks

My Tracks

Currently Active

Recently Used

Category

- 1 Comparative Genomics
- 4 Epigenomics
- 3 Gene Models
- 9 Genetic Variation
- 19 Sequence Analysis
- 1,539 Transcriptomics

Subcategory

- 2 (no data)
- 11 Array Probes
- 3 BLAT and Blast Alignments
- 4 ChIP-Seq
- 9 DNA polymorphism
- 1 Orthology and Synteny
- 1,538 RNA-Seq
- 1 Sequence assembly
- 3 Sequence composition, complexity and repeats
- 1 Sequence sites, features and motifs
- 2 Transcripts

Dataset

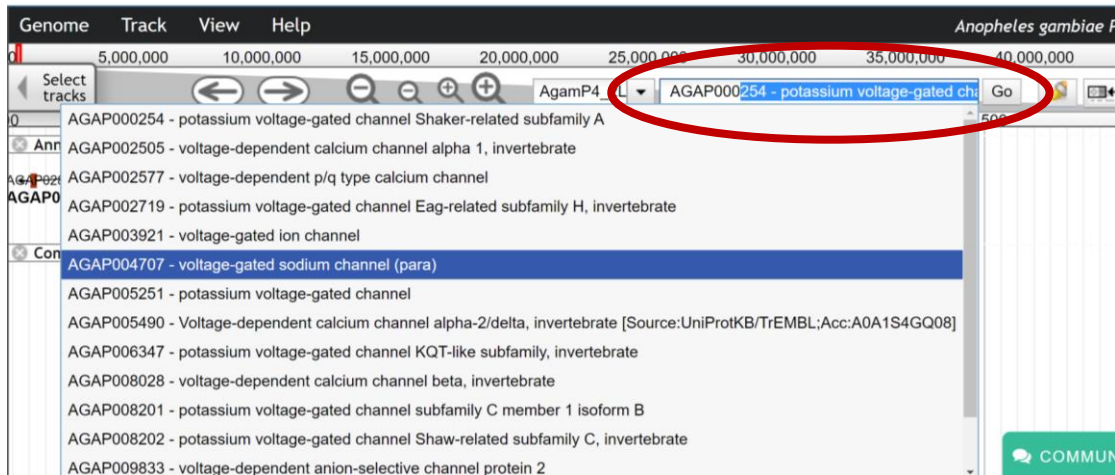
- 14 (no data)
- 20 Aedes aegypti Transcriptome

1,575 tracks

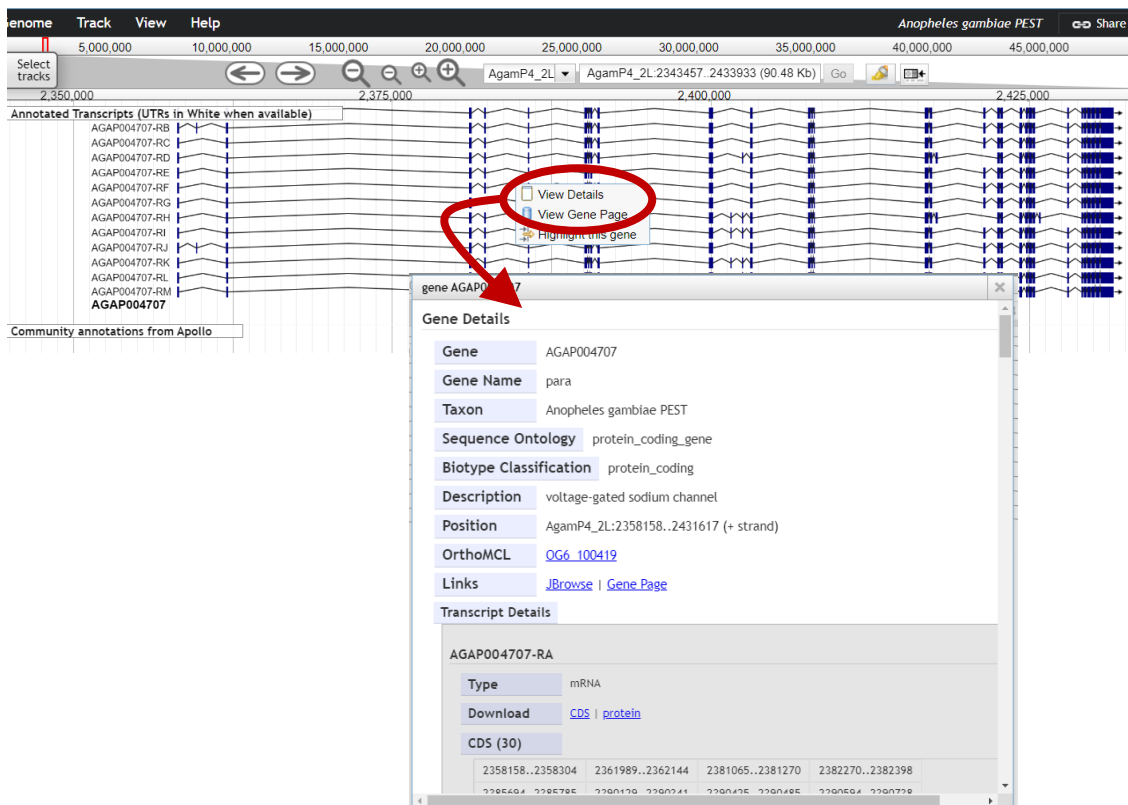
Name	Category	Subcategory	Dataset	Track Type	RNA-Seq Alignment	RNA-Seq Strand
<input type="checkbox"/> Aedes aegypti Transcriptome Sequencing - F11 (non-unique forward) Coverage	Transcriptomics	RNA-Seq	Aedes aegypti Transcriptome Sequencing	Coverage	non-unique	forward
<input type="checkbox"/> Aedes aegypti Transcriptome Sequencing - F11 (non-unique reverse) Coverage	Transcriptomics	RNA-Seq	Aedes aegypti Transcriptome Sequencing	Coverage	non-unique	reverse
<input type="checkbox"/> Aedes aegypti Transcriptome Sequencing - F11 (unique forward) Coverage	Transcriptomics	RNA-Seq	Aedes aegypti Transcriptome Sequencing	Coverage	unique	forward
<input type="checkbox"/> Aedes aegypti Transcriptome Sequencing - F11 (unique reverse) Coverage	Transcriptomics	RNA-Seq	Aedes aegypti Transcriptome Sequencing	Coverage	unique	reverse
<input type="checkbox"/> Aedes aegypti Transcriptome Sequencing - F12 (non-unique forward) Coverage	Transcriptomics	RNA-Seq	Aedes aegypti Transcriptome Sequencing	Coverage	non-unique	forward

Select Tracks menu

3. **Navigate to a specific gene in JBrowse.** The goal of this step is to navigate to the voltage-gated sodium channel gene of *Anopheles gambiae* PEST.
 - a. Make sure the *Anopheles gambiae* PEST genome is selected from the genome menu.
 - b. Start typing the word voltage in the search box. After a few seconds you should see a list of results (do not hit enter). Select the gene called AGAP004707- voltage-gated sodium channel (para) from the search dropdown. Then click GO



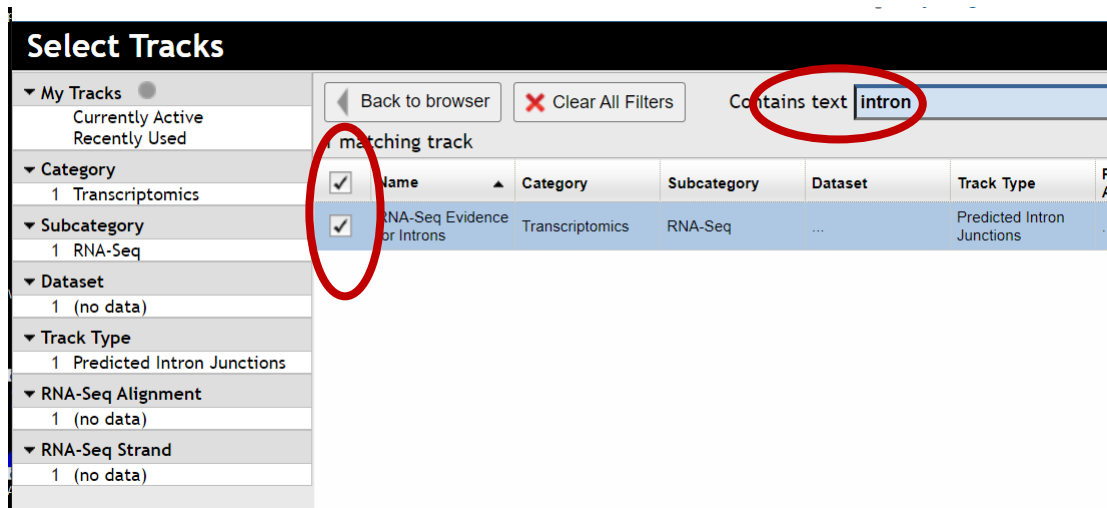
- c. Details about most features are available in pop-up panels. Click (or control click) on the gene feature to view the details panel. You can also right click to choose the same details panel, a link to the gene page, or highlight the gene in yellow. What information is available in the popup?



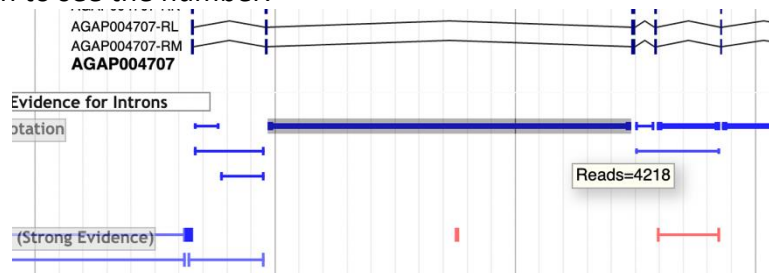
- d. How many isoforms does this gene have? What genes are immediately upstream and downstream of this gene? (Hint: if needed use the zoom out button in the navigation bar). What is the difference between the small and large zoom buttons? (Tip1: another

way to zoom in and out is by clicking on shift and the up or down arrows. What happens if you click shift and left or right arrows? *Tip2:* you can also zoom in by clicking and dragging your cursor in the location ruler in the navigation bar).

4. **Exploring intron evidence.** VectorBase contains many RNAseq datasets which can help you explore the gene model. VectorBase also provides a metric for all possible predicted introns based on the RNAseq data. Turn on the track called “RNA-Seq Evidence for Introns”. To do this click on the “Select Tracks” tab then type the word intron in the search box – this will filter the tracks to the ones that contain the word “intron”. Select this track and close the “Select Tracks” window.



- a. Explore the intron evidence tracks. Notice there are two subtracks turned on by default: Matches Annotation and Unannotated (Strong evidence). What is the difference between these subtracks?
- b. How do you determine the number of reads that support an intron? Mouse over the intron span to see the number.



- c. The track titles also contain dropdown menus with actions or information about the track. Hover over the track title and then click the down arrow that appears to access more track functions.
- d. What happens if you click on one of the intron spanning reads? Notice the popup which contains a lot of information about this intron and the experiments that support it.

AgamP4_2L_2362145_2381064_0 details

Intron Junction Details

Intron Location

AgamP4_2L:2362145..2381064 (+ strand)

Intron Spanning Reads (ISR)

4218

ISR per million (ISRPM)

364.34

Gene assignment

AGAP004707 - annotated intron

% of Most Abundant Intron (MAI)

33.67

Sample Details

Experiment	Sample	Unique	ISRPM	ISR/Cov	% MAI
Antennae vs maxillary palps (An. gambiae)	Female_antenna	6	1.45	0.11	11.12
	Female_maxillary_palp	5	1.36	0.18	19.21
Antennal transcriptome expression profiles following a blood meal	Blood_fed_+36h	1	0.23	0.11	6.17
	Non-blood_fed_+1h	2	0.84	0.27	22.28
	Non-blood_fed_+24h	8	2.85	0.92	66.59
	Non-blood_fed_+36h	17	5.39	1.05	73.94
	Non-blood_fed_+48h	5	1.74	0.49	35.8
Chemosensory appendages, male and female	Female_antenna	5	5.87	0.23	9.61
	Female_maxillary_palp	2	4.05	0.07	4.88

e. Are there any introns in the “unannotated (strong evidence)” subtrack that have good support? Hint: the darker the color, the stronger the support.

5. **Examining synteny tracks.** JBrowse in VEuPathDB includes a configurable track containing sequence alignment between genomes where genes are shaded between genomes based on orthology. This provides a nice way to examine co-linearity between genomes. In this example we will set up the synteny track to display synteny between *Anopheles gambiae* PEST and *Drosophila melanogaster* iso-1.

a. Turn on the track called “Syntenic Sequences and Genes (Shaded by Orthology)”

Back to browser

Clear All Filters

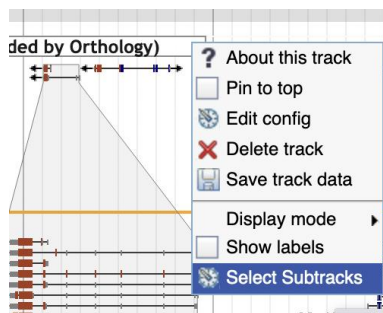
Contains text

synten

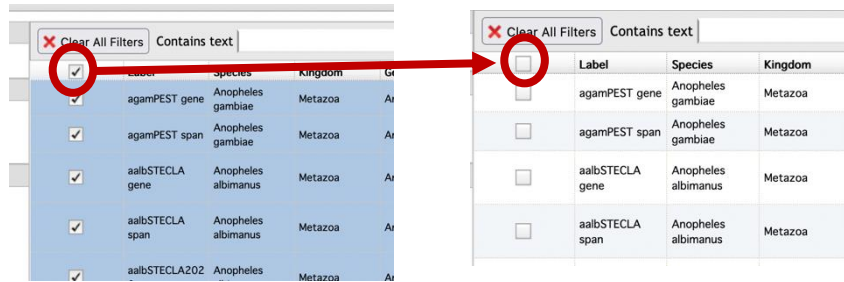
1 matching track

<input checked="" type="checkbox"/>	Name	Category	Subcategory	Dataset	Track Type
<input checked="" type="checkbox"/>	Syntenic Sequences and Genes (Shaded by Orthology)	Comparative Genomics	Orthology and Synteny	...	Segments

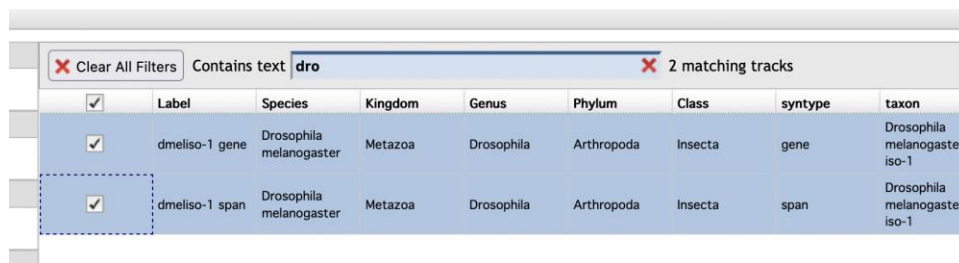
b. Select synteny subtracks by clicking on the down arrow on the track name and selecting “Select subtracks”.



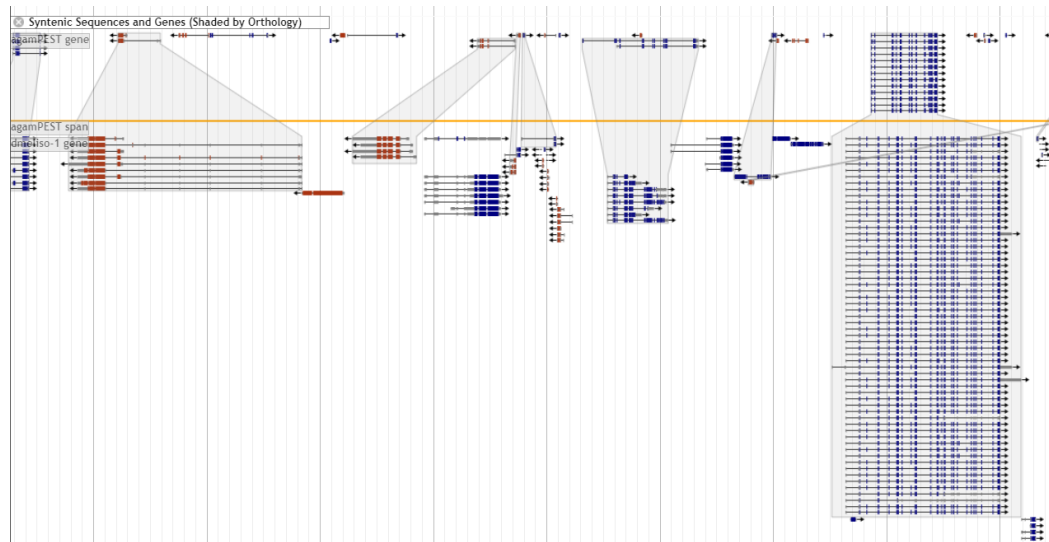
c. Unselect all the tracks (the easiest way is to use the top check box to select all then unselect all)



- d. Select the tracks for *Anopheles gambiae* PEST and *Drosophila melanogaster* iso-1 and then click SAVE at the bottom of the Select Subtracks popup. (note you can use the search box at the top to find your organism of interest).



- e. View the track. Does *Drosophila* have an ortholog of the *Anopheles* sodium channel? Does it have a similar number of isoforms or does it have many more?
- f. What does synteny looked like around this gene? Zoom out a bit (this might be slow to load). Is co-linearity relatively preserved between these two species?

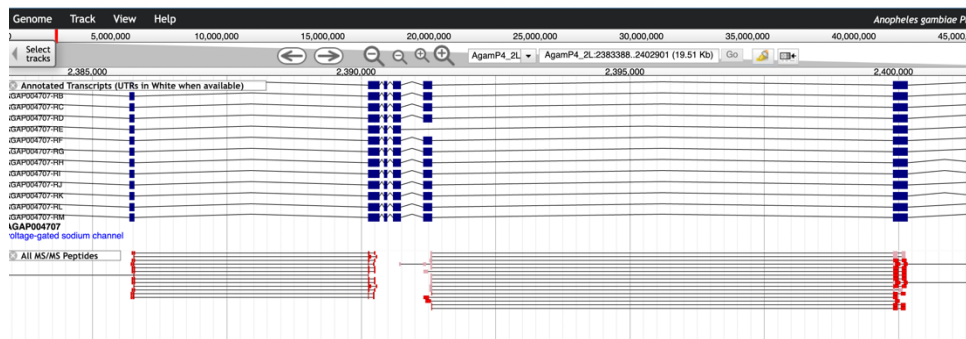


6. **Adding more data tracks.** JBrowse contains many data tracks that you can load by selecting them from the select tracks tab. Each track represents sequence data or features aligned to

- Load the track called All MS/MS Peptides. This track displays all peptides from mass spectroscopy experiments we have in the database mapped to the genome.

Select Tracks								
* My Tracks		Back to browser Clear All Filters		Contains text	mass	6 matching tracks		
Currently Active Recently Used		<input type="checkbox"/>	Name	Category	Subcategory	Dataset	Track Type	RNA-Seq Alignment
* Category	6 Proteomics	<input checked="" type="checkbox"/>	All MS/MS Peptides	Proteomics	Protein Expression	...	Segments	...
* Subcategory	Protein Expression	<input type="checkbox"/>	Antennae and total head appendages (THA; maxillary palps, antennae, and proboscises) MS/MS Peptides Rund et al 2013	Proteomics	Protein Expression	Antennae and total head appendages (THA; maxillary palps, antennae, and proboscises)	Segments	...
* Dataset	1 (no data)	<input type="checkbox"/>	Brain Proteomics MS/MS Peptides Duvet et al 2014	Proteomics	Protein Expression	Brain Proteomics	Segments	...
1	Antennae and total head appendages (THA; maxillary palps, antennae, and proboscises)	<input type="checkbox"/>	Head samples: 1) whole head including all appendages (maxillary palps, proboscises and antennae); 2) only antenna; 3) only eyes MS/MS Peptides Champion et al 2016	Proteomics	Protein Expression	Head samples: 1) whole head including all appendages (maxillary palps, proboscises and antennae); 2) only antenna; 3) only eyes	Segments	...
1	Brain Proteomics	<input type="checkbox"/>	Immature stages (larvae & pupae) and various tissues from adults (male and female) MS/MS Peptides Charnay et al 2011	Proteomics	Protein Expression	Immature stages (larvae & pupae) and various tissues from adults (male and female)	Segments	...
1	Head samples: 1) whole head including all appendages (maxillary palps, proboscises and antennae); 2) only antenna; 3) only eyes	<input type="checkbox"/>	Pupal cuticles and larval head capsules MS/MS Peptides He et al 2007	Proteomics	Protein Expression	pupal cuticles and larval head capsules	Segments	...
* Track Type	6 Segments							
* RNA-Seq Alignment	6 (no data)							
* RNA-Seq Strand	6 (no data)							

b. Does the voltage-gated channel gene have mapped peptides? You may want to zoom in to a region with peptide evidence to see more details.



7. **(Optional) Retrieving sequence data from JBrowse** Sequence data from a region of interest can be downloaded from JBrowse in FASTA format. Download is a function of the Reference Sequence track, so that must be turned on. And the region to download must be highlighted in yellow.

- Make sure the “annotated transcripts” and the “reference sequence” tracks are turned on.

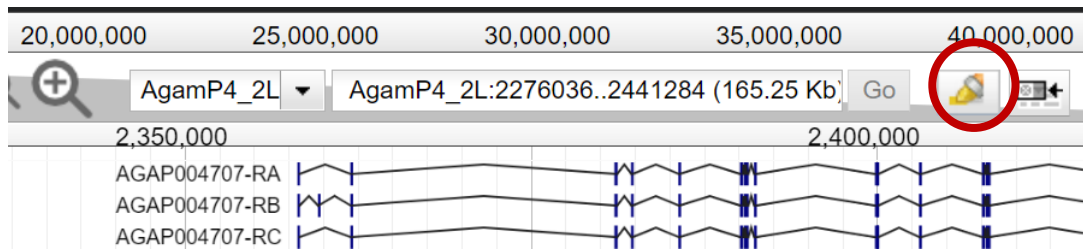
Back to browser

✖ Clear All Filters

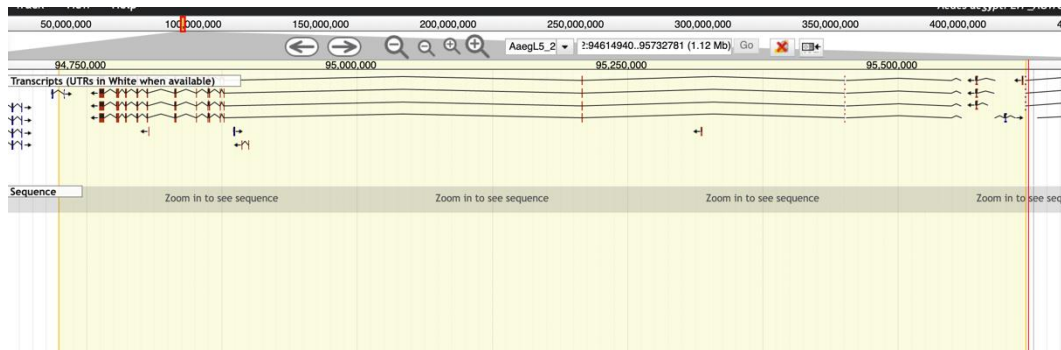
Contains text ✖ 1

<input type="checkbox"/>	Name	Category	Subcategory	Dataset	Track Type	RNA-Seq Align
<input type="checkbox"/>	Reference Sequence	Sequence Analysis	Reference Sequence	...

b. Click on the “highlight a region” button in the navigation bar. It should turn yellow when activated.



- c. Click and drag in the genome view region and select the area you would like to highlight.



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

