

Browsing Genomes II: Visualizing and adding tracks

The screenshot shows the VectorBase website interface. At the top, there is a navigation bar with the VectorBase logo and the text "Bioinformatics Resource for Invertebrate Vectors of Human Pathogens". Below this, a dropdown menu for "Anopheles gambiae" is visible, and a location bar shows "Location: 2L:39,221,590-39,222,203". The main content area is divided into two panels. The left panel, titled "Location-based displays", contains a tree view with options like "Whole genome", "Chromosome summary", "Region overview", "Region in detail" (which is selected), "Comparative Genomics", "Genetic Variation", "Markers", and "Other genome browsers". The right panel, titled "Chromosome 2L", shows a map of the chromosome with a region of interest highlighted. Below the map, there is a section titled "Region in detail" which includes a table of "Assembly exceptions" and a "Configure this page" button.

Location tab: from here you can access:

- Various views for this region:
 - Whole chromosome
 - Chromosome summary
 - Region overview
 - Region detail
- Comparative data
- Genetic variation and marker data (if available)
- You can also configure the page to visualize data from VectorBase, external sources and your own

The following exercises focus on how to configure the page to visualize different types of data.

1. Go to VectorBase search box and type this *A. aegypti* gene ID: AAEL006498.

The screenshot shows the VectorBase search bar. The search bar is located at the top right of the page, with the text "AAEL006498" entered. To the right of the search bar is a "GO" button. Below the search bar, there is a "LOGIN" button.

Click on "Genome" or in the "top hit" in the results, both links are the same and will take you to the Genome Browser.

The screenshot shows the VectorBase search results page. The search bar at the top contains the text "AAEL006498" and a "GO" button. Below the search bar, there is a "Filter Results" table and a "Search results" section. The "Filter Results" table has two columns: "Domain" and "Hits". The "Search results" section shows "Showing 1 to 20 of 23 results found." and a list of results. The first result is "GPROP1 (AAEL006498)" with a description: "long wavelength sensitive opsin".

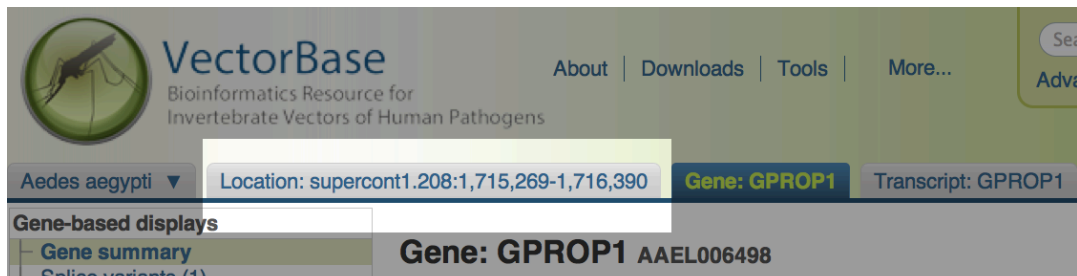
Domain	Hits
Expression	18
Transcriptome	3
Comparative	1
Genome	1
Species	Hits

Search results
Showing 1 to 20 of 23 results found.

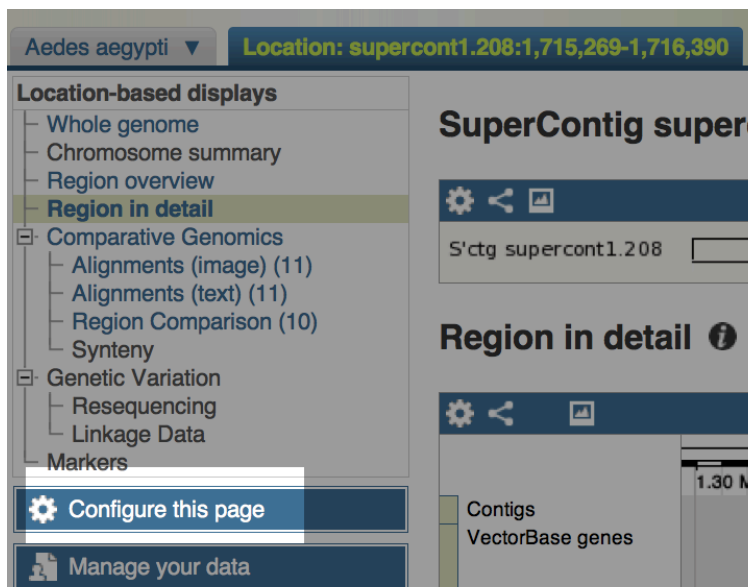
GPROP1 (AAEL006498)
Description: long wavelength sensitive opsin

2. Activate tracks:

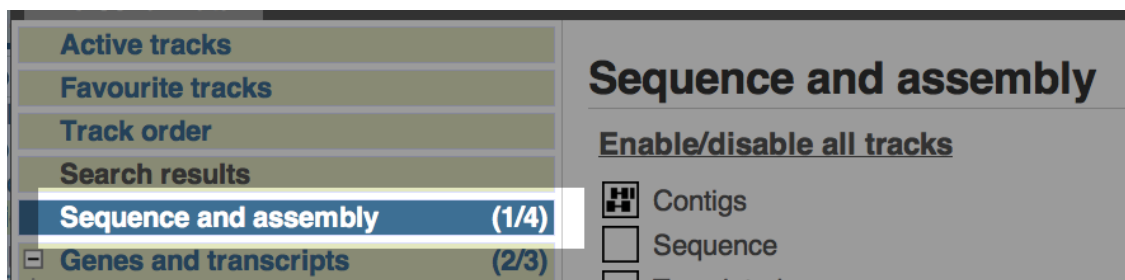
- Once in the genome browser, go to the “Location” tab



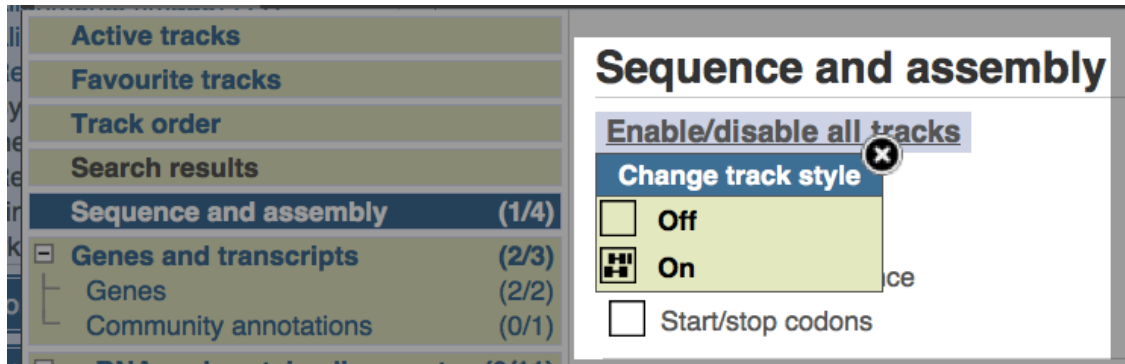
- Click on “Configure this page”



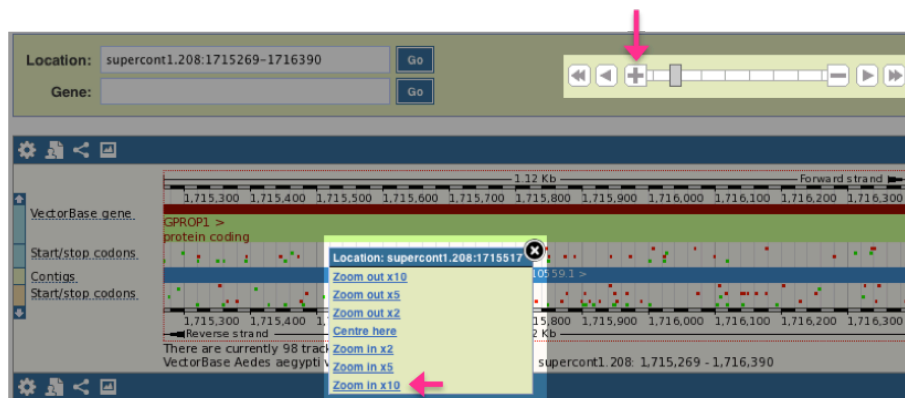
- On the pop-out window click on “sequence and assembly”, located on the left hand side menu



- Enable all four tracks by clicking on the empty squares next to each track name or click on “Enable/disable all tracks” and click on ON.



- To save and close click on the check mark, on the top right hand side
- Scroll down to the bottom of the page.
- Zoom in until you see the nucleotide and amino acid sequences. **Hint:** Click on the white area below the gene and on the pop out window click “zoom in x10”. Later, click on the “zoom in” plus icon.



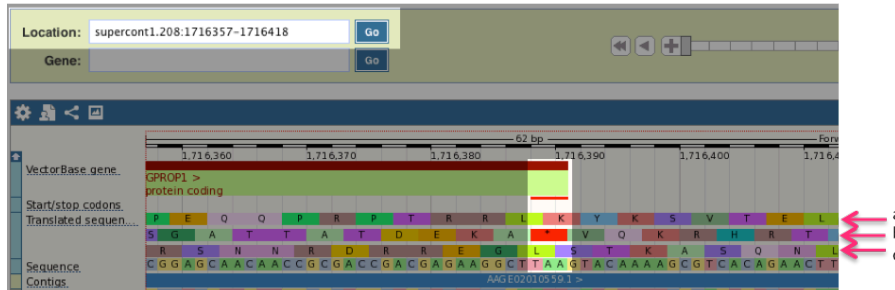
3. What is the open reading frame (ORF) of this gene, a, b or c?

- **Hint:** To see the **beginning** of the gene type this region in the location box: supercont1.208:1715241-1715302



Which one is the START codon in nucleotide and amino acids?

- **Hint:** To see the **end** of the gene, type this region in the location box: supercont1.208:1716357-1716418



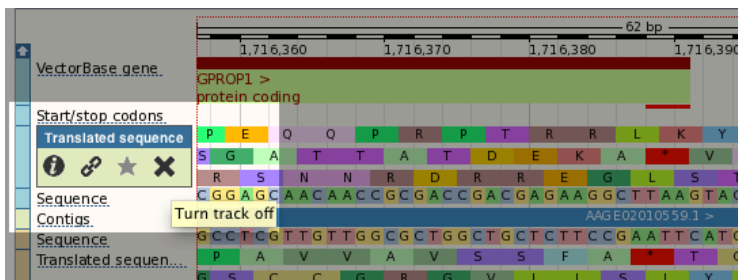
Which one is the STOP codon in nucleotides? Which are the other two possible STOP codons?

4. Type this region in the location box: supercont1.208:1715800-1715860. What do the red and green lines show in the “Start/Stop codons” track?

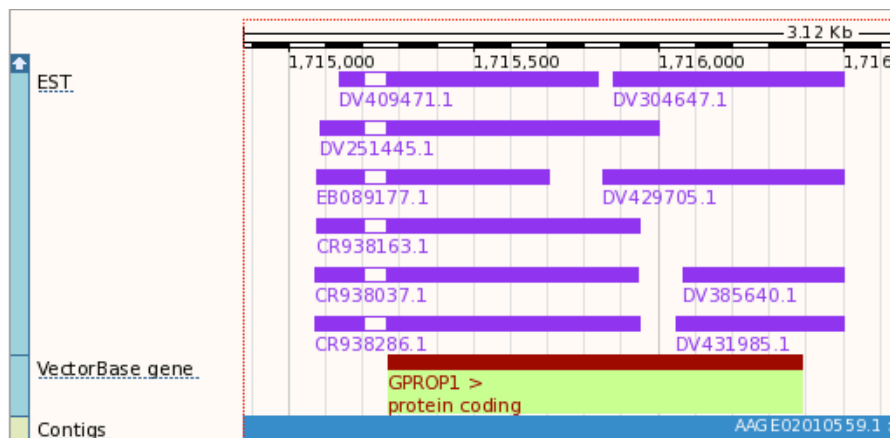
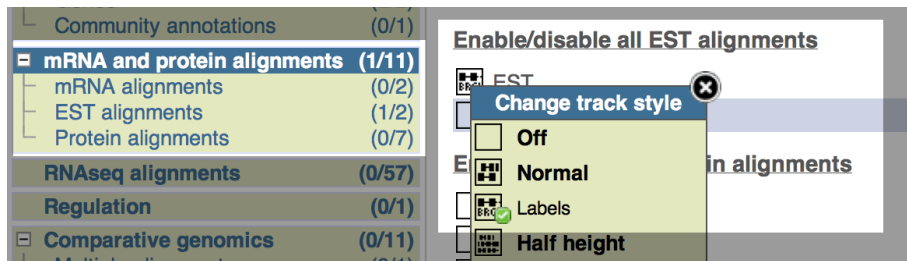


5. Turn OFF the tracks “Sequence”, “Translated sequence” and “Start/Stop codons”.

- Hover over these track names
- On the pop out window click on the “x mark”



- Go to this region supercont1.208:1714876-1718000 and activate the EST track with labels.

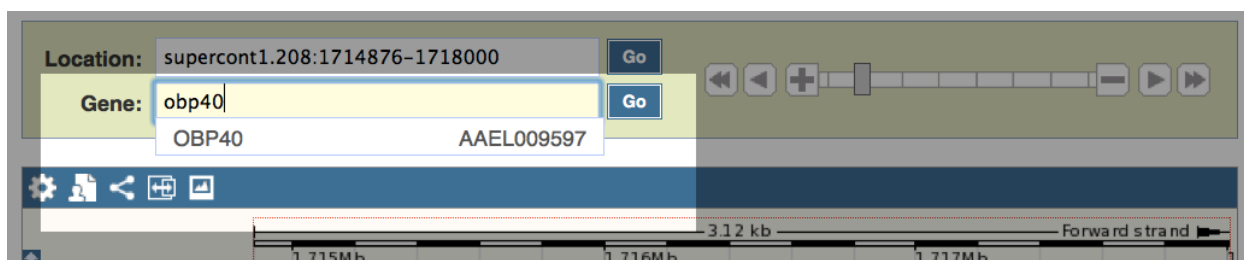


- What the EST labels mean? **Hint:** Click on one EST and follow the link. According to these ESTs what can you infer about this gene?

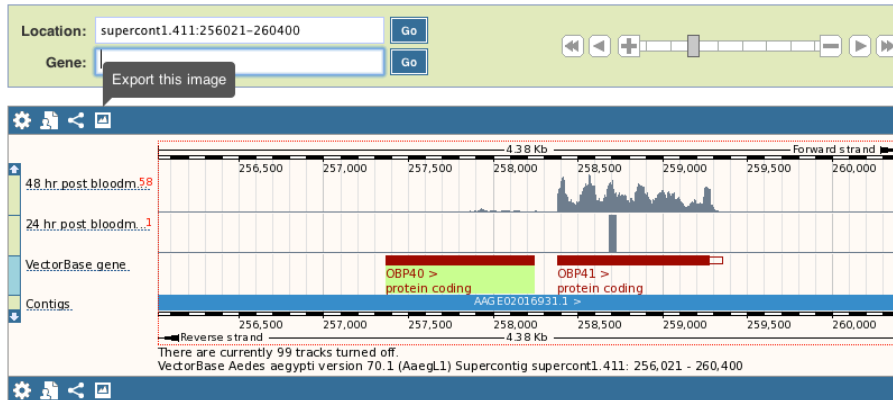
- Turn OFF the EST tracks.

7. Use the search box to find tracks:

- Type “OBP40” (or the gene ID) and click “Go” in the Gene search box.



- Click on “Configure this page”
- Type “blood” in the search box of the pop out window
- Select “coverage only” for the track style of these two “RNAseq alignments” tracks
 - ✓ 24 hr post blood meal female carcass
 - ✓ 48 hr post blood meal female carcass
- To save and close click on the check mark, on the top right hand side
- In the location box type: supercont1.411:256021-260400



- According to this RNAseq data what can you infer about these two genes?

8. RNAseq available tracks:

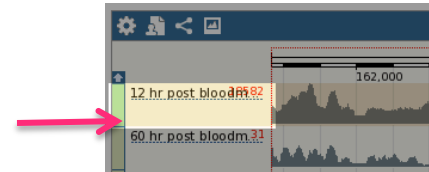
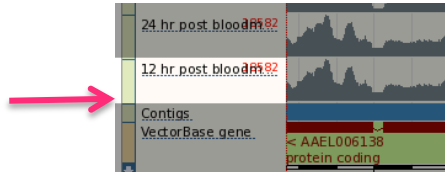
- Currently only available for eight *Anopheles* species and *A. aegypti* LVP genomes
 - How many RNAseq tracks are available for these two genomes? **Hint:** you can also go to this page for answers: www.vectorbase.org/rna-seq-data-sets
-

9. Do not close the genome browser. In another window go to the FAQ page www.vectorbase.org/faqs and click on “What styles are available for the data tracks in the Genome Browser? Keep this new page open to answer questions 10, 11 and 12.

10. To organize the tracks there are two options:

Option 1:

- Select, with a click, the light green bar of the track of interest and move it to your preferred location.



What is option 2?

Hint: In the FAQ page scroll to the bottom of the page.

11. Including the one above two are three options to close tracks. Describe these two options using two bullet points for each.

Option 1:

-
-

Option 2:

-
-

Hint: In the FAQ page look at the top six images and its legends.

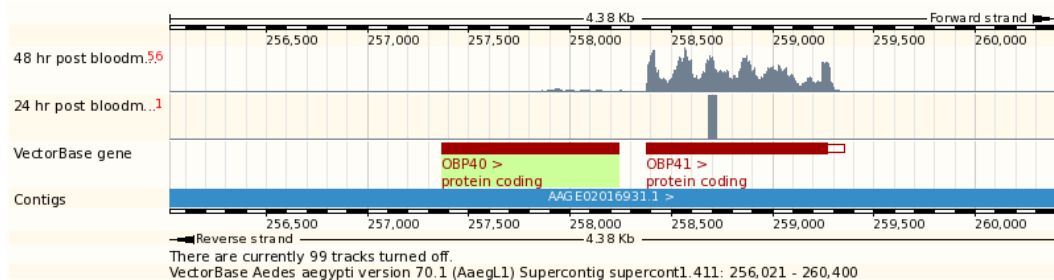
12. Track styles. Do all tracks have the same styles? Mention at least three track styles.

Hint: Skim through the In the FAQ page.

13. Export and publish images.

- Click on the “landscape” icon, when you hover over it, it reads “Export this image” (shown on the image of question 7)
- Select “export as PNG” and select “download” (the arrow icon)
- Insert the PNG into a word document.
- You are encouraged to use these figures for your gene or genes of interest, for your thesis or paper. Please do not forget to cite VectorBase’s most recent paper, in the text and/or the caption of the figure.

Megy, K. et al. (2012) VectorBase: improvements to a bioinformatics resource for invertebrate vector genomics. *Nucleic Acids Research*. 40: D729-734.



Note: Remember that answers may change because VectorBase is updated every two months (www.vectorbase.org/releases).