Exploring the Gene Page

Note: *this exercise uses* ***VectorBase.org*** *as an example database, but the same functionality is available on all VEuPathDB resources.*

What is a Gene Page on VectorBase?

- The gene page conveniently consolidates links to all the data available on VectorBase for a particular gene into a single page

Learning objectives

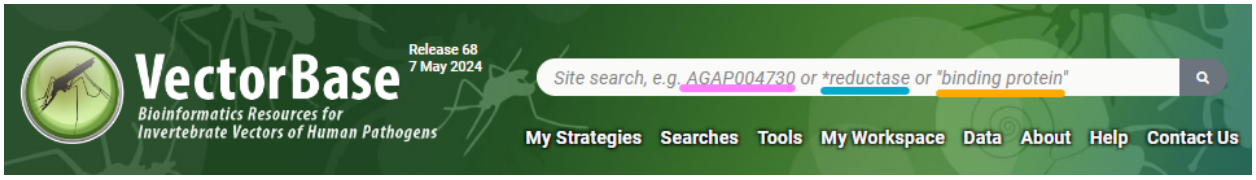
Gene pages:

- Become familiar with the information on gene pages
- Navigate to and from the gene pages
- Use the contents section of the gene page
- Interact with gene page subsections

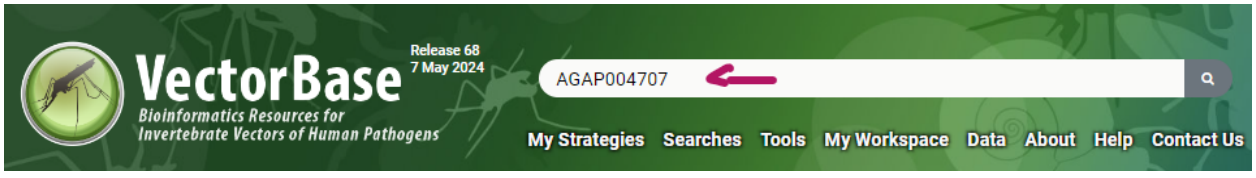
1. Navigation to the Gene pages

For this exercise visit the gene page for AGAP004707 (voltage-gated sodium channel). How can you do it? Hint: Go to the **Site Search** box, at the top of the website. Notice you can use different types of keywords:

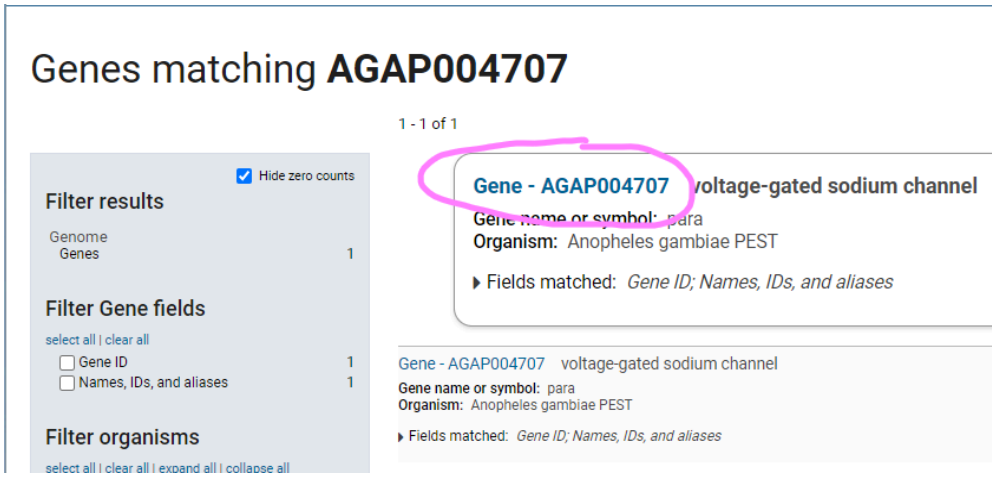
- gene ID
- gene name or function, with a wild card/asterisk
- gene name or function, with quotation marks



Type the gene ID: AGAP004707

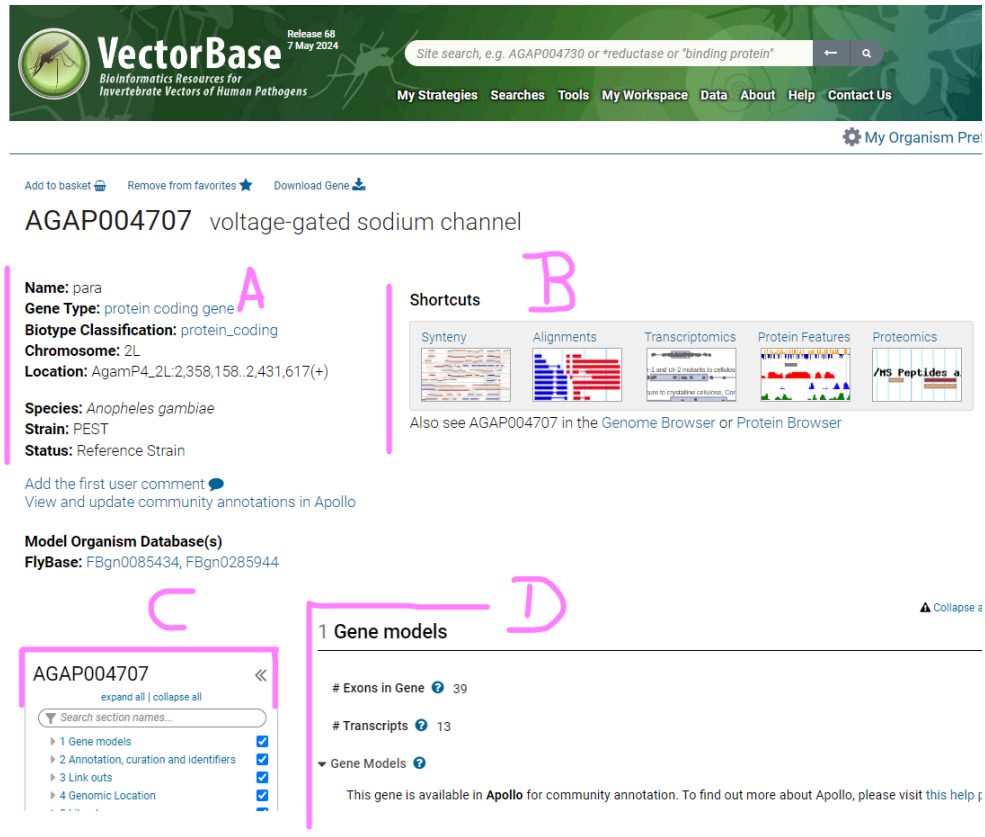


Click on the gene ID on the results page



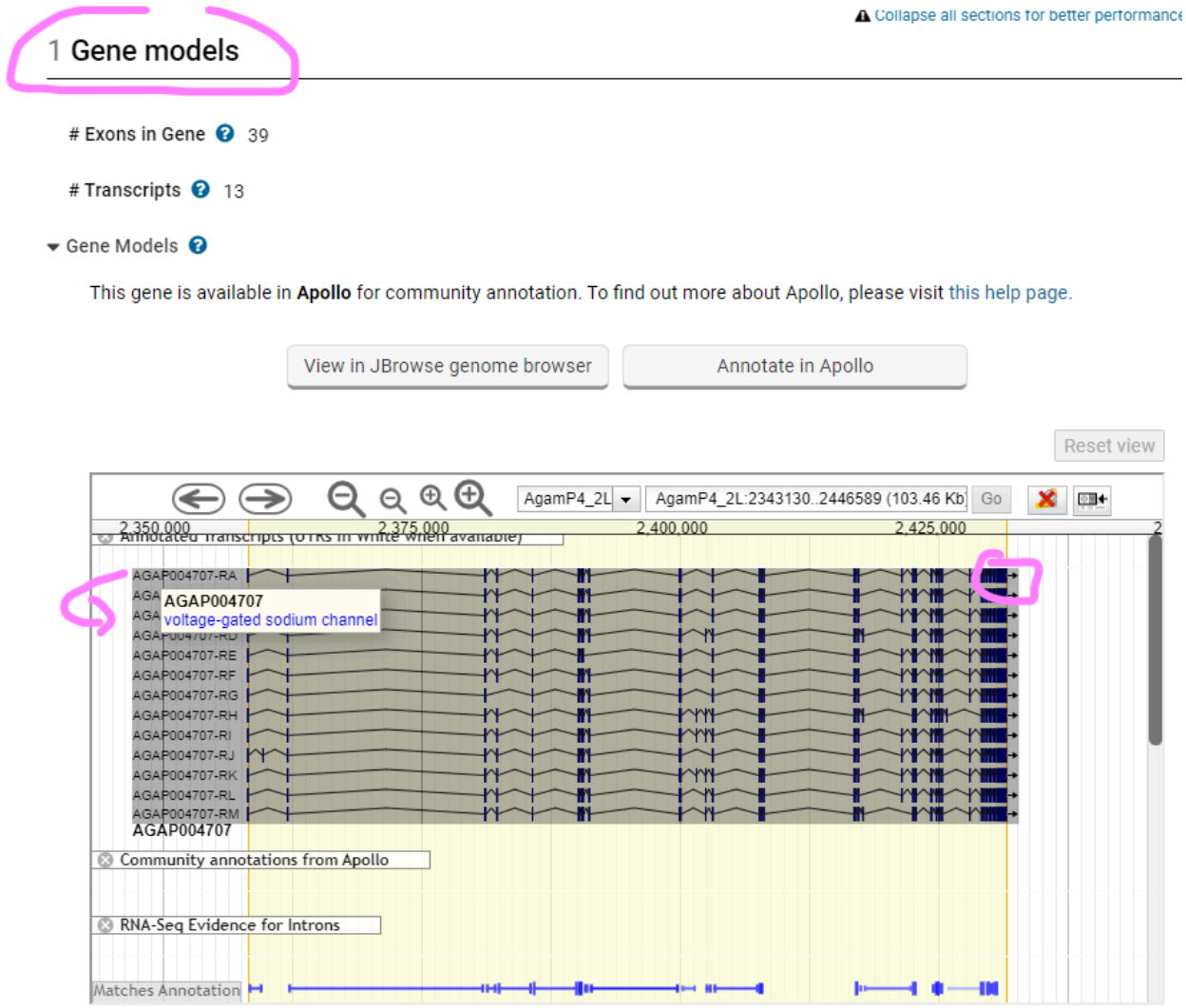
2. Explore the layout of the page

a. What information is in the top left (section A in the image below)?
   i. Can you easily find which chromosome this gene is located on?
   ii. Is this gene protein coding?
b. What do the shortcuts do? (section B)
c. Where can you find the contents of this page? (Section C)
d. Scroll down and examine what section D contains.

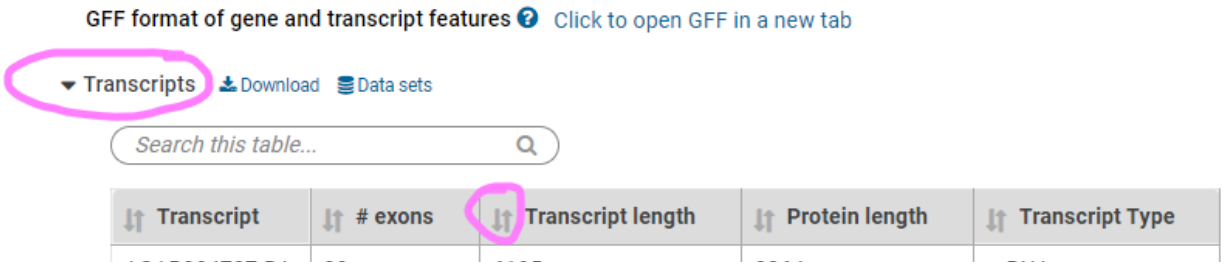


3. Explore the gene model section. Scroll down to the gene model section of the gene page.

a. What direction is the transcript relative to the chromosome?
b. Does the gene have UTRs?
c. How many exons does the gene have?
d. Does this gene have any available community annotation from Apollo?

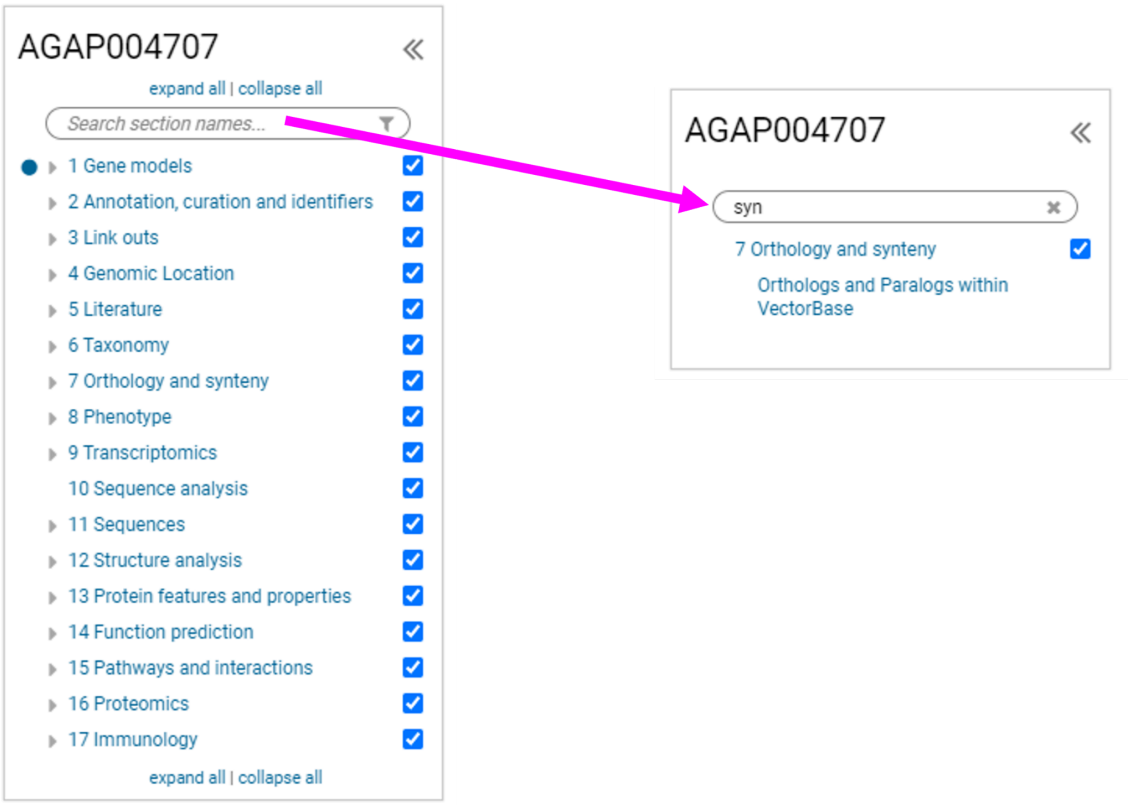


e. How long is the longest transcript? You can determine transcript length by expanding the Transcripts table. Click on the appropriate column header to sort the rows



3. Content navigation

How do you find/navigate to the different sections of the page? Use the "Table of contents'' menu on the left side. Gene page content is organized by data type and the section titles serve as links to data within the page. When expanded, each section reveals more navigation links. The content menu can also be filtered using the search function as shown below. Begin typing the 'synteny' in the filter to collapse the content menu.

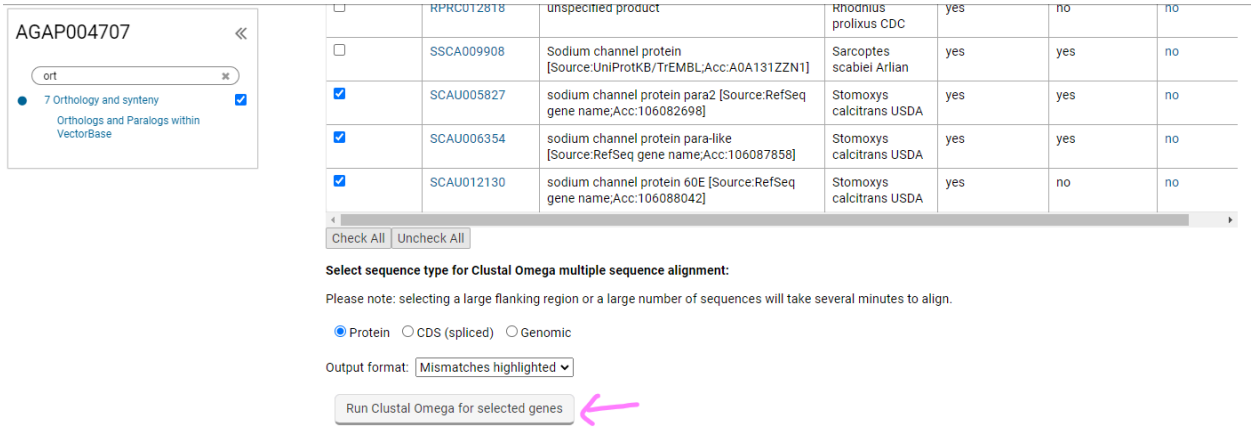


a. Click to navigate to the **Synteny** section. Does *Anopheles gambiae* share synteny in this region? If you get an error message reload the page
b. Navigate to the **Transcript Expression** table and open the row for the experiment called "Chemosensory appendages, male and female" (Notice the filter box at the top of the table)

i. What data type was used to produce this data (microarray or RNAseq)?
ii. Is the expression level equal in male and female maxillary palps? What about male and female bodies and antennae?
iii. Open the Data table for this experiment. What is the TPM expression value for unique reads mapped to female antenna?

4. Running an alignment of selected sequences

a. Expand the "**Orthologs and Paralogs within VectorBase**" section.
b. Select a few genes from the table using the checkbox.
c. Scroll to the bottom of the table and click on the Run Clustal Omega button.



5. Explore other sections of the gene page.

Feel free to scroll around the gene page and ask questions for clarification. Here are some questions you may want to ask about the following two genes:

Gene AGAP001633

a. Is there any mass spec evidence for this gene? (hint: go to the Mass Spec. Expression Evidence table in the **Proteomics** section).
b. In the Mass Spec. Expression Evidence in Protein Browser (**Proteomics** section), What do the different color lines mean in the protein browser?

Gene AAEL012109

c. Is there any microarray or RNA-seq evidence for AAEL012109?
d. Search for "alphaFold" within the section names navigation pane. What data can you obtain in this section? (hint: it will take you to the **Structure Analysis** section)