

Exploring the Gene Page

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

Learning objectives

Gene pages:

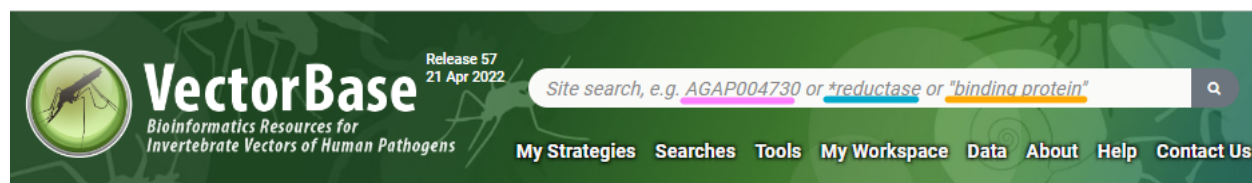
- Become familiar with the information in gene pages
- Navigate to and from the gene pages
- Use the contents section of the gene page
- Interact with gene page subsections

1. Navigation to the Gene pages

For this exercise visit the gene page for AGAP004707 (voltage-gated sodium channel). How can you do it? 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 in the results page

Genes matching **AGAP004707**

1 - 1 of 1

Filter results

Genome
Genes 1

Filter Gene fields

select all | clear all

☐ Gene ID 1

☐ Names, IDs, and aliases 1

Filter organisms

select all | clear all | expand all | collapse all

Gene - AGAP004707 voltage-gated sodium channel

Gene name or symbol: para

Organism: Anopheles gambiae PEST

► Fields matched: Gene ID; Names, IDs, and aliases

Gene - AGAP004707 voltage-gated sodium channel

Gene name or symbol: para

Organism: Anopheles gambiae PEST

► Fields matched: Gene ID; Names, IDs, and aliases

2. Explore the top section of the gene page

- What information is in this section?
- Can you find which chromosome this gene is located on?
- Is this gene protein coding?
- What do the shortcuts do?

Remove from basket Remove from favorites Download Gene

AGAP004707 voltage-gated sodium channel

Name: para
Gene Type: protein coding gene
Biotype Classification: protein_coding
Chromosome: 2L
Location: AgamP4_2L:2,358,158..2,431,617(+)
Species: *Anopheles gambiae*
Strain: PEST
Status: Reference Strain

Shortcuts

Synteny

Alignments

Transcriptomics

Protein Features

Proteomics

Also see AGAP004707 in the [Genome Browser](#) or [Protein Browser](#)

3. Explore the Gene models section

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

- What direction is the transcript relative to the chromosome?
- Does the gene have UTRs?
- How many exons does the gene have?
- Does this gene have any available “Community annotations from Apollo”?

1 Gene models

Exons in Gene 39

Transcripts 13

▼ Gene Models

This gene is available in **Apollo** for community annotation. To find out more about Apollo, please visit [this help page](#).

[View in JBrowse genome browser](#) [Annotate in Apollo](#)

scroll and zoom

AGAP028437 unspecified product

Community annotations from Apollo

- What is the length of this gene longest transcript? You can determine transcript length by expanding the Transcripts section. Click on the appropriate column header to sort the rows

GFF format of gene and transcript features [Click to open GFF in a new tab](#)

▼ **Transcripts** [Download](#) [Data sets](#)

Transcript	# exons	Transcript length	Protein length	Transcript Type

4. Content navigation

How do you find/navigate to the different sections of the page? Use the "Table of contents" menu on the left side, type a keyword and click on the menu, click on the work to navigate to it on the page. In the example below the word "synteny" is used. You can also click on the images in the Shortcuts section in the top of the page.

AGAP004707

expand all | collapse all

- 1 Gene models ☒
- 2 Annotation, curation and identifiers ☒
- 3 Link outs ☒
- 4 Genomic Location ☒
- 5 Literature ☒
- 6 Taxonomy ☒
- 7 Orthology and synteny ☒
- 8 Phenotype ☒
- 9 Transcriptomics ☒
- 10 Sequence analysis ☒
- 11 Sequences ☒
- 12 Structure analysis ☒
- 13 Protein features and properties ☒
- 14 Function prediction ☒
- 15 Pathways and interactions ☒
- 16 Proteomics ☒
- 17 Immunology ☒

expand all | collapse all

AGAP004707

7 Orthology and synteny ☒

Orthologs and Paralogs within VectorBase

5. Running an alignment of selected sequences

- Expand the "Orthologs and Paralogs within VectorBase" section.
- 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.

AGAP004707 <<

ort x

● 7 Orthology and synteny ☒

Orthologs and Paralogs within VectorBase

<input type="checkbox"/>	RPRC012818	unspecified product	Rhodnius prolixus CDC	yes	no	no
<input type="checkbox"/>	SSCA009908	Sodium channel protein [Source:UniProtKB/TrEMBL;Acc:A0A131ZZN1]	Sarcoptes scabiei Arlian	yes	yes	no
<input checked="" type="checkbox"/>	SCAU005827	sodium channel protein para2 [Source:RefSeq gene name;Acc:106082698]	Stomoxys calcitrans USDA	yes	yes	no
<input checked="" type="checkbox"/>	SCAU006354	sodium channel protein para-like [Source:RefSeq gene name;Acc:106087858]	Stomoxys calcitrans USDA	yes	yes	no
<input checked="" type="checkbox"/>	SCAU012130	sodium channel protein 60E [Source:RefSeq gene name;Acc:106088042]	Stomoxys calcitrans USDA	yes	no	no

Check All Uncheck All

Select sequence type for Clustal Omega multiple sequence alignment:

Please note: selecting a large flanking region or a large number of sequences will take several minutes to align.

☒ Protein ☐ CDS (spliced) ☐ Genomic

Output format: Mismatches highlighted

Run Clustal Omega for selected genes
←

6. Explore expression evidence of the gene page.

- Go to the gene page of AGAP008802
- Is there any RNA-Seq data available for this gene? Go to the Transcriptomics section and expand the subsections called RNA-Seq transcription summary and Transcript Expression. Are there any studies comparing males vs. females?
- Is there any mass spec. data available for this gene? Go to the Proteomics section. What do the different color lines mean in the genome browser? trans

Feel free to scroll around the gene page
and ask questions for clarification!