

Exploring the Gene Page

Note: this exercise uses ToxoDB (<https://ToxoDB.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 TGME49_222020 (phosphoglycerate kinase PGKII). How did you get to this gene? (hint: copy and paste the ID in the site search, then click on the gene in the results).

Genes matching **TGME49_222020**

1 - 1 of 1

Filter results: Hide zero counts

Genome: Genes 1

Filter Gene fields: select all | clear all

Gene ID: 1

Filter organisms: select all | clear all | expand all | collapse all

Gene - **TGME49_222020** phosphoglycerate kinase PGKII

Gene name or symbol: PGKII

Organism: Toxoplasma gondii ME49

Fields matched: Gene ID

2. Explore the top section of the gene page

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

Add to basket | Add to favorites | Download Gene

TGME49_222020 phosphoglycerate kinase PGKII

Name: PGKII

Type: protein coding gene

Chromosome: II

Location: TGME49_chrlI:761,396..767,399(+)

Species: Toxoplasma gondii

Strain: ME49

Status: Reference Strain

Add the first user comment | View and update community annotations in Apollo

Shortcuts

Synteny | Alignments | Phenotype | SNPs | Transcriptomics | Protein Features | Proteomics

Also see TGME49_222020 in the Genome Browser or Protein Browser

COMMUNITY CHAT

3. Explore the gene model 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 an available community annotation?
- How long is the transcript? You can determine transcript length by expanding the

TGME49_222020

expand all | collapse all

Search section names...

- 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 Genetic variation
- 10 Transcriptomics
- 11 Sequences
- 12 Sequence analysis
- 13 Structure analysis
- 14 Protein features and properties
- 15 Protein targeting and localization
- 16 Function prediction
- 17 Pathways and interactions
- 18 Proteomics
- 19 Immunology

expand all | collapse all

1 Gene models

Exons in Gene 6

Transcripts 1

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

Community annotations from Apollo

RNA-Seq Evidence for Introns

Matches Annotation

Reads=1715 Reads=1304 Reads=2160 Reads=1532 Reads=917 Reads=1147 Reads=3209 Reads=1080

Reads=2127 Reads=1033 Reads=808 Reads=904 Reads=2375 Reads=2084 Reads=739 Reads=840 Reads=2472 Reads=2994 Reads=109

COMMUNITY CHAT

GFF format of gene and transcript features Click to open GFF in a new tab

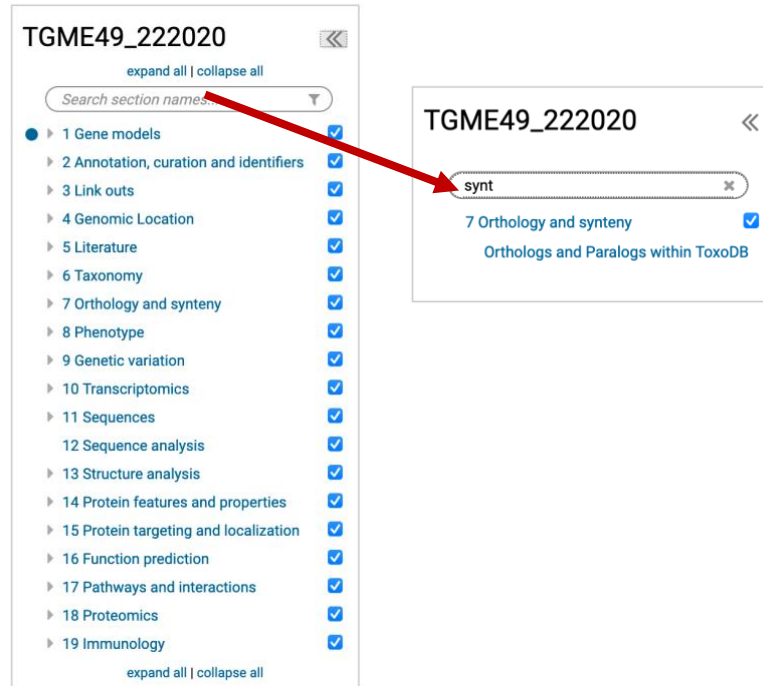
Transcripts Download Data sets

Transcript	# exons	Transcript length	Protein length	Transcript Type
TGME49_222020-t26_1	6	2638	593	mRNA

Transcripts section.

4. Content navigation.

How do you find/navigate to the different sections of the page? Use the “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.



5. Running an alignment of selected sequences

- Expand the “Orthologs and Paralogs in ToxoDB” section.
- Select a few genes from the table using the checkbox.
- Scroll to the bottom of the table and click on the Run Clustal Omega button.

6. Exploring the genetic variation section

			TgCatPRC2			
<input type="checkbox"/>	TGVAND_222020	phosphoglycerate kinase PGKII	Toxoplasma gondii VAND	no	yes	no
<input checked="" type="checkbox"/>	TGVAND_318230	phosphoglycerate kinase PGKI	Toxoplasma gondii VAND	no	no	no
<input checked="" type="checkbox"/>	TGVEG_222020	phosphoglycerate kinase PGKII	Toxoplasma gondii VEG	no	yes	no
<input type="checkbox"/>	TGVEG_318230	phosphoglycerate kinase PGKI	Toxoplasma gondii VEG	no	no	no
<input type="checkbox"/>	TGP89_222020	phosphoglycerate kinase PGKII	Toxoplasma gondii p89	no	yes	no
<input type="checkbox"/>	TGP89_318230	phosphoglycerate kinase PGKI	Toxoplasma gondii p89	no	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.

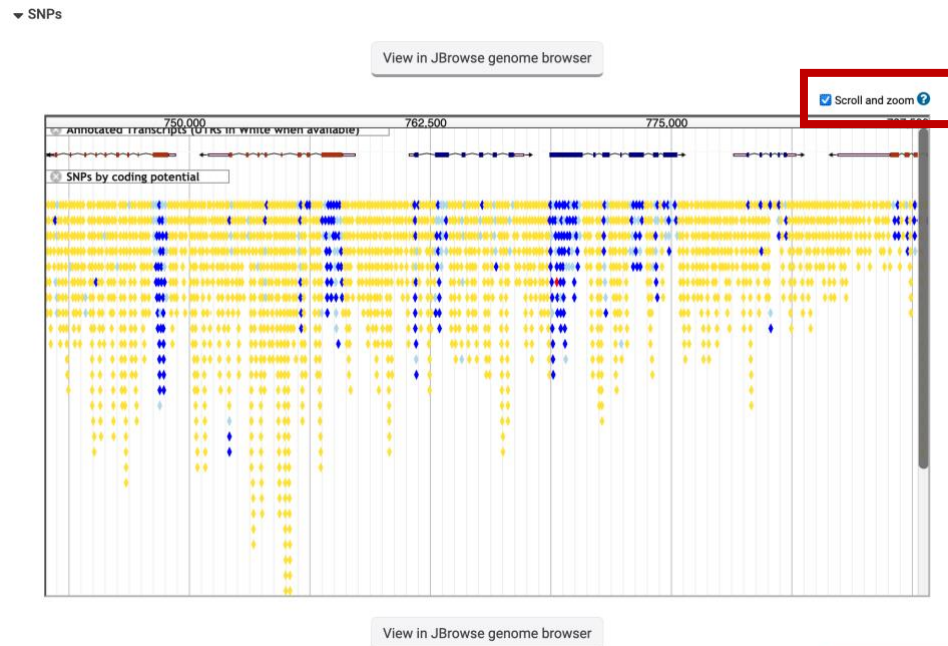
☒ Protein ☐ CDS (spliced) ☐ Genomic

Output format: Mismatches highlighted

Run Clustal Omega for selected genes

COMMUNITY CHAT

Go to the Genetic variation section of the gene page and expand the SNP section. Notice that by default you cannot scroll within the embedded browser window. To enable scrolling, select the “Scroll and Zoom” check box in the upper right-hand side of the browser window. To scroll down within the browser window, you click and drag or use two-finger scrolling. You can also double click in an area to zoom in.



SNP color code: Dark blue (non-synonymous), light blue (synonymous), Yellow (non-coding), Red (nonsense). What kind of SNPs are in this gene? Can you see any non-synonymous SNPs? How does this compare to the neighboring genes?

7. 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 this gene:

- Is there evidence that this protein is phosphorylated? (hint: go to the proteomics section and expand the Post Translational Modification section).
- Where is the protein localized? (hint: go to the Protein Targeting and Localization section and expand the cellular localization section).
- Is there any phenotypic data available for this gene? (hint: go to the Phenotype section and expand its subsections).
- Is there any RNA-Seq data available for this gene? (hint: go to the Transcriptomics section and expand the subsections called RNA-Seq transcription summary and Transcript Expression).