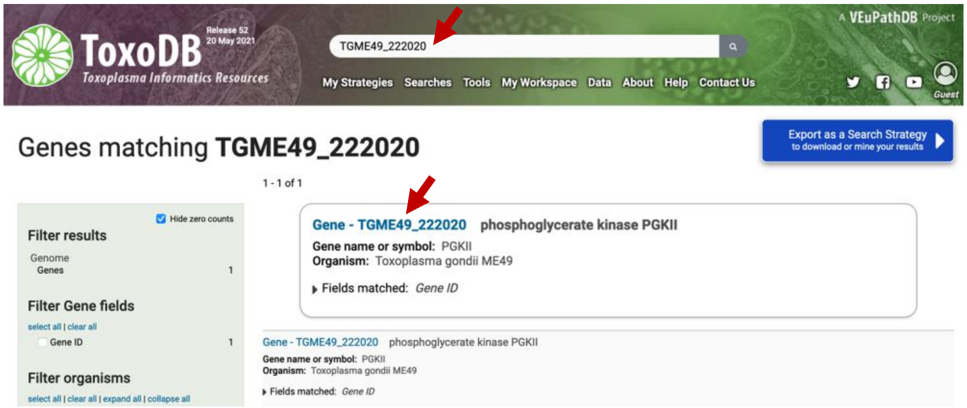
Exploring the Gene Page

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

Learning objectives

- 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 a Gene page:** For this exercise visit the gene page for TGME49_222020 (phosphoglycerate kinase PGKII). How did you get to this gene? (hint: use the site search to retrieve the gene page link, then click on the gene ID in the results.)



TGME49_222020

Genes matching **TGME49_222020**

1 - 1 of 1

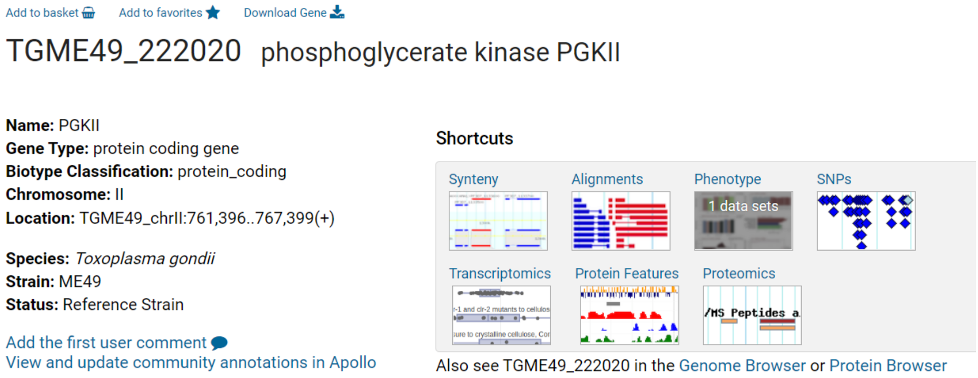
Genome
Genes 1

Gene ID 1

Gene - TGME49_222020 phosphoglycerate kinase PGKII
Gene name or symbol: PGKII
Organism: Toxoplasma gondii ME49
Fields matched: *Gene ID*

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**
   a. What information is in this section?
   b. Can you easily find which chromosome this gene is located on?
   c. Is this gene protein coding?
   d. What do the shortcuts do?



TGME49_222020 phosphoglycerate kinase PGKII

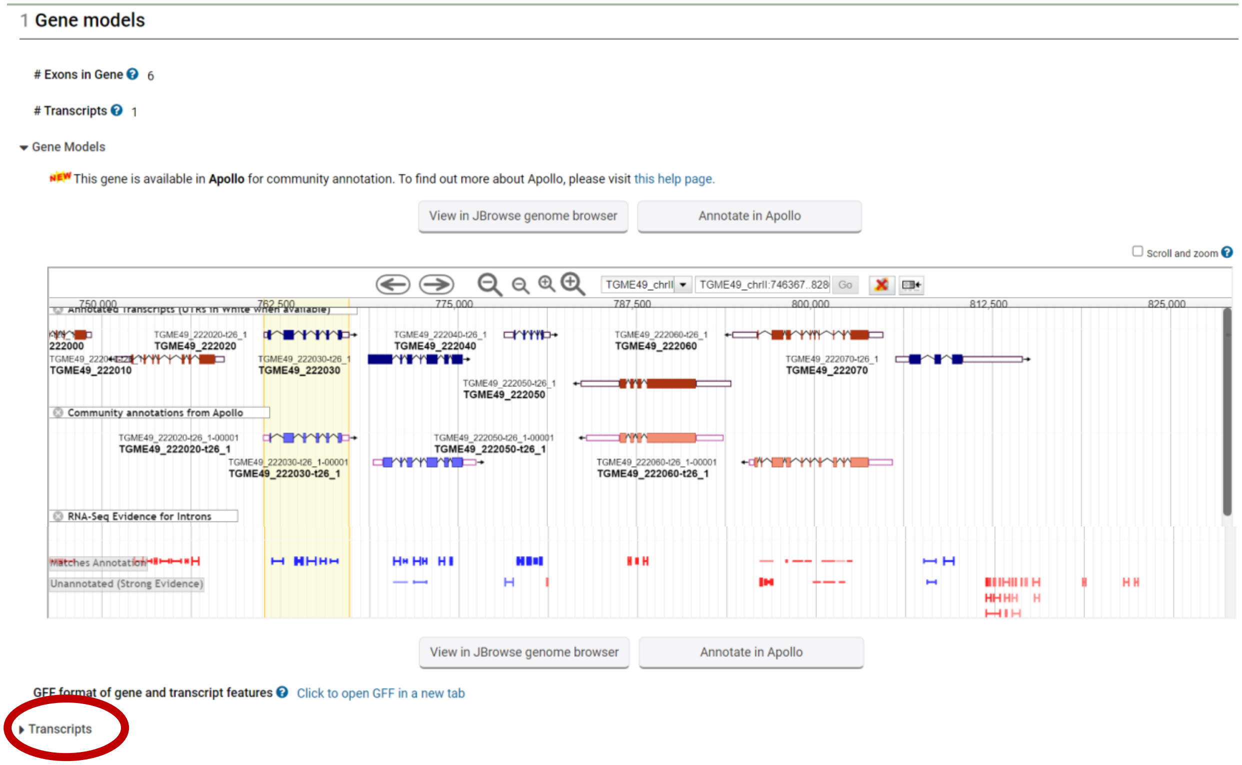
Name: PGKII
Gene Type: protein coding gene
Biotype Classification: protein_coding
Chromosome: II
Location: TGME49_chrII:761,396..767,399(+)

Species: *Toxoplasma gondii*
Strain: ME49
Status: Reference Strain

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

Also see TGME49_222020 in the Genome Browser or Protein Browser

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?
   e. How long is the transcript? You can determine transcript length by expanding the Transcripts table.



1 Gene models

Exons in Gene 6

Transcripts 1

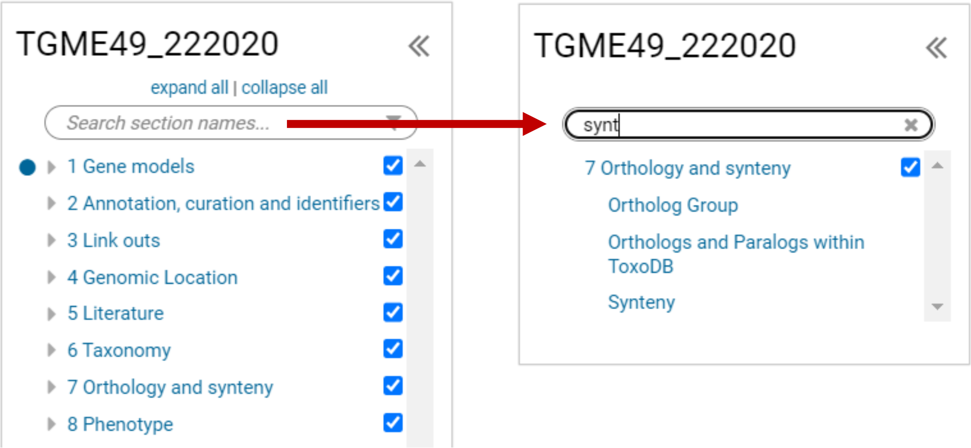
Gene Models

NEW This gene is available in **Apollo** for community annotation. To find out more about Apollo, please visit this help page.

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

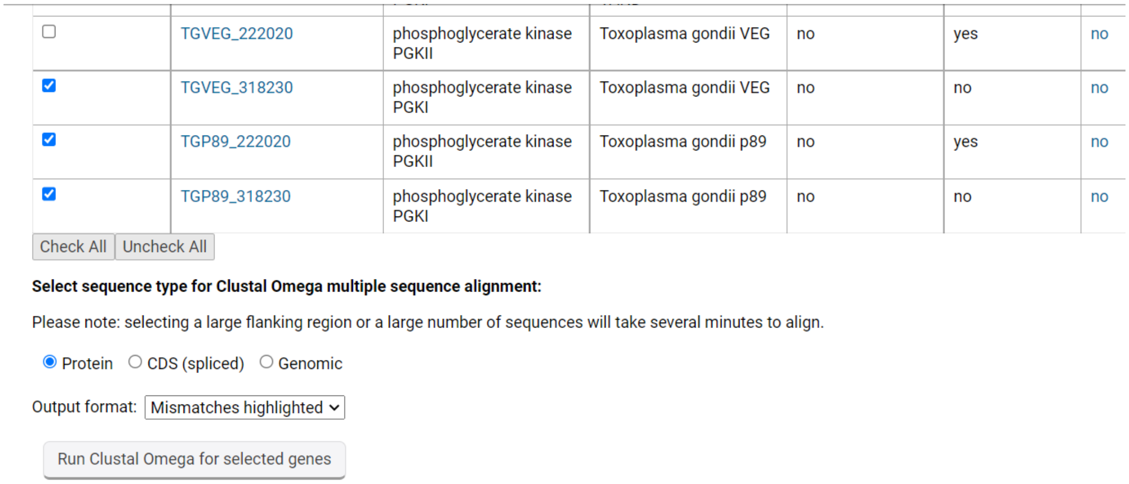
Transcripts

4. **Content navigation.** How do you navigate to the different sections of the page? Use the "Contents" navigation tool on the left side of the page. 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. Navigate to the synteny section. Does Cyclospora cayetanensis isolate NF1_C8 share synteny in this region?



b. Navigate to the Transcript Expression table and open the row for the experiment called “mRNA Illumina sequences from 4 different mouse cell types infected with T. gondii and aligned to the ME49 Genome.” (Notice the filter box at the top of the table)
   i. What data type was used to produce this data?
   ii. In what cell type does this gene have the highest expression level?
   iii. Open the Data table for this experiment. What is the TPM expression value for unique reads mapped to the Neurons?

5. **Run an alignment of selected protein sequences.** The Orthologs and Paralogs in ToxoDB table shows all orthologs for TGME49_222020. This table is also a tool for running Clustal Omega on protein, CDS and genomic sequence for this gene and the orthologs you choose from the table. Within all VEuPathDB sites, OrthoMCL is used to determine orthology.
   a. Expand the “Orthologs and Paralogs in ToxoDB” 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.



☐	TGVEG_222020	phosphoglycerate kinase PGKII	Toxoplasma gondii VEG	no	yes	no
☑	TGVEG_318230	phosphoglycerate kinase PGKI	Toxoplasma gondii VEG	no	no	no
☑	TGP89_222020	phosphoglycerate kinase PGKII	Toxoplasma gondii p89	no	yes	no
☑	TGP89_318230	phosphoglycerate kinase PGKI	Toxoplasma gondii p89	no	no	no

Select sequence type for Clustal Omega multiple sequence alignment:

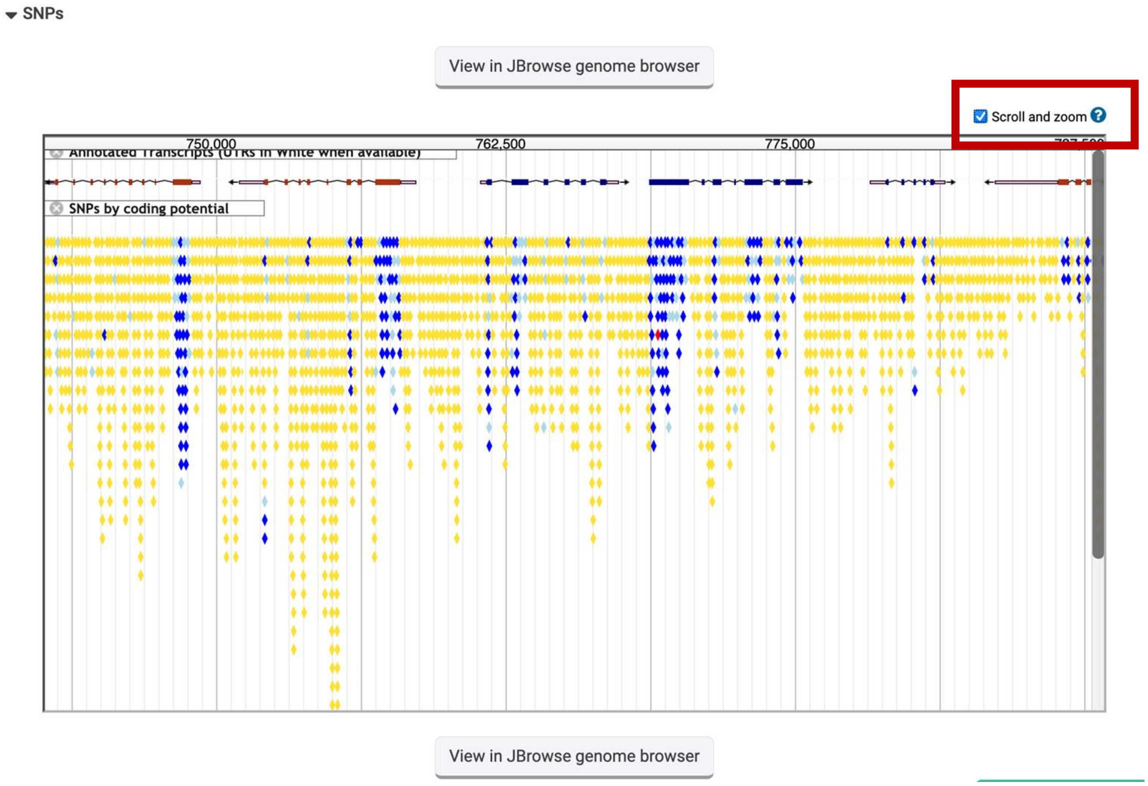
Please note: selecting a large flanking region or a large number of sequences will take several minutes to align.

● Protein ○ CDS (spliced) ○ Genomic

Output format: Mismatches highlighted

6. Explore the genetic variation section

The genetic variation section contains a graphic and a tool for aligning isolate sequences within the region of the gene. 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, 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:

a. Is there evidence that this protein is phosphorylated? (hint: go to the proteomics section and expand the Post Translational Modification section).
b. Where is the protein localized? (hint: go to the Protein Targeting and Localization section and expand the cellular localization section).
c. Is there any phenotypic data available for this gene? (hint: go to the Phenotype section and expand its subsections).

d. What other 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).