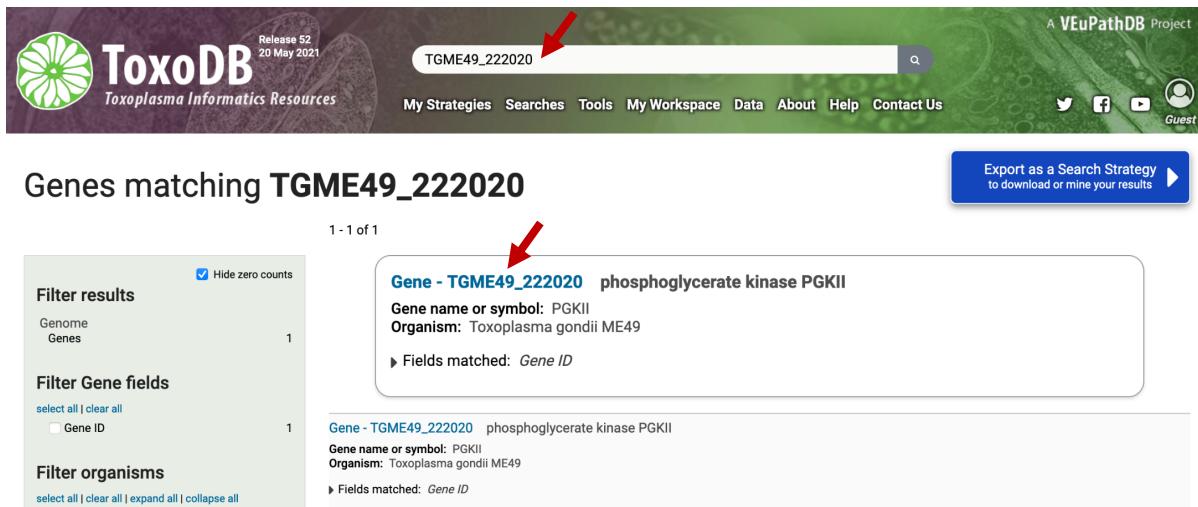


Exploring the Gene Page

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, go to ToxoDB.org¹ and 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.)



ToxoDB Release S2
20 May 2021

TGME49_222020

My Strategies Searches Tools My Workspace Data About Help Contact Us

Guest

Export as a Search Strategy to download or mine your results

Genes matching **TGME49_222020**

1 - 1 of 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

¹ Note: this exercise uses ToxoDB (<https://ToxoDB.org>) as an example database, but the same functionality is available on all VEuPathDB resources.

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
Gene Type	protein coding gene
Biotype Classification	protein_coding
Chromosome	II
Location	TGME49_chrl:761,352..767,399(+)
Species	Toxoplasma gondii
Strain	ME49
Status	Reference Strain
User Comments	Add the first
Community Annotations	View / Update in Apollo editor

Shortcuts

- Synteny
- Alignments
- SNPs
- Transcriptomics
- Protein Features
- Proteomics
- /MS Peptides

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.

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



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 datatype, 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 in the screenshot. Begin typing the word ‘synteny’ in the filter to collapse the content menu.

- Navigate to the synteny section. *Does Cyclospora cayetanensis isolate NF1_C8 share synteny in this region?*

TGME49_222020

expand all | collapse all

Search section names...

- Gene models
- Annotation, curation and identifiers
- Link outs
- Genomic Location
- Literature
- Taxonomy
- Orthology and synteny
- Phenotype
- Genetic variation
- Transcriptomics
- Sequence analysis
- Sequences
- Structure analysis
- Protein features and properties
- Protein targeting and localization
- Function prediction
- Pathways and interactions
- Proteomics
- Immunology

TGME49_222020

synt

Orthology and synteny

Ortholog Group (OrthoMCL 6)
Ortholog Group (OrthoMCL 7)
Orthologs and Paralogs within ToxoDB
Synteny

- Navigate to the “Transcriptomics” section of the gene page.

TGME49_222020

transcriptomi

Transcriptomics

- # Microarray Datasets 11
- # RNASeq Datasets 25
- > AI Expression Summary
- > RNA-Seq Transcription Summary
- > Transcript Expression
- > User Dataset Transcriptomics Graphs
- > Single Cell RNA-Seq (scRNA-Seq)

AI-assisted summary

Summary graph of all RNAseq data

Explore individual expression graphs

- How many transcriptomic datasets are available?
- VEuPathDB resources offer various tools to facilitate data exploration and summarization. For transcriptomic data, there are three ways to interact with the data:
 - You can examine individual experiments by expanding the “Transcript Expression” section.
 - You can explore a summary graph that represents all available RNA-seq experiments.
 - You can utilize an AI-assisted summary for expression data.
- Take a moment to compare the various methods for exploring transcriptomics data. **Is the AI-assisted summary useful?**
- In addition to bulk transcriptomics data, when available, VEuPathDB resources also include single-cell RNA sequencing (scRNASeq) data. Explore the scRNASeq section; we will also discuss this as a separate module during the workshop.

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

- 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.

<input type="checkbox"/>	TGVEG_222020	phosphoglycerate kinase PGKII	Toxoplasma gondii VEG	no	yes	no
<input checked="" type="checkbox"/>	TGVEG_318230	phosphoglycerate kinase PGKI	Toxoplasma gondii VEG	no	no	no
<input checked="" type="checkbox"/>	TGP89_222020	phosphoglycerate kinase PGKII	Toxoplasma gondii p89	no	yes	no
<input checked="" 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

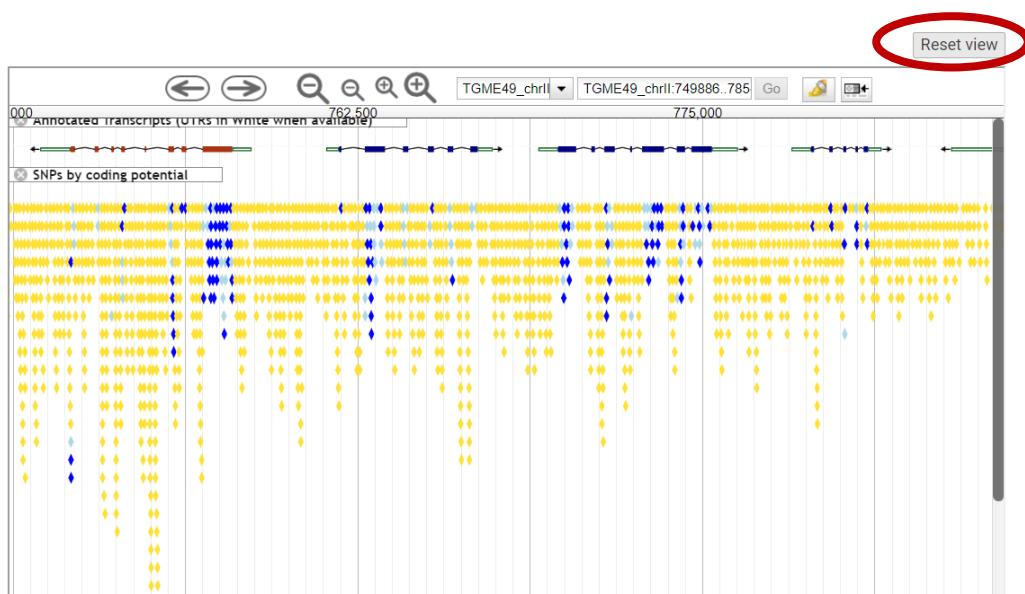
Output format: [Mismatches highlighted](#) ▾

[Run Clustal Omega for selected genes](#)

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. Scrolling is enabled by default. To reset the image to the default position, choose Reset to Default. 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).
-