

## Site Search

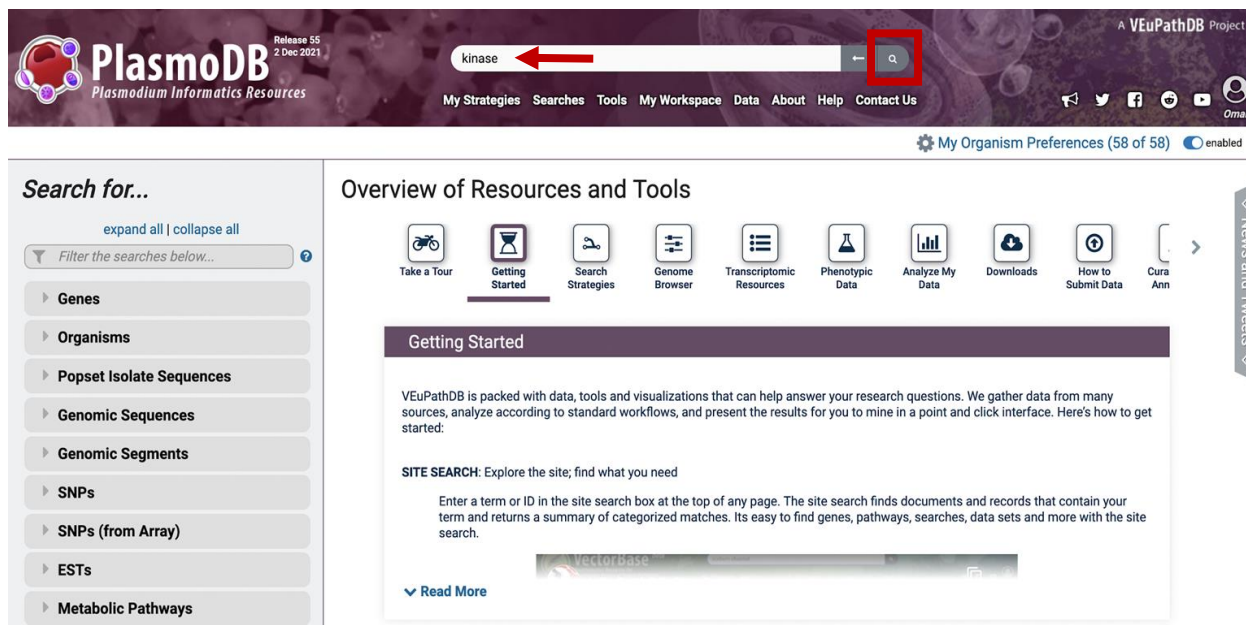
**Note:** this exercise uses *PlasmoDB.org* as an example database, but the same functionality is available on all VEuPathDB resources.

### Learning objectives:

- Search by keywords or identifiers
- Filter site search results by categories and fields
- Export results to a search strategy
- Find a specific gene using its ID in site search
- Navigate to and from the site search result
- Explore searches using wild cards (\*)

The site search is located in the header of any VEuPathDB site and is available from every page. The site search queries the databases for your term or ID and returns a list of pages and documents that contain your query term.

1. **Search for a keyword.** Enter the word *kinase* in the site search window (arrow in the image below). Then click enter on your keyboard or click on the search icon (square in the image below).



2. **Site Search result format:** The site search returns a categorized list of pages and documents that contain your term. Site search results are summarized on the left with a details panel on the right. Changing the panel on the left will populate the details panel with that result. What is the total number of results with the word kinase? Are all the results genes?

## All results matching kinase

The screenshot shows a search results interface for the term 'kinase'. On the left, a 'Filter results' sidebar lists categories like Genome (17,924), Population biology (352), Metabolism (352), and Data access (3). A red arrow points to the 'Genes' link under the 'Genome' category. Below this, 'Filter fields' and 'Filter organisms' sections are visible. The main panel displays a list of genes, with the first four shown in detail. Each entry includes the gene name, type, organism, and a list of matched fields. A red bracket at the bottom of the sidebar is labeled 'Results are summarized by category'. Another red bracket at the bottom of the main panel is labeled 'Details panel with information about each item returned'.

Export as a Search Strategy  
to download or mine your results

1 - 20 of 18,825

**Filter results** ☒ Hide zero counts

Genome  
Genes 17,924  
Population biology 352  
Metabolism 352  
Data access 3

**Filter fields**  
Select a result filter above

**Filter organisms**  
select all | clear all | expand all | collapse all  
Type a taxonomic name     
☐ Haemoproteidae 324  
☐ Plasmodiidae 17,601

Gene - PCYB\_132500 kinase  
Gene type: protein coding gene  
Organism: Plasmodium cynomolgi strain B  
Fields matched: GO terms; InterPro domains; Preferred product description; Product descriptions

Gene - PKNOH\_S07456300 Kinase  
Gene type: protein coding gene  
Organism: Plasmodium knowlesi strain Malayan Strain Pk1 A  
Fields matched: GO terms; InterPro domains; Orthologs; Preferred product description; Product descriptions

Gene - PKNOH\_S140234600 Kinase  
Gene name or symbol: IPK2  
Gene type: protein coding gene  
Organism: Plasmodium knowlesi strain Malayan Strain Pk1 A  
Fields matched: EC descriptions and numbers; GO terms; InterPro domains; Orthologs; PDB chains; Preferred product description; Product descriptions

Gene - AK88\_00505 pantothenate kinase  
Gene type: protein coding gene  
Organism: Plasmodium fragile strain nilgiri  
Fields matched: EC descriptions and numbers; GO terms; InterPro domains; Orthologs; PDB chains; Preferred product description; Product descriptions

Gene - AK88\_01656 phosphoglycerate kinase  
Gene type: protein coding gene  
Organism: Plasmodium fragile strain nilgiri  
Fields matched: EC descriptions and numbers; GO terms; InterPro domains; Orthologs; PDB chains; Preferred product description; Product descriptions

Gene - AK88\_02186 pyridoxal kinase  
Gene type: protein coding gene  
Organism: Plasmodium fragile strain nilgiri  
Fields matched: EC descriptions and numbers; GO terms; InterPro domains; Orthologs; PDB chains; Preferred product description; Product descriptions

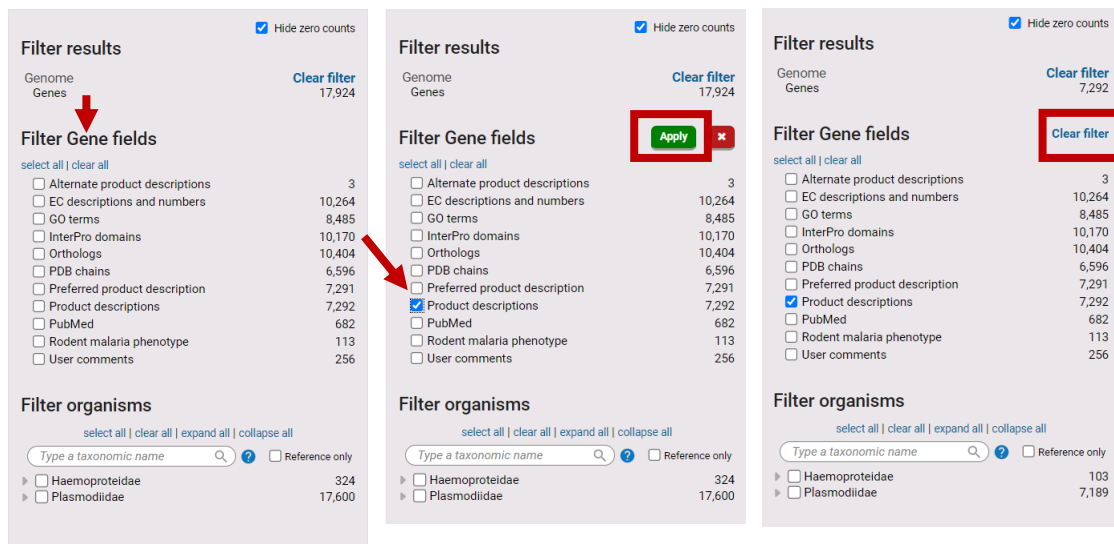
COMMUNITY CHAT

Results are  
summarized  
by category

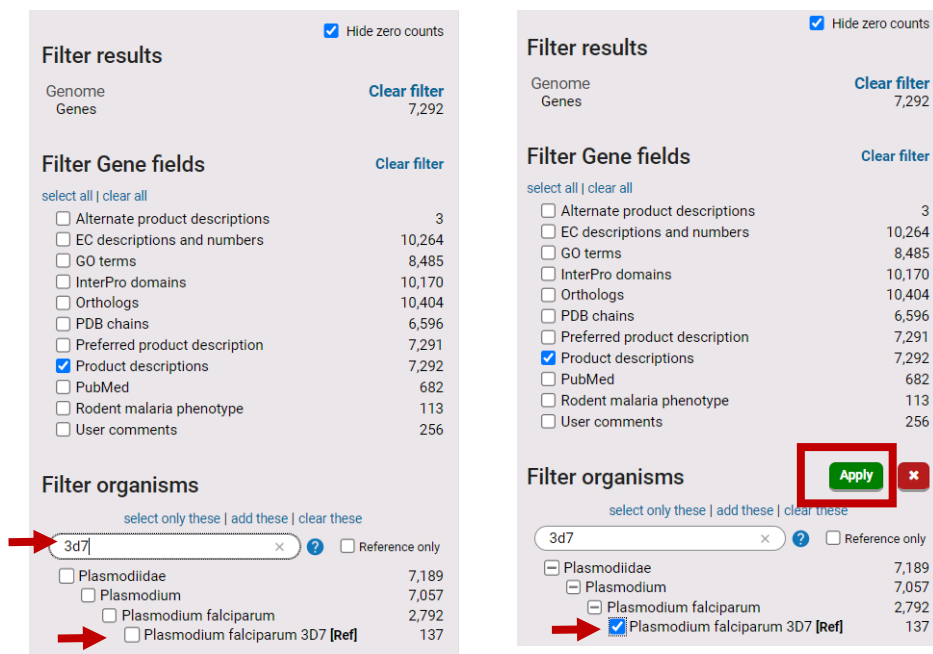
Details panel with  
information about  
each item returned

3. **Filter the site search result by category:** How many of the genes included the word kinase in their product descriptions?

Filter the results so that you only view gene results (hint: click on the word *genes* in the *Filter results* section; arrow in image above) and the Filter Fields section expands to reveal additional filtering options. Select the *Product descriptions* field and choose *Apply* (middle panel below). Once a filter is applied it can be removed by clicking on *Clear filter* (right panel below).



4. **Filter the site search result by organism:** How many of the above genes are found in *Plasmodium falciparum* 3D7? Explore the *Filter organisms* section of the results filter and use the search filter to navigate the tree.



5. **Export the results to a search strategy.** Use the blue *Export as a search strategy* button at the top right-hand side of the results. Once exported you will be able to take advantage of over 100 specialized searches using the Add Step button. We will learn more about this in a future exercise.

The screenshot shows the search results for 'kinase' filtered by fields and organisms. At the top right, there is a blue button labeled 'Export as a Search Strategy to download or mine your results'. A red arrow points from this button to the 'Add a step' button in the 'Unnamed Search Strategy' window. The search results page shows 137 genes, with details for Gene PF3D7\_0616000 (pyridoxal kinase) visible. The search strategy builder window shows a 'Text' step with 137 genes and an 'Add a step' button.

Genes matching **kinase** (filtered by fields and organisms)

1 - 20 of 137

Filter results: Hide zero counts, Clear filter, 137

Filter Gene fields: Clear filter

Gene - PF3D7\_0616000 pyridoxal kinase  
Gene name or symbol: PDXK  
Gene type: protein coding gene  
Organism: Plasmodium falciparum 3D7  
Fields matched: EC descriptions and numbers; GO terms; InterPro domains; Orthologs; PDB chains; Preferred product description; Product descriptions; PubMed

Unnamed Search Strategy \*

Text 137 Genes Step 1 Add a step

137 Genes (113 ortholog groups) Revise this search

Gene Results Genome View Analyze Results

Genes: 137 Transcripts: 138 (hiding 1) Show Only One Transcript Per Gene

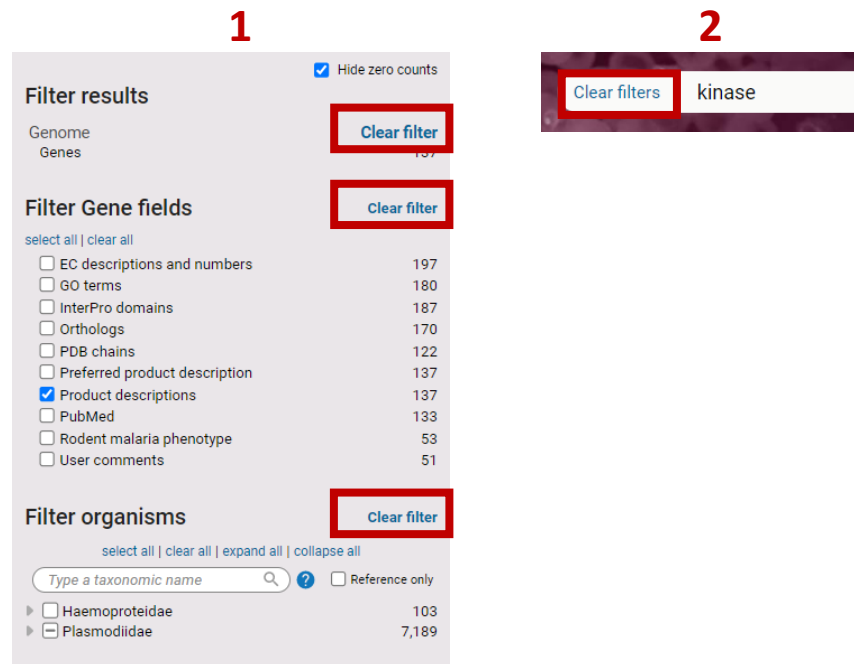
Rows per page: 1000 Download Send to... Add Columns

Gene ID	Transcript ID	Organism	Genomic Location (Gene)	Product Description
PF3D7_0102600	PF3D7_0102600.1	Plasmodium falciparum 3D7	PF3D7_01_v3:118,812..122,534(-)	serine/threonine protease
PF3D7_0103700	PF3D7_0103700.1	Plasmodium falciparum 3D7	PF3D7_01_v3:166,497..168,687(+)	L-seryl-tRNA(Sec) kinase
PF3D7_0107600	PF3D7_0107600.1	Plasmodium falciparum 3D7	PF3D7_01_v3:313,824..319,525(+)	eukaryotic translation initiation factor 4E
PF3D7_0110900	PF3D7_0110900.1	Plasmodium falciparum 3D7	PF3D7_01_v3:419,727..420,942(-)	adenylate kinase-like protein
PF3D7_0111500	PF3D7_0111500.1	Plasmodium falciparum 3D7	PF3D7_01_v3:439,395..442,195(-)	UMP-CMP kinase, putative

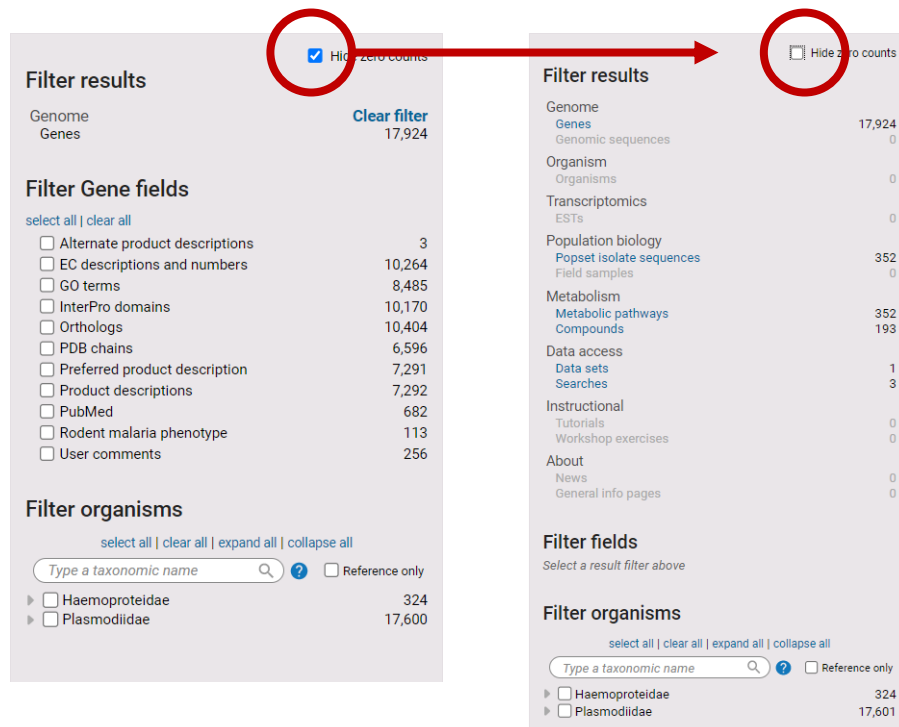
6. Return to the site search results page. You can achieve this in two ways: 1. Your previous results and filter settings were preserved and can be accessed by clicking on the 'back to results' arrow in the site search window. 2. Click on your browser's back arrow.



7. **Clear all filters.** You can achieve this in two ways: 1. You can click on each of the clear filter options in the filter results panel (boxes below). 2. You can click on the *clear filters* option in the site search window, which serves to Clear All filters.



8. Click the *Hide zero counts* check box in the *Filter results* panel. What does this do?



9. **Run a wild card search.** The wild card (denoted by an asterisk \*) can be used alone to retrieve all site search results or combined with a word such as *\*kinase* to retrieve compound words ending with the word kinase like phosphofructokinase. As usual results can then be explored using the filters in the *Results filter* on the left side of the website. Feel free to compare the results you get when you run a search or the word *kinase* to a search with a wild card *\*kinase* or *\*kinase\**.

All results matching \*

1 - 20 of 957,899

Export as a Search Strategy to download or mine your results

Filter results ☒ Hide zero counts

Genome	336,439
Genes	21,810
Genomic sequences	
Organism	60
Transcriptomics	
ESTs	287,336
Population biology	
Popset isolate sequences	153,109

Compound - CHEBI:100000 (2S,3S,4R)-3-[4-(3-cyclopentylprop-1-ynyl)phenyl]-4-(hydroxymethyl)-1-(2-methoxy-1-oxoethyl)-2-azetidinecarboxamide

Compound - CHEBI:100001 N-[(2R,3S,6R)-2-(hydroxymethyl)-6-[2-[[oxo-4-(trifluoromethyl)anilino]methyl]amino]ethyl]-3-oxanyl]-3-pyridinecarboxamide

Compound - CHEBI:100002 3-chloro-N-[(5S,6S,9S)-5-methoxy-3,6,9-trimethyl-2-oxo-11-oxa-3,8-diazabicyclo[10.4.0]hexadeca-1(12),13,15-trien-14-yl]benzenesulfonamide

Compound - CHEBI:100003 (4R,7S,8R)-8-methoxy-4,7,10-trimethyl-11-oxo-14-(1-oxobutylamino)-N-propyl-2-oxa-5,10-diazabicyclo[10.4.0]hexadeca-1(12),13,15-triene-5-carboxamide

Compound - CHEBI:100004 1-(2,5-difluorophenyl)-3-[(5S,6S,9S)-5-methoxy-3,6,9-trimethyl-2-oxo-8-oxo(2-pyrazinyl)methyl]-11-oxa-3,8-diazabicyclo[10.4.0]hexadeca-1(12),13,15-trien-14-yl]urea

Compound - CHEBI:100005 N-[(1S,3S,4aS,9aR)-1-(hydroxymethyl)-3-[2-oxo-2-(1-piperidinyl)ethyl]-3,4,4a,9a-tetrahydro-1H-pyranol[3,4-b]benzofuran-6-yl]-3-methoxybenzenesulfonamide

All results matching \*kinase

1 - 20 of 21,371

Export as a Search Strategy to download or mine your results

Filter results ☒ Hide zero counts

Genome	19,406
Genes	
Population biology	1,273
Popset isolate sequences	
Metabolism	484
Metabolic pathways	204
Compounds	
Data access	1
Data sets	3
Searches	

Filter fields Select a result filter above

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

Type a taxonomic name   ☐ Reference only

☐ Haemoproteidae 358

☐ Plasmodidae 19,049

Gene - AK88\_00104 CK1/CK1/CK1-D protein kinase

Gene type: protein coding gene

Organism: Plasmodium fragile strain nilgiri

Fields matched: EC descriptions and numbers; GO terms; InterPro domains; Orthologs; PDB chains; Preferred product description; Product descriptions

Gene - AK88\_00479 CAMK protein kinase

Gene type: protein coding gene

Organism: Plasmodium fragile strain nilgiri

Fields matched: EC descriptions and numbers; GO terms; InterPro domains; Orthologs; PDB chains; Preferred product description; Product descriptions

Gene - AK88\_00505 pantothenate kinase

Gene type: protein coding gene

Organism: Plasmodium fragile strain nilgiri

Fields matched: EC descriptions and numbers; GO terms; InterPro domains; Orthologs; PDB chains; Preferred product description; Product descriptions

Gene - AK88\_00565 Atypical/ABC1 protein kinase

Gene type: protein coding gene

Organism: Plasmodium fragile strain nilgiri

Fields matched: GO terms; InterPro domains; Orthologs; Preferred product description; Product descriptions

Gene - AK88\_00580 CMGC protein kinase

Gene type: protein coding gene

Organism: Plasmodium fragile strain nilgiri

10. **Search for a specific gene ID.** Enter the gene ID in the site search window: *PF3D7\_0310100*. When there is an exact match for an ID in the database, the site search offers a card in the details panel to draw attention to the direct link to the gene page. Although your search for PF3D7\_0310100 does return a direct link to the gene in *P. falciparum* 3D7, it also returns a link to the *P. gaboni* strain gene. Why?

PF3D7\_0310100

My Organism Preferences (60 of 60)

enabled

Genes matching **PF3D7\_0310100**

Export as a Search Strategy  
to download or mine your results

1 - 2 of 2

Filter results

Genome

Genes 2

Filter Gene fields

select all | clear all

External links 1

Gene ID 1

Names, IDs, and aliases 1

Notes from annotators 1

Filter organisms

select all | clear all | expand all | collapse all

Type a taxonomic name

Reference only

Plasmodiidae 2

Plasmodium 2

Gene - **PF3D7\_0310100** calcium-dependent protein kinase 3

Gene name or symbol: CDPK3

Gene type: protein coding gene

Organism: Plasmodium falciparum 3D7

Fields matched: External links; Gene ID; Names, IDs, and aliases

Gene - PF3D7\_0310100 calcium-dependent protein kinase 3

Gene name or symbol: CDPK3

Gene type: protein coding gene

Organism: Plasmodium falciparum 3D7

Fields matched: External links; Gene ID; Names, IDs, and aliases

Gene - **PGSY75\_0310100** calcium-dependent protein kinase 3

Gene type: protein coding gene

Organism: Plasmodium gaboni strain SY75

Fields matched

Notes from annotators: gap found within coding sequence~ort

1 - 2 of 2

Direct link to PF3D7\_0310100

Why is this gene returned by a search for PF3D7\_0310100?

7