

Site Search

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

What is the VEuPathDB Site Search?

- The site search queries the database for your search term (keyword or identifier) and returns a list of pages and documents that contain your query term
- The site search is a convenient starting point for mining the vast amount of data available in VEuPathDB and identifying and filtering information relevant to you
- The site search box is located in the header of every VEuPathDB site and is available on every page

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 the gene ID in site search
- Navigate to and from the site search result
- Explore searches using wildcards (*)

1. **Go to PlasmoDB.org and search for a keyword.** Enter the word *kinase* in the site search window (red arrow in the image below) at the top of the home page. Then click <enter> on your keyboard or click on the search icon (red circle in the image below).



2. **Site Search result format:** The site search returns a categorized list of pages and documents that contain your search term. Site search results are presented with a **summary panel on the left** and a **details panel on the right**. Clicking on a particular result type on the left will populate the details panel with that result.
 - a. What is the total number of results with the word kinase?
 - b. Are all the results genes? (Red circle in image below)

All results matching **kinase**

1 - 20 of 20,497

◀ 1 2 3 ... 1,025 ▶

Filter results

☒ Hide zero counts

Genome
Genes 19,596

Population biology
Popset isolate sequences 352

Metabolism
Metabolic pathways 352
Compounds 193

Data access
Data sets 1
Searches 3

Filter fields
Select a result filter above

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

Type a taxonomic name ☐ Reference only

☐ Haemoproteidae 331
☐ Plasmodiidae 19,266

Summary panel with categories of results

Gene - PCYB_132500 kinase
Gene type: protein coding gene
Organism: Plasmodium cynomolgi strain B
Fields matched: GO terms; InterPro domains; Orthologs; Product description; Product descriptions (all)

Gene - PKNOH_S07456300 Kinase
Gene type: protein coding gene
Organism: Plasmodium knowlesi strain Malaysian Strain Pk1 A
Fields matched: GO terms; InterPro domains; Orthologs; Product description; Product descriptions (all)

Gene - PKNOH_S140234600 Kinase
Gene name or symbol: IPK2
Gene type: protein coding gene
Organism: Plasmodium knowlesi strain Malaysian Strain Pk1 A
Fields matched: EC descriptions and numbers; GO terms; InterPro domains; Orthologs; PDB chains; Product description; Product descriptions (all)

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; Product description; Product descriptions (all)

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; Product description; Product descriptions (all)

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; Product description; Product descriptions (all)

Details panel with information about each item returned

3. **Filter the site search result by category:** Click on “genes” (red arrow in image above) to view only results from that category. Notice that the **Filter Fields** section expands to reveal additional filtering options (categories).
 - a. How many of the genes included the word kinase in their product descriptions? To find out, select the “Product description” field (red arrow, left panel in image below) and choose Apply (red arrow, middle panel in image below)
 - b. Select the *Product descriptions* field and choose *Apply*. Once a filter is applied it can be removed by clicking on *Clear filter* (red circle, right panel in image below).

Genome
Genes 7,630

Filter Gene fields
select all | clear all

☐ EC descriptions and numbers 11,643
☐ GO terms 8,874
☐ InterPro domains 10,652
☐ Notes from annotators 1
☐ Orthologs 11,026
☐ PDB chains 6,901
☒ Product description 7,630
☐ Product descriptions (all) 7,633
☐ PubMed 682
☐ Rodent malaria phenotype 119
☐ User comments 256

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

Type a taxonomic name ☐ Reference only

☐ Haemoproteidae 103
☐ Plasmodiidae 7,527

Clear filter

Apply ✖

Genome
Genes 19,596

Filter Gene fields
select all | clear all

☐ EC descriptions and numbers 11,643
☐ GO terms 8,874
☐ InterPro domains 10,652
☐ Notes from annotators 1
☐ Orthologs 11,026
☐ PDB chains 6,901
☒ Product description 7,630
☐ Product descriptions (all) 7,633
☐ PubMed 682
☐ Rodent malaria phenotype 119
☐ User comments 256

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

Type a taxonomic name ☐ Reference only

☐ Haemoproteidae 331
☐ Plasmodiidae 19,266

Genome
Genes 7,630

Filter Gene fields
select all | clear all

☐ EC descriptions and numbers 11,643
☐ GO terms 8,874
☐ InterPro domains 10,652
☐ Notes from annotators 1
☐ Orthologs 11,026
☐ PDB chains 6,901
☒ Product description 7,630
☐ Product descriptions (all) 7,633
☐ PubMed 682
☐ Rodent malaria phenotype 119
☐ User comments 256

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

Type a taxonomic name ☐ Reference only

☐ Haemoproteidae 103
☐ Plasmodiidae 7,527

Clear filter

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 summary results panel

- Start typing “3d7” in the “filter organisms” search box (red arrow in first panel below) and explore the resulting organism tree (red box in first panel below)
- Apply the filter (red arrow in second panel below)
- How many genes in *P. falciparum* contain “kinase” in product descriptions?

The image shows two panels from a bioinformatics interface. The left panel, titled 'Filter results', has a 'Hide zero counts' checkbox checked. It contains three sections: 'Filter Gene fields' with a list of fields and their counts (e.g., EC descriptions and numbers: 11,643), 'Filter organisms' with a search box containing '3D7' and a tree view showing the selection of 'Plasmodium falciparum 3D7 [Ref]' (137 genes). The right panel, titled 'Filter results', shows the 'Genome' filter set to 'Genes' (7,630) and the 'Filter Gene fields' section with 'Product description' selected (7,630). The 'Filter organisms' section shows the same search box and tree view, with 'Plasmodium falciparum 3D7 [Ref]' selected. Red arrows point to the search box in the left panel and the 'Apply' button in the right panel.

- Export the results to a search strategy.** Click the blue *Export as a search strategy* button at the top right-hand side of the results (see image below).

The image shows a search strategy page titled 'Genes matching kinase (filtered by fields and organisms)'. It displays a list of genes with their names, symbols, and descriptions. The first gene is 'PF3D7_0615000 pyridoxal kinase'. The page includes a sidebar with filter results and a top right button labeled 'Export as a Search Strategy'. The main content area shows a list of genes with their names, symbols, and descriptions, along with a 'Fields matched' section for each gene.

Once exported, the results open into a search strategy (image below). From here 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.

My Search Strategies

Opened (1) All (1) Public (43) Help

Unnamed Search Strategy *

Test 137 Genes

Step 1

137 Genes (113 ortholog groups) [Revise this search](#)

Gene Results Genome View **Analyze Results**

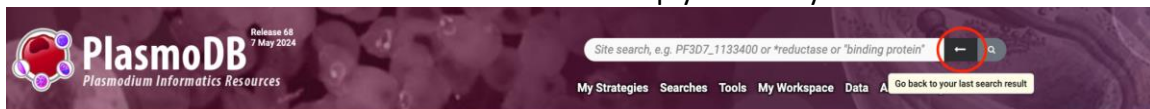
Genes: 137 Transcripts: 138 ☐ Show Only One Transcript Per Gene

Rows per page: 20

Download Send to... Add Columns

Gene ID	Transcript ID	Organism	Genomic Location (Gene)	Product Description	Score
PF3D7_0102600	PF3D7_0102600.1	Plasmodium falciparum 3D7	PF3D7_01_v2:118,812..122,534(-)	serine/threonine protein kinase, FIKK family	23.71
PF3D7_0103700	PF3D7_0103700.1	Plasmodium falciparum 3D7	PF3D7_01_v2:166,497..168,687(+)	L-seryl-tRNA(Sec) kinase, putative	23.71
PF3D7_0107600	PF3D7_0107600.1	Plasmodium falciparum 3D7	PF3D7_01_v3:313,824..319,525(+)	eukaryotic translation initiation factor 2-alpha kinase 2, putative	20.41
PF3D7_0110900	PF3D7_0110900.1	Plasmodium falciparum 3D7	PF3D7_01_v3:419,727..420,942(-)	adenylate kinase-like protein 1	25.79
PF3D7_0111500	PF3D7_0111500.1	Plasmodium falciparum 3D7	PF3D7_01_v3:439,395..442,195(-)	UMP-CMP kinase, putative	28.27
PF3D7_0203100	PF3D7_0203100.1	Plasmodium falciparum 3D7	PF3D7_02_v3:148,914..157,296(-)	protein kinase, putative	31.28
PF3D7_0211700	PF3D7_0211700.1	Plasmodium falciparum 3D7	PF3D7_02_v3:469,408..474,167(+)	tyrosine kinase-like protein, putative	25.79
PF3D7_0213400	PF3D7_0213400.1	Plasmodium falciparum 3D7	PF3D7_02_v3:540,403..543,807(+)	protein kinase 7	31.28
PF3D7_0214600	PF3D7_0214600.1	Plasmodium falciparum 3D7	PF3D7_02_v3:593,374..600,035(+)	serine/threonine protein kinase STK2, putative	23.71
PF3D7_0217500	PF3D7_0217500.1	Plasmodium falciparum 3D7	PF3D7_02_v3:719,266..723,617(+)	calcium-dependent protein kinase 1	25.79

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. Or simply 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. 2. You can click on the *clear filters option* in the site search window, which serves to clear all filters.



Filter results 1 ☒ Hide zero counts

Genome 137 [Clear filter](#)

Genes

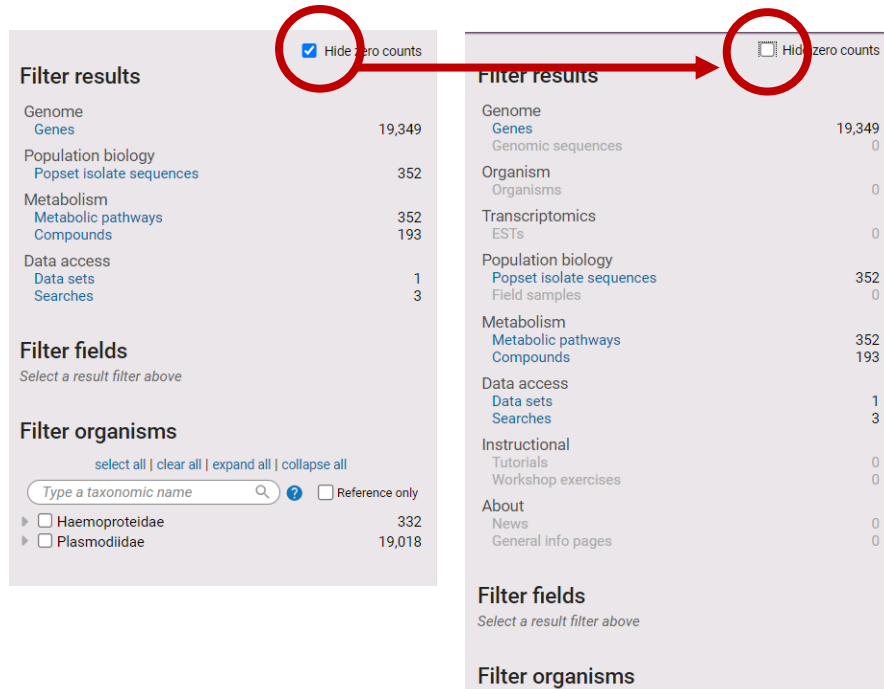
Filter Gene fields [Clear filter](#)

select all | clear all

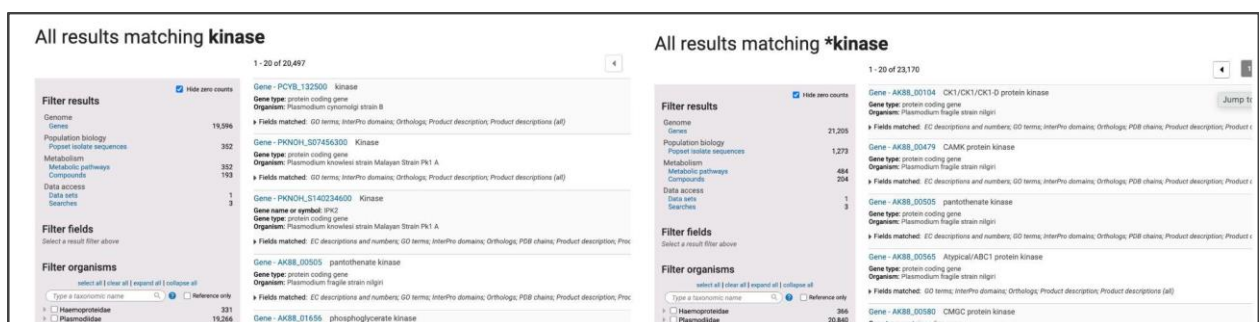
<input type="checkbox"/> EC descriptions and numbers	212
<input type="checkbox"/> GO terms	180
<input type="checkbox"/> InterPro domains	187
<input type="checkbox"/> Orthologs	170
<input type="checkbox"/> PDB chains	122
<input checked="" type="checkbox"/> Product description	137
<input type="checkbox"/> Product descriptions (all)	138
<input type="checkbox"/> PubMed	133
<input type="checkbox"/> Rodent malaria phenotype	56
<input type="checkbox"/> User comments	51

Filter organisms [Clear filter](#)

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



9. **Run a wildcard search.** The wildcard (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**.



10. **Search for a specific gene ID.** Enter a 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 (red arrow in image below). 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 (red box in image below). Why?

Genes matching PF3D7_0310100

1 - 2 of 2

Filter results

Genome

Genes

2

Filter Gene fields

select all | clear all

☐ External links

☐ Gene ID

☐ Names, IDs, and aliases

☐ Notes from annotators

1

1

1

1

Filter organisms

select all | clear all | expand all | collapse all

Type a taxonomic name

☐ Plasmodiidae

☐ Plasmodium

2

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--ortholog of PF3D7_0310100

1 - 2 of 2