

Site Search

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

Learning objectives:

- Use keywords in site search
- Explore site search results
- Filter site search results by categories
- Filter site search results by organisms
- Filter site search results by category fields
- Export results to a search strategy
- Find a specific gene using its ID in site search

1. Enter the word *kinase* in the site search window (top center of the page, arrow in the image below). Then click enter on your keyboard or click on the search icon (square in the image below).

The screenshot shows the PlasmoDB beta homepage. At the top, there is a search bar with the word "kinase" typed into it. A red arrow points to the search bar, and a red square highlights the search icon (a magnifying glass) to its right. Below the search bar, there is a navigation menu with links like "My Strategies", "Searches", "Tools", "My Workspace", "Data", "About", "Help", and "Contact Us". To the right of the menu, there is a user profile icon for "Omar". The main content area has a section titled "Overview of Resources and Tools" with various links to tools like "Take a Tour", "Getting Started", "Search Strategies", "Genome Browser", "Translomic Resources", "Phenotypic Data", "Analyze My Data", "Downloads", and "How to Submit Data". There is also a "Getting Started" section with a "Read More" link. On the left, there is a sidebar titled "Search for..." with a "Filter the searches below..." button. Under this button, there is a list of categories: Genes, Organisms, Popset Isolate Sequences, Genomic Sequences, Genomic Segments, SNPs, SNPs (from Array), ESTs, and Metabolic Pathways. The "Genes" option is currently selected.

2. How many results with the word kinase did you get? Are all the results genes? Explore the filter panel on the left side of the webpage. Filter the results so that you only view gene results (hint: click on the word *genes* in the *Filter results* section; arrow in image below).

The screenshot shows the search results page for the query "kinase". At the top, it says "All results matching kinase" and "1 - 20 of 17,367". There is a "Export as a Search Strategy" button on the right. The results are listed in a table with columns for "Gene name or symbol", "Organism", and "Fields matched". The first result is "Gene - PCYB_132500 kinase" from Plasmodium cynomolgi strain B. The second result is "Gene - PKNH_S07456300 Kinase" from Plasmodium knowlesi strain Malayan Strain Pk1 A. The third result is "Gene - PKNH_S140234600 Kinase" from Plasmodium knowlesi strain Malayan Strain Pk1 A. The fourth result is "Gene - AK88_00505 pantothenate kinase" from Plasmodium fragile strain nigrificans. The fifth result is "Gene - AK88_01656 phosphoglycerate kinase" from Plasmodium fragile strain nigrificans. The sixth result is "Gene - AK88_02186 pyridoxal kinase" from Plasmodium fragile strain nigrificans. On the left, there is a "Filter results" sidebar with a red arrow pointing to the "Genes" option under "Genome". Other options include "Population biology", "Popset isolate sequences", "Metabolism", "Metabolic pathways", "Concepts", "Data access", "Data sets", and "Searches". Below this is a "Filter fields" section with a "Select a result filter above" link. At the bottom, there is a "Filter organisms" section with a "Type a taxonomic name" input field and a "select all | clear all | expand all | collapse all" button.

3. How many of the genes included the word kinase in their product descriptions? Notice that once you filter the result by genes (click on the *Genes* filter), the fields section expands to reveal additional filtering options. Once you select the *Product descriptions* field you are provided the option to *apply* this filter or cancel it (box middle panel below). Once a filter is applied it can be cleared by clicking on *Clear filter* (box left panel below).

Filter results

Hide zero counts

Genome Genes

Clear filter 16,684

Filter Gene fields

[select all](#) | [clear all](#)

- Alternate product descriptions 7
- EC descriptions and numbers 13,510
- GO terms 5,432
- InterPro domains 6,955
- Orthologs 7,136
- PDB chains 4,814
- Product descriptions 6,007
- PubMed 689
- Rodent malaria phenotype 87
- User comments 255

Filter organisms

[select all](#) | [clear all](#) | [expand all](#) | [collapse all](#)

Type a taxonomic name

- Plasmodium adleri 322
- Plasmodium berghei 370
- Plasmodium billcollinsi 375

Filter results

Hide zero counts

Genome Genes

Clear filter 16,684

Filter Gene fields

[select all](#) | [clear all](#)

- Alternate product descriptions 7
- EC descriptions and numbers 13,510
- GO terms 5,432
- InterPro domains 6,955
- Orthologs 7,136
- PDB chains 4,814
- Product descriptions 6,007
- PubMed 689
- Rodent malaria phenotype 87
- User comments 255

Filter organisms

[select all](#) | [clear all](#) | [expand all](#) | [collapse all](#)

Type a taxonomic name

- Plasmodium adleri 322
- Plasmodium berghei 370
- Plasmodium billcollinsi 375

Filter results

Hide zero counts

Genome Genes

Clear filter 6,007

Filter Gene fields

[select all](#) | [clear all](#)

- Alternate product descriptions 7
- EC descriptions and numbers 13,510
- GO terms 5,432
- InterPro domains 6,955
- Orthologs 7,136
- PDB chains 4,814
- Product descriptions 6,007
- PubMed 689
- Rodent malaria phenotype 87
- User comments 255

Filter organisms

[select all](#) | [clear all](#) | [expand all](#) | [collapse all](#)

Type a taxonomic name

- Plasmodium adleri 157
- Plasmodium berghei 121
- Plasmodium billcollinsi 147
- Plasmodium blacklocki 144
- Plasmodium chabaudi 122
- Plasmodium coatneyi 117
- Plasmodium cynomolgi 244
- Plasmodium falciparum 2,353
- Plasmodium falciparum 3D7 145
- Plasmodium falciparum 7G8 147
- Plasmodium falciparum CD01 146
- Plasmodium falciparum Dd2 146
- Plasmodium falciparum GA01 147
- Plasmodium falciparum GB4 149
- Plasmodium falciparum GN01 146
- Plasmodium falciparum HB3 145
- Plasmodium falciparum IT 147
- Plasmodium falciparum KE01 146
- Plasmodium falciparum KH01 146
- Plasmodium falciparum KH02 148
- Plasmodium falciparum ML01 151
- Plasmodium falciparum SD01 146
- Plasmodium falciparum SN01 147
- Plasmodium falciparum TG01 151
- Plasmodium fragile 94
- Plasmodium gaboni 302
- Plasmodium gallinaceum 129
- Plasmodium iui 111
- Plasmodium knowlesi 241
- Plasmodium malariae 125
- Plasmodium ovale curtisi 125
- Plasmodium praefalciparum 144
- Plasmodium reichenowi 291
- Plasmodium relictum 128
- Plasmodium vinckei 213
- Plasmodium vivax 247

4. How many of the above genes are found in *Plasmodium falciparum* 3D7? How did you find this number? (hint: explore the *Filter organisms* section of the results filter). Select the correct organism and apply the filter.

Filter organisms

[select all](#) | [clear all](#) | [expand all](#) | [collapse all](#)

Type a taxonomic name

Plasmodium falciparum 3D7

[Add these](#) | [Clear these](#) | [Select only these](#)

[Select all](#) | [Clear all](#)

- Plasmodium falciparum 2,238
- Plasmodium falciparum 3D7 138

Filter organisms

[select all](#) | [clear all](#) | [expand all](#) | [collapse all](#)

Type a taxonomic name

- Plasmodium adleri 157
- Plasmodium berghei 121
- Plasmodium billcollinsi 147
- Plasmodium blacklocki 144
- Plasmodium chabaudi 122
- Plasmodium coatneyi 117
- Plasmodium cynomolgi 244
- Plasmodium falciparum 2,353
- Plasmodium falciparum 3D7 145
- Plasmodium falciparum 7G8 147
- Plasmodium falciparum CD01 146
- Plasmodium falciparum Dd2 146
- Plasmodium falciparum GA01 147
- Plasmodium falciparum GB4 149
- Plasmodium falciparum GN01 146
- Plasmodium falciparum HB3 145
- Plasmodium falciparum IT 147
- Plasmodium falciparum KE01 146
- Plasmodium falciparum KH01 146
- Plasmodium falciparum KH02 148
- Plasmodium falciparum ML01 151
- Plasmodium falciparum SD01 146
- Plasmodium falciparum SN01 147
- Plasmodium falciparum TG01 151
- Plasmodium fragile 94
- Plasmodium gaboni 302
- Plasmodium gallinaceum 129
- Plasmodium iui 111
- Plasmodium knowlesi 241
- Plasmodium malariae 125
- Plasmodium ovale curtisi 125
- Plasmodium praefalciparum 144
- Plasmodium reichenowi 291
- Plasmodium relictum 128
- Plasmodium vinckei 213
- Plasmodium vivax 247

5. Export the results to a search strategy. (hint: to achieve this click on the blue *Export as a search strategy* button at the top right-hand side of the results).

The screenshot shows the 'My Search Strategies' interface. At the top, there is a blue button labeled 'Export as a Search Strategy' with a white arrow pointing right, followed by the text 'to download or data mine'. A red arrow points downwards from this button to the search results page below.

The search results page has a header 'My Search Strategies' with links to 'Opened (1)', 'All (403)', 'Public (42)', and 'Help'. Below this is a section titled 'Unnamed Search Strategy *' with a 'Text' tab selected, showing '145 Genes'. There is also a 'Add a step' button.

The main results table is titled '145 Genes (121 ortholog groups) [Revise this search]'. It includes columns for 'Gene ID', 'Transcript ID', 'Organism', 'Genomic Location (Gene)', and 'Product Description'. The results show three entries:

| Gene ID | Transcript ID | Organism | Genomic Location (Gene) | Product Description |
|---------------|-----------------|------------------------------|---------------------------------|------------------------------|
| PF3D7_0102600 | PF3D7_0102600.1 | <i>Plasmodium falciparum</i> | Pf3D7_01_v3:119,041..121,249(-) | serine/threonine kinase, FIK |
| PF3D7_0103700 | PF3D7_0103700.1 | <i>Plasmodium falciparum</i> | Pf3D7_01_v3:166,513..168,687(+) | L-seryl-RN putative |
| PF3D7_0107600 | PF3D7_0107600.1 | <i>Plasmodium falciparum</i> | Pf3D7_01_v3:314,618..319,405(+) | eukaryotic factor 2-alp |

6. Return to the site search results page. How did you do this? (hint: you can achieve this in two ways: 1. Click on your browser's back arrow. 2. Click on the back to results arrow in the site search window. Notice that your previous results and filter settings were preserved.



7. Clear all filters. How did you do this? (hint: you can achieve this in two ways: 1. You can click on each of the clear filter options in the filter results panel on the left (boxes below). 2. You can click on the single *clear filters* option in the site search window.

The screenshot shows the Biogrid search interface. On the left, there are three filter panels: 'Filter results' (with a red box around the 'Clear filter' button), 'Filter Gene fields' (with a red box around the 'Clear filter' button), and 'Filter organisms' (with a red box around the 'Clear filter' button). At the top right, there is a search bar containing the word 'kinase' and a 'Clear filters' button (also highlighted with a red box).

Filter results

Genome Genes 145

Hide zero counts

Clear filter

Filter Gene fields

select all | clear all

| | | |
|-------------------------------------|--------------------------------|-----|
| <input type="checkbox"/> | Alternate product descriptions | 3 |
| <input type="checkbox"/> | EC descriptions and numbers | 320 |
| <input type="checkbox"/> | GO terms | 156 |
| <input type="checkbox"/> | InterPro domains | 165 |
| <input type="checkbox"/> | Orthologs | 164 |
| <input type="checkbox"/> | PDB chains | 112 |
| <input checked="" type="checkbox"/> | Product descriptions | 145 |
| <input type="checkbox"/> | PubMed | 126 |
| <input type="checkbox"/> | Rodent malaria phenotype | 42 |
| <input type="checkbox"/> | User comments | 50 |

Clear filter

Filter organisms

select all | clear all | expand all | collapse all

| | | |
|-------------------------------------|------------------------|-------|
| <input type="checkbox"/> | Plasmodium adleri | 157 |
| <input type="checkbox"/> | Plasmodium berghei | 121 |
| <input type="checkbox"/> | Plasmodium bilcollinsi | 147 |
| <input type="checkbox"/> | Plasmodium blacklocki | 144 |
| <input type="checkbox"/> | Plasmodium chabaudi | 122 |
| <input type="checkbox"/> | Plasmodium coatneyi | 117 |
| <input type="checkbox"/> | Plasmodium cynomolgi | 244 |
| <input checked="" type="checkbox"/> | Plasmodium falciparum | 2,353 |
| <input type="checkbox"/> | Plasmodium fragile | 0 |

Clear filter

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

The image shows two side-by-side 'Filter results' panels. A red arrow points from the first panel to the second, highlighting the difference in the count for 'Genomic sequences'.

Filter results (Left Panel, Hide zero counts checked)

- Genome Genes 16,684
- Population biology Popset isolate sequences 249
- Metabolism Metabolic pathways Compounds 345 85
- Data access Data sets Searches 1 3

Filter results (Right Panel, Hide zero counts unchecked)

- Genome Genes 16,684 0
- Genomic sequences 0
- Organism Organisms 0
- Transcriptomics ESTs 0
- Population biology Popset isolate sequences Field samples 249 0
- Metabolism Metabolic pathways Compounds 345 85
- Data access Data sets Searches 1 3
- Instructional Tutorials Workshop exercises 0 0
- About News General info pages 0 0

9. Try running a search with a wild card. The wild card is denoted by an asterisk *. The wild card can be used alone to retrieve all results available to the site search 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.

All results matching *

1 - 20 of 516,501

| Compound - CHEBI:10000 | Vismione D | |
|---|----------------------|--|
| Compound - CHEBI:10001 | Visnadin | |
| Compound - CHEBI:10002 | Visnagin | |
| Compound - CHEBI:10003 | ribostamycin sulfate | |
| Definition: An aminoglycoside sulfate salt resulting from the reaction of ribostamycin with sulfuric acid. | | |
| Compound - CHEBI:100147 | naldixic acid | |
| Definition: A monocarboxylic acid comprising 1,8-naphthyridin-4-one substituted by carboxylic acid, ethyl and methyl groups at positions 3, 1, and 7, respectively. | | |
| Compound - CHEBI:10014 | Voacamine | |
| Compound - CHEBI:10015 | vobasine | |
| Definition: An indole alkaloid that is vobasine in which the bridgehead methyl group is substituted by a methoxycarbonyl group and an additional oxo substituent is present in the 3-position. | | |
| Compound - CHEBI:10016 | vobtusine | |
| Compound - CHEBI:10017 | volementol | |
| Definition: A heptitol that is heptane-1,2,3,4,5,6,7-heptol that has R-configuration at positions 2, 3, 5 and 6. | | |
| Compound - CHEBI:10018 | volkenin | |
| Definition: A cyanogenic glycoside that is (4R)-4-hydroxycyclopent-2-ene-1-carbonitrile attached to a beta-D-glucopyranosyloxy at position 1. | | |
| Compound - CHEBI:10019 | Vomicine | |

All results matching *kinase

1 - 20 of 18,073

| Gene - AK88_00104 | CK1/CK1/CK1-D protein kinase | |
|--|------------------------------|--|
| Organism: Plasmodium fragile strain nilgiri | | |
| ▶ Fields matched: EC descriptions and numbers; GO terms; InterPro domains; Orthologs; PDB chains; Product descriptions | | |
| Gene - AK88_00479 | CAMK protein kinase | |
| Organism: Plasmodium fragile strain nilgiri | | |
| ▶ Fields matched: EC descriptions and numbers; GO terms; InterPro domains; Orthologs; PDB chains; Product descriptions | | |
| Gene - AK88_00505 | pantothenate kinase | |
| Organism: Plasmodium fragile strain nilgiri | | |
| ▶ Fields matched: EC descriptions and numbers; GO terms; Orthologs; PDB chains; Product descriptions | | |
| Gene - AK88_00565 | Atypical/ABC1 protein kinase | |
| Organism: Plasmodium fragile strain nilgiri | | |
| ▶ Fields matched: InterPro domains; Orthologs; Product descriptions | | |

10. Try searching for a specific gene ID. Enter the gene ID below in the site search window:

PF3D7_0310100

The screenshot shows the PlasmoDB search results for the gene ID PF3D7_0310100. The search bar at the top contains the query "PF3D7_0310100". The main content area displays "Genes matching PF3D7_0310100" with a count of 1-2 of 2. On the left, there is a sidebar titled "Filter results" showing counts for Genomes (2) and Genes (2). Below it are sections for "Filter Gene fields" (External links, Gene ID, Notes from annotators) and "Filter organisms" (Plasmodium falciparum, Plasmodium gaboni). The right side shows the detailed results for the top hit: "Gene - PF3D7_0310100 calcium-dependent protein kinase 3". It provides the gene name or symbol (CDPK3), organism (Plasmodium falciparum 3D7), and notes that fields matched External links; Gene ID. There is also a link to "Gene - PGSY75_0310100 calcium-dependent protein kinase 3" which has the same details. A blue button in the top right corner says "Export as a Search Strategy" with a download icon.

Notice that the gene of interest appears at the top for easy access. You can click on the Gene ID to go the gene page.