

Site Search in VEuPathDB

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 (*)

Introduction

The site search is located in the header of any VEuPathDB site and is available on 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. Go to [PlasmoDB.org](#)¹ and search for a keyword. Enter the word *kinase* in the site search window. Then click enter on your keyboard or click on the search icon.



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 number² of results with the word kinase?

¹ Note that this exercise uses PlasmoDB.org as an example database, but the same functionality is available on all VEuPathDB resources.

² VEuPathDB is updated regularly, so the numbers you see while completing the exercise may differ slightly from those shown in the screenshots.

Are all the results genes?

All results matching kinase

[Export as a Search Strategy](#) ►

1 - 20 of 20,497

1 2 3 ... 1,025

The screenshot shows a search results page for the term "kinase". On the left, there is a sidebar titled "Filter results" with sections for Genome, Population biology, Metabolism, Data access, and Filter fields. The "Genes" section is highlighted with a red arrow. Below it, there is a "Filter organisms" section with a search bar and a list of taxonomic names. The main content area displays a list of items, each with a title, gene type, organism, and a "Fields matched" section. The first item is "Gene - PCYB_132500 kinase".

Results are summarized by category.

Details panel with information about each item returned.

2. Filter the site search result by category.

How many genes included the word kinase in their product descriptions?

Filter the results so that you can only view gene results, and the Filter Fields section expands to reveal additional filtering options. Select the *Product descriptions* field and choose *Apply*. Once a filter is applied, it can be removed by clicking on *Clear filter* (right panel below).

The three screenshots illustrate the filtering process:

- Left Panel:** Shows the initial search results with the "Genes" filter selected. A red arrow points to the "Genes" link in the "Filter results" section.
- Middle Panel:** The "Filter Gene fields" section is expanded. The "Product description" checkbox is selected, and the "Apply" button is highlighted with a red box. A red arrow points to the "Product description" checkbox.
- Right Panel:** The results are now filtered to show only genes with product descriptions. The "Clear filter" button is highlighted with a red box.

3. Filter the site search result by organism: How many of the above genes are found in *Plasmodium falciparum* 3D7?

(Hint: Explore the *Filter organisms* section of the results filter and use the search filter to navigate the tree.

The figure consists of two side-by-side screenshots of a search interface. Both screenshots show a 'Filter results' section at the top with a 'Clear filter' button and a count of 7,630. Below this is a 'Filter Gene fields' section with various checkboxes and counts. The bottom section is 'Filter organisms'.

Left Screenshot: The 'Organism' dropdown is set to '3d7'. An arrow points to this dropdown. Below it, under 'select only these', 'Plasmodiidae' is checked. Another arrow points to this checkbox. Under 'Reference only', 'Plasmodium falciparum 3D [Ref]' is checked. A third arrow points to this checkbox.

Right Screenshot: The 'Organism' dropdown is set to '3d7'. An arrow points to the 'Apply' button. Below it, under 'select only these', 'Plasmodiidae' is checked. Another arrow points to this checkbox. Under 'Reference only', 'Plasmodium falciparum 3D [Ref]' is checked. A third arrow points to this checkbox.

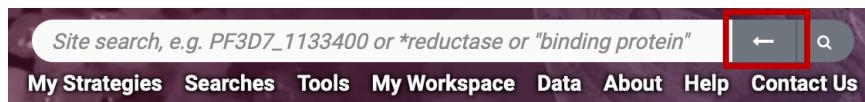
4. Export the results to a search strategy. Use the blue *Export as a Search Strategy* button at the top right-hand side of the results. Once exported, you can take advantage of over 100 specialized searches using the Add Step button. We will learn more about this in a future exercise.

A screenshot of the search results page for 'kinase'. The title is 'Genes matching kinase (filtered by fields and organisms)'. The results count is 1 - 20 of 137. At the top right is a blue button labeled 'Export as a Search Strategy' with a red circle around it. Below the results, there is a summary for 'Gene - PF3D7_0616000 pyridoxal kinase'.

A screenshot of the 'My Search Strategies' page. It shows an 'Opened (1)' strategy named 'Unnamed Search Strategy'. The strategy contains one step: 'Text 137 Genes'. A red arrow points to the 'Add a step' button next to the text input field.

A screenshot of the search results page for 'kinase' genes. The title is '137 Genes (113 ortholog groups)'. The results are displayed in a table with columns for Gene ID, Transcript ID, Organism, and Interpro Description. Two specific genes are highlighted: PF3D7_0102600.1 and PF3D7_0103700.1.

5. 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.



6. Clear all filters. You can achieve this in two ways: 1. click on each clear filter option in the filter results panel. 2. You can click on the *clear filters* option in the site search window to Clear All filters.

Screenshot 1: Filter results panel. Three "Clear filter" buttons are highlighted with red boxes: one under "Filter results" (Genome Genes), one under "Filter Gene fields" (Product descriptions), and one under "Filter organisms" (Haemoproteidae).

Screenshot 2: Site search results page for "kinase". The "Clear filters" button is highlighted with a red box.

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

Left Panel (Checked):

Category	Value	Count
Genome	Genes	19,349
Population biology	Popset isolate sequences	352
Metabolism	Metabolic pathways	352
	Compounds	193
Data access	Data sets	1
	Searches	3

Right Panel (unchecked):

Category	Value	Count
Genome	Genes	19,349
	Genomic sequences	0
Organism	Organisms	0
Transcriptomics	ESTs	0
Population biology	Popset isolate sequences	352
	Field samples	0
Metabolism	Metabolic pathways	352
	Compounds	193
Data access	Data sets	1
	Searches	3
Instructional	Tutorials	0
	Workshop exercises	0
About	News	0
	General info pages	0

8. 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 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 when you run a search or the word *kinase* to a search with a wild card **kinase* or **kinase**.

All results matching *		Export as a Search Strategy to download or mine your results					
1 - 20 of 978,719		◀ 1 2 3 ... 48,936 ▶					
<input checked="" type="checkbox"/> Hide zero counts							
Filter results							
Genome	356,359						
Genes	356,359						
Genomic sequences	22,633						
Organism	63						
Organisms	63						
Transcriptomics	287,336						
ESTs	287,336						
Population biology	153,109						
Popset isolate sequences	153,109						
Compound - CHEBI:100000 (2S,3S,4R)-3-[4-(3-cyclopentenylprop-1-ynyl)phenyl]-4-(hydroxymethyl)-1-(2-methoxy-1-oxyethyl)-2-azetidinocarbonitrile							
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-triene-2-yl}methyl]amino]propan-1-ol							
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-2-yl]methanol							
Compound - CHEBI:100004 1-(2,5-difluorophenyl)-3-[{(5S,6S,9S)-5-methoxy-3,6,9-trimethyl-2-oxo-8-[{[2-pyrazinyl]methyl}]-11-oxa-3,8-diazabicyclo[10.4.0]hexadeca-1(12),13,15-triene-2-yl]methyl]propan-1-ol							
Compound - CHEBI:100005 N-[{(1S,3S,4S,9aR)-1-(hydroxymethyl)-3-[2-oxo-2-(1-piperidinyl)ethyl]-3,4,4,9a-tetrahydro-1H-pyran]3-[{[2-(2,5-difluorophenyl)-oxomethyl]amino}-2-(hydroxymethyl)}-3,6-dihydro-2H-pyran-2-yl]benzofuran-5-carboxamide							
Compound - CHEBI:100006 N-[{(1,3-benzodioxol-5-yl)methyl}-2-[(2R,3S,6S)-3-[[{[2-(2,5-difluorophenyl)-oxomethyl]amino}-2-(hydroxymethyl)}-3,6-dihydro-2H-pyran-2-yl]benzofuran-5-carboxamide]							
Compound - CHEBI:100007 LSM-11386							

All results matching *kinase		Export as a Search Strategy to download or mine your results				
1 - 20 of 23,170		◀ 1 2 3 ... 1,159 ▶				
<input checked="" type="checkbox"/> Hide zero counts						
Filter results						
Genome	21,205	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; Product description; Product descriptions (all)				
Genes	21,205	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; Product description; Product descriptions (all)				
Population biology	1,273	Gene - AK88_00505 pantothenate kinase Gene type: protein coding gene				
Popset isolate sequences	1,273					
Metabolism	484					
Metabolic pathways	484					
Compounds	204					
Data access	1					
Data sets	1					
Searches	3					

9. 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

Hide zero counts

Genome	Count
Genes	2

Filter Gene fields

select all | collapse all

Field	Count
External links	1
Gene ID	1
Names, IDs, and aliases	1
Notes from annotators	1

Filter organisms

select all | collapse all | expand all | all

Type a taxonomic name Reference only

Taxon	Count
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 - PGYS7_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—orf

Direct link to

Why is this a search for

[Direct link to PE3D7_0310100](#)

Why is this gene returned by
a search for PE3D7_0310100?