

Creating and modifying search strategies in FungiDB: Site Search

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).



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 the image below).

All results matching **kinase**

Export as a Search Strategy to download or mine your results

1 - 20 of 157,362

1 2 3 ... 7,869

Filter results

☒ Hide zero counts

Genome
Genes ← 151,162
Population biology
Popset isolate sequences 5,796
Metabolism
Metabolic pathways 309
Compounds 80
Data access
Data sets 10
Searches 4
About
News 1

Filter fields

Select a result filter above

Filter organisms

select all | clear all | expand all | collapse all

Type a taxonomic name

☐ Fungi 119,392
☐ Oomycota 31,770

Data set - C. neoformans kinase phenome database

Summary: Links to Cryptococcus neoformans kinase phenome database.

Fields matched: Data set; Description; Name; Principal Investigator and Collaborators; Summary

Data set - Analysis of the protein kinase A-regulated proteome of Cryptococcus neoformans

Fields matched: Associated publications; Data set; Description; Name

Gene - AN7743 Predicted G-protein coupled pheromone receptor

Gene name or symbol: preA
Organism: Aspergillus nidulans FGSC A4

Fields matched: Phenotype

Gene - CAGL0L03520g Ortholog(s) have MAP kinase kinase kinase activity

Gene name or symbol: BCK1
Organism: Candida glabrata CBS 138

Fields matched: GO terms; Notes from annotators; Orthologs; PDB chains; Product descriptions

Gene - CGB_I0230W MAP kinase kinase kinase, putative

Organism: Cryptococcus gattii WM276

Fields matched: EC descriptions and numbers; GO terms; Orthologs; PDB chains; Product descriptions

Gene - ASPNIDRAFT2_1098583 MAP kinase kinase kinase

Organism: Aspergillus niger ATCC 1015

COMMUNITY CHAT

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 with 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).

The sequence of panels illustrates the filtering workflow:

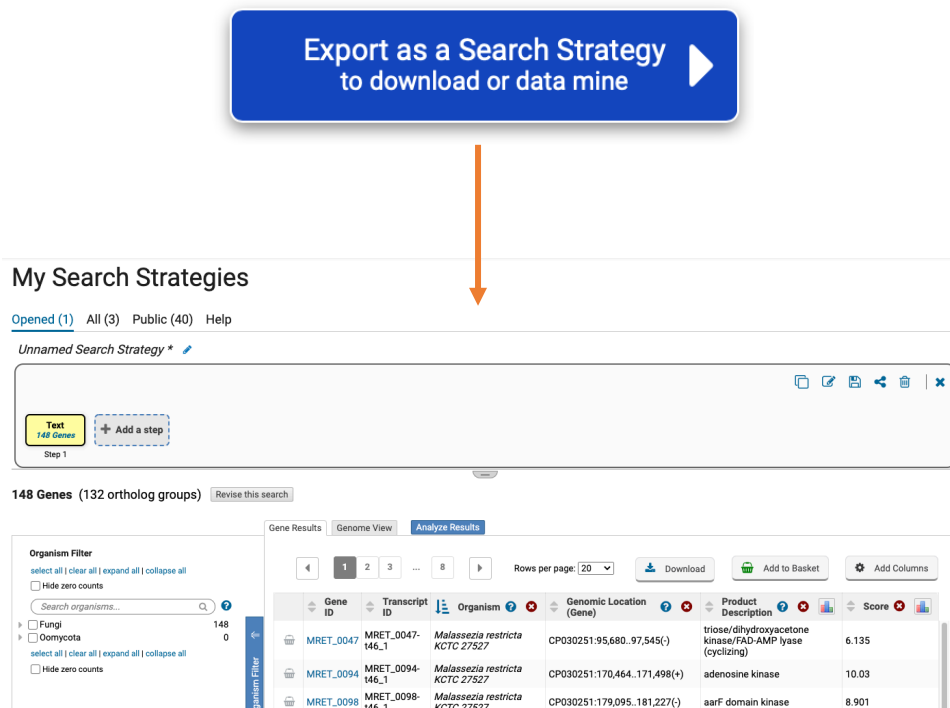
- Panel 1:** 'Filter results' shows 'Genome' (119,392) and 'Genes' (151,162). 'Filter Gene fields' is collapsed.
- Panel 2:** 'Filter Gene fields' is expanded, showing fields like 'Alternate product descriptions' (11), 'EC descriptions and numbers' (50,235), etc. 'Product descriptions' is selected, showing 27,159 results. The 'Apply' button is highlighted.
- Panel 3:** The 'Apply' button is clicked, and the 'Clear filter' button is highlighted.
- Panel 4:** The 'Filter organisms' panel shows a taxonomic tree with 'Malasseziomycetes' and 'Malassezia restricta' selected.

- How many of the above genes are found in *Malassezia restricta* species? Can you find gene numbers for individual strains? (hint: explore the *Filter organisms* section of the results filter). Select *Malassezia restricta* KCTC 27527 only and apply the filter.

The final filtering steps are shown in two panels:

- Panel 1:** 'Filter organisms' panel shows 'Malassezia restricta' selected, showing 283 results. The 'Apply' button is highlighted.
- Panel 2:** The 'Filter organisms' panel shows 'Malassezia restricta' selected, showing 283 results. The 'Apply' button is highlighted.

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).



Export as a Search Strategy
to download or data mine

My Search Strategies

Opened (1) All (3) Public (40) Help

Unnamed Search Strategy *

Text 148 Genes Step 1

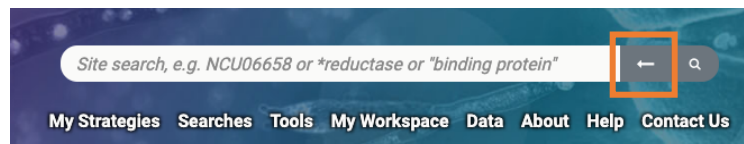
148 Genes (132 ortholog groups) [Revise this search](#)

Gene ID	Transcript ID	Organism	Genomic Location (Gene)	Product Description	Score
MRET_0047	MRET_0047-146_1	Malassezia restricta KCTC 27527	CP030251:95,680..97,545(-)	triase/dihydroxyacetone kinase/PAD-AMP lyase (cyclizing)	6.135
MRET_0094	MRET_0094-146_1	Malassezia restricta KCTC 27527	CP030251:170,464..171,498(+)	adenosine kinase	10.03
MRET_0098	MRET_0098-146_1	Malassezia restricta KCTC 27527	CP030251:179,095..181,227(-)	aarF domain kinase	8.901

6. Return to the site search results page.

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 (see below). Notice that your previous results and filter settings were preserved.



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 on the left (boxes below).
2. You can click on the single *clear filters* option in the site search window.

1

2

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

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.

Filter results

☒ Hide zero counts

Genome 1,885,291
Genes 162,441
Organism 186
Transcriptomics 1,709,817
ESTs 1,077,920
Population biology 3,045
Popset isolate sequences 61,998
Metabolism 381
Metabolic pathways 435
Compounds 15
Data access 1
Data sets 2
Searches 16
Instructional 1
Tutorials 1
Workshop exercises 2
About 16
News
General info pages

Filter fields
Select a result filter above

Filter organisms
select all | clear all | expand all | collapse all
Type a taxonomic name

1 - 20 of 4,901,548

Export as a Search Strategy to download or mine your results

Compound - CHEBI:10000 Vismione D
Compound - CHEBI:10001 Vianadin
Compound - CHEBI:10002 Vianagin
Compound - CHEBI:10003 ribostamycin sulfate
Definition: An aminoglycoside sulfate salt resulting from the reaction of ribostamycin with sulfuric acid.
Compound - CHEBI:10017 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 Vocamine
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 volemitol
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
Compound - CHEBI:10022 Vomoxin
Compound - CHEBI:10023 voriconazole
Definition: A triazole-based antifungal agent used for the treatment of esophageal candidiasis, invasive pulmonary aspergillosis, and serious fungal infections caused by *Scedosporium apiospermum* and *Fusarium* spp. It is an inhibitor of cytochrome P450 3C9 (CYP2C9) and CYP3A4.
Compound - CHEBI:100241 ciprofloxacin
Definition: A quinolone that is quinolin-4(1H)-one bearing cyclopropyl, carboxylic acid, fluoro and piperazin-1-yl substituents at positions 1, 3, 6 and 7, respectively.

COMMUNITY CHAT

If the wild card is combined with a word such as **kinase*, the search will 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 ***kinase**

1 - 20 of 167,035

Filter results

☒ Hide zero counts

Genome
Genes 159,776

Population biology
Popset isolate sequences 6,731

Metabolism
Metabolic pathways
Compounds 425
88

Data access
Data sets 10
Searches 4

About
News 1

Filter fields
Select a result filter above

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

Type a taxonomic name

☐ Fungi 125,830
☐ Oomycota 33,946

Data set - C. neoformans kinase phenome database
Summary: Links to Cryptococcus neoformans kinase phenome database.
Fields matched: Data set; Description; Name; Principal Investigator and Collaborators; Summary

Data set - Analysis of the protein kinase A-regulated proteome of Cryptococcus neoformans
Fields matched: Associated publications; Data set; Description; Name

Gene - ACLA_001680 nucleoside diphosphate kinase
Organism: Aspergillus clavatus NRRL 1
Fields matched: EC descriptions and numbers; GO terms; Orthologs; PDB chains; Product descriptions

Gene - ACLA_001720 thymidylate kinase
Organism: Aspergillus clavatus NRRL 1
Fields matched: EC descriptions and numbers; GO terms; Orthologs; PDB chains; Product descriptions

Gene - ACLA_001930 pyruvate dehydrogenase kinase, putative
Organism: Aspergillus clavatus NRRL 1
Fields matched: GO terms; Orthologs; PDB chains; Product descriptions

Gene - ACLA_001990 protein kinase, putative
Organism: Aspergillus clavatus NRRL 1
Fields matched: EC descriptions and numbers; GO terms; Orthologs; PDB chains; Product descriptions

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

CNAG_00047

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

FungiDB Release 31 16 Mar 2021
Fungal & Oomycete Informatics Resources

Search: CNAG_00047 Clear filters

My Strategies Searches Tools My Workspace Data About Help Contact Us

Genes matching **CNAG_00047** (filtered by organisms)

1 - 1 of 1

Filter results

☒ Hide zero counts

Genome
Genes 1

Filter Gene fields
select all | clear all

☐ Gene ID 1

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

Type a taxonomic name

☒ Fungi 1
☒ Basidiomycota 1

Gene - CNAG_00047 pyruvate dehydrogenase kinase
Gene name or symbol: PKP1
Organism: Cryptococcus neoformans var. grubii H99
Fields matched: Gene ID

Gene - CNAG_00047 pyruvate dehydrogenase kinase
Gene name or symbol: PKP1
Organism: Cryptococcus neoformans var. grubii H99
Fields matched: Gene ID

1 - 1 of 1