

Site Search

Learning objectives:

- Deploy site search using a keyword.
- Filter site search results by categories, by organisms, and by category fields.
- Export results to a search strategy.
- Find a specific gene using its ID in site search.
- Understand how organism preferences filter works.

The site search is located in the header of the site and is available from every page. The site search queries the database for a term (e.g., text) or ID and returns a list of pages and documents that contain the query term.

Site search: deploy site search with a keyword

1. Enter the word kinase in the site search window (at the top centre of the page). Click on the "enter" key on your keyboard or on the search icon as shown in the screenshot below.



2. How many results were returned by your search? Are all of these records genes?
3. Filter the results to view results for genes only (hint: click on the word “Genes” in the *Filter results* section):

A screenshot of the search results page for 'kinase'. The title is 'All results matching kinase'. It shows 1 - 20 of 337,992 results. A blue button in the top right corner says 'Export as a Search Strategy'. On the left, there are two sections: 'Filter results' and 'Filter organisms'. In 'Filter results', under 'Genome', the 'Genes' option is selected and highlighted with a red arrow. In 'Filter organisms', there are checkboxes for 'Fungi' and 'Oomycota'. The main area lists search results with details like gene name, type, and organism. At the bottom, there is a 'Filter Gene fields' sidebar with a list of checkboxes and their counts.

Notice that clicking on the “Genes” category reveals additional filtering options.

A screenshot of the 'Filter Gene fields' sidebar. It has a 'select all | clear all' button at the top. Below it is a list of checkboxes with their counts: Alternate product descriptions (15), Apollo Annotations (1), EC descriptions and numbers (81,470), GO terms (80,460), InterPro domains (105,190), Notes from annotators (911), Orthologs (283,464), PDB chains (66,887), Phenotype (14), Preferred product description (44,673), Product descriptions (44,202), PubMed (5,955), and User comments (178).

4. Apply the “Product descriptions” filter to look for genes that have “kinase” in the “product description” field only.

Note: Any applied filter can be easily cleared by clicking on “Clear filter” option.

The left screenshot shows the initial state of the 'Filter Gene fields' section with the 'Product descriptions' checkbox checked. The right screenshot shows the result after applying the filter, where the count has changed from 329,564 to 44,202.

Filter Gene fields	Count
Product descriptions	44,202

5. Filter gene results for a single organism - *Cryptococcus neoformans* var. *grubii* H99.

Hint: use “H99” in the filter box to quickly identify the genome of interest.

The screenshot shows the 'Filter organisms' interface with the search bar set to 'H99'. The 'Cryptococcus neoformans' var. 'grubii' H99 genome is selected, indicated by a pink checkmark.

6. Export search results to a search strategy in FungiDB.

To achieve this, click on the blue button called “Export as a search strategy...” at the top right-hand side of the results page.

Notice that before the Genes category was selected this button was inactive. This is because the search strategy can be deployed on a signle category only (e.g. Genes or Go terms, but not both).

The screenshot shows the 'Export as a Search Strategy' button being clicked, which leads to the 'Unnamed Search Strategy' interface. The 'Genes' category is selected, and the search results table shows various genes related to Cryptococcus neoformans var. grubii H99.

Gene ID	Transcript ID	Organism	Genomic Location (Gene)	Product Description	Score
CNAG_00047	CNAG_00047-26_1	<i>Cryptococcus neoformans</i> var. <i>grubii</i> H99	CP003820.1;139,553..141,970(-)	pyruvate dehydrogenase kinase	9.174
CNAG_00106	CNAG_00106-26_1	<i>Cryptococcus neoformans</i> var. <i>grubii</i> H99	CP003820.1;291,793..296,797(-)	Two-component-like sensor kinase	7.673
CNAG_00130	CNAG_00130-26_1	<i>Cryptococcus neoformans</i> var. <i>grubii</i> H99	CP003820.1;333,191..338,332(-)	CAMK/CAMK1/CAMK1-ROCK protein kinase	8.357
CNAG_00363	CNAG_00363-26_1	<i>Cryptococcus neoformans</i> var. <i>grubii</i> H99	CP003820.1;922,612..928,950(-)	Two-component-like sensor kinase	7.673
CNAG_00394	CNAG_00394-26_1	<i>Cryptococcus neoformans</i> var. <i>grubii</i> H99	CP003820.1;1,039,957..1,042,941(-)	AGC/PKA protein kinase	8.357
CNAG_00405	CNAG_00405-26_1	<i>Cryptococcus neoformans</i> var. <i>grubii</i> H99	CP003820.1;1,064,014..1,067,321(-)	ste/ste20/ysk protein kinase	9.174
CNAG_00415	CNAG_00415-26_1	<i>Cryptococcus neoformans</i> var. <i>grubii</i> H99	CP003820.1;1,089,398..1,093,237(-)	CMGC/CDK protein kinase	8.357
CNAG_07359	CNAG_07359-26_1	<i>Cryptococcus neoformans</i> var. <i>grubii</i> H99	CP003820.1;1,122,440..1,126,595(-)	Abi/p65/ABC1/ABC1-C protein kinase	8.357

At this point you can name the strategy and save it to your account. If you don't have a FungiDB account, it is easy to create. FungiDB is free and your information will never be shared with other parties.

7. Try running the same search but this time use a wild card (*) (e.g., kinase*).

The significance of the wild card: When the wild card is combined with a word (e.g., kinase* or *kinase), the search will retrieve compound words ending or beginning with the word kinase (e.g. *kinase = phosphofructokinase). The wild card (*) can be also used alone to retrieve all records available to the site search (see screenshot below).

All results matching *

1 - 20 of 4,901,548

Filter results

Filter fields

Filter organisms

Export as a Search Strategy

COMMUNITY CHAT

Site search: deploy site search with a GeneID.

1. Run a site search with the following gene id: C1_00800C_A

Notice that the output of the search produces a gene record card for C1_00800C_A (see screenshot below). Click on the gene link in blue to be redirected to the gene record page for this gene.

Genes matching C1_00800C_A

1 - 1 of 1

Gene - C1_00800C_A C-8 sterol isomerase [Source:UniProtKB/TrEMBL;Acc:A0A1D8PCB9]

Gene name or symbol: ERG2
Gene type: protein coding gene
Organism: Candida albicans SC5314

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

Filter results

Filter Gene fields

Filter organisms

Export as a Search Strategy

2. Export the result as a search strategy. This feature may be useful if you are interested in cross-referencing different types of data (e.g., transcriptomics, SNPs, etc.) for a single gene or interest.

(optional) Site search: find a dataset using a PubMedID.

1. Run a site search to identify records for the “26453029”, which is a PubMedID associated with Geddes et al. 2015 publication - <https://pubmed.ncbi.nlm.nih.gov/26453029/>.
2. Filter results on “Data sets”.
3. Click on the first dataset to be redirected to a dataset record page.

The screenshot shows the FungiDB homepage with a search bar containing "26453029". A red circle labeled "1" highlights the search bar. Below it, a navigation bar includes "My Strategies", "Searches", "Tools", "My Workspace", "Data", "About", "Help", "Contact Us", and a user profile for "Evelina". A red circle labeled "2" highlights the "Data" menu item. The main content area displays "Data sets matching 26453029" with a red circle labeled "3" highlighting the first result. The result is for "Secretome profiling of Cryptococcus neoformans" and includes a summary, fields matched (Associated publications), and a link to "Export as a Search Strategy".

FungiDB dataset record pages contain information about the integrated datasets, associated publications, and links how to explore this data within the site.

The screenshot shows a detailed dataset record. At the top, it lists organism details: *Cryptococcus neoformans* var. grubii H99, category Protein expression, primary publication Geddes et al. BMC Microbiol 2015 Oct 09;15():206, and source version 2015-10-09. A red circle labeled "1" highlights the "Data Provenance" section, which contains a table of associated publications and a description of the experimental setup. A red circle labeled "2" highlights the "Link outs" section, which includes links to associated data sets, exploration within the website, and external resources. The "External Resources" section notes "No data available".

My Organism Preferences

The screenshot shows the FungiDB homepage with the title "FungiDB" and subtitle "Fungal & Oomycete Informatics Resources". The top navigation bar includes links for "My Strategies", "Searches", "Tools", "My Workspace", "Data", "About", "Help", and "Contact Us". On the right, there are social media icons for Twitter, Facebook, GitHub, YouTube, and LinkedIn, along with a user icon labeled "Evelina". A banner at the top right indicates "VEuPathDB Project". Below the header, a search bar says "Site search, e.g. NCU06658 or *reductase or *binding protein*". A link "My Organism Preferences (1 of 300)" is visible, followed by a toggle switch labeled "disabled".

The “My Organism Preferences” tool allows to cherry-pick any combination of organisms and apply these organism preferences globally across the site. This means that data, search menus and tools will be filtered based on the selected organism/s helping you focus on your work.

Set custom parameters for the “My Organism Preferences” filter.

1. Click on the My Organism Preference link at the top of the page.
2. Click on the “clear all” to remove any default filters.
3. Use the filter box to search for *Cryptococcus* and select all *Cryptoccaceae*
4. Click on the “Apply” button to activate the filter.

Screenshot 1: Shows the "My Organism Preferences (1 of 300)" section with a note: "You can always disable this preference to bring back the full set of organisms, and explore the data, tools and searches they offer." A toggle switch is shown as "disabled".

Screenshot 2: Shows the "Choose organisms to keep" interface. A red arrow points from the "My Organism Preferences" section to this screen. A red circle with the number "2" is on the left side of the interface. The search bar contains "cryptoc". Under "Fungi", "Basidiomycota", "Tremellomycetes", and "Tremellales" are expanded. Under "Tremellales", "Cryptoccaceae" is expanded, and several *Cryptococcus* species are listed with checkboxes. A green "Apply" button is at the bottom.

Screenshot 3: Shows the "Preview of My Organism Preferences (20 of 300)" section. A red arrow points from the "Choose organisms to keep" interface to this screen. A red circle with the number "3" is on the left side. The preview lists 20 *Cryptococcus* genomes. A green "Apply" button is at the bottom. The toggle switch is now "enabled".

Notice that the filter uses 19 genomes. This means that any searches deployed within FungiDB will be applied only to the 19 *Cryptococcus* genomes. If you want to run search in Candida or other organisms, you can add genomes to the filter or disable the filter via the toggle button on the right.

Explore how the My Organism Preference parameters affect your experience on FungiDB.

1. Use the “Search” menu at the top to navigate to the “RNA-Seq Evidence” search page. Notice that all available datasets are pre-filtered for Cryptococcus.
2. Disable “My Organism Preferences” filter.
3. Notice the change in the number of available RNA-Seq studies.

My Organism Preferences

1

Identify Genes based on RNA-Seq Evidence

Organism	Data Set
Cryptococcus neoformans var. grubii H99	Fluconazole Sensitivity of Wild-type vs cfo1 Mutant (Kim et al. 2012)
Cryptococcus neoformans var. grubii H99	Transcriptomic profiling of the Cryptococcus neoformans transition to titan-like cells (Trevijano-Contador et al. 2018)
Cryptococcus neoformans var. grubii H99	RNA-Seq of cells treated with eltrombopag (Ko et al. 2019)
Cryptococcus neoformans var. grubii H99	Virulence of WT vs rim101 Mutant Cells (O'Meara et al. 2013)
Cryptococcus neoformans var. grubii H99	The Cryptococcal central nervous system transcriptome during human disease (Yu et al. 2021)

27 rows

2

My Organism Preferences (19 of 300) enabled

3

My Organism Preferences (19 of 300) disabled

Identify Genes based on RNA-Seq Evidence

Organism	Data Set
Aspergillus fumigatus A1163	Transcriptomics analysis of Aspergillus fumigatus co-cultivated with Pseudomonas aeruginosa (Valiente 2018)
Aspergillus fumigatus Af293	Transcriptome analysis of conidium germination of Aspergillus fumigatus in different growth conditions (Danion et al. 2021)
Aspergillus fumigatus Af293	Gene expression in WT, hrmA deletion, hrmA OE, hrmA_REV, EVOL under hypoxia and normoxia conditions (Kowalski et al. 2019)
Aspergillus fumigatus Af293	Transcriptomes of WT, nctA, and nctB mutants in response to itraconazole. (Furukawa et al. 2020)
Aspergillus fumigatus Af293	Gene expression under oxidative and iron stresses (Kurucz et al. 2018)

164 rows