

Home

Home:

Description of the role of CopciDB.

News:

Welcoming the users. Brief summary of available genomes (contig/scaffold number, gene number). Description and indication of the source of the transcriptome datasets available on the website.

Search

Blast

It provides an interface to start blast searches in various databases. If the default setting of 0.001 E-value is too permissive for your search, you can select a stricter option in the 'Show Advanced Options' menu.

Available databases:

Assembly: perform a nucleotide sequence blastn search against a selected genome database

Transcripts: perform a nucleotide sequence blastn search against a selected transcript database

Proteins: perform a protein sequence blastp search against a selected protein database

Text Search

It provides an interface to start a search using keywords in the various subsections of the integrated gene annotation database.

Before searching, you can select all or one of the available *Coprinopsis cinerea* strains listed in the 'Organism' subsection. In the 'Find in' subsection, you can select any number of elements to further specify the searching scope (using all elements is the default behavior or use the Ctrl+A keyword command to select all the elements in the box if any deviation from the basic setting has occurred).

The search returns a table containing the 'Systematic name' of the identified database entry with the list of subsections name (Find in) that contain the word you are looking for. If a match is found, you can download the identified gene, transcript, CDS and protein sequence as a Fasta file by ticking the box in front of the result name and selecting the 'Download' menu above the results table. If you need information about the identified gene's expression, after selecting 'GET EXPRESSION DATA' and the desired *Coprinopsis* strain, you can easily plot the mean CPM for the different experimental groups and display it in tabular form by selecting the button DATA TABLE below each graph.

Description of the selectable element of the gene annotation database:

- **Systematic name:**
 - Search for entries that have a specified id (protein id)

- Example: 'CC1G_08232' for Okayama strain, 'CopciAB_61699' for the AmutBmut strain
- **Standard name:**
 - Search for entries that have a standard gene name
 - Example: 'ACT1', 'creA', 'GLN1'
- **Synonyms:**
 - Search for entries that have a specific alternative id (id from another database)
- **Uniprot id:**
 - Search for entries that have a specific Uniprot id
 - Example: 'Q9UVX4' (Currently only the Okayama strain has such)
- **Functional description:**
 - Searches for entries with functional descriptions that contain the specified text. In this case a non-exact match is performed.
 - Example: 'Aryl-alcohol oxidase'
- **Fungi model prote data:**
 - Search for entries that have an orthologous protein in a particular fungal model strain.
 - Example: 'NCU00414', 'ado1'
- **Fungi model strain name long:**
 - Search for entries that have an orthologous protein in a particular fungal model strain
 - Example: 'Schizosaccharomyces pombe'
- **Mushroom models target:**
 - Search for entries that have an orthologous protein in a particular mushroom forming fungal model strain. The query should look like this: short species name (based on JGI scheme)_protein id.
 - Example: 'Agrae_CAA7260863'
- **Mushroom models long name:**
 - Search for entries that have an orthologous protein in a particular mushroom forming fungal model strain.
 - Example: 'Flammulina velutipes'
- **Conserve domains signature accession:**
 - Search for entries that have a specific Pfam or CDD accession number
 - Exampe: 'PF03732', 'cd00303'
- **InterPro accession:**
 - Search for entries that have a specific InterPro accession
 - Example: 'IPR021109'
- **GO id:**
 - Search for entries that belong to a specified GO group
 - Example: 'GO:0003676'
- **KEGG Orthology:**
 - Search for entries that belong to a specified KEGG category
 - Example: 'K10901'
- **Cazy annotations:**
 - Search for entries that belong to a specified Cazy family
 - Example: 'GH18'
- **Transcription factor group:**

- Search for entries that belong to a specified transcriptional factor family
- Example: 'zinc finger, c2h2-type'

Functional annotation groups

Transcription factors

It contains manually curated *Coprinopsis cinerea* TFs grouped into transcription factor families.

Kinases

It contains the kinome of *Coprinopsis cinerea* determined by the kinase.com kinome project. Source: <http://kinase.com/web/current/fungi/coprinopsis/>

CAZYmes

It contains the manually curated CAZYmes of the *Coprinopsis cinerea* strains, which are organized into Cazy families.

Expression group

It provides an interface for visualizing the expression of multiple genes for the purpose of better comparability. The comma-separated 'Systematic names' of the entries can be used as a query to generate a figure for the available experiment describing the change in mean CPM per sample. Pre-generated gene sets are available in the 'Gene Sets' section of the Download tab.

JBrowse

It provides an interface for visualizing the composition of the genome of the selected *Coprinopsis cinerea* strains with various annotations, feature layers such as gene model layer, long read alignment layer, polyA tail containing Quantseq and RNA-Seq read coverage. If you would like a larger view of JBrowse, click on the "Share" link at the top right of the JBrowse window and then click on the "Preview" link to open the view in a new tab.

Download

On this page, most of the data used by the website is available for download in a clearly arranged form.

About

A brief description of how to cite and use CopciDB.

Protein/Gene table

Summary:

General data:

Systematic name: protein id (Source: gene annotation file). By selecting the name, you can get to the related FungiDB page (currently only works with the *C. cinerea Okayama7* strain)

Standard name: well-recognised name of the gene within the research community (Example: glnA, Source: EggNOG v5.0, eggNOG-mapper v2)

Uniprot id: UniProt id of the gene if available. By selecting the entry, you can get to the related UniProt webpage (currently only works with the *C. cinerea Okayama7* strain). (Example: A8NS14, Source: UniProt)

Strain: This entry leads to the indicated strain's orthologous genes, gene page.

Synonyms: Alternative protein id occurring in different annotation/databases (for *C. cinerea AmutBmut* the JGI protein id)

Functional description: Short functional description of the protein (Example: Zinc finger, C2H2 type, Source: EggNOG v5.0, eggNOG-mapper v2)

Location: Genomic coordinate of the gene. By selecting this entry, you can display the gene within the JBrowse.

Strand, Gene length, Transcript length, CDS length, Protein length: Brief summary of the annotation entries of various aspects.

Reciprocal best hits in model fungi

The orthologue of the gene in the well recognized model fungal strains identified by the reciprocal best hit method.

Saccharomyces cerevisiae

Schizosaccharomyces pombe

Aspergillus nidulans

Neurospora crassa

Penicillium marneffeii

Sordaria macrospora

Ustilago maydis

Cryptococcus neoformans

Blastomyces dermatitidis

Aspergillus niger

Histoplasma mississippiense
Candida albicans
Fusarium graminearum
Aspergillus fumigatus
Fomitopsis palustris
Magnaporthe (Pyricularia) oryzae

Orthologs in mushroom models

The orthologue of the gene in the well recognized mushroom forming fungal strains identified by the reciprocal best hit method.

Agaricus bisporus var bisporus (H97)
Agaricus bisporus var. burnettii JB137-S8
Agrocybe aegerita
Ganoderma sp. 10597 SS1 v1.0
Lentinula edodes B17
Lentinula edodes W1-26 v1.0
Pleurotus eryngii ATCC 90797
Pleurotus ostreatus PC15
Flammulina velutipes
Pleurotus ostreatus PC9
Auricularia subglabra
Grifola frondosa
Lentinula edodes NBRC 111202
Schizophyllum commune H4-8
Hypsizygus marmoreus strain 51987-8

Protein:

General data:

Systematic name: as in the Summary table General data section.

Protein id: The protein id of the used isoforme (in all cases the T0 isoforme was used)

Description: as in the Summary table General data section.

Annotation summary:

Visualization of the identifiable conserved domains (pfam, CDD), signal peptide and cleavage site and transmembrane domains.

Conserved domains:

Listing of general information on the identified Pfam, CDD domains. By selecting the entry, you can get to the related Pfam, CDD page. (Source: InterProScan)

SignalP:

List general information about the identified signal peptide and cleavage site. (Source: SignalP 5)

Transmembrane domains:

List general information about the identified transmembrane domains (Source: TMHMM2)

Functional annotation

InterPro:

List of Interpro accessions related to the protein and their description. Selecting the accession will take you to the InterPro page of the corresponding accession. (Source: InterProScan 93)

GO:

List of Gene Ontology (GO) accessions related to the protein and their description with ontology information. Selecting the accession will take you to the Gene Ontology page of the corresponding accession. (Source: InterProScan 93)

KEGG:

List of KEGG orthology entries related to protein. Selecting the accession will take you to the KEGG page of the corresponding entry. (Source: EggNOG v5.0, eggNOG-mapper v2)

EggNOG:

List of COG categories related to the protein and their description. (Source: EggNOG v5.0, eggNOG-mapper v2)

CAZy:

Class, Family and Subfamily level classification of the Cazyme proteins. (Source: manual curation)

Transcription factor:

Specify the group of the transcription factor. (Source: manual curation)

Expression:

It contains a wide selection of expression studies. If you open the study sheets by selecting the blue down arrow located at the beginning of the study description header, you can view the gene expression on CPM in a bar chart where the whiskers represent the standard

deviation of the samples. You can check the exact value by selecting the 'DATA TABLE' menu below the bar chart.

Conservation:

It shows the conservation of the selected protein entry across fungi.

Sequences:

It lists the Protein, Coding DNA, Transcript and Gene sequence of the selected entry.