



## FungiDB Workshop

32<sup>nd</sup> Fungal Genetics Conference in Asilomar, USA  
March 14, 2024

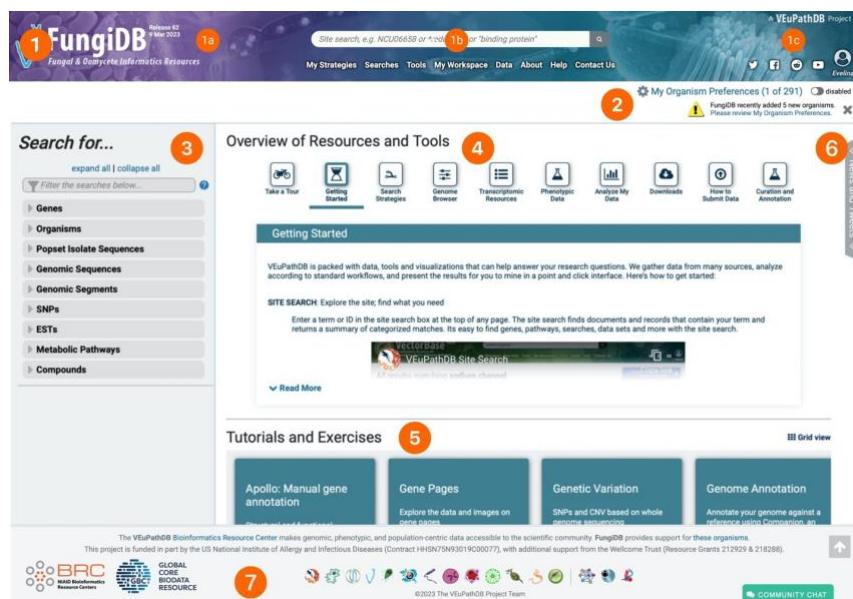
## **Table of contents**

Anatomy of the landing page	3
Site search	4
Creating advanced search strategies	7
Enrichment analysis	16
Exploring evidence on gene record pages	19
Exploring evidence in JBrowse	28

## Anatomy of the FungiDB main page

FungiDB is a component of the Eukaryotic Pathogen, Vector and Host Informatics Resource (VEuPathDB.org) that enables browsing, querying, and mining of genomic-scale datasets across diverse groups of organisms including hosts (HostDB.org), invertebrate vectors of human disease, pathogenic and non-pathogenic species, and environmental and epidemiological studies (ClinEpiDB.org).

FungiDB (<https://fungidb.org/>) is a free online informatics resource for exploration of fungal and oomycete omics data. It provides an easy-to-use, interactive interface to explore genomes, annotation, functional data (transcriptomics or proteomics), metabolic pathways, and results from numerous genome-wide analyses (i.e. AlphaFold, InterPro scan, signal peptide and transmembrane domain predictions, orthology, etc.). FungiDB also contains an expanding number fungal and oomycete genomes, including but not limited to plant, animal, and human pathogens. Additionally, FungiDB offers capabilities for private data analysis via the integrated VEuPathDB Galaxy – My Workspace platform. The database also permits genome improvements by capturing expert knowledge through the User Comments system and the Apollo genome annotation editor for structural and functional gene curation.



**1. Banner:** It includes 1a) release cycle information on the left, 1b) site search and menus centrally, providing access to nearly all searches, tools, and records in FungiDB, and 1c) account management and social media links on the right.

**2. My Organism Preferences Filter:** It is located directly below the banner and allows users to tailor database searches to a specified list of organisms.

**3. Left the “Search for...” panel:** Offers easy access to a categorized list of searches mirroring the options available in the "Searches" menu within the banner for consistent accessibility during site navigation.

**4. Overview of Resources & Tools Section:** Centrally located, it provides short tutorials on how to use FungiDB.

**5. Tutorials and Exercises:** Positioned immediately under the Overview of Resource & Tools section. It offers comprehensive educational resources featuring in-depth overviews and exercises.

**6. News & Tweets Section:** Positioned on the right, this expandable section keeps users informed about new datasets, tools, features, and X (formerly Twitter) updates.

**7. The footer:** It contains links to other VEuPathDB databases and relevant resources. Additionally, there is a link to the community chat app where users can ask questions about how to use FungiDB resources.

# 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.
- Use site search and other types of searches to create a multi-step query across different types of records

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

## Site search: text, term or gene id.

- 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.



- How many results with the word kinase did you get? Are all these records genes?
- Explore the filter panel on the left side of the page. Filter the results to view gene results only (hint: click on the word **Genes** in the “Filter results” section):

A screenshot of a search results page titled "All results matching kinase". The page shows 1 - 20 of 394,386 results. On the left, there is a "Filter results" sidebar with a "Genes" button highlighted with a red arrow. The main content area displays search results for "kinase", including a dataset analysis and individual gene entries. An "Export as a Search Strategy" button is visible at the top right.

Notice that clicking on the “Genes” category reveals additional filtering options (on the left) and activates the “Export as a Search Strategy” button on the top right, which is now shown in dark blue color. This is because the search strategy can be deployed on a single category only (e.g. Genes or Data sets, but not both).

- Select and apply the “Product descriptions (all)” filter.

Note: The applied filter can be easily cleared by clicking on “Clear filter” option as shown in the screenshot below.

- In the “Filter organisms” section, select to filter gene results by ***Malassezia restricta* KCTC 27527**. How many genes contain “kinase” in the product description field in this organism?
- Export the results to a search strategy.

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

Gene ID	Transcript ID	Organism	Genomic Location (Gene)	Product Description
MRET_0047	MRET_0047-146_1	Malassezia restricta KCTC 27527	CP030251:95,680..97,545(−)	triose/dihydroxyacetone kinase/FAD-AMP lyase (cyclizing)
MRET_0094	MRET_0094-146_1	Malassezia restricta KCTC 27527	CP030251:170,464..171,989(+)	adenosine kinase
MRET_0098	MRET_0098-146_1	Malassezia restricta KCTC 27527	CP030251:179,095..181,227(−)	aarF domain kinase
MRET_0099	MRET_0099-146_1	Malassezia restricta KCTC 27527	CP030251:181,386..181,844(+)	nucleoside-diphosphate kinase
MRET_0136	MRET_0136-146_1	Malassezia restricta KCTC 27527	CP030251:231,306..233,297(+)	pseudouridylate synthase/pseudouridine kinase
MRET_0167	MRET_0167-146_1	Malassezia restricta KCTC 27527	CP030251:270,552..272,270(−)	meiosis induction protein kinase IME2/SME1
MRET_0178	MRET_0178-146_1	Malassezia restricta KCTC 27527	CP030251:288,959..289,339(+)	tyrosine-protein kinase srms
MRET_0205	MRET_0205-146_1	Malassezia restricta KCTC 27527	CP030251:331,297..333,045(+)	type II pantothenate kinase

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

When the wild card is combined with a word (**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 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**

Genome	1,885,291
Genes	162,441
Genomic sequences	
Organism	
Organisms	186
Transcriptomics	
ESTs	1,709,817
Population biology	
Population isolate sequences	1,077,320
Metabolism	
Metabolic pathways	3,045
Compounds	61,998
Data access	
Data sets	381
Searches	435
Instructional	
Tutorials	
Workshop exercises	15
About	1
News	2
General info pages	16

**Filter fields**  
Select a result filter above

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

**Export as a Search Strategy**  
to download or mine your results

Compound - CHEBI:10000 Vismidine D  
Compound - CHEBI:10001 Visnagin  
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 nalidixic 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 Vacamidine  
Compound - CHEBI:10015 vobasine  
Definition: An indole alkaloid that is vobasan 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 volbutusine  
Compound - CHEBI:10017 volenitol  
Definition: A heptitol that is heptane-1,2,3,4,5,6,7-heptol that has R-configuration at positions 2, 3, and 6.  
Compound - CHEBI:10018 volkenine  
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 Vornicicine  
Compound - CHEBI:10022 Vomitoxin  
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 *Candida* spp. It is an inhibitor of cytochrome P450 2C9 (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

- The site search also works with gene ids. Run a site search for the following gene id: Afu2g13260

The gene id search will return the gene record card for [Afu2g13260](#).

Genes matching Afu2g13260

1 - 1 of 1

**Filter results**

Genome	1
Genes	1

**Filter Gene fields**  
select all | clear all  
External links  
Gene ID  
Names, IDs, and aliases  
User comments

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

**Export as a Search Strategy**  
to download or mine your results

**Gene - Afu2g13260** Developmental regulator medA, putative  
Gene name or symbol: medA  
Gene type: protein coding gene  
Organism: *Aspergillus fumigatus* Af293  
Fields matched: External links; Gene ID; Names, IDs, and aliases; User comments

**Gene - Afu2g13260** Developmental regulator medA, putative  
Gene name or symbol: medA  
Gene type: protein coding gene  
Organism: *Aspergillus fumigatus* Af293  
Fields matched: External links; Gene ID; Names, IDs, and aliases; User comments

1 - 1 of 1

Clicking on the gene link in blue within the card will bring up the gene record page for this gene.

Clicking on the “Export as a Search Strategy” button will create a search strategy with a single gene ID. This may be useful if you are interested in cross-referencing different types of data for one gene.

Search strategy links:

**kinase** - <https://fungidb.org/fungidb/app/workspace/strategies/import/9c47e36cfaf7790f6>

**kinase\*** - <https://fungidb.org/fungidb/app/workspace/strategies/import/eee9e7d2dfb3e7c1>

**Afu2g13260** -

<https://fungidb.org/fungidb/app/workspace/strategies/import/6fc6b7e52a15b76b>

## Advanced Search Strategies

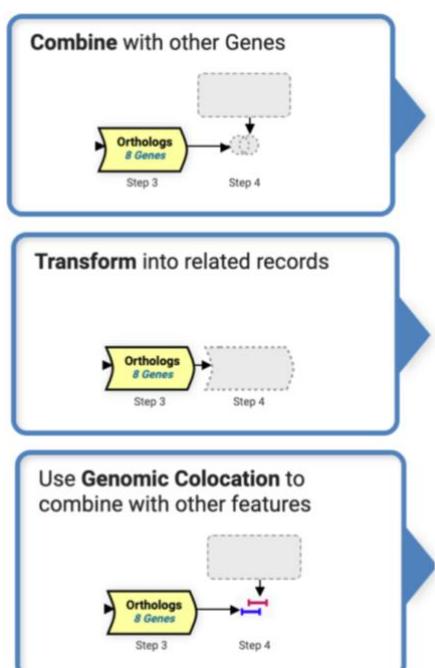
### Learning objectives

- Deploy search for different types of data and create advanced search strategy in FungiDB.

The strategy system offers a unique system of structured searches that can be combined to create multi-step *in-silico* experiments. As seen above, searches can be deployed from the site search, or the ‘Search For...’ menu on the home page, and from the ‘Searches’ dropdown menu in the header of every page.

Searches listed under the “Genes” category will return a list of gene IDs, while searches listed under the ‘SNPs’ or ‘Metabolic Pathways’ will return record relevant to SNPs data (e.g., sequences) and metabolic pathways, respectively.

When creating multi-step search strategy, the search strategy steps can be combined via three methods:



**Combine with other Genes:** compares results that are gene lists.

**Transform into related records:** transforms results into orthologs (e.g. *Aspergillus* > *Candida*), metabolic pathways or compounds.

**Use Genomic Colocation** to combine with other features: cross-references different types of data – e.g., gene lists with metabolic pathways.

Within the search strategy, each step is connected via the system of Boolean operators that can intersect, unite, or subtract similar records (e.g., gene lists) and cross-references different types of data via the genomic colocation option. Additionally, steps can be masked off from the strategy with the help of “ignore step” Boolean operators that allow quick strategy modification without the need of step deletion.

Revise as a boolean operation

1 INTERSECT 2     1 UNION 2     1 MINUS 2     2 MINUS 1

Revise as a span operation

1 RELATIVE TO 2, using genomic colocation

Ignore one of the inputs

IGNORE 2     IGNORE 1

Revise

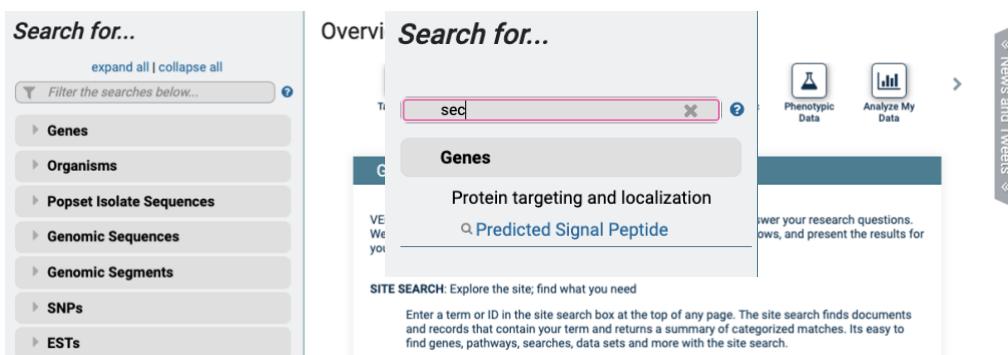
## Creating advanced search strategies in FungiDB.

**In silico experiment:** Identify *Aspergillus fumigatus* Af293 genes that have a signal peptide and non-synonymous mutations identified by whole genome sequencing (WGS) of clinical isolates. Next, determine which genes may be putative vaccine targets (are known epitopes that do not have orthologs in humans).

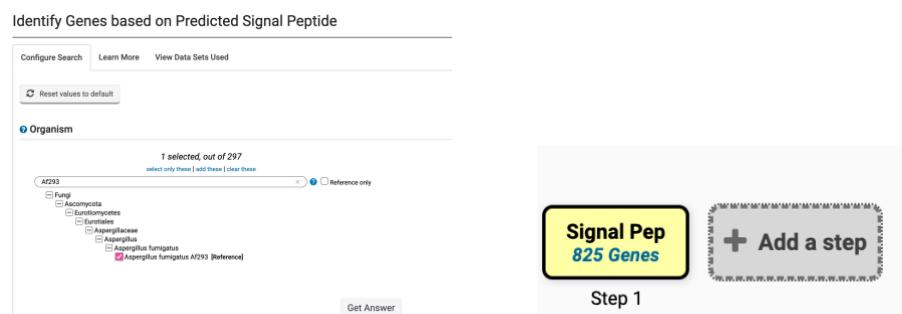
- **Deploy the “Predicted Signal Peptide” search**

1. Select the search from the “Search for...” panel (shown below) or the “Searches” menu from the top menu.

Hint: use the filter box to quickly bring up relevant search.

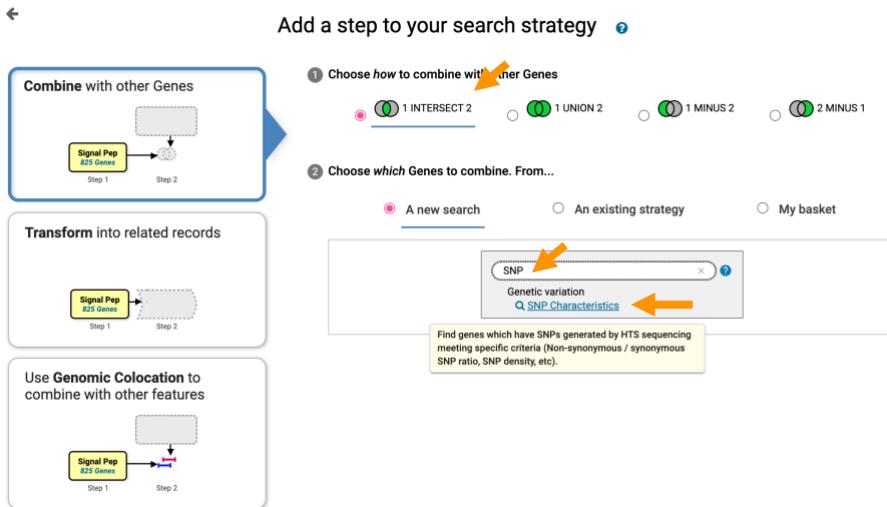


2. Set the organism to “*Aspergillus fumigatus* Af293 [Reference]” and click on the “Get Answer” button.



- **Find genes with non-synonymous SNPs using WGS data of clinical isolates aligned to the reference genome *Aspergillus fumigatus* Af293.**

1. Click on the “Add a step” button with the search strategy.
2. Within the “Combine with other Genes”, use the “1 INTERSECT 2” Boolean operator and filter available searches to identify and deploy the “SNP Characteristics” search.
3. Click on the “SNP Characteristics” link in blue to deploy the search.



## Next, specify parameters of the SNP characteristics search:

- Select Aspergillus fumigatus Af293 as the genome of interest.

- Select datasets.

Within the Set of Samples section, click on the Data Set category and select two datasets:

- SNP call on WGS of Aspergillus fumigatus drug-resistant clinical isolates
- SNP calls on A. fumigatus strains isolated from patients with PA and CNPA

### Set of Samples

1,118 Set of Samples Total		39 of 1,118 Set of Samples selected		Data Set
expand all	collapse all			
<input type="checkbox"/> culture medium				
<input type="checkbox"/> Fungal organism				
<input checked="" type="checkbox"/> Data Set				
Sample				
Sample collection				
Geographic location				
Sample source				
Organism under investigation				
<input type="checkbox"/> Keep checked values at top		1,118 (100%) of 1,118 Set of Samples have data for this variable		
	Data Set	Remaining Set of Samples	Set of Samples	Distribution
		1,118 (100%)	1,118 (100%)	%
	Aligned SNPs - Aspergillus fumigatus Af1163 strain	1 (< 1%)	1 (< 1%)	(100%)
	Aspergillus fumigatus Af293 Genome Sequence and Annotation	1 (< 1%)	1 (< 1%)	(100%)
	Aspergillus fumigatus LH-EVOL strains	4 (< 1%)	4 (< 1%)	(100%)
	Genomic Context of Azole-Resistance Mutations in Aspergillus fumigatus	24 (2%)	24 (2%)	(100%)
	<input checked="" type="checkbox"/> SNP call on WGS of Aspergillus fumigatus drug-resistant clinical isolates	22 (2%)	22 (2%)	(100%)
	<input checked="" type="checkbox"/> SNP calls on A. fumigatus strains isolated from patients with PA and CNPA	17 (2%)	17 (2%)	(100%)

3. Indicate specific SNP characteristics.

Scroll down the parameter selection window and choose to deploy the search using the following:

**SNP Class** = Non-Synonymous

**Number of SNPs of above class >=1**

**SNP Class**

Non-Synonymous

**Number of SNPs of above class >=**

1

```

graph LR
    A[Signal Pep  
825 Genes] --> B[SNPs  
7,726 Genes]
    B --> C[681 Genes]
    D[Add a step]
  
```

- **Identify known *Aspergillus* epitopes.**

Epitopes are recognized by immune system and can be used for vaccine development.

1. Click on the “Add a step” button.
2. Within the “Combine with other Genes”, use the “2 INTERSECT 3” Boolean operator and filter available searches for “epitope” to identify and deploy the “Epitope Presence (IEDB)” search.

**Combine with other Genes**

Step 1: SNPs 7,726 Genes

Step 2: 681 Genes

Step 3: (dashed box)

**① Choose how to combine with other Genes**

2 INTERSECT 3    2 UNION 3    2 MINUS 3    3 MINUS 2

**② Choose which Genes to combine. From...**

A new search    An existing strategy    My basket

epitope

Immunology

Epitope Presence (IEDB)

3. Set organism to *Aspergillus fumigatus* Af293.
4. Set Confidence to “high” and “Medium” and click on the Run Step button.

## Organism

1 selected, out of 232  
[select only these](#) | [add these](#) | [clear these](#)

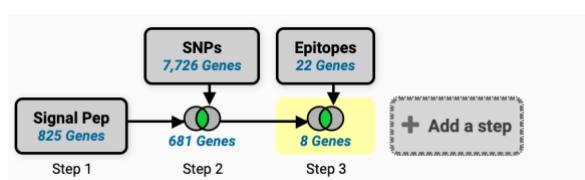
Af293   Reference only

- Fungi
- Ascomycota
- Eurotiomycetes
- Eurotiales
- Aspergillaceae
- Aspergillus
- Aspergillus fumigatus Af293 [Reference]

## Confidence

- High  
 Medium  
 Low
- [select all](#) | [clear all](#)

Run Step



- Identify Aspergillus epitopes that do not have orthologs in humans.

- Click on the “Add a step” button.
- Select to deploy the “Orthology Phylogenetics Profile” search.

This search uses OrthoMCL algorithm to identify fungal orthologs across all species in VEuPathDB. Run this search if you want to explore species outside those supported in FunigDB.org.

**Combine with other Genes**

Epitopes 22 Genes → 8 Genes

**Transform into related records**

Epitopes 22 Genes → 8 Genes

**Use Genomic Colocation to combine with other features**

Epitopes 22 Genes → 8 Genes

**Choose how to combine with other Genes**

① Choose how to combine with other Genes

3 INTERSECT 4     3 UNION 4     3 MINUS 4     4 MINUS 3

**Choose which Genes to combine. From...**

② Choose which Genes to combine. From...

A new search     An existing strategy     My basket

orthology

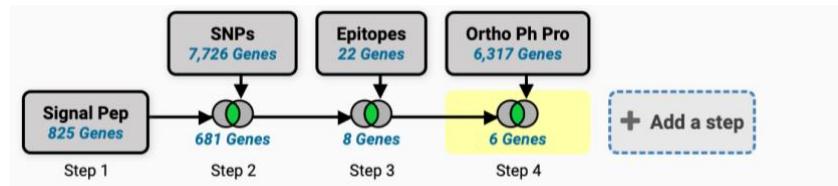
Orthology and synteny  
 Orthology Phylogenetic Profile  
 Paralog Count

3. Set parameters for the “Search for Gene by Orthology Phylogenetic Profile”:
- Find genes in these organisms:** Aspergillus fumigatus Af293
- Select orthology profile:** Homo sapiens REF (hsap) must not be in group

The screenshot shows two main sections of the search interface:

- Find genes in these organisms:** A dropdown menu titled "Af29" lists organisms under "Fungi". The "Aspergillus fumigatus" option is selected with a green checkmark, while "Aspergillus fumigatus Af293" is selected with a red checkmark. A checkbox for "Reference only" is unchecked.
- Select orthology profile:** A dropdown menu titled "sapiens" lists various taxonomic groups. The "Homo sapiens REF (hsap)" option is selected with a red cross, indicating it must not be in the orthology profile.

How many genes did you get?



Search strategy links:

FGC2024 advanced search strategy 1:

<https://fungidb.org/fungidb/app/workspace/strategies/import/741f30239b2cc68d>

## Advanced search strategies (optional)

The next block of exercises will be carried out in [HostDB.org](#)

SC5314 induces rapid transcriptional response in the host, while 101 has slower kinetic. The commensal 101 strain has also reduced filamentation when compared to SC5314.

- **Identify genes up-regulated in mice infected with SC5314 at 1d.**

1. Navigate to the RNA-Seq Evidence search and filter RNA-Seq datasets for “Kirch” to examine the dataset by Kirchner et al. 2019.
2. Click on the “DE” button.
3. Choose to examine the sense strand.
4. Select reference sample: naïve.
5. Select comparator sample: SC5314\_infected\_1d.
6. Look for up-regulated genes.
7. Select magnitude of upregulation: 4 fold.

The screenshot shows the HostDB.org search interface with the following steps highlighted:

1. In the search bar, type "ma". Below the search bar, under "RNA-Seq Evidence", click the "DE" (Differential Expression) button.
2. In the legend, click the "DE" button. In the "Organism" dropdown, select "Mus musculus C57BL/6J". In the "Data Set" dropdown, select "Mouse transcriptomes during oropharyngeal candidiasis infection (Kirchner, et al. 2019)".
3. Under "Experiment", select "naive" as the reference sample.
4. Under "Comparator Sample", select "SC5314\_infected\_1d".
5. Under "Direction", select "up-regulated".
6. Under "fold difference >=", enter "4".
7. Under "adjusted P value less than or equal to", enter "0.1".

At the bottom right, there is a yellow box labeled "Calb\_Galleria\_mouse (de) 857 Genes" with an "Edit" button. A blue "Add a step" button is also present. A red arrow points from the search parameters to the results table.

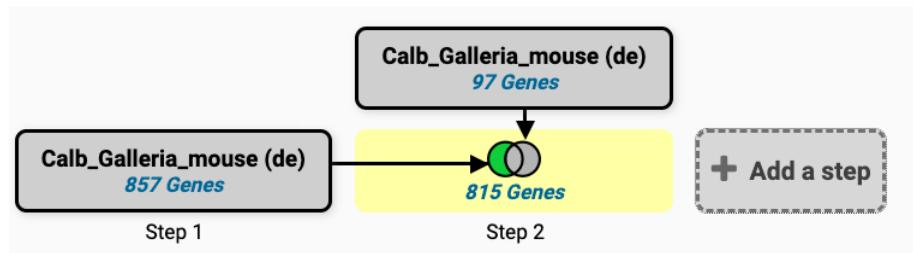
- Identify genes up-regulated in response to 101 persistent strain at 1d of infection.
  1. Click on the “Add Step” button.
  2. Navigate to the RNA-Seq Evidence search, select “1 minus 2” Boolean operator, filter for “Kirch” to quickly identify the dataset and click on the “DE” button.
  3. Choose to examine the sense strand.
  4. Select reference sample: naïve.
  5. Select comparator sample: 101\_infected\_1d.
  6. Look for up-regulated genes.
  7. Select magnitude of upregulation: 4 fold.

The screenshot shows the Galleria software interface with numbered steps indicating the process:

- Step 1:** A yellow box labeled "Cells\_Galleria\_mouse (de)" is selected. An orange circle labeled "1" is over the "Add a step" button.
- Step 2:** The "Combine with other Genes" panel is open. An orange circle labeled "2" is over the "Transform into related records" section. The "Choose how to combine with other Genes" dropdown is set to "1 MINUS 2". The "Choose which Genes to combine, From..." dropdown has "A new search" selected. The "Gene models" dropdown includes "Gene Model Characteristics", "Transcriptome", and "RNA-Seq Evidence".
- Step 3:** The search results are displayed. An orange circle labeled "3" is over the results table. It shows two entries:
  - Mouse transcriptomes during oropharyngeal candidiasis infection in mouse - Sense
  - Mouse transcriptomes during oropharyngeal candidiasis infection in mouse - Antisense
- Step 4:** The "Reference Sample" section is shown. An orange circle labeled "4" is over the "naive" option in the list of samples.
- Step 5:** The "Comparator Sample" section is shown. An orange circle labeled "5" is over the "101\_infected\_1d" option in the list of samples.
- Step 6:** The "Direction" section is shown. An orange circle labeled "6" is over the "up-regulated" dropdown menu.
- Step 7:** The "fold difference >=" section is shown. An orange circle labeled "7" is over the "4" input field.

At the bottom right, there is a "Get Answer" button with an orange arrow pointing to it.

- **Modify the Boolean operator to determine genes that are upregulated in response to SC5314 but not 101 strain.**

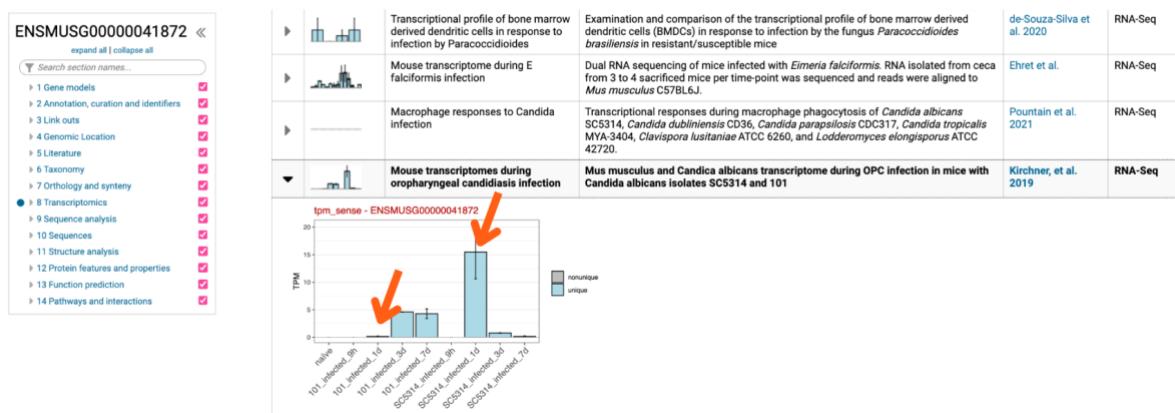


- **Examine the results in HostDB:**

1. Click on the [Gene ID](#) link for “interleukin 17F” and navigate to the transcriptomics expression section.

Gene ID	Transcript ID	Product Description	# Transcripts
ENSMUSG000000104379	ENSMUST00000084867	predicted gene, 37509 [Source:MIgI Symbol;Acc:MIgI:5610737]	1
ENSMUSG00000067780	ENSMUST00000088476	peptidase inhibitor 15 [Source:MIgI Symbol;Acc:MIgI:1934659]	1
ENSMUSG00000025929	ENSMUST00000027061	interleukin 17A [Source:MIgI Symbol;Acc:MIgI:107364]	1
ENSMUSG00000041872	ENSMUST00000039046	interleukin 17F [Source:MIgI Symbol;Acc:MIgI:2676631]	4
ENSMUSG00000041872	ENSMUST00000189301	interleukin 17F [Source:MIgI Symbol;Acc:MIgI:2676631]	4
ENSMUSG00000041872	ENSMUST00000190692	interleukin 17F [Source:MIgI Symbol;Acc:MIgI:2676631]	4
ENSMUSG00000041872	ENSMUST00000191111	interleukin 17F [Source:MIgI Symbol;Acc:MIgI:2676631]	4
ENSMUSG000000104358	ENSMUST00000192924	predicted gene, 37127 [Source:MIgI Symbol;Acc:MIgI:5610355]	1
ENSMUSG000000417180	ENSMUST00000056946	neuronalized E3 ubiquitin protein ligase 3 [Source:MIgI Symbol;Acc:MIgI:2429944]	2
ENSMUSG000000417180	ENSMUST00000188666	neuronalized E3 ubiquitin protein ligase 3 [Source:MIgI Symbol;Acc:MIgI:2429944]	2
ENSMUSG00000037447	ENSMUST00000037778	AT rich interactive domain 5A (MRFI-like) [Source:MIgI Symbol;Acc:MIgI:2443039]	15
ENSMUSG00000037447	ENSMUST00000115029	AT rich interactive domain 5A (MRFI-like) [Source:MIgI Symbol;Acc:MIgI:2443039]	15
ENSMUSG00000037447	ENSMUST00000115031	AT rich interactive domain 5A (MRFI-like) [Source:MIgI Symbol;Acc:MIgI:2443039]	15
ENSMUSG00000037447	ENSMUST00000115032	AT rich interactive domain 5A (MRFI-like) [Source:MIgI Symbol;Acc:MIgI:2443039]	15
ENSMUSG00000037447	ENSMUST00000116629	AT rich interactive domain 5A (MRFI-like) [Source:MIgI Symbol;Acc:MIgI:2443039]	15
ENSMUSG00000037447	ENSMUST00000124280	AT rich interactive domain 5A (MRFI-like) [Source:MIgI Symbol;Acc:MIgI:2443039]	15
ENSMUSG00000037447	ENSMUST00000126413	AT rich interactive domain 5A (MRFI-like) [Source:MIgI Symbol;Acc:MIgI:2443039]	15
ENSMUSG00000037447	ENSMUST00000137906	AT rich interactive domain 5A (MRFI-like) [Source:MIgI Symbol;Acc:MIgI:2443039]	15
ENSMUSG00000037447	ENSMUST00000140218	AT rich interactive domain 5A (MRFI-like) [Source:MIgI Symbol;Acc:MIgI:2443039]	15
ENSMUSG00000037447	ENSMUST00000141121	AT rich interactive domain 5A (MRFI-like) [Source:MIgI Symbol;Acc:MIgI:2443039]	15

In summary, we identified genes upregulated in response to SC5314 infection. Notice that the interleukin 17F response is much stronger at 1d in response SC5314 infection. This is consistent with mouse response to *C. albicans* strain 101 being delayed compared to strain SC5314. Now, you may want to go back and look at gene enrichment signatures in fungi to learn more about SC5314 and 101-driven responses.



Search strategy links:

FGC2024 advanced search strategy 2:

<https://hostdb.org/hostdb/app/workspace/strategies/import/ec36d02df763b3d5>

## Enrichment analysis

### Learning objectives

- Run enrichment analysis on RNA-Seq analysis results.

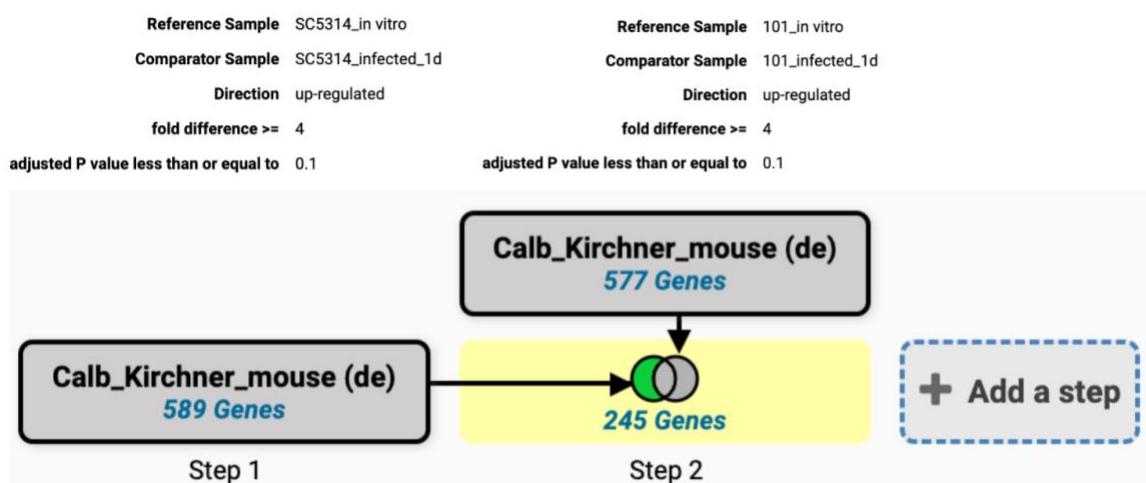
The enrichment analysis tools can be accessed under the blue Analyze Results tab and it includes Gene Ontology, Metabolic Pathway, and Word Enrichment tools. The three types of analysis apply Fisher's Exact test to evaluate ontology terms, over-represented pathways, and product description terms. Enrichment is carried out with Fisher's exact test to assess the proportion of GO terms in one set of genes versus all the genes in the query organism. The test produces P-values which are then corrected for multiple testing with the Benjamini-Hochberg false discovery rate or the Bonferroni test.

In the previous exercise, you examined host response to infection with Candida strains. In this section, we will learn how to perform enrichment analysis on the fungal component of the Kirchner et al dataset. SC5314 induces rapid transcriptional response in the host, while 101 has slower kinetic. The commensal 101 strain has also reduced filamentation when compared to SC5314.

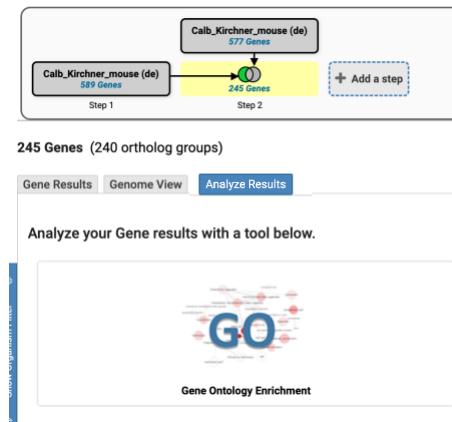
To begin, click on the search strategy link below:

<https://fungidb.org/fungidb/app/workspace/strategies/import/34f998f05745cbc3>

This strategy identifies genes up-regulated in SC5314 at 1d post infection (Step 1) and subtracts up-regulated genes in common with the persistent 101 strain (Step2).



- Perform GO Enrichment analysis (Molecular function)
  - Select Step 2 results (they will become highlighted in yellow).
  - Click on the Analyze Results button located about the gene results table.



GO enrichment analysis can be performed on the following ontology groups: molecular function, cellular component, and biological processes. Also, other parameters allow users to limit their analysis on either “Curated” or “Computed” annotations, or both. Those with a GO evidence code inferred from electronic annotation (IEA) are denoted “Computed”, while all others have some degree of curation. The default P-value is set to 0.05 but can be adjusted manually.

3. Select the “Molecular function” option.
4. Run the enrichment analysis on both computed and curated evidence (GO terms) and leave other parameters at default.

Note: When the GO Slim option is chosen, both the genes of interest and the background are limited to GO terms that are part of the generic GO Slim subset

**Organism** ? Candida albicans SC5314

**Ontology** ?  Biological Process  Cellular Component  Molecular Function

**Evidence** ?  Computed  Curated select all | clear all

**Limit to GO Slim terms** ?  No  Yes

**P-Value cutoff** ? 0.05 (0 - 1)

**Submit**

5. Click on the “Submit” button.

- Examine your results. Looking at the enriched terms, do they make sense in terms of what you know about the Kirchner et al. 2019 dataset?

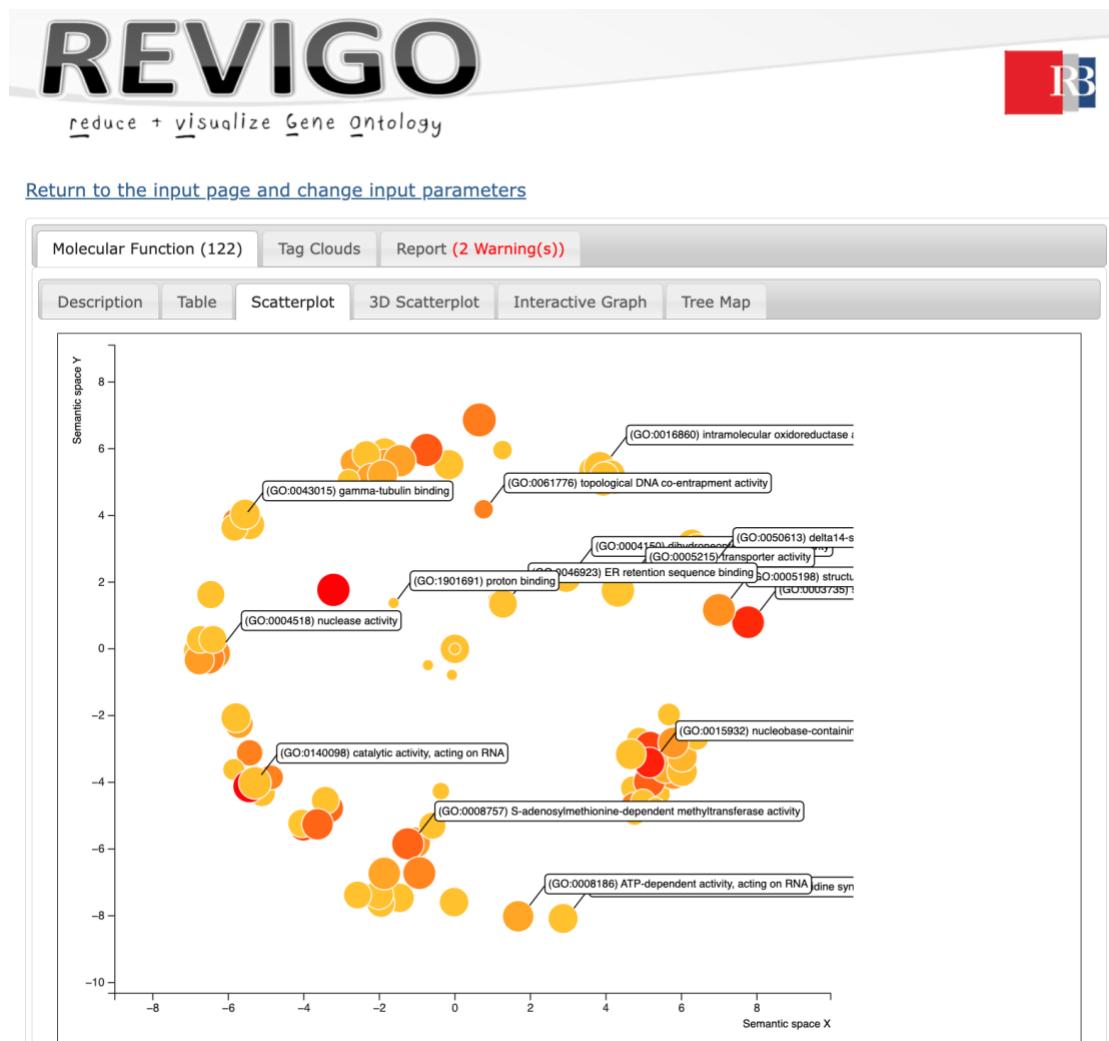
Analysis Results:										
GO ID	GO Term	Genes in the bkgd with this term	Genes in your result with this term	Percent of bkgd genes in your result	Fold enrichment	Odds ratio	P-value	Benjamini	Bonferroni	
GO:0140098	catalytic activity, acting on RNA	262	28	10.7	2.70	3.15	1.19e-6	6.18e-4	6.18e-4	
GO:0140640	catalytic activity, acting on a nucleic acid	403	36	8.9	2.26	2.62	2.56e-6	6.68e-4	1.34e-3	
GO:0015932	nucleobase-containing compound transmembrane transporter activity	40	9	22.5	5.68	7.28	1.92e-5	2.26e-3	1.00e-2	
GO:0005347	ATP transmembrane transporter activity	23	7	30.4	7.69	10.90	1.90e-5	2.26e-3	1.04e-2	
GO:0015215	nucleotide transmembrane transporter activity	32	8	25.0	6.31	8.33	2.48e-5	2.26e-3	1.29e-2	
GO:0003735	structural constituent of ribosome	166	19	11.4	2.89	3.32	2.60e-5	2.26e-3	1.36e-2	

The results table includes several additional statistical measurements:

- **Fold enrichment** - The ratio of the proportion of genes in the list of interest with a specific GO term over the proportion of genes in the background with that term.
- **Odds ratio** - Determines if the odds of the GO term appearing in the list of interest are the same as that for the background list.
- **P-value** - The probability of a result occurring by chance under a null hypothesis.
- **Benjamini-Hochberg false discovery rate** - A method for controlling false discovery rates for type 1 errors.
- **Bonferroni adjusted P-values** - A method for correcting significance based on multiple comparisons.

Note: you can sort genes in your results using the sort options within a column.

- **Visualize enrichment in REVIGO.**
1. Click on the “Open in Revigo” button and follow prompts to complete this step.



The table tab provides a detailed overview of the GO terms, P-values and also parent GO terms used to describe a group of related GO terms (<http://geneontology.org/docs/ontology-relations/>).

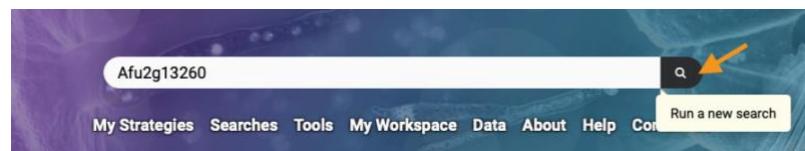
More about REVIGO:

<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0021800>

## Exploring the gene record page

### Learning objectives

- Become familiar with gene page structure and content
  - Navigate to and from the gene pages
- 
- Use the site search to navigate to the gene record page of [Afu2g13260](#), which is a gene known to be important for the virulence of *Aspergillus fumigatus*.



Genes matching **Afu2g13260**

1 - 1 of 1

**Filter results**

Genome Genes **1**

**Filter Gene fields**

select all | clear all  
External links **1**  
Gene ID **1**

**Filter organisms**

select all | clear all | expand all | collapse all  
Type a taxonomic name  **1**

Fungi **1**  
Ascomycota **1**

**Gene - Afu2g13260** Putative regulator of adherence, host cell interactions and virulence

Gene name or symbol: medA  
Organism: Aspergillus fumigatus Af293

Fields matched: External links; Gene ID

Gene - Afu2g13260 Putative regulator of adherence, host cell interactions and virulence

Gene name or symbol: medA  
Organism: Aspergillus fumigatus Af293

Fields matched: External links; Gene ID

3 Export as a Search Strategy →  
to download or mine your results

- 1 The panel on the left provides a summary of all record types that match Afu2g13260.
- 2 Click on the gene link to navigate to the gene record page for Afu2g13260.
- 3 Clicking on this button will transform your search into a search strategy.  
Note: If the button is shaded/inactive, limit your search to a single data type using the Filter results panel on the left.

## Gene page components

The top section of the gene record page provides a snapshot of the information available for this gene and offers several shortcuts:

The screenshot shows the gene record page for Afu2g13260. At the top, there are three orange circles numbered 1, 2, and 3. Below them are three small buttons: "Add to basket" (with a shopping cart icon), "Add to favorites" (with a star icon), and "Download Gene" (with a download icon). The gene name is Afu2g13260, followed by its description: "Developmental regulator medA, putative".

Component 1: "Add to basket" button.

Component 2: "Add to favorites" button.

Component 3: "Download Gene" button.

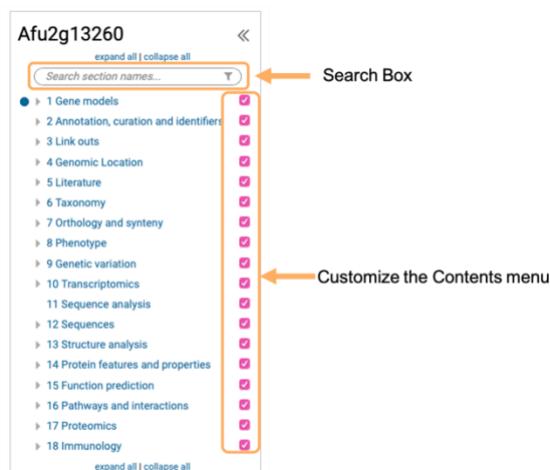
Component 4: A section titled "Model Organism Database(s)" which lists "CGD: C3\_01180C\_A".

Component 5: A "Shortcuts" panel containing six items: Synteny, Alignments, SNPs, Transcriptomics, Protein Features, and Proteomics. Below the panel is a link: "Also see Afu2g13260 in the [Genome Browser](#) or [Protein Browser](#)".

- 1 Add to basket: Save to basket if you want to download gene-specific information for selected genes.
- 2 Add to favorites: Saves genes in the private My favorites section, where you can add notes or keep track of your project.
- 3 Download Gene: Redirects to a download options page where gene records can be exported at text, FASTA, and GFF3 formats.
- 4 Submit a comment or annotate gene in Apollo, a web-based structural and functional gene annotation platform.
- 5 Shortcuts panel provides quick access to the selected section within the gene record page.

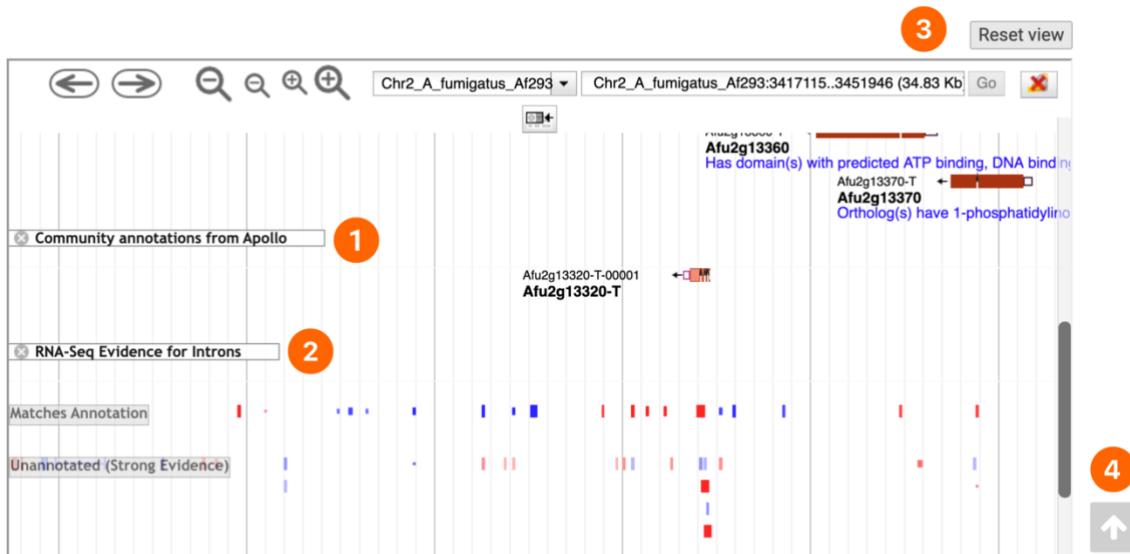
- **Explore the content of this gene record page.** Below are the navigational highlights.

The **Contents** section contains a list of links to various sections in the gene record page and is searchable. Sub-sections of the Contents menu can be hidden by checking the box to the right.



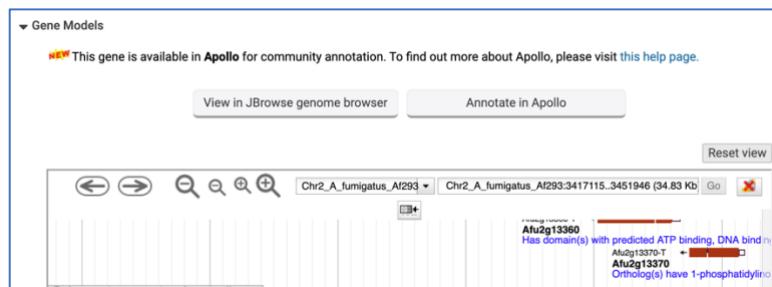
- Explore the gene model section.

The **Gene Models** section is the first section of the gene record page, and it contains information about the structure of the gene (e.g. exon count, transcript number, annotated UTRs, community gene annotation in Apollo, introns, etc.) displayed within the genome browse JBrowse wrapper.



- 1 Community annotation from Apollo provides the latest annotation updates (within 24-48hrs after the changes were submitted to the curation team in Apollo)
- 2 Introns that are matching transcript annotation for which there is an abundance of supporting data from aligned RNA-Seq reads are displayed in bold colours (Blue for a forward gene and red for a gene located on the reverse strand).
- 3 If you navigated away from the gene of interest while scrolling, click on the “Reset view” button to return to the default position within the JBrowse wrapper.
- 4 Click this button to navigate to the top of the gene record page.

The “View in JBrowse genome browser” and “Annotate in Apollo” buttons open in separate tabs. In JBrowse, you can activate additional tracks and build custom evidence views. In Apollo, you can modify and create new genes to improve the genome annotation.



- Explore the RNA-Seq Evidence for Introns track.

Click on the intron. The pop-up window contains a table showing all experiments and samples that provide evidence for this intron junction (generated by in-house automated pipelines):

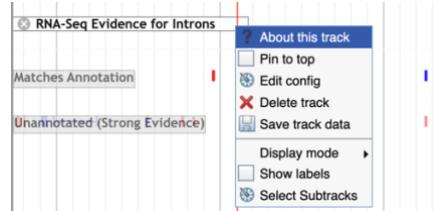
*Intron Spanning Reads (ISR)*: The total number of uniquely mapped reads (all samples) which map across the junction and are on the appropriate strand. GSNAP uses splice site consensus sequences to determine the strand of the mapped read.

*ISR per million (ISRPm)*: Intron Spanning Reads Per Million intron spanning reads and thus represents a normalized count of unique reads.

Chr2_A_fumigatus_Af293_3434523_3434611_0 details				
Intron Junction Details				
Intron Location	Chr2_A_fumigatus_Af293:3434523..3434611 (+ strand)			
Intron Spanning Reads (ISR)	7027			
ISR per million (ISRPm)	2140.85			
Gene assignment	Afu2g13295 - annotated intron			
% of Most Abundant Intron (MAI)	100			
Sample Details				
Experiment	Sample	Unique ISRPM	ISR/Cov	% MAI
Gene expression in WT, hrmA deletion, hrmA OE, hrmA_REV, EVOL under hypoxia and normoxia conditions	Delta hrmA hypoxia	64	33.07	5.61
	Delta hrmA normoxia	61	28.65	4.78
				61.61

*% of Most Abundant Intron (MAI)*: The percentage (ISRPm of this junction / ISRPm of maximum junction for this gene) of this junction over the maximum for this gene.

Note that the tracks within the JBrowse wrapper have a drop-down menu for further track customization:



- Explore other contents within the page.

The **Annotation, curation and identifiers** section offers alternate product descriptions, previous identifiers, and aliases, and is populated using data from internal curation, other fungal resources (*e.g.* AspGD, Ensembl, *etc.*), or user-submitted data (user comments).

2 Annotation, curation and identifiers		
Community annotations from Apollo	<a href="#">Data sets</a>	
Product Descriptions	<a href="#">Download</a>	<a href="#">Data sets</a>
Alternate Product Descriptions	<a href="#">Data sets</a>	
Gene Name or Symbol	medA	
Names, Previous Identifiers, and Aliases	<a href="#">Download</a>	<a href="#">Data sets</a>
<input type="text" value="Search this table..."/> <span style="float: right;">?</span>		
Name/ID/Alias	Type	Source
746128.CADAFAUBP00002828	alternate ID	STRING
EAL93620.1	alternate ID	protein_id
Q4X0J5	alternate ID	Uniprot/SPTREMBL
UPI000051EE09	alternate ID	UniParc
XM_750565.1	alternate ID	RefSeq_dna
XP_755658.1	alternate ID	RefSeq_peptide
medA	name	N/A
AFUA_2G13260	previous ID	N/A
CADAFAUG00004866	previous ID	N/A

The **Link outs** section offers redirection to other resources (*e.g.*, CGD, Ensembl, MycoCosm, *etc.*).

The **Orthology and synteny** section provides a table of Orthologs and Paralogs within FungiDB produced by OrthoMCL ([www.orthomcl.org](http://www.orthomcl.org)).

The table has a search box for creating a custom display of orthologs and also deploy the ClustalOmega analysis. The output of this tool can be used to build phylogenetic trees (*e.g.* iTOL).

**Orthology and Synteny**

Ortholog Group: CGD\_19863

Orthologs and Paralogs within FungiDB: 2 new sets

Ceolus

Showing 8 of 125 rows

Use search box to limit to a species of interest

Protein cluster	Gene	Organism	Product	In synteny	has comments
CGO_23707	Coccidioides immitis H5284		transcriptional regulator, Module	yes	no
CGMO_06073	Coccidioides immitis RS		transcriptional regulator, Module	yes	no
CGR2_09048	Coccidioides pseudotropicalis 348B		transcription factor	yes	no
CPCT08_086370	Coccidioides pseudotrop C755		hypothetical protein	yes	no
CPG05_080167	Coccidioides pseudotrop		hypothetical protein	yes	no
CPG05_080168	Coccidioides pseudotrop RS		transcription factor, Module	yes	no
CPG05_080169	Coccidioides pseudotrop RS		transcriptional regulator, Module	yes	no
PAAG_17199	Paracoccidioides lutzii PAF1		hypothetical protein	yes	no
PAIB_06303	Paracoccidioides brasiliensis		hypothetical protein	yes	no
PN05_10519	Paracoccidioides brasiliensis		hypothetical protein	yes	no

Run clustal Omega for selected genes: Check All

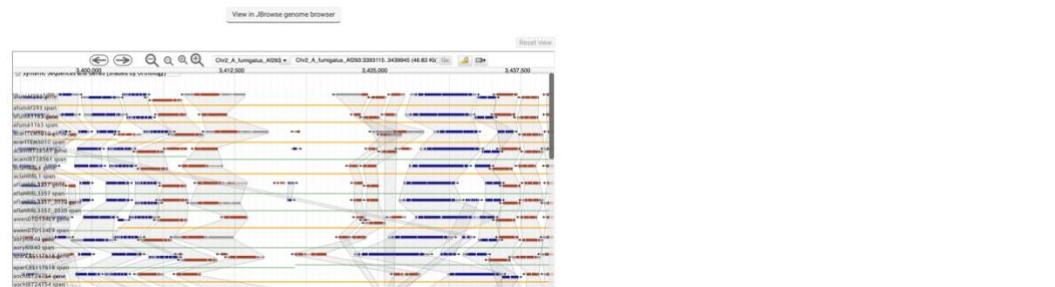
Run ClustalOmega to generate protein sequence alignments and a .dnd file for building phylogenetic trees

**Clustal Omega 1.2.3 Multiple Sequence Alignments**

**.dnd file**

```
{
    "PAGD_12199-t30_l-p1": 0.0101626,
    "PAGD_05304-t30_l-p1": 0.00135501,
    "PAGD_12215-t30_l-p1": 0.00135501
} ; 0.09880759 ; 0.271152
```

The **Orthology and Synteny** section also contains synteny graphs in JBrowse:



In the screenshot above, the synteny genes are highlighted in grey.

The **Phenotype** section offers curated information, including annotations from the Pathogen-Host Interactions database, COFUN project (selected transcription factors knockouts) and other sources.

Aspergillus fumigatus transcription factor KO collection Data sets

No data available

PH-base curated phenotypes Download Data sets

PH-base entity	Essential gene	Multiple mutations	Pathogen species	Pathogen strain	Host species	Host strain	Tissue	Mutant phenotype	Disease
PHB.2661	no	no	Aspergillus fumigatus	Af293	N/A	larva	reduced virulence	invasive pulmonary aspergillosis	no

Phenotype (qualities or directionality + entity or biological process) Download Data sets

Modification	Allele	Phenotype	Further Information	PubMed	CHEBI Annotation Extension
null mutant	medA::hygR	abnormal colony color	Details:delay in brown pigmentation	19889083	N/A
null mutant	medA::hygR	decreased amount biological adhesion	Condition:fibronectin coated wells	19889083	N/A
null mutant	medA::hygR	decreased amount conidium formation	N/A	19889083	N/A
null mutant	medA::hygR	decreased amount virulence	Virulence model:immunosuppressed mouse pulmonary infection	19889083	N/A
null mutant	medA::hygR	decreased amount virulence	Virulence model:insect infection (Galleria mellonella larva)	19889083	N/A
null mutant	medA::hygR	increased amount cell growth	Details:slightly larger conidia and conidiophores	19889083	N/A

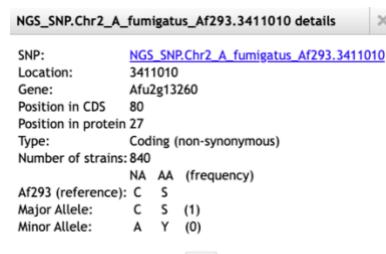
The **Genetic variation** section summarizes integrated SNP data for a given region and classifies SNPs based on the effect on gene function:

- noncoding (yellow diamonds)
- non-synonymous (dark blue)
- synonymous (light blue)
- nonsense (red)



Note that you can interact with the SNP records by using left and right clicking options on your mouse/touch pad.

Left click brings up a pop-up window containing more information about a particular SNP:



The SNP record linked in blue is linked to the SNP record page, which contains summary of the SNP across different isolates and samples.

Add to basket Add to favorites Download SNP

### SNP: NGS\_SNP.Chr2\_A\_fumigatus\_Af293.3411010

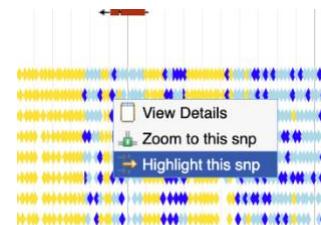
Organism: Aspergillus fumigatus Af293  
Location: Chr2\_A\_fumigatus\_Af293: 3,411,010  
Type: coding  
Number of Strains: 840  
Gene ID: Afu2g13260  
Gene Strand: reverse  
Major Allele: G (1)  
Minor Allele: T (0)  
Distinct Allele Count: 2  
Reference Allele: G  
Reference Product: S 27  
Allele (gene strand): C  
SNP context: AGCCGATCCGTCTGCCCTGCATTTGCCAAAGGAGCAGCAGTGCTCAAGGAAGAAAGAGGGCA  
SNP context (gene strand): TGCCCTCTTCTCCCTTGAGCACTGCTGCTCTTTGCCAAATCGAGCAGACGGATCGGCT

**Major allele** is the most common allele in the studied population/isolates.

**Minor allele** frequency is the frequency of the second most common allele. Minor allele

frequency is useful when looking for rare variants or disease-causing SNPs. However, there can be exceptions when minor allele frequency increases under selective pressure (e.g., development of drug resistance).

Right click provides more options for JBrowse view:



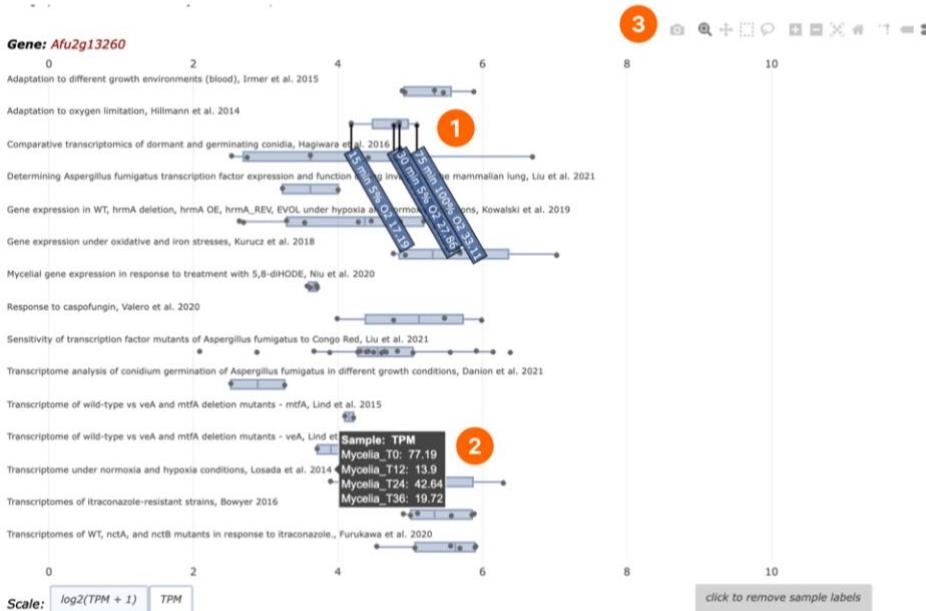
## The Transcriptomics section (RNA-Seq and microarray data).

The Transcript Expression Summary section provides a big picture of gene expression across different samples and experiments, and helps identify experiments in which the current gene is highly regulated.

▼ RNA-Seq Transcription Summary    Data sets

▼ Summary of expression values. Each row represents a distinct RNA-Seq experiment. [Click to read more...](#)

- Each black dot represents expression in a sample. A boxplot is shown, with the box representing the median and upper/lower quartiles, and the whiskers representing the max/min values (or 1.5 times the interquartile range, in which case values beyond the whiskers are considered outliers).
- Hover over the experiment name to show a table of sample names and values.
- Click on a boxplot to show labels and values for each sample in an experiment. Click again on the boxplot to hide labels and values. A button at the bottom-right removes labels and values for all experiments.
- Use the toggle button at the bottom-left to switch from Log Scale to Linear Scale.
- Log Scale values are  $\log_2(\text{TPM} + 1)$  for these reasons:
  - TPM+1: to de-emphasize low noisy TPM values (i.e., <1).
  - log<sub>2</sub>: so that each unit on the x-axis represents a 2-fold difference.
- Navigation buttons appear at the above-right when hovering over the graph. If the buttons do not appear, reload the page.
- Zoom in with the navigation button or click and drag within the graph. Zoom out with the navigation button or double-click within the graph.
- This graph was created with Plotly. [Get more help at their website.](#)



1

Clicking on the box plot will bring up sample labels.

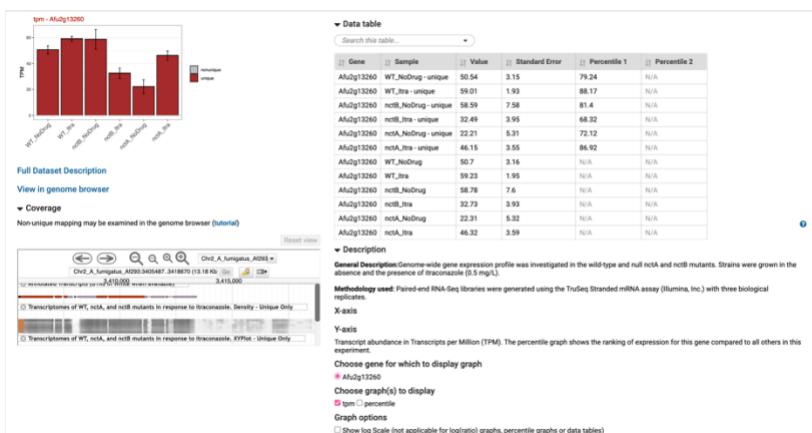
2

Hovering over the experiments will display sample names.

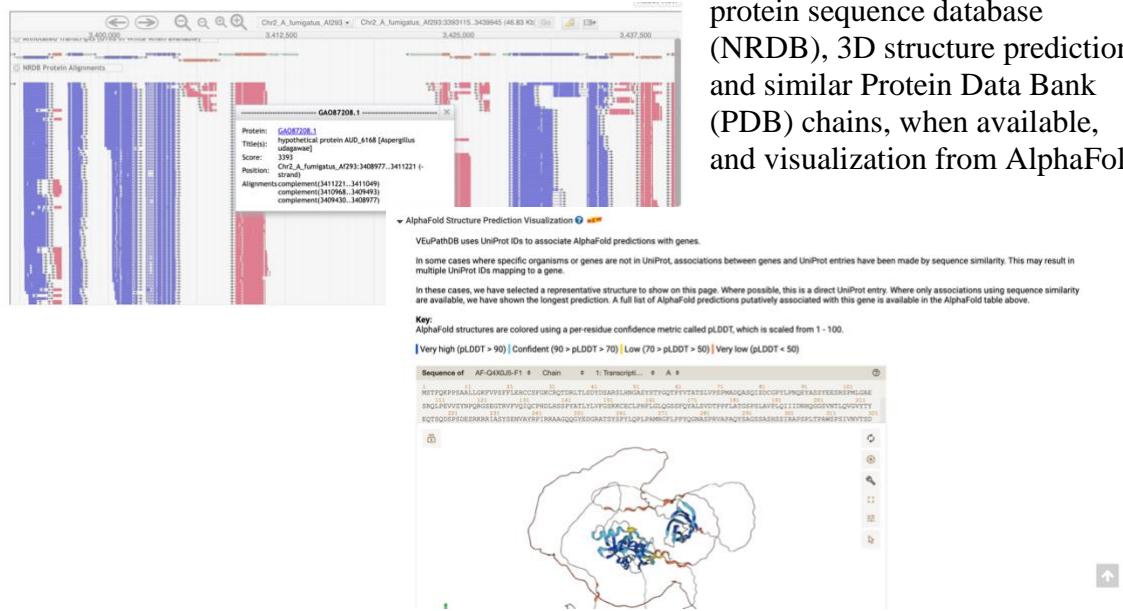
3

The buttons above the summary graph provide additional options (e.g., download data in PNG, zoom, pan, etc.).

The **Transcript Expression** section, which is located under the RNA-Seq summary section, can be expanded to view the expression graph (TPM), data table, full dataset description, coverage plots, a link to the dataset in JBrowse.



The **Sequence analysis, Sequences, and Structure analysis** sections offer sequence information (DNA, RNA, and protein), an interactive summary of EST alignments and BLAT hits against the nonredundant protein sequence database (NRDB), 3D structure predictions and similar Protein Data Bank (PDB) chains, when available, and visualization from AlphaFold.



The **Protein features and properties** section provides access to information about protein domains identifications (InterPro), signal and transmembrane predictions graphics, BLASTP hits, and other tools that can be deployed directly from the gene record page using the amino acid sequence of interest:

<b>14 Protein features and properties §</b>
▶ Attributes and Protein Brower
▶ BLASTP (protein-protein BLAST)
▶ GPI anchor prediction: big-PI Predictor
▶ InterPro Domains
▶ InterProScan: Run on EBI site
▶ MitoProt
▶ STRING: functional protein association networks
▶ WoLF PSORT

The **Function Prediction** section features Gene Ontology (GO) assignments that have been either downloaded from Gene Ontology databases and manually curated by FungiDB. GO Slim terms, Enzyme Commission (EC) numbers with links to EC number and GO terms description and relevant publications are available as well.

Gene ontology provides statements for describing the functions of genes along with three aspects: Molecular function, Biological process, and Cellular component, and it is organized in hierarchies. The GO terms table above provides GO IDs and terms associated with a particular gene and additional metadata that are available for these associations such as source, evidence code (e.g. IDA, IMP, etc.) and reference (PubMed ID).

There are three classes of GO terms in FungiDB:

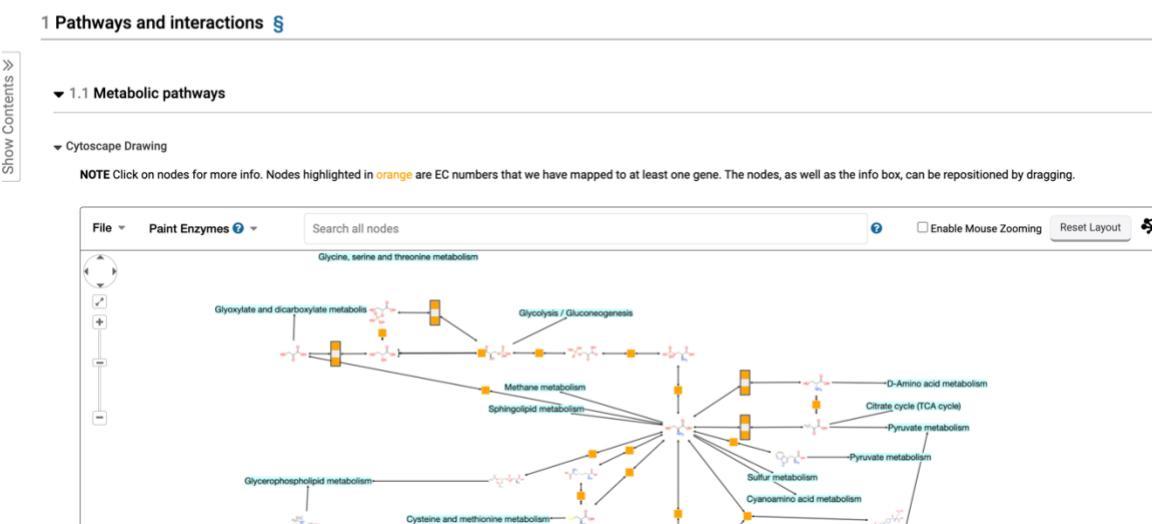
1. Automatically assigned by InterPro2GO
2. Assigned by FungiDB curators
3. Obtained from external resources such as AspGD, MIPS, and others

For some genes, the **Pathways and interactions** section provides information about metabolic pathways loaded from the KEGG and MetaCyc repositories. Genes are linked to individual pathways by EC numbers when this data is available and clicking on any of the links redirects to an interactive metabolic pathway viewer where the user can explore individual reactions or export all known data for a given pathway. EC numbers are inferred by the OrthoMCL pipeline automatically. As with any automated analysis, use caution when interpreting this data.

For example, Afu2g13260 is not associated with metabolic pathways but its neighbor ([Afu2g13250](#)) does:

### Glycine, serine and threonine metabolism

Pathway Source: KEGG  
 Pathway Id: [ec00260](#)  
 Total Pathway Enzymes: 81  
 Total Pathway Compounds: 48



# Exploring records in JBrowse

## Learning objectives

- Become familiar with JBrowse layout and navigation menus

JBrowse can be accessed from the main menu and also gene record pages.

Accessing JBrowse from gene record pages will pre-select the genome automatically. If navigating directly from the main page, the genome of choice can be chosen under the Genome tab.

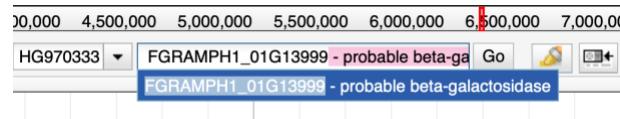
The screenshot shows the JBrowse interface for the gene Afu2g13260. The sidebar on the left lists various sections such as Gene models, Annotations, and Genomic features. The main panel displays a genomic track for the Chz2\_A\_Sumigatake\_A0293393180\_3423443 (40.29 Kb) region. A red arrow points to the 'Gene browser' button in the top right corner of the sidebar.

## 1. Navigate to the gene **FGRAMPH1\_01T12283** in JBrowse.

The screenshot shows the JBrowse interface for the Fusarium graminearum PH-1 genome. Numbered callouts point to: 1. Main menu options (Track, View, Help). 2. Current genome selection (Fusarium graminearum PH-1). 3. Navigation bar (zoom, pan, highlighter, sequence selector, search, overview bar). 4. Select tracks dropdown. The main panel displays a genomic track for the FGRAMPH1\_01T12283 gene, which encodes a hypothetical protein.

- 1 The main menu provides options to select genomes, upload sequence or tracks, customize JBrowse view, and access several help articles.
- 2 Current genome selection id indicated on the right. The Share link generate a custom URL specific to the JBrowse session & tracks loaded.
- 3 The navigation bar contains zooming (magnifying glass icons), panning (left/right arrows) and highlighting (yellow highlighter) buttons, reference sequence selector (drop down with sequences from the selected genome sorted by length), a text box to search for features such as gene ids, and overview bar showing the location of the region in view.
- 4 Select additional tracks (transcriptomics, SNPs, sequence, ploidy, etc.) to customise your JBrowse view.

If you navigated away from the gene, you could type **FGRAMPH1\_01T12283** directly in the location search box and select the highlighted match to return to your original position.

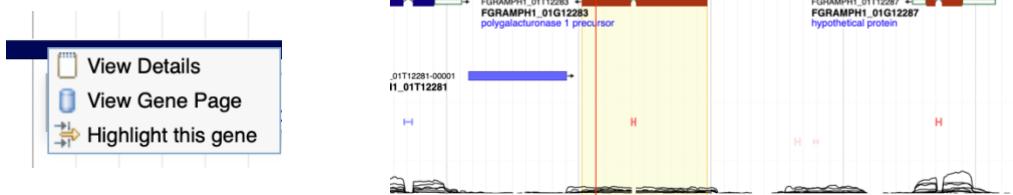


- **Displaying more information about a gene**

Left clicking on a gene of interest will display a pop-up window that provides additional information about gene classification, position, orthology, links to JBrowse and gene record pages, as well CDS and protein sequences.

- **Highlighting gene for easy tracking**

Right-click (or control click) on the gene feature to display the context menu. The context menu offers gene highlighting and also options to display the gene record page and left-click details menu described above.



- **Practice activating tracks in JBrowse** by clicking on the Select tracks tab and choosing to display the following tracks:

1. Intron Evidence (if not selected by default already)
2. RNA-Seq (mycelium and 6days post-infection of wheat with Fusarium (6dpi))
3. Chip-Seq tracks of chromatin marks in WT and kmt6 mutant (KMT6 is Histone H3 K27 Methyltransferase)
4. Syntenic Sequences and Genes (Shaded by Orthology)

- Select the **Transcriptomics** category, **RNA-seq** subcategory and choose tracks:
  1. Click on the “**Transcriptomic analysis during vegetative and infectious growth of Fusarium graminearum PH-1**” dataset
  2. Select **unique** in the RNA-Seq Alignment category
  3. Select two tracks: **infected 6dpi** and **mycelia** as shown below

<input checked="" type="checkbox"/>	Transcriptomic analysis during vegetative and infectious growth of <i>Fusarium graminearum</i> PH-1 - infected 6dp (unique) Coverage	Transcriptomics	RNA-Seq	Transcriptomic analysis during vegetative and infectious growth of <i>Fusarium graminearum</i> PH-1	Coverage	unique	not strand specific
<input checked="" type="checkbox"/>	Transcriptomic analysis during vegetative and infectious growth of <i>Fusarium graminearum</i> PH-1 - mycelia (unique) Coverage	Transcriptomics	RNA-Seq	Transcriptomic analysis during vegetative and infectious growth of <i>Fusarium graminearum</i> PH-1	Coverage	unique	not strand specific

- Clear your search by clicking on the “Clear All Filters” button.
- Select the **Epigenomics** category, **ChIP-Seq** subcategory and select H3K27me3 and H3K4me3 methylation marks in kmt6 deletion mutant and WT in low nitrogen growth conditions (low). The exact track names are provided below. Use the “Contains text” filter window to search for the tracks (copy and paste names listed below (text only) into the filter window).

For kmt6 deletion mutant:

- H3K27me3\_DELkmt6\_neoR\_low-R2
- H3K4me2\_DELkmt6\_neoR\_low-R2

For WT:

- H3K27me3\_WT\_low
- H3K4me2\_WT\_low

- Activate the **Syntenic Sequences and Genes (Shaded by Orthology)** track, which is located under the **Comparative Genomics** category, **Orthology and Synteny** subcategory but can be also searched for in the “Contains text” filter box at the top.

- Click on the **Back to browser** button to return to JBrowse.

- Customize the JBrowse syntenic view to display *Fusarium* species only.

Hint: Click on the drop-down menu in the **Syntenic Sequences and Genes (Shaded by Orthology)** track, choose Select Subtracks, unselect current tracks, use the filter to identify “*Fusarium*” and then click on the “Save” button at the bottom.

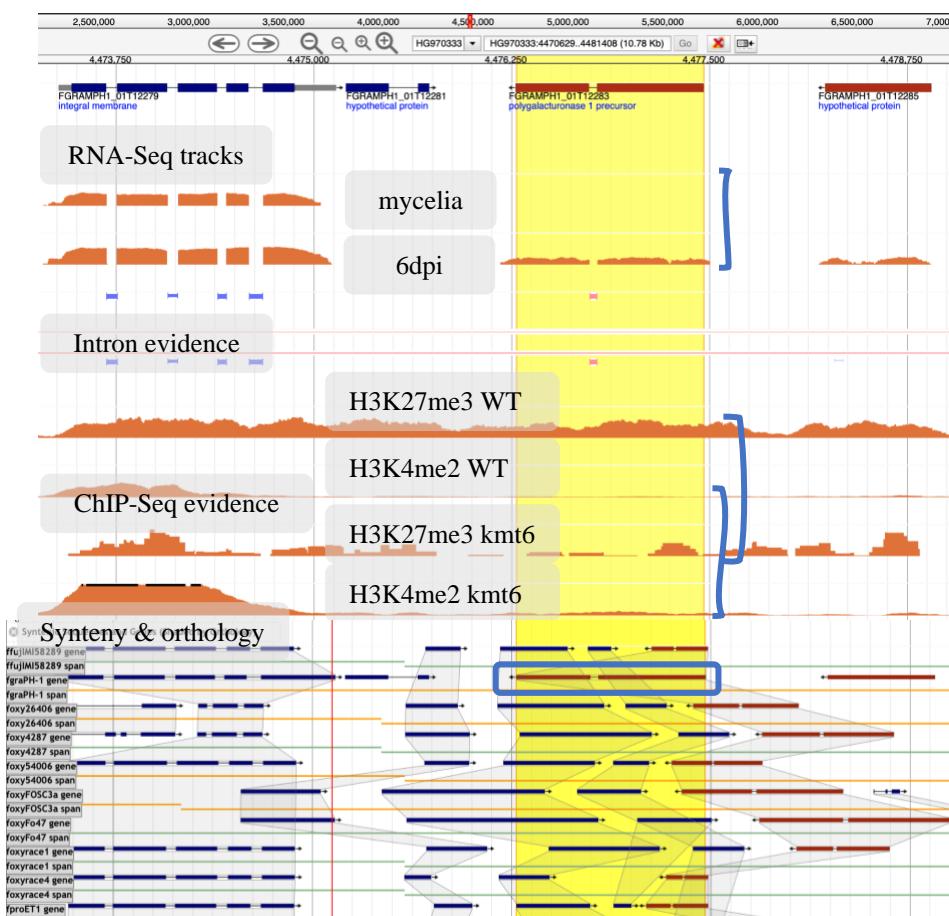
- Re-arrange tracks within JBrowse in the following order:
  - Pin the highlighted gene and gene model track to the top.

Note: you can use the drag and drop function in JBrowse to rearrange tracks in the order that helps you to evaluate the data.

In *Fusarium* and other fungi, H3K4me2 and H3K27me3 are found in large, mutually exclusive, gene-rich blocks of the genome. About one-third of the *F. graminearum* genome is associated with H3K27me3 when the fungus is grown in a minimal medium with low nitrogen. Kmt6 encodes a DNA methylase that deposits repressive chromatin mark H3K27me3. kmt6 mutants in *Fusarium* and other species demonstrate reorganization of chromatin marks (e.g., absence of the repressive chromatin mark H3K27me3) and upregulation of genes that are normally suppressed.

Having this information in hand and using JBrowse track view, answer the following questions:

- What can you tell about FGRAMPH1\_01T12283 expression mycelia and 6 days post-infection?
- Does this gene show de-regulation of repressive chromatin marks (H3K27me3) in the kmt6 mutant? Would you expect the expression of this gene to be up-regulated or down-regulated in the kmt6 mutant?
- Is this gene conserved in *Fusarium* species?
- How would you generate a unique URL to this JBrowse view?



URL: [https://fungidb.org/fungidb/jbrowse/index.html?loc=HG970333%3A4474707..4479028&data=%2Ffungidb%2Fservice%2Fbrowse%2Ftracks%2FfgraPH-1&tracks=+gene%2CCCommunity%20annotations%20from%20Apollo%2CRNA-Seq%20Evidence%20for%20Introns%2CFgraminearumPH-1%20combined%20RNASeq%20plot%2CfgraPH-1\\_Wang\\_VegetativeAndInfectiousGrowth\\_ubi\\_rnaSeq\\_RSRC%201\\_infected\\_6dpi\\_unique%20Coverage%2CfgraPH-1\\_Wang\\_VegetativeAndInfectiousGrowth\\_ubi\\_rnaSeq\\_RSRC%201\\_mycelia\\_unique%20Coverage%2CfgraPH-1\\_Freitag\\_histonemod\\_nitrogenLevel\\_chipSeq\\_RSRC%20H3K27me3\\_DELkmt6\\_neoR\\_low-R2%20Coverage%2CfgraPH-1\\_Freitag\\_histonemod\\_nitrogenLevel\\_chipSeq\\_RSRC%20H3K27me3 WT\\_low%20Coverage%2CfgraPH-1\\_Freitag\\_histonemod\\_nitrogenLevel\\_chipSeq\\_RSRC%20H3K4me2 WT\\_low%20Coverage%2CfgraPH-1\\_Freitag\\_histonemod\\_nitrogenLevel\\_chipSeq\\_RSRC%20H3K4me2\\_low%20Coverage%2CfgraPH-1\\_Freitag\\_histonemod\\_nitrogenLevel\\_chipSeq\\_RSRC%20Sequences%20and%20Genes%20Shaded%20by%20Orthology&highlight=HG970333%3A4476280..4477469](https://fungidb.org/fungidb/jbrowse/index.html?loc=HG970333%3A4474707..4479028&data=%2Ffungidb%2Fservice%2Fbrowse%2Ftracks%2FfgraPH-1&tracks=+gene%2CCCommunity%20annotations%20from%20Apollo%2CRNA-Seq%20Evidence%20for%20Introns%2CFgraminearumPH-1%20combined%20RNASeq%20plot%2CfgraPH-1_Wang_VegetativeAndInfectiousGrowth_ubi_rnaSeq_RSRC%201_infected_6dpi_unique%20Coverage%2CfgraPH-1_Wang_VegetativeAndInfectiousGrowth_ubi_rnaSeq_RSRC%201_mycelia_unique%20Coverage%2CfgraPH-1_Freitag_histonemod_nitrogenLevel_chipSeq_RSRC%20H3K27me3_DELkmt6_neoR_low-R2%20Coverage%2CfgraPH-1_Freitag_histonemod_nitrogenLevel_chipSeq_RSRC%20H3K27me3 WT_low%20Coverage%2CfgraPH-1_Freitag_histonemod_nitrogenLevel_chipSeq_RSRC%20H3K4me2 WT_low%20Coverage%2CfgraPH-1_Freitag_histonemod_nitrogenLevel_chipSeq_RSRC%20H3K4me2_low%20Coverage%2CfgraPH-1_Freitag_histonemod_nitrogenLevel_chipSeq_RSRC%20Sequences%20and%20Genes%20Shaded%20by%20Orthology&highlight=HG970333%3A4476280..4477469)