

DAY1

Introduction to FungiDB Site Search and Search Strategies

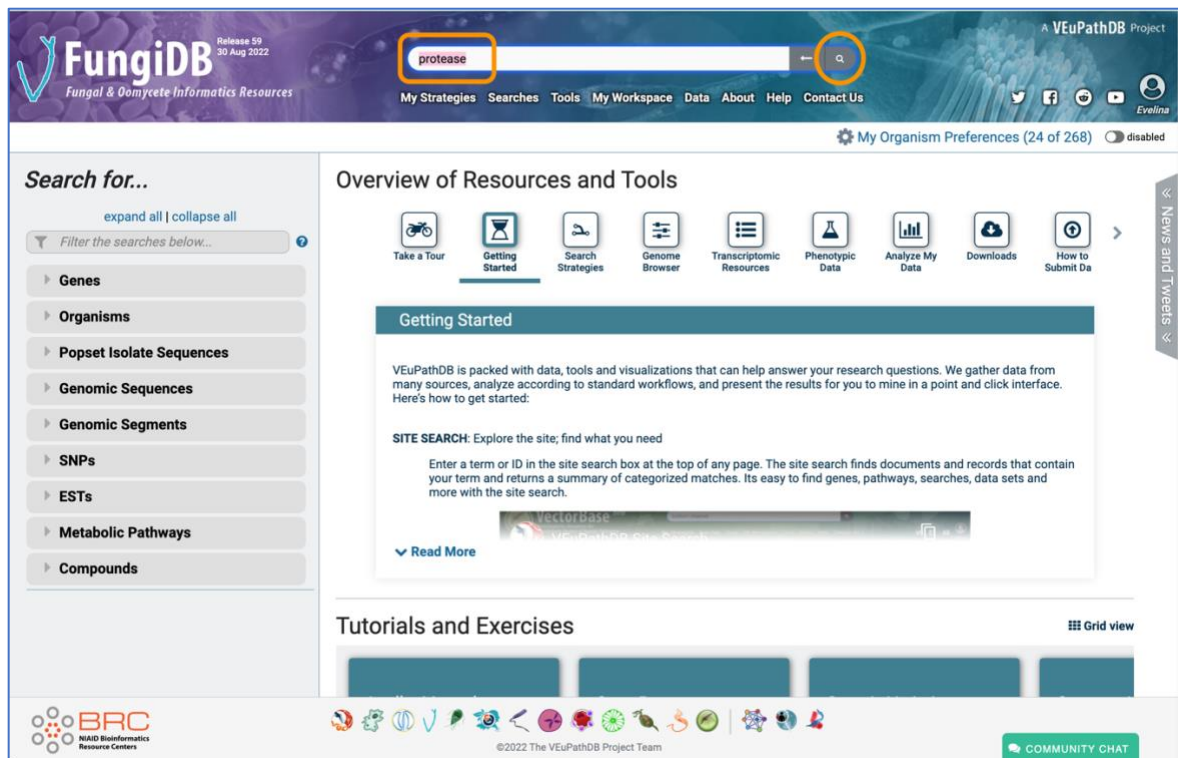
Learning objectives:

- Deploy Site Search using terms
- Filter site search results by categories and fields
- Export results to a search strategy
- Use Site Search to find a gene by its ID
- Navigate to and from the site search result
- Explore searches using wild cards (*)

The site search is in the center of the header, which is available from every page and throughout navigation of the genomics site. The site search queries the databases for a term or ID and returns a list of pages and documents that contain this query term.

1. Deploy Site Search using a term.



- a. Enter the word **protease** in the site search window
- b. Click enter on a keyboard or click on the search icon (magnifying glass icon).



2. Understand the Site Search result format.

The site search returns a categorized list of pages and documents that contain the search term. The categorized summary of the results shown in the left panel. The details panel is on the right. Filtering the summary table on the left will populate the details panel on the right accordingly.

- a) How many data categories have “protease” word?

 My Organism Preferences (24 of 268)  disabled

All results matching **protease**

Export as a Search Strategy
to download or mine your results

Categorised summary of
the Site Search results

Details panel that reflects selections in the
Categorised summary panel on the left

☒ Hide zero counts

Filter results


Genome	
Genes	75,428
Population biology	
Popset isolate sequences	2,035
Metabolism	
Compounds	73
Data access	
Data sets	1
About	
General info pages	1

Filter fields

Select a result filter above

Filter organisms

select all | clear all | expand all | collapse all

Type a taxonomic name 

☐ Fungi

55,986

☐ Oomycota

19,443

Gene - AKAW_09677 protease

Organism: Aspergillus luchuensis IFO 4308

Fields matched: InterPro domains; Orthologs; PDB chains; Product descriptions

Gene - BO70DRAFT_297753 protease

Organism: Aspergillus heteromorphus CBS 117.55

Fields matched: Orthologs; Product descriptions

Gene - BO72DRAFT_369796 protease

Organism: Aspergillus fijiensis CBS 313.89

Fields matched: Orthologs; Product descriptions

Gene - BO83DRAFT_443944 protease

Organism: Aspergillus eucalypticola CBS 122712

Fields matched: InterPro domains; Orthologs; PDB chains; Product descriptions

Gene - CIHG_00878 protease

Organism: Coccidioides immitis H538.4

Fields matched: Product descriptions

Gene - CIMG_06482 protease

Organism: Coccidioides immitis RS

Fields matched: Orthologs; Product descriptions



For this exercise, make sure that the “My Organism Preferences” is disabled. We will come back to this filter in a few steps.

3. Filter the site search results by category.

- a. Select **Genes** category to filter the site Search results.



Notice that the “Export as a Search Strategy” button turned dark blue when you applied “Genes” category. The color change occurs when this feature become available/active. For this to happen, a *single filter category* must be selected. For example, you can export all results in the “Genes” or “Popset isolate sequences” (but not both).

- b. Use filter options to limit your results where the word “protease” occurs in **Product descriptions**.
- c. How many of the genes included the word “protease” in their product descriptions?

Filter results ☒ Hide zero counts

Genome	
Genes	75,428
Population biology	
Popset isolate sequences	2,035
Metabolism	
Compounds	73
Data access	
Data sets	1
About	
General info pages	1

Filter fields
Select a result filter above

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

Type a taxonomic name

☐ Fungi 55,986
☐ Oomycota 19,443

Filter results ☒ Hide zero counts

Genome [Clear filter](#)
Genes 75,428

Filter Gene fields [Apply](#) ☒

select all | clear all

<input type="checkbox"/> EC descriptions and numbers	201
<input type="checkbox"/> GO terms	350
<input type="checkbox"/> InterPro domains	25,105
<input type="checkbox"/> Notes from annotators	246
<input type="checkbox"/> Orthologs	66,625
<input type="checkbox"/> PDB chains	9,052
<input type="checkbox"/> Phenotype	7
<input checked="" type="checkbox"/> Product descriptions	6,489
<input type="checkbox"/> PubMed	903
<input type="checkbox"/> User comments	5

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

Type a taxonomic name

☐ Fungi 55,985
☐ Oomycota 19,443



To return to the original results summary view or to apply new filters, click on the **Clear filter** option.

Filter results ☒ Hide zero counts

Genome [Clear filter](#)
Genes 75,428

Filter Gene fields [Clear filter](#)

select all | clear all

<input type="checkbox"/> EC descriptions and numbers	201
<input type="checkbox"/> GO terms	350
<input type="checkbox"/> InterPro domains	25,105
<input type="checkbox"/> Notes from annotators	246
<input type="checkbox"/> Orthologs	66,625
<input type="checkbox"/> PDB chains	9,052
<input type="checkbox"/> Phenotype	7
<input checked="" type="checkbox"/> Product descriptions	6,489

4. Filter the site search result by Organism.

Filter organisms [Apply](#) ☒

select only these | add these | clear these

☐ Fungi 3,814
☐ Ascomycota 3,018
☐ Sordariomycetes 794
☐ Hypocreales 483
☐ Nectriaceae 389
☒ Fusarium graminearum 44
☒ Fusarium graminearum PH-1 44

a. Identify how many “protease” genes are annotated in *Fusarium graminearum* PH-1.



Use the filter box to quickly bring up the genome of interest. Click on the “Apply” button to apply your selection.

5. Export the results to a search strategy.

- Click on the “Export as a Search Strategy” button to export your results into a search strategy.

The screenshot shows a blue button labeled "Export as a Search Strategy to download or data mine" with a right-pointing arrow. An orange arrow points down to a screenshot of the FungiDB search strategy interface. The interface has a header "Unnamed Search Strategy" with a link to rename it to "protease PH-1". Below is a "Text" box with "44 Genes" and an "Add a step" button. The main section shows "44 Genes (37 ortholog groups)" with tabs for "Gene Results", "Genome View", and "Analyze Results". A table lists gene results with columns for Gene ID, Transcript ID, Organism, Genomic Location, Product Description, and Score. The table contains 5 rows of data for *Fusarium graminearum* PH-1 genes.

Gene ID	Transcript ID	Organism	Genomic Location (Gene)	Product Description	Score
FGRAMPH1_01G00411	FGRAMPH1_01T004...	<i>Fusarium graminearum</i> PH-1	HG970332:504,157..504,345(-)	micromolar calcium activated neutral protease 1	10.15
FGRAMPH1_01G00869	FGRAMPH1_01T008...	<i>Fusarium graminearum</i> PH-1	HG970332:1,044,183..1,046,989(+)	adam protease adm-b	11.95
FGRAMPH1_01G01423	FGRAMPH1_01T014...	<i>Fusarium graminearum</i> PH-1	HG970332:1,770,694..1,773,671(+)	26S protease regulatory subunit 7	10.97
FGRAMPH1_01G02967	FGRAMPH1_01T029...	<i>Fusarium graminearum</i> PH-1	HG970332:3,949,107..3,950,453(+)	26S protease subunit rpt4	11.95
FGRAMPH1_01G03623	FGRAMPH1_01T036...	<i>Fusarium graminearum</i> PH-1	HG970332:4,899,278..4,902,642(+)	cell division protease ftsH	11.95

Congratulations!! You just created your first search strategy in FungiDB. The results can now be combined with other specialized searches using the Add Step button.



Note that all strategy can be saved or annotated with additional information. You can also generate a unique URL to share the strategy with your colleagues.



6. Return to the site search results page and run a new query.

- The filter settings used the previous search were stored in the search. Click on the “back to results” **arrow icon** as shown below to return to the original site search filter params:

The screenshot shows the FungiDB homepage with a search bar. A red arrow points to a "back to results" icon (a left-pointing arrow) in the top right of the search bar. Below the search bar is a "My Organism Preferences (24 of 268)" section. To the right is a "Filter results" panel with sections for "Filter Gene fields" and "Filter organisms". The "Filter Gene fields" section has checkboxes for InterPro domains (119), Orthologs (269), PDB chains (53), Product descriptions (44), and User comments (1). The "Filter organisms" section has checkboxes for Fungi (3,814) and Oomycota (2,675).



Internet browser's back arrow can be also used to return to the site search filter settings.

b) **Clear all filters.**

Filters can be cleared individually or in bulk. *Use the bulk setting to clear all filters from the previous search.*

Clearing filters individually:

Filter results ☒ Hide zero counts

Genome **Clear filter**

Genes 44

Filter Gene fields **Clear filter**

select all | clear all

<input type="checkbox"/> InterPro domains	119
<input type="checkbox"/> Orthologs	269
<input type="checkbox"/> PDB chains	53
<input checked="" type="checkbox"/> Product descriptions	44
<input type="checkbox"/> User comments	1

Filter organisms **Clear filter**

select all | clear all | expand all | collapse all

Type a taxonomic name

<input checked="" type="checkbox"/> Fungi	3,814
<input type="checkbox"/> Oomycota	2,675

Clearing filter in bulk:

FungiDB Release 55 30 Aug 2022
Fungal & Oomycete Informatics Resources

Clear filters

My Strategies Searches Tools My Workspace Data About Help Contact Us

c) **Run a wild card search.**

The wild card (denoted by an asterisk *) can be used alone to retrieve all site search results or combined with a word (e.g., **kinase*) to retrieve compound words ending with the word *kinase* like *carbohydrate purine kinase* or *phosphofructokinase*.

- Run a search for ***kinase**
- Filter on **Genes** and **Product descriptions**
- Limit your search to **Phytophthora infestans T30-4** only
- Export your search as a strategy

**kinase in Phytophthora infestans T30-4 **

Text 559 Genes

Step 1

7. Search for a specific gene ID.

Site search can be also used to search for specific IDs.

- Copy and paste one of the IDs from the previous search into the Site Search box
- Click enter on the keyboard or “look up” icon next to the site search box

The screenshot shows the FungiDB search results for the query 'PITG_00005'. The search bar at the top contains the query. Below the search bar, there is a table of results. The first row of the table is highlighted with an orange box, and an orange arrow points from the search bar to this row. The table has columns for Gene ID, Transcript ID, Organism, Genomic Location (Gene), and Product Description. The highlighted row shows the gene PITG_00005, transcript PITG_00005-t26, organism *Phytophthora infestans* T30-4, and product description 'protein kinase, putative'.

Gene ID	Transcript ID	Organism	Genomic Location (Gene)	Product Description
PITG_00005	PITG_00005-t26...	<i>Phytophthora infestans</i> T30-4	DS028118:38,775..39,071(+)	protein kinase
PITG_00004	PITG_00004-t26...	<i>Phytophthora infestans</i> T30-4	DS028118:39,526..41,494(+)	protein kinase, putative
PITG_00005	PITG_00005-t26...	<i>Phytophthora infestans</i> T30-4	DS028118:42,014..43,659(+)	protein kinase, putative

When there is an exact match for an ID in the database, the site search offers a card in the details panel to draw attention to the direct link to the gene page.

The screenshot shows the FungiDB search results for the query 'PITG_00005'. The search results are displayed in a table. The first row of the table is highlighted with a blue box, and a blue arrow points from the search bar to this row. The table has columns for Gene ID, Transcript ID, Organism, Genomic Location (Gene), and Product Description. The highlighted row shows the gene PITG_00005, transcript PITG_00005-t26, organism *Phytophthora infestans* T30-4, and product description 'protein kinase, putative'.

Gene ID	Transcript ID	Organism	Genomic Location (Gene)	Product Description
PITG_00005	PITG_00005-t26...	<i>Phytophthora infestans</i> T30-4	DS028118:38,775..39,071(+)	protein kinase
PITG_00004	PITG_00004-t26...	<i>Phytophthora infestans</i> T30-4	DS028118:39,526..41,494(+)	protein kinase, putative
PITG_00005	PITG_00005-t26...	<i>Phytophthora infestans</i> T30-4	DS028118:42,014..43,659(+)	protein kinase, putative



This search results can be also exported as a search strategy when the results are filtered on “Genes”. This is quite handy if you wish to start your search with a single GeneID.

My Organism Preferences

Learning objectives:

- Set custom My Organism Preferences parameters
- Enable and disable the tool

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 only data, search menus and tools will be filtered based on the selected organism/s helping you focus on your work.

1. Set custom “My Organism Preferences” parameters.

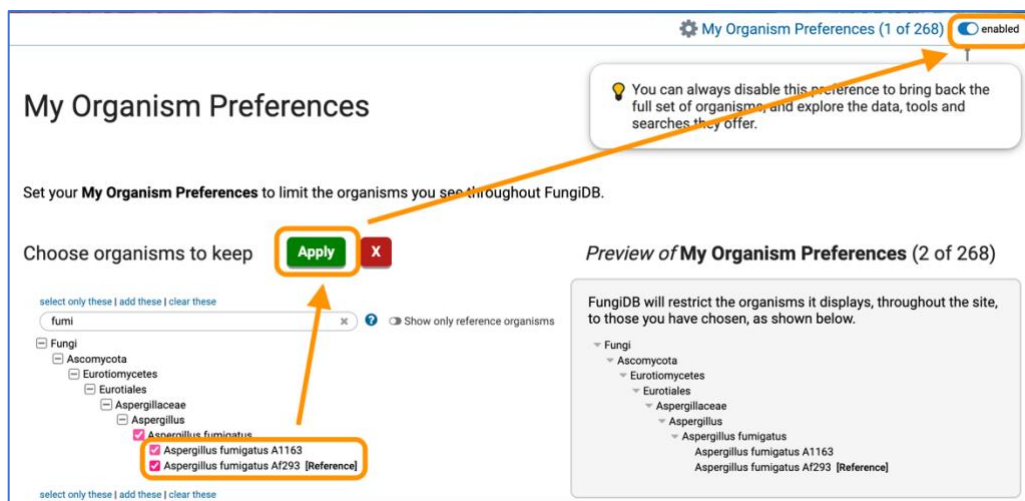
- a) Click on the My Organism Preference link at the top of the page

 My Organism Preferences

- Click on the “clear these” to remove any default filters

select only these | add these **clear these**

- Filter on **fumi** for “fumigatus”
- Select both strains (A1163 and Af293)
- Click on the **Apply** button
- **Enable** the toggle to activate your preferences



My Organism Preferences (1 of 268) **enabled**

You can always disable this preference to bring back the full set of organisms, and explore the data, tools and searches they offer.

Set your **My Organism Preferences** to limit the organisms you see throughout FungiDB.

Choose organisms to keep **Apply** **X**

select only these | add these | clear these

fumi

☐ Fungi

- ☐ Ascomycota
 - ☐ Eurotiomycetes
 - ☐ Eurotiales
 - ☐ Aspergillaceae
 - ☐ Aspergillus
 - ☒ Aspergillus fumigatus
 - ☒ Aspergillus fumigatus A1163
 - ☒ Aspergillus fumigatus Af293 [Reference]

select only these | add these | clear these

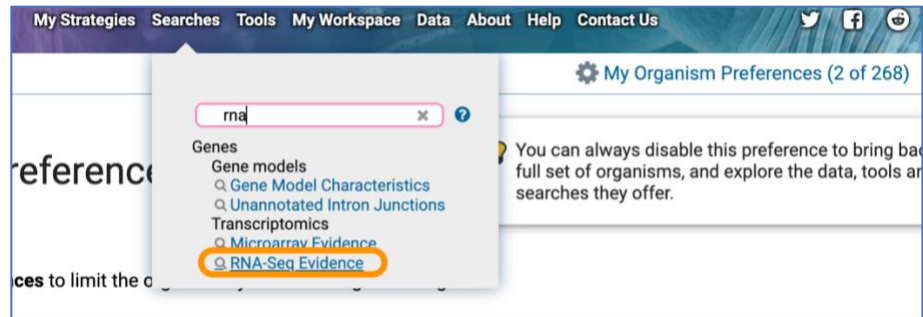
Preview of My Organism Preferences (2 of 268)

FungiDB will restrict the organisms it displays, throughout the site, to those you have chosen, as shown below.

- Fungi
 - Ascomycota
 - Eurotiomycetes
 - Eurotiales
 - Aspergillaceae
 - Aspergillus
 - Aspergillus fumigatus
 - Aspergillus fumigatus A1163
 - Aspergillus fumigatus Af293 [Reference]

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

- a) Navigate to the “RNA-Seq Evidence” search page.



Organism	Data Set	Choose a Search
<i>Aspergillus fumigatus</i> Af293	Transcriptome analysis of conidium germination of <i>Aspergillus fumigatus</i> in different growth conditions (Danion et al. 2021)	DE FC P SA
<i>Aspergillus fumigatus</i> Af293	Gene expression in WT, hrma deletion, hrma OE, hrma_REV, EVOL under hypoxia and normoxia conditions (Kowalski et al. 2019)	DE FC P SA
<i>Aspergillus fumigatus</i> Af293	Transcriptomes of WT, nctA, and nctB mutants in response to itraconazole. (Furukawa et al. 2020)	DE FC P
<i>Aspergillus fumigatus</i> Af293	Gene expression under oxidative and iron stresses (Kurucz et al. 2018)	DE FC P
<i>Aspergillus fumigatus</i> Af293	Transcriptomes of itraconazole-resistant strains (Bowyer 2016)	DE FC P
<i>Aspergillus fumigatus</i> Af293	Transcriptome under normoxia and hypoxia conditions (Losada et al. 2014)	FC P
<i>Aspergillus fumigatus</i> Af293	Response to caspofungin (Valero et al. 2020)	DE FC P SA
<i>Aspergillus fumigatus</i> Af293	Comparative transcriptomics of dormant and germinating conidia (Hagiwara et al. 2016)	FC P
<i>Aspergillus fumigatus</i> Af293	Adaptation to different growth environments (blood) (Irmer et al. 2015)	DE FC P

Notice that the available datasets are only for *A. fumigatus* Af163 and Af293, which are the organisms selected in the My Organism Preferences parameters.

3. Disable “My Organism Preferences” with a toggle.



Complex Search strategies

Learning objectives:

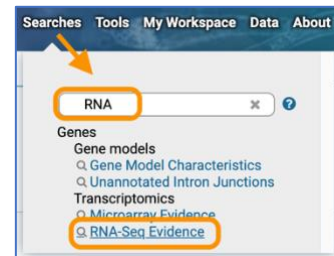
- Create multi-step search strategy
- Enrich results with GO Term Enrichment Analysis

1. Use the integrated RNA-Seq data to identify genes up-regulated in *F. graminearum* PH-1 during infection of wheat

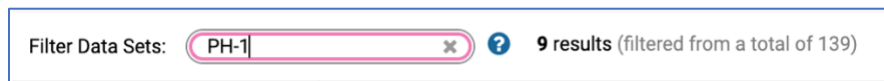


For this exercise, make sure that My Organisms Preference filter is inactive.

a) Navigate to the RNA-Seq Evidence dataset page

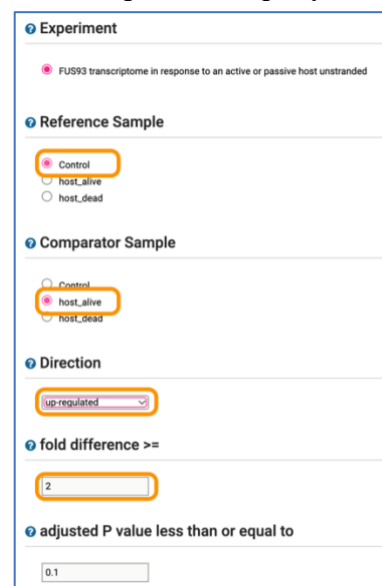


b) Filter for PH-1 (*Fusarium graminearum* PH-1)

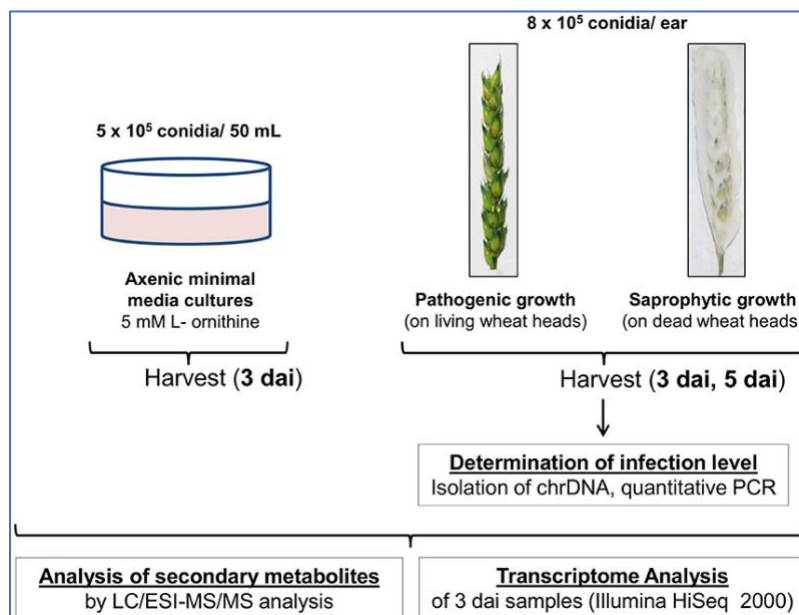


c) Identify *Fusarium* genes up-regulated during infection of a live host

- **Identify** the dataset titled “Transcriptome in response to an active or passive host (Boedi et al. 2016)”
- **Click** on the **DE** button to run the “Differential Expression” query
- **Select parameters:**
 - Control** (Reference)
 - host_live** (Comparator)
 - up-regulated** (Direction)
 - 2 fold** (Fold Difference)
- **Click** on the “Get Answer” button



In this dataset, Boedi al. looked at transcriptomes of *Fg* infecting a living, actively defending plant and dead plant tissues (cold-killed flowering wheat heads). The first condition represented the mixed live style (pathogenic growth on living flowering wheat heads), while the second condition represented strictly saprophytic conditions.



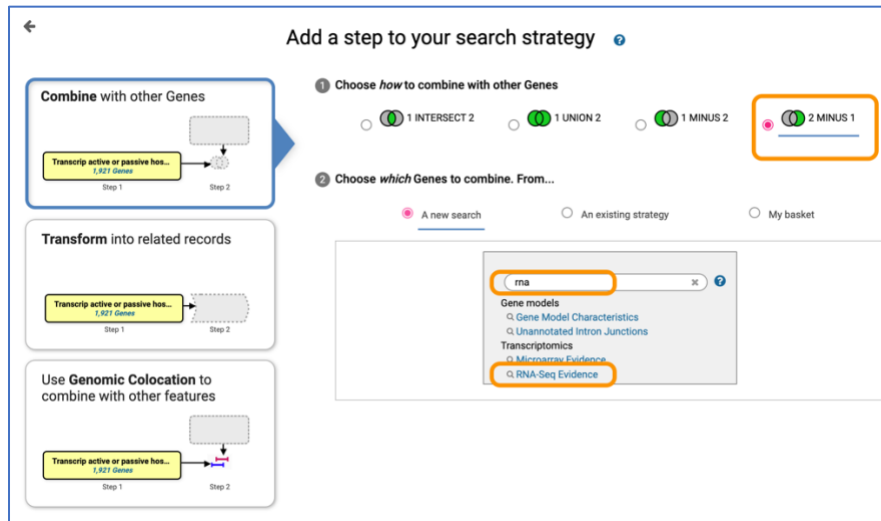
The strategy outlined above looked for the up-regulated genes while *Fusarium* growing on the live plants with the control = Axenic minimal media cultures. How many genes were returned by the search?

Transcript active or passive hos...
1,921 Genes
Step 1

d) Use Boedi et al. dataset to identify *Fusarium* genes up-regulated during saprotrophic growth conditions only.

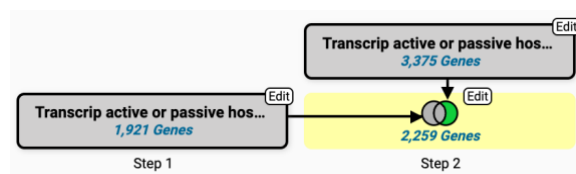
- Click on the “Add Step” button
- Select an appropriate Boolean operator and navigate to the RNA-Seq evidence dataset page as shown below

+ Add a step

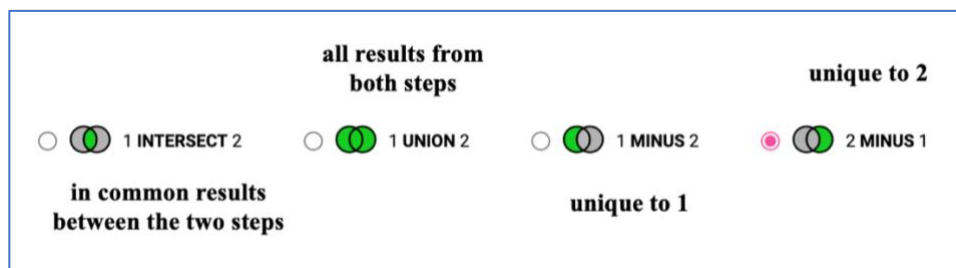


- **Identify** the dataset titled “Transcriptome in response to an active or passive host (Boedi et al. 2016)”
 - **Click** on the **DE** button to run the “Differential Expression” query
 - **Select parameters:**
 - Control** (Reference)
 - host_dead** (Comparator)
 - up-regulated** (Direction)
 - 2 fold** (Fold Difference)
 - **Click** on the “Get Answer” button

How many genes were returned but the search?



When two searches are combined, the two result sets (e.g., list of IDs) are merged. There are different ways to merge strategy results:



Important: Name and then save your FungiDB strategy. We will come back to it tomorrow to learn how to create complex search strategies using different types of data and tools.