

Gene Ontology Enrichment

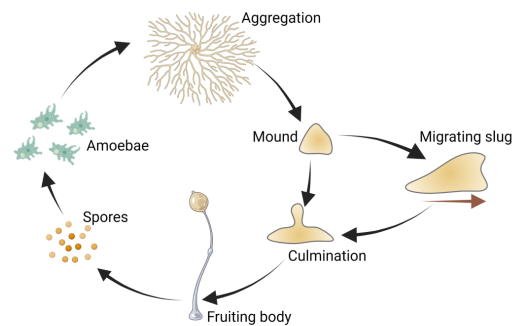
Learning objectives

- Identify a biologically meaningful set of regulated genes
- Analyze results using Gene Ontology (GO) enrichment
- Practice using tools to summarize big lists of GO terms
- Use Revigo to simplify and visualize GO results

Introduction



Dictyostelium (slime mold) is a model system for developmental biology. Its life cycle transitions from free-living single cells to a multicellular structure. Upon starvation, single cells (amoebae) begin to cooperate in a developmental process that leads to multicellular fruiting body



formation 24 hours later. These striking morphological changes are driven by large shifts in gene expression as different biological processes turn on and off.

In this exercise, we will focus on the **genes that become upregulated in the later stages of development**. By analyzing these genes, we can ask questions such as: What biological processes become active as *Dictyostelium* builds a multicellular structure?

To answer this, you will learn how to explore patterns in gene expression, use Gene Ontology (GO) enrichment to identify major biological themes, and visualize the results through tools like word clouds and Revigo.

1. Go to AmoebaDB.org¹ and search for RNA-Seq experiments.
2. In the list of datasets, look for the “**D. discoideum developmental program transcriptome**”.
3. Choose the Differential Expression (DE) search for this dataset.

¹ Note that this exercise uses AmoebaDB.org as an example database, but the same functionality is available on all VEuPathDB resources.



Search for...

ma

Genes

Gene models

- Gene Model Characteristics
- Unannotated Intron Junctions

Transcriptomics

- Microarray Evidence
- RNA-Seq Evidence

Identify Genes based on RNA-Seq Evidence

Legend: DE Differential Expression FC Fold Change P Percentile SA SenseAntisense

Filter Data Sets: 19 rows

Organism	Data Set	Choose a Search
Acanthamoeba castellanii str. Neff	Trophozoite transcriptome of A. castellanii (Wojtkawska et al.)	P
Acanthamoeba castellanii str. Neff	Transcriptome during encystment of A. castellanii Neff (Bernard C. et al. 2022)	DE FC P
Dictyostelium discoideum AX4	D. discoideum developmental program transcriptome (Parikh et al. 2010) NEW	DE FC P
Dictyostelium discoideum AX4	Intracellular M. marinum infection of D. discoideum (Kjellin et al. 2019)	DE FC P
Dictyostelium discoideum AX4	Intracellular L. pneumophila infection of D. discoideum (Kjellin et al. 2019)	DE FC P SA
Dictyostelium purpureum QSDP1	D. purpureum developmental program transcriptome (Parikh et al. 2010) NEW	DE FC P

4. This opens the Differential Expression search page which has data on the RNA sequence transcriptome of *Dictyostelium discoideum* during a 24-hour developmental program. *Hint: this information can be found in the “View Data Sets Used tab”*
5. Run a search to find genes that were upregulated in late stages, i.e., at 24 hr (comparator sample) compared to 16 hr (reference sample). **How many genes do you find?**

Identify Genes based on D. discoideum AX4 D. discoideum developmental program transcriptome RNA-Seq (Differential Expression)

Configure Search Learn More View Data Sets Used

Reset values to default

Experiment

D. discoideum developmental time course transcriptome unstranded

Reference Sample

0 hr
4 hr
8 hr
12 hr
16 hr
20 hr
24 hr
NC4 prespore
NC4 prestalk

Comparator Sample

0 hr
4 hr
8 hr
12 hr
16 hr
20 hr
24 hr
NC4 prespore
NC4 prestalk

Direction

up-regulated

fold difference >=

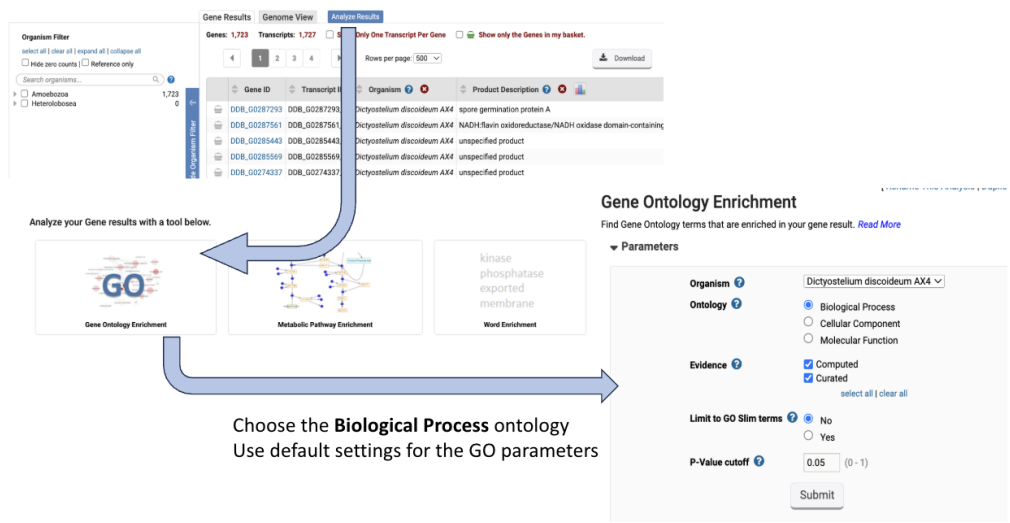
2

adjusted P value less than or equal to

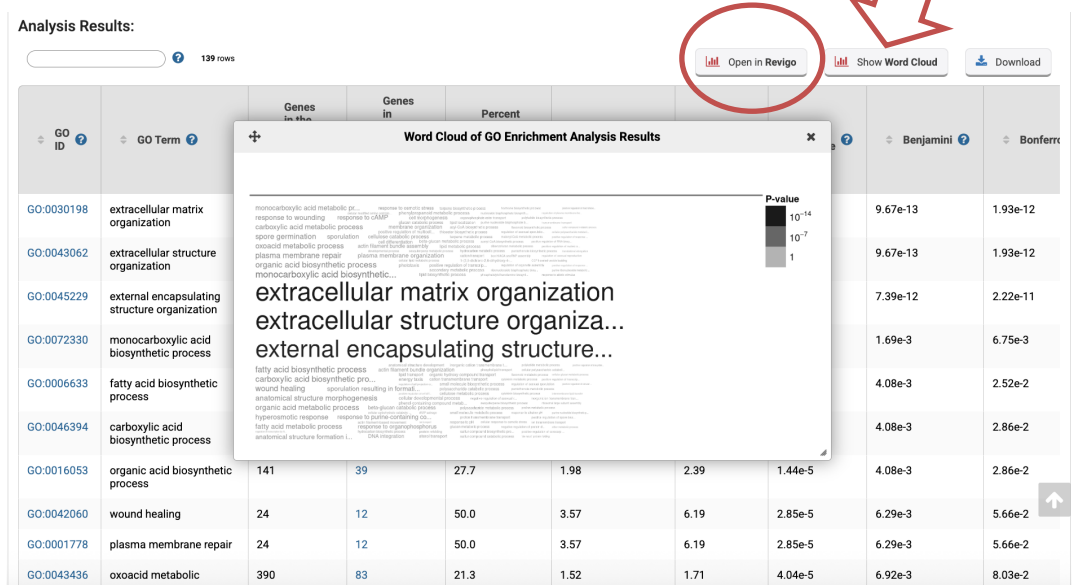
0.1

Get Answer

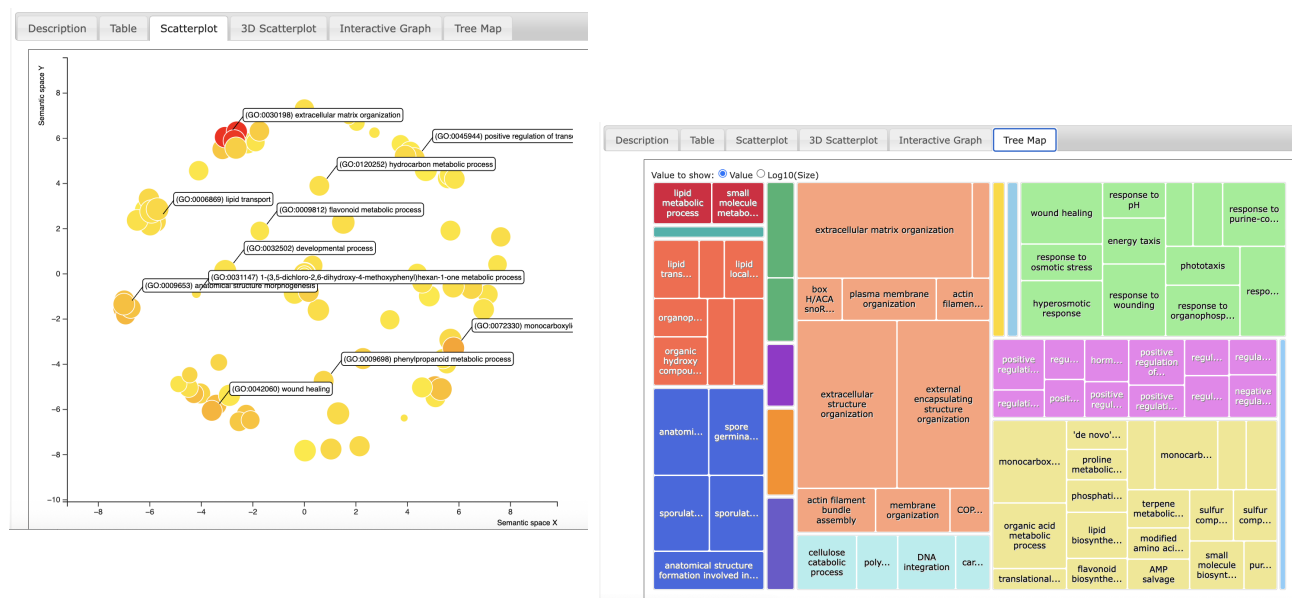
6. Analyze the results for GO Enrichment



7. What functions are enriched in the gene set?



8. Visualize related functions in Revigo. What are the major biological processes at this stage of Dictyostelium development?



9. What happens to the results if you increase the stringency of DE search to 4-fold upregulation and p-value of 0.01? How does this affect the GO enrichment?