PLAZA workshop

Workbench functional data analysis

October 25-26, 2021

# Topics Covered

In this series of exercises, you will learn how to use the PLAZA workbench. This involves performing functional analyzes of gene sets, the comparison of gene sets, as well as data retrieval operations. These data operations contain finding orthologs, gene descriptions and sequences.

# Data Content

In this series of exercises, we will make use of the workbench experiments created in the previous set of exercises: the up- and down-regulated auxin-responsive genes from *Oryza sativa ssp. Japonica*.

In this series of exercises, we will make use of the workbench in the **Monocot 4.5** PLAZA.

# Exercises

## Part 1: Functional Analyses

1. Explore the functional annotation of the up-regulated experiment [Experiment page 🡪Toolbox 🡪 View the functional annotation]
   1. Looking at the GO terms, what percentage of genes is associated with stress response? What percentage is associated with auxin response? [GO-group: Biological process]
   2. Looking at the descriptions of the GO terms, which cells part(s) is/are associated with auxin response? [GO-group: Cellular Component]
   3. Do the descriptions of the associated InterPro domains match those of the GO terms? Can this be explained?
2. View the GO enrichment of the up-regulated experiment [Experiment page 🡪Toolbox 🡪View the GO enrichment]
   1. Do the GO terms with the highest p-value in the ‘Biological process’ category make sense?
   2. What is the first GO term in the ‘Molecular Function’ category that is indicated as being depleted within the data set?
3. View the GO enrichment of the down-regulated experiment [Experiment page 🡪Toolbox 🡪View the GO enrichment]
   1. Reduce the input GO terms by using only the primary GO data. Can you see a difference in the results compared to using all (primary + projected) GO data?
   2. Make the GO enrichment statistical testing more stringent by requiring a more stringent p-value (0.001). Can you see a difference in the results when compared with the default settings?

## Part 2: Data Retrieval

1. Export the gene content for all genes in the up-regulated experiment [Experiment page 🡪Actions 🡪Export]
   1. Export the descriptions for the genes
   2. Export the protein sequences for the genes
   3. Export the InterPro functional annotation for the genes
2. View the orthologs for the up-regulated experiment using the integrative orthologs tool [Experiment page 🡪Toolbox 🡪View the orthologous genes]
   1. Find the orthologs in *Triticum aestivum* using a minimum of 2 evidence types. For how many genes can we NOT find orthologs? [hint: scroll down the page and look for common error message]
   2. Clearly, there is a problem with the number of paralogs in *Triticum aestivum*. Perform the same operation as in (a) but limit the orthologous/paralogous relationship to 1-1.
      1. For how many *Oryza sativa* genes do we find a 1-1 ortholog with 2 evidence types in *Triticum aestivum*?
      2. Is there an orthologous relationship with more than 2 evidence types?
   3. Download all orthologous relationships (for the genes in the workbench experiment) between *Oryza sativa* and *Zea mays*. [Integrative orthology page 🡪View results🡪Download results]
      1. When scrolling down the result-file, which evidence type is clearly overrepresented?

## Part 3: Experiment Comparison

1. Compare the content of the up- and down-regulated workbench experiments [Experiment page 🡪Toolbox 🡪Compare with other PLAZA workbench experiments]
   1. Is the Excel sheet coherent? Are there genes that are (incorrectly) indicated as being both up- and down-regulated? [Compare page 🡪Toolbox 🡪Compare genes]
   2. Are the two experiments functionally different? Are there GO terms common to both experiments? [Compare page 🡪Toolbox 🡪Compare GO terms]
2. Create a new experiment based on the content of an existing experiment [Experiment page 🡪Actions 🡪Copy experiment]
   1. Give the experiment an easily identifiable name
   2. Identify in this experiment the largest group of genes for a family [Experiment page🡪Toolbox🡪View the associated gene families ; Sort by associated genes]
      1. Delete these genes from the copied experiment [Click on associated genes; Toolbox🡪Remove all genes in this subset from this experiment]
      2. Compare the GO enrichment of the truncated experiment with the GO enrichment of the original experiment [Experiment page 🡪Toolbox 🡪Compare with other PLAZA workbench experiments]
         1. Which GO terms that were NOT enriched in the original experiment, have been annotated as being enriched in the truncated experiment?

## Part 4: Extra

1. When working with a co-worker you don’t want to have to synchronize your experiment content all the time
   1. Share the experiment with another workshop participant[Experiment page 🡪Actions 🡪Share experiment]
   2. Try to perform some common actions in the PLAZA workbench using the shared experiment access.