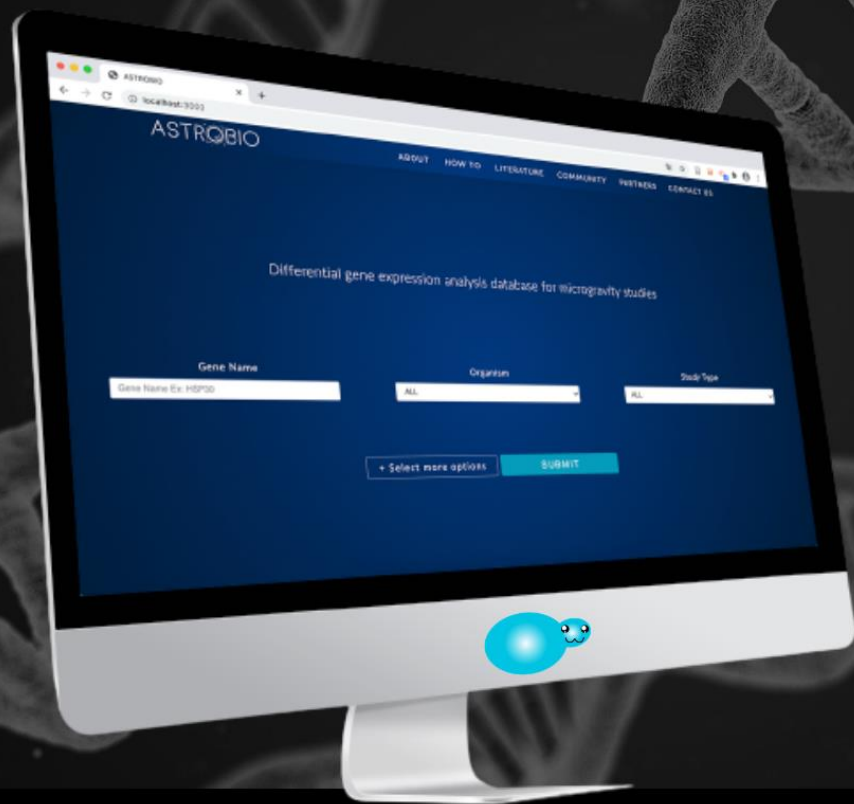


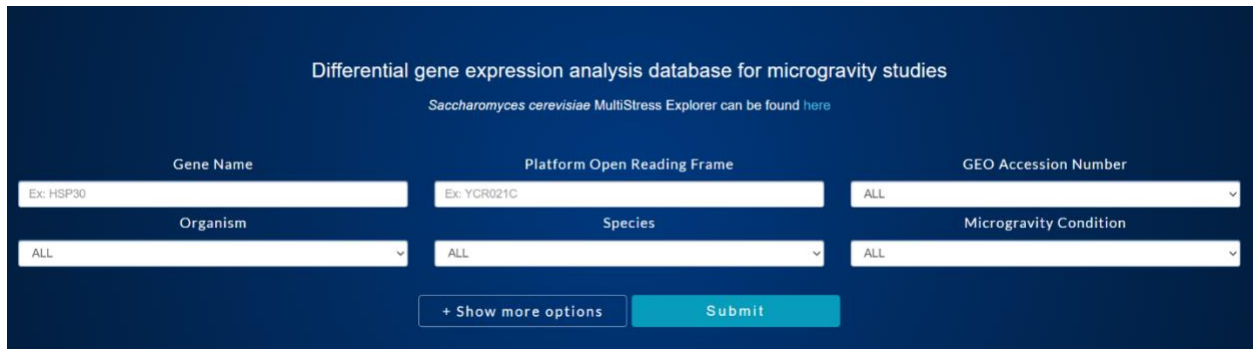
ASTROBIO USER GUIDE



iGEM CONCORDIA

This is a User Guide for AstroBio, an open-source, differential gene expression analysis software and database for microgravity researchers.

Step 1. Enter your desired search options:

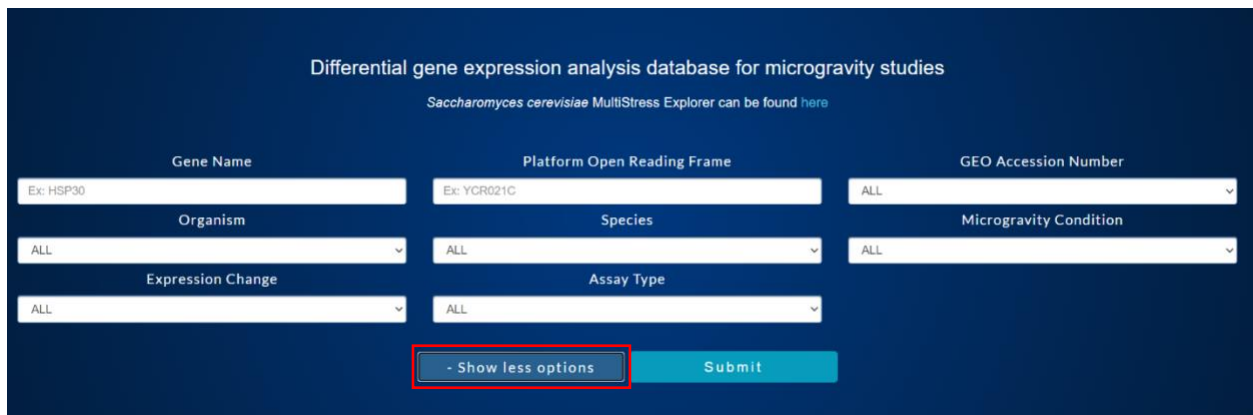


The screenshot shows the AstroBio search interface. At the top, it reads "Differential gene expression analysis database for microgravity studies" and includes a link for "Saccharomyces cerevisiae MultiStress Explorer". Below this are six search fields: "Gene Name" (with example "HSP30"), "Platform Open Reading Frame" (with example "YCR021C"), "GEO Accession Number", "Organism", "Species", and "Microgravity Condition". Each field has a dropdown arrow. At the bottom, there are two buttons: "+ Show more options" and "Submit".

Figure 1: AstroBio Search Parameters

- Gene Name: HSP30
- Platform Open Reading Frame: YCR021C
- GEO Accession Number: GSE4136
- Organism: Yeast
- Species: *Saccharomyces cerevisiae*
- Microgravity Condition: HARV

Step 2. Click on + **Show more options** if you want to refine your search parameters. You can also opt to just - **Show less options**.



This screenshot shows the same AstroBio search interface but with more options visible. In addition to the fields from Figure 1, there are "Expression Change" and "Assay Type" dropdown menus. The "+ Show more options" button is now replaced by a "- Show less options" button, which is highlighted with a red box. The "Submit" button remains.

Figure 2: AstroBio Search Parameters - Options

- Expression Change: Up Regulation
- Assay Type: Microarray

Step 3. Once you are satisfied with your search criteria, click on **Submit**.

The following **Search Results** should be available to you – in this example, we looked for the HSP30 Gene using the Gene Name parameter

SEARCH RESULTS Export/Print

4 results found.

Note: Default results are sorted in ascending order by Adjusted P Value.

Results per page: 10 | Sort by: Adjusted P Value | [Submit](#)

Gene Name	Author(s)	GEO Accession	Microgravity Condition	Expression Change (Log ₂ FC)	Species	Generation	Adjusted P Value	
HSP30	Kate and McInerney	GSE4136	HARV	4.19743717	Saccharomyces cerevisiae	25	0.000321223	More Info
HSP30	Kate and McInerney	GSE4136	HARV	3.976342201	Saccharomyces cerevisiae	5	0.029966079	More Info
HSP30	Sheila and Nielsen	GSE50881	SPACEFLOWN	-0.053092408	Candida albicans	NA	0.03884708	More Info
HSP30	Sheila and Nielsen	GSE50881	SPACEFLOWN	0.074622055	Candida albicans	NA	0.042765309	More Info

Figure 3: AstroBio Search Results

Step 4. If your research results multiple results, you can choose how many studies are displayed by clicking on **Results per page**. You can also sort the results by clicking on **Sort by**. This will allow you to sort studies by *Adjusted P Value*, *P Value*, *Descending Expression Change*, as well as *Ascending Expression Change*.

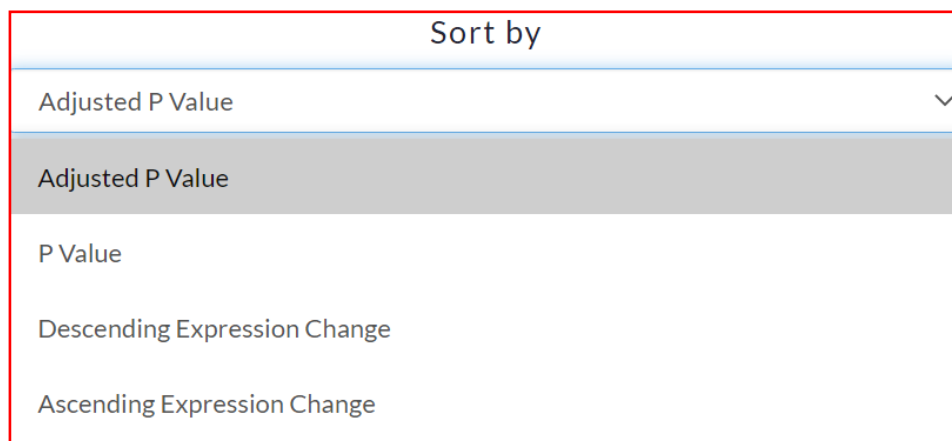


Figure 4: Sort by Options

Step 5. To obtain more information on a particular study, simply click on **More Info**.

ASTROBIO LITERATURE CURRENT PRACTICES HOME PAGE

SEARCH RESULTS Export/Print

4 results found.

Note: Default results are sorted in ascending order by Adjusted P Value.

Results per page: 10 Sort by: Adjusted P Value Submit

Gene Name	Author(s)	GEO Accession	Microgravity Condition	Expression Change (Log ₂ FC)	Species	Generation	Adjusted P Value	
HSP30	Kate and McInerney	GSE4136	HARV	4.197433717	Saccharomyces cerevisiae	25	0.000321223	More Info

Figure 5: More Info Option

The following window will appear:

ASTROBIO

OVERVIEW

Overview

- Gene Ontology
- Gene Location
- Spaceflight Essential Genes
- Study Overview
- Study Information

Gene Name: HSP30
 Name Description: Hsp30p
 ORF: YCR021C
 Log₂FC: 4.197433717
 t value: 7.773290873
 B value: 2.096816683
 Organism: Yeast
 Species: Saccharomyces cerevisiae
 Strain: BY4743
 Generation: 25
 Essential for spaceflight: FALSE
 ID: 1775364_at
 P Value: 0.0000461
 Adjusted P Value: 0.000321223
 GEO Accession Number: GSE4136
 Link to study: [Click here](#)
 Advanced Stress Analysis: [AtstroBio MultiStress Explorer](#)

Figure 6: Overview Tab

Note: Click on **AstroBio MultiStress Explorer** for *Saccharomyces cerevisiae*

AstroBio MultiStress Explorer

Choose a Metastudy: HeatShock

Gene Forest Plot: HSP30

Choose threshold values to download genes and/or display on a pca biplot

Summary Log2-Fold Change	Number of Top Ranking Genes	Sign Consistency
<input type="text"/>	<input type="text"/>	<input type="text"/>

Enter common gene names from *S. cerevisiae* genome. Multiple entries can be entered using ',' as a delimiter with no spaces between and in all caps. Ex : ADR1,UPC2

Plot Specific Gene Values:

Show IDs on PCA-Biplot

Number of Queried Genes:

About | Graphs

Summary

AstroBio MultiStress Explorer is a small webapp intended for exploring results of a series of pairwise comparisons of *Saccharomyces cerevisiae* transcriptomics analyses in simulated microgravity (HARV biosensor) with various other environmental stressors. Prior to the meta-analysis, differential expression analysis was performed using linear mixed models provided by the R package lme4 for each condition. The adjusted p-values for each experiment were combined using a random effects model in addition to summarizing the fold change across conditions. The environmental stresses include:

- Oxidative Stress from 0.19mM Cumene hydroperoxide. RNA was extracted after 6 minutes of treatment. (GSE26169, Sha et al. 2013)
- Heat Shock at 37C for 7 minutes. (GSE132186, Muhlhofer et al. 2019)
- Hyperosmotic Shock from 0.4M of NaCl. RNA was extracted after 6 minutes of treatment.(GSE13097, Romero-Santacruz et al. 2009)

All of the above datasets were then combined both pairwise and pooled with data derived from the condition:

- Simulated low shear modeled microgravity (LSMMG) from growth in a High Aspect Ratio Vessel (HARV). RNA extracted after 5 generations of growth.(GSE4136, Sheehan et al.)

Guide

The MultiStress explorer provides three different plots to review the results of the meta-analysis. The first is an interactive **Volcano Plot** which displays the combined p-values versus the summarized log-fold change in expression for each gene. Colour intensity is derived from sign consistency across conditions, a sign consistency of 2 for instance would indicate that the gene is upregulated across two conditions. Users can hover over the datapoints to view information concerning the gene or click on them to change the forest plot display below.

The **Gene Forest Plot** shows the expression of particular gene across conditions when a name is typed in the window or when clicked on the volcano plot. The input expects a common gene name in most cases or an ORF identifier when none are available. All genes displayed in the volcano plot are valid entries for this view.

The **PCA-Biplot** is the results of a principal component analysis of the log-fold difference in expression prior to meta-analysis. Eigenvectors are plotted with their direction indicative of their similarity to other groups and their length corresponding to the magnitude of variance the group contributes. Individual genes can be plotted and labeled by using the three query variables which select genes on the basis of their inclusion in the random effects model. These query variables are:

- Log2 Summary Fold Change:** the threshold for summary fold change values as shown in the volcano plot. The genes are queried using an *at* heuristic of selection. This means that an input of 1 will select all genes at one or above (≥ 1) while an input of -1 will select all genes at -1 or below (≤ -1).
- Number of Top Ranking Genes:** Genes are ranked from smallest to largest based on their summary p-values. An input of 10 for example would select the ten top-most genes in the volcano plot.
- Sign Consistency:** The direction of expression across stress conditions are combined into a single metric by using the sign of the fold change value. A gene which is upregulated in one study contributes +1 to the sum and vice versa. The formal definition is:

$$\sum_{i=1}^n \frac{\log FC_i}{|\log FC_i|}$$
 where n is the number of studies and $\log FC$ is the log2 fold change for a particular gene.

Plot Specific Gene Values: A gene name or a series of gene names can be entered to be shown on the biplot. If multiple entries are desired then the user must use a comma as separator with no spaces included. Gene names are expected in all-caps and are not present in the biplot data if one or more name is not found in a study.

The **Heatmap** is also generated from the result of the query input variables showing the intensity of log2 Fold-Change from each dataset. Genes are only drawn if the gene is found in every study.

The data which generates the PCA-Biplot and the heatmap can be downloaded in csv format. All datapoints are log2 Fold-Change values.

Figure 7: AstroBio MultiStress Explorer for *Saccharomyces cerevisiae*

Step 6. Browse through the different tabs to obtain more information.

Overview | **Gene Ontology** | Gene Location | Spaceflight Essential Genes | Study Overview | Study Information

GENE ONTOLOGY

Molecular Function

- ion channel activity,molecular_function

Biological Processes

- cellular response to DNA damage stimulus
- cellular response to ethanol
- cellular response to heat
- cellular response to hydrogen peroxide
- cellular response to osmotic stress
- ion transport
- negative regulation of ATPase activity

Cellular Component

- endoplasmic reticulum
- integral component of membrane
- integral component of membrane
- membrane
- plasma membrane
- plasma membrane

Figure 8: Gene Ontology Tab

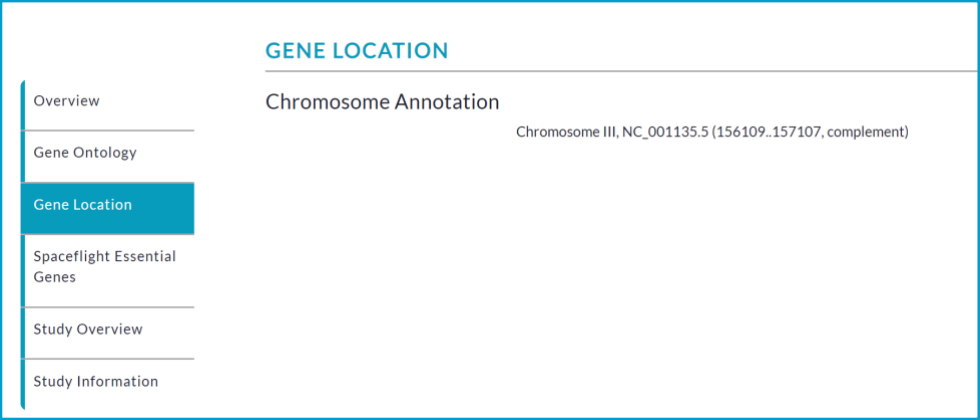


Figure 9: Gene Location Tab

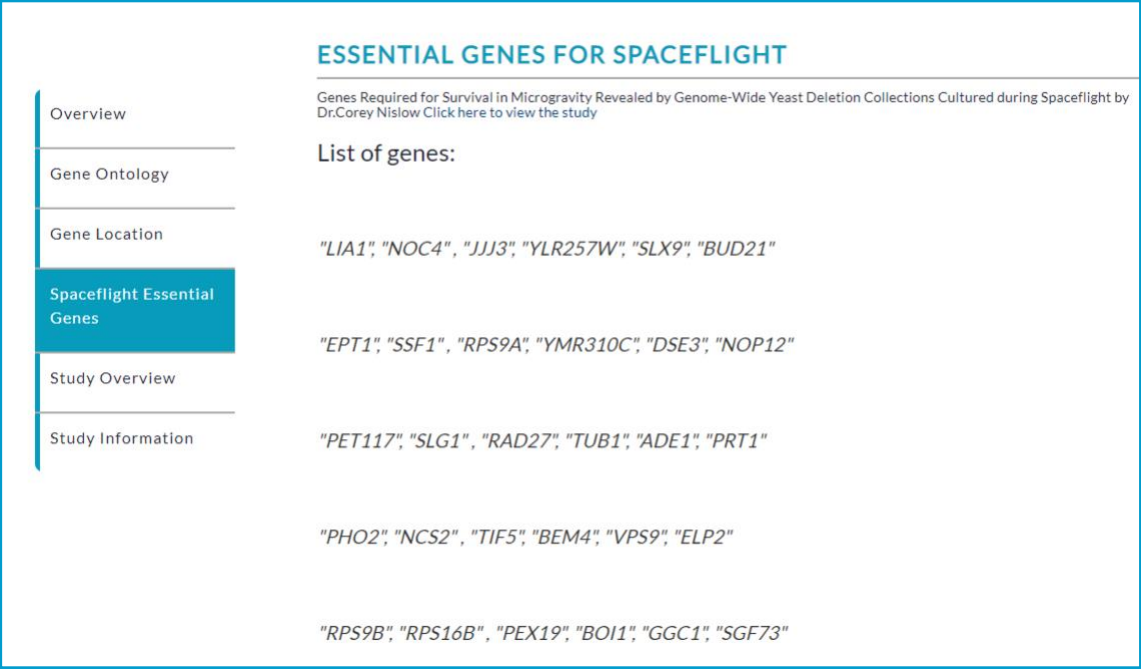


Figure 10: Spaceflight Essential Genes

STUDY OVERVIEW		
Overview	Author(s)	Kate and McInerney
Gene Ontology	Study Type	HARV
Gene Location	Assay Type	Microarray
Spaceflight Essential Genes	PMID	17201921
Study Overview	Institute	Montana State University
Study Information	Platform ID	GPL2529 GSM94610 GSM94611 GSM94612 GSM94616 GSM94617 GSM94618
	Samples	
	Treatments	Control sample 25th generation_rep1 Control sample 25th generation_rep2 Control sample 25th generation_rep3 Low shear modeled microgravity sample 25th generation_rep1 Low shear modeled microgravity sample 25th generation_rep2 Low shear modeled microgravity sample 25th generation_rep3

Figure 11: Study Overview Tab

STUDY		
Overview	Study Title	Yeast Genomic Expression Patterns in Response to Low-Shear Modeled Microgravity
Gene Ontology	Study Type	Expression profiling by array
Gene Location	Design	Four conditions are compared with three replicates each: yeast grown in low-shear modeled microgravity (HARV bioreactor) for 5 and 25 generations; yeast grown in a horizontal (non-LSMMG) HARV bioreactor for 5 and 25 generations
Spaceflight Essential Genes	Summary	The goal of this study was to assess whether low shear-modeled microgravity (LSMMG) effects yeast genomic expression patterns using the powerful tool of whole genome microarray hybridization. We determined changes in the yeast model organism, <i>Saccharomyces cerevisiae</i> , when grown in LSMMG using the rotating High Aspect Ratio Vessel (HARV). A significant number of genes were up- or down-regulated by at least two fold in cells that were grown for 5 generations or 25 generations in HARVs. We identified genes in cell wall integrity signaling pathways containing MAP kinase cascades that may provide clues to novel physiological responses of eukaryotic cells to the external stress of a low-shear modeled microgravity environment. A comparison of the microgravity response to other environmental stress response (ESR) genes showed that 26% of the genes that respond significantly to LSMMG are involved in a general environmental stress response, while 74% of the genes may represent a unique transcriptional response to microgravity. In addition, we found changes in genes involved in budding, cell polarity establishment, and cell separation that confirm our hypothesis that exposure to LSMMG causes changes in gene transcription resulting in a phenotypic response. The results of the study provide interesting clues to potential mechanisms involved in the response to, adaptation to, and survival of eukaryotic cells in a microgravity environment and our findings may have important health implications for human spaceflight.
Study Overview	Link	Click here
Study Information		

Figure 7: Study Information Tab

To access the study directly, go to the Link provided in the **Study Information** tab.

ASTROBIO

iGEM CONCORDIA

igem.concordia@gmail.com