Purdue Biomedical Engineering

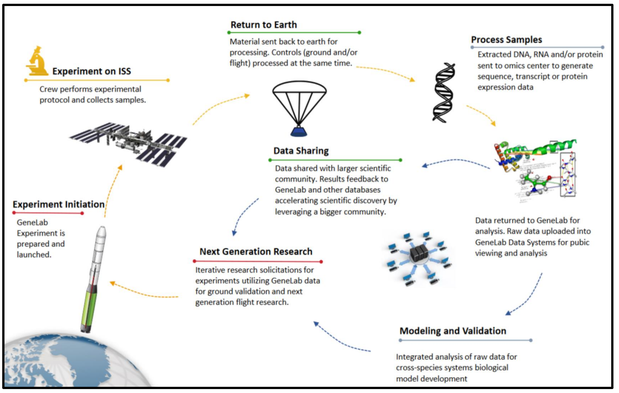
PURDUE UNIVERSITY | The Data Mine

Corporate Partnership between The Data Mine and GeneLab

GeneLab — Open Science for Life in Space

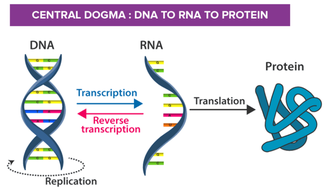
Introduction and Motivation

NASA GeneLab Data Repository



Why Gene Data?

- Genes are the fundamental building blocks of life.
- Genes -> mRNA -> Proteins



Why experiments in space matter?

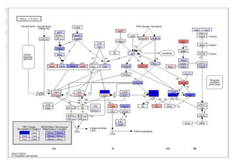
- Long-term space exploration and Habitation
- Diseases are accelerated in space. Studying their molecular underpinnings improves life on earth.
- Innovation on Earth: Challenging fundamental assumptions opens a new perspective.

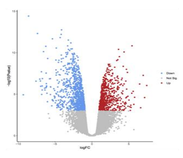
Microgravity Radiation Genetics Analytics

Research Methodology

Pre-Processing (Spring 2023)
Fastq -> Normalized counts

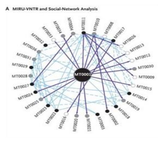
Pathway Analysis (Fall 2022)





Data Processing (Fall 2022)

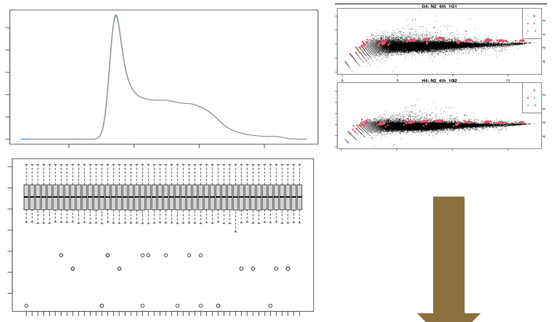
Meta Analysis & Network Analysis (Spring 2023)



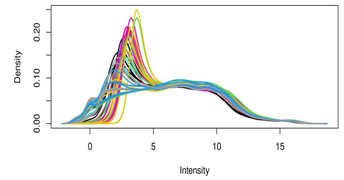
- Relevant datasets of samples in space and those in ground control, 1G or vivarium control are compared.
- Gene expression counts are then processed and normalized. The statistically significant differences are visualized.
- The differences are then mapped to relevant biological pathways in the organism and it confirms or disproves the hypothesis
- Finally, multiple differences clubbed together present the overall picture for conclusion

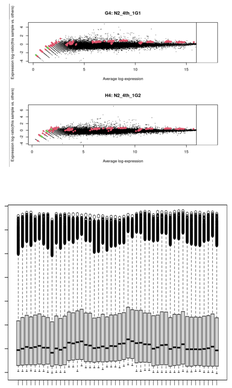
Data Visualization

Normalized Data



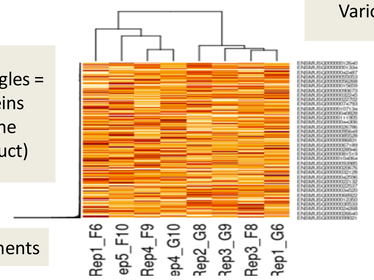
Raw Data



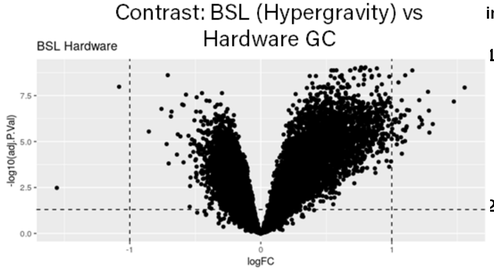


Differential Gene Expression

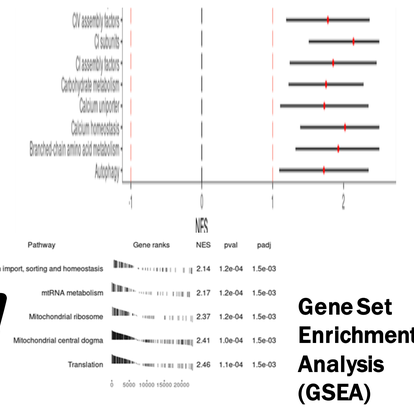
Heatmap



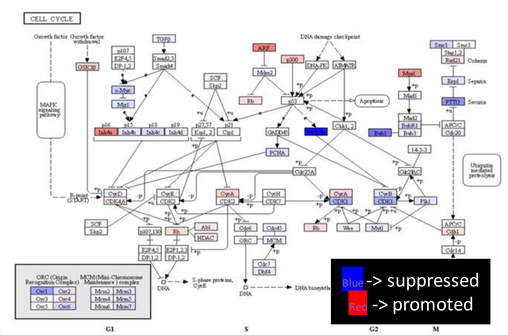
Volcano Plots



Pathway Analysis



Gene Set Enrichment Analysis (GSEA)



KEGG Pathway Analysis

Project Recap

What Did We Learn?

- ✓ Relevant tools to draw relevant conclusions from unprocessed data
- ✓ Different methods of visualizing differential gene expression data
- ✓ Statistical and pathway analysis for differential gene expression in R
- ✓ Understanding of the mechanism of obtaining gene expression data from biological samples

Future Goals

- ➢ Continue to process new data being added to the GeneLab repository
- ➢ Formulate new hypotheses with the relevant data for deeper understanding
- ➢ Development of professional relationships with GeneLab project groups

References and Acknowledgement

We'd like to thank our partners at the Bioinformatics Lab (Universtiy of Nevada, Reno); Dr. Marhsall Porterfield and Dr. Marshall Tabetah (Purdue University) and the Purdue Data Mine team for their inputs during the project

1. Overbey, E. G., Saravia-Butler, A. M., Zhang, Z., Rathi, K. S., Fogle, H., da Silveira, W. A., ... & Galazka, J. M. (2021). NASA GeneLab RNA-seq consensus pipeline: standardized processing of short-read RNA-seq data. *Iscience*, *24*(4), 102361.
2. Thompson, T., Gibbs, K., Rask, J., Coughlan, J., & Smith, J. (2014). *NASA GeneLab Concept of Operations* (No. ARC-E-DAA-TN18494).
3. https://genelab.nasa.gov/about