Mini Project: Transcriptomics (RNA-Seq Differential Expression Analysis)

Author: Itunuoluwa Isewon PhD

Email: itunu.isewon@covenantuniversity.edu.ng

Introduction

In this mini-project, participants will perform differential gene expression analysis using RNA-seq data from the NCBI Gene Expression Omnibus (GEO) dataset GSE292521. Raw counts will be provided in a CSV format along with sample group information. This exercise will allow participants to gain hands-on experience with transcriptomics data analysis.

Summary of GSE292521 Experiment

Title: *Genomics and Transcriptomics of 3ANX (NX-2) and NX (NX-3) producing isolates of Fusarium graminearum*

Organism: *Fusarium graminearum*

Scope: Analysis of 20 fungal isolates from different regions in Manitoba, characterizing both genomic variations and gene expression profiles linked to mycotoxin chemotypes (3ANX and NX). The data illuminate differential expression patterns related to pathogenicity and suggest the 3ANX chemotype may be more widespread in Canada than previously recognised.

Platform: Illumina NovaSeq 6000 sequencing.

Objectives

By the end of this mini-project, participants will be able to:

- Import RNA-seq count data and metadata into R.
- Conduct differential expression analysis using DESeq2.
- Visualize results using MA plots, heatmaps, and volcano plots.
- Interpret biological significance of differentially expressed genes.

Dataset Information

Dataset: GSE292521 (Available on GEO:
https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE292521)
Provided Files:
- Raw counts in CSV format
- Metadata (sample groupings, e.g., control vs treatment)

Tasks

Participants are expected to carry out the following steps:

1. Load count data and metadata into R.
2. Normalize data using DESeq2.
3. Conduct differential expression analysis.
4. Generate visualization plots (volcano plot, heatmap).

Author: Itunuoluwa Isewon PhD   Email: itunu.isewon@covenantuniversity.edu.ng

5. Identify top differentially expressed genes and their biological roles.

Required R Packages

The following R packages will be needed for this project:

- DESeq2 (for differential expression analysis)
- pheatmap (for heatmaps)
- EnhancedVolcano (for volcano plots)
- ggplot2 (for visualization)
- dplyr (for data manipulation)

Expected Output

By completing this project, participants should produce:

- A heatmap of top differentially expressed genes.
- A volcano plot highlighting significant genes.
- A list of significantly upregulated and downregulated genes.
- A Group presentation showing the rationale, all plots, the top differentially expressed genes, their associated enriched KEGG pathways and Link findings to mycology/mycotoxicology relevance (e.g., stress response, secondary metabolite genes, toxin biosynthesis genes if present)

Author: Itunuoluwa Isewon PhD   Email: itunu.isewon@covenantuniversity.edu.ng