Author: Prince Mishra

TCGA BRCA Shiny App

An interactive R Shiny web application to explore **gene expression** and **biological insights** in **breast cancer (BRCA)** using **TCGA (The Cancer Genome Atlas)** RNA-Seq data.

This app allows users to:

- View expression of a selected gene in **Tumor vs Normal** samples (Boxplot)

- Explore **differentially expressed genes (DEGs)** using a Volcano Plot

- Understand affected **biological processes** (GO Enrichment)

- Explore **disrupted pathways** in cancer (KEGG Enrichment)

Project Objective

To build a simple web app that:

- Visualizes and compares gene expression between cancer and normal tissue

- Identifies important genes and biological mechanisms involved in BRCA

- Makes bioinformatics data accessible to **non-coders and biologists**

Background

Cancer changes how genes behave. Some genes become overactive, while others get turned off. Understanding these changes helps in:

- Finding biomarkers for diagnosis

- Discovering targets for therapy.

- Learning how cancer grows or spreads

But analyzing RNA-Seq data requires coding. This app simplifies that using an intuitive web interface powered by R Shiny.

Materials and Methods

Data Source

-TCGA BRCA RNA-Seq gene expression data (Tumor + Normal samples)

- Preprocessed into:

  - `count_matrix_with_symbols_clean.csv`: expression counts

  - `DEG_results_mapped.csv`: DEGs with gene names (SYMBOLs)

Samples Used

This analysis was performed on:

- 20 Tumor samples (primary tumor)

- 20 Normal samples (solid tissue normal)

All samples were selected from the TCGA BRCA dataset using `TCGAbiolinks`.

For exact sample IDs and patient info, refer to:

- `data/raw/metadata.csv`

Tools Used

- `DESeq2`: For identifying DEGs (Differentially Expressed Genes)

- `EnhancedVolcano`: To plot DEGs visually

- `clusterProfiler`: To perform GO and KEGG enrichment

- `org.Hs.eg.db`: Human gene ID mapping

- `Shiny`: To build the web interface

App Features

1. Boxplot

-Visualize **expression levels** of any gene in Tumor vs Normal samples.

2. Volcano Plot

-Displays all DEGs. Helps identify which genes are most different between cancer and normal tissues.

3. GO Enrichment

-Shows which biological processes (e.g., "cell cycle", "immune response") are enriched among the DEGs.

4. KEGG Pathway Enrichment

-Highlights disrupted biological pathways (e.g., "p53 signaling", "DNA repair").