

# **Deciphering the Gene Regulatory Networks Controlling Somatic Embryogenesis in *Coffea arabica***

## **Description**

This project involves a Gene Coexpression Network (GCN) analysis to investigate the molecular mechanisms regulating somatic embryogenesis (SE) in *Coffea arabica*. The analysis aims to identify significantly correlated gene modules and uncover hub genes potentially involved in SE regulation.

## **Objectives**

- To analyse unprocessed expression data across different developmental stages of SE
- To construct a GRN and identify key gene modules associated with SE
- To identify hub genes and transcription factors involved in the regulation of SE
- To functionally characterize key genes and modules related to SE

## **Data**

- Source: The transcriptome data covering somatic embryo developmental stages of the *Coffea arabica* GPFA116 intraspecific hybrid variety were retrieved from the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) under accession PRJNA744419.
- Type: RNA-seq data (Illumina paired-end)

## **Tools & Technologies**

- Languages: R, Bash
- Libraries / Packages:
  - WGCNA (R) – for gene coexpression network (GCN) construction
  - Cufflinks, Cuffmerge, Cuffnorm, Cuffdiff – for transcript assembly and differential expression analysis
  - ggplot2, ComplexHeatmap, igraph – for visualization and network representation
- Software / Platforms:
  - SRA Toolkit – for data retrieval
  - RStudio – for statistical computing and visualization

## References

- Awada, R., Lepelley, M. et al. (2023). Global transcriptome profiling reveals differential regulatory, metabolic and hormonal networks during somatic embryogenesis in *Coffea arabica*. *BMC Genomics*, 24(1), 41.
- Langfelder, P., & Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*, 9(1), 559.