**Supplementary File:**

To run the three-step analysis, the following codes and parameters were used:

Here, where it says “Transcriptomic” is replaced by the name of the RNA-Seq file downloaded from NCBI.

1. **Fastp**

```
fastp --detect_adapter_for_pe --overrepresentation_analysis --correction \
    --cut_mean_quality 30 --cut_right --cut_front --cut_tail --thread 10 \
    -i Transcriptomic_1.fastq.gz \
    -I Transcriptomic_2.fastq.gz \
    -o Transcriptomic-fastp_R1.fq.gz \
    -O Transcriptomic-fastp_R2.fq.gz \
```

2. **Star**

```
ulimit -n 102400;STAR --runThreadN 10 --outSAMtype BAM SortedByCoordinate
--twopassMode Basic \
--genomeDir Elaeis_guineensis_genome \
--outFileNamePrefix Transcriptomic \
--readFilesCommand zcat \
--readFilesIn Transcriptomic-fastp_R1.fq.gz Transcriptomic-fastp_R2.fq.gz
## indexing bam
samtools index TranscriptomicAligned.sortedByCoord.out.bam
```

3. **HTSeq-Count**

```
htseq-count -n 10 -r pos --mode=union --format=bam --type=CDS \
 --idattr=protein_id --stranded=no \
Transcriptomic_Aligned.sortedByCoord.out.bam  \
Elaeis_guineensis_genomic.gff > Transcriptomic_count_table.txt
```

## Genie3

**Score Definition**: The Genie3 score represents the **relative importance** of each **regulatory gene** for the expression of a **target gene**. This score is derived from the contribution of the predictor variables (regulatory genes) in the division of the nodes of the Random Forest model trees.

The score is calculated by following these steps:

• For each target gene, a Random Forest model is trained, where the other genes are used as predictor variables.

• Each tree in the Random Forest divides the data into nodes, using the regulatory genes as explanatory variables.

• The importance of each regulatory gene is measured by the reduction in error (or impurity) in each division of the nodes along the trees.

• The contribution of each regulatory gene in all trees of the model is accumulated to generate a final score.

• The final score is normalized to reflect the relative influence of each regulatory gene on the target gene.

The higher the score, the greater the evidence that the regulatory gene influences the expression of the target gene, in the **context of specific tree**.

## Parameters used in Genie3

For the inference of gene networks, we used Genie3 with the following parameters:

• **tree-method**: Random Forest (RF);

• **expr_data**: matrix containing gene expression values. Each row corresponds to a condition and each column corresponds to a gene;

• **gene_names**: list containing the names of the genes;

• **regulators**: list containing the names of the candidate regulators;

• **K**: Defines the number of attributes selected in each node of a tree (all)

• **ntrees**: number of trees (1000 - Default)

• **nthreads**: number of threads used for parallel computing (10).

**The main code is as follows:**

```
# Genie3 is obtained from: https://github.com/vahuynh/GENIE3
from Genie3 import *
import pandas as pd
import numpy as np

data = pd.read_csv("Carvalho da Silva et al_Table_S4.csv")
data.set_index("id", inplace = True)
data = data.T

gene_names = list(data.columns)
data_dende = np.array(data)

reg = pd.read_excel("Carvalho da Silva et al_Table_S2.xlsx", skiprows = 2)
reg = reg[reg['Description'].str.contains('Present')]
reg_list = reg['Regulator'].to_list()

g3 = GENIE3(data_dende,
            gene_names = gene_names,
            regulators = reg_list,
            K = 'all',
            nthreads = 10)
get_link_list(g3, gene_names = gene_names, file_name = "EG_Net.txt")
```