

## **BioProject: Differentially regulated genes in TOR knockdown Arabidopsis lines**

Rapamycin-sensitive transgenic Arabidopsis lines (BP12) expressing yeast FK506 Binding Protein12 (FKBP12) were developed. Inhibition of TOR in BP12 plants by rapamycin resulted in slower overall root, leaf and shoot growth and development leading to poor nutrient uptake and light energy utilization. Genetic and physiological studies together with RNA-Seq and metabolite analysis of TOR-suppressed lines revealed that TOR regulates development and lifespan in Arabidopsis by restructuring cell growth, carbon and nitrogen metabolism, gene expression, ribosomal RNA and protein synthesis. Overall design: Arabidopsis WT (Col) and BP12-2 (TOR knockdown line) seedlings at 15 DAG were treated with rapamycin for 3 days by transferring from 0.5 MS medium to 0.5 MS+10 µg/ml rapamycin. Triplicate samples of rapamycin treated WT and BP12-2 seedlings were used for RNA-Seq analysis (Illumina Hiseq 2000)

### **Dados**

Os arquivos de *reads* estão no diretório:

/data/homedir/userXX/libs\_normalizadas

São seis arquivos no formato fastq, sendo as bibliotecas SRR634969, SRR634970, SRR634971 referentes à “BP12-2”, e as bibliotecas SRR634972, SRR634973, SRR634974 são referente à “WT”.

### **1 – Montagem *ab initio***

#### **Objetivos:**

1. Nesta etapa vamos fazer a análise sem o auxílio de uma referência genômica ou de transcritos. Por mais que o organismo que estamos trabalhando (*Arabidopsis thaliana*) esteja com o genoma e os genes bem definidos, esta etapa tem conteúdo didático. Vale lembrar que a montagem *ab initio* só é indicada em casos onde não há uma referência

disponível, visto que ela tende a gerar transcritos “picados” e de tamanho menor que os originais.

2. Calcular a expressão dos transcritos da montagem.
3. Identificar os transcritos diferencialmente expressos.

## **Pipeline**

1. Antes de iniciar essa montagem apague a pasta toda da montagem\_guiada por questão de espaço.
2. Criar diretório com o nome “montagem\_abinitio” no seu home.
3. Entrar dentro do diretório “montagem\_abinitio” e realizar a montagem de todas as bibliotecas ao mesmo tempo utilizando o software Trinity.

## **Dicas e Sugestões**

- Parâmetros sugeridos:

`--max_memory 50G`

- Essa montagem vai levar mais ou menos 20 horas para terminar, então rode o comando utilizando “nohup” e “&”.

4. Calcular a expressão dos transcritos da montagem utilizando o software kallisto

## **Dicas e Sugestões**

- Primeiramente crie o índice da referência (arquivo fasta com a montagem dos transcritos).
- Não se esqueça de mudar o nome do diretório de saída para cada biblioteca.
- Rode o kallisto utilizando os comandos Linux “nohup” e “&”. Ex “nohup comando &”. Dessa forma o terminal é liberado e o

processo não fica preso a sua sessão, assim é possível fechar o terminal sem que o processo seja perdido.

5. Criar a tabela com os read counts utilizando o script “roda\_cat.sh”.

### **Dicas e Sugestões**

- Você deve passar para esse script os arquivos de expressão de cada biblioteca. Ex:

```
rodaCat.sh SRR634969/abundance.tsv SRR634970/abundance.tsv  
SRR634971/abundance.tsv SRR634972/abundance.tsv SRR634973/abundance.tsv  
SRR634974/abundance.tsv > read_counts.txt
```

- Ele irá criar no seu diretório atual um arquivo chamado “read\_counts.txt” que tem o formato tabular, sendo a primeira coluna o nome do transcrito e as demais colunas os reads counts para cada uma das bibliotecas.
- A ordem das bibliotecas no arquivo de saída é a mesma dos argumentos.

6. Identificar os genes diferencialmente expressos utilizando o pacote em R “DESeq2”.

### **Dicas e Sugestões**

- Fazer a análise de expressão no computador local.
- Para copiar o arquivo readcounts.txt para a máquina local, baixe e utilize o programa WinSCP:

(<https://winscp.net/download/winscp574.zip>)

- Primeiramente fazer um PCA e/ou HEATMAP das bibliotecas
- Seguir o manual do DESeq2 para fazer a análise e criar um arquivo no formato csv com os resultados.