

# Analysis of differential gene expression in wild and cultivated rice under drought stress

Ingo Giebel

Math.-Nat. Fakultät, Heinrich-Heine-Universität Düsseldorf

QBio304: Applied Bioinformatics

Prof. Dr. Björn Usadel

Dr. Jędrzej Jakub Szymanski

May 1, 2023

## Contents

<b>1</b>	<b>Abstract</b>	<b>2</b>
<b>2</b>	<b>Introduction</b>	<b>2</b>
2.1	Background . . . . .	2
2.2	Objectives . . . . .	2
<b>3</b>	<b>Materials and methods</b>	<b>2</b>
3.1	Selection of publicly available plant RNA-seq data . . . . .	2
3.2	Quality evaluation . . . . .	2
3.3	Mapping to the respective genome . . . . .	2
3.4	Statistical evaluation and differential expression analysis . . . . .	2
3.5	Functional enrichment analysis . . . . .	3
<b>4</b>	<b>Results</b>	<b>3</b>
4.1	Quality evaluation . . . . .	3
4.2	Mapping efficiency and coverage . . . . .	3
4.3	Exploratory data analysis . . . . .	3
4.4	Differentially expressed genes . . . . .	3
4.5	Functional enrichment analysis . . . . .	3
<b>5</b>	<b>Discussion</b>	<b>3</b>
5.1	Critical evaluation of the results . . . . .	3
5.2	Biological implications . . . . .	3
5.3	Limitations and future directions . . . . .	4
<b>6</b>	<b>Conclusion</b>	<b>4</b>

## **1 Abstract**

Provide a brief summary of the purpose of the assignment, the methods used, the main findings, and the significance of the results. Limit the abstract to 200-250 words.

## **2 Introduction**

### **2.1 Background**

Introduce and explain the study and give a rationale for the RNA-seq analysis. Discuss why of RNA sequencing in understanding gene expression and regulation in plants and why it is used in this study.

### **2.2 Objectives**

State the specific aims of the assignment, which include obtaining plant RNA-seq data, evaluating its quality, mapping to a respective genome, performing statistical evaluation, differential expression analysis, functional enrichment analysis, and critically evaluating and discussing the results.

## **3 Materials and methods**

### **3.1 Selection of publicly available plant RNA-seq data**

Explain the criteria for selecting the data and the source of the published study.

### **3.2 Quality evaluation**

Describe the tools and methods used to assess the quality of the selected RNA-seq data.

### **3.3 Mapping to the respective genome**

Detail the reference genome used and the bioinformatics tools employed for the mapping process.

### **3.4 Statistical evaluation and differential expression analysis**

Explain the statistical methods and software used for evaluating the data and identifying differentially expressed genes.

### **3.5 Functional enrichment analysis**

Describe the tools and databases used to perform functional enrichment analysis to interpret the biological significance of the differentially expressed genes.

## **4 Results**

### **4.1 Quality evaluation**

Present the findings from the quality evaluation of the selected RNA-seq data.

### **4.2 Mapping efficiency and coverage**

Report the results of the mapping process, including the mapping efficiency and coverage.

### **4.3 Exploratory data analysis**

Discuss the dominating variance components, reproducibility, possible batch effects and confounding variables.

### **4.4 Differentially expressed genes**

Discuss the identified differentially expressed genes and their potential biological significance.

### **4.5 Functional enrichment analysis**

Present the results of the functional enrichment analysis, highlighting the enriched functional categories.

## **5 Discussion**

### **5.1 Critical evaluation of the results**

Discuss the quality and reliability of the RNA-seq data and the downstream analyses.

### **5.2 Biological implications**

Discuss the potential implications of the findings for plant biology and the broader scientific community.

### **5.3 Limitations and future directions**

Address the limitations of the current analysis and suggest possible future directions to expand on the findings.

## **6 Conclusion**

Summarize the main findings of the assignment, reiterating the significance of the results, and provide a final statement on the overall outcome of the study.