# GENETICS AND GENOMICS STUDIES FOR WATER-RELATED STRESSES IN RICE (ORYZA SATIVA L.)

A Dissertation

by

YU-YA LIANG

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

| | |
|---|---|
| Chair of Committee, | Endang M. Septiningsih |
| Committee Members, | Alan Dabney |
| | Lee Tarpley |
| | Michael Thomson |
| Head of Department, | David D. Baltensperger |

August 2019

Major Subject: Plant Breeding

# ABSTRACT

Rice is one of the most important crops that feeds more than half of the world's population. Along with the escalating problems with climate change, drought and submergence events have increased in recent years and have challenged the sustainability of rice production. This dissertation focused on these two water-related stresses, and studied the tolerance mechanisms underlying.

Reproductive stage is the most water-sensitive period for rice cultivation, therefore the study was conducted to investigate the molecular and physiological responses of two rice genotypes, 'Rondo' and '4610', to drought stress during reproductive stage under field conditions. Leaf samples were collected for RNA-Seq. Additionally, 10 agronomic traits and chlorophyll fluorescence were measured. The results showed that 4610 had better performance than Rondo under moderate drought stress conditions. The results indicated that 4610 had more stress responsive GO terms and several known genes families related to drought stress were also identified in 4610 up-regulated DEGs, including the LEA proteins, HSPs, APXs, and GSTs.

The second goal was to characterize the region of *qSub8.1*, a new submergence tolerance QTL during vegetative stage, which will enable us to develop suitable DNA markers and identify the gene(s) underlying *qSub8.1* for further functional characterization. Three rice varieties, Ciherang-Sub1, Ciherang, and IR64-Sub1, were sequenced with 150 bp pair-end WGS. The results showed that Ciherang-Sub1 genome is composed of 59% Ciherang, 24% of IR64-Sub1, and 17% of unknown sources; and

the *qSub8* region is mainly from Ciherang with a few introgressed segments from IR64-Sub1 and unknown sources.

There are challenges in conducting plant transformation and regeneration required to introduce the CRISPR reagents into the plant cell for gene-editing. Here, we used mature seeds as explants and reported a high-efficiency transformation and regeneration protocol for a recalcitrant *indica* rice cultivar Ciherang-Sub1 using particle bombardment to deliver CRISPR/Cas9 gene-editing vector. Our protocol successfully gives an optimal condition for shoot regeneration with 95% of regeneration rate and by using this protocol, CRISPR gene-edited plants can be generated and validated within approximately 12 weeks. Using this protocol, CRISPR gene-edited plants can be generated and validated within 12 weeks.

# DEDICATION

For my mother.

# ACKNOWLEDGEMENTS

I am grateful for everyone who has helped me in my struggle to achieve my dream of earning a Ph.D. degree. First, I would first like to thank my committee chair, Dr. Endang M. Septiningsih for her continual support and guidance of my graduate studies with endless patience. My sincere thanks to my committee members, Dr. Alan Dabney, Dr. Lee Tarpley, Dr. Michael Thomson, and Dr. Shichen Wang, who have given much time, effort and knowledge to aid in the completion of this dissertation.

Many others have been instrumental in this process. I would like to acknowledge their role in the completion of my work. I appreciate the generous assistance of Dr. Rodante Tabien for supporting the research in Beaumont Center; Chersty Harper, Patrick Carre, and Kyle Jones for managing the field and harvesting plant materials; Dr. Abdul Mohammed for teaching me all the physiological measurement. I would also like to thank Dr. John Cason and Michael Baring for teaching me everything about peanut breeding although I did not include peanut research in this dissertation.

The faculty, staff, and students in the Soil and Crop Sciences Department and the Crop Genome Editing Lab also deserve recognition: LeAnn Hague, Taylor L. Atkinson, and Amada Ray for helping me deal with document processing and unexpected situations as an international student. Dr. Ana Barrientos Velazquez and Dr. Sakiko Okumoto for expanding my knowledge in soil science and plant physiology. Dr. Backki Kim for teaching me everything about CRISPR and for being a great example as an excellent breeding scientist. Oneida Ibarra, Dr. Nithya Subramanian, Sudip Biswas,

Sruja Iyer, Ranjita Thapa, Benjamin Rogers, Stephon Warren, Dr. Nancy Wahl, Mark Brooks, Cooper Svajda, Wardah K. Mustahsan, Sejuti Mondal, and A.S.M. Faridul Islam for the sharing of their knowledge and expertise, camaraderie and rapport. I would also like to thank the Soil and Crop Sciences Department for all the support not only for my graduate studies but also for many student activities such as travel funding, SOGO, and the Plant Breeding Symposium.

Many thanks go to my fellow students and friends, Yu-Ming Lin, Yi-Pin Lai, Szu-Ting Kuo, Yu-Lyu Yeh, Wei-Ju Chen, Naining Chi, Ruei-Ping Chang, Hsin-Yi Li, Shu-Hui Chuang, Shen-Yu Hu, Long Chen, Ang Xu, Yuanyuan Luan, Ted Chang, Jean Hsu, Yong-Yu Jhan, Yi-Hua Lee, Morries Cheng, Justine Chu, Saul Perez, Shirley Arbizu, Tabby Liu, Jorge Valenzuela Antelo, Diana Zapata, Steve Anderson, Drutdaman Bhangu, Cynthia Sias, Blake Young, Martin Carlos Costa, Heather Baldi, Ammani Kyanam, and all the members in Herman Heepers softball team, Average Jose's soccer team, TSA basketball team and TSA badminton team, who have provided friendship and support, and with whom I have shared laughter, frustration, and companionship.

Thanks also go to Guosong Wang, Dr. Fu-Jin Wei, Hung-Yu Dai, Wei-Ting Chen, Peter Wu, James Chen, Chi Kao, Jen-Hau Yeh, Sheng-Kai Hsu, Candice Chu, and Nathália Penna Cruzato who shared their knowledge with me without hesitation.

Finally, I cannot begin to express my gratitude to my family for all of the love and support they have sent my way along this journey. To my parents, your unconditional love and support have meant the world to me, I hope that I have made you proud. To my little brother, thanks for always being by my side.

# CONTRIBUTORS AND FUNDING SOURCES

## Contributors

This work was supervised by a dissertation committee consisting of Professors Endang M. Septiningsih, Lee Tarpley, and Michael Thomson of the Department of Soil and Crop Sciences, Professor Alan Dabney of the Department of Statistics, and special appointment committee Dr. Shichen Wang of Texas A&M Genomics and Bioinformatics Service.

The field management and sample harvesting for Chapter 2 were assisted by Dr. Rodante Tabien, Chersty Harper, Patrick Carre, and Kyle Jones; and the chlorophyll fluorescence measurement was collected by Dr. Lee Tarpley and Dr. Abdul Mohammed of Texas A&M AgriLife Research Center at Beaumont station. The genomic data analyzed for Chapter 3 was guided by Dr. Shichen Wang. The construction of CRISPR/Cas9 vector in Chapter 4 was guided by Dr. Backki Kim.

All other work conducted for the dissertation was completed by the student independently.

## Funding Sources

Graduate study was supported by a graduate research assistantship from Texas A&M University and a dissertation research project from Texas A&M AgriLife Research.

This work was also made possible in part by USDA NIFA under Grant Number # 2017-67013-26194. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the USDA NIFA.

# NOMENCLATURE

| | |
|---|---|
| ABA | Abscisic Acid |
| APXs | Ascorbate Peroxidases |
| BP | Biological Process |
| bZIP | basic leucine Zipper motif |
| Cas9 | CRISPR associated protein 9 |
| CAT | Catalase |
| CC | Cellular Component |
| CDPK | Calcium-Dependent Protein Kinase |
| CIM | Callus Induction Medium |
| CRISPR | Clustered Regularly Interspaced Short Palindromic Repeats |
| CTAB | Cetyl Trimethylammonium Bromide |
| DEGs | Differentially Expressed Genes |
| DREB | Drought Responsive Element Binding proteins |
| EDTA | Ethylenediaminetetraacetic Acid |
| ERF | Ethylene-Response Factor |
| EtOH | Ethanol |
| FDR | False Discovery Rate |
| gRNA | Guide RNA |
| GO | Gene Ontology |
| GR | Glutathione Reductase |

| | |
|---|---|
| GSTs | Glutathione S-transferases |
| HR | Homologous Recombination |
| HSPs | Heat Shock Proteins |
| $K^+$ | Potassium |
| KEGG | Kyoto Encyclopedia of Genes and Genomes |
| LEA | Late Embryogenesis Abundant |
| LRT | Likelihood Ratio Test |
| MF | Molecular Function |
| MH63 | Minghui 63 |
| MSM | Selection Medium |
| NAA | 1-Naphthaleneacetic Acid |
| NAC | NAM, ATAF, and CUC |
| NGS | Next-Generation Sequencing |
| NHEJ | Non-Homologous End-Joining |
| OSM | high-Osmotic Medium |
| PR | Pathogenesis-related |
| qPCR | Quantitative Polymerase Chain Reaction |
| QTL | Quantitative Trait Locus |
| RCBD | Randomized Complete Block Design |
| RNA-Seq | RNA Sequencing |
| ROS | Reactive Oxygen Species |
| R1 | Shooting Medium (Regeneration 1) |

| | |
|---|---|
| R2 | Rooting Medium (Regeneration 2) |
| SNPs | Single-Nucleotide Polymorphisms |
| SOD | Superoxide Dismutase |
| SSR | Simple Sequence Repeat |
| Tris | Tris(hydroxymethyl)aminomethane |
| WGS | Whole Genome Shotgun Sequencing |
| 2,4-D | 2,4-dichlorophenoxyacetic Acid |
| 6BA | 6-Benzylaminopurine Solution |

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

# CHAPTER I

## INTRODUCTION

Rice (*Oryza sativa* L.) is the third largest crop produced in 2018, with 495.87 million tons of total production worldwide (https://www.statista.com/statistics/263977/world-grain-production-by-type/). Rice is also one of the most important staple food that feeds 3.5 billion, more than half of the world's population (http://ricepedia.org/rice-as-food/the-global-staple-rice-consumers). Along with the increase of global population which is expected to reach to 9.1 billion by 2050 (http://www.fao.org/fileadmin/templates/wsfs/docs/expert_paper/How_to_Feed_the_World_in_2050.pdf); the demand of increased rice production is also pressing. Hence, in order to meet this increasing demand, different strategies must be applied to increase rice production, including the development of improved high-yielding rice cultivars that are more resilience under climate change scenarios.

More than 70% of rice production are in Asian countries, including China, India, Indonesia, Bangladesh, Vietnam, Philippines, Thailand, Burma, and Japan (https://www.statista.com/statistics/255971/top-countries-based-on-rice-consumption-2012-2013/). Brazil and the United States are the 10$^{th}$ and 13$^{th}$ largest rice produced countries, which account for 1.6% and 0.9% of total rice production in 2018, respectively. (https://www.statista.com/statistics/255971/top-countries-based-on-rice-consumption-2012-2013/). The United States, however, is the fifth largest rice exporter, which account for more than 6% of global exports

(https://www.ers.usda.gov/topics/crops/rice/rice-sector-at-a-glance/). All rice in the United States is grown in the irrigated fields in four main rice production regions, Arkansas Grand Prairie, Mississippi Delta (parts of Arkansas, Mississippi, Missouri, and Louisiana), Gulf Coast (Texas and Southwest Louisiana), and California (Sacramento Valley). Arkansas is the major rice-producing state, which account for 48% of rice production in the United States, followed by California (19%), Louisiana (14%), Missouri (8%), Texas (7%), and Mississippi (4%). More than 70% of rice production in the US is long-grain, followed by 26% of medium-grain, and 1% of short-grain. All the states are mainly producing long-grain types, except for California, with 92% of medium-grain and 7% of short-grain (https://www.nass.usda.gov/Publications/Todays_Reports/reports/cropan19.pdf). The rice planting season in the Gulf Coast region begins at the end of March or early April, followed by the Delta regions in April, and California in May. Some of the fields in Gulf Coast region are able to be ratooned for the second harvest before winter comes.

Among major abiotic stresses that affect rice production are cold, heat, salinity, flooding, and drought. Cold stress decreases seed germination, plant growth and also affects yield (Cruz, Milach, & Federizzi, 2006; Farrell, Fox, Williams, Fukai, & Lewin, 2006). Heat stress, especially high nighttime temperature, has negative impacts on membrane stability, pollen viability, spikelet fertility, and ultimately yield (Mohammed & Tarpley, 2009; Peng et al., 2004). Salinity is a major problem for rice growing particularly in coastal regions such as Bangladesh which can damage plants and cause low yield and poor grain quality (Haque, 2006). Among the abiotic stresses, drought

stress is the most problematic in rice production; about 50% of rice production worldwide is affected by this stress alone (Bouman, Peng, Castañeda, & Visperas, 2005). However, drought is not the only water-related stress for rice production. The submergence due to rapid heavy rain and severe floods are also destructive to the plant growth and yield (Ismail, 2018; Septiningsih & Mackill, 2018). Along with the escalating problems from climate change, the extreme weather events have also increased globally in recent years causing more severe drought and flooding.

Rice is a diploid crop (2n = 2x = 24) with a genome size of 379-392 Mb (Sakai et al., 2013; Zhang et al., 2016). The International Rice Genome Sequencing Project (IRGSP), started to decode Nipponbare, a japonica rice variety, in 1999 (Sasaki & Burr, 2000). Thereafter, as the first completly sequenced crop, rice genetic and genomic researches are blooming in the past two decades. There are more than a thousand accessions of *Oryza* species genomic sequences now available in the public database (https://www.ncbi.nlm.nih.gov/genome/?term=Oryza). The *Oryza* database provides powerful resources for further genetic and genomic researches in rice and the grasses family. Therefore, we are able to perform more genetics and genomic studies efficiently such as performing whole-genome RNA-Sequencing analysis with gene annotation, conducting referenced-guided genome assembly, and search cloned gene sequences for gene-editing.

The aims of this dissertation were to dissect the genetics and genomics profiles of two water-related stresses in rice, drought and submergence, and to investigate the tolerance mechanism underlying these stresses. In chapter 2, RNA-Seq analysis was

performed to obtain the genome-wide expression profile. Along with the measurement of phenotypic traits, investigated the drought tolerance molecular mechanism of two rice genotypes. In chapter 3, a whole genome shotgun sequencing and a reference-guided *de novo* genome assembly were performed in order to fully understand the genome profile and genome-wide nucleotide information of a submergence tolerance cultivar, Ciherang-Sub1. In chapter 4, a transformation and regeneration protocol was optimized for Ciherang-Sub1 using mature seeds as explants and using a biolistic delivery of CRISPR/Cas9 construct.

CHAPTER II

# TRANSCRIPTOMIC ANALYSIS OF TWO RICE GENOTYPES UNDER DROUGHT CONDITIONS DURING REPRODUCTIVE STAGE

## Introduction

Rice (*Oryza sativa* L.) is one of the most important staple foods that feeds 3.5 billion, more than a half of the world's population (Ricepedia, http://ricepedia.org/rice-as-food/the-global-staple-rice-consumers). Global population is projected to reach 9.1 billion by 2050 (http://www.fao.org/fileadmin/templates/wsfs/docs/expert_paper/How_to_Feed_the_World_in_2050.pdf); therefore, the increased of rice production is a necessity to keep pace with the rising global demand for food. However, rice production is challenged by various factors, including abiotic and biotic stresses. Drought stress is one of the main obstacles for crop production, especially for rice; 50% of the rice production is affected worldwide (Bouman et al., 2005). The damages caused by drought stress includes limiting nutrients uptake, cell dehydration, and the production of excessive metabolic compounds of reactive oxygen species (ROS) (Choudhury, Rivero, Blumwald, & Mittler, 2017; Price, Atherton, & Hendry, 1989). A simple cell dehydration may damage cells, including membrane dysfunction, unfolded enzyme, and cytoskeletal damage. Likewise, ROS may cause protein and membrane damages, and damage at the DNA level as well (Mittler, 2002; Tenhaken, 2014).

The drought tolerance mechanisms are complex and this abiotic stress can occur in different developing stages of rice plant growth (Tripathy, Zhang, Robin, Nguyen, & Nguyen, 2000), with the reproductive stage being the most susceptible (Garrity & O'Toole, 1994). Drought stress during reproductive stage can reduce yield dramatically. Drought tolerance mechanism mainly can be divided into two categories, i.e. preventing water loss and preventing damage caused by dehydration (Farooq, Wahid, Kobayashi, Fujita, & Basra, 2009; McDowell et al., 2008). For water loss prevention, ABA regulation (Munemasa et al., 2015), K+ concentration and membrane potential (Corratge-Faillie et al., 2017) can regulate stomatal conductance to reduce water loss (Chaves, Pinheiro, & Flexas, 2008). Previous studies also revealed that over-expression of drought-induced genes such as calcium-dependent protein kinase (CDPK), NAM, ATAF, and CUC (NAC) transcription factor, and drought responsive element binding proteins (DREBs) transcription factor contribute to drought tolerance in rice (Chen, Meng, Zhang, Xia, & Wang, 2008; Hu et al., 2006; Saijo, Hata, Kyozuka, Shimamoto, & Izui, 2000). Hence, to understand the molecular mechanisms to drought during reproductive stage in rice is important to ensure good survival during this stress period and good harvest.

The aims of this study were to evaluate agronomic traits under drought stress and to investigate the drought tolerance molecular mechanism during rice reproductive stage. To reach these goals, agronomic traits measurement and RNA-Seq analysis were performed. RNA-seq is a next-generation sequencing (NGS) based technique, providing genome-wide expression profile, which allow us to dissect the whole genome picture of

how plants respond to drought. Two rice genotypes, Rondo, a susceptible variety, and 4610, a moderately drought tolerant breeding line, were selected for the study. My results showed that a total of 1040 genes were differentially expressed in 4610, with 555 genes up-regulated and 485 genes down-regulated, and 856 genes were differentially expressed in Rondo, with 460 genes up-regulated and 396 genes down-regulated. The differentially expressed genes (DEGs) including late embryogenesis abundant (LEA) proteins, heat shock proteins (HSP), and some known-stress related genes, which are involved in several stress response Gene Ontology (GO) categories.

## Materials and Methods

### *Plant Materials and Drought Treatment*

Two rice genotypes with different performance under drought stress, 'Rondo' and '4610', were selected for this study. Rondo is an elite indica rice cultivar, regularly evaluated in Texas and has high yield and resistance to various diseases (Yan & McClung, 2010), but it is susceptible to drought stress. On the other hand, 4610 is a breeding line, from Texas A&M AgriLife Beaumont Research Center (Beaumont station), that was previously reported as less affected by drought (Tabien, Harper, Carre, & Jones, 2015). A field experiment was performed in Beaumont station in 2016 and 2017. The two genotypes were directly sowed at the rice experimental field in Beaumont, Texas, using randomized complete block design (RCBD) with three replications in both drought and irrigated (control) conditions. All replicates were

planted in three-row blocks, each block has 6 m length and 18 cm between rows. Drought treatment was applied after more than 50% plants flowering by draining the field completely.

*Measurement of Chlorophyll Fluorescence and the Agronomic Characters*

The plants for sampling were randomly selected from each plot and the third leaf from the top on the main stem was collected to measure the chlorophyll fluorescence. The capability of reaction center ($F_v/F_m$), and quantum yield (Y), which reflects the efficiency of PSII, were assessed by measuring fluorescence with a pulse-modulated fluorometer (OS5P; Opti-Sciences, Hud- son, NH, USA). For $F_v/F_m$ measurements, the leaves were dark adapted for 30 min (Mohammed, Cothren, Chen, & Tarpley, 2015). The chlorophyll fluorescence was measured 14 days after field being drained.

A total of 10 agronomic traits were measured in the current study. They were filled grain number per panicle, unfilled grain number per panicle, spikelet fertility, filled grain weight per plot, unfilled grain weight per plot, total grain weight per plot, plant dry weight, hundred-seed weight, panicle length, and yield. Yield was calculated in kilogram per hectare, spikelet fertility was calculated in percentage (%) and the rest of the units were measured in gram (g) or centimeter (cm).

Agronomic data were collected from the same plants as for the chlorophyll fluorescence measurement. The ten best panicles from each plot were used to count the grain number including filled and unfilled grain number respectively. Spikelet fertility was calculated from filled and unfilled grain number on each of the selected panicle. All

panicles in each plot were manually threshed and manually separated into the filled and unfilled grain. Filled and unfilled grain data was used to calculate filled grain weight, unfilled grain weight, and total grain weight per plot. Plant dry weight was measured for the above ground part of the plant after air-drying for 7 days according to Beaumont Center Rice Breeding Lab protocol. Filled grain number, unfilled grain number, and spikelet fertility were the averages of 10 panicles from each plot. Hundred-seed weight was the average of three times measurement of 100 randomly selected filled grains. Panicle length was measured as average from 20 panicles in each plot. Yield was multiplied from total grain weight per plot. In order to test the differences, panicle length, filled grain number, unfilled grain number, and spikelet fertility were examined by t-test with $\alpha = 0.05$.

### *RNA Extraction and Sequencing*

The second leaves, counted from the top, on the main stem were collected from the same plant as for the chlorophyll fluorescence measurement at the 14 days after field was drained. Leaf samples were wrapped in aluminum foil, plunged into liquid nitrogen immediately and stored in -80°C freezer before extraction. RNA was extracted by using TRIzol reagent followed by QIAgen RNeasy Plant Mini Kit. RNA quality was determined by 28S to 18S ribosomal RNA ratio by using Agilent 2100 Bioanalyzer. Samples with ratio within range of 1.8-2.2 were used for the library preparation. TruSeq Stranded RNA-Seq libraries were prepped at Texas A&M AgriLife Genomics and Bioinformatics Service (TxGen) as per Standard Operating Protocol (Illumina). Libraries

was run on multiple lanes of an Illumina HiSeq 4000 to provide at least 25 million reads (75 nt pair-end) per sample.

### *Data Processing and Statistical Analysis*

In total, twelve cDNA libraries were made for sequencing including three biological replicates for each treatment. Trimmomatic version 0.36 (Bolger, Lohse, & Usadel, 2014) PE –Phred 33 command was used for quality control using the following steps: (1) raw sequencing reads were trimmed to remove adaptors; (2) low quality bases with quality score less than 20 on the ends and tails of reads were removed; (3) scan the read with a 5 bp sliding window, and remove them when the average quality per bp drops below 20; (4) reads below the 25 bases long were dropped; (5) reads without correspondence read pairs were dropped.

*O. sativa* spp. japonica genome (IRGSP-1.0), Nipponbare, was used as a reference in this study. The reference genome sequence FASTA file and gene annotation GFF file was downloaded from EnsemblPlants (http://plants.ensembl.org). HISAT2 version 2.1.0 (Kim, Langmead, & Salzberg, 2015) was used to align the reads to the reference genome sequence. Thereafter, StringTie v1.3.4d (Pertea et al., 2015) was used to assemble the transcripts within the regions and to obtain the gene counts. Differential expression analysis was performed in R software package DESeq2 v1.20 (Love, Huber, & Anders, 2014) with reads normalizing and variance stabilizing transformation to account for library size and sequencing depth differences. In order to determine: (1) whether the genes express differently under drought condition, and (2) whether gene

expression under drought condition was different between the two genotypes; experimental data was tested in Wald test and likelihood ratio test (LRT), respectively. Internally, P-values was adjusted for multiple testing using the Benjamini-Hochberg method in DESeq2 package. Genes with false discovery rate (FDR) adjusted p-value (padj) < 0.05 were identified as DEGs.

### *GO Enrichment and KEGG Pathway Analysis*

GO enrichment was performed online using PANTHER v14.1 (Mi et al., 2019) and KEGG pathway analysis was performed using *clusterProfiler* package in R (Yu, Wang, Han, & He, 2012). In PANTHER v14.1, Fisher's test was used for main statistical analysis and KEGG enrichment was calculated based on hypergeometric distribution. Both analyses were performed with FDR adjusted by Benjamini-Hochberg method internally. GO terms with both P-value < 0.05 and FDR < 0.05 was identified as significant. KEGG pathway with both P-value < 0.05 and FDR < 0.1 was identified as significant.

## Results

### *Agronomic Characters and Chlorophyll Florescence of Two Rice Genotypes*

Agronomic characters were collected from a total of 12 plots, covering 3 replications for each genotype under each condition. Under irrigated condition, the panicle length of Rondo was longer than 4610. 4610 had higher hundred-seed weight than Rondo (Table 1). Under drought treatment, most of agronomic characters were

reduced in both 4610 and Rondo with the exceptions of hundred-seed weight of Rondo and panicle length of 4610 which were the same in two conditions (Table 2). Filled grain number and unfilled grain number of 4610 were increased 1.9% and 47.2% under drought treatment. Filled grain number, spikelet fertility, and hundred-seed weight were significantly different between 4610 and Rondo under drought treatment. Under irrigated condition, Rondo had higher filled grain number than 4610. However, under drought treatment, 4610 turned out to be having higher filled grain number than Rondo. Under drought treatment, 4610 had significantly higher spikelet fertility than Rondo whereas there was no significant difference between the two genotypes under irrigated condition.

Chlorophyll fluorescence was measured with two parameters $F_v/F_m$ and quantum yield. $F_v/F_m$ represents the maximum photochemical capacity of Photo-System II photochemistry and quantum yield indicates the number of fluorescent events for each photon absorbed. The ratio of $F_v/F_m$ can be used to estimate the maximum quantum yield. Under irrigated condition, 4610 had 0.73 of $F_v/F_m$ and quantum yield of 0.37, and Rondo had 0.75 of $F_v/F_m$ and quantum yield of 0.22. Under drought treatment, 4610 had 0.71 of $F_v/F_m$ and quantum yield of 0.26, and Rondo had 0.71 of $F_v/F_m$ and quantum yield of 0.26 (Table 3). Both 4610 and Rondo had reduced values of $F_v/F_m$ under drought treatment. However, the quantum yield was increased in Rondo under drought treatment.

**Table 1. Effect of variety on yield-related traits under drought treatment.**

| | Cultivar | |
|---|---|---|
| | 4610 | Rondo |
| (1) Filled grain number/ panicle* | 69.5 | 58.9 |
| (2) Unfilled grain number/ panicle | 31.2 | 30.5 |
| (3) Spikelet fertility (%)* | 69.5 | 65.9 |
| (4) Filled grain weight (g)/ plot | 55.3 | 43.7 |
| (5) Unfilled grain weight (g)/ plot | 4.1 | 4.6 |
| (6) Total grain weight (g)/ plot | 59.5 | 48.3 |
| (7) Plant dry weight (g) | 126.6 | 117.3 |
| (8) Hundred-seed weight (g)** | 2.9 | 2.6 |
| (9) Panicle length (cm) | 19.5 | 19.5 |
| (10) Yield (kg/ha) | 9850.04 | 9035.18 |

[a)] Traits with "*" and "**" indicates that the difference between the two cultivars is significant by $p < 0.05$ and $p < 0.001$, respectively.

**Table 2. Effect of variety on yield-related traits under irrigated treatment.**

| | Cultivar | |
|---|---|---|
| | 4610 | Rondo |
| (1) Filled grain number/ panicle* | 68.2 | 80.1 |
| (2) Unfilled grain number/ panicle* | 21.2 | 27.5 |
| (3) Spikelet fertility (%) | 76.9 | 74.3 |
| (4) Filled grain weight (g)/ plot | 77.3 | 63.8 |
| (5) Unfilled grain weight (g)/ plot | 5.7 | 6.9 |
| (6) Total grain weight (g)/ plot | 83.0 | 70.7 |
| (7) Plant dry weight (g) | 200.76 | 163.9 |
| (8) Hundred-seed weight (g)** | 2.9 | 2.7 |
| (9) Panicle length (cm)** | 19.5 | 21.6 |
| (10) Yield (kg/ha) | 10666 | 10351 |

[a)] Traits with "*" and "**" indicates that the difference between the two cultivars is significant by $p < 0.05$ and $p < 0.001$, respectively.

**Table 3. Chlorophyll fluorescence under irrigated and drought conditions.**

| Genotype | Condition | Chlorophyll fluorescence | |
|---|---|---|---|
| | | $F_v/F_m$ | Quantum yield (Y) |
| Rondo | Irrigated | 0.75 | 0.22 |
| Rondo | Drought | 0.71 | 0.26 |
| 4610 | Irrigated | 0.73 | 0.37 |
| 4610 | Drought | 0.71 | 0.26 |

*Characterization of Drought Stress Responsive DEGs*

The expression levels of 46046 annotated rice genes were subjected to differential expression test in DESeq2 package in R. By comparing irrigated control and drought treatment of the two genotypes, 1040 and 856 genes were identified to be differentially expressed under drought stress for 4610 and Rondo with FDR less than 0.05, respectively (Fig. 1). Among all DEGs, 4610 and Rondo had 555 and 460 up-regulated genes, and 485 and 396 down-regulated genes, respectively (Fig. 2). While for the two genotypes, there were 67 up-regulated and 33 down-regulated DEGs in common. The commonly up-regulated genes including many known drought responsive genes including bZIP transcription factor (OS01G0658900), LEA proteins (OS01G0705200, OS11G0454200), HSPs (OS02G0232000, OS03G0277300), ROS scavenger ascorbate peroxidase (OS04G0434800), and other drought tolerance proteins such as the rare cold-inducible 2 family protein (RCI2, OS03G0286900) and the drought and salt stresses responses 1 (OsDSSR1, OS09G0109600). On the other hand, the commonly down-regulated genes included the regulation of grain size (*GS5*, OS05G0158500), branching, photosynthesis, and grain weight (*MIR8*, OS01G0322700), cytokinin-activating (*LOG1*, OS01G0588900), and cytokinin signaling (*CKT1*, OS02G0738400).

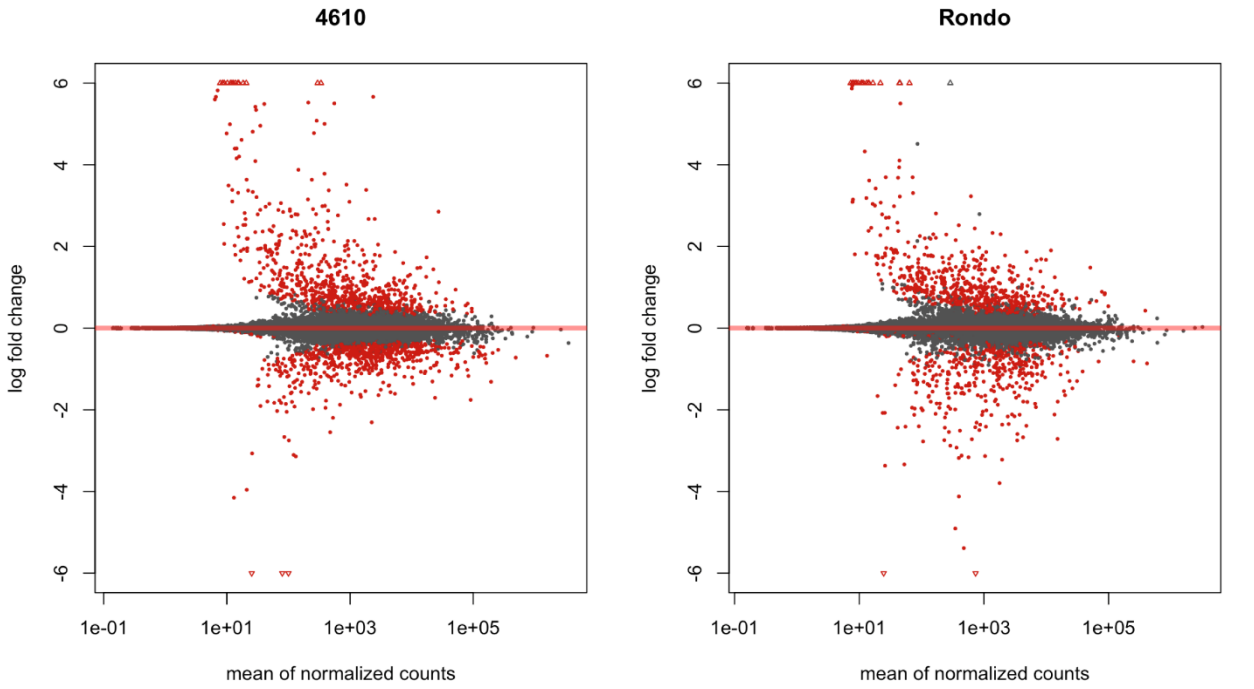


**Figure 1. Differentially expressed genes (DEGs) in two genotypes. Genes with false discovery rate (FDR) adjusted p-value (padj) < 0.05 were identified as DEGs.**

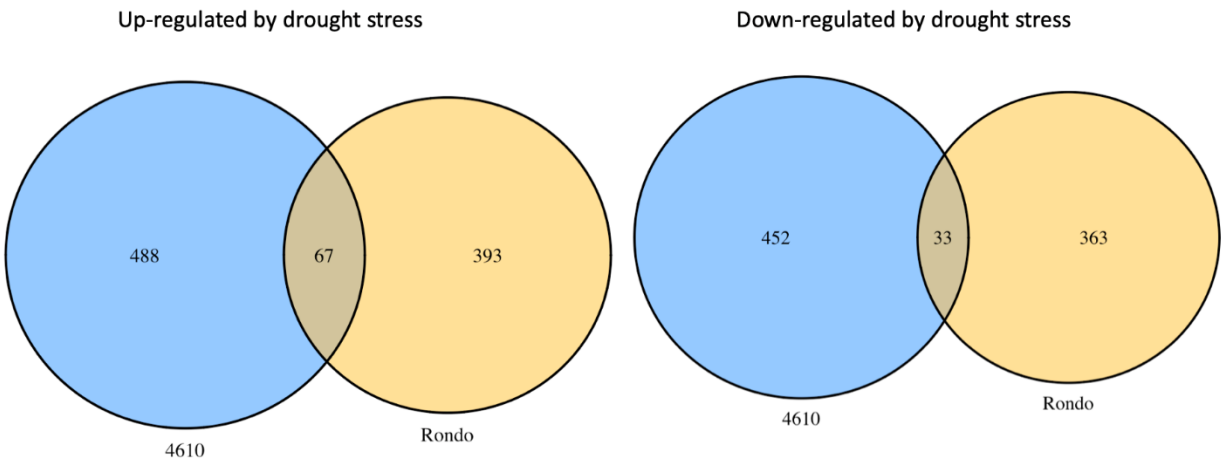


**Figure 2. Overlapping among DEGs regulated by stress between moderate-tolerant genotype 4610 and susceptible genotype Rondo.**

*GO Enrichment Analysis*

To explore the common and specific responses to drought, GO enrichment analyses of all the DEGs were performed, and a total of 127 gene categories were significantly enriched. 4610 had 36 up-regulated categories including 25 Biological Process (BP) and 11 Molecular Function (MF), and Rondo had 37 up-regulated categories including 11 BP, 21 MF, and 5 Cellular Component (CC). For down-regulated categories, 4610 had a total of 24 categories with 7 BP, 12 MF, and 5 CC; whereas Rondo had 10 BP, 18 MF, and 2 CC. The enrichment analysis was not significant for both up-regulated and down-regulated common DEGs. Notably, 4610 had many DEGs categorized as biological processes, including response to stress, response to oxidative stress, response to stimulus, response to ROS, glutathione metabolic process (Table 4). On the contrary, Rondo had no significant enrichment for these stress responsive categories. Among the enriched biological process categories of Rondo were signal transduction, regulation of biological quality, and ion homeostasis. DEGs involved in the stress responsive GO terms in Table 3A including APXs, Dehydrin family proteins, heat stress transcription factors, and HSPs.

Down-regulated DEGs also enriched in distinct categories between 4610 and Rondo (Table 5). 4610 had DEGs involved in photosynthesis, cellular component assembly and metabolic process; whereas Rondo had DEGs enriched in different stress responses and a hormone-mediated signaling pathway. Among the photosynthesis genes of 4610, one was Oxygen-evolving enhancer protein (OS07G0544800); while the Chlorophyll a-b binding protein (OS02G0764500) appeared in both 4610 and Rondo. The DEGs of Rondo enriched in the hormone-mediated signaling pathway, including the ABA receptor (OS05G0213500) and four pathogenesis-related (PR) genes. The same five DEGs in hormone-mediated signaling pathway were also enriched in response to abscisic acid of Rondo (Table 3B).

**Table 4. Gene ontology (GO) of biological process classification of up-regulated DEGs between 4610 and Rondo.**

| Biological process | GO ID | 4610 | FDR | Rondo | FDR |
|---|---|---|---|---|---|
| Response to stress | GO:0006950 | 24 | 8.27E-08 | 6 | ns |
| Response to oxidative stress | GO:0006979 | 10 | 9.86E-07 | 1 | ns |
| Response to osmotic stress | GO:0006970 | 7 | 1.03E-04 | 1 | ns |
| Response to stimulus | GO:0050896 | 29 | 2.23E-05 | 14 | ns |
| Response to reactive oxygen species | GO:0000302 | 10 | 1.70E-09 | 1 | ns |
| Response to cold | GO:0009409 | 4 | 7.28E-03 | 2 | ns |
| Response to drug | GO:0042493 | 7 | 1.72E-05 | 0 | ns |
| Response to heat | GO:0009408 | 14 | 2.83E-08 | 3 | ns |
| Response to toxic substance | GO:0009636 | 11 | 7.22E-04 | 2 | ns |
| Fatty acid catabolic process | GO:0009062 | 4 | 1.59E-02 | 0 | ns |
| Protein folding | GO:0006457 | 17 | 4.96E-07 | 3 | ns |
| Glutathione metabolic process | GO:0006749 | 6 | 2.00E-02 | 0 | ns |
| Organic substance catabolic process | GO:1901575 | 11 | 1.38E-03 | 4 | ns |

**Table 5. Gene ontology (GO) of biological process classification of down-regulated DEGs between 4610 and Rondo.**

| Biological process | GO ID | 4610 | FDR | Rondo | FDR |
|---|---|---|---|---|---|
| Photosynthesis | GO:0015979 | 5 | 2.14E-02 | 2 | ns |
| Organic substance biosynthetic process | GO:1901576 | 16 | 4.74E-02 | 2 | ns |
| Plastid organization | GO:0009657 | 5 | 4.19E-02 | 0 | ns |
| Cellular component assembly | GO:0022607 | 7 | 2.64E-04 | 0 | ns |
| Metabolic process | GO:0008152 | 76 | 2.17E-02 | 51 | ns |
| Response to stimulus | GO:0050896 | 7 | ns | 21 | 3.03E-03 |
| Response to chemical | GO:0042221 | 4 | ns | 12 | 1.95E-02 |
| Response to abscisic acid | GO:0009737 | 0 | ns | 5 | 7.62E-03 |
| Hormone-mediated signaling pathway | GO:0009755 | 0 | ns | 5 | 1.78E-02 |

*Pathway Enrichment Analysis*

Among the four categories of DEGs: 4610 up-regulated, 4610 down-regulated, Rondo up-regulated, and Rondo down-regulated, only 4610 down-regulated DEGs and Rondo down-regulated DEGs had the enriched pathways (Fig. 3). The 4610 down-regulated DEGs had 7 enriched pathways including photosynthesis, carbon metabolism, glyoxylate and dicarboxylate metabolism, carbon fixation in photosynthesis organisms, porphyrin and chlorophyll metabolism, sulfur metabolism, and ribosome (Fig. 3A). Rondo down-regulated DEGs had only MAPK signaling pathway enriched (Fig. 3B). Among the seven enriched pathways of 4610 down-regulated DEGs, there were three pathways, carbon metabolism, glyoxylate and dicarboxylate metabolism, and carbon fixation in photosynthesis organisms that interacted with each other (Fig. 3C). However, another 3 photosynthesis-related pathways, photosynthesis, carbon fixation in photosynthesis organisms, and porphyrin and chlorophyll metabolism that did not share any common genes between pathways.

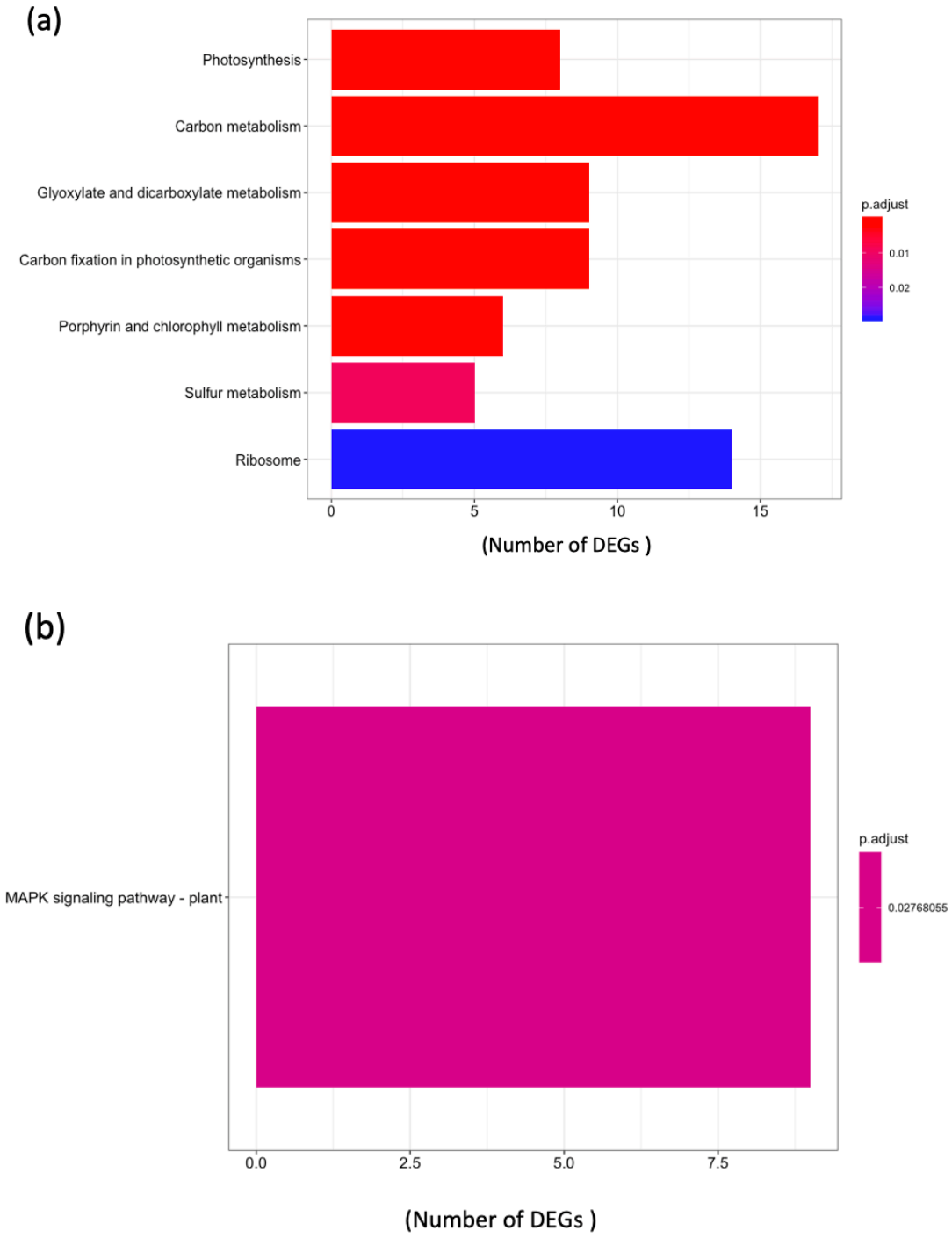


**Figure 3. KEGG enrichment analysis of down-regulated DEGs. (a) Pathway of down-regulated DEGs in 4610. (b) Pathway of down-regulated DEGs in Rondo. (c) Interaction between enriched pathways in 4610.**

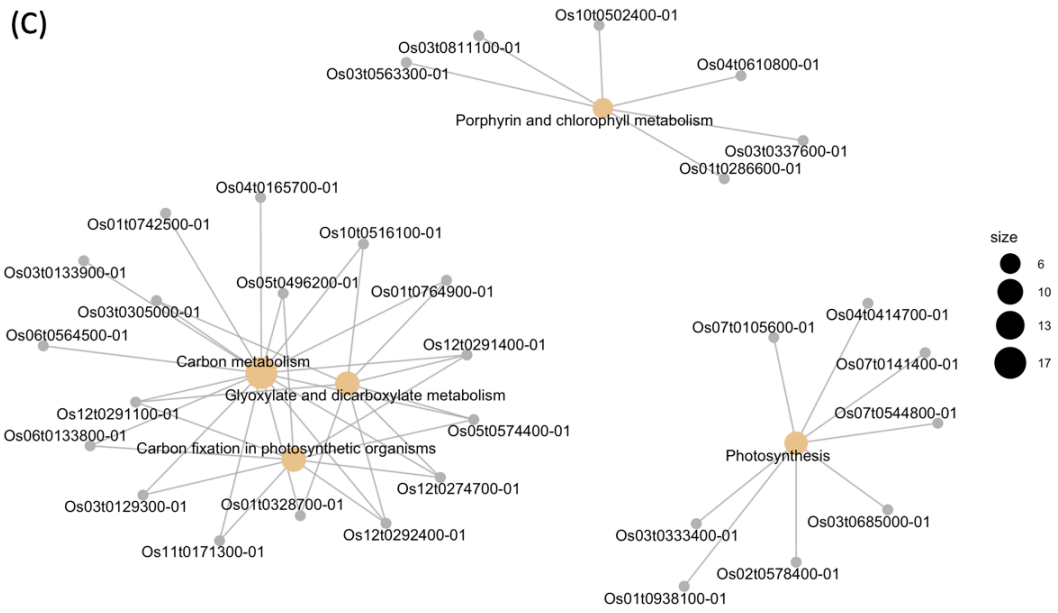


**Figure 3. Continued.**

## Discussion

*Agronomic and Physiological Performance and Gene Regulation under Drought Stress*

Previous researches indicated that water stress reduces plant biomass, grain size, panicle development, and fertility (Garrity & O'Toole, 1994; van der Weijde et al., 2017; Zhang et al., 2018). The results also showed that panicle number, filled grain number, spikelet fertility, grain weight, plant dry weight, panicle length, and yield were decreased under drought stress. After further evaluation of the DEGs and enrichment analysis, I found that genes involved in the regulations of branching, grain weight, grain size, and photosynthesis were significantly decreased in both genotypes. According to Rice Annotation Project Database (RAP-DB, https://rapdb.dna.affrc.go.jp), *MIR408* (OS01G0322700) is involved in regulation of panicle branching, grain weight, grain yield and photosynthesis. *GW5* (OS05G0158500) is a serine carboxypeptidase, which acts as a positive regulator of grain size. The expression of *MIR408* was decreased 2 and 4 times in Rondo and 4610, respectively. Comparing to the panicle numbers, Rondo had 10% reduction of panicle numbers; whereas that of 4610 was 35% (Appendix Table1). Although the panicle number is determined before the drought stress was imposed, the drought stress may still affect the initiation of the secondary panicles. *GW5* had decreased expression level with 2 times in both genotypes under drought stress. Therefore, our study suggests that the down-regulated of *MIR408* and *GW5* may partially contribute on the reduction of panicle numbers and hundred-seed weight under drought stress condition.

### *Diagnosis of DEG analysis*

The MA plot and the p-value distribution histogram are two classic diagnostic plots for DEG analysis that can be used to diagnosis whether the results and the methods are appropriate. The MA plots were symmetrical with the horizontal line log2 fold-change equal to 0 (Fig. 1). This indicated that the read counts of all samples had been normalized as expected. The MA plot also showed that genes expressed with no significant differences (black dot) were clustered around log2 fold-change equal to 0, indicating the relationship between fold-change and p-value was held. The p-value distribution histogram showed the proportion of p-value, which can be used to diagnosis whether the multiple testing approach is reliable. The results showed that the raw p-value of genotype-by-treatment effect and treatment effect in Rondo, were U-shaped, with p-values dense near 0 and 1, and spread uniformly in between (Appendix Fig. 1 and 3). The p-value distribution of treatment effect in 4610 was most dense near 0 and became less dense as the p-values increase (Appendix Fig. 2). Both two shapes indicated the results of multiple testing were reliable. The green histogram showed that the proportion of padj less than 0.05 were very few in three DEG analysis, which indicated that the DEGs analysis with FDR control was appropriate and stricter (Appendix Fig. 1-3).

### *Expression Patterns of Drought Responsive Gene Families*

The DEGs profile showed that bZIP transcription factors, LEA proteins, HSPs, and other drought tolerance proteins were up-regulated in both 4610 and Rondo. GO

enrichment results also showed that 4610 had many more DEGs involved in the stress responsive categories. To further study the expression pattern between the two genotypes, I developed gene lists based on drought-related categories and compared the log2 fold-change of expression level between 4610 and Rondo. The log2 fold-change was obtained from the differential gene expression analysis which indicated the gene expression difference between the drought and control conditions.

LEA protein is an unfolded aqueous solution acting as the substitute of water, that can prevents protein aggregation in a cell during water stress, which allows the cell to maintain its function (Goyal, Walton, & Tunnacliffe, 2005). Therefore, LEA proteins were annotated as drought and cold tolerance protein and have been overexpressed in many different plants, including brassica and rice, to improve their tolerance to such stresses (Dalal, Tayal, Chinnusamy, & Bansal, 2009; Xiao, Huang, Tang, & Xiong, 2007). The results showed that under drought condition, many LEA family proteins had higher expression level in 4610 than Rondo (Fig. 4A). *OsLEA23*, *OsLEA14*, *OsLEA29*, *OsLEA28*, and *OsLEA3-1* were differentially expressed under drought, and clearly show that *OsLEA29*, *OsLEA28*, and *OsLEA3-1* were expressed 2 to 8 times higher in 4610 than Rondo. Other than LEA proteins, HSPs also play a role in helping with the protein folding and damage prevention under the stress (Lee et al., 2005). The results showed that a number of HSPs also expressed higher level in 4610, especially *Hsp20*, *HSP16.9A*, *Hsp 90.1* and transcription factor *A2A*. Notably, a chaperone protein *Hsp40* was significantly decreased in Rondo (Fig. 4B). Some of bZIP transcription factors have also been known to enhance the gene regulation under stress in rice (Liu et al., 2014). The

results showed that a few members of the bZIP transcription factors were up-regulated in both 4610 and Rondo (Fig. 4C). *OsbZIP23* (OS02G0766700), was known to be involved in drought resistance was more upregulated in 4610. Notably, two *bZIP* transcription factors related to the ABA signaling were more up-regulated, and a flowering promotion gene (OS05G0518000) was down-regulated in 4610, indicating that 4610 might delay the subsequent flowering to escape drought stress. This strategy has been demonstrated in orange (Juan Carlos, Jill, Albrigo, & James, 2010) and rice (Zhang et al., 2016). Notably, bZIP transcription factor *05* was actually expressed with higher level in Rondo.

Aside from gene families above, DREB genes also annotated as drought and cold tolerance gene by regulating drought and cold stress response gene in the ABA-independent pathway (Lata & Prasad, 2011; Shinozaki & Yamaguchi-Shinozaki, 2000). However, in the current transcription profile among a total of 19 annotated DREB family genes in rice, none of them was significant differentially expressed.

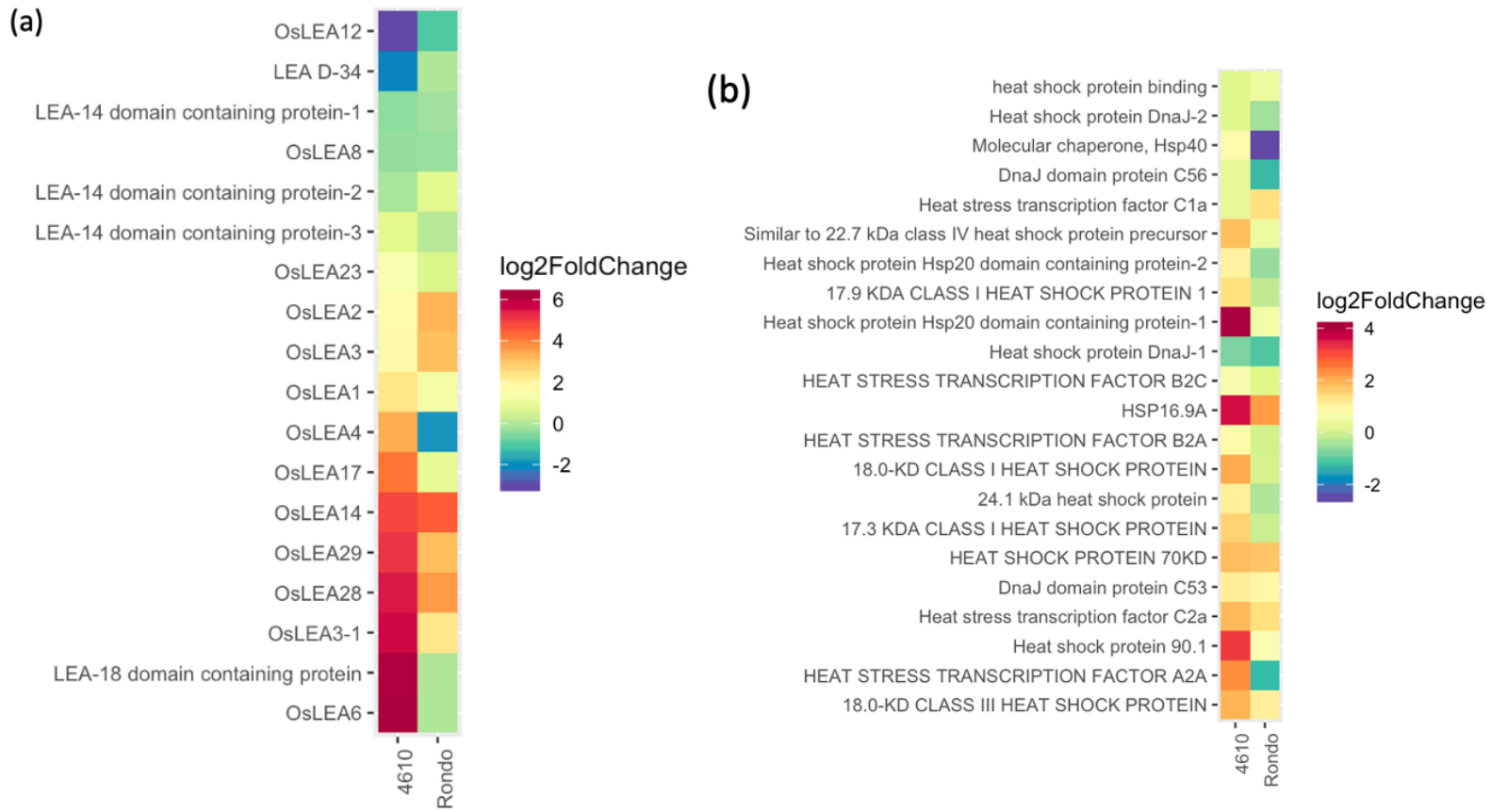


**Figure 4. The expression differences between 4610 and Rondo. (a) Late embryogenesis abundant (LEA) proteins. (b) Heat shock proteins (HSPs). (c) bZIP transcription factors.**

(c)

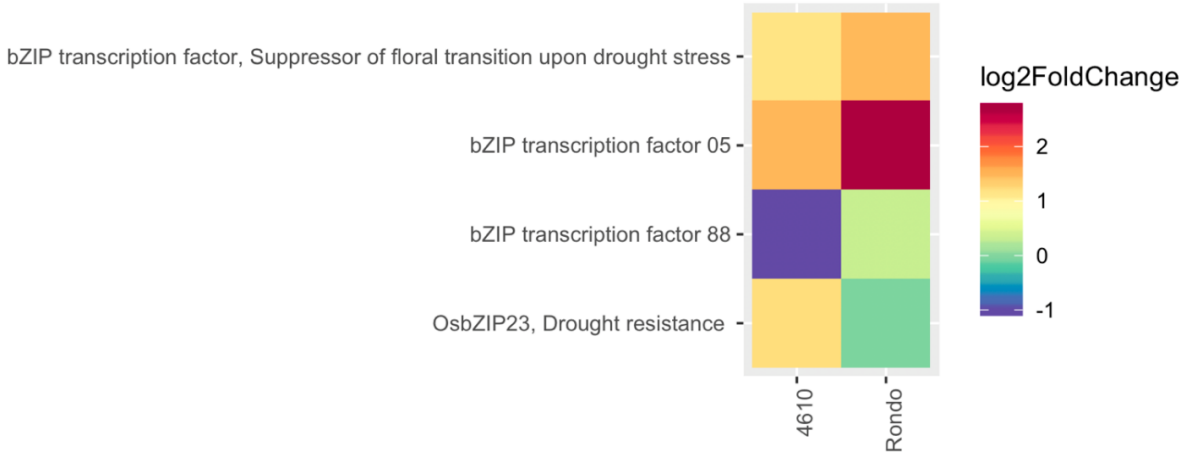


**Figure 4. Continued.**

*Expression Patterns of Other Drought Tolerance Genes*

Some of the genes related to drought tolerance or drought resistance were included NAC transcription factors (Zheng, Chen, Lu, & Han, 2009), Calcium-dependent protein kinase (Zou et al., 2010), Calcium signaling (Cheong et al., 2003), potassium channel (Mian, Ahmad, & Maathuis, 2016), ABA signaling (Kizis & Pagès, 2002), RCI2 protein (Li et al., 2014), and beta-carotene hydrolase (Du et al., 2010). To better understand the differences of gene expressions between 4610 and Rondo under drought stress, those related-tolerance genes were obtained from RAP-DB and were investigated. The log2 fold-change results showed that the basic helix-loop-helix transcription factor (OS03G0741100), *OsAHL1* (OS03G0741100), and *Leucine-rice receptor like kinase* (OS02G0154000) were significantly down-regulated in Rondo while there were no differences in expressions in 4610 under drought stress (Fig. 5). For other genes, such as the two-pore $K^+$ channel (OS07G0108800), *OsDSSR1*, *RIC2*, and *beta-carotene hydrolase,* were up-regulation in both genotypes with higher expression level in 4610 was observed. Thereby, the drought tolerance of 4610 might be significantly enhanced. Further, a negative regulator of the drought tolerance transcription factor, *NACTF* (OS08G0200600) was significantly up-regulated in Rondo (Fig. 5). The down-regulated MAPK pathway, which transmits the signals to activate tolerance response (Mizoguchi, Ichimura, & Shinozaki, 1997), was also enriched in Rondo (Fig. 3B). Interestingly, a negative regulator of the cuticular wax biosynthesis gene (OS02G0682300), which reduces drought tolerance, was also significantly up-regulated in 4610.

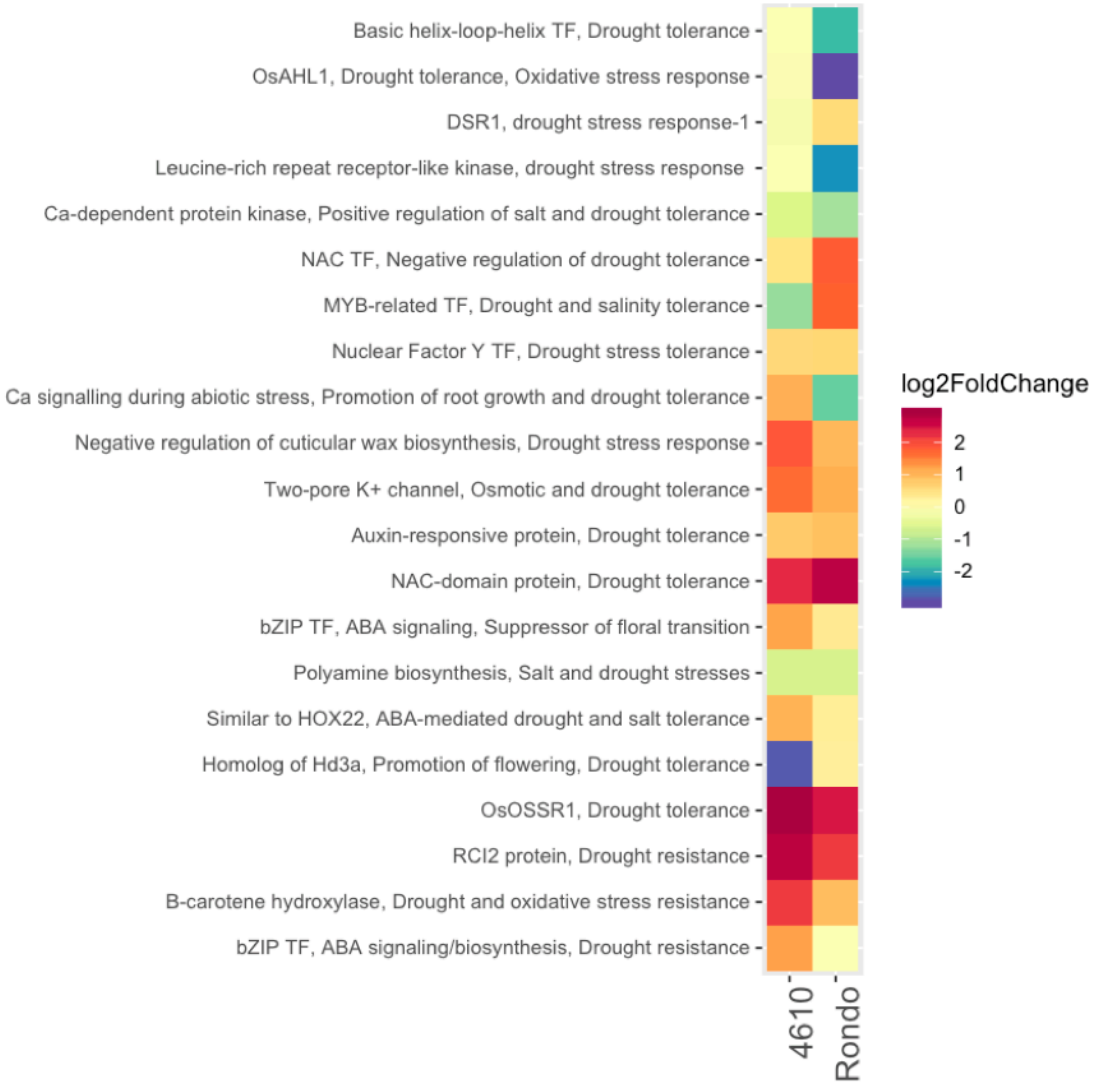


**Figure 5. The expression difference of annotated drought tolerance genes in 4610 and Rondo.**

*ROS Scavengers Enhanced Drought Tolerance*

ROS scavengers play an important role in stress tolerance mechanisms in plants (Das & Roychoudhury, 2014). Under drought stress, the increased ROS accumulation (Mittler, 2002) cause oxidative stress and damage at different levels such as proteins, lipids, DNA and RNA (Chool Boo & Jung, 1999; Moran et al., 1994; Sgherri, Pinzino, & Navari-Izzo, 1993). There are ROS scavenging mechanisms in plants that are able to protect the cells under drought stress, such as the scavenging enzymes and non-enzyme antioxidants. There are two main enzymes that act in the ascorbate/glutathione scavenging pathway, the APXs and the glutathione reductases (GR), which can convert toxic $O_2^-$ to $H_2O$ (Foyer & Noctor, 2011). Other than APXs and GRs, other ROS scavenging enzymes including superoxide dismutase (SOD), catalase (CAT), and glutathione S-transferases (GSTs) (Cruz de Carvalho, 2008; Kumar & Trivedi, 2018). Therefore, we subtracted the log2 fold-change of ROS scavenging enzymes from significant DEGs to further evaluate their expression levels. The results showed that *APX7* (OS04G0434800) was up-regulated in both genotype with higher expression level in 4610 (Fig. 6A). Interestingly, *peroxidase 11 precursor* (OS06G0274800) was expressed higher in Rondo. For GSTs, *GSTU1* (OS03G0785900), *GST28* (OS03G0785900), *GST39* (OS03G0785900), *GST6* (OS01G0692000), and *mitochondrial sulfur dioxygenase* (OS01G0667200) were expressed higher in 4610 (Fig. 6B).

Although *GST* is not directly involved in ascorbate/glutathione scavenging pathway, previous study has revealed that GST can protect cell from oxidative damage by quenching reactive molecules with the addition of glutathione (Kumar & Trivedi, 2018), and GSTs also have been proved to enhance chilling and drought tolerance in rice (Guo, Ou, Lu, & Zhong, 2006). For SOD and CAT, the results did not show any significant DEGs belonging to these two categories. Therefore, this study suggests that ROS scavengers APXs and GSTs play an important role in enhancing drought tolerance in 4610.

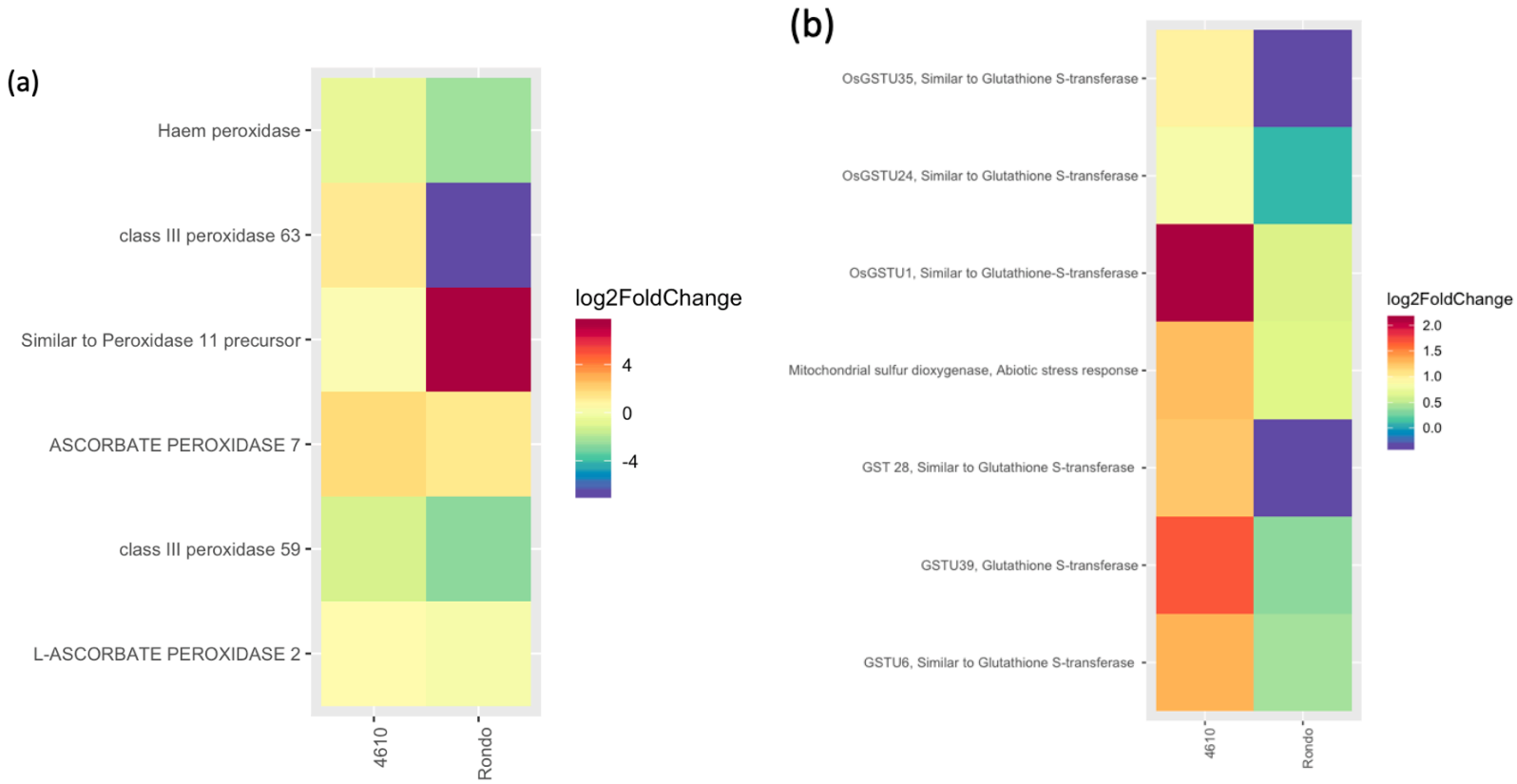


**Figure 6. The expression difference of ROS scavenging genes in 4610 and Rondo. (a) Peroxidases. (b) Glutathione S-transferases (GSTs).**

## Conclusion

The study showed that filled grain number, spikelet fertility, grain weight, plant dry weight, and yield were decreased for both genotypes under drought stress. Based on the relative agronomic performances of the two rice genotypes under control and drought conditions, 4610 had better performance than Rondo under drought stress condition. The results from the transcriptomic study also indicated that 4610 had more stress responsive genes that were differentially expressed than its counterpart. Furthermore, several known genes families related to drought stress or stress-related conditions were also identified in 4610 up-regulated genes, including the LEA proteins and HSPs. The ROS scavengers, especially APXs and GSTs, also had higher expression in 4610, which may enhance the tolerance of 4610. Further validation of drought-related DEGs can be validated by qPCR.

# CHAPTER III

# REFERENCE-GUIDED DE NOVO GENOME ASSEMBLY TO DISSECT A QTL REGION FOR SUBMERGENCE TOLERANCE IN CIHERANG-SUB1

## Introduction

Rice (*Oryza sativa* L.) is a staple food that feeds more than half of the world population (Ricepedia, http://ricepedia.org/rice-as-food/the-global-staple-rice-consumers). Submergence induces hypoxia stress for plants which inhibits photosynthesis, leads to anaerobic metabolism, and may reduce crop productivity dramatically (Fukao & Bailey-Serres, 2004). In the U.S., for example, the recent hurricane Harvey brought heavy rain and rapid flood and caused $7.5 million in losses for rice and soybean in Texas in 2017 (Fannin, 2017). Due to the climate change, extreme weather events have increased the frequency of severe flooding which significantly affect agriculture, including rice production both in the US and internationally. A strong submergence tolerance gene, *Submergence 1* (*SUB1*) from a landrace FR13A had been cloned and introgressed into a few popular rice varieties (Septiningsih et al., 2009; Xu et al., 2006). *Sub1A* is an ethylene-response factor (ERF) which is activated by ethylene accumulation caused by water jacket effect during hypoxia circumstances (Glinski, 2018). Recently, a novel submergence tolerance QTL, *qSub8.1* was identified and mapped on chromosome 8 from a SUB1-tolerant allele fixed RIL mapping population, derived from Ciherang-Sub1 and IR10F365 (Gonzaga et al.,

2017). Ciherang-sub1 was developed from a cross of Ciherang, a popular variety from Indonesia, and IR64-Sub1. IR10F365 is an advanced breeding line having the *SUB1* gene. It was hypothesized that this QTL may have a complimentary effect with the *SUB1A* gene in the events of submergence during vegetative stage. A variety with prolonged submergence tolerant would be beneficial to maintain rice production when we face the increasing effects of climate change.

The goals of this study were to characterize the region of *qSub8.1*, and obtain sequence information of the three cultivars to facilitate candidate gene identification and validation, and DNA marker development for recombinant identification and molecular breeding. In order to fully understand the genome profile and genome-wide nucleotide information, a whole genome shotgun sequencing and a reference-guided *de novo* genome assembly were performed in the current study. WGS is a NGS method that cuts the whole genome into small pieces and sequences each of them in a particular way. DNA was isolated from leaf samples of Ciherang-Sub1 and its parental cultivars, Ciherang and IR64-Sub1. The high quality DNA samples were then sent for 150 bp pair-end WGS. Reads were first *de novo* assembly to construct scaffolds; thereafter the scaffolds were aligned to a reference genome to obtain the relative orders to build the genome. Based on the alignment to the reference sequence Nipponbare, Ciherang-Sub1, Ciherang and IR64-Sub1 have genome sizes of 345.4 Mb, 343.7 Mb, and 344.7 Mb, respectively. On the other hand, based on the alignment to the *indica* reference, Minghui 63 (MH63) (Zhang et al., 2016), Ciherang-Sub1, Ciherang and IR64-Sub1 have genome sizes of 354.9 Mb, 353.9 Mb, and 353.9 Mb, respectively.

## Materials and Methods

### *Plant Materials*

Ciherang-Sub1, which carries the novel submergence tolerance QTL *qSub8.1* (Gonzaga et al., 2017), was derived from a cross of the submergence susceptible parent Ciherang, a popular variety from Indonesia, and the submergence resistant parent IR64-Sub1 (Septiningsih et al., 2014). In order to characterize *qSub8.1*, Ciherang-Sub1, Ciherang, and IR64-Sub1 were selected for the whole genome sequencing and assembly. The seeds of Ciherang-Sub1, Ciherang, and IR64-Sub1 were imported from the International Rice Research Institute (IRRI), Philippines for the experiment and were planted in a quarantined greenhouse in Texas A&M AgriLife Research Center in Beaumont, Texas in 2017 for seed productions.

### *Whole Genome Sequencing and Assembly*

DNA was collected from healthy leaves from the three cultivars and stored in -80°C freezer before extraction. A modified CTAB method was used to obtain high quality of DNA (Doyle, 1987), with 2% CTAB, 100 mM Tris (Tris(hydroxymethyl)aminomethane), 700 mM NaCl, 20 mM EDTA, 0.9% sodium bisulfate, 4% polyvinylpyrrolidone (PVP-40) and 0.5% β-mercapto-ethanol. DNA quality was accessed using Agilent 2100 Bioanalyzer and libraries were prepared at the Texas A&M AgriLife Genomics and Bioinformatics Service (TxGen). Whole-genome shotgun sequencing of the three rice genomes was performed on Illumina HiSeq4000

platform to provide at least 80-100 million reads per sample, with 150 bp pair-end. The quality control of raw reads was performed using Trimmomatic version 0.36 (Bolger et al., 2014) with PE –Phred 33 command. The quality control was determined using the following steps: (1) raw sequencing reads were trimmed to remove adaptors; (2) low quality bases with quality score less than 20 on the ends and tails of reads were removed; (3) reads were scanned with a 5 bp sliding window, and an average quality of sequences per bp drops below 20 were removed ; (4) reads less than the 25 bases long were dropped; and (5) reads without correspondence read pairs were dropped. SOAPdenovo2 version r240 (Luo et al., 2012) was used to perform the *de novo* assembly. In SOAPdenovo2, both *-63mer* and *-127mer* were used. MUMmer version 3.23 (Kurtz et al., 2004) was used for aligning *de novo* scaffolds and contigs to (1) *O. sativa* L. spp. japonica reference genome (IRGSP-1.0); and (2) *O. sativa* L. spp. indica reference genome to decide the order of all scaffolds and contigs. Reference genome of indica rice cultivar MH63 version 2 (MH63RS2) was downloaded from the Rice Information GateWay (RIGW, http://rice.hzau.edu.cn/rice/). Following this, three genomes were constructed by re-orienting and connecting all the scaffolds and contigs according to their order.

### *Variants Calling*

The *O. sativa* L. spp. japonica reference genome (IRGSP-1.0) was used as a reference genome in the assembly. Reads were sorted by individual sample and aligned to the reference genome using the Bowtie2 v2.2.9 (Langmead & Salzberg, 2012) with

default parameters for end-to-end mode. The likelihood of each genotype was computed and summarized using *samtools mpileup* function in SAMtools v0.1.19 (H. Li et al., 2009), and the actual variant calling was performed using *bcftools -bvcg* function. SNPs were identified when there were polymorphisms between the three cultivars.

To make the genome structure graph, SNPs of Ciherang-Sub1 were compiled with two different window sizes, 50k and 100k. That is, the physical distance 50k and 100k were used as a block, calculated total SNP types (Ciherang-alike or IR64-Sub1-alike) within the block and decided the dominant SNP to represent each block.

## Results

### *Genome Assembly and Validation*

The Illumina sequencing results showed that there were more than 100 million raw reads in each cultivar and more than 93% of the raw read-pairs were good quality that can be used in *de novo* assembly (Table 6). A total of 102,213,313 reads of Ciherang-Sub1 genome, corresponding to 15,331,996,950 bp were generated, and represented 38-fold sequencing depth and covered 91% of the Nipponbare reference genome or 90.5% of the MH63 reference genome (Table 7). Reference-guided *de novo* assembly of three cultivars was combined *de novo* assembler SOAPdenovo2 with reference genome alignment (Fig. 7). In the first step, reads were *de novo* assembly with two different functions, *63mer* and *127mer*, separately. The results showed that all three cultivars had better average scaffold length, length of longest scaffold, and N50, which

demonstrate the length of the 50$^{th}$ percentile of scaffolds in determining the assembly quality were also assessed, with *63mer* (Table 8). The average scaffold length (bp) of Ciherang-Sub1, Ciherang and IR64-Sub1 were 6907, 7456, and 7521 with *127mer*; whereas, 7819, 8115, and 8321 with *63mer*, respectively. The longest scaffold (bp) of Ciherang-Sub1, Ciherang and IR64-Sub1 were 80935, 101859, and 120411 with *127mer*; and 212257, 210201, 210221 with *63mer*, respectively. The N50 of Ciherang-Sub1, Ciherang and IR64-Sub1 were 11298, 12635, and 12524 with *127mer*; and it was more than twice as long in *63mer* with 25138, 25390, and 26139 in the three cultivars, respectively. Therefore, the scaffolds assembled with *63mer* were used in the next step. All scaffolds were aligned to the Nipponbare reference genome and the MH63 reference genome separately with *nucmer* function. The scaffolds were then connected based on their orders to become a complete genome. The results showed that Ciherang-Sub1, Ciherang, and IR64-Sub1 had genome size of 345.4 Mb, 343.7 Mb, and 344.7 Mb with 91%, 90.6%, and 90.8% coverage based on the Nipponbare reference genome (Table 7). For alignment based on MH63 reference genome, Ciherang-Sub1, Ciherang, and IR64-Sub1 had genome size of 354.9 Mb, 353.9 Mb, and 353.9 Mb with 90.5%, 90.2%, and 90.2% coverage, respectively.

**Table 6. Results of raw reads cleaning.**

| Variety | Number of raw read pairs | Number of read pairs after trimming |
|---|---|---|
| Ciherang-Sub1 | 107,957,981 | 102,213,313 (94.68%) |
| Ciherang | 122,304,187 | 115,172,273 (94.17%) |
| IR64-Sub1 | 122,140,197 | 114,085,140 (93.41%) |

**Table 7. General sequencing statistics and summary of assembly.**

| Variety | Number of reads | Total read length (bp) | Sequencing depth (X) | Genome size (bp) | |
|---|---|---|---|---|---|
| | | | | Nipponbare reference | MH63 reference |
| Ciherang-Sub1 | 102,213,313 | 15,331,996,950 | 38 | 345,442,284 (91.0%) | 354,934,762 (90.5%) |
| Ciherang | 115,172,273 | 17,275,840,950 | 43 | 343,737,849 (90.6%) | 353,859,778 (90.2%) |
| IR64-Sub1 | 114,085,140 | 17,112,771,000 | 43 | 344,678,967 (90.8%) | 353,859,751 (90.2%) |

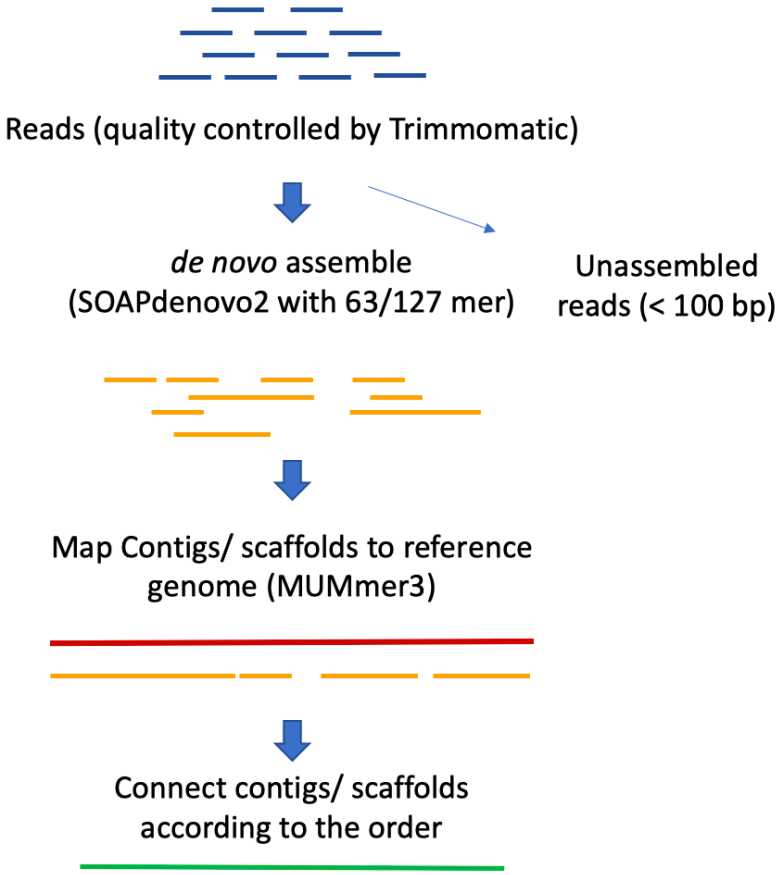


**Figure 7. Reference-guided *de novo* assembly pipeline.**

**Table 8. Summary of *de novo* assembled scaffold.**

| | 127mer | | | 63mer | | |
|---|---|---|---|---|---|---|
| | Ciherang-Sub1 | Ciherang | IR64-Sub1 | Ciherang-Sub1 | Ciherang | IR64-Sub1 |
| Number of scaffold | 43595 | 40208 | 39881 | 40341 | 38595 | 37731 |
| Average length (bp) | 6907 | 7456 | 7521 | 7819 | 8115 | 8321 |
| Length of longest scaffold (bp) | 80935 | 101859 | 120411 | 212257 | 210201 | 210221 |
| N50 (bp) | 11298 | 12635 | 12524 | 25138 | 25390 | 26139 |
| N90 (bp) | 1372 | 1515 | 1595 | 514 | 354 | 468 |

The three genome assemblies were validated by the position of known genes, *GW5*, *SD1*, *Sub1A*, and *Sub1B*. Results showed that the positions of *SD1* and *GW5* were correct in all three cultivars and aligned well to both Nipponbare (Table 9) and MH63 (Table 10). *SD1* is located at 38,382,385 to 38,385,469 on chromosome 1 in the Nipponbare genome. For the three cultivars, *SD1* is located on chromosome 1 at 42,999,152 to 42,501,761 in Ciherang-Sub1, 43,076,841 to 43,079,530 in Ciherang, and 42,596,482 to 42,597,334 in IR64-Sub1. *GW5* is positioned on chromosome 5 at the physical position of 4,832,476 to 4,834,024 in Ciherang-Sub1, 4,873,155 to 4,874,734 in Ciherang, and 5,262,475 to 5,264,054 in IR64-Sub1. For *Sub1B*, Ciherang-Sub1 and IR64-Sub1 had similar position with Nipponbare reference, both located on chromosome 9 and at the position of 5,266,747 to 5,267,932 of Ciherang-Sub1, and 4,932,683 to 4,933,868 of IR64-Sub1. However, for Ciherang, *Sub1B* is located at 4,258,518 to 4,259,713 on chromosome 1. Interestingly, *Sub1A* is located at 3,892,209 to 3,896,051 on chromosome 11 of Ciherang-Sub1, 4,281,291 to 4,284,501 on chromosome 10 of Ciherang, and 8,519,270 to 8,523,113 on chromosome 8 of IR64-Sub1.

Table 10 shows the position of four genes in the three genomes based on the MH63 reference. *SD1* is located on chromosome 1 for three cultivars, at the position of 42,731,953 to 42,734,562 of Ciherang-Sub1, 42,731,953 to 42,734,562 of Ciherang, and 42,731,953 to 42,734,562 of IR64-Sub1. *GW5* is also on chromosome 5 for three cultivars, at the position of 5,278,241 to 5,279,789 of Ciherang-Sub1, 5,178,118 to 5,179,697 of Ciherang, and 4,864,637 to 4,866,216 of IR64-Sub1. For *Sub1B*, Ciherang-Sub1 and IR64-Sub1 had similar position with MH63 reference, both located on

chromosome 9 and at the position of 4,262,753 to 4,263,938 of Ciherang-Sub1, and 5,209,441 to 5,210,626 of IR64-Sub1. For Ciherang, *Sub1B* is located at 7,911,678 to 7,912,873 on chromosome 12. *Sub1A* is on chromosome 11 of three cultivars, at the position of 17,912,307 to 17,916,149 of Ciherang-Sub1, 8,533,735 to 8,556,945 of Ciherang, and 18,314,723 to 18,318,566 of IR64-Sub1.

**Table 9. Summary of position of known gene in the three genomes based on the Nipponbare reference genome.**

| | Ref. genome | Ciherang-Sub1 | Ciherang | IR64-Sub1 |
|---|---|---|---|---|
| SD1 | chr01:38,382,385 | chr01:42,499,152 | chr01:43,076,841 | chr01:42,596,482 |
| (Os01g0883800) | ..38,385,469 | ..42.501,761 | ..43,079,530 | ..42,597,334 |
| GW5 | chr05:5,365,122 | chr05:4,832,476 | chr05:4,873,155 | chr05:5,262,475 |
| (Os05g0187500) | ..5,366,701 | ..4,834,024 | ..4,874,734 | ..5,264,054 |
| Sub1A | chr08:6,319,062 | chr11:3,892,209 | chr10:4,281,291 | chr08:8,519,270 |
| | ..6,321,934 | ..3,896,051 | ..4,284,501 | ..8,523,113 |
| Sub1B | chr09:6,388,205 | chr09:5,266,747 | chr1:4,258,518 | chr09:4,932,683 |
| (Os09g0287000) | ..6,389,719 | ..5,267,932 | ..4,259,713 | ..4,933,868 |

**Table 10. Summary of position of known gene in the three genomes based on the MH63 reference genome.**

| | Ref. genome | Ciherang-Sub1 | Ciherang | IR64-Sub1 |
|---|---|---|---|---|
| SD1 | chr01:39,643,093 | chr01:42,731,953 | chr01:42,731,953 | chr01:42,731,953 |
| (OsMH_01G0636900) | ..39,644,426 | ..42,734,562 | ..42,734,562 | ..42,734,562 |
| GW5 | chr05:5,428,533 | chr05:5,278,241 | chr05:5,178,118 | chr05:4,864,637 |
| (OsMH_05G0081900) | ..5,430,112 | ..5,279,789 | ..5,179,697 | ..4,866,216 |
| Sub1A | chr06:22,422,489 | chr11:17,912,307 | chr11:8,553,735 | chr11:18,314,723 |
| | ..22,419,199 | ..17,916,149 | ..8,556,945 | ..18,318,566 |
| Sub1B | chr09: 7,179,132 | chr09: 4,262,753 | chr12: 7,911,678 | chr09: 5,209,441 |
| (OsMH_09G0114700) | ..7,180,335 | ..4,263,938 | ..7,912,873 | .. 5,210,626 |

*Genome Structure of Ciherang-Sub1*

The genome structure of Ciherang-Sub1 was examined by the characteristics of the genome-wide SNPs. A total of 636,243 SNPs were detected between Ciherang-Sub1, Ciherang, and IR64-Sub1. There were 83,555 SNPs on chromosome 8, and 5,665 SNPs were located within the *qSub8.1* region. Based on the assessment of SNPs with the 50k window size, a total of 7,453 SNPs was identified with 4,406 SNPs Ciherang-alike, and 1,791 SNPs IR64-Sub1 alike. Notably, there were 1,256 SNPs neither Ciherang nor IR64-Sub1. In summary, Ciherang-Sub1 has 59% of genetic background from Ciherang, 24% from IR64-Sub1, and 17% came from unknown sources. To better visualize the genome structure of Ciherang-Sub1, 7,453 SNPs were plotted according to their physical position with different colors representing their sources (Fig. 8). The genome structure graph illustrated that the majority fragments of chromosome 2, 5, 6, 7, 8, 10 and 12 of Ciherang-Sub1 were from Ciherang; whereas chromosome 1, 3, 4, and 11 were mixtures between Ciherang and IR64-Sub1. Notably, a clear region on the proximal end of chromosome 9 was from IR64-Sub1.

With the scope zoomed-in to *qSub8* region, we found that *qSub8* region of Ciherang-Sub1 was mainly from Ciherang background (Fig. 9). With 50 kb window size, there were two segments from IR64-Sub1, and two segments from other sources within *qSub8* region, and one segment from other sources slightly outside of *qSub8* region (Fig. 9A). However, with 100k window size, there was only one segment from other sources within *qSub8* region (Fig. 9B).

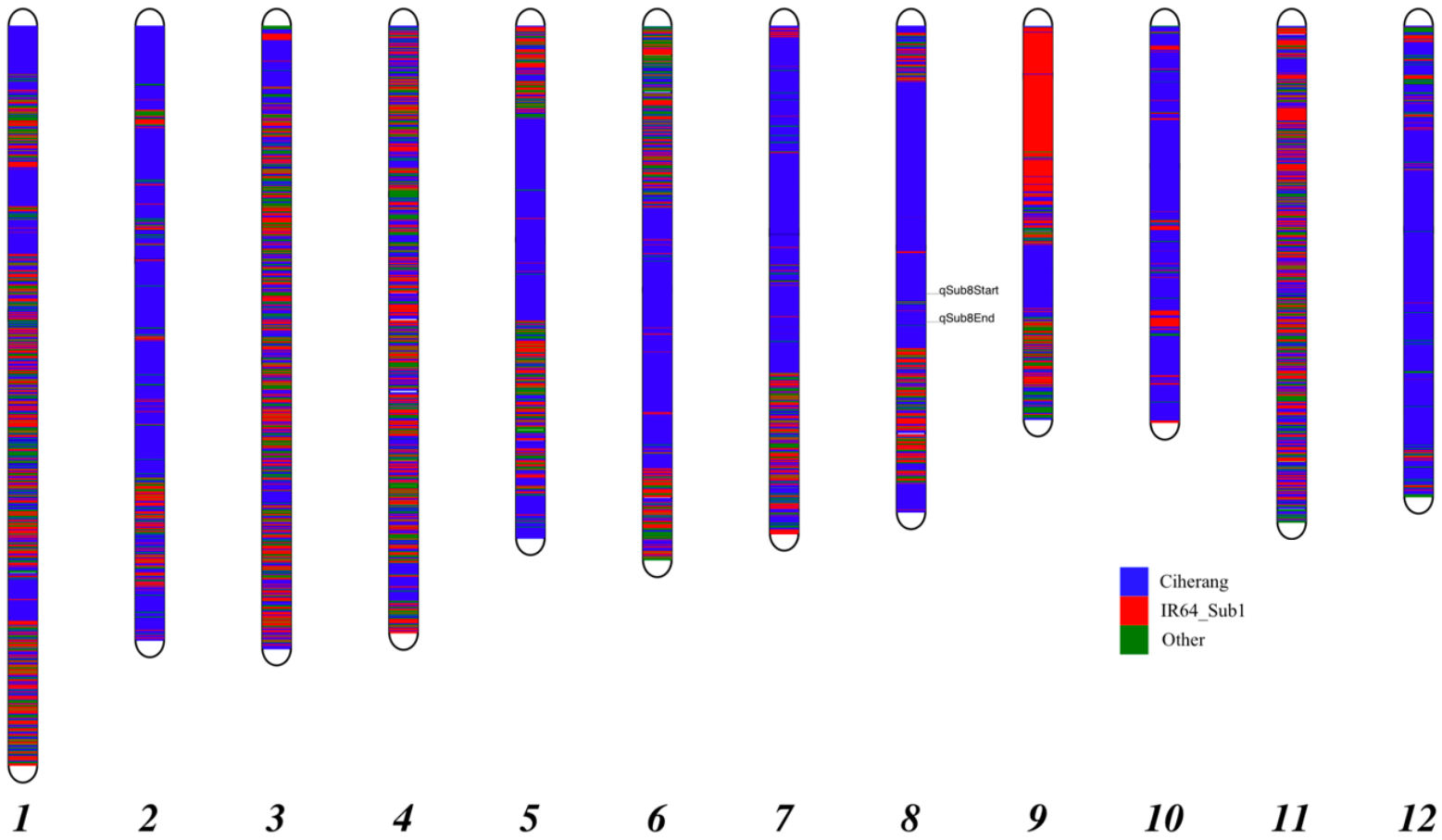


**Figure 8. Genome structure of Ciheramg-Sub1 with the window size of 50k.**

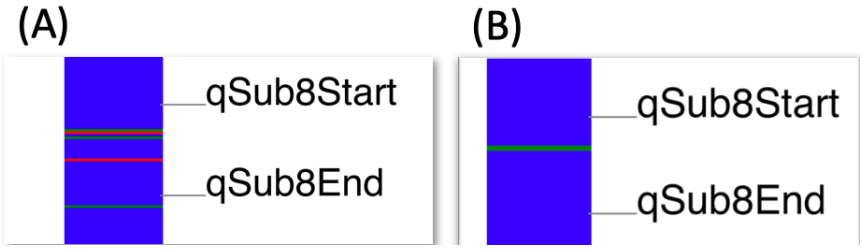


**Figure 9. *qSub8* region using different SNP block window sizes. (A) *qSub8* region with 50 kb window size. (B) *qSub8* region with 100 kb window size.**

## Discussion

### *Gene in Unexpected Position*

The results showed that some genes have different positions in the assembled genome than we expected such as *Sub1A* and *Sub1B* (Table 9 and 10). *Sub1A* has been cloned and revealed that it is near the centromere of chromosome 9 (Xu et al., 2006). However, in the three genomes, either based on Nipponbare or MH63 references, the position of *Sub1A* was not on chromosome 9. For genomes aligned with Nipponbare reference, *Sub1A* was on chromosome 11 for Ciherang-Sub1, chromosome 10 for Ciherang, and chromosome 8 for IR64-Sub1 (Table 9). This might be due to the fact that the *Sub1A* gene was actually not present in Japonica rice, including Nipponbare. Previous research indicated that *Sub1A* gene is absent in japonica rice and some other *Oryza* species such as *O. rhizomatis* and *O. eichingeri* (Niroula et al., 2012; Xu et al., 2006), therefore the scaffold contains *Sub1A* would be assigned to another location with higher sequence similarity. For the alignments with the *indica* MH63 reference, *Sub1A* was on chromosome 11 for the three cultivars (Table 10). However, it seems that *Sub1A* gene was not readily annotated in the MH63 genome (Zhang et al., 2016). Therefore, I searched the position with *Sub1A* complete sequence to further deliberate if MH63 has Sub1A locus. The results showed that *Sub1A* is corresponding to the distal end of chromosome 6. Previous research indicated that some indica varieties may also lack Sub1A locus such as IR24, Swarna, IR50, and Habiganj aman (Xu et al., 2006). Therefore, the unexpected location of Sub1A in MH63 genome could be due to its

genetic background or assembling error. Hence, we used Nipponbare as a reference genome in variant calling process even though all of the three cultivars belong to the *indica* rice subgroup.

*Sub1B* is on the upper region of chromosome 9 in Ciherang-Sub1, IR64-Sub1 in both reference genomes, whereas Ciherang has *Sub1B* position on a different chromosome (Table 9 and 10). This result revealed that Ciherang-Sub1 and IR64-Sub1 has high similarity in *Sub1B* locus. In fact, Ciherang-Sub1 was derived from Ciherang and IR64-Sub1, and IR64-Sub1 was Sub1 donor. When Sub1 was introgressed into Ciherang, it actually included Sub1A, Sub1B, and Sub1C. Therefore, Ciherang-Sub1 not only has Sub1A from IR64-Sub1 background but also Sub1B and Sub1C (Septiningsih et al., 2014). Sub1B position in Ciherang, however, the top match was at different chromosome. This might be due to the scaffold size that was a bit too short compared to those of Ciherang-Sub1 and IR64-Sub1.

### *Genome Structure*

The statistics indicated that Ciherang-Sub1 genome is composed of 59% Ciherang, 24% IR64-Sub1, and 17% unknown sources. I can also easily define that the majority of Ciherang-Sub1 genome was from Ciherang (Fig. 8). Notably, almost the entire proximal end of chromosome 9 is composed of IR64-Sub1 background. This result was clearly corresponding to the breeding process of Ciherang-Sub1. Ciherang-Sub1 was derived from marker-assisted breeding with two Sub1 flanking SSR markers ART5 and RM8300 (Septiningsih et al., 2014). Therefore, only this region was IR64-

Sub1 background and the rest of the part was mainly from Ciherang or a mixture with both parents. Ciherang-Sub1 was released as a cultivar at $BC_1F_2$ generation, and it has been cultivated in the paddy fields in IRRI and farmers' fields for a few years. This might explain the less purity of the seeds of this cultivar as well.

### *Source of qSub8.1 QTL*

*qSub8* region is delimited to a physical position of 15,666,434 to 17,319,167 on chromosome 8 in the Nipponbare genome. According to the genome structure of Ciherang-Sub1, *qSub8* region was mainly from Ciherang background with possible segment from IR64-Sub1 (Fig. 9A) or unknown sources (Fig. 9B). The structures profiled with 50 kb window size and 100 kb window size were slightly different. The one with 50 kb window size has another two segments from IR64-Sub1, and the other one with 100 kb window size has only one segment from unknown sources. Obviously, enlarged windows can reduce the noise from sequencing errors or ambiguous variant callings. However, some useful information could also be lost in this scenario. The ideal window size largely depends on the size of the linkage block within the region of interest. The sizes of linkage block vary from 75 kb to 500 kb in rice,and the linkage block is about 75 kb in the *indica* subgroup (Mather et al., 2007). Therefore, based on my analysis, the source of submergence tolerance QTL *qSub8.1* can be gene(s) from unknown sources or IR64-Sub1, or as a result of an epistatic interaction between gene(s) that are normally susceptible in Ciherang background, but then contribute to tolerance

with the Ciherang-Sub1 background. Hence, further studies are needed to elucidate this QTL.

## Conclusion

The Trimmomatic-SOAPdenovo2-MUMmer3 pipeline for genome assembly resulted in approximate genome sizes of 354.4 Mb, 343.7 Mb, and 344.7 Mb, with N50 values of 25.1 kb, 25.4 kb and 26.1 kb, respectively. The results showed that Ciherang-Sub1 genome is composed of 59% Ciherang, 24% of IR64-Sub1, and 17% of unknown sources. The genome profile showed that *qSub8* region is mainly from Ciherang with few introgressed segments from IR64-Sub1 and unknown sources. For the results with 50 kb window size, there were two introgressed segments from IR64-Sub1 and two introgressed segments from unknown sources; however, with the 100 kb window size only one introgressed segment from unknown sources was present. The SNPs obtained from variant calling can be used to develop *qSub8.1* NILs. To further narrow down and clone *qSub8.1*, a longer sequencing technology, such as PacBio sequencing, needs to be used.

# CHAPTER IV

# OPTIMIZATION OF PROTOCOL FOR EFFICIENT TRANSGENIC PLANT DEVELOPMENT OF AN ELITE INDICA RICE CULTIVAR, CIHERANG-SUB1 USING A BIOLISTIC DELIVERY OF CRISPR/CAS9 CONSTRUCT

## Introduction

In recent years, the clustered regularly interspaced short palindromic repeats/ Cas9 nuclease (CRISPR/ Cas9) technique has been considered as the most powerful tool for crop improvement with its precise gene editing ability. CRISPR technique has been applied to many crops to develop the gene edited plants, including wheat (Wang et al., 2014), soybean (Cai et al., 2015), and rice (Li et al., 2016; Miao et al., 2013). The CRISPR/Cas9 gene editing vector assembly of gRNA promotor, gRNA, and a Cas9 expression vector in one final destination vector (Lowder et al., 2015). Therefore, once the destination vector is delivered in the host plant, gRNA-Cas9 complex will be expressed, Cas9 will cut the host genome at the specific targeted site, and can cause the mutation through non-homologous end-joining (NHEJ) or homologous recombination (HR) (Paul & Qi, 2016). Agrobacterium-mediated transformation with T-DNA insertion is the most common and stable method to deliver the vector into the host plants (Gelvin, 2003). However, there are still some challenges remain for gene editing in plants, including the technically demanding and time-consuming aspect of tissue culture; and the low regeneration efficiency, especially for recalcitrant species or genotypes (Khanna

& Raina, 1998; Mao, Botella, Liu, & Zhu, 2019). In addition, in order to completely avoid the USDA GMO regulation, the delivery systems that bypass the bacterium delivery system also need to be established (https://www.aphis.usda.gov/aphis/ourfocus/biotechnology/brs-news-and-information/2018_brs_news/pbi-details).

In rice (*Oryza sativa* L.), the standard tissue and transformation protocol of japonica cultivars has been established using mature seeds as explants (Hiei, Ohta, Komari, & Kumashiro, 1994); on the contrary, *indica* cultivars are more recalcitrant for such a method with poor callus induction and regeneration and a low transformation efficiency (Aldemita & Hodges, 1996; Hiei & Komari, 2008). Previous study revealed that using immature embryos as explants can provide higher efficiency in *indica* cultivar (Hiei & Komari, 2008). However, there are some disadvantages of using immature embryos as explants, for example, inconvenience for dissecting the materials and also it is time consuming. Several studies have reported the development of an efficiency *indica* rice transformation protocol using mature seeds (Kumar, Maruthasalam, Loganathan, Sudhakar, & Balasubramanian, 2005; Lin & Zhang, 2005; Sahoo, Tripathi, Pareek, Sopory, & Singla-Pareek, 2011). However, *indica* cultivars are more genetically diverse than *japonica* cultivars, many high-efficiency protocols are genotype-specific, and the transformation efficiency may decrease due to different genotypes (Hiei & Komari, 2006; Zaidi et al., 2006).

Hence, the aim of the current study was to optimize the protocol for transformation and regeneration of the indica rice cultivar Ciherang-Sub1 using mature

seeds as explants and the biolistic bombardment as a delivery method for the CRISPR/Cas9 gene-editing construct. In this study, CRISPR/Cas9 gene editing vector was constructed following a published rice gene editing toolbox (Lowder et al., 2015) and the Biolistic® PDS-1000/He particle delivery system was used. Different carbon sources, auxin/cytokinin ratio, and callus age were also evaluated in this study. Results suggest that this protocol can accelerate the tissue culture process to 80 days with a high regeneration efficiency of 95%.

## Materials and Methods

### *Plant Materials and Callus Induction*

Mature dry seeds of an indica rice cultivar, Ciherang-Sub1 was used in this study to induce embryogenetic callus. Husks were removed from healthy seeds of Ciherang-Sub1 and the seeds were then sterilized with 1.5% sodium hypochlorite and 1 drop of Tween 20 for 20 min in mini rotator (Grant-bio, PTR-25) at speed of 3. Seeds were rinsed by autoclaved tap water for 7-8 times until all sodium hypochlorite solution had been removed. The seeds were then dried on autoclaved filter paper for 5-10 min to remove the excess liquid. For callus induction, 19-20 seeds were placed onto callus induction medium (CIM) in 15 mm deep petri dishes (90 x 15 mm) and incubated at 28°C in the dark. CIM contains 4.4 g/l Murashige and Skoog Basal Medium (Sigma-Aldrich, M404), 30 g/l maltose, 0.6 g/l proline, 0.3 g/l casein hydrolysate, 3 mg/l 2,4-dichlorophenoxyacetic acid (2,4-D), 0.3 mg/l 6-Benzylaminopurine solution (6BA), and 3 g/l phytagel. The pH was adjusted to 5.8 with 1N NaOH and the whole media then was autoclaved at 121°C for 15 min.

After the first incubation at 29°C in the dark for 14 days, calli were removed from roots and cut into 2-4 pieces depending on the size of the calli. These calli were then sub-cultured onto fresh CIM in 20 mm petri dishes (90 x 20 mm) at 29°C in the dark for another 14-20 days before the bombardment treatment.

*CRISPR/ Cas9 Vector Construction*

The full sequence of the *SUB1A* of Ciherang-Sub1 will be obtained from the genome assembly results. The second exon of *SUB1A* gene (Fig. 10) was targeted for knock-out through gene editing. Two gRNAs were designed using two online gRNA design tools CRISPRdirect and Cas-OFFinder (Bae, Park, & Kim, 2014; Bono, Hino, Ui-Tei, & Naito, 2014). The sequences of the two gRNAs were gRNA1: 5'-CCGGCGAGGAGGCTGTCCATCAC-3' and gRNA2: 5'-ACGGCCGCTGCCGGATGCCGTGG-3'. The CRISPR/Cas9 vectors assembly procedures of the *SUB1A* knock-out were performed according to a published protocol (Lowder et al., 2015). Two Golden Gate entry vectors, pYPQ131C and pYPQ132C, that contain OsU6 promoter were used to clone the two gRNAs to form gRNA expression cassette. The two gRNA expression cassettes were then assembled into the Golden Gate recipient vector pYPQ142. After that, both pYPQ142 and Cas9 expression vector (pYPQ150) were assembled into a binary vector, pMDC32, through the LR Clonase II recombination reaction. The final destination vector pMDC32 was then transformed into *Escherichia coli* (*E. coli*) competent cell DH5α through heat shock transformation at 42°C for 30s.

```
Query  1    GGTGTTGAGGTCACCAGGTGAAAATGATGCAGGCCGGGGCGCCGCCGCCACCATGTCCAT  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  675  GGTGTTGAGGTCACCAGGTGAAAATGATGCAGGCCGGGGCGCCGCCGCCACCATGTCCAT  616
Query  61   GCCGCTGGACCCCGTGACCGAGGAGGCCGAGCCGGCGGTGGCTGAGAAGCCTCGCCGGCG  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  615  GCCGCTGGACCCCGTGACCGAGGAGGCCGAGCCGGCGGTGGCTGAGAAGCCTCGCCGGCG  556
Query  121  CCGGCCGAGGCGGAGCTACGAGTACCACGGCATCCGGCAGCGGCCGTGGGGGCGGTGGTC  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  555  CCGGCCGAGGCGGAGCTACGAGTACCACGGCATCCGGCAGCGGCCGTGGGGGCGGTGGTC  496
Query  181  GTCGGAGATCCGCGACCCCGTCAAGGGCGTCCGCCTCTGGCTCGGCACCTTCGACACCGC  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  495  GTCGGAGATCCGCGACCCCGTCAAGGGCGTCCGCCTCTGGCTCGGCACCTTCGACACCGC  436
Query  241  CGTCGAAGCCGCGCTCGCCTACGACGCCGAGGCCCGCCGCATCCACGGCTGGAAAGCCCG  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  435  CGTCGAAGCCGCGCTCGCCTACGACGCCGAGGCCCGCCGCATCCACGGCTGGAAAGCCCG  376
Query  301  GACAAACTTCCCACCCGCCGATCTTTCTTCGCCGCCGCCGCCGTCGCAGCCGCTCTGCTT  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  375  GACAAACTTCCCACCCGCCGATCTTTCTTCGCCGCCGCCGCCGTCGCAGCCGCTCTGCTT  316
Query  361  CTTGCTCAACGACAACGGCCTCATCACAATCGGAGAAgcgccgaccgacgacgccgcgtc  420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  315  CTTGCTCAACGACAACGGCCTCATCACAATCGGAGAAGCGCCGACCGACGACGCCGCGTC  256
Query  421  gacgtcgacgtcgacgacggaggcgtccggcgacgcgcgcATACAACTGGAGTGCTGCTC  480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  255  GACGTCGACGTCGACGACGGAGGCGTCCGGCGACGCGCGCATACAACTGGAGTGCTGCTC  196
Query  481  GGACGACGTGATGGACAGCCTCCTCGCCGGCTACGACGTGGCCAGCGGCGACGACATATG  540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  195  GGACGACGTGATGGACAGCCTCCTCGCCGGCTACGACGTGGCCAGCGGCGACGACATATG  136
Query  541  GACATGGACATCTGGAGCCTCCTCCACCTCTGTTAACCAAGAGATCAAGACCCCATCGAT  600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  135  GACATGGACATCTGGAGCCTCCTCCACCTCTGTTAACCAAGAGATCAAGACCCCATCGAT  76
```

**Figure 10. Sub1A sequence alignment between Ciherang-Sub1 and Sub1A exon2 on NCBI database with the two gRNAs marked as yellow: (1) gRNA1: CCGGCGAGGAGGCTGTCCATCAC; (2) gRNA2: ACGGCCGCTGCCGGATGCCGTGG.**

*Plasmid Isolation and Biolistic Bombardment*

The transformed *E. coli* competent cell were cultured overnight in 5 ml liquid LB media with 5 ul kanamycin as selecting agent, incubated at 37°C at 125 rpm. Afterwards plasmids were isolated by QIAprep Spin Miniprep Kit which can then be stored in -20°C freezer for up to 2 weeks before using. The biolistic bombardment was performed using the PDS-1000/He particle bombardment system with 1.0 um gold particle at a helium pressure of 1,100 psi with a target distance of 6 cm.

The bombardment procedures followed the Biolistic® PDS-1000/He Particle Delivery System user manual (http://www.bio-rad.com/webroot/web/pdf/lsr/literature/M1652249.pdf). Before coating gold particles with DNA, 30 mg of 1.0 μm gold particles were sterilized with 70% EtOH and resuspended with 50% glycerol. The gold particles can be store at -20°C for up to 2 weeks. For the DNA coating, all steps need to be done while vortexing continually. For each of 50 μl gold particle, brought 5 μg of DNA to a total volume of 55 μl with 2.5M icy $CaCl_2$ and added 20 μl 0.1M spermidine. The mix was then vortexed vigorously for 3 minutes to coat DNA onto gold particle. After that, I spun down the gold particles and discarded the solution. Gold particles were again sterilized with absolute (100%) EtOH and resuspended with 48 μl 100% EtOH by tapping. Gold particles need to be vortexed at low speed continuously before using. Macrocarriers were sterilized with 70% EtOH for 1 minute and dried on sterile filter paper. For each shot, 7 μl of gold particles were applied onto sterilized macrocarriers. In this study, two shots were applied to each plate under 1,100 psi.

### *Cultivation Before and After Bombardment*

Callus was transferred onto high-osmotic medium (OSM) (Liang et al., 2018) at least 4 h before bombardment and incubated at 23°C in dark. OSM was prepared using 4.4 g/l Murashige and Skoog Basal Medium, 5 mg/l 2,4-D, 72.87 g/l mannitol, pH adjusted to 5.8 by KOH, and 3.2 g phytagel. The OSM was autoclaved at 121°C for 15 min and poured as 40 ml aliquots into petri dishes (90 x 15 mm). OSM can be stored at 4°C in dark for up to 6 weeks. After bombardment, calli were incubated in OSM overnight in the dark at 23°C. In the next day, calli were transferred onto CIM without any antibiotic at 29°C in dark for 7-10 days recovery.

### *Selection and Regeneration of Transformed Callus*

The selection and regeneration protocol was modified mainly from two previous studies (Hiei & Komari, 2008; Sahoo et al., 2011). After recovery, callus were transferred onto selection medium (MSM), which was prepared using CIM supplied with 50 mg/l hygromycin, and cultured at 27°C in dark for 12 days. MSM can be stored in dark at 4°C for up to 4 weeks. After the first selection, brown or black callus were removed and only yellow or creamish calli were transferred onto fresh MSM for the second selection for 10-12 days. After the second selection, yellow or creamish calli should propagate some small new callus. These calli can be transferred onto regeneration medium in petri dishes. If the callus condition is ambiguous, the third selection can be performed optionally.

Regeneration can be divided into two phases, shooting and rooting. After the selection process, calli were transferred onto shooting medium (R1), which has composition of 4.4 g/l Murashige and Skoog Basal Medium, 30 g/l maltose, 3 mg/l kinetin, 0.1 mg/l 1-Naphthaleneacetic acid (NAA), and 7 g/l agarose. The pH was adjusted to 5.8 with 1N NaOH and autoclaved at 121°C for 15 min. After autoclaving, medium needs to be cooled down to 50°C and 30 mg/l hygromycin was added. The medium was poured as 40 ml aliquots into petri dishes (90 x 15 mm) or 100 ml aliquots into Magenta GA-7 vessels. In the first stage of shooting, calli were transferred onto R1 with petri dishes for 14 days until the green shoots reached the lid. After that, callus with green shoots was transferred into fresh R1 with Magenta GA-7 vessel for 7-10 days until the plants had more than 3 tillers and more than 5 cm height. For regeneration, callus were cultured at 29°C with day/ night period 14/ 10 hr.

For root development, plants were transferred to rooting medium (R2), containing 2.2 g/l Murashige and Skoog Basal Medium, 30 g/l sucrose, 3 g/l phytagel with the pH of 5.8. Thereafter, 30 mg hygromycin was added after autoclaving at 121°C for 15 min. Medium was poured as 100-ml aliquots into Magenta GA-7 vessels and can be stored in dark at room temperature for up to 2 weeks. At the first stage of rooting, plants were cultured in R2 for 7-10 days until the leaves reached the lid. After that, the lid was removed, and replaced by an empty, inverted Magenta GA-7 vessel to allow plant to have more growing space. After 5-7 days, the leaves will reach the ceiling; at this point the plant can be removed from the medium. The roots were cleaned properly with tap water, removed of all the medium residues, and put in a Magenta GA-7 vessel

with fresh tap water to cover all the roots. For the 7-10 days acclimation period, plants were kept under cool (23-25°C), well ventilated condition (no need to cover), and protected from direct sunlight. During the acclimation period, I made sure water was clean and covered all the roots. If fungus growth was noticed, the water was changed right away. After the new roots appeared, the plants were transferred to soil, with Metro Mix 820 potting Mix and Turface Athletics MVP 1:1 ratio mix.

## Results

### *CRISPR/ Cas9 Vector Construction*

After cloning gRNAs into two Golden Gate entry vectors, pYPQ131C and pYPQ132C, three single colonies of each vector were cultured in liquid LB with tertracycline and isolated plasmids for Sanger sequencing. The sequencing results of the three vectors were compared with its corresponding vector to confirm that the gRNA was successfully cloned into the vectors. In entry vector, the nucleotides in lower case indicated the insertion region for gRNA. Once gRNA was ligased, the region should be replaced by the gRNA sequence (Fig. 11). The results showed that three gRNA1 vectors, gRNA1-1, 1-2, and 1-3, were ligased into pYPQ131C entry vector (Fig. 11A). For gRNA2, two out of three vectors were ligased into pYPQ132C entry vector, while gRNA2-3 had no gRNA inserted (Fig. 11B).


**Figure 11. Sanger sequencing results of the two gRNAs cloning into the entry vectors. Yellow marks indicated the gRNA ligased sequences and the crossed-out regions were the regions replaced by the gRNA insertions. (A) Alignment of three gRNA1 sequences with its entry vector pYPQ131C. (B) Alignment of three gRNA2 sequences with its entry vector pYPQ132C.**

gRNA1-1 and gRNA2-1 were assembled into pYPQ142 recipient vector, selected with blue-white screen, and validated with restriction enzyme digestion. The enzyme digestion results showed that only α2 and Hα2 were assembled successfully (Fig. 12) since α1 had no digestion occurred and Hα1 had non-expected bands appeared. pYPQ142 vector α2, Cas9 vector pYPQ150, and binary vector pMDC32 were assembled through LR reaction. The final destination vector was pMDC32 containing gRNA expression cassette and Cas9 expression sequence (Fig. 13). The final expression vector pMDC32 was validated by restriction enzyme digestion. The expected fragments of *EcoRI* and *EcoRV* digestion are 9760 bp, 4756 bp, 1043 bp, and 683; and 6059 bp, 5432 bp, 2624 bp, and 1646 bp. The results showed that only sample1 had the expected fragments in both *EcoRI* and *EcoRV* digestion (Fig. 14). Therefore, only sample1 was successfully assembled and could be used for further processes. Sample1 was stored at -80°C and was used to prepare plasmids for biolistic bombardment.

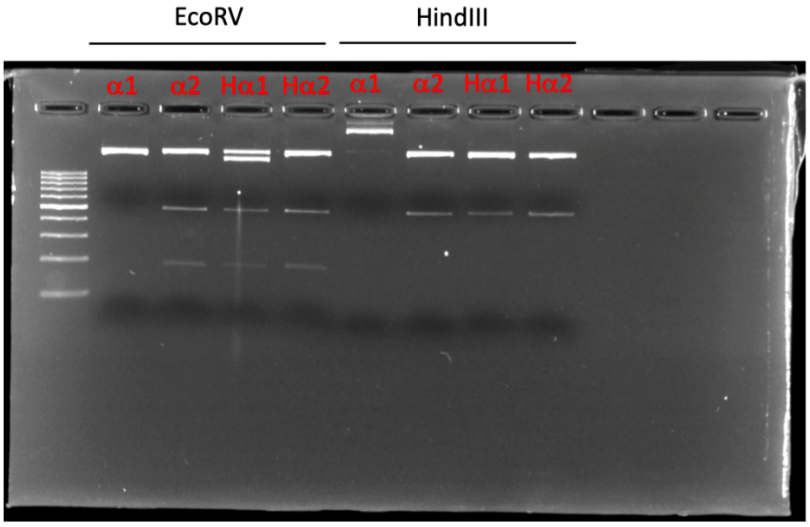


**Figure 12. Validation of *E. coli* recipient vector (pYPQ142) using restriction enzyme. The expected fragments of *EcoRV* digestion were 3186 bp, 481 bp, and 183 bp; while the expected fragments of *HindIII* digestion were 3369 bp and 481 bp.**

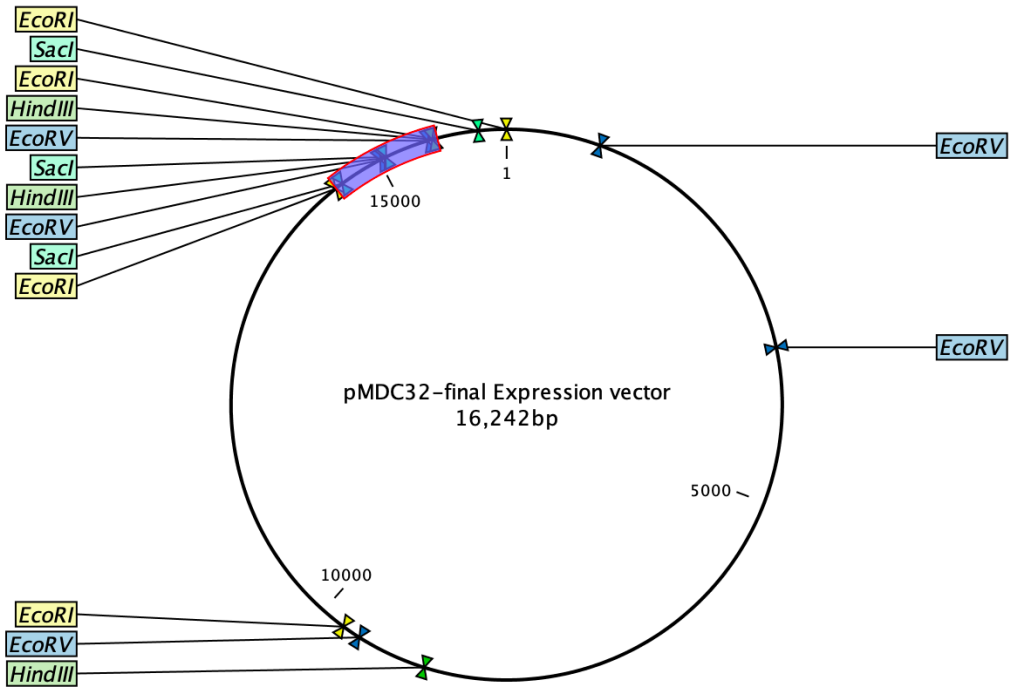


**Figure 13. Illustration of restriction enzyme cutting sites and expected sizes. The purple box indicated the gRNA expression cassette (pYPQ142) insertion region. If the gRNA expression cassette inserted successfully, the expected fragments of *EcoRI* digestion were 9760 bp, 4756 bp, 1043 bp, and 683; and 10180 bp and 1909 bp if no LR recombination happened. The expected sizes for *EcoRV* fragments were 6059 bp, 5432 bp, 2624 bp, 1646 bp, and 481 bp; and 6059 bp, 3406 bp, and 2624 bp for non-recombinants. The expected sizes for *HindIII* fragments were 9612 bp, 6149 bp, and 481; and 16242 bp for non-recombinants. For *SacI*, the expected sizes were 480 bp, 906 bp, and 14856 bp; and 16242 bp for non-recombinants.**

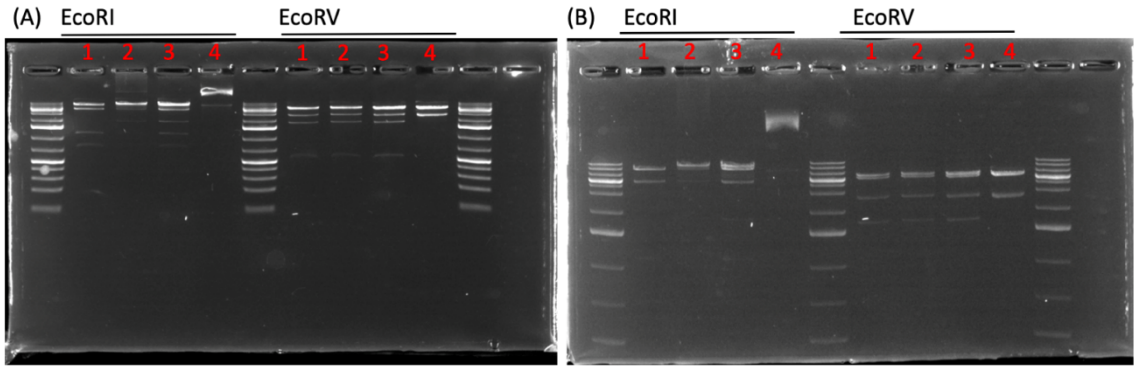


**Figure 14. Validation of *E. coli* destination vector (pMDC32) using different restriction enzymes. (A) 1.3% agarose 20 min at 135V. (B) 1.3% agarose 40 min at 135V.**

*Optimization of Callus Induction Medium and Callus Status*

This study compared six different compositions of CIM including (1) with 3 mg 2,4-D and 0.25 mg 6BA (Sahoo et al., 2011), (2) 2.5 mg 2,4-D and 0.15 mg 6BA (Sahoo & Tuteja, 2012), (3) 3 mg 2,4-D and 0.2 mg 6BA, (4) 3 mg 2,4-D and 0.2 mg 6BA with phytagel increased, (5) with 0.75 mg kinetin as cytokinin source instead of 6BA, and (6) 3 mg 2,4-D and 0.3 mg 6BA. Callus induced from 2.5 mg 2,4-D and 0.15 mg 6BA and 3 mg 2,4-D and 0.2 mg 6BA, were smaller (Fig. A4) and slower comparing with the optimal recipe. Some of callus induced from 3 mg 2,4-D and 0.25 mg 6BA were tended to be darker (Fig. A4). Calli induced from 0.75 mg kinetin medium grew faster than other media but it tended to generate more small pieces rather than a bigger calli with pretty shape. As for increasing phytagel from 3 g/l to 4 g/l, the results did not show difference between two groups.

Comparing with 5 other recipes above, this study suggested that medium with 3 mg 2,4-D and 0.3 mg 6BA (see Method section) has the best callus induction rate of Ciherang-Sub1. Callus induced from 2.5 mg 2,4-D and 0.15 mg 6BA were smaller; and callus induced from 3 mg 2,4-D and 0.25 mg 6BA tended to be darker (Fig. A4) and also generated more small pieces rather than a bigger calli with pretty shape. At the first culturing, almost all the seeds can be induced for embryogenic callus (Fig. 15A). This study also suggested that removing roots, sub-culturing and cutting callus into small pieces can facilitate callus propagation. More than 80% of seeds had propagated new callus after 16 days of sub-culturing in CIM (Fig. 15B). However, not all of the calli can be used to perform transformation. There were about 10% of calli turned brown and dry,

which cannot be used (Fig. 15C). The optimal callus for transformation should have creamy or yellow color, with visible differentiated shape on the surface, and cannot be too wet, such as juicy surface or visible water drop on the surface, or too fragile, which can be easily crushed by tweezers (Fig. 15D).

In order to confirm whether the calli were ideal for transformation and validate the biolistic bombardment delivery system, calli were observed and screened under microscope to perform biolistic bombardment. The selected calli were transformed with pPTN-EYFP control vector via biolistic bombardment. The results showed that the green fluorescence can be observed under fluorescence microscope 5 days after biolistic bombardment (Fig. 16). The result indicated that the delivery protocol was ready to use and the callus condition was optimal for transformation.

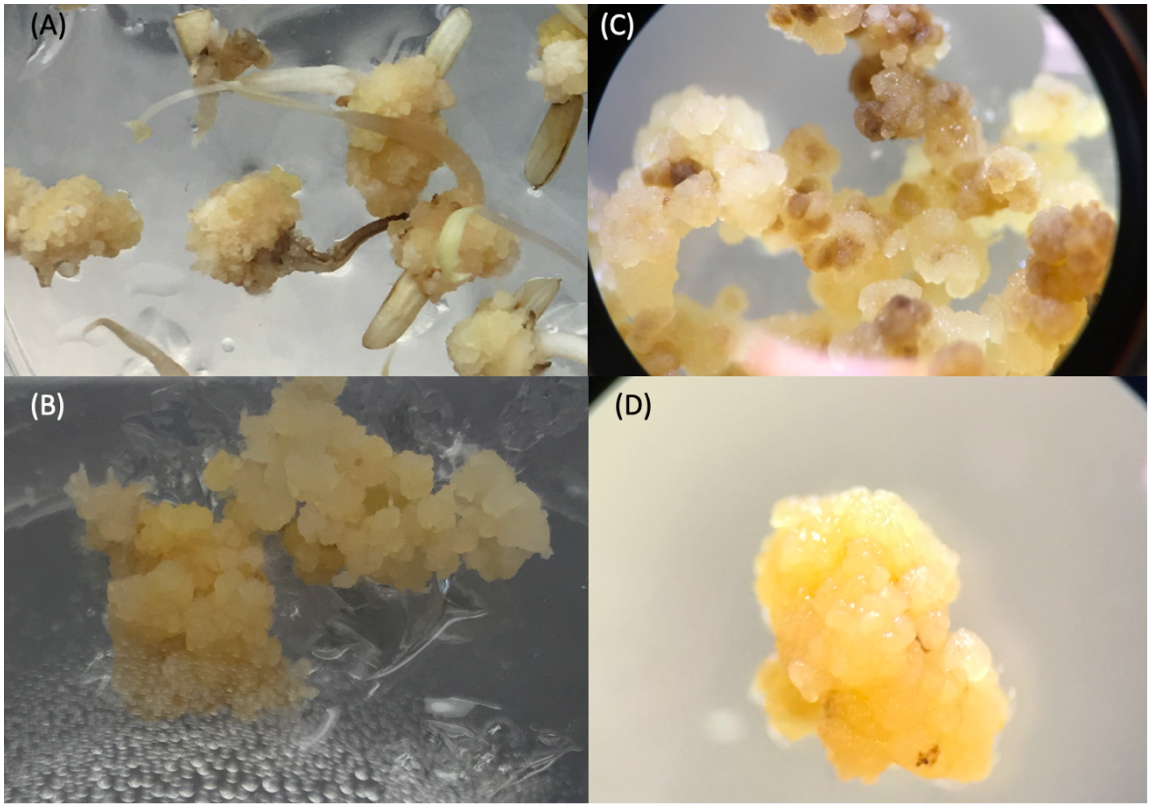


**Figure 15. Callus induction and the optimal callus for bombardment. (A) Callus induced from mature seeds after 14 days cultured in CIM. (B) Callus after 16 days sub-cultured in CIM. (C) Observation of callus under microscope. (D) The optimal callus shape for biolistic bombardment.**

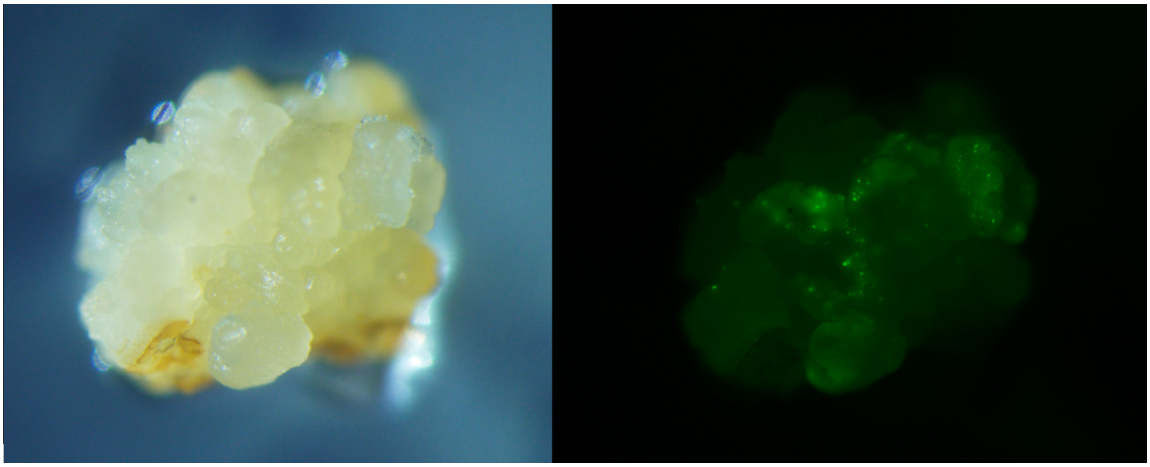

**Figure 16. Bombardment with pPTN-EYFP control vector under the fluorescence microscope.**

*Optimization of Regeneration Medium*

For testing the regeneration medium, three different formulas were tested including (1) 0.2 mg kinetin and 2 mg NAA, (2) 0.1 mg kinetin and 3 mg NAA, and (3) 0.1 mg kinetin and 3 mg NAA with agarose reduced from 8 g/l to 7 g/l. The optimal calli without transformation were selected under microscope and transferred onto regeneration medium, without any antibiotic added, to calculate the regeneration rate. With 0.2 mg kinetin and 2 mg NAA medium, green tissues were observed as early as 10 days after transferred onto regeneration medium; however, some calli became dry and dark after turning green in this medium. The results suggested that the regeneration medium (R1) with 0.1 mg kinetin and 3 mg NAA with 7 g agarose to solidify had the highest percentage of green callus and had a faster progress were observed.

After being transferred onto R1 medium, the yellow or creamy color callus would partially turn light green from the corner. After that, the green became darker, a part expanded, and the tiny spikes started to appear (Fig. 17). The spikes would elongate and differentiate to leaves. The earliest light green formation can be observed 5-7 days after being transferred onto the R1 medium. We observed 19 out of 20 calli were regenerated after 5-14 days after being transferred onto R1 medium (Fig. 17A-C), indicating that we have a 95% regeneration rate of Ciherang-Sub1. However, not all the calli turned to light green can develop spikes, and not all the spikes can differentiate leaves. We also observed that about 5-10% of green callus turned brown and died eventually. Notably, once the spikes elongate and the leaf formation appears (Fig. 17D), they will not turn brown.

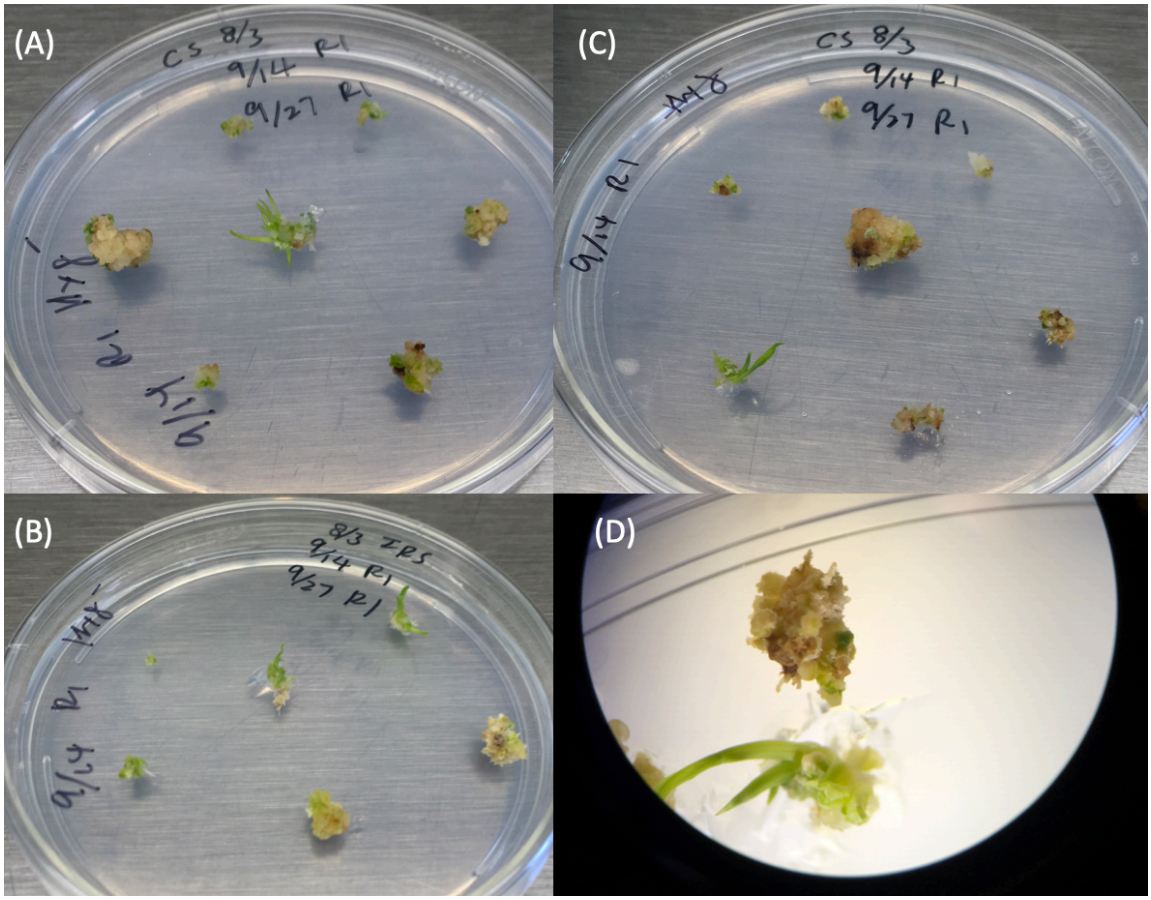


**Figure 17. (A), (B), (C) Ciherang-Sub1 callus in regeneration medium without any antibiotic. (D) Observation of regeneration of callus under microscope.**

*Confirmation of Transgenic Plants*

After bombardment and recovery, from calli that survived under three times of selection processes and regenerated from R1, R2 regeneration media with hygromycin added, three plants were obtained. After taking out plants from R2 medium, plants were transferred to clean water for acclimation. In the meantime, 2 cm leaf pieces from each plant were cut and used for DNA samples. Two leave samples were collected from plant1, and DNA was extracted with a modified CTAB method (Liang, Baring, Wang, & Septiningsih, 2017). Two selective markers, hygromycin phosphotransferase (HPT) and Cas9 sequence for pYPQ150 vector, were used to confirm the transformation. The expected product size of marker HPT375 is 375 bp, while the expected product size of Cas9 marker is 218 bp. Ciherang, Ciherang-Sub1 and autoclaved water were used as negative control, whereas an *E. coli* destination vector, which had been validated with restriction enzyme digestion, was used as positive control. DNA of transgenic samples, positive control and negative controls were diluted to 100-150 ng for the PCR reaction with annealing temperature of 55°C. The results showed that all the negative controls had no PCR products amplified, while positive control, plant2, plant3, and both leaf samples of plant1 had amplified PCR products with the expected sizes (Fig. 18).

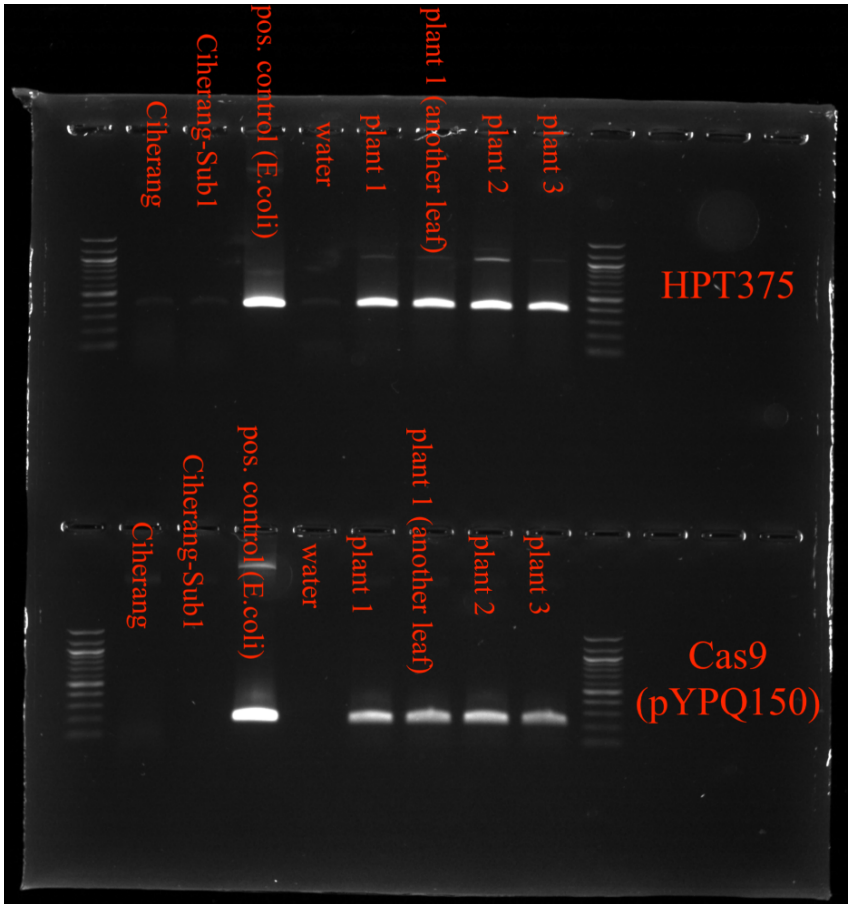


**Figure 18. Transformation validation with selective marker hygromycin phosphotransferase (HPT) and Cas9 sequence.**

## Discussion

### *Optimization of Tissue Culture and Regeneration Protocol of Ciherang-Sub1 Using Mature Seeds*

The optimal protocol of Ciherang-Sub1 is illustrated in Figure 19. It takes 80 days to obtain a transgenic plant from mature seed to transfer to soil. For this improved protocol, several medium components were modified from current japonica and indica rice tissue culture protocols (Hiei & Komari, 2008; Sahoo et al., 2011). In callus induction stage, maltose is the optimal carbon sucrose of Ciherang-Sub1. We observed that callus turned brown, and dry and the shoot grew longer when using sucrose instead of maltose (Fig. 20A-B). Previous studies had revealed not only different plant hormone ratio but also different carbon sources have significant effects on tissue culture due to different reasons (Mendoza & Kaeppler, 2002), for example, sucrose cause greater osmotic potential than maltose and has different organogenesis rate (de Paiva Neto & Otoni, 2003). Previous studies also showed that replacement of sucrose with maltose can improve callus induction for spring wheat (*Triticum aestivum* L.) and indica rice (Orshinsky, McGregor, Johnson, Hucl, & Kartha, 1990; Sahoo et al., 2011). However, for the callus induction in japonica rice cultivars, sucrose is used as the carbon source (Hiei & Komari, 2008).

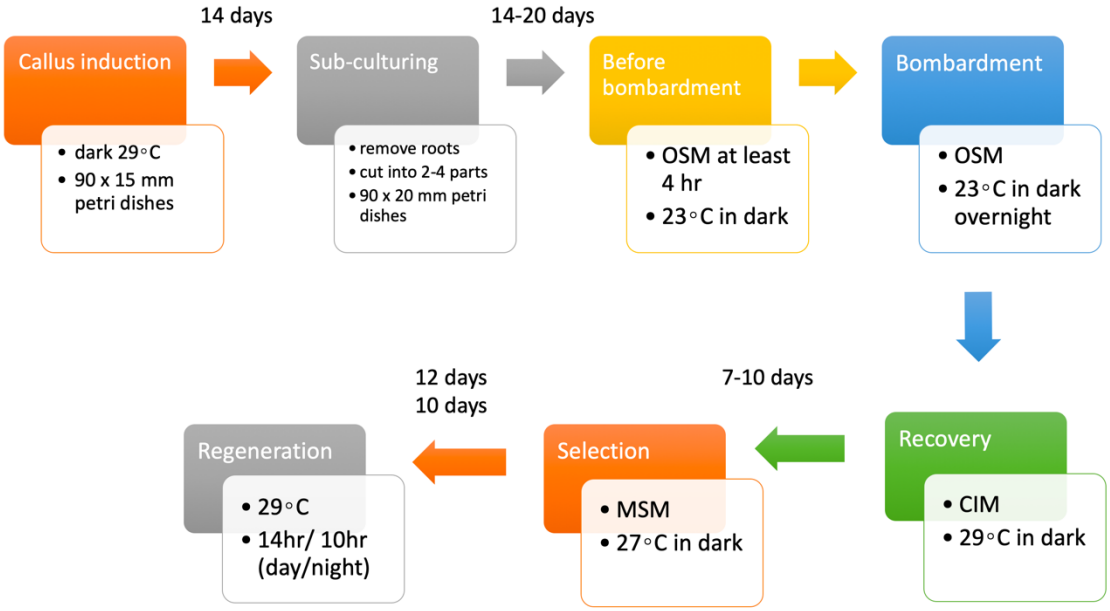


**Figure 19. Workflow of the transformation protocol in Ciherang-Sub1.**

In organogenesis, auxin and cytokinin are the major phytohormones to control the different organogenesis stages, for example, auxin to cytokinin ratio equal to 3:0.2 will induce callus, 0.03:1 can induce shoot regeneration and 3.0:0.02 can induce roots (Skoog & Miller, 1957).

For callus induction, I recommend auxin to cytokinin ratio of 3:0.3 as optimal for Ciherang-Sub1. Compared with other indica cultivars such IR64 and IR64-Sub1, Ciherang-Sub1 has slower callus formation and propagation processes. Therefore, I slightly increased the concentration of 2,4-D, a synthetic auxin, and BAP, a cytokinin source, separately from the Sahoo *et al.* (2011) protocol. With higher 2,4-D supplied, the callus became smaller and there were more lateral roots differentiated from the main root derived from mature seeds. On the other hand, when I slightly increased 6BA concentration from 0.25 mg/l to 0.3 mg/l, the callus formation and propagation were more vigorous. This study also considers that callus with ideal shape developed after 30 days of callus induction is the optimal callus for biolistic bombardment.

Previous studies have revealed that low auxin/cytokinin ratio promotes shoot formation from callus (Gaspar et al., 1996; Skoog & Miller, 1957; Skoog & Tsui, 1948). Therefore, I increased the kinetin concentration and reduced the NAA concentration, a natural auxin, in the regeneration medium (R1). I also noticed that with agarose concentration reduced from 8 g/l to 7 g/l can also improve the callus regeneration of Ciheramg-Sub1. During the testing process, we also tested the possibility of inducing root before shoot regeneration. However, the result showed that callus, without shoot regenerated, died after being transferred to rooting medium directly (Fig. 20C).

Previous studies apply fungicide (Benomyl) to root or culture in hydroponic solution after plants are moved out from rooting medium (Deb & Imchen, 2010; Hauptmann, Widholm, & Paxton, 1985; Thielges & Hoitink, 1972). In this study, I observed that fresh tap water is adequate for rice acclimation. The new roots can be observed 5-7 days after culturing in water (Fig. 20D). For transferring to soil, the air temperature and aeration are critical points. Air temperature that is too high or poor aeration will cause plant death.

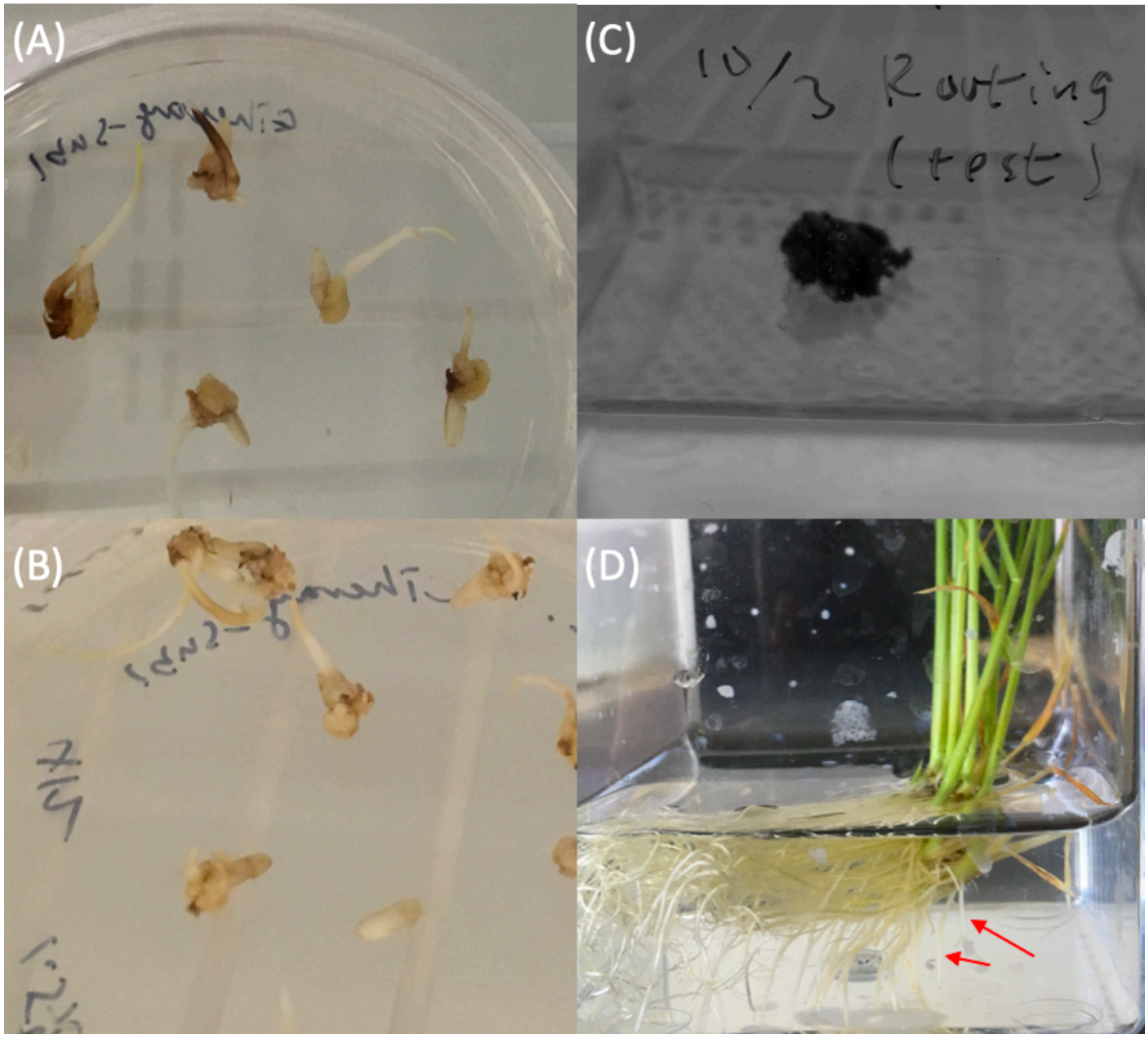


**Figure 20. Optimization of protocol. (A), (B) Callus induction with sucrose supplied. (C) Calli was transferred onto rooting medium before shoot formation. (D) Plant acclimation in fresh tap water. Red arrows indicate the new grown roots.**

### *Agrobacterium-mediated Transformations Using this Protocol*

I also tried to use the optimal protocol and callus for agrobacterium-mediated transformation with *Agrobacterium tumefaciens* EHA105 strain. However, no callus was propagated under the selection medium. Previous studies show that the callus age has crucial impact on transformation efficiency. In japonica rice and Basmati indica rice, the ideal callus age is 3 weeks (Hiei & Komari, 2008; Rashid, Yokoi, Toriyama, & Hinata, 1996), but is 8 weeks callus for IR64 and IR72 indica rice (Kumar et al., 2005). Therefore, the optimal callus age for agrobacterium-mediated transformation needs further investigation.

## Conclusion

This study demonstrated that the auxin/cytokinin ratio of 3:0.3 is the optimal condition for callus induction; while the auxin/cytokinin ratio of 0.1:3 is the optimal condition for shoot regeneration. Leaf samples were collected from the regenerated plants for DNA extraction. The confirmation of transformation was validated by two selection markers, the hygromycin resistance gene, hygromycin phosphotransferase (HPT), and the Cas9. Sanger sequencing to reveal the gene-edited region is currently underway. The results showed that 28-35 day-old calli were ideal for particle bombardment, and the regeneration medium (R1) had a regeneration rate of 95%. Using this protocol, CRISPR gene-edited plants can be generated and validated within 12 weeks.

# CHAPTER V

# CONCLUSIONS

This dissertation provided some insight into the molecular regulation and genomics of drought and submergence stresses in rice. In the drought stress part, the results suggested that under moderate drought condition during reproductive stage, rice genotype 4610 had better performance than Rondo, possessing higher spikelet fertility, hundred-seed weight, and yield. The transcriptomics analysis results revealed that the drought tolerance of 4610 can be partially attributed by the higher expression of drought stress response molecular mechanisms such as the LEA proteins, HSPs and ROS scavengers such as APXs and GSTs. In the second part, to better understand the target QTL for submergence tolerance during vegetative stage, *qSub8.1*, whole genome sequencing of three varieties was performed. Based on the genome assembly of the three varieties, Ciherang-Sub1, Ciherang, and IR64-Sub1, the results inferred that Ciherang-Sub1 genome is composed of 59% of Ciherang, 24% of IR64-Sub1, and 17% of unknown sources. According to the SNPs profiles, the *qSub8.1* QTL region of Ciherang-Sub1 is mainly from Ciherang background with a few introgressed segments from IR64-Sub1 and unknown sources. The SNPs flanking the *qSub8.1* region can be used to design DNA markers for further characterization of the *qSub8.1* region. In addition, in preparation of the gene validation of the candidate genes underlying the *qSUB8.1* QTL, an experiment to optimize transformation and transgenic development using Ciherang-Sub1 was conducted. In summary, a protocol which allows us to generate CRISPR gene-

edited plants within 12 weeks was developed with helium-driven biolistic bombardment delivery system. This is the first transformation and tissue culture protocol reported for Ciherang-Sub1 genetic background using mature seed as explants. This protocol demonstrates that the auxin to cytokinin ratio of 3:0.3 is the optimal condition for callus induction for mature seed; while the auxin to cytokinin ratio of 0.1:3 is the optimal condition for shoot regeneration.

# REFERENCES


Aldemita, R. R., & Hodges, T. K. (1996). Agrobacterium tumefaciens-mediated transformation of japonica and indica rice varieties. *Planta, 199*(4), 612-617. doi:10.1007/BF00195194

Bae, S., Park, J., & Kim, J. S. (2014). Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics, 30*(10), 1473-1475. doi:10.1093/bioinformatics/btu048

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics, 30*(15), 2114-2120. doi:10.1093/bioinformatics/btu170

Bono, H., Hino, K., Ui-Tei, K., & Naito, Y. (2014). CRISPRdirect: software for designing CRISPR/Cas guide RNA with reduced off-target sites. *Bioinformatics, 31*(7), 1120-1123. doi:10.1093/bioinformatics/btu743

Bouman, B. A. M., Peng, S., Castañeda, A. R., & Visperas, R. M. (2005). Yield and water use of irrigated tropical aerobic rice systems. *Agricultural Water Management, 74*(2), 87-105. doi:10.1016/j.agwat.2004.11.007

Cai, Y., Chen, L., Liu, X., Sun, S., Wu, C., Jiang, B., . . . Hou, W. (2015). CRISPR/Cas9-Mediated Genome Editing in Soybean Hairy Roots. *PLOS ONE, 10*(8), e0136064. doi:10.1371/journal.pone.0136064

Chaves, M. M., Pinheiro, C., & Flexas, J. (2008). Photosynthesis under drought and salt stress: regulation mechanisms from whole plant to cell. *Annals of Botany, 103*(4), 551-560. doi:10.1093/aob/mcn125

Chen, J.-Q., Meng, X.-P., Zhang, Y., Xia, M., & Wang, X.-P. (2008). Over-expression of OsDREB genes lead to enhanced drought tolerance in rice. *Biotechnology Letters, 30*(12), 2191-2198. doi:10.1007/s10529-008-9811-5

Cheong, Y. H., Kim, K.-N., Pandey, G. K., Gupta, R., Grant, J. J., & Luan, S. (2003). CBL1, a Calcium Sensor That Differentially Regulates Salt, Drought, and Cold Responses in Arabidopsis. *The Plant Cell, 15*(8), 1833. doi:10.1105/tpc.012393

Chool Boo, Y., & Jung, J. (1999). Water Deficit — Induced Oxidative Stress and Antioxidative Defenses in Rice Plants. *Journal of Plant Physiology, 155*(2), 255-261. doi:https://doi.org/10.1016/S0176-1617(99)80016-9

Choudhury, F. K., Rivero, R. M., Blumwald, E., & Mittler, R. (2017). Reactive oxygen species, abiotic stress and stress combination. *Plant J, 90*(5), 856-867. doi:10.1111/tpj.13299

Corratge-Faillie, C., Ronzier, E., Sanchez, F., Prado, K., Kim, J. H., Lanciano, S., . . . Xiong, T. C. (2017). The Arabidopsis guard cell outward potassium channel GORK is regulated by CPK33. *FEBS Lett, 591*(13), 1982-1992. doi:10.1002/1873-3468.12687

Cruz de Carvalho, M. H. (2008). Drought stress and reactive oxygen species. *Plant Signaling & Behavior, 3*(3), 156-165. doi:10.4161/psb.3.3.5536

Cruz, R. P. d., Milach, S. C. K., & Federizzi, L. C. (2006). Inheritance of rice cold tolerance at the germination stage. *Genetics and Molecular Biology, 29*(2), 314-320.

Dalal, M., Tayal, D., Chinnusamy, V., & Bansal, K. C. (2009). Abiotic stress and ABA-inducible Group 4 LEA from Brassica napus plays a key role in salt and drought tolerance. *Journal of Biotechnology, 139*(2), 137-145. doi:https://doi.org/10.1016/j.jbiotec.2008.09.014

Das, K., & Roychoudhury, A. (2014). Reactive oxygen species (ROS) and response of antioxidants as ROS-scavengers during environmental stress in plants. *Frontiers in Environmental Science, 2*(53). doi:10.3389/fenvs.2014.00053

de Paiva Neto, V. B., & Otoni, W. C. (2003). Carbon sources and their osmotic potential in plant tissue culture: does it matter? *Scientia Horticulturae, 97*(3), 193-202. doi:https://doi.org/10.1016/S0304-4238(02)00231-5

Deb, C., & Imchen, T. (2010). An Efficient In vitro Hardening Technique of Tissue Culture Raised Plants. *Biotechnology, 9*(1), 79-83.

Doyle, J. J. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull, 19*, 11-15.

Du, H., Wang, N., Cui, F., Li, X., Xiao, J., & Xiong, L. (2010). Characterization of the β-Carotene Hydroxylase Gene <em>DSM2</em> Conferring Drought and Oxidative Stress Resistance by Increasing Xanthophylls and Abscisic Acid Synthesis in Rice. *Plant Physiology, 154*(3), 1304. doi:10.1104/pp.110.163741

Fannin, B. (2017). Texas agricultural losses from Hurricane Harvey estimated at more than $200 million. *AgriLife TODAY*. Retrieved from https://today.agrilife.org/2017/10/27/texas-agricultural-losses-hurricane-harvey-estimated-200-million/

FAO, IFAD, UNICEF, WFP and WHO. 2018. The State of Food Security and Nutrition in the World 2018. Building climate resilience for food security and nutrition. Rome, FAO.

Farooq, M., Wahid, A., Kobayashi, N., Fujita, D., & Basra, S. M. A. (2009). Plant Drought Stress: Effects, Mechanisms and Management. In E. Lichtfouse, M. Navarrete, P. Debaeke, S. Véronique, & C. Alberola (Eds.), *Sustainable Agriculture* (pp. 153-188). Dordrecht: Springer Netherlands.

Farrell, T., Fox, K., Williams, R., Fukai, S., & Lewin, L. (2006). Minimising cold damage during reproductive development among temperate rice genotypes. II. Genotypic variation and flowering traits related to cold tolerance screening. *Australian Journal of Agricultural Research, 57*(1), 89-100.

Foyer, C. H., & Noctor, G. (2011). Ascorbate and glutathione: the heart of the redox hub. *Plant Physiol, 155*(1), 2-18. doi:10.1104/pp.110.167569

Fukao, T., & Bailey-Serres, J. (2004). Plant responses to hypoxia – is survival a balancing act? *Trends in Plant Science, 9*(9), 449-456. doi:https://doi.org/10.1016/j.tplants.2004.07.005

Garrity, D. P., & O'Toole, J. C. (1994). Screening rice for drought resistance at the reproductive phase. *Field Crops Research, 39*(2), 99-110. doi:https://doi.org/10.1016/0378-4290(94)90012-4

Gaspar, T., Kevers, C., Penel, C., Greppin, H., Reid, D. M., & Thorpe, T. A. (1996). Plant hormones and plant growth regulators in plant tissue culture. *In Vitro Cellular & Developmental Biology - Plant, 32*(4), 272-289. doi:10.1007/BF02822700

Gelvin, S. B. (2003). Agrobacterium-Mediated Plant Transformation: the Biology behind the "Gene-Jockeying" Tool. *Microbiology and Molecular Biology Reviews, 67*(1), 16. doi:10.1128/MMBR.67.1.16-37.2003

Glinski, J. (2018). *Soil Physical Conditions and Plant Roots* (1 ed.). Boca Raton: CRC Press.

Gonzaga, Z. J. C., Carandang, J., Singh, A., Collard, B. C. Y., Thomson, M. J., & Septiningsih, E. M. (2017). Mapping QTLs for submergence tolerance in rice using a population fixed for SUB1A tolerant allele. *Molecular Breeding, 37*(4), 47. doi:10.1007/s11032-017-0637-5

Goyal, K., Walton, L. J., & Tunnacliffe, A. (2005). LEA proteins prevent protein aggregation due to water stress. *The Biochemical Journal, 388*(Pt 1), 151-157. doi:10.1042/BJ20041931

Guo, Z., Ou, W., Lu, S., & Zhong, Q. (2006). Differential responses of antioxidative system to chilling and drought in four rice cultivars differing in sensitivity. *Plant*

*Physiology and Biochemistry, 44*(11), 828-836. doi:https://doi.org/10.1016/j.plaphy.2006.10.024

Haque, S. (2006). Salinity problems and crop production in coastal regions of Bangladesh. *Pakistan Journal of Botany, 38*(5), 1359-1365.

Hauptmann, R. M., Widholm, J. M., & Paxton, J. D. (1985). Benomyl: a broad spectrum fungicide for use in plant cell and protoplast culture. *Plant Cell Reports, 4*(3), 129-132.

Hiei, Y., & Komari, T. (2006). Improved protocols for transformation of indica rice mediated by Agrobacterium tumefaciens. *Plant Cell, Tissue and Organ Culture, 85*(3), 271-283. doi:10.1007/s11240-005-9069-8

Hiei, Y., & Komari, T. (2008). Agrobacterium-mediated transformation of rice using immature embryos or calli induced from mature seed. *Nature Protocols, 3*, 824. doi:10.1038/nprot.2008.46

Hiei, Y., Ohta, S., Komari, T., & Kumashiro, T. (1994). Efficient transformation of rice (Oryza sativa L.) mediated by Agrobacterium and sequence analysis of the boundaries of the T-DNA. *The Plant Journal, 6*(2), 271-282. doi:10.1046/j.1365-313X.1994.6020271.x

Hu, H., Dai, M., Yao, J., Xiao, B., Li, X., Zhang, Q., & Xiong, L. (2006). Overexpressing a NAM, ATAF, and CUC (NAC) transcription factor enhances drought resistance and salt tolerance in rice. *Proceedings of the National Academy of Sciences, 103*(35), 12987. doi:10.1073/pnas.0604882103

Ismail, A. M. (2018). Submergence tolerance in rice: resolving a pervasive quandary. *New Phytologist, 218*(4), 1298-1300. doi:10.1111/nph.15188

Juan Carlos, M., Jill, M. D., Albrigo, L. G., & James, P. S. (2010). Winter Drought Stress Can Delay Flowering and Avoid Immature Fruit Loss during Late-season Mechanical Harvesting of 'Valencia' Oranges. *HortScience horts, 45*(2), 271-276. doi:10.21273/HORTSCI.45.2.271

Khanna, H. K., & Raina, S. K. (1998). Genotype x culture media interaction effects on regeneration response of three indica rice cultivars. *Plant Cell, Tissue and Organ Culture, 52*(3), 145-153.

Kim, D., Langmead, B., & Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat Methods, 12*(4), 357-360. doi:10.1038/nmeth.3317

Kizis, D., & Pagès, M. (2002). Maize DRE-binding proteins DBF1 and DBF2 are involved in rab17 regulation through the drought-responsive element in an ABA-dependent pathway. *The Plant Journal, 30*(6), 679-689. doi:10.1046/j.1365-313X.2002.01325.x

Kumar, K. K., Maruthasalam, S., Loganathan, M., Sudhakar, D., & Balasubramanian, P. (2005). An improvedAgrobacterium-mediated transformation protocol for recalcitrant elite indica rice cultivars. *Plant Molecular Biology Reporter, 23*(1), 67-73. doi:10.1007/BF02772648

Kumar, S., & Trivedi, P. K. (2018). Glutathione S-Transferases: Role in Combating Abiotic Stresses Including Arsenic Detoxification in Plants. *Frontiers in Plant Science, 9*, 751. doi:10.3389/fpls.2018.00751

Kurtz, S., Phillippy, A., Delcher, A. L., Smoot, M., Shumway, M., Antonescu, C., & Salzberg, S. L. (2004). Versatile and open software for comparing large genomes. *Genome Biology, 5*(2), R12. doi:10.1186/gb-2004-5-2-r12

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods, 9*(4), 357-359. doi:10.1038/nmeth.1923 http://www.nature.com/nmeth/journal/v9/n4/abs/nmeth.1923.html#supplementary-information

Lata, C., & Prasad, M. (2011). Role of DREBs in regulation of abiotic stress responses in plants. *Journal of Experimental Botany, 62*(14), 4731-4748. doi:10.1093/jxb/err210

Lee, U., Wie, C., Escobar, M., Williams, B., Hong, S.-W., & Vierling, E. (2005). Genetic Analysis Reveals Domain Interactions of Arabidopsis Hsp100/ClpB and Cooperation with the Small Heat Shock Protein Chaperone System. *The Plant Cell, 17*(2), 559. doi:10.1105/tpc.104.027540

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., . . . Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics, 25*. doi:10.1093/bioinformatics/btp352

Li, L., Li, N., Song, S. F., Li, Y. X., Xia, X. J., Fu, X. Q., . . . Deng, H. F. (2014). Cloning and characterization of the drought-resistance OsRCI2-5 gene in rice

(*Oryza sativa* L.). *Genet Mol Res, 13*(2), 4022-4035. doi:10.4238/2014.May.23.13

Li, Q., Zhang, D., Chen, M., Liang, W., Wei, J., Qi, Y., & Yuan, Z. (2016). Development of japonica photo-sensitive genic male sterile rice lines by editing carbon starved anther using CRISPR/Cas9. *Journal of Genetics and Genomics*. doi:10.1016/j.jgg.2016.04.011

Liang, Y., Baring, M., Wang, S., & Septiningsih, E. M. (2017). Mapping QTLs for Leafspot Resistance in Peanut Using SNP-Based Next-Generation Sequencing Markers. *Plant Breeding and Biotechnology, 5*(2), 115-122. doi:10.9787/pbb.2017.5.2.115

Liang, Z., Chen, K., Zhang, Y., Liu, J., Yin, K., Qiu, J.-L., & Gao, C. (2018). Genome editing of bread wheat using biolistic delivery of CRISPR/Cas9 in vitro transcripts or ribonucleoproteins. *Nature Protocols, 13*, 413. doi:10.1038/nprot.2017.145 https://www.nature.com/articles/nprot.2017.145#supplementary-information

Lin, Y. J., & Zhang, Q. (2005). Optimising the tissue culture conditions for high efficiency transformation of indica rice. *Plant Cell Reports, 23*(8), 540-547. doi:10.1007/s00299-004-0843-6

Liu, C., Mao, B., Ou, S., Wang, W., Liu, L., Wu, Y., . . . Wang, X. (2014). OsbZIP71, a bZIP transcription factor, confers salinity and drought tolerance in rice. *Plant Molecular Biology, 84*(1), 19-36. doi:10.1007/s11103-013-0115-3

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol, 15*(12), 550. doi:10.1186/s13059-014-0550-8

Lowder, L. G., Zhang, D., Baltes, N. J., Paul, J. W., Tang, X., Zheng, X., . . . Qi, Y. (2015). A CRISPR/Cas9 Toolbox for Multiplexed Plant Genome Editing and Transcriptional Regulation. *Plant Physiology, 169*(2), 971-985. doi:10.1104/pp.15.00636

Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., . . . Liu, Y. (2012). SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *GigaScience, 1*. doi:10.1186/2047-217x-1-18

Mao, Y., Botella, J. R., Liu, Y., & Zhu, J.-K. (2019). Gene Editing in Plants–Progress and Challenges. *National Science Review*.

Mather, K. A., Caicedo, A. L., Polato, N. R., Olsen, K. M., McCouch, S., & Purugganan, M. D. (2007). The extent of linkage disequilibrium in rice (*Oryza sativa* L.). *Genetics, 177*(4), 2223-2232. doi:10.1534/genetics.107.079616

McDowell, N., Pockman, W. T., Allen, C. D., Breshears, D. D., Cobb, N., Kolb, T., . . . Yepez, E. A. (2008). Mechanisms of plant survival and mortality during drought: why do some plants survive while others succumb to drought? *New Phytologist, 178*(4), 719-739. doi:10.1111/j.1469-8137.2008.02436.x

Mendoza, M. G., & Kaeppler, H. F. (2002). Auxin and sugar effects on callus induction and plant regeneration frequencies from mature embryos of wheat (*Triticum*

*aestivum* L.). *In Vitro Cellular & Developmental Biology - Plant, 38*(1), 39-45. doi:10.1079/IVP2001250

Mi, H., Muruganujan, A., Huang, X., Ebert, D., Mills, C., Guo, X., & Thomas, P. D. (2019). Protocol Update for large-scale genome and gene function analysis with the PANTHER classification system (v.14.0). *Nature Protocols, 14*(3), 703-721. doi:10.1038/s41596-019-0128-8

Mian, A., Ahmad, I., & Maathuis, F. J. M. (2016). Overexpression of the rice AKT1 potassium channel affects potassium nutrition and rice drought tolerance. *Journal of Experimental Botany, 67*(9), 2689-2698. doi:10.1093/jxb/erw103

Miao, J., Guo, D., Zhang, J., Huang, Q., Qin, G., Zhang, X., . . . Qu, L.-J. (2013). Targeted mutagenesis in rice using CRISPR-Cas system. *Cell Research, 23*(10), 1233-1236. doi:10.1038/cr.2013.123

Mittler, R. (2002). Oxidative stress, antioxidants and stress tolerance. *Trends in Plant Science, 7*(9), 405-410. doi:https://doi.org/10.1016/S1360-1385(02)02312-9

Mizoguchi, T., Ichimura, K., & Shinozaki, K. (1997). Environmental stress response in plants: the role of mitogen-activated protein kinases. *Trends in Biotechnology, 15*(1), 15-19. doi:https://doi.org/10.1016/S0167-7799(96)10074-3

Mohammed, A., Cothren, J., Chen, M. H., & Tarpley, L. (2015). 1-Methylcyclopropene (1-MCP)-Induced Alteration in Leaf Photosynthetic Rate, Chlorophyll Fluorescence, Respiration and Membrane Damage in Rice (*Oryza sativa* L.) Under High Night Temperature. *Journal of Agronomy and Crop Science, 201*(2), 105-116.

Mohammed, A., & Tarpley, L. (2009). High nighttime temperatures affect rice productivity through altered pollen germination and spikelet fertility. *Agricultural and Forest Meteorology, 149*(6-7), 999-1008.

Moran, J. F., Becana, M., Iturbe-Ormaetxe, I., Frechilla, S., Klucas, R. V., & Aparicio-Tejo, P. (1994). Drought induces oxidative stress in pea plants. *Planta, 194*(3), 346-352. doi:10.1007/BF00197534

Munemasa, S., Hauser, F., Park, J., Waadt, R., Brandt, B., & Schroeder, J. I. (2015). Mechanisms of abscisic acid-mediated control of stomatal aperture. *Current Opinion in Plant Biology, 28*, 154-162. doi:https://doi.org/10.1016/j.pbi.2015.10.010

Niroula, R. K., Pucciariello, C., Ho, V. T., Novi, G., Fukao, T., & Perata, P. (2012). SUB1A-dependent and -independent mechanisms are involved in the flooding tolerance of wild rice species. *The Plant Journal, 72*(2), 282-293. doi:10.1111/j.1365-313X.2012.05078.x

Orshinsky, B. R., McGregor, L. J., Johnson, G. I. E., Hucl, P., & Kartha, K. K. (1990). Improved embryoid induction and green shoot regeneration from wheat anthers cultured in medium with maltose. *Plant Cell Reports, 9*(7), 365-369. doi:10.1007/BF00232400

Paul, J. W., 3rd, & Qi, Y. (2016). CRISPR/Cas9 for plant genome editing: accomplishments, problems and prospects. *Plant Cell Reports, 35*(7), 1417-1427. doi:10.1007/s00299-016-1985-z

Peng, S., Huang, J., Sheehy, J. E., Laza, R. C., Visperas, R. M., Zhong, X., . . . Cassman, K. G. (2004). Rice yields decline with higher night temperature from global warming. *Proceedings of the National Academy of Sciences, 101*(27), 9971-9975.

Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T.-C., Mendell, J. T., & Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nature Biotechnology, 33*, 290. doi:10.1038/nbt.3122 https://www.nature.com/articles/nbt.3122#supplementary-information

Price, A. H., Atherton, N. M., & Hendry, G. A. F. (1989). Plants Under Drought-Stress Generate Activated Oxygen. *Free Radical Research Communications, 8*(1), 61-66. doi:10.3109/10715768909087973

Rashid, H., Yokoi, S., Toriyama, K., & Hinata, K. (1996). Transgenic plant production mediated by Agrobacterium in Indica rice. *Plant Cell Reports, 15*(10), 727-730. doi:10.1007/BF00232216

Sahoo, K. K., Tripathi, A. K., Pareek, A., Sopory, S. K., & Singla-Pareek, S. L. (2011). An improved protocol for efficient transformation and regeneration of diverse indica rice cultivars. *Plant Methods, 7*(1), 49. doi:10.1186/1746-4811-7-49

Sahoo, R. K., & Tuteja, N. (2012). Development of Agrobacterium-mediated transformation technology for mature seed-derived callus tissues of indica rice cultivar IR64. *GM crops & food, 3*(2), 123-128.

Saijo, Y., Hata, S., Kyozuka, J., Shimamoto, K., & Izui, K. (2000). Over-expression of a single Ca2+-dependent protein kinase confers both cold and salt/drought

tolerance on rice plants. *The Plant Journal, 23*(3), 319-327. doi:10.1046/j.1365-313x.2000.00787.x

Sakai, H., Lee, S. S., Tanaka, T., Numa, H., Kim, J., Kawahara, Y., . . . Abe, T. (2013). Rice Annotation Project Database (RAP-DB): an integrative and interactive database for rice genomics. *Plant and Cell Physiology, 54*(2), e6-e6.

Sasaki, T., & Burr, B. (2000). International Rice Genome Sequencing Project: the effort to completely sequence the rice genome. *Current Opinion in Plant Biology, 3*(2), 138-142.

Septiningsih, E. M., Hidayatun, N., Sanchez, D. L., Nugraha, Y., Carandang, J., Pamplona, A. M., . . . Mackill, D. J. (2014). Accelerating the development of new submergence tolerant rice varieties: the case of Ciherang-Sub1 and PSB Rc18-Sub1. *Euphytica, 202*(2), 259-268. doi:10.1007/s10681-014-1287-x

Septiningsih, E. M., & Mackill, D. J. (2018). Genetics and breeding of flooding tolerance in rice *Rice Genomics, Genetics and Breeding* (pp. 275-295): Springer.

Septiningsih, E. M., Pamplona, A. M., Sanchez, D. L., Neeraja, C. N., Vergara, G. V., Heuer, S., . . . Mackill, D. J. (2009). Development of submergence-tolerant rice cultivars: the Sub1 locus and beyond. *Annals of Botany, 103*(2), 151-160. doi:10.1093/aob/mcn206

Sgherri, C. L. M., Pinzino, C., & Navari-Izzo, F. (1993). Chemical changes and $O2^{\cdot -}$ production in thylakoid membranes under water stress. *Physiologia Plantarum, 87*(2), 211-216. doi:10.1111/j.1399-3054.1993.tb00144.x

Shinozaki, K., & Yamaguchi-Shinozaki, K. (2000). Molecular responses to dehydration and low temperature: differences and cross-talk between two stress signaling pathways. *Current Opinion in Plant Biology, 3*(3), 217-223. doi:https://doi.org/10.1016/S1369-5266(00)80068-0

Skoog, F., & Miller, C. (1957). *Chemical regulation of growth and organ formation in plant tissues cultured.* Paper presented at the Vitro, Symp. Soc. Exp. Biol.

Skoog, F., & Tsui, C. (1948). Chemical Control of Growth and Bud Formation in Tobacco Stem Segments and Callus Cultured in vitro. *American Journal of Botany, 35*(10), 782-787. doi:10.1002/j.1537-2197.1948.tb08148.x

Tabien, R. E., Harper, C. L., Carre, P., & Jones, K. (2015). Reproductive Drought Tolerance of Herbicide Tolerant Rice. *Texas Rice Special Section*, 4-5. Retrieved from https://beaumont.tamu.edu/eLibrary/Newsletter/2015_Highlights_in_Research.pdf

Tenhaken, R. (2014). Cell wall remodeling under abiotic stress. *Frontiers in Plant Science, 5*, 771. doi:10.3389/fpls.2014.00771

Thielges, B. A., & Hoitink, H. A. J. (1972). Fungicides Aid Rooting of Eastern White Pine Cuttings. *Forest Science, 18*(1), 54-55. doi:10.1093/forestscience/18.1.54

Tripathy, J. N., Zhang, J., Robin, S., Nguyen, T. T., & Nguyen, H. T. (2000). QTLs for cell-membrane stability mapped in rice (*Oryza sativa* L.) under drought stress. *Theoretical and Applied Genetics, 100*(8), 1197-1202. doi:10.1007/s001220051424

van der Weijde, T., Huxley, L. M., Hawkins, S., Sembiring, E. H., Farrar, K., Dolstra, O., . . . Trindade, L. M. (2017). Impact of drought stress on growth and quality of miscanthus for biofuel production. *GCB Bioenergy, 9*(4), 770-782. doi:10.1111/gcbb.12382

Wang, Y., Cheng, X., Shan, Q., Zhang, Y., Liu, J., Gao, C., & Qiu, J.-L. (2014). Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nature Biotechnology, 32*, 947. doi:10.1038/nbt.2969 https://www.nature.com/articles/nbt.2969#supplementary-information

Xiao, B., Huang, Y., Tang, N., & Xiong, L. (2007). Over-expression of a LEA gene in rice improves drought resistance under the field conditions. *Theoretical and Applied Genetics, 115*(1), 35-46. doi:10.1007/s00122-007-0538-9

Xu, K., Xu, X., Fukao, T., Canlas, P., Maghirang-Rodriguez, R., Heuer, S., . . . Mackill, D. J. (2006). Sub1A is an ethylene-response-factor-like gene that confers submergence tolerance to rice. *Nature, 442*(7103), 705-708. doi:10.1038/nature04920

Yan, W., & McClung, A. M. (2010). 'Rondo', a Long-Grain Indica Rice with Resistances to Multiple Diseases. *Journal of Plant Registrations, 4*, 131-136. doi:10.3198/jpr2009.07.0404crc

Yu, G., Wang, L.-G., Han, Y., & He, Q.-Y. (2012). clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *OMICS: A Journal of Integrative Biology, 16*(5), 284-287. doi:10.1089/omi.2011.0118

Zaidi, M., Narayanan, M., Sardana, R., Taga, I., Postel, S., Johns, R., . . . Loit, E. (2006). Optimizing tissue culture media for efficient transformation of different indica rice genotypes. *Agron Res, 4*(2), 563-575.

Zhang, C., Liu, J., Zhao, T., Gomez, A., Li, C., Yu, C., . . . Lin, C. (2016). A Drought-Inducible Transcription Factor Delays Reproductive Timing in Rice. *Plant Physiol, 171*(1), 334-343. doi:10.1104/pp.16.01691

Zhang, J., Chen, L.-L., Xing, F., Kudrna, D. A., Yao, W., Copetti, D., . . . Zhang, Q. (2016). Extensive sequence divergence between the reference genomes of two elite <em>indica</em> rice varieties Zhenshan 97 and Minghui 63. *Proceedings of the National Academy of Sciences, 113*(35), E5163. doi:10.1073/pnas.1611012113

Zhang, J., Zhang, S., Cheng, M., Jiang, H., Zhang, X., Peng, C., . . . Jin, J. (2018). Effect of Drought on Agronomic Traits of Rice and Wheat: A Meta-Analysis. *Int J Environ Res Public Health, 15*(5). doi:10.3390/ijerph15050839

Zheng, X., Chen, B., Lu, G., & Han, B. (2009). Overexpression of a NAC transcription factor enhances rice drought and salt tolerance. *Biochemical and Biophysical Research Communications, 379*(4), 985-989. doi:https://doi.org/10.1016/j.bbrc.2008.12.163

Zou, J.-J., Wei, F.-J., Wang, C., Wu, J.-J., Ratnasekera, D., Liu, W.-X., & Wu, W.-H. (2010). Arabidopsis Calcium-Dependent Protein Kinase CPK10 Functions in Abscisic Acid- and Ca<sup>2+</sup>-Mediated Stomatal Regulation

in Response to Drought Stress. *Plant Physiology, 154*(3), 1232. doi:10.1104/pp.110.157545

APPENDIX A

Table A1. Effect of variety on yield-related traits under drought treatment.

| | 4610 | | Rondo | |
|---|---|---|---|---|
| | Irrigated | Drought | Irrigated | Drought |
| Panicle number | 49.7 | 36.7 | 47.7 | 43.0 |

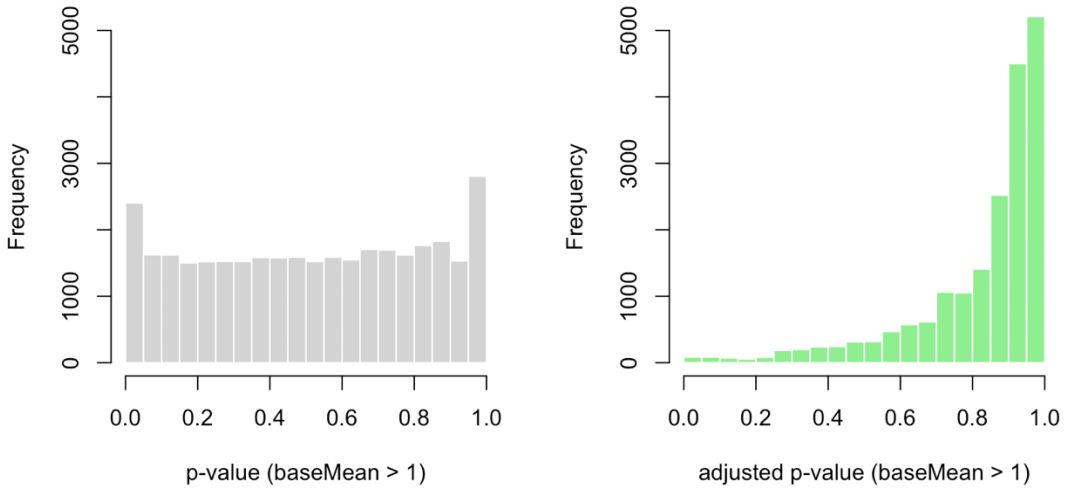


Figure A1. The p-value distribution of likelihood ratio test for genotype-by-treatment effect.

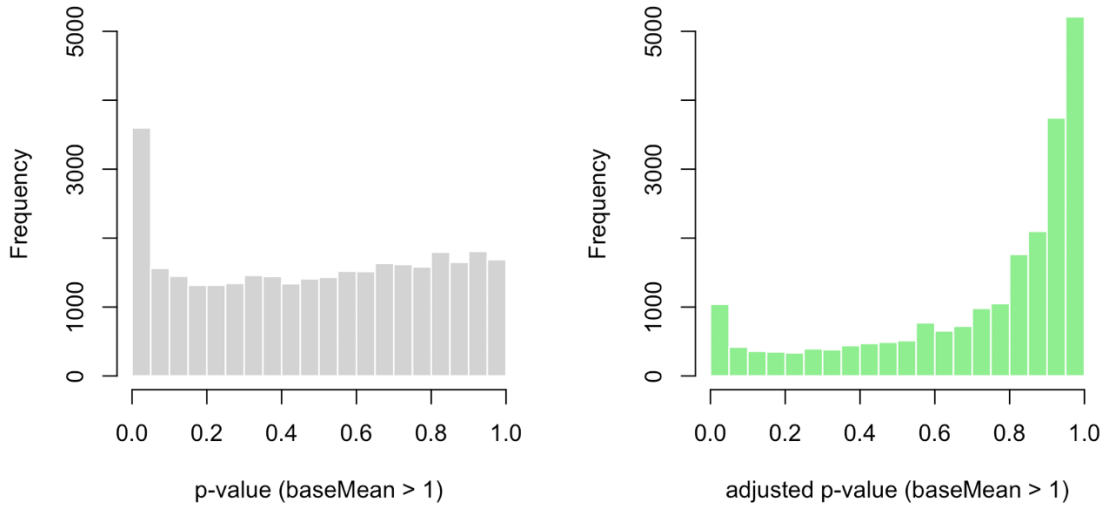


Figure A2. The p-value distribution of Wald test for the treatment effect in 4610.

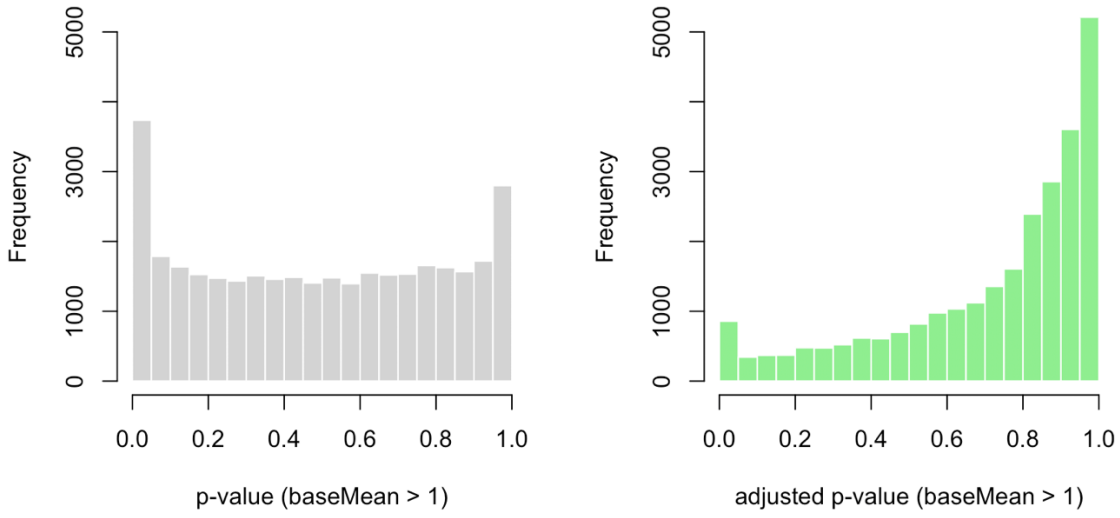


Figure A3. The p-value distribution of Wald test for the treatment effect in Rondo.

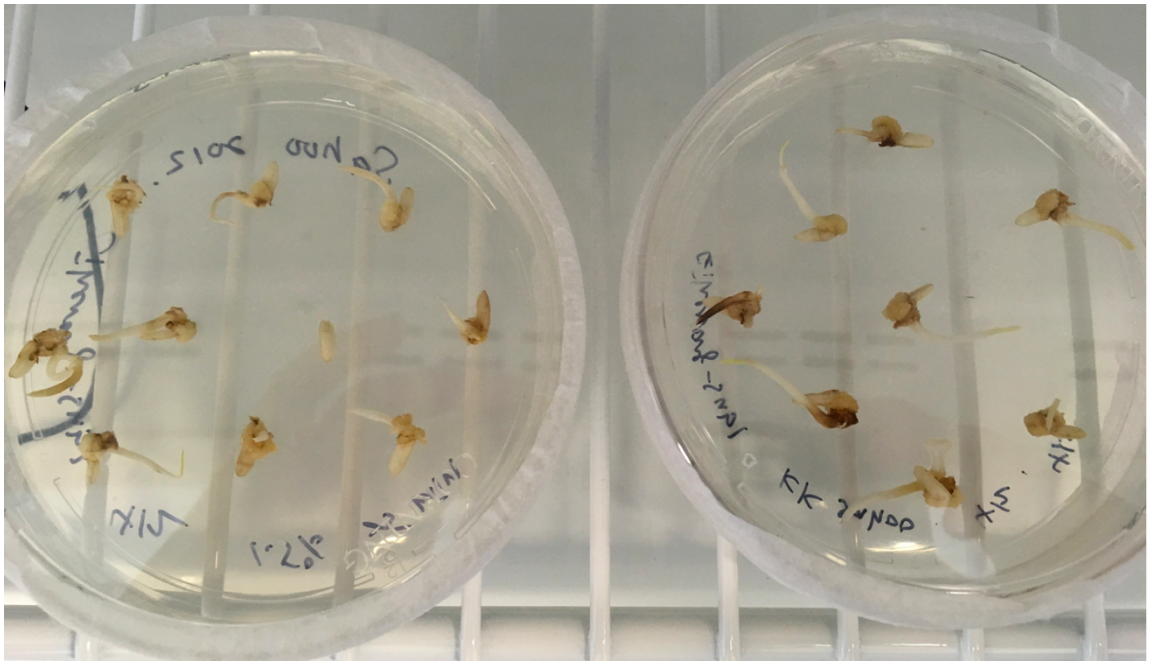

Figure A4. Callus induced from mature seeds after 10 days cultured in CIM with 2.5 mg 2,4-D and 0.15 mg 6BA (left); and with 3 mg 2,4-D and 0.25 mg 6BA (right).