GENOME-WIDE ASSOCIATION STUDIES FOR DRY-DIRECT-SEEDED RICE TRAITS AND GENE EDITING TO VALIDATE THE GENES UNDERLYING PURPLE LEAF COLOR IN RICE

A Dissertation

by

SUMEET PRABHAKAR MANKAR

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 Septiningsih
Committee Members,	Michael Thomson
	William Rooney
	Dirk Hays
Head of Department,	David D. Baltensperger

May 2021

Major Subject: Plant Breeding

ABSTRACT

Rice (*Oryza sativa* L.) is the major source of calories for more than half of the world population. Dry direct-seeded rice is becoming increasingly popular in rainfed and irrigated ecosystems worldwide mainly due to labor and water scarcities. Addressing constraints in the DDSR ecosystem is a key to develop high-yielding cultivars suitable for this ecosystem. Toward this goal, we performed a genome-wide association study (GWAS) using a subset of 300 *indica*/*aus* accessions from the 3,000 Rice Genomes Project. The 300 accessions were phenotyped for 23 traits, including nutrient, root, grain yield and yield-related traits components at the International Rice Research Institute, Philippines experimental fields in replicated trials during the dry and wet seasons of 2018. A total of 265,650 SNPs were used for the GWAS analysis using the Compressed Mixed Linear Model in R/GAPIT. For the 23 traits evaluated, a total of 55 QTLs were detected with the false discovery rate (FDR) values of < 0.001. Many previously reported genes and QTLs potentially colocalized with our significant GWAS sites, while some potentially novel QTLs were also detected. More QTLs were detected during the dry season compared to the wet season, partly due to a less favorable environment which leads to lower heritability for nearly all the traits examined. This study offers key insights into the prospective links between grain yield, yield components, and its environment using high-resolution association mapping. Our results demonstrate the complex nature of the genetic architecture of yield and related traits, which may assist lay the foundation to develop high-yielding varieties.

In another experiment, we designed an approach to validate the genes underlying purple leaf color. Out of the various color pigments produced by plants, anthocyanins are the essential secondary metabolites that protect plants against biotic and abiotic stresses. Toward this goal, we have successfully designed and validated guide *RNAs* for *OSB1* and *OSB2,* which subsequently will be used to knock out the purple leaf color via gene editing.

DEDICATION

To Khushboo for her unwavering support and encouragement over the years.

ACKNOWLEDGEMENTS

My deepest gratitude goes to Almighty God who has blessed me abundantly to complete my research work with success. With great reverence and pleasure, I obediently consider myself highly privileged to have Dr. Endang Septiningsih, Associate Professor, Department of Soil and Crop Sciences, Texas A&M University as chairman of the advisory committee. With a deep sense of lifelong indebtedness, I express my heartfelt gratitude to Dr Septiningsih for her valuable guidance, care, constant encouragement and help throughout the course of this study and in bringing out this dissertation.

I extend my thanks to dissertation committee members, Dr. Michael Thomson (Professor, Department of Soil and Crop Sciences), Dr. William Rooney (Professor and Borlaug-Monsanto Chair for Plant Breeding and International Crop Improvement) and Dr Dirk Hays (Professor and Chair Molecular and Environmental Plant Sciences) for their valuable suggestions throughout the course of my study. Thanks also go to the faculty and staff for making my time at Texas A&M University a great experience.

I am sincerely obliged to Dr. Arvind Kumar (Interim Head, Rice Breeding Platform), Dr Shailesh Yadav and Dr. Nitika Sandhu (Postdoctoral Fellows), International Rice Research Institute, Philippines for their valuable ideas and facilities rendered during the study. My special thanks to Tess Gerpacio Sta Cruz and Margaret Catolos, who were always helping me in field and lab works during my IRRI stay.

I record my gratitude and thanks towards respected Dr. Edward Runge Program Director and Judging Panel Chair, Monsanto/Bayer Beachell-Borlaug International Scholars Program for the support, guidance and care.

With much pleasure and respect, I extend my countless gratitude to my Lab members Dr. Backki Kim, Dr Nikolaos Tsakirpaloglou, Onieda Ibarra, Khushboo Rastogi, Sudip Biswas and Karina for their wholehearted help, critical suggestions, cordial teamwork and care rendered throughout research program.

On a personal note, I wish to place my thanks on book, to the giant pillars, my beloved wife (Khushboo), caring father, affectionate mother and compassionate in-laws for their motivations and prayers to keep me in high spirits to pursue the program successfully. I also express my deep sense of gratitude to my Brother (Amit Dada), and sisters (Dr Vrushali and Dr Pallavi) and Jiju (Dr Jayant Gawande and Mr. Shridhar Umale) for their inordinate support and constant encouragement throughout my study period and I dedicate this humble piece of work to them.

Finally, thanks to Khushboo Rastogi my wife for her encouragement, support, help, dedication, patience and love during the time of my PhD.

(Sumeet Prabhakar Mankar)

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a dissertation committee consisting of Professor Dr. Endang Septiningsih [Advisor] and Dr. Michael Thomson (Professor, Department of Soil and Crop Sciences), Dr. William Rooney (Professor and Borlaug-Monsanto Chair for Plant Breeding and International Crop Improvement) and Dr Dirk Hays (Professor and Chair Molecular and Environmental Plant Sciences).

The research experiment at International Rice Research Institute was designed by Dr Nitika Sandhu. All the research work conducted, and data analyzed for the dissertation was completed by the student independently.

Funding Sources

Graduate study was supported by a PhD fellowship from the Bayer's Beachell-Borlaug International Scholarship program previously known as Monsanto Beachell-Borlaug International Scholarship program.

This work was also funded in part by Texas A&M AgriLife Research.

NOMENCLATURE

CMLM	Compressed Mixed Linear Model
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
DDSR	Dry Direct-Seeded Rice
DS	Dry Season
DSB	Double-Strand Breaks
FAO	Food and Agriculture Organization of the United Nations
FDR	False Discovery Rate
GAPIT	Genomic Association and Prediction Integrated Tool
GWAS	Genome-Wide Association Studies
HDR	Homology-directed repair DNA repair
IRRI	International Rice Research Institute, Philippines
LCC	Leaf Chlorophyll Content
MAF	Minor Allele Frequency
MAS	Marker-Assisted Selection
NHEJ	Non-Homologous End Joining
PC	Principal Components
QTL	Quantitative Trait Loci
SPAD	Soil Plant Analysis Development Chlorophyll Meter
TASSEL	Trait Analysis by aSSociation, Evolution, and Linkage
WS	Wet Season

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

1. INTRODUCTION

Rice (*Oryza sativa* L.) is the principal staple food for almost half of the world's population. Rice is the major source of peoples' livelihood, food security, income, and employment worldwide. Rice is cultivated in more than 120 nations, with 158 million hectares of overall area grown worldwide, which supplies over 700 M tons of rice annually. Rice is primarily produced by China, India, Indonesia, Bangladesh, Vietnam, Myanmar and Thailand, together accounting for more than 80% of world rice production. In 2018, the United States of America ranked 1st in the average yield per hectare of rice as high as 8.05 t/ha and overall 12th in world rice production. In 2018-19, the USA produced 11 million metric tons (MT) of rice, which is 5% of world rice production (FAO, 2018). Because of climatic changes, rice cultivation is at risk, which can be detrimental to rice productivity (Swaminathan & Jana, 1992).

Due to climate change and the shortage of water and labor in the rice-growing areas, puddled transplanted rice (PTR) is being replaced with dry direct-seeded rice (DDSR) as a practical substitute in some irrigated environments (Kumar & Ladha, 2011). As a result of climate change, unpredictable rainfall adversely affects the crop establishment at seedling and reproductive stages by reducing nutrient accessibility and eventually grain yield, resulting in increasing irrigation and labor cost (Pathak *et al.*, 2011). According to several studies, DDSR mitigates 30-57% of the water requirement, approximately 70% of the labor requirement, around 45% production cost and 43% to 75% emission of greenhouse gas (GHG) by producing similar yields and higher net economic returns for farmers (Kumar & Ladha, 2011; P K Sharma *et al.*, 2002; A. Singh

et al., 2002). To reduce the yield losses in DDSR conditions, early uniform germination and emergence benefit the increase of early plant establishment and weed competitiveness in DDSR conditions (Azhiri-Sigari *et al.*, 2005). Rapid early vigor helps in more rapid field cover (Shipley, 2006), which decreases the rate of evaporation and hastened root growth for soil water and nutrients uptake(Zhao *et al.*, 2006). Superior seedling vigor is positively correlated with yield stability potential in dry direct-seeded conditions (Okami *et al.*, 2011).

Because of the semiaquatic ancestral foundation of rice and the diversity of environments and cultivation ecosystems, existing rice production methods are dependent on abundant water, therefore, remain more highly susceptible to water stress than other cropping systems. Rice is commonly cultivated by transplanting seedlings into puddled soil and flooded water conditions. Puddled transplanted rice (PTR) is traditionally grown by farmers all over the world. Overall, PTR uses 30% of the world's freshwater resource to yield 1 kg of rice using 1600 liters of water (Gleick, 1993; Pimentel *et al.*, 2004). Despite numerous advantages of using PTR, including higher yield, nutrient accessibility and improved weed control, it uses more water and labor, making it practically non-profitable and unsustainable (Farooq *et al.*, 2015).

One of the most important bottlenecks associated with DDSR cultivation system is maintaining nutrient dynamics (Kumar & Ladha, 2011). To address this, the root system architecture needs longer roots with higher nodal roots (Rose *et al.*, 2012) for better nutrient and water acquisition. Despite several advantages, DDSR has some shortcomings viz., more difficult weed control, increases in soilborne pathogens such as

nematodes and reduced availability of major soil nutrients (Nitrogen, Phosphorus, Potassium, Iron, and Zinc). Grain yield of the cereal crops is so far the principal driving force for the past and present breeding programs. The genetic architecture of grain yield and yield-related traits is complex with relatively low heritability, which needs to be dissected to meet the growing yield demands of the future. These traits are often governed by numerous small-effect quantitative trait loci (QTLs) affected by environmental factors. Up to now, very limited studies reported QTLs correlated with grain yield and the rest of the traits in DDSR conditions (Sandhu *et al.*, 2015; Singh *et al.*, 2017; Subedi *et al.*, 2019; Yadav *et al.*, 2017). Addressing constraints in the DDSR ecosystem is a key to develop high-yielding cultivars suitable for sustainable rice production for the future.

1.1. Genome-wide association mapping study

To address the constraints in the DDSR ecosystem, the extensive genetic diversity and variation untapped within the rice gene pool needs to be exploited for improved adaptation to DDSR conditions. The gene-tagging efforts based on traditional QTL mapping or genome-wide association study (GWAS) have been powerful tools in helping to clone genes of interest, including those that control grain yield and yield-related traits. However, GWAS offers several key benefits over linkage analysis due to its power to identify the hidden ancestral historical events instead of the family pedigree. Some of these benefits include very high-resolution mapping, a higher number of alleles, wider genetic diversity, no need for the development of expensive and tedious bi-parental mapping populations (Flint-Garcia *et al.*, 2003; Yu & Buckler, 2006).

The recent availability of high-throughput sequenced data for the 3K rice genome panel further increases the statistical power in detecting the causal genomic regions of many key agronomic traits, which currently limits its full exploitation as a mapping and pre-breeding resources. The ability to use this genomic data to identify QTLs and candidate regions for future fine-mapping and exploitation via marker-assisted selection (MAS) and also to identify lines that combine favorable alleles provides an approach for the exploitation of advanced germplasm and technologies (Wang *et al.*, 2018).

1.2. CRISPR/CAS-based genome editing

Conventional and recent plant breeding methods utilize natural or artificially generated heritable variations, which demands time and labor; it involves crossing and characterization of the multiple generations of progenies to modify even a single gene. Historically plant breeders have been trying to develop new technologies to manipulate the genome at faster rates. Following the discovery of fundamental biological processes like transcription and translation (central dogma of molecular biology), academics have been trying to develop new technologies to deliver rapidly and accurately modify genomes of elite mega-varieties. Genome editing is a precise method utilized to assess the function of a genomic region or genes. The CRISPR/Cas9 system is a genome editing tool that enables alteration of the genomic DNA by using a customizable single guide RNA to specifically target a nucleic acid sequence (Belhaj *et al.*, 2013). Cas9 is a nuclease protein complex that can be used to cleave the double-stranded DNA. The enzyme cleaves the complementary strand by following the guide RNA for the complementary DNA sequence of the target-specific sequence recognition; it then

creates double-strand breaks (DSB) in the region of interest. These double-strand breaks produce genomic rearrangements because of the non-homologous end joining (NHEJ) and homology-directed repair (HDR) DNA repair processes (Belhaj *et al.*, 2013). NHEJ utilizes DNA ligase IV to rejoin the broken ends, while HDR uses template strands (Belhaj *et al.*, 2015). Repair by NHEJ may lead to frameshift mutations by introducing insertions or deletions, most commonly knocking out the gene function (Song *et al.*, 2016). The combination of the NHEJ repair system and multiplex gRNAs can also be employed to achieve chromosomal deletions targeting large genomic regions. Similarly, HDR mechanisms can lead to the creation of insertions, base substitutions, or deletions (Belhaj *et al.*, 2015). The NHEJ repair mechanism is currently more routinely used in plant genome editing (Belhaj *et al.*, 2013).

1.3. OBJECTIVES

There are two major objectives of this proposal, which include performing a genome-wide association study (GWAS) for DDSR traits and genome editing in rice. The specific objectives are:

1. Identify significant genomic regions controlling yield and yield-related traits.
2. Identify significant genomic regions controlling root and seedling vigor traits at different growth stages.
3. Identify significant genomic regions controlling nutrient uptake.
4. Design an approach to validate the genes underlying purple leaf color.

2. IDENTIFY SIGNIFICANT GENOMIC REGIONS CONTROLLING GRAIN YIELD AND RELATED TRAITS

2.1. Introduction

Rice (Oryza sativa L.) is a staple food that feeds more than half of the world population. Sustainable production will have to overcome some challenges, including a decline in arable land, global water shortage, and global climate change (Royal Society, 2009). A decrease in rice production caused by biotic and abiotic stresses is one of the major concerns. To feed 9 billion people in 2050, the world's annual rice production, however, will have to increase markedly over the next 30 years to keep up with population growth and income-induced demand for food (Carriger & Vallée, 2007). Because of climatic changes, rice cultivation is at risk, which can be detrimental to rice productivity (Swaminathan & Jana, 1992).

Grain yield of the cereal crops has been the principal driving force for the past and present breeding programs. The genetic architecture of grain yield and yield-related traits is extremely complex with relatively low heritability, which needs to be dissected to meet the growing yield demands of the future. These traits are often governed by numerous small-effect quantitative trait loci (QTLs) affected by environmental factors. Numerous genetic studies on dissecting grain yield and its components using QTL mapping or GWAS have been reported (Begum *et al.*, 2015; Septiningsih *et al.*, 2003; U. Singh *et al.*, 2017a; Thomson *et al.*, 2003; Yonemaru *et al.*, 2010; Zhao *et al.*, 2011). However, limited studies on the identification of QTLs correlated with grain yield and

related traits in dry direct-seeded rice (DDSR) conditions have been done (Sandhu *et al.*, 2014, 2019; Singh *et al.*, 2017a; Subedi *et al.*, 2019; Yadav *et al.*, 2017).

To address the constraints in the DDSR ecosystem, the untapped wealth within the rice gene pool needs to be exploited for improved adaptation to DDSR conditions, which leads to higher rice production. GWAS offers some major advantages over linkage mapping analysis with higher mapping resolution due to higher recombination events, greater allele number, broader reference population, and no need of biparental mapping populations; which makes the approach more time saving and cost-effective in establishing an association (Flint-Garcia *et al.*, 2003; Yu & Buckler, 2006).

Toward this goal, we performed a GWAS using a subset of 300 *indica/aus* accessions from the 3,000 Rice Genomes (3K RG) diversity rice panel for grain yield and related traits under dry direct-seeded cultivation. The specific aims of this study are to identify significant QTLs controlling grain yield and related traits under DDSR and to identify potential candidate genes for the QTL targets. It is hoped that the results of this study will assist in developing high-yielding rice varieties under DDSR.

2.2. Materials and Methods

2.2.1. Plant materials

A subset of 300 lines of mostly *indica* and *aus* accessions from the 3K RG diversity panel were used for this current study (Appendix A). The seeds were requested from the International Rice Genebank at the International Rice Research Institute (IRRI), Los Baños, Philippines. A total of fifteen checks were added to the experiment: IRRI 154, IR 87707-446-B-B-B:7, IRRI148, Vandana, KALI AUS, IR74371-70-1-1, UPL RI 7, IRRI123, IR 94225-B-82-B, IR 94226-B-177-B, IR 91648-B-153-B-B, IR 91648-B-32-B-B, AUS BAK TULSI, IR 91648-B-230-B-B, and IR 91648-B-289-B-B.

2.2.2. Field experiment and management

The diversity rice panel was evaluated during the 2018 dry season (2018 DS) and wet season (2018 WS) in the upland experiment fields at IRRI, Los Baños, Philippines. Seeds were direct seeded at ~2 cm depth with a uniform spacing of 20 × 20 cm and 3m rows using 2 gm of seeds per accessions in an alpha lattice design (Serpentine) with two replications (5 progenies per block). Seeds were sown directly in the dry soil without raising nursery beds in non-puddled conditions. For better germination and weed control, the field was prepared about a month before the sowing date. Weed control management and fertilizer applications followed the method of Sandhu *et al.* (2019). Following a month of sprinkler irrigation at the seedling stage, surface irrigation was used as needed.

2.2.3. Phenotyping and data analysis

All the phenotypes for grain yield and related traits were sampled by choosing five plants randomly within each plot. Days to flowering (DTF) was noted after 50 % of the plants in a plot showed spikelet anthesis. At maturity, the plant height (PHT) was recorded in cm as the length from the ground-level to the uppermost tip of the panicle. The number of tillers (TN) and the number of productive tillers (PT) were counted manually at the maturity stage. The length from the panicle neck to the panicle tip was recorded in cm as panicle length (PL). For panicle weight (PW), five randomly selected panicles were weighed. The number of unfilled grains per panicle, number of filled grains per panicle, and the total number of grains per panicle were counted manually, and spikelet fertility is calculated using the formula: Spikelet fertility (SF) (%) = number of fertile grains × 100/total number of grains. For grain yield (GY), five plants were randomly selected and harvested and oven-dried for 3 days at 50ºC to 14% moisture content, then weighed and converted to grain yield per plant. 100 filled grains (SW) were randomly selected per panicle, and the weight was measured in gm.

R (v. 3.4.4) program was used to measure correlations between the traits using *corrplot.mixed* function in *corrplot* package. To get more precise estimates of genotypic values, we used the best linear unbiased predictions (BLUPs), which can eliminate the environmental deviation and estimate the real individual breeding value in GENSTAT V17.1. Broad sense heritability (H^2) was calculated using $H2 = \sigma^2G/(\sigma^2G + \sigma^2E/r)$, where r is replication number, and σ^2E is the error variance.

2.2.4. Genotyping

The high-resolution resequenced genotypic data are available for the subset of 300 rice accessions from the 3K RG panel (http://iric.irri.org/resources/3000-genomes-project). The initial number of SNPs that were selected for our panel was 559,297. After setting a minor allele frequency (MAF) criterion of 5%, heterozygosity of 0.2, and SNPs call rates ≥90%, a total of 265,650 SNPs were used in our GWAS.

2.2.5. Population structure and association mapping

The population structure on our panel was accessed using the 265,650 selected SNPs with the python-based algorithm fastStructure (http://rajanil.github.io/fastStructure). To choose the appropriate model complexity essential to describe the population structure, multiple values of K ranging 1 to 10 were executed using a variational Bayesian framework with simple model prior. The best K value (Q matrix) for the average admixture proportion in the population was predicted using the python algorithm chooseK.py, which predicts model components and complexity, explaining the population ancestry contribution and variation.

The marker-based kinship (K) matrix was estimated by averaging genotypic score at that locus using "Centered_IBS" function in TASSEL (Trait Analysis by aSSociation, Evolution, and Linkage) to predict relationship with additive genetic variance. The neighbor-joining distance matrix output from Tassel 5.0 (Bradbury *et al.*, 2007) was used to visualize the phylogenetic tree using FigTree v1.4.4. GWAS was performed using CMLM (Compressed Mixed Linear Model) (Zhang *et al.*, 2010) using the R package of GAPIT (Genomic Association and Prediction Integrated Tool) (Lipka

et al., 2012). The kinship matrix (K) and population structure (Q) were incorporated into the model to improve the statistical power of GWAS. The QQ plot will be used as a visual diagnostic in accessing the accuracy of the GWAS analysis. If necessary, suitable phenotypic transformation including log, square root, or box cox transformations will be used accordingly prior to the analysis.

For association tests, a threshold value for declaring marker-trait association was set at p-value < 0.00001 and FDR<0.01. A Bonferroni multiple test correction for the genome-wide threshold at a significant level 1% of 7.42×10^{-7} was accessed using the formula of "-log10 (0.01/effective number of SNPs)".

Candidate genes and colocalized QTLs were predicted by using the significant causal SNPs associated with the trait of interest using previous published reports, the QTARO database (Yonemaru *et al.* 2010; http://qtaro.abr.affrc.go.jp) and the MSU Rice Genome Browser release 7 (http://rice.plantbiology.msu.edu/cgi-bin/gbrowse/rice/). The average LD for *indica* and *aus* panel is reported to be approximately 100kb.

2.3. Results

2.3.1. Phenotypic correlations and heritability

In general, the correlations among traits are higher in 2018 WS compared to the 2018 DS (**Fig. 2.1**). Overall, the correlation coefficients in the 2018 DS were weak (-1< r <1). Among the traits, panicle length (PL) and the plant height (PHT) had the highest correlation coefficient in both the seasons (0.19 and 0.33). Panicle length (PL) was positively correlated with panicle weight (PW) (0.01 and 0.03), tiller number (0.09 and 0.03), and the number of productive tillers (0.17 and 0.14) in both the seasons. Similarly, PL was also positively correlated with spikelet fertility (SF) (0.14) during WS and grain yield (GY) (0.09) during the 2018 DS. Spikelet fertility (SF) was positively correlated with GY (0.08 and 0.12) during both the seasons; SF showed a positive correlation with PHT (0.14), PL (0.14), and PW (0.09) during the wet season; these correlations are negative in DS. In both seasons, a positive correlation was observed between PHT and HI; though, it was almost similar in both seasons (0.16 in DS; 0.15 in WS).

The trend for trait heritability is exactly opposite to the correlation coefficient, where in general, the trait heritability in 2018 dry (DS) was higher compared to those of 2018 wet season (WS) (**Table 2.1**). Among all traits, DTF and GY constantly had the highest heritability across seasons (DTF: 0.69 in DS, 0.63 in WS and GY: 0.66 in DS, 0.64 in WS); while PHT and PL had higher heritability during the dry season compared to the wet season (PHT: 0.66 in DS, 0.53 in WS and GY: 0.72 in DS, 0.29 in WS).

2.3.2. Population structure analysis

A total of 265,650 SNPs present across 12 rice chromosomes from 300 progenies were used to estimate the population structure and linkage disequilibrium (LD) decay. Using the python-based fastStructure algorithm, estimation of population structure i.e., Q matrix was predicted through the best K value of K = 6 (Fig. 2.2A). The estimation population structure with higher precision(Q) matrix (K=6) and kinship matrix (K) i.e., Q + K were utilized as covariates. The scree plot constructed using R/GAPIT showed a gradual decline after the first two principal components (PCs) (Fig. 2.2B). The first four PCs explain 24% of total variance; PC1, PC2, PC3 and PC4 explained 11.7%, 3.7%, 3.3%, and 2.8%, respectively. Phylogenetic analysis and kinship separated the population into two key diverse clusters, which is further divided into subpopulations (**Fig. 2.2A and Fig. 2.2D**). The VanRanden kinship heatmap signified very weak relatedness as the kinship value concentrated at *0.0 level (**Fig. 2.2D**). The unweighted NJ tree showed the phylogenetic structure of the GWAS panel, where most of the population were dominated by indica and aus, along with several aus/indica admixed and aromatic rice (Fig. 2.2E).

2.3.3. Identification of QTLs by GWAS

A total of 31 QTLs were detected in our GWAS of the ten agronomic traits across two seasons based on FDR<0.01 (**Table 2.2**); among those, 18 QTLs reached FDR<0.001. Those 18 QTLs were: a QTL for panicle weight (*qPW-12*), four QTLs for grain yield (*qGY-6, qGY-7, qGY-10*, and *qGY-7*), four QTLs for harvesting index (*qHI-1 qHI-2 qHI-5* and *qHI-10*), two QTLs for the number of unfilled grains (*qDTF-6-1* and

qDTF-6-2), three QTLs for plant height (*qPHT-1-1, qPHT-1-2* and *qPHT-5*), two QTLs for the number of productive tillers (*qPT-2* and *qPT-10*), two QTLs for the 100-seed weight (*qSW-2* and *qSW-9*), and a QTL for number of tillers (*qTN-12*) (**Table 2.2**).

From the total 31 QTLs identified during 2018 DS and WS, it is worth noted that 12 QTLs were detected in both seasons. Some examples of GWAS results on grain yield (GY), plant height (PHT), days to flowering (DTF), and panicle length (PL) were presented in Manhattan and Q-Q (quantile-quantile) plots depicting -log10 (P-values) and observed against expected data of marker-trait association as shown in Fig. 2.3. Candidate genes and QTLs from previous reports potentially colocalized with our QTLs were also listed in **Table 2.2**.

Out of the 31 QTLs detected in this study, 23 QTLs were identified in 2018 DS, while only 17 QTLs reached the threshold of FDR<0.01 in 2018 WS. QTLs were detected in all the ten traits; no common QTL was detected for days to flowering (DTF), plant height (PHT), and harvesting index (HI) during the dry and wet seasons. The largest number of QTLs detected on chromosome 2 with 6 QTLs, the least ones were on chromosome 9 with 1 QTL, and no QTLs detected on chromosome 4 and 8. The least number of QTLs was detected for panicle weight with only one QTL, i.e., *qPW-12*, which was seen at S12_19259233 with an FDR value of 0.00088 in 2018 DS and 0.00095 in 2018 WS.

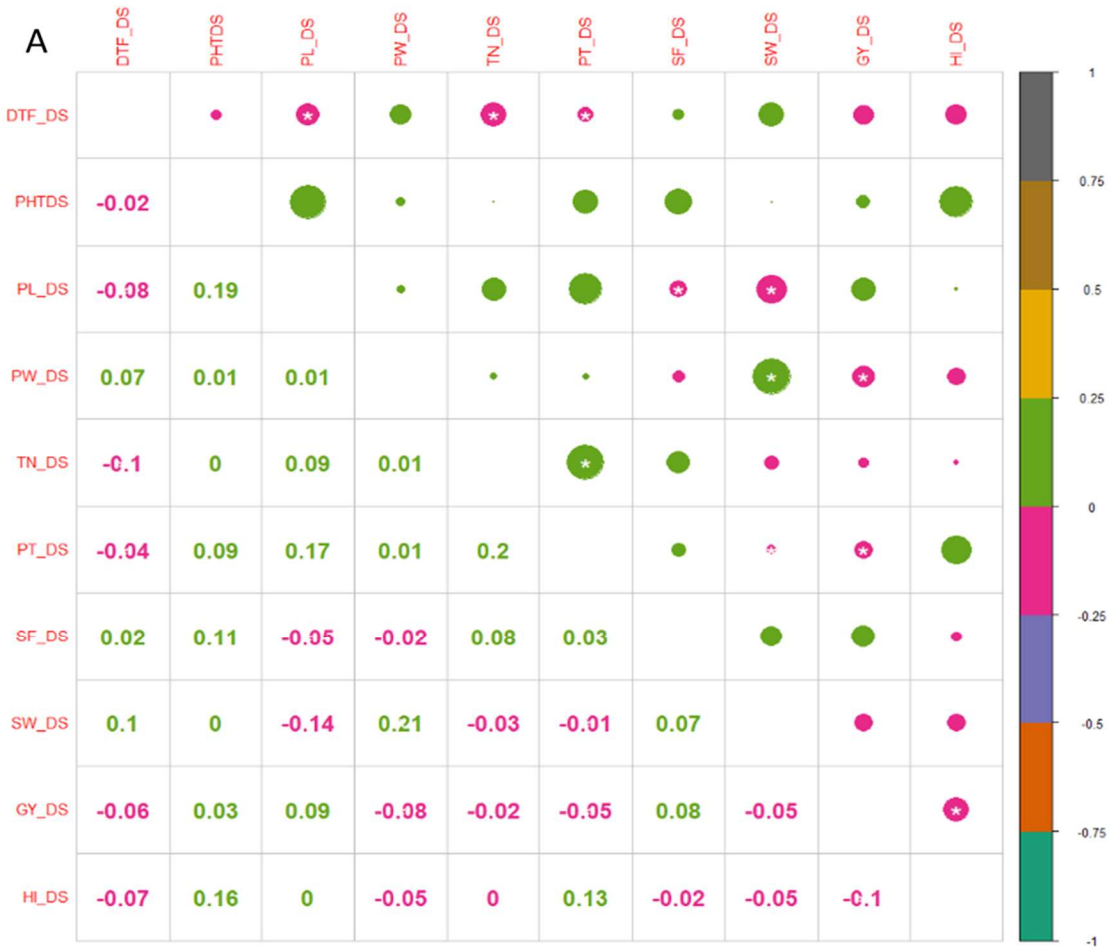

Fig. 2.1 Correlation plot of grain and grain yield related traits (P < 0.05 *, P < 0.01 **, P < 0.001***): (A) 2018DS. (B) 2018WS. The eight color indicates the correlation scale from -1 to 1 among different traits. DTF: days to 50% flowering (days), PHT: plant height (cm), PL: panicle length (cm), PW: panicle weight (gm), TN: number of tillers per plant, PT: number of productive tillers per plant, SF: spikelet fertility (%), SW: 100-grain weight (g), GY: grain yield per plant, HI: harvesting index.

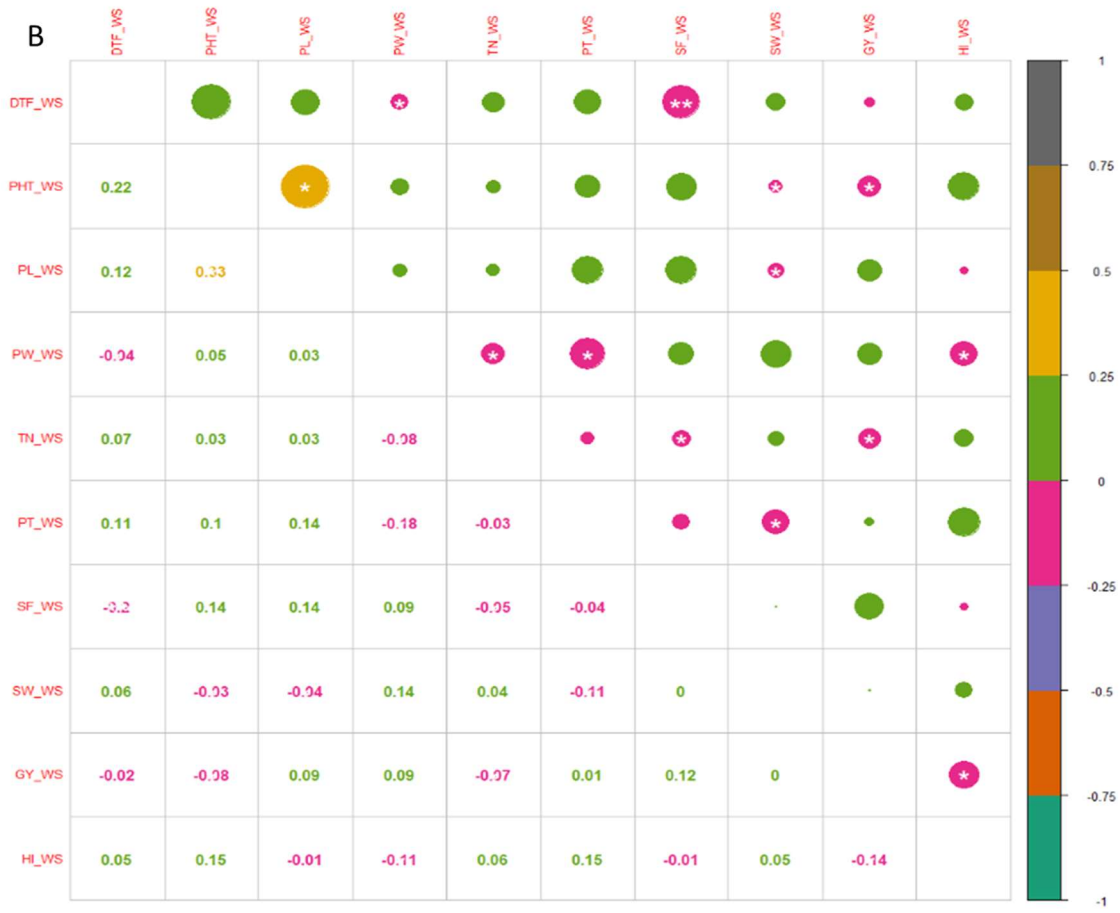

Fig. 2.1 Continued

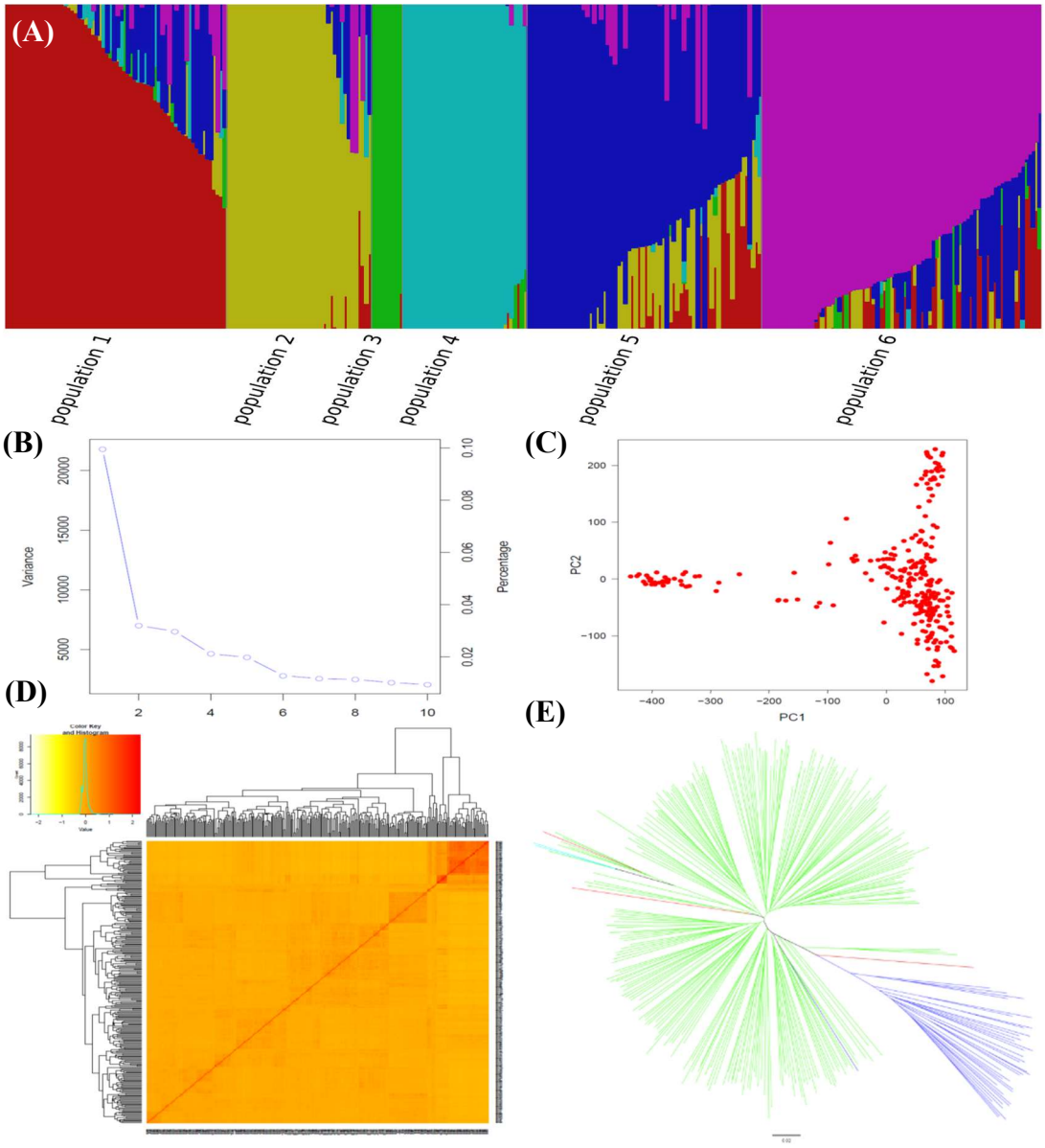

Fig. 2.2 Population structure, principal components and phylogenetic analysis of the association panel. (A) fastStructure cluster plots for the K= 6 of the association panel. (B) Scree plot showing most of the variability explained by PCs for association study. (C) Variation explained by first two principal components. (D) The genetic clustering heat map for evaluating the genetic differences among 300 rice varieties. (E) Phylogenetic neighbor-joining tree of association panel used for phenotyping under dry direct-seeded cultivation conditions. Colors indicates - Green: indica, Blue: aus; Teal: aromatic and Red: Admixture.

Table 2.1. Mean and trait heritability of checks and GWAS population for grain yield and related traits in Dry (DS) and wet season (WS) 2018.

Traits	Seasons	DTF	PHT	PL	PW	TN	PT	SW	SF (%)	HI	GY
IRRI 154	DS	73	102	22.0	2.91	13	8	2.65	90	0.80	4724
	WS	92	135	24.9	3.20	13	8	3.37	84	0.76	3198
IR 87707-446-B-B-B:7	DS	79	92	19.6	1.65	16	14	1.92	86	0.70	4233
	WS	85	113	24.8	2.48	19	19	3.35	73	0.65	3916
IRRI148	DS	--	--	--	--	--	--	--	--	--	--
	WS	79	115	22.9	3.42	9	9	2.4	94	0.63	2072
Vandana	DS	83	99	19.9	3.90	17	15	1.66	92	0.88	4377
	WS	93	111	22.2	2.54	12	10	2.31	90	0.76	3110
KALI AUS	DS	68	95	22.5	3.40	19	18	1.39	84	0.90	4623
	WS	69	107	22.9	2.65	19	18	3.21	78	0.86	4383
IR74371-70-1-1	DS	96	113	22.6	3.37	8	6	2.01	86	0.84	4482
	WS	84	112	25.3	3.75	9	9	2.49	81	0.79	3117
UPL RI 7	DS	73	117	24.0	2.78	19	16	2.21	71	0.89	4771
	WS	88	116	25.7	3.83	11	11	3.92	78	0.86	4167
IRRI123	DS	67	104	22.3	2.56	18	17	3.9	94	0.88	3542
	WS	84	109	23.4	2.29	18	17	3.19	80	0.73	2154
IR 94225-B-82-B	DS	72	120	22.2	2.92	10	9	3.33	95	0.76	3868
	WS	76	127	23.8	2.50	7	7	3.36	93	0.69	3273
IR 91648-B-289-B-B	DS	72	96	20.2	1.88	12	8	5.02	92	0.75	3168
	WS	102	92	21.1	3.49	13	11	1.27	88	0.69	2567
IR 94226-B-177-B	DS	96	131	24.2	3.21	7	7	2.97	93	0.83	3268
	WS	85	142	27.2	3.71	16	15	2.56	85	0.79	2862
IR 91648-B-153-B-B	DS	75	99	23.5	3.40	7	7	4.71	88	0.81	3836
	WS	96	124	23.0	2.44	7	7	1.43	80	0.65	3275
IR 91648-B-230-B-B	DS	68	98	21.3	2.83	19	18	1.23	96	0.79	3739
	WS	101	107	21.7	1.94	19	18	1.51	79	0.58	3143
AUS BAK TULSI	DS	73	137	22.1	2.48	12	11	2.27	91	0.72	3933
	WS	83	140	23.9	2.20	12	11	2.33	86	0.60	2707
IR 91648-B-32-B-B	DS	79	116	25.4	3.13	18	17	3.42	61	0.84	3476
	WS	87	120	22.9	2.95	17	16	2.02	72	0.78	3137
Population mean	DS	78 ± 10	114 ± 17	23.08 ± 2.5	2.68 ± 0.68	11.3 ± 4.6	11.4 ± 4.5	2.5 ± 0.22	67 ± 14	0.52 ± 0.07	3284 ± 834
	WS	86 ± 14	127 ± 22	24.26 ± 3.2	2.69 ± 0.69	9.6 ± 2.7	12.9 ± 2.9	2.3 ± 0.32	67 ± 8	0.72 ± 0.16	3134 ± 1292
Heritability	DS	0.69	0.66	0.72	0.32	0.13	0.12	0.33	0.19	0.65	0.66
	WS	0.64	0.53	0.29	0.11	0.13	0.11	0.24	0.36	0.63	0.64

DTF: Days to 50% flowering (days), PHT: plant height (cm), PL: panicle length (cm), PW: panicle weight (g), TN: number of tillers per plant, PT: number of productive tillers per plant, SF: spikelet fertility (%), SW: 100-grain weight (g), GY: grain yield per plant, HI: harvesting index.

A total of 4 QTLs were detected for days to flowering; they were positioned on chromosomes 6 (2 QTLs), 7, and 12. The QTL detected on chromosome 6, *qDTF-6-2*, has the highest SNP peak (FDR= 0.00065). There were 4 QTLs detected for grain yield, positioned on chromosomes 6, 7, 10, and 12. Interestingly, all the grain yield QTLs reached the threshold of FDR < 0.001 expect *qGY-12*. For tiller number, 3 QTLs were found on chromosomes 2, 7, and 12. Out of these three, only one QTL on chromosome 12 (*qTN-12*) reached the threshold of FDR>0.001. Only 2 QTLs were identified for the number of productive tillers; they were on chromosomes 2 and 10. The largest number of QTLs that were detected for the harvesting index was 5.

Four of the QTLs on chromosomes 1 (*qPHT-1-1* and *qPHT-1-2*), 3 (*qPHT-3*), and 5 (*qPHT-5*) reached the threshold of FDR > 0.01. Moreover, one of these three QTLs (*qPHT-1*) detected above the threshold of Bonferroni correction. Only 2 QTLs were identified for panicle length, located on chromosomes 3 and 6 (2 QTLs). Both the QTLs were detected during the 2018 DS and WS. A total of 5 QTLs were detected for harvesting index; they were on chromosomes 1, 2, 5, 10, and 11. Except for the QTL detected in 2018 WS on chromosome 1 (*qHI-11*), all other QTLs reached the threshold of FDR>0.001 in 2018 DS. There were 4 QTLs found for spikelet fertility on chromosomes 1, 2 (2 QTLs), and 11; out of these 4, 3 were detected during the 2018 DS.

A few of QTL hot spots harboring QTLs of more than one trait were identified across the chromosomes. For examples, *qSW-2* at S2_24382081 from 2018 DS and WS; *qHI-2* at S2_24389464 from 2018 DS; *qSF-2-2* at S2_24502637 from DS and WS; and *qPT-2* at S2_24999100 (chromosome 1; 617 kb apart between these 4 SNP peaks).

Second QTL hotspot was detected on chromosome 7 between *qDTF-7* at S7_20900202 and *qGY-7* at S7_21503408 (chromosome 7; 600 kb apart). Another QTL hotspot was detected on chromosome 12 between *qTN-12* (DS and WS) at S12_19115829; *qPW-12* (DS and WS) at S12_19259233 and *qGY-12* (DS) at S12_19508942 (chromosome 12; 393 kb apart). The proximity of those QTLs suggests that the genetic causal controlling those associated traits are either pleiotropic or tightly linked.

2.4. Discussion

2.4.1. Phenotypic correlations and heritability

The significant positive correlation of harvesting index with plant height and number of productive tillers may suggest the role of semi-dwarfing genes contributing to the plant biomass and grain yield. Days to 50% flowering was significantly and positively correlated with SW, which may explain the role of early flowering in reproductive and grain filling stage under water stress condition. The heritability of grain yield, SW, PHT, and DTF in the 2018 DS was higher than the 2018 WS. Additionally, lower trait heritability of traits observed in the 2018WS compared to 2018DS, which possibly explains the reason for the higher number of QTLs detected in the dry season (**Table 2.1 and 2.2**).

2.4.2. QTL identification

Grain yield of the cereal crops is so far the prime motivating force for the previous and current breeding programs. Low economic inputs, yield stability, and adaptability to withstand the variable growing environments are the primary cause for deploying the water-saving technologies viz, a dry direct-seeded condition in rice. DDSR significantly decreases the water use by ~50%, labor requirement by ~75%, cost of production by 45%, and greenhouse gas by 43% to 75%, while increasing net economic returns with similar yields (Kumar & Ladha, 2011). Several studies have been reported to identify QTLs and traits contributing toward grain yield in dry direct-seeded conditions (Sandhu *et al.*, 2014, 2019; Singh *et al.*, 2017b; Subedi *et al.*, 2019; Yadav *et al.*, 2017a). In the present study, a novel subset of 3K rice genome project i.e., a panel

of an aus/indica with 300 rice accessions has been assessed for grain yield under upland rice DDSR conditions. In our DDSR GWAS experiments, we have identified QTLs for all the ten target traits, as discussed below (**Fig 2.3 and Table 2.2**).

Table 2.2. QTLs for grain yield and relate traits detected by GWAS and potentially colocalized genes/QTLs.

QTL ID	Season	Chr	Position	P-value	FDR	Effect	Gene/QTL	Start	End	Function and Reference
qPHT-1-1	DS	1	35100345	$1.49x10^{-08}$	0.00029	10.42	*OsGH3.1*	35064294	35066779	Dwarf (Domingo *et al.*, 2009)
qPHT-1-2	WS	1	40108301	$2.96x10^{-07}$	0.0038	7.31	*sd-1*	40138232	40141316	Dwarf (Sasaki *et al.*, 2002)
qPHT-3	DS	3	33799373	$8.87x10^{-06}$	0.0010	10.56	*pla3/gp*	33710350	33719780	Dwarf (Kawakatsu *et al.*, 2009)
qPHT-5	DS	5	2346830	$7.1 x10^{-07}$	0.00021	-9.05	*OsY14a*	2311101	2311248	Dwarf (Gong & He, 2014)
qDTF-6-1	DS	6	15874051	$9.04 x10^{-07}$	0.00024	-7.97	*qHD-6-1*	6927624	20691040	Heading date (QTL) (Cui *et al.*, 2004)
qDTF-6-2	WS	6	3241538	$2.48 x10^{-07}$	0.00065	-4.46	*OsMADS5*	3161804	3168332	Floral organ formation (Cui *et al.*, 2010)
qDTF-7	WS	7	22400202	$9.98 x10^{-06}$	0.0043	-4.15	*OsUDT1*	22474472	22476169	Panicle flower (Gong & He, 2014)
							dth7.1	5512628	22532504	Days to heading (QTL) (Thomson *et al.*, 2003)
qDTF-12	DS	12	26073976	5.06x10-06	0.0067	5.03	*OsPID*	26281668	26283429	Panicle flower (Morita & Kyozuka, 2007)
qGY-6	DS	6	4552736	$3.52 x10^{-06}$	0.00067	-1085	*qLW-6*	3459492	6023472	Grain shape (QTL) (Yan *et al.*, 2003)
qGY-7	Both	7	24303408	$3.15 x10^{-09}$	0.00065	-560	*oswrky78*	24314295	24319767	Grain size (Zhang *et al.*, 2011)
qGY-10	Both	10	12904831	$5.17 x10^{-06}$	0.00009	860	*ssd10*	12044545	19623828	Spikelet setting density (QTL) (Xiao *et al.*, 1996)
							qSPB2-1, qSPBp10-2	9888577	15136649	Number of spikelet's (QTL) (Cui *et al.*, 2002)
qGY-12	DS	12	19508942	$1.12 x10^{-05}$	0.0016	959	*1000-GW*	5820051	24012742	1000 grain weight (QTL) (Zhuang *et al.*, 2001)
qPL-3	Both	3	12789514	$1.01 x10^{-08}$	0.0026	-2.3	*OsNAP*	12024635	12027052	Source activity (C. Liang *et al.*, 2014)
qPL-6	Both	6	7855170	$3.60 x10^{-08}$	0.0095	-1.76				
qSF-1	WS	1	5577035	$1.80 x10^{-06}$	0.0023	5.47	*OsMADS3*	5558512	5567904	Panicle flower (H.-G. Kang *et al.*, 1998)
qSF-2-1	DS	2	6648836	$5.20 x10^{-06}$	0.0058	5.04	*Cga1*	6734753	6736433	Sterility (Hudson *et al.*, 2013)
qSF-2-2	Both	2	24802637	$7.08 x10^{-06}$	0.0049	7.64	*OsCOL4*	24847664	24849132	Flowering (Lee *et al.*, 2010)
qSF-11	DS	11	24530882	$4.10 x10^{-06}$	0.0058	-7.78	*fon2*	24506800	24507550	Floral organ number (Suzaki *et al.*, 2006)
qTN-2	WS	2	9077449	$1.69 x10^{-06}$	0.0032	2.77	*lp*	9071132	9075127	Larger panicle (Li *et al.*, 2011)
							ep3	9071132	9075127	Erect panicle (Piao *et al.*, 2009)
qTN-7	WS	7	25406880	$9.60 x10^{-06}$	0.0010	2.01	*OsMADS18*	25448633	25453836	Flowering time (Fornara *et al.*, 2004)

QTL ID	Season	Chr	Position	P-value	FDR	Effect	Gene/QTL	Start	End	Function and Reference
qTN-12	Both	12	19115829	1.86×10^{-06}	0.00029	-2.36	--			
qPT-2	Both	2	24999100	1.15×10^{-06}	0.00016	3.19	--			
qPT-10	Both	10	10328386	5.22×10^{-07}	0.00015	2.07	--			
qPW-12	Both	12	19259233	4.67×10^{-06}	0.00088	-0.31	--			
qSW-2	Both	2	25382081	3.27×10^{-06}	0.00052	-0.09	*OsMPS*	25438318	25440388	Grain size (Schmidt *et al.*, 2013)
qSW-9	DS	9	13809469	3.72×10^{-06}	0.00073	-0.28	--			
qHI-1	DS	1	22991070	7.50×10^{-07}	0.00099	-0.03	--			
qHI-2	DS	2	24389464	1.60×10^{-07}	0.00042	0.04	*yld2.1*	21658702	26758298	Plot yield (QTL) (Marri *et al.*, 2005)
qHI-5	DS	5	9051221	1.52×10^{-06}	0.00013	0.04	--			
qHI-10	DS	10	5347152	6.76×10^{-06}	0.00024	0.04	--			
qHI-11	WS	11	12661168	1.78×10^{-07}	0.017	-0.09				

Table 2.2. Continued

DTF: Days to 50% flowering (days), PHT: plant height (cm), PL: panicle length (cm), PW: panicle weight (g), TN: number of tillers per plant, PT: number of productive tillers per plant, SF: spikelet fertility (%), SW: 100-grain weight (g), GY: grain yield per plant, HI: harvesting index.

2.4.2.1. Grain yield

This study detected QTLs *qGY-6, qGY-7, qGY-10*, and *qGY-12* for grain yield per plant (GY). These QTLs were detected in the proximity of known related QTL or gene. Our study identified the co-location of *qGY6* QTL on chromosome 6 for grain yield. Singh *et al.*, (2017) have reported this QTL hotspot region to be associated with early vigor, early uniform emergence and shoot length in an interval of 11.7- 27.6 cM under dry direct-seeded rice population of 3*Swarna/Moroberakan. Similar region was also reported to be colocalized with QTL for one of the most desirable traits under DDSR conditions i.e., culm strength identified from Swarna*3/Moroberekan mapping population, which elucidate the significance of lodging resistance for higher yields (Yadav *et al.*, 2017b).

The colocalization SNP S7_24303408 for grain yield (*qGY-7*) and S7_22400202 for DTF (*qDTF-7*) on chromosome 7 was detected in the close proximity with the previously identified important gene i.e., *OsWRKY78* contributing for seed development in rice (**Table 2.2**). T-DNA insertion mutant showed small seed phenotype due to reduced cell size resulting in decreased grain weight (Zhang *et al.*, 2011).

Two other grain yield QTLs identified on chromosome 10 (*qGY-10*) and 12 (*qGY-12*) are colocated with the reported QTLs for grain yield-related traits in rice viz., spikelet setting density (ssd10; Xiao *et al.*, 1996), number of spikelets on primary branches per primary branches (*qSPB2-1* and *qSPBp10-2*; Cui *et al.*, 2002), and 1000-grain weight (Zhuang *et al.*, 2001), respectively (Table 4). Xiao et al, (1996) reported 3 major QTLs positively contributing to grain yield and yield components identified from

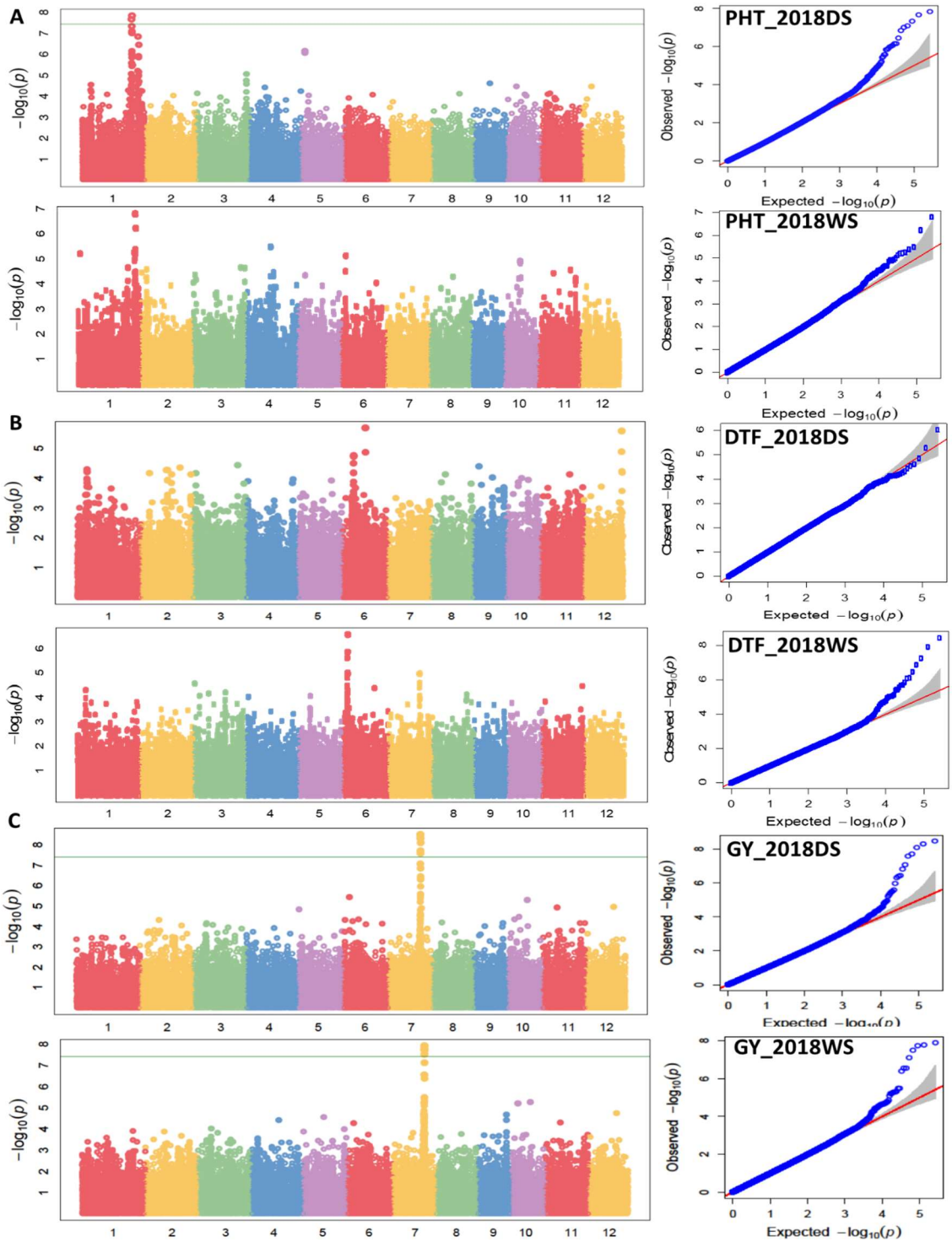

Figure 2.3 Manhattan and Q-Q plots of genome wide association mapping of Plant Height (PHT) (A); Days to flowering (DTF) (B), and Grain Yield (GY) (C) from 2018 Dry (DS) and Wet (WS) seasons.

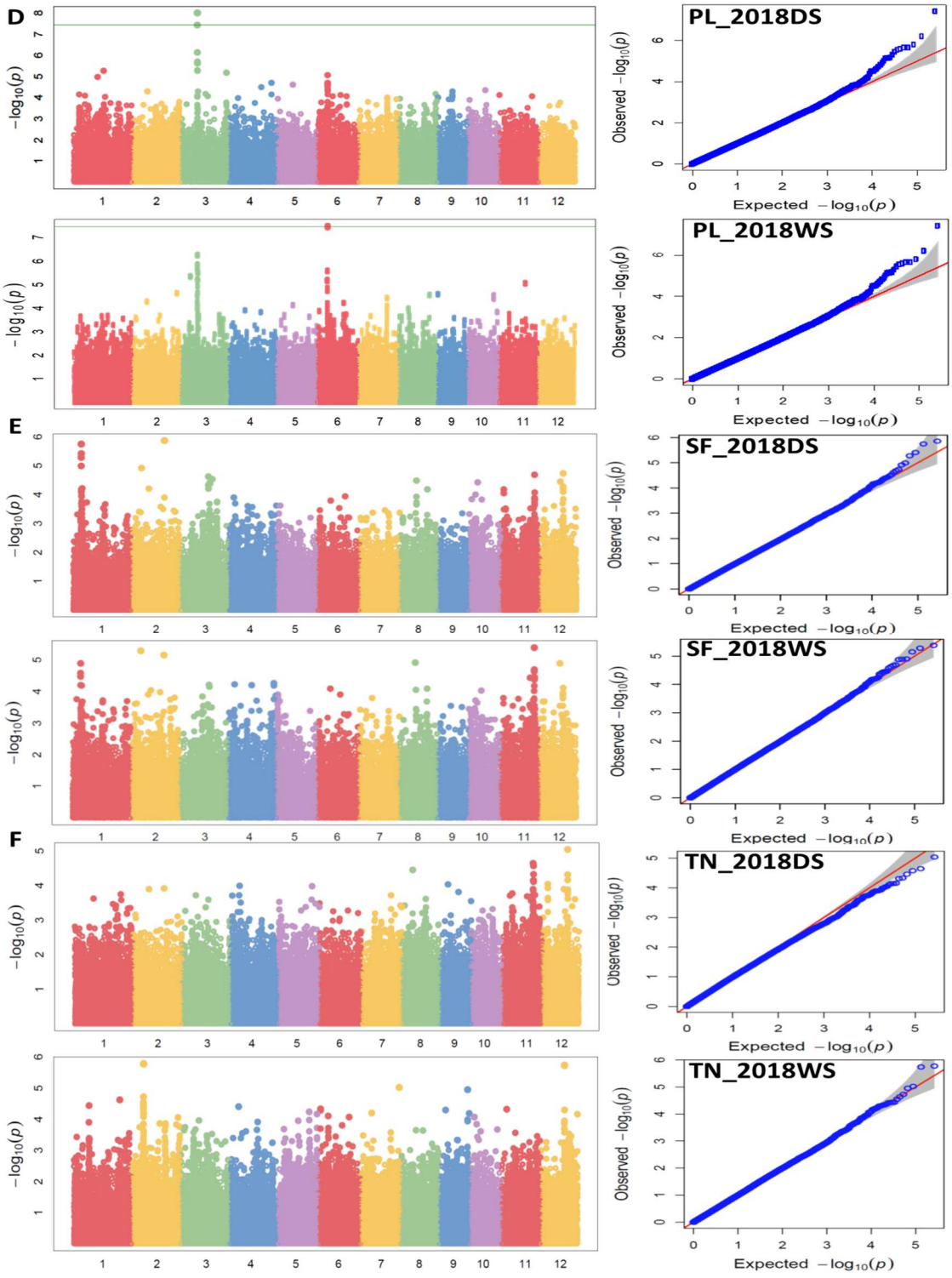

Figure 2.3 Continued.
Panicle length (PL) (D); Spikelet fertility (SF) (E), and Tiller number (TN) (F) from 2018 Dry (DS) and Wet (WS) seasons.

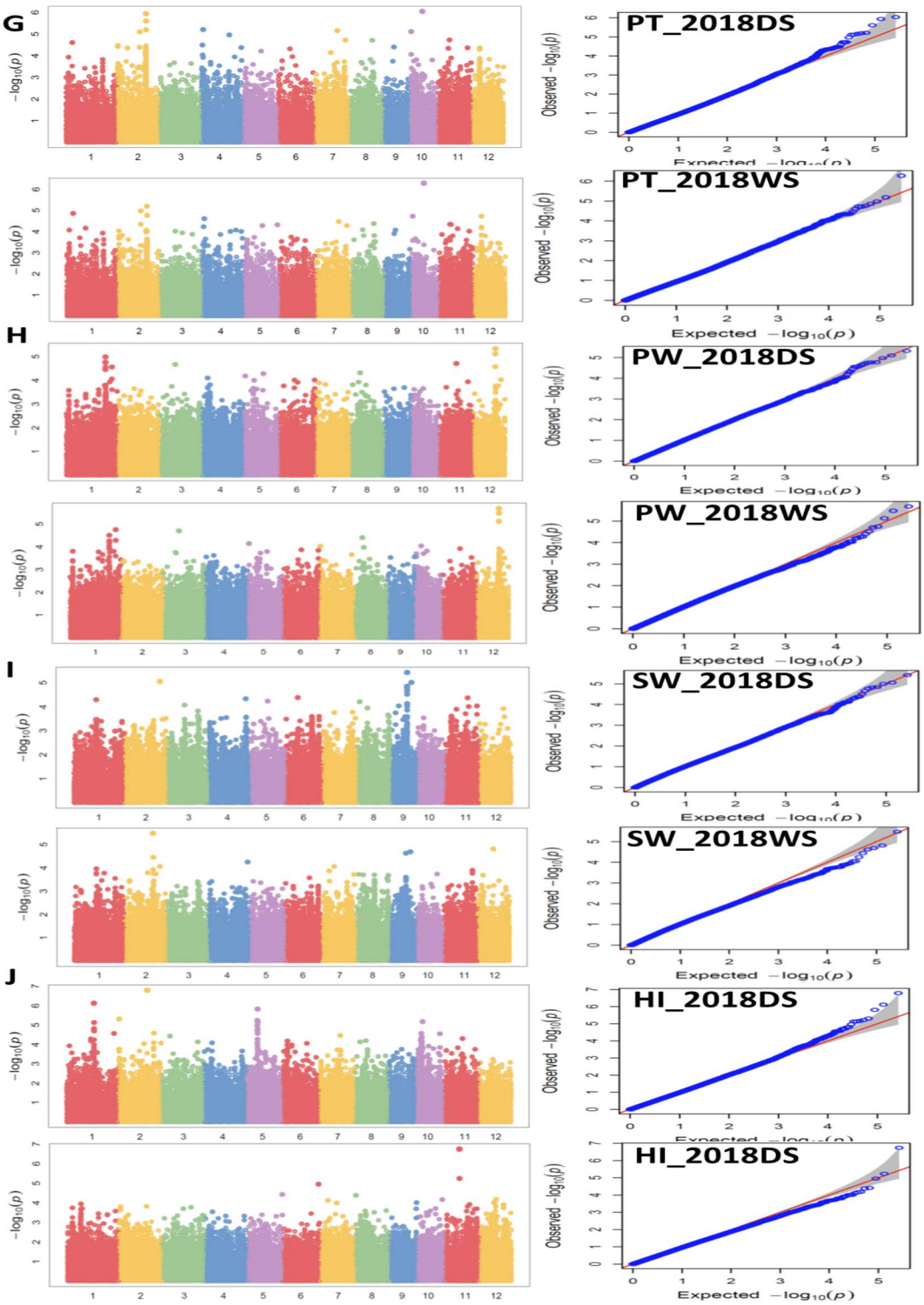

Figure 2.3 Continued.
Productive tillers (PT) (G); Panicle weight (PW) (H), Seed weight (SW) (I), and Harvesting index (HI) (J) from 2018 Dry (DS) and Wet (WS) seasons.

a subspecific of Oryza sativa *indica* (9024) × Oryza sativa *japonica* (LH422) F7 population assessed in an upland environment.

2.4.2.2. Day to flower

Numerous factors affect grain yield of rice plant, plant height, flowering time and are the vital determinants when it comes to adapt in variable growing environments. Xiao *et al.*, (1996) stated that "days to heading in rice parallels to flowering traits". Flowering time is an important determinant for adaptation to stress in reproductive stages. Early flowering helps to escape the water stress in reproductive stage and ensures higher fertility, which ultimately converge into superior yield benefits. Flowering time is negatively correlated with grain yield (-0.06 and -0.02 in 2018DS and 2018WS, respectively). In the present study, days to 50% flowering (DTF) associated SNP S6_3241538 present on chromosome 6 colocalized with flowering and floral organ regulator (*OsMADS5*; Cui *et al.*, 2010); induces the homeotic transformation of all floral organs and induces the flowering. Exploiting these QTLs/genes through marker assisted selection (MAS) in the future breeding programs may aid in improving grain yield and biomass under DDSR conditions.

The colocalization SNP S7_ 24303408for grain yield (*qGY-7*) and S7_22400202 for DTF (*qDTF-7*) on chromosome 7 was detected in the close proximity with the previously identified QTLs/genes for flowering time in rice viz., *dth7.1* (Thomson *et al.*, 2003) (**Table 2.2**). On chromosome 7, *OsUDT1* a key regulator gene in stamen development in angiosperm was reported by Gong & He, (2014), suggesting its role in adaptive evolution of rice.

2.4.2.3. Plant height

For plant height, SNPs S1_35100345and S1_40108301 on chromosome 1 colocalized with genes *OsGH3.1* and Semi-Dwarf1 (*sd-1)*, respectively, which the dwarf genes that are known to be associated with plant height and grain yield (Sasaki *et al.*, 2002). *sd-1* and *OsGH3.1* are known to be involved in the stem and leaf polar cell elongation (Domingo *et al.*, 2009). *OsGH3.1* is an GH3 family protein which is responsible for the inactivation of active IAA. Overexpression of *OsGH3.1* produces a dwarf phenotype in rice due to reduction in auxin (Domingo *et al.*, 2009).

sd-1 is also known as the green revolution semi-dwarf genes, which reduces the plant height identified in Dee-geo-woo-gen (Sasaki *et al.*, 2002). The *sd-1* mutation resulted plants were had plant architecture favorable for lodging resistance and use of heavy use of nitrogen fertilizer to improve grain yield. *sd1* gene encodes biosynthesis enzyme *GA20ox* an oxidase enzyme involved in the biosynthesis of gibberellic acid.

The SNP S3_33799373on chromosome 3 is also associated with the previously known gene *pla3/gp* (Plastochron3/Goliath), which severely reduces uppermost internode forming dwarfism in rice (Kawakatsu *et al.*, 2009). A similar genomic region was previously reported to improve grain yield significantly in *Oryza rufipogon* × Oryza sativa cultivar Jefferson BC2F2 population (Thomson *et al.*, 2003)

Another QTL for plant height on chromosome 5 was colocalized with *OsY14a* gene. *OsY14a-RNAi* transgenic plants were drastically shorter as compared to control plants, which occurred due to reduced culm length and absence of internode 5 (Gong & He, 2014),

2.4.2.4. Number of tillers and productive tillers

Tiller number (TN), S7_25406880 showed a significant marker-trait association detected on chromosome 7 which was located in close proximity to *OsMADS18* candidate gene controlling flowering time in rice (Fornara *et al.*, 2004). *OsMADS18* is necessary for flower development and floral transition (Fornara *et al.*, 2004). The rest of the two QTLs for number of tillers are not colocalized with any previously reported QTLs/genes, which may show their novelty. Similarly, there are no QTLs or genes have been reported in the similar region of the two QTLs for number of productive tillers.

2.4.2.5. Spikelet fertility

A total of 4 QTLs for spikelet fertility were identified; 3 for DS, 2 for WS, and 1 for both seasons, on chromosome 1, 2 (2 QTLS), and 11. The QTL *qSF-1* detected on chromosome 1 was in the proximity of known gene *OsMADS3* (Kang *et al.*, 1998). The *OsMADS3* gene is essential for the normal development of the internal two whorls, the stamen, and carpel, of the flower. Another two QTLs for SF were detected on chromosome 2 (*qSF-2-1* and *qSF-2-2*). The SNP S2_6648836 (*qSF-2-1*) was identified on chromosome 2 in the proximity with *Cga1* (Cytokinin-Responsive GATA Transcription Factor 1) gene directly contributing to spikelet fertility in rice (Hudson *et al.*, 2013). *Cga1* overexpression inhibits nutrient remobilization to the grains and inhibits grain filling in the panicles impairing grain filling process.

Another SNP S2_24802637 was identified on chromosome 2 in the close proximity with *OsCOL4* (CONSTANS-like 4) (Lee *et al.*, 2010), this gene is known to positively contribute toward flowering. *OsCOL4* acts autonomously from formerly

reported flowering genes. The last QTL, *qSF-11* at SNP S11_24530882 on chromosome 11 was detected to be colocalized with *fon2* (Floral Organ Number-2); plays a role in increasing floral organ number (Suzaki *et al.*, 2006).

2.4.2.6. Panicle length and weight

For panicle length, a total of 2 QTLs were identified for both seasons on chromosome 3 and 6. The SNP S3_12789514 on chromosome 3 was also detected to be colocated with several source activity and panicle development genes/QTLs like *OsNAP* (Liang *et al.*, 2014), *ccfs3*, and *ccf3b*, which are necessary for cumulative chlorophyll contents of the flag leaves (Yoo *et al.*, 2007). *OsNAP* encrypts a NAC transcription factor, positively regulating JA pathway to facilitate the foliage senescence process.

The QTL on chromosome 6 (*qPL-6*) is mot colocalized with previously reported QTLs or genes, which may suggest that this QTL could be a novel one. The QTL only for panicle weight (PW) was detected on chromosome 12 with an FDR value of 0.00088, which is not colocalized with any previously reported QTLs or genes.

2.4.2.7. Harvest index

For the harvest index, this study detected five QTLs, i.e., *qHI-1, qHI-2, qHI-5, qHI-10,* and *qHI-11*. A total of 4 QTLs were identified for DS, 1 for WS on chromosome 1, 2, 5, 10, and 11. No common QTLs identified between the two seasons. On chromosome 2, SNP S2_24389464 is also associated with plant height related semi-dwarfing QTL *qPH-2* (Marri *et al.*, 2005); grain yield *yld2.1* (Marri *et al.*, 2005) and flowering time QTL *qPTD-2* (Sheng *et al.*, 2002). No other previously reported QTLs or genes were detected for the rest of the harvest index QTLs.

3. IDENTIFY SIGNIFICANT GENOMIC REGIONS CONTROLLING ROOT AND SEEDLING VIGOUR TRAITS AT DIFFERENT GROWTH STAGES

3.1. Introduction

Rice cultivation is one of the most important agricultural activities on earth, with nearly 90% of it being produced in Asia and supplies more than 50% of calories consumed by the world's population. Food security in Asia is challenged by increasing food demand and threatened by declining water availability. More than 75% of the rice supply comes from 79 million ha of irrigated land, which utilizes 90% of total diverted freshwater. However, the water-use efficiency of rice is very low, and growing rice requires large amounts of water. An increasing scarcity of water has threatened traditional rice cultivation practices all over the world. The situation is further aggravated by drought, global warming, methane emission, adverse climatic changes, over-pumping of groundwater, causing aquifer resources to decline, and the high 'cost' of water (Tuong & Bouman, 2003).

Overexploitation of groundwater has caused serious problems in many parts of world, including America and Asia; groundwater tables have dropped on average by 0.5 - 1m per year (Bouman & Tuong, 2001). By 2025, it is expected that the irrigated wet season rice will experience "physical water scarcity" ; most of ~22 Mha of irrigated dry season rice will suffer "economic water scarcity" (Tuong & Bouman, 2003).

3.1.1. Root system architecture

In comparison to other cereal crops, rice has a shallow root architecture which is limited to the depth of 60 cm (Fukai & Inthapan, 1988). Root architecture of rice also

differs with the irrigation method (Yoshida & Hasegawa, 1982). Root architectural plasticity, the ability to exhibit morphological and physiological responses to a changing environment, may play an essential role in plant adaptation to heterogeneous environments (Bradshaw, 1965; Pigliucci, 2001; Sultan, 2000) and is vital in maintaining crop productivity (Wang *et al.*, 2006) with low inputs. Root system plasticity has been implicated in causing genotype x environment interactions (MacMillan *et al.*, 2006). Interest has increased in understanding the molecular control of root development in rice because of the importance of root system establishment and plasticity in plant performance in the field under direct-seeded cultivation conditions.

During the vegetative stage, rapid groundcover achieved with early vegetative vigor (Michael Dingkuhn *et al.*, 1999; Poorter & De Jong, 1999; Shipley, 2006) can reduce soil evaporation and accelerate root access to soil water and nutrients (Zhao *et al.*, 2006). Direct seeded conditions decrease the number of nodal roots per internode and stimulate the elongation of nodal roots (Kondo *et al.*, 1999). Other traits that may be important for direct-seeded rice are root hair growth and lateral root growth. Nutrient uptake is closely related to root hair length, and plants grown under phosphate-limiting conditions form longer root hairs (Bates & Lynch, 1996; Zhang *et al.*, 2003). Interest has increased in understanding the molecular control of root development in rice because of the importance of root system establishment and plasticity in plant performance in the field under direct-seeded cultivation conditions.

Root traits and early vegetative vigor are typically complex and controlled by many genes, each with a small genetic effect; such traits are typically controlled by

quantitative trait loci (QTL; Sharma *et al.*, 2011). Identifying genetic variation and QTL for root traits can contribute to our understanding of their role in plant performance under direct-seeded conditions.

3.1.2. Seedling vigor

Compared with transplanted rice fields, DSR conditions may be more favorable for the growth of weeds, which compete with rice for nutrients, moisture, and sunlight and can cause large yield losses (Pathak *et al.*, 2011). To improve initial crop establishment and competitiveness of direct-seeded rice, varieties with higher germination and faster seedling emergence with more vigorous seedlings under upland direct-seeded conditions must be selected to minimize the risks encountered in direct-seeding (Azhiri-Sigari *et al.*, 2005).

During the vegetative stage, rapid groundcover achieved with early vegetative vigor (Michael Dingkuhn *et al.*, 1999; Poorter & De Jong, 1999; Shipley, 2006) can reduce soil evaporation and accelerate root access to soil water and nutrients (Zhao *et al.*, 2006). Early vegetative vigor can be defined as a high relative growth rate (RGR) during exponential growth before canopy closure (Michael Dingkuhn *et al.*, 1999; Poorter & De Jong, 1999; Shipley, 2006). Early vegetative vigor depends on the assimilate source (light capture and photosynthetic rate) as well as sink constituted by structural growth (leaf appearance rate, potential size, and tiller outgrowth).

Early vegetative vigor may also accelerate the depletion of soil water reserves, making less water available for later crop stages (Zhang *et al.*, 2005). However, in dry direct-seeded environments, early vegetative vigor is associated with yield stability

(Okami *et al.*, 2011). Differences in early vegetative vigor of rice cultivars affecting crop establishment under direct-seeded conditions have earlier been reported in Asia (Caton *et al.*, 2003; Garrity *et al.*, 1992), Latin America, and Africa (Michael Dingkuhn *et al.*, 1999; Fischer *et al.*, 1997; Johnson *et al.*, 2008).

The objectives of this study were to investigate genetic variability in root traits and early vegetative vigor; to identify traits that could be used as selection criteria for improved weed competitiveness, crop establishment, and stress tolerance leading to high yield.

3.2. Materials and Methods

3.2.1. Plant materials, field experiment and management, genotyping, population structure and association mapping

The procedure and plant material used in Chapter 2 for Genotyping, Population structure and Association mapping analysis is similar to Chapter 3.

3.2.2. Phenotyping

All the phenotypes for early vegetative vigor and root traits were sampled by choosing three plants randomly within each plot. The plants were uprooted from the soil at each sampling to measure various root traits by digging a hole (40 cm deep) surrounding each plant. The roots and shoots were then separated by cutting the plants from the topsoil line. The separated root samples at 15, 22, 29 DAS, and DTF were washed gently and properly with running tap water. The number of nodal roots (NR), i.e., NR1 (number of nodal roots at 15 DAS), NR2 (number of nodal roots at 22 DAS), NR3 (number of nodal roots at 29 DAS) and NR4 (number of nodal roots at DTF), was counted manually. Maximum root length (RL). i.e., RL1 (maximum root length at 15 DAS), RL2 (maximum root length at 22 DAS), RL3 (maximum root length at 29 DAS), and RL4 (maximum root length at DTF), was measured using a ruler in centimeter. The roots were then oven-dried at 60°C for 72 hours for root dry weight (RDW) measurements, i.e., RDW1, RDW2, RDW3, and RDW4.

Vegetative vigor in terms of relative growth rate (RGR) was documented from three randomly selected seedlings per plot. The seedlings were uprooted using a trowel at 15, 22, 29 DAS, and DTF. The roots and shoots were separated, and the shoots were

oven-dried at 60 °C for 72 hours with each sampling. The oven-dried shoots were weighed (DSW) immediately after being taken out from the oven to calculate RGR. RGR was calculated at different time points, i.e., RGR1 (from 15 to 22 DAS), RGR2 (from 22 to 29 DAS, RGR3 (from 15 to 29 DAS, and RGR4 (from 29 DAS to DTF, using the following equation:

$$\frac{\text{Log (dry shoot weight at sampling 2) – log (dry shoot weight at sampling 1)}}{\text{(date of sampling 2}-\text{date of sampling 1)}}$$

3.3. Results

3.3.1. Phenotypic correlations and heritability

In general, the correlations among traits are higher in 2018 WS compared to the 2018 DS (Fig. 3.1). Overall, the correlation coefficients in 2018 WS were weak (-1< r <1). Among the traits, root dry weight (RDW4_DS) and the shoot dry weight (SDW4_DS) had the highest correlation coefficient in both the seasons (0.87 and 0.46). The number of nodal roots (NR4_DS) was positively correlated with root length (RL4) (0.42 and 0.24), root dry weight (RDW4_DS) (0.57 and 0.45) and shoot dry weight (SDW4_DS) (0.40 and 0.81) in both the seasons. Similarly, root length (RL4) was also positively correlated with root dry weight (RDW4) (0.82) and shot dry weight (SDW4) (0.81) during the dry season (2018DS). Root length (RL3) was positively correlated with root dry weight (RDW4) (0. 89). Root length (RL2) shown a positive correlation with shoot dry weight (SDW2) (0.41) during the dry season; these correlations are non-significant in WS. In the wet season, a positive correlation (0.64) was observed between root dry weight (RDW3) and shoot dry weight (SDW3), which is negative in dry season.

The trend for trait heritability is similar to the correlation coefficient, where in general the trait heritability in 2018 DS was higher compared to those of 2018 WS (Table 3.1). Among all traits, root dry weight (RDW4) and shoot dry weight (SDW3) constantly had the highest heritability across seasons (RDW4: 0.77 in DS, 0.71 in WS and SDW3: 0.79 in DS, 0.63 in WS); while root dry weight (RDW3) and root dry weight (RDW1) has higher heritability during the dry season compared to the wet season (RDW3: 0.73 in DS, 0.56 in WS and RDW1: 0.64 in DS, 0.59 in WS).

3.3.2. Identification of QTLs by GWAS

A total of 16 QTLs were detected in our GWAS of the 16 agronomic traits across two seasons based on FDR<0.001 (Table 3.2). Among those, 16 QTL were: a QTL for root length at 22 DAS (*qRL2-1*), two QTLs for nodal roots at 29 DAS (*qNR3-12* and *qNR4-6*), one QTL for shoot dry weight at 29 DAS (*qSDW3-8*), and twelve QTLs for root dry weight (*qRDW1-3, qRDW1-4, qRDW1-5, qRDW1-7, qRDW1-10, qRDW2-3, qRDW2-4, qRDW2-7, qRDW3-7, qRDW3-3, qRDW3-4* and *qRDW3-5*).

From the total 12 QTLs identified for root dry weight during 2018 DS and WS, it is worth noted that among those 3 QTLs were detected in 15, 22 and 29 DAS. Some examples of GWAS results on root dry weight (RDW), nodal roots (NR3), root length at 22 DAS (RL2), and shoot dry weight at 29 DAS (SDW3) were presented in Manhattan and *Q-Q* (quantile-quantile) plots depicting -log10(P-values) and observed against expected data of marker-trait association as shown in **Fig. 3.2**. Candidate genes and QTLs from previous reports potentially colocalized with our QTLs were also listed in Table 3.2.

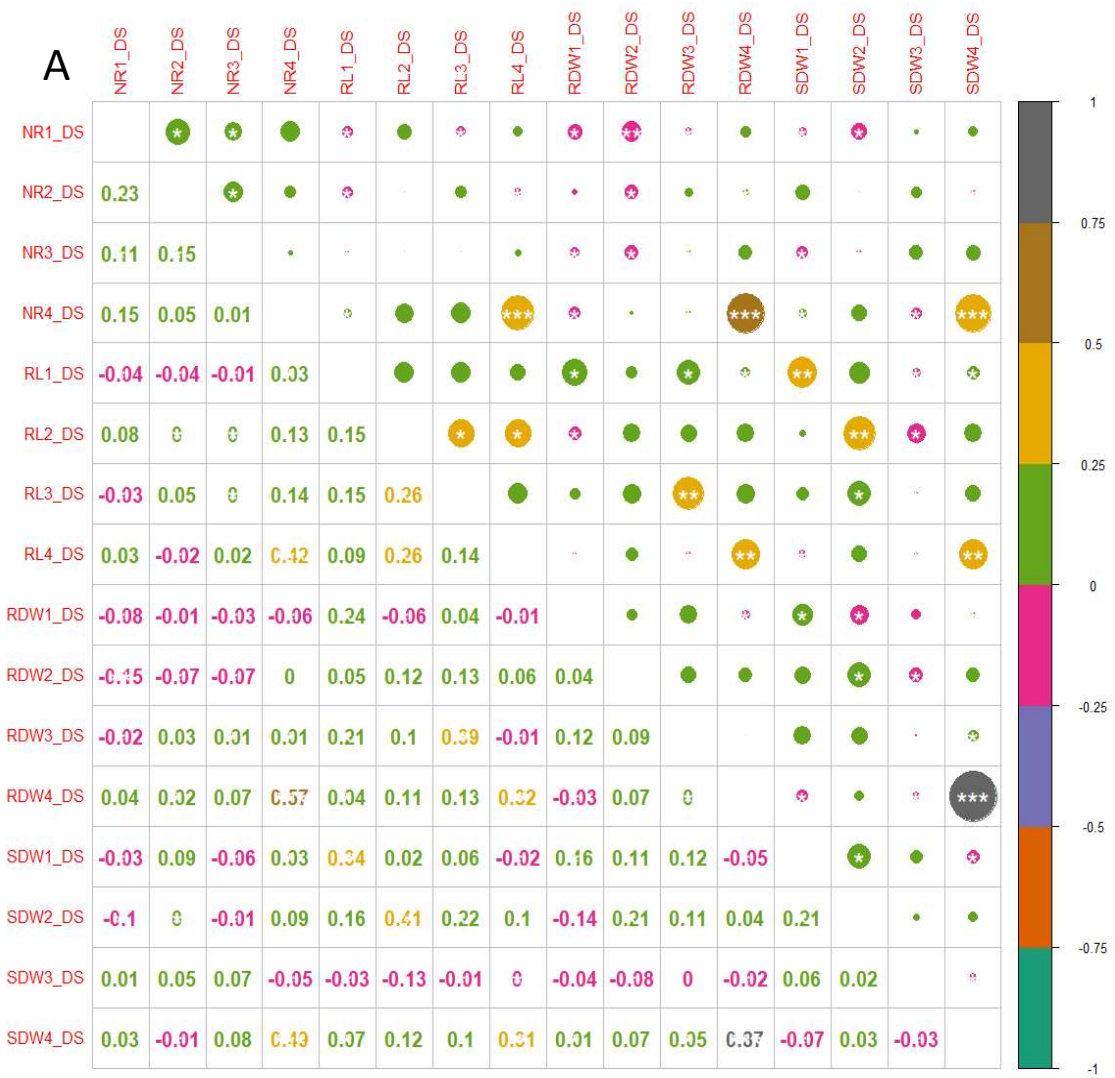

Fig. 3.1 Correlation plot of root and early vegetative vigor traits (P < 0.05 *, P < 0.01 **, P < 0.001); (A) 2018DS. (B) 2018WS (P < 0.05 *, P < 0.01 **, P < 0.001). The eight color indicates the correlation scale from -1 to 1 among different traits. Number of nodal roots (NR) [NR1, NR2, NR3, NR4: 15, 22, 29 DAS and DTF]. Maximum root length (RL) [RL1, RL2, RL3, RL4: 15, 22, 29 DAS and DTF]. Root dry weight (RDW) [RDW1, RDW2, RDW3, RDW4: 15, 22, 29 DAS and DTF]. Shoot dry weight (SDW) [SDW1, SDW2, SDW3, SDW4: 15, 22, 29 DAS and DTF].

Fig. 3.1 Continued.

Table 3.1. Mean data and trait heritability of checks and GWAS population for root traits and seedling vigor in dry and wet season 2018.

Traits	Season	RDW1	RDW2	RDW3	RDW4	RL1	RL2	RL3	RL4	NR1	NR2	NR3	NR4	SDW1	SDW2	SDW3	SDW4
IRRI 154	DS	0.20	0.24	0.33	38.08	85.17	92.22	130.78	297.09	15	22	41	336	0.30	0.34	0.90	208.53
	WS	0.20	0.21	0.22	25.91	48.44	54.33	102.11	263.67	10	17	23	273	0.21	0.26	0.41	177.48
IR 87707-446-B-B-B:7	DS	0.15	0.19	0.41	36.08	83.67	88.50	144.83	309.22	12	21	35	325	0.29	0.33	0.69	225.77
	WS	0.18	0.25	0.31	28.39	58.42	67.92	111.33	256.75	11	17	26	320	0.21	0.35	0.50	104.48
IRRI148	DS	0.20	0.21	0.31	24.32	64.50	95.00	133.88	226.75	13	19	28	241	0.35	0.42	0.69	131.78
	WS	0.24	0.25	0.28	8.48	52.50	65.17	89.46	234.58	11	19	24	168	0.22	0.42	0.53	30.74
Vandana	DS	0.17	0.25	0.31	21.35	66.92	73.67	136.45	263.33	13	16	34	244	0.32	0.38	0.65	149.33
	WS	0.20	0.24	0.26	15.83	59.58	75.33	103.08	244.60	10	20	23	189	0.22	0.34	0.58	51.23
KALI AUS	DS	0.16	0.42	0.25	21.06	65.11	92.83	149.17	286.00	13	26	34	401	0.39	0.49	0.95	129.56
	WS	0.20	0.21	0.27	13.76	48.29	79.56	91.83	262.14	11	19	19	295	0.20	0.33	0.48	52.43
IR74371-70-1-1	DS	0.13	0.40	0.31	30.46	64.64	76.22	133.58	279.14	12	21	30	366	0.30	0.41	0.83	153.95
	WS	0.18	0.22	0.25	22.44	56.33	59.08	92.29	252.67	11	18	20	195	0.25	0.30	0.56	84.84
UPL RI 7	DS	0.15	0.47	0.31	36.40	64.50	80.44	149.50	328.44	15	23	36	397	0.18	0.36	1.02	173.32
	WS	0.19	0.25	0.32	28.57	41.13	82.11	86.83	282.30	10	17	28	271	0.23	0.34	0.78	100.46
IRRI123	DS	0.19	0.31	0.33	52.63	64.17	84.09	140.83	263.67	14	28	46	385	0.33	0.41	0.85	257.49
	WS	0.19	0.24	0.29	37.27	49.54	61.92	101.58	240.63	11	18	27	252	0.19	0.30	0.58	192.14
IR 94225-B-82-B	DS	0.25	0.45	0.37	50.68	82.33	92.33	96.67	342.80	12	28	32	361	0.36	0.36	0.40	151.60
	WS	0.17	0.23	0.23	20.94	53.83	64.50	90.17	280.00	13	17	21	242	0.19	0.27	0.46	122.99
IR 91648-B-289-B-B	DS	0.12	0.34	0.75	37.98	62.67	91.83	154.33	316.00	10	26	31	344	0.35	0.43	1.51	178.47
	WS	0.20	0.24	0.26	26.37	52.17	67.17	102.67	294.33	8	14	19	207	0.17	0.35	0.45	118.21
IR 94226-B-177-B	DS	0.08	0.35	0.47	35.12	78.67	82.83	154.67	343.50	12	19	29	270	0.37	0.47	0.79	159.20
	WS	0.19	0.26	0.26	28.43	70.17	73.08	110.83	247.00	14	21	29	189	0.21	0.37	0.75	149.47

IR 91648-B-153-B-B	DS	0.11	0.49	0.56	20.55	87.01	90.00	157.33	314.00	13	19	46	300	0.43	0.46	1.12	189.09
	WS	0.17	0.23	0.29	36.97	46.33	58.50	90.33	215.50	10	12	18	284	0.25	0.29	0.54	115.84
IR 91648-B-230-B-B	DS	0.14	0.33	0.40	24.57	80.17	84.33	90.67	334.83	11	16	24	262	0.40	0.49	1.44	200.12
	WS	0.15	0.24	0.27	12.56	66.33	83.67	96.83	266.00	10	16	21	236	0.25	0.34	0.57	42.89
AUS BAK TULSI	DS	0.09	0.52	0.67	19.46	58.50	78.00	145.17	289.00	11	22	32	282	0.34	0.42	1.34	133.48
	WS	0.14	0.29	0.57	17.73	58.53	73.53	101.83	244.80	15	24	33	163	0.24	0.49	0.67	66.17
IR 91648-B-32-B-B	DS	0.12	0.23	0.46	40.56	68.00	68.83	132.67	302.83	11	19	28	373	0.31	0.41	1.23	278.23
	WS	0.16	0.21	0.31	28.05	55.50	81.50	106.50	287.00	11	20	22	282	0.20	0.39	0.70	85.10
Population mean	DS	0.26±0.06	0.33±0.09	0.3±0.19	35±17	4.03±0.7	7.9±1.1	13.5±1.8	286±40	13.9±2.4	32±5.3	42±6	322±96	0.34±0.07	0.39±0.17	0.99±0.47	88±35.4
	WS	0.2±0.04	0.27±0.04	0.36±0.08	30±16	4.01±0.5	7.8±1	12.9±1.6	267±28	13.5±2.2	32±5.5	45±8	266±29	0.21±0.05	0.35±0.06	0.68±0.22	81±27.7
Heritability	DS	0.64	0.37	0.73	0.77	0.37	0.75	0.41	0.33	0.47	0.23	0.61	0.69	0.21	0.24	0.79	0.36
	WS	0.59	0.14	0.56	0.71	0.31	0.71	0.39	0.31	0.38	0.14	0.35	0.47	0.17	0.20	0.63	0.25

Table 3.1. Continued.

Number of nodal roots (NR) [NR1, NR2, NR3, NR4: 15, 22, 29 DAS and DTF]. Maximum root length (RL) [RL1, RL2, RL3, RL4: 15, 22, 29 DAS and DTF]. Root dry weight (RDW) [RDW1, RDW2, RDW3, RDW4: 15, 22, 29 DAS and DTF]. Shoot dry weight (SDW) [SDW1, SDW2, SDW3, SDW4: 15, 22, 29 DAS and DTF].

QTLs were detected in all 16 traits; 3 common QTLs were detected for RDW at 15, 22, and 29 DAS during the dry and wet seasons. A total of 2 QTLs were detected on chromosome 5; while only 1 QTL each was detected on chromosome 1, 8, and 12. The least number of QTLs was detected for nodal roots (NR3), nodal roots (NR3), root length at DTF (NR4), and shoot dry weight at 29 DAS (SDW3) only one, i.e., *qNR3-12, qNR4-6, qRL2-1* and *qSDW3-8*.

A total of 2 QTLs detected for number of nodal roots; they were positioned on chromosomes 6 and 12. The QTL detected on chromosome 6, *qNR4-6*, has the highest SNP peak (FDR= 0.00013). There was 1 QTL detected for maximum root length at 22 DAS, positioned on chromosome 1.

There were 5 QTLs detected for root dry weight at 15 DAS, positioned on chromosomes 3, 4, 5, 7 and 10. Interestingly, all the root dry weight QTLs reached the threshold of FDR < 0.01. Out of these five only two QTL on chromosome 3 (*qRDW1-3*) and 5 (*qRDW1-5*) reached the threshold of FDR>0.01. For root dry weight, 4 QTLs were found in dry season. The largest number of QTLs were detected for the root dry weight at 15 DAS are 5.

There were 5 QTLs detected for root dry weight at 15 DAS, positioned on chromosomes 3, 4, 5, 7 and 10. Interestingly, all the root dry weight QTLs reached the threshold of FDR < 0.01. Out of these five only two QTL on chromosome 3 (*qRDW1-3*) and 5 (*qRDW1-5*) reached the threshold of FDR>0.01. For root dry weight, 4 QTLs

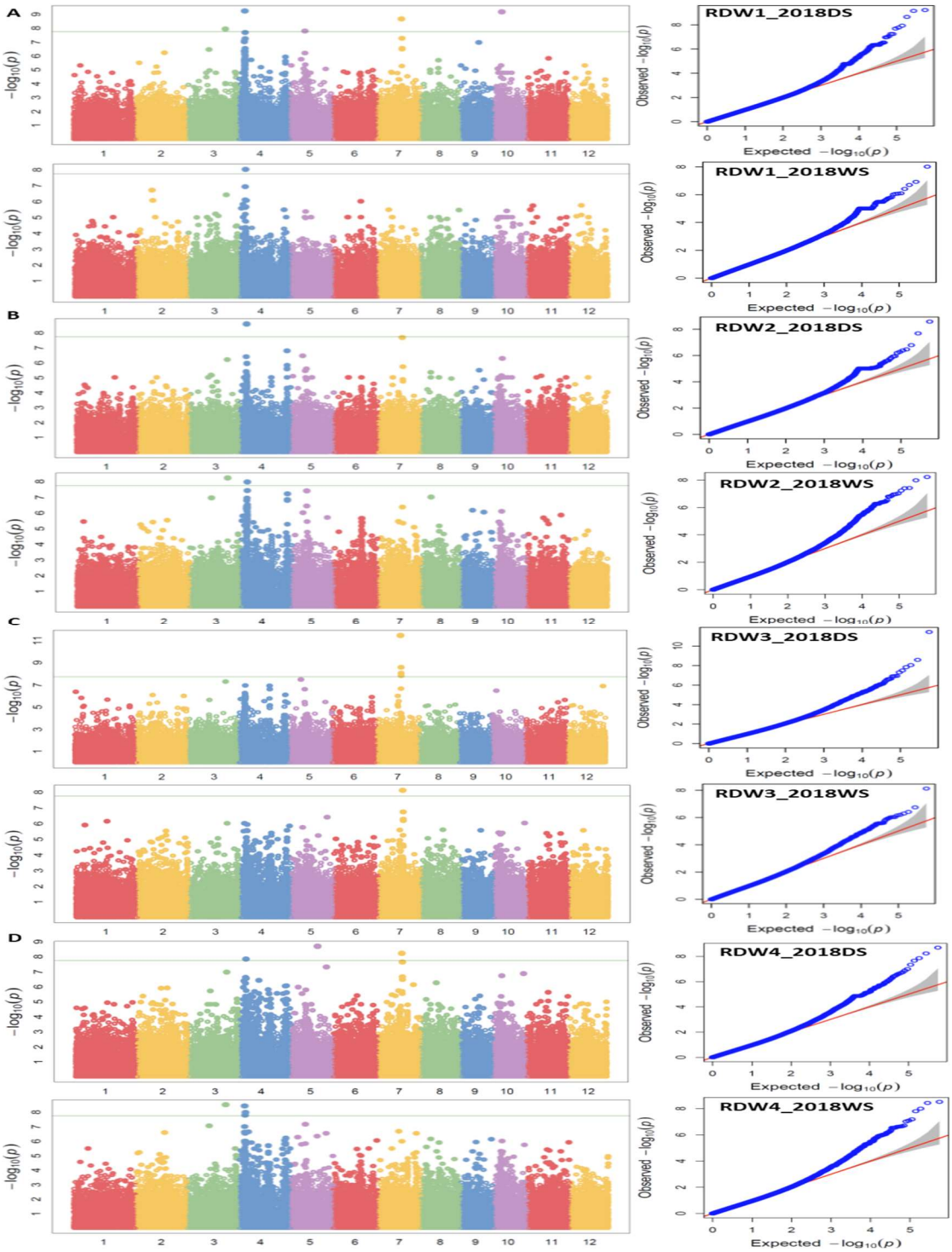

Figure 3.2 Manhattan and Q-Q plots of genome wide association mapping of root and early vegetative vigor traits. Root dry weight (RDW) [RDW1: 15 DAS (A), RDW2: 22 DAS (B), RDW3: 29 DAS (C), RDW4: DTF].

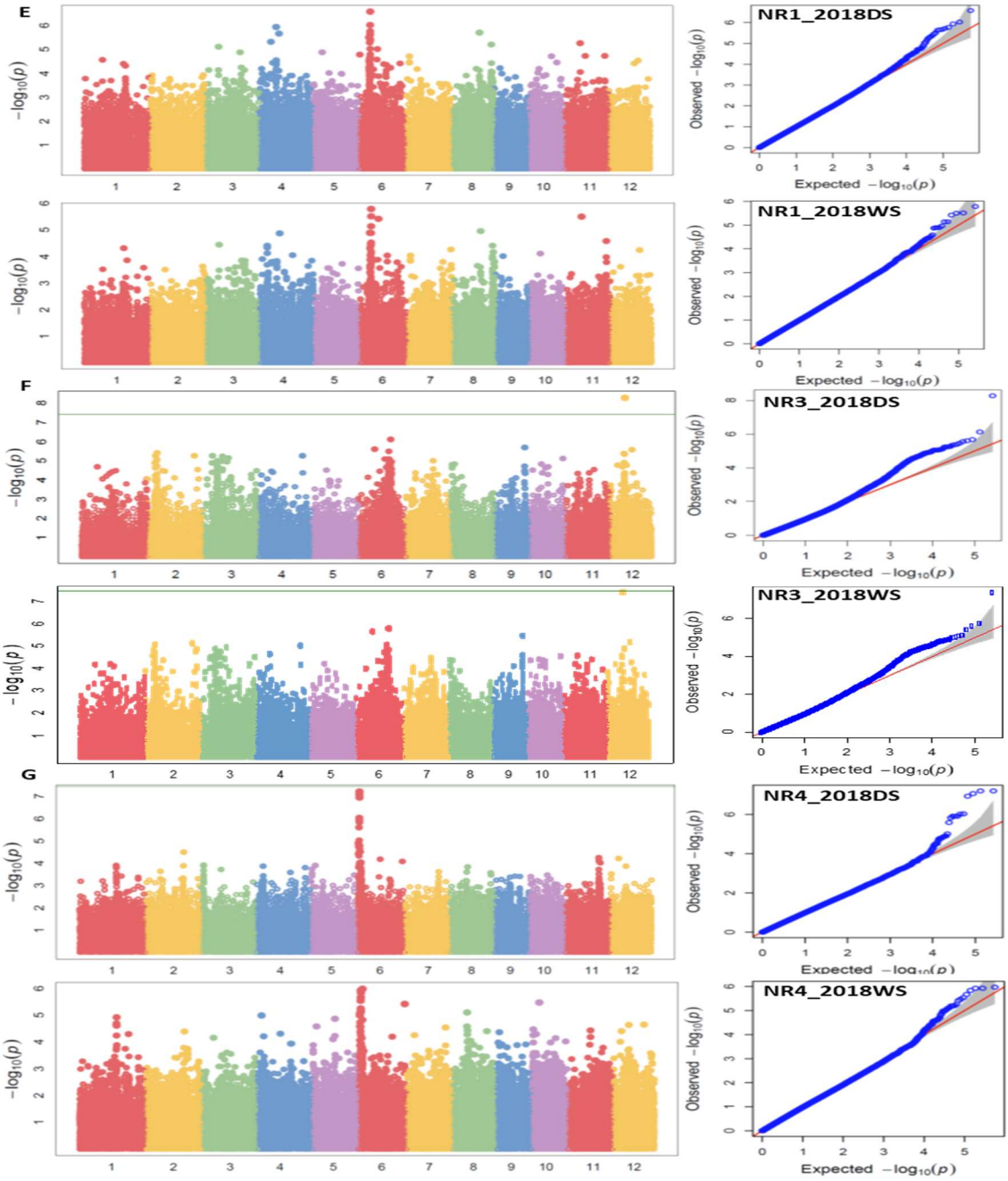

Figure 3.2 Continued.
Root and early vegetative vigor traits. Number of nodal roots [NR1: 15 DAS (E), NR3: 29 DAS (F), NR4: DTF (G)].

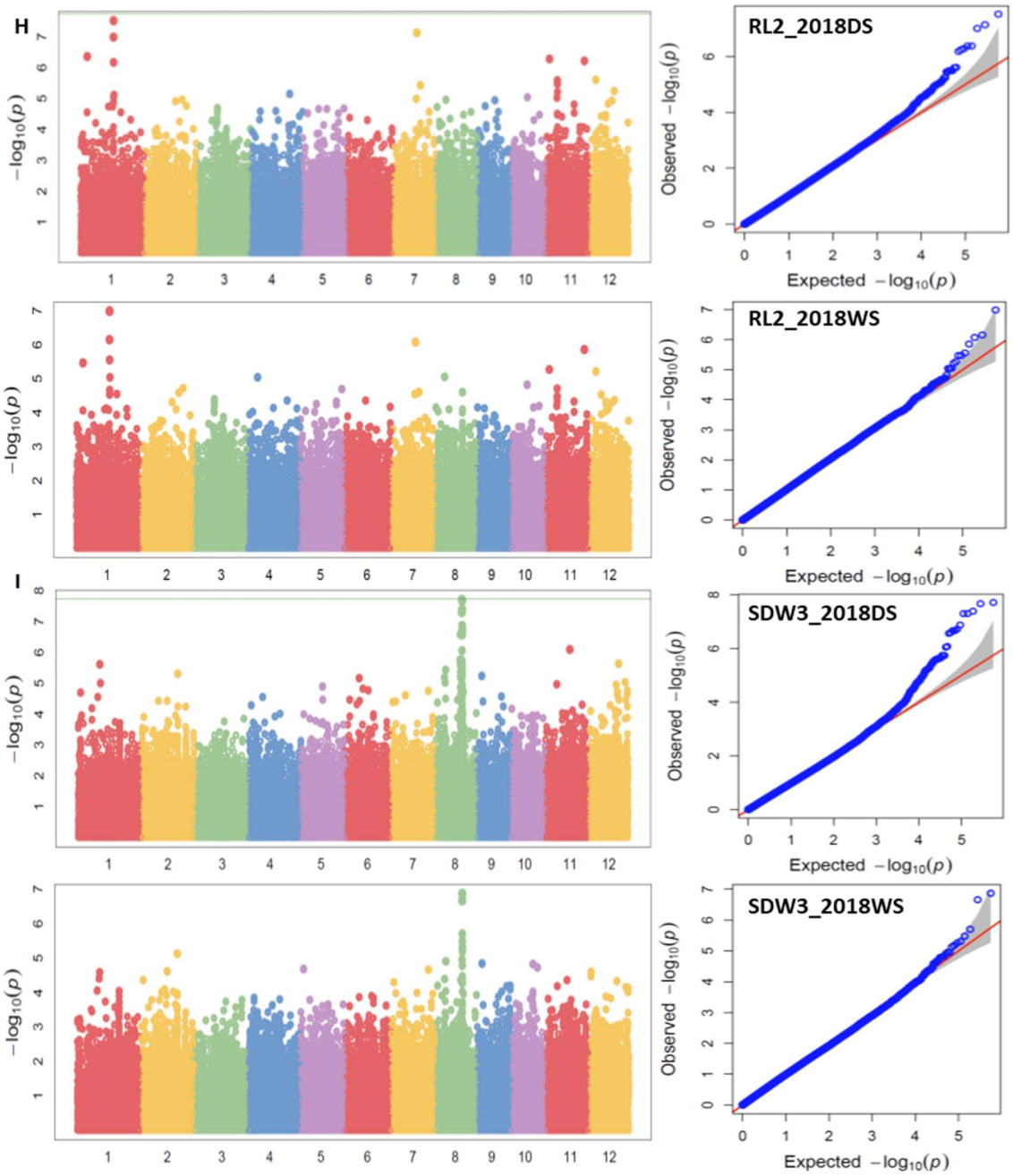

Figure 3.2 Continued.
Root and early vegetative vigor traits. Maximum root length [RL2: 22 DAS (H)]. Shoot dry weight [SDW3: 29 DAS (I)].

were found in dry season. The largest number of QTLs were detected for the root dry weight at 15 DAS are 5.

A few of QTL hot spots harboring QTLs of more than one trait were identified across the chromosomes. For examples, *qRDW1-3* from 2018 DS; *qRDW2-3* and *qRDW3* from WS at S3_26904570. A second QTL hotspot was detected on chromosome 4 for *qRDW1-4, qRDW2-4* and *qRDW3-4* at S4_215776 from 2018 DS and WS. Another QTL hotspot was detected on chromosome 7 between *qRDW1-7* (DS); *qRDW2-7* (DS) and *qRDW3-4* at S7_14096600. The same detected SNPs of those QTLs suggests that the genetic causal controlling those associated traits are tightly linked.

3.4. Discussion

3.4.1. Phenotypic correlations and heritability

The overall trait heritabilities in the dry season were higher than the wet season. On average, lower trait heritability of traits was observed in the wet season (2018WS) compared to the dry season (2018DS), which possibly partly explains the reason for the higher number of QTLs detected in the dry season (**Table 3.2**). During days to 50% flowering, the significant positive correlation of the number of nodal roots with root length, root dry weight and shoot dry weight may suggest the role of roots architecture system associated with the source and sink activity leading to increasing the plant biomass.

3.4.2. QTL identification

The development of rice varieties with improved root systems for upland dry direct seeded ecosystems, however, has been slow and problematic because of the difficulty in the phenotyping of target root traits. Water uptake in rice depends on the root system (Nguyen *et al.*, 1997); consequently, studying the root system is an integral part of understanding its role in the establishment and production of rice in various systems of planting, including direct seeded rice. Roots are the principal plant organ for early vegetative vigor and nutrient and water uptake. Therefore, improving our understanding of the associations between root architecture/plasticity and water scarcity/availability in rice could have a significant impact on rice production and global food security.

3.4.2.1. Root dry weight

This study detected QTLs - *qRDW1-3, qRDW1-4, qRDW1-5, qRDW1-7, qRDW1-10, qRDW2-3, qRDW2-4, qRDW2-7, qRDW3-7, qRDW3-3, qRDW3-4* and *qRDW3-5* for root dry weight at different plant growth stages. Direct associations can be interpreted between root system architecture and early vegetative vigor in the form of phenotype-phenotype relationship and co-located QTLs/genes. For root system architecture numerous known QTLs/genes were detected in close proximity of known genes/QTLs for early vegetative vigor and morphological traits detected in the previous for direct-seeded rice conditions.

Our study identified the first QTL hotspot region of *qRDW1-3, qRDW2-3* and *qRDW3-3* QTLs to be co-localized on chromosome 3 for root dry weight at 15, 22 and 29 DAS with genes - *Osgnome1* and *crl4*. Kitomi *et al.*, (2008) and Liu *et al.*, (2009) have reported *Osgnome1* (guanine nucleotide exchange factor for ADP-ribosylation factor: *OsGNOM1*) and *crl4* (crown rootless 4) genes affects the formation of adventitious roots (ARs) through regulating polar auxin transport (PAT); thereby reducing number of lateral roots (LRs) and partial loss of gravitropism. OsGNOM1 and *CRL4* mutants alters the *OsPIN1, OsPIN2, OsPIN5b* and *OsPIN9* transcription, indicating that OsGNOM1 and *CRL4* is necessary to sustain the natural distribution of auxin during the initiation of adventitious roots.

For root dry weight, we detected QTLs *qRDW1-4, qRDW2-4* and *qRDW3-4* at SNP S4_215776 on chromosome 4 colocalized with gene *SOR1*; identified from a M2 population that was regenerated from seed calli of a japonica rice cultivar - Nipponbare,

Table 3.2. QTLs for root traits and seedling vigor traits detected by GWAS and potentially colocalized genes.

QTL	Season	Chr	Position	P-value	FDR	Effect	Gene	Start	End	Function and Reference
qRDW1-3	DS	3	26904570	1.15×10^{-08}	0.0016	4.29	*Osgnome1*	26962650	26968135	Crown root formation (Liu *et al.*, 2009)
							crl4	26962650	26968135	Crown root formation (Kitomi *et al.*, 2008)
qRDW1-4	Both	4	215776	5.77×10^{-10}	0.00018	13.9	*SOR1*	112189	115776	Root growth angle (Hanzawa *et al.*, 2013)
qRDW1-5	DS	5	11362086	1.66×10^{-08}	0.0019	11.33	*OsEXPA3*	11405956	11404176	Root hair growth, cell size of root vascular bundles by regulating cell expansion (Qiu *et al.*, 2014)
qRDW1-7	DS	7	14096600	2.16×10^{-09}	0.00040	12.89	*RePRP2.1*	14026574	14027700	Root cell elongation and growth (Tseng *et al.*, 2013)
							RePRP2.2	14014393	14015451	Root cell elongation and growth (Tseng *et al.*, 2013)
qRDW1-10	DS	10	1449904	6.72×10^{-10}	0.00018	13.8	*OsSNDP1*	1426573	1430357	Root hair development (Huang *et al.*, 2013)
qRDW2-3	WS	3	26904570	6.01×10^{-09}	0.0029	4.29	*Osgnome1*	26962650	26968135	Crown root formation (Liu *et al.*, 2009)
							crl4	26962650	26968135	Crown root formation (Kitomi *et al.*, 2008)
qRDW2-4	Both	4	215776	2.54×10^{-09}	0.0014	13.07	*SOR1*	112189	115776	Root growth angle (Hanzawa *et al.*, 2013)
qRDW2-7	DS	7	14096600	2.00×10^{-08}	0.00056	11.49	*RePRP2.1*	14026574	14027700	Root cell elongation and growth (Tseng *et al.*, 2013)
							RePRP2.2	14014393	14015451	Root cell elongation and growth (Tseng *et al.*, 2013)
qRDW3-7	Both	7	14096600	3.47×10^{-12}	0.0000019	18.38	*RePRP2.1*	14026574	14027700	Root cell elongation and growth (Tseng *et al.*, 2013)
							RePRP2.2	14014393	14015451	Root cell elongation and growth (Tseng *et al.*, 2013)
qRDW3-3	WS	3	26904570	6.01×10^{-09}	0.0029	4.29	*Osgnome1*	26962650	26968135	Crown root formation (Liu *et al.*, 2009)
							crl4	26962650	26968135	Crown root formation (Kitomi *et al.*, 2008)
qRDW3-4	Both	4	215776	3.77×10^{-09}	0.0011	12.49	*SOR1*	112189	115776	Root growth angle (Hanzawa *et al.*, 2013)
qRDW3-5	DS	5	23760502	2.07×10^{-09}	0.0011	13.94	*eui*	23775727	23785531	Root gravitropism. Starch granule generation in root cap (Zhang *et al.*, 2008)
							OsRPK1	23952607	23946639	Growth of primary embryonic root, lateral and adventitious root number (Zou *et al.*, 2014)
							OsCCaMK	24122248	24118350	CH4 oxidation, N2 fixation activation (Bao *et al.*, 2014)

QTL	Season	Chr	Position	P-value	FDR	Effect	Gene	Start	End	Function
qRL2-1	Both	1	27608104	3.05×10^{-08}	0.0017	6.13	*OsYUCCA1*	27750515	27752683	Root development. Auxin biosynthesis (Yamamoto *et al.*, 2007)
							OsPIN3t	27619548	27624190	Crown root development and gravitropism (Zhang *et al.*, 2012)
qNR1-6	Both	6	887151	6.26×10^{-08}	0.0074	79.79	*OsEXPA17*	510890	511792	Root hair elongation (ZhiMing *et al.*, 2011)
qNR3-12	Both	12	9802344	5.16×10^{-09}	0.00014	10.91	**Unknown**			
qNR4-6	Both	6	887151	6.16×10^{-08}	0.0074	79.79	*OsEXPA17*	510890	511792	Root hair elongation (ZhiMing *et al.*, 2011)
qSDW3-8	Both	8	18312619	1.95×10^{-08}	0.00555	12.76	**Unknown**			

Table 3.2. Continued.

Root dry weight (RDW) [RDW1, RDW2, RDW3, RDW4: 15, 22, 29 DAS and DTF], Maximum root length (RL2) – 22DAS, Number of nodal roots (NR) [NR1, NR3, NR4: 15, 29 DAS and DTF]. Shoot dry weight (SDW3) – 29DAS.

which is known to be associated with the root growth angle. *SOR1* is known to be involved in soil-surface rooting and root gravitropic response at an early growth stage.

A QTL hotspot region was found to be associated with root dry weight at 15, 22 and 29 DAS; a similar region was also reported to be colocalized with genes associated with the root growth, root cell elongation and ABA sensitivity (Tseng *et al.*, 2013). *RePRP2.1* and *RePRP2.2* overexpression reduces root cell elongation in the absence of ABA. Conversely, knockdown of *RePRP2.1* and *RePRP2.2* genes reduces the root sensitivity to ABA implying that both the genes have critical function in stress adjustment of root growth and development.

Two major root dry weight QTLs were identified on chromosome 5 (*qRDW1-5*) and 10 (*qRDW1-10*) are colocated with the key root architecture genes *OsEXPA3* and *OsSNDP1* reported by Qiu *et al.*, (2014) and Huang *et al.*, (2013), respectively. Qiu *et al.*, (2014) reported that the *OsEXPA3* gene is necessary for normal root system growth and development at seedling stage. The RNAi plants had significant reduction in the primary root length, lateral root density, shorter root hair, decreased cell length of root vascular bundles and cell growth, indicating *OsEXPA3* has an important role in root cell-wall loosening. *OsSNDP1* is the key gene involved in the normal root hair elongation and growth in rice (Huang *et al.*, 2013).

One QTL (*qRDW3-5*) on chromosome 5 was detected for root dry weight at 29 DAS in 2018 DS, peaked at S5_23760502 with p-value 2.07×10^{-9}, FDR=0.0011 and with an effect size of 13.94 gm. This QTL is colocalized with important root system architecture gene i.e., *eui* (elongated uppermost internode) (Zhang *et al.*, 2008), *OsRPK1*

(Zou *et al.*, 2014) and *OsCCaMK* (Ca^{2+}/calmodulin-dependent protein kinase) (Bao *et al.*, 2014), respectively. These gene are directly involved in root gravitropism, starch granule generation in root cap, growth of primary embryonic root, lateral root number, adventitious root number, CH4 oxidation and N2 fixation activation, suggesting the role of these genes in increasing the root dry weight in rice for normal root development at different growth stages.

3.4.2.2. Root length

Numerous factors affect the root system architecture of the rice plant; plant root length is one of the vital determinants when it comes to adaptation in variable growing environments. Flowering time is negatively correlated to grain yield (-0.06 and -0.02 in 2018DS and 2018WS, respectively). In this study, root length at 22 DAS (*qRL2-1*) associated SNP S1_27608104 present on chromosome 1 colocalized with crown root development and gravitropism regulator (*OsYUCCA1*; Yamamoto *et al.*, 2007 and *OsPIN3t* (PIN-FORMED 3t; Zhang *et al.*, 2012); induces the auxin (IAA) biosynthesis by producing extensive crown hairy roots.

3.4.2.3. Number of nodal roots

Number of nodal roots (*qNR4-6*), S6_887151 showed a significant marker-trait association detected on chromosome 6 which was located in close proximity to *OsEXPA17* (rice EXPANSIN-17; ZhiMing *et al.*, 2011) candidate gene controlling root hair elongation. *OsEXPA17* is expressed exclusively in roots playing vital role in the cell wall loosening protein which is essential for the root hair elongation in rice. Moreover, SNP S12_9802344 is also colocalized with a major QTL *qtl12.1* (Bernier *et al.*, 2009)

associated with drought resistance in rice. No known gene was detected in this region for root system architecture.

3.4.2.4. Shoot dry weight

For shoot dry weight at 29 DAS, 1 QTL was identified for both the seasons (DS and WS) on chromosome 8 (*qSDW3-82*). This QTL region on chromosome 8 detected to be associated with shoot dry weight do not contain any previously known QTLs/gene for shoot dry weight traits.

3.4.2.5. QTL hotspots for root traits and early vegetative vigor

Chromosome regions related with multiple traits are extremely valuable as they may well represent important plant growth regulators. These genomic regions indicate the presence of a pleiotropic effect gene affecting multiple biological processes simultaneously (Pelgas *et al.*, 2011). We identified 3 such QTL hotspots: on chromosome 3, on chromosome 4 and on chromosome 7. These genomic regions were identified across the seasons which illustrates the consistency of these detected QTLs. Similar chromosome regions have been reported by several previous studies in diverse rice mapping populations. These QTLs can be further characterized for candidate gene identification and validation, which can be used in marker-assisted breeding programs to DDSR conditions Therefore, the QTL hotspot detected can further be utilized in rice genetic improvement to dissect the complex architecture of root traits, not only in DDSR conditions but also in other water related stresses.

4. IDENTIFY SIGNIFICANT GENOMIC REGIONS CONTROLLING NUTRIENT UPTAKE

4.1. Introduction

In the twenty-first century, rice production is anticipated to be more restricted due to decreased accessibility and higher expenditure of nutrient and water reserves (Lal *et al.*, 2007). Nutrient and water uptake dynamics in rice depend on the root system architecture (Nguyen *et al.*, 1997), but poor knowledge of nutrient and water management further reduces the production. Puddled transplanted rice (PTR) establishes conditions to boost nutrient availability (SANCHEZ, 1973); however, it affects soil properties adversely by producing hard-pans, which reducing rice production (Sharma *et al.*, 2003; Tripathi *et al.*, 2007).

Dry direct-seeded rice (DDSR) production system is limited due to the bottlenecks of maintaining the nutrients dynamics, which is the foundation of effective biological and economic yield. DDSR faces the challenge of decreased nutrient availability, particularly nitrogen, phosphorus and iron, compared to PTR, which adversely affects the agronomic traits and yield.

During the grain-filling phase, Katsura *et al.*, (2010) reported having high grain yield in DDSR aerobic rice compared to PTR due to elevated nitrogen accumulation. Poor foliage nitrogen concentration at heading concludes in greater spikelet sterility, lower panicle density and fewer grains per panicle (Dingkuhn *et al.*, 1990; Kabir *et al.*, 2008). Phosphorus uptake is positively correlated with lateral root growth and negatively correlated with root hair length, where fields under reduced phosphorus conditions show

extensive root hair growth (Lynch & van Beem, 1993; Zhang *et al.*, 2003). A few studies reported QTLs correlated with nutrient uptake in rice (Sandhu *et al.*, 2019; Singh *et al.*, 2017; Subedi *et al.*, 2019), wheat (Su *et al.*, 2009), maize (Zhu *et al.*, 2005), and soybean (Liang *et al.*, 2010). A better understanding of root-related traits and how root-trait QTLs interact to affect soil resource acquisition across various environments will be important in breeding direct-seeded rice varieties. This highlights the urgency of exploring the genetic diversity to develop the nutrient efficient and low input rice production system. The present study aims to dissect the genetic control of root traits associated with nutrient uptake under DDSR.

4.2. Materials and Methods

4.2.1. Phenotyping

All the phenotypes for nitrogen and chlorophyll content were sampled by choosing 3 plants randomly within each plot. Randomly, 3 plants were used to check Soil Plant Analysis Development (SPAD) chlorophyll meter and Leaf Chlorophyll Content (LCC) scale.

For nutrient uptake estimation, a total of 100 accessions (50 high and 50 with low leaf chlorophyll content) along with 10 checks were selected at booting-stage for nutrient analysis. Total Nitrogen is determined by Kjeldahl Nitrogen determination method. Inductively Coupled Plasma-Optical Emission Spectrometer (ICP-OES) is used to determine trace elements (P, K, Fe and Zn) in the booting stage. Nutrient uptake (N, P, K, Fe and Zn) estimation is performed at Analytics service lab, IRRI.

4.2.2. Plant materials, field experiment and management, genotyping, population structure and association mapping

The procedure and plant material used in Chapter 4 for genotyping, population structure and association mapping analysis are similar to Chapter 2.

4.3. Results

4.3.1. Phenotypic correlations and heritability

Overall, the correlation coefficients for nutrient traits were weak (-1< r <1) (Fig 4.1). Among the traits, SPAD (2018WS) had the highest correlation coefficient with LCC in both the seasons (0.75 and 0.76). SPAD (2018WS) also have positive correlation with SPAD (2018WS), LCC (DS and WS). Phosphorus (P) was positively correlated with SPAD (DS and WS), LCC (DS and WS), Nitrogen (N), Potassium (K), and Iron (Fe); except zinc content (Zn) (-0.1). On the other hand, Zinc (Zn) is negatively correlated to Nitrogen (N), Potassium (K), Iron (Fe) and Phosphorus (P). Among the nutrient traits, correlation between Phosphorus (P) and Potassium (K) was highest (.27).

For SPAD and LCC, the trend for trait heritability is exactly same to the correlation coefficient, where in general the trait heritability in 2018 WS was higher compared to those of 2018 DS (Table 4.1). Among all the nutrient traits, Nitrogen (N) and Iron (Fe) had the highest heritability across seasons (N: 0.58 in DS and Fe: 0.54 in DS); while Phosphorus (P), Potassium (K) and Zinc (Zn) has lower heritability during the dry season (P: 0.24, K: 0.20 and Zn: 0.28).

4.3.2. Identification of QTLs by GWAS

A total of 8 QTLs were detected in our GWAS of the 7 traits across two seasons for SPAD and LCC; and one season for nutrient uptake based on FDR<0.01 (**Table 4.2**); among those, 2 QTLs reached FDR<0.001. Those 8 QTLs were: a QTL for leaf chlorophyll content (*qLCC-9*), two QTLs for SPAD (*qSPAD-1-1* and *qSPAD-1-2*), one

QTL for nitrogen (*qN-6*), a QTL for phosphorus (*qP-12*), one QTL for Iron (*qFe-2*), and two QTL for zinc content (*qZN-7-1* and *qZN-7-2*) (**Table 4.2**).

From the total 3 QTLs identified during 2018 DS and WS for SPAD and LCC, it is worth noted that among those 2 QTLs were detected in both seasons. Some examples of GWAS results on SPAD, LCC and nutrient uptake traits were presented in Manhattan and Q-Q (quantile-quantile) as shown in **Fig. 4.2**. Candidate genes and QTLs from previous reports potentially colocalized with our QTLs were also listed in Table 4.2.

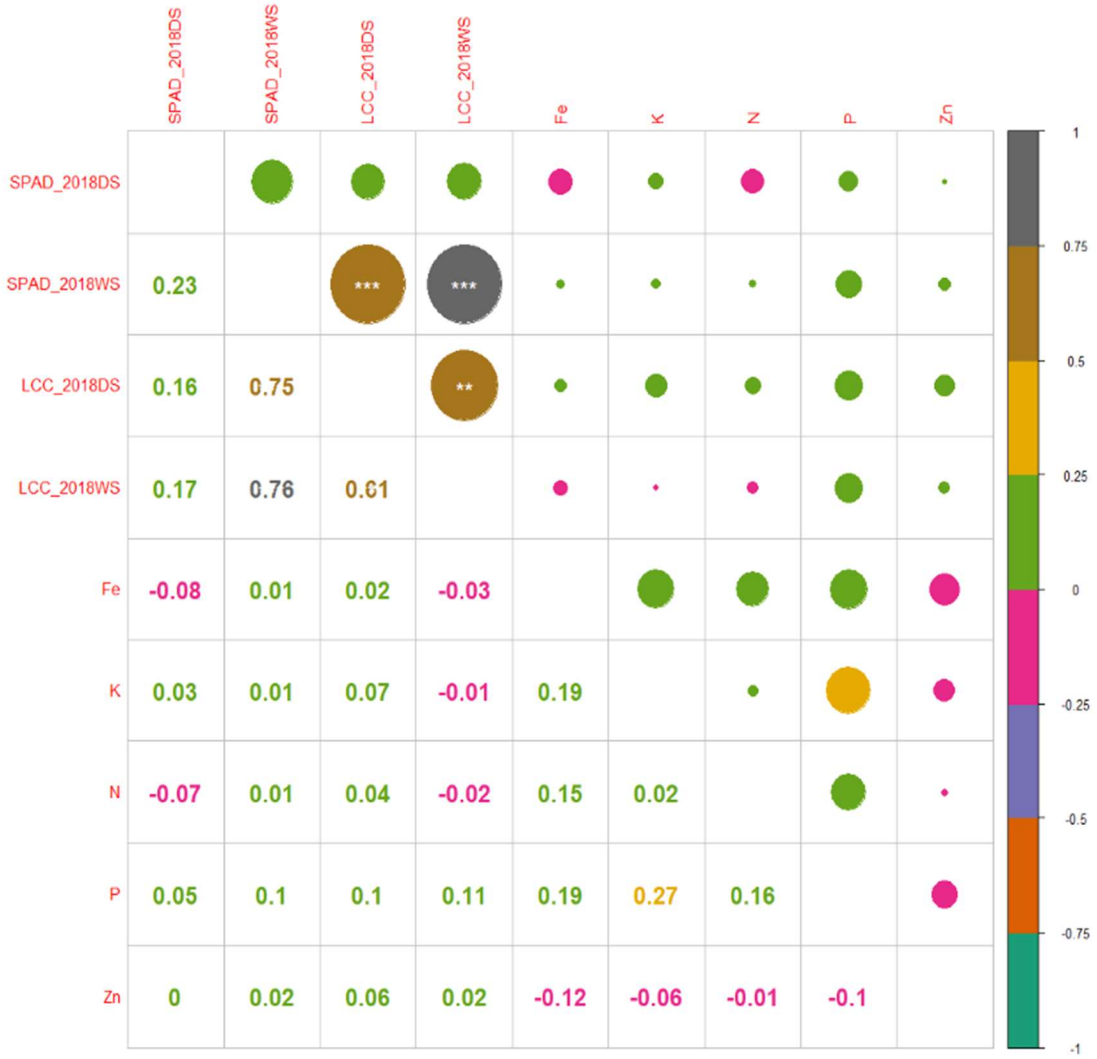

Fig. 4.1 Correlation plot of SPAD, LCC and nutrients uptake traits: P < 0.05 *, P < 0.01 **, P < 0.001 ***. The eight color indicates the correlation scale from -1 to 1 among different traits. SPAD: Soil Plant Analysis Development chlorophyll meter, LCC: leaf chlorophyll content (Scale: 2-5), N: Nitrogen (%), P: Phosphorus (%), K: Potassium (%), Fe: Iron (mg/kg), Zn: Zinc (mg/kg).

Table 4.1. Mean data and trait heritability of checks and GWAS population for SPAD, LCC and nutrient uptake traits in dry and wet season 2018.

Traits	Season	SPAD	LCC	N	P	K	Fe	Zn
IRRI 154	DS	32.62	3.4	0.991	0.200	1.80	192.0	32.5
	WS	12.75	2.92	-	-	-	-	-
IR 87707-446-B-B-B:7	DS	32.33	3.1	0.770	0.177	2.32	505.5	19.5
	WS	16.33	3.5	-	-	-	-	-
IRRI148	DS	24.82	2.8	-	-	-	-	-
	WS	13.64	3.3	-	-	-	-	-
Vandana	DS	31.49	3	0.817	0.162	1.72	206.5	21.5
	WS	12.43	2.8	-	-	-	-	-
KALI AUS	DS	27.54	2.8	1.078	0.192	1.54	332.0	26.5
	WS	11.04	2.6	-	-	-	-	-
IR74371-70-1-1	DS	30.67	2.8	0.957	0.261	2.25	469.5	22.0
	WS	12.25	2.9	-	-	-	-	-
UPL RI 7	DS	34.56	3	1.160	0.195	2.17	462.5	15.0
	WS	14.08	3.1	-	-	-	-	-
IRRI123	DS	35.85	3.3	0.988	0.207	2.18	253.5	24.5
	WS	13.74	3.3	-	-	-	-	-
IR 94225-B-82-B	DS	29.27	3.1	0.804	0.155	1.95	105.0	33.0
	WS	11.87	2.9	-	-	-	-	-
IR 91648-B-289-B-B	DS	27.87	3.4	-	-	-	-	-
	WS	15.43	3.6	-	-	-	-	-
IR 94226-B-177-B	DS	28.08	2.7	0.639	0.200	2.49	530.0	27.5
	WS	13.72	3.1	-	-	-	-	-
Traits	**Season**	**SPAD**	**LCC**	**N**	**P**	**K**	**Fe**	**Zn**
IR 91648-B-153-B-B	DS	31.27	3.1	-	-	-	-	-
	WS	15.17	3.2	-	-	-	-	-
IR 91648-B-230-B-B	DS	29.1	3.3	-	-	-	-	-
	WS	17.53	3.8	-	-	-	-	-

AUS BAK TULSI	DS	34.52	3.3	-	-	-	-	-
	WS	13.18	3.0	-	-	-	-	-
IR 91648-B-32-B-B	DS	36.56	3.4	0.928	0.190	1.88	128.5	38.0
	WS	19.70	3.6	-	-	-	-	-
Population means	DS	32.39 ± 8	3.05 ± 0.55	0.94 ± 0.27	0.21 ± 0.04	1.96 ± 0.4	545.6 ± 376.8	28.6 ± 10.7
	WS	31.09 ± 4	1.74 ± 0.12	-	-	-	-	-
Heritability	DS	0.50	0.42	0.58	0.24	0.20	0.54	0.28
	WS	0.59	0.44	-	-	-	-	-

Table 4.1. Continued.

SPAD: Soil Plant Analysis Development chlorophyll meter, LCC: leaf chlorophyll content (Scale: 2-5), N: Nitrogen (%), P: Phosphorus (%), K: Potassium (%), Fe: Iron (mg/kg), Zn: Zinc (mg/kg).

4.4. Discussion

4.4.1. Phenotypic correlations and heritability

The positive significant correlation of SPAD and LCC with nutrient uptake indicating the role of SPAD and LCC techniques in the identification of plants with nutrients uptake traits. LCC was significantly and positively correlated with SPAD: supporting the use LCC as an optional technique to SPAD meter to detect chlorophyll content in reproductive and grain filling stage in water stress condition. The heritability's of SPAD and LCC in wet season was higher than dry season. Lower trait heritability of traits observed in wet season for SPAD (2018WS) compared to dry season (2018DS), which possibly explains the reason for two QTLs detected in dry season (**Table 4.2**).

4.4.2. QTL identification

4.4.2.1. Soil Plant Analysis Development (SPAD: chlorophyll meter) and leaf chlorophyll content (LCC)

This study detected QTLs *qSPAD-1-1* and *qSPAD-1-2* for SPAD. Our study identified the co-location of *qSPAD-1-1* QTL on chromosome 1 for SPAD. Miyoshi *et al.*, (2003) have reported this genomic region to be associated with chloroplast development. The *OsHAP3A-RNAi* lines demonstrated pale green leaves with reduced chloroplast development, suggesting its role in regulating the normal development of chloroplasts.

The colocalization SNP S1_26268177 for SPAD (*qSPAD-1-2*) on chromosome 1 was detected in the close proximity with the previously identified gene for nitrogen and carbon content in rice viz., *OsAAT5* (rice amino acid transporters-5) (Lu *et al.*, 2012)

(**Table 4.2**). On chromosome 1, *osaat5 showed increased* levels of and carbon and reduced nitrogen content in the whole plants (Lu *et al.*, 2012); suggesting its role nitrogen uptake and homeostasis in rice plants.

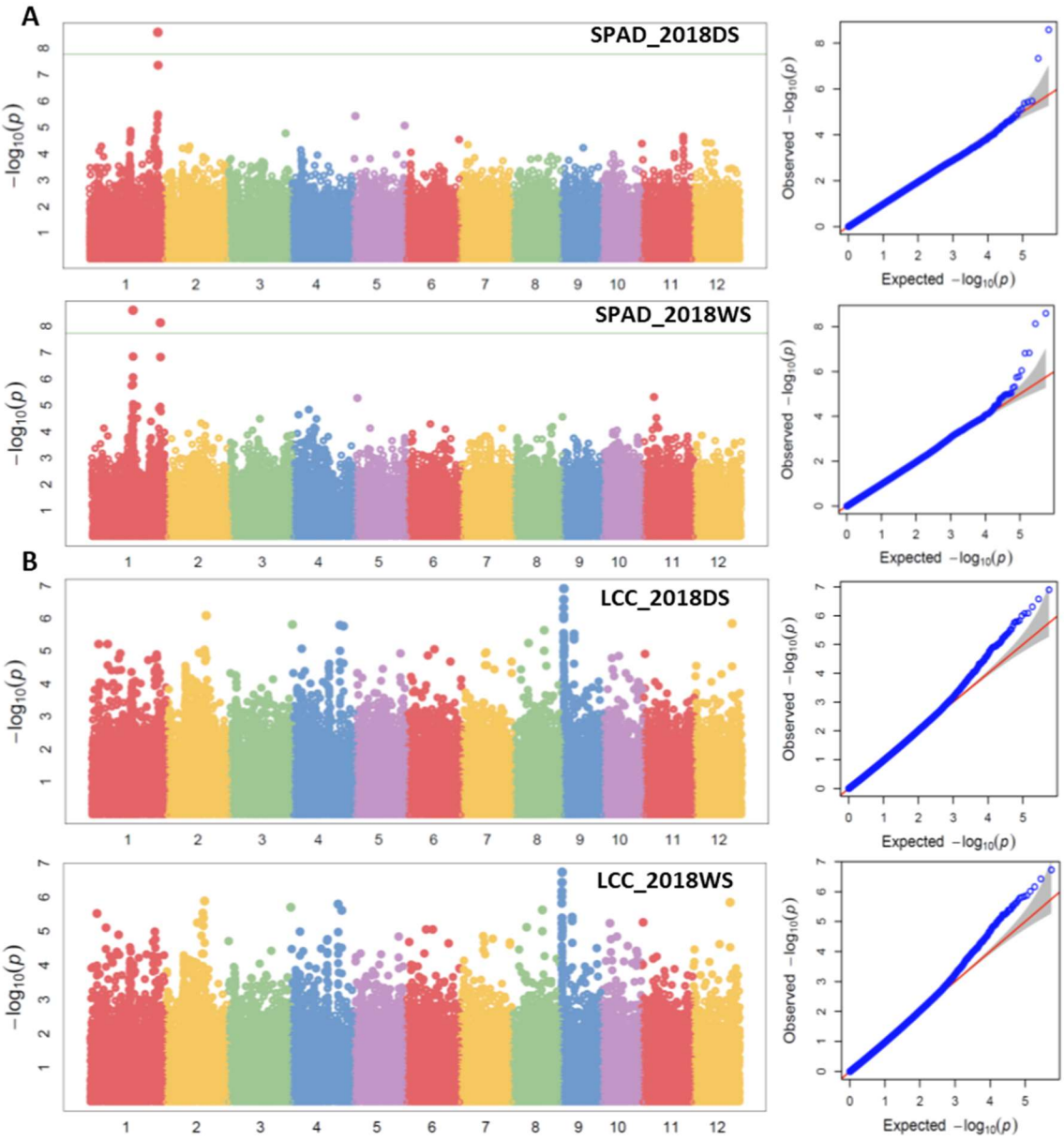

Figure 4.2 Manhattan and Q-Q plots of genome wide association mapping of SPAD (A), and LCC (B) from 2018 Dry (DS) and Wet (WS) seasons.

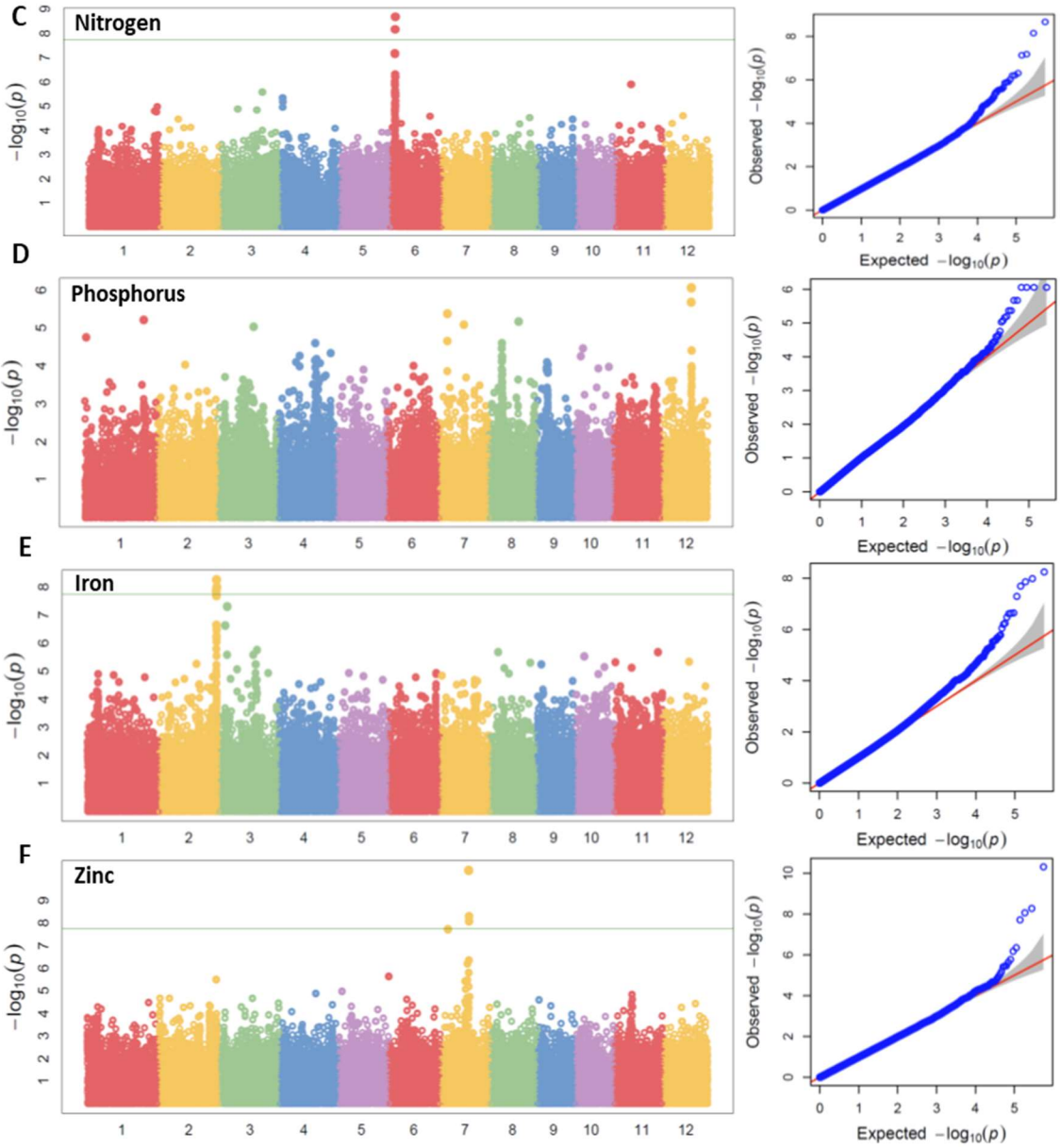

Figure 4.2 Continued.
Nutrient uptake traits - Nitrogen (C), Phosphorus (D), Iron (E) and Zinc (F) from 2018 dry (DS) seasons.

In the present study, leaf chlorophyll content (LCC) was found to be associated with SNP S9_412246 present on chromosome 9, which is not colocalized with any known QTL/genes, encouraging further study for chlorophyll content in the region.

4.4.2.2. Nitrogen uptake

For nitrogen uptake (N), S6_6591538 showed a significant marker-trait association detected on chromosome 6 which was located in close proximity to *OsAAT49* (Lu *et al.*, 2012) major gene controlling nitrogen accumulation in rice. *OsAAT49* expression is up regulated in the shoots due to nitrogen starvation in *OsAAT49* mutant lines. *OsAAT49* mutant lines showed significant reduction in the grain yield (37%) explaining importance of *OsAAT49* in nitrogen accumulation. Furthermore, *OsAAT49* is necessary for nitrogen accumulation and transport (Lu *et al.*, 2012).

4.4.2.3. Phosphorus uptake

For phosphorus uptake, a QTL was identified on chromosome 12 for DS. This genomic region on chromosome 12 was associated with the previously known gene for phosphate uptake *PSTOL1* (Gamuyao *et al.*, 2012). The phosphorus-starvation tolerance-1 (*PSTOL1*) gene from Kasalath is essential for the significant enhancement root growth facilitating rice to absorb more phosphorus and other nutrients in phosphorus deficient soil, which improves rice grain yield productivity considerably.

Another gene at same genomic region has been found to be associated with lateral root development i.e., *OsORC3* (origin recognition complex3). The *orc3* knockdown mutant demonstrated a temperature-dependent abnormal lateral root growth, developing a dwarf phenotype. In DDSR conditions, phosphorus is known to be positively correlated with lateral root growth under reduced phosphorus conditions, plants show extensive lateral root growth.

4.4.2.4. Iron uptake

For iron uptake panicle length, a QTL was identified on chromosome 2. The SNP S2_26968483 on chromosome 2 was also detected to be co-located with two iron uptake and translocation genes like *OsYSL2* (Ishimaru *et al.*, 2010) and *OsYSL15* (Inoue *et al.*, 2009), which are necessary for iron (Fe) uptake and homeostasis for increased Fe accumulation in the grain. The *OsYSL2i-RNAi* lines had decreased iron content in the shoots and seeds, which is monitored by the sucrose transporter promoter. *OsYSL2* is a critical Fe-nicotianamine transporter important for Fe translocation, especially in the shoots and endosperm (Ishimaru *et al.*, 2010). *OsYSL15* knockout plants demonstrated reduced germination and seedling growth, which can be compensated by providing elevated iron source (Inoue *et al.*, 2009).

4.4.2.5. Zinc uptake

For zinc uptake, this study 2 detected QTLs were *qZN-7-1* and *qZN-7-2* on chromosome 7. On chromosome 7, SNP S7_7443671 is also associated with zinc uptake and translocation. One of the major gene for zinc uptake was detected on chromosome 7 i.e., *OsZIP8*, which translates a plasma membrane zinc transporter in rice (Lee *et al.*, 2010). Overexpression lines of *OsZIP8* gene showed varied zinc allocation in rice, increased accumulation and transport of zinc in roots suggests its implication in zinc homeostasis in the regulation of rice growth and development.

No other known gene was detected to be colocalized with the QTL *qZN-7-2*.

4.2. QTLs for SPAD, LCC and nutrient uptake traits detected by GWAS and potentially colocalized genes/QTLs.

QTL ID	Season	Chr	Position	P-value	FDR	Effect	Gene	Start	End	Function and Reference
qSPAD-1-1	Both	1	37529255	$2.53x10^{-9}$	0.00141	14.9	*OsHAP3A*	37512183	37514517	Chloroplast development (Miyoshi *et al.*, 2003)
qSPAD-1-2	WS	1	26268177	$2.49x10^{-9}$	0.00139	3.8	*OsAAT5*	26208906	26210836	Nitrogen content (Lu *et al.*, 2012)
qLCC-9	Both	9	412246	$1.24x10^{-7}$	0.00493	14.7	--			
qN-6	DS	6	6591538	$2.09x10^{-9}$	0.00116	0.48	*OsAAT49*	6681438	6684948	Nitrogen content (Lu *et al.*, 2012)
qP-12	DS	12	15442977	$8.68x10^{-7}$	0.00107	-12.9	*PSTOL1*	15463109	15618170	Phosphate uptake (Gamuyao *et al.*, 2012)
							OsORC3	15475562	15477583	Lateral root development (Chen *et al.*, 2013)
qFe-2	DS	2	26968483	$5.58x10^{-9}$	0.0025	180	*OsYSL2*	27058478	27061460	Iron translocation (Ishimaru *et al.*, 2010)
							OsYSL15	27086997	27091329	Iron uptake (Inoue *et al.*, 2009)
qZn-7-1	DS	7	7443671	$1.90x10^{-8}$	0.0026	0.13	*OsZIP8*	7426642	7429733	Zinc uptake and translocation (Lee *et al.*, 2010)
qZn-7-2	DS	7	16931388	$4.66x10^{-11}$	0.000026	0.17	--			

SPAD: Soil Plant Analysis Development chlorophyll meter, LCC: leaf chlorophyll content (Scale: 2-5), N: Nitrogen (%), P: Phosphorus (%), K: Potassium (%), Fe: Iron (mg/kg), Zn: Zinc (mg/kg).

5. DESIGN AND VALIDATION OF GUIDE RNAS FOR GENE EDITING OF PURPLE LEAF COLOR IN RICE

5.1.Introduction

Plants make an extraordinary range of various color pigments utilized in defense and photosynthesis. One of the essential secondary metabolites pigment from the flavonoid class is anthocyanins. Accumulation of anthocyanins in rice leaf tissues is involved in various biological roles, for instance, hormonal responses, protection against biotic and abiotic stress, and safety against UV radiation (Ithal & Reddy, 2004). Several studies have reported the health benefits of anthocyanins in human ingestion, which are - reduced risk of diabetes, melanomas, cardiac disease, and other chronic illnesses (Deng *et al.*, 2013).

The biosynthetic pathway for anthocyanin comprises conserved proteins accountable for the synthesis of a sequence of metabolites together with regulatory proteins that control the expression of anthocyanin in plant tissues (Yuan & Grotewold, 2015). The anthocyanin biosynthetic pathway of purple rice involves *OsC1* (the homolog of *ZmC1* in maize; the *MYB-R2R3*-type transcription factor Colored 1), *OsB1*, and *OsB2* (the homologs of Booster 1 of maize; basic helix-loop-helix [bHLH]-type transcription factor). This complex triggers the biosynthetic pathway for anthocyanin in rice leaves (Sakamoto *et al.*, 2001). Also, *OsB2* promotor diverged gene *Kala4* is responsible for the black pericarp of black rice (Oikawa *et al.*, 2015).

5.2.Materials and Methods

5.2.1. Plant Materials

The purple rice variety (*Oryza sativa* L. *indica*) was collected from the Beaumont Research Center, Texas. Rice plants were grown in a greenhouse at 30°C day and 24 °C night with 12 hours of light and dark cycle.

5.2.2. Genomic DNA extraction, PCR assay and designing of gRNAs

Rice genomic DNA was extracted as described previously using the CTAB method (Nekrasov *et al.*, 2017). DNA sequences for both *OSB1* and *OSB2* genes were retrieved from the RAP-DB and MSU databases (Os04g0557800 - LOC_Os04g47080.1 and Os04g0557500 -LOC_Os04g47059.1, respectively). The genomic sequences from both the databases were then compared to find a shared conserved region for designing sgRNA. Site-specific primers were designed to confirm the accuracy of these sgRNA sequences. The target regions were amplified with specific primers (IDT, Inc) (**Appendix B**) using Q5 high fidelity Taq-DNA polymerase (New England Biolabs). The PCR product was separated using a 1% agarose gel with SYBR Safe DNA gel staining dye (Thermo Fisher Scientific) and then quantified using the Gel Doc (Azure Biosystems (c200)). Selected PCR products were purified and cloned into the Zero Blunt TOPO PCR cloning vector (Invitrogen) to confirm the target genomic region by DNA sequencing. Five clones per line were used to verify the sgRNA region. The pair of sgRNAs for each gene were preferred to be designed in the first two exons to confer gene disfunction. For this experiment, we have used two different sgRNA designing software, i.e., CRISPR-Direct and CRISPR-P, to select the common sgRNA with the

least number off-targets. To confirm the sgRNA efficiency, the gRNAs were ordered from Synthego, and once confirmed, the gRNA-tRNA construct was ordered from Genscript.

5.2.3. *In vitro assay*

The individual sgRNAs ordered from Synthego were subjected to check the efficacy of each individual gRNA using Mali *et al.*, (2013) *in vitro* digestion protocol. The *in vitro* digestion protocol includes a reaction of PCR purified substrate DNA (3 µl), NE buffer (3 µl), 1 µM Cas9 nuclease and 300nM gRNA.

5.2.4. Rice protoplast transfection

The purple rice protoplast was isolated and transfected, as previously reported by Shan *et al.*, (2014). To test the transfection efficiency of these isolated purple rice protoplasts, approximately 15 µg of GFP plasmid DNA was used to transfect 2×10^5 protoplasts in 10% PEG. The protoplast samples were visualized in a microscope after 48 hours of transfection.

5.2.5. Construction of gene and plasmid vector

The gRNA-tRNA construct ordered from Genscript will be phosphorylated, annealed, and then ligated into the pMOD_B2103 Golden Gate cloning vector. From the cloning vector, the synthesized PTGs will then be inserted into the Bsal-digested for stable rice transformation. The plasmid vector pRGEB32 (cloning vector) will be used to express transiently with three individual promotors U3, CmYLCV, and U6 along with the PTG and Cas9. The destination vector, pRGEB32 (binary vector) will be used for transformation.

5.2.6. Methods of transformation

Agrobacterium-mediated transformation (AMT) will be used to deliver our desired construct (Sahoo & Tuteja, 2012). Alongside AMT, a biolistic particle delivery system using callus shoot apical meristems (SAMs) will also be used for transformation, as this method of delivery will save months in the regeneration process (Hamada *et al.*, 2017). For transient expression, GFP will be used to confirm the in vitro expression and transformation into the callus/SAMs.

5.2.7. Screening and identification of mutants

The preliminary screening and identification of mutant will be performed using PCR, restriction enzyme digestion, and DNA sequencing. Later the Blunt-end TOPO clones will be used to confirm the mutant plants. Moreover, the green leaf can be utilized as the visual marker to screen the mutants.

5.3.Results and Discussion

5.3.1. Sequence confirmation

To confirm the *OSB1* and *OSB2* transcript sequences, first, the genomic DNA was used to amplify targeted sequences using the designed primers (**Appendix B**). The PCR product sizes were confirmed on 1% agarose gel with the band size of 794 bp for *OSB1* and 1021 bp for *OSB2*, respectively (Fig 5.1).

For *OSB1*, after comparing the sequenced PCR cloned product and sequence retrieved from both the databases (RAP-DB and MSU V.7), we found a 98.81% similarity in the rice genomic DNA extracted and given database sequence. Here, we observed three base pair differences in the second exon of *OSB1* (**Fig 5.2**). Using this sequenced shared region between RAP-DB and MSU database for *OSB1*, we developed 2 gRNAs with the help of the gRNA designing software, i.e., CRISPR-Direct and CRISPR-P (**Appendix C**). Here, we selected the gRNAs, which are common from both software with the least off-target effect.

For *OSB2*, a similar procedure was repeated. Here, we found 97.7% and 92.1% similarity in the rice genomic DNA extracted and given database sequence for exon 1 and 3, respectively. We have observed two base pair differences in the first exon and three base pair differences in the third exon of *OSB2* (**Fig 5.3**). Using this sequenced shared region between RAP-DB and MSU database for *OSB2*, we developed 3 gRNAs using CRISPR-Direct and CRISPR-P (**Appendix C**). We selected the gRNAs which are common from both software tools with the least off-target effect.

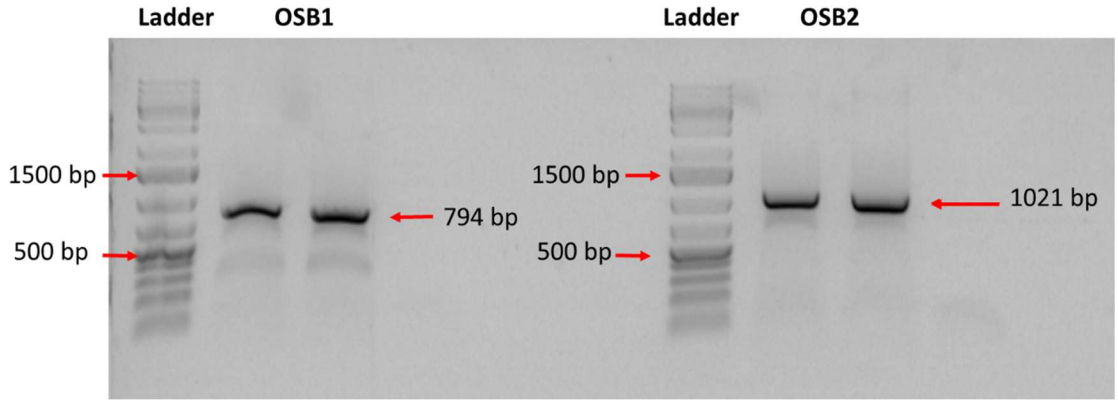

Figure 5.1 PCR product amplification and band size confirmation of *OSB1* (794 bp) and *OSB2* (1021 bp).

5.3.2. *In vitro assay*

To confirm the efficacy of the designed gRNAs from the confirmed target sequence in *OSB1* and *OSB2*, we performed an *in vitro* assay using Mali *et al.*, (2013) *in vitro* digestion protocol. The *in vitro* products were confirmed using 5% agarose gel for better separation. For *OSB1* and *OSB2*, gRNAs 1, 2, 3, 4, and 5 with the expected sized products were detected from the *in vitro* reaction expressing sgRNAs (**Fig 5. 4 and 5.5**). For *OSB2*, except gRNA 3, other gRNAs worked perfectly and were selected to include in the final polycistronic tRNA-gRNA construct (PTG).

5.3.3. Multiplex genome editing in rice protoplasts via *pTRANS100* [pr35S::GFP]

To experiment with the transfection efficiency of the *pTRANS100* [pr35S::GFP] and purple rice protoplast viability, we used an empty vector with GFP for transient expression. We transfected the purple rice protoplasts with the plasmid DNA comprising *pTRANS100* [pr35S::GFP]. The protoplast viability was checked using Fluorescein Diacetate (FDA) stain after 48 hours of transfection (**Fig 5.6. A**). The protoplast transfection was performed using 10 µg of the plasmid DNA (15 µl), and its efficiency

was tested after 48 hours of transfection (**Fig 5.6. B- E**). The protoplast expressed the respective GFP plasmid with high efficiency (**Fig 5.6. B- E**).

After the successful protoplast transfection, we will use the four gRNAs which were arranged to assemble PTGs for targeting two rice genes (*OSB1* and *OSB2*). These gRNAs are distributed into two pairs for each gene. After 48 hours of transfection, the protoplast will be used to check the chromosomal fragment deletions at *OSB1* and *OSB2* loci employing PCR products with target-specific primers.

The gRNAs were validated and future studies can be performed to confirm that these gRNAs will knock out the purple leaf color through transformation and regeneration of rice plants.

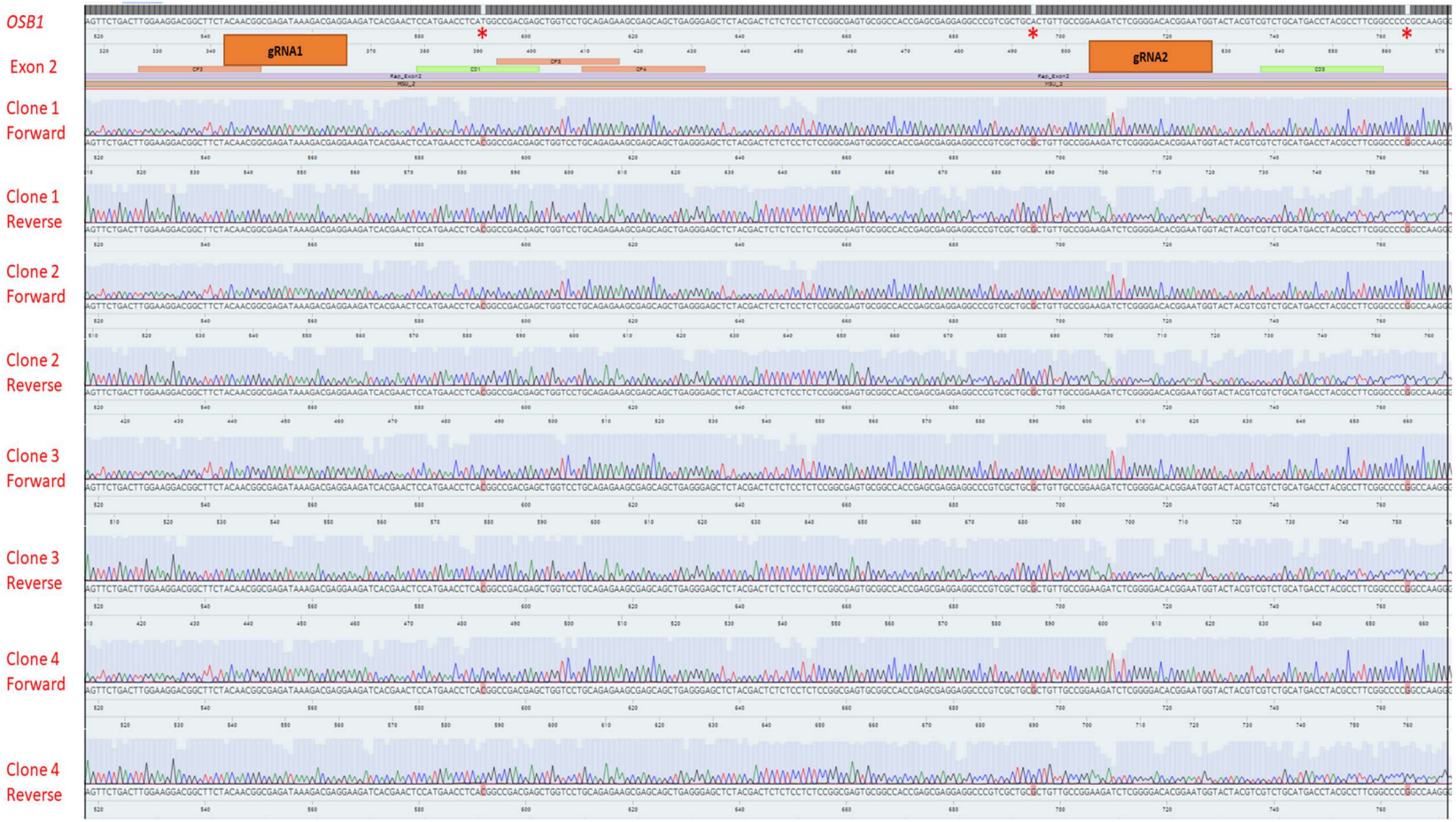

Figure 5.2 Alignment of reference genomic region of *OSB1* with TOPO Blunt-end clones (clones 1 – 4 out of 5) confirms the sequence of exon 2. gRNA1 and gRNA2: designed gRNA for *OSB1*. *: Base pair change in the purple rice exon 2 compared to RAP-DB and MSU database sequences.

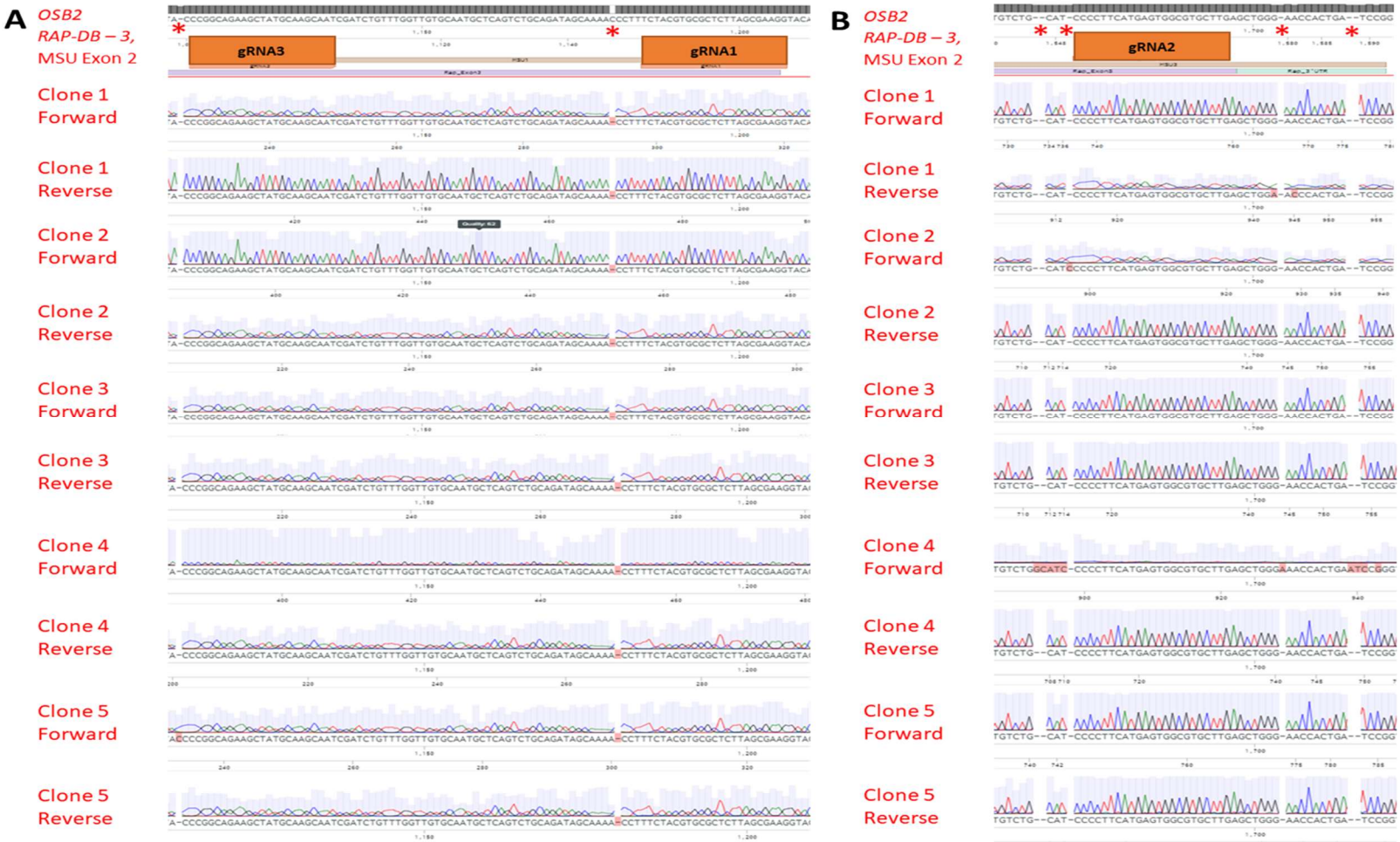

Figure 5.3 Alignment of reference genomic region of *OSB2* with TOPO Blunt-end clones (clones 1 – 5 out of 5) confirms the sequence of gRNA1 and gRNA3 (A) and gRNA2 (B): designed gRNA for *OSB2*. *: Base pair change in the purple rice exon 2 compared to RAP-DB and MSU database sequences.

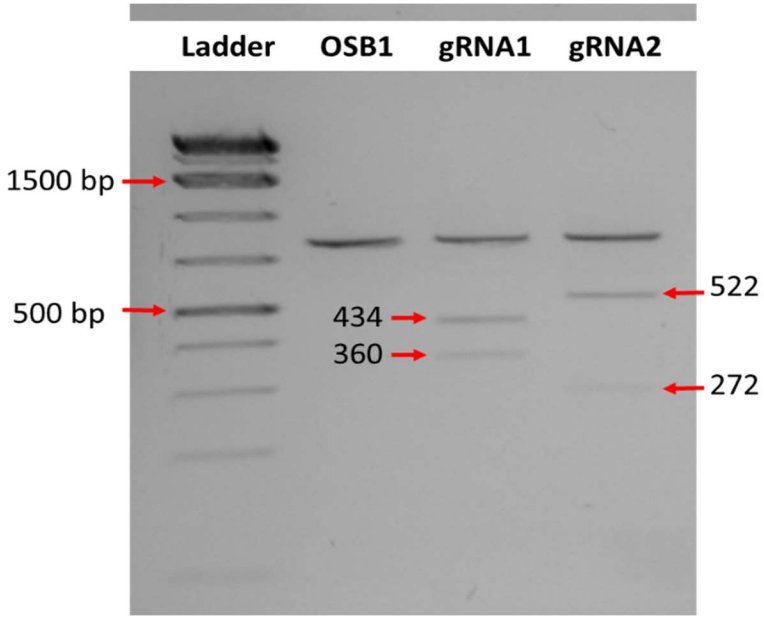

Figure 5.4 *In vitro assay* using gRNA1 and gRNA2 using PCR purified substrate DNA of *OSB1*.

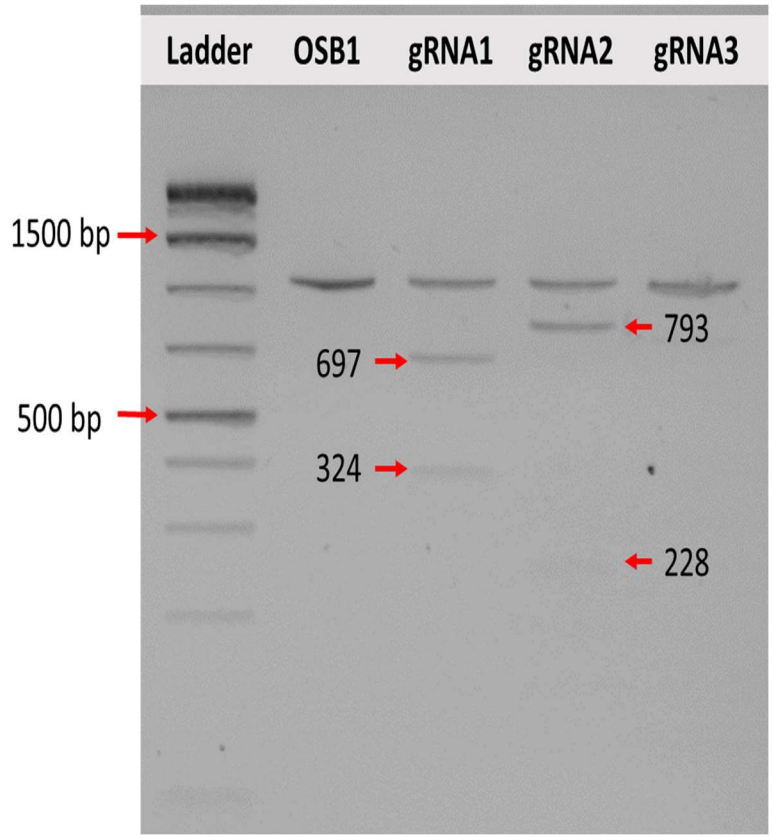

Figure 5.5 *In vitro* assay using gRNA1, gRNA2 and gRNA3 using PCR purified substrate DNA of *OSB2*.

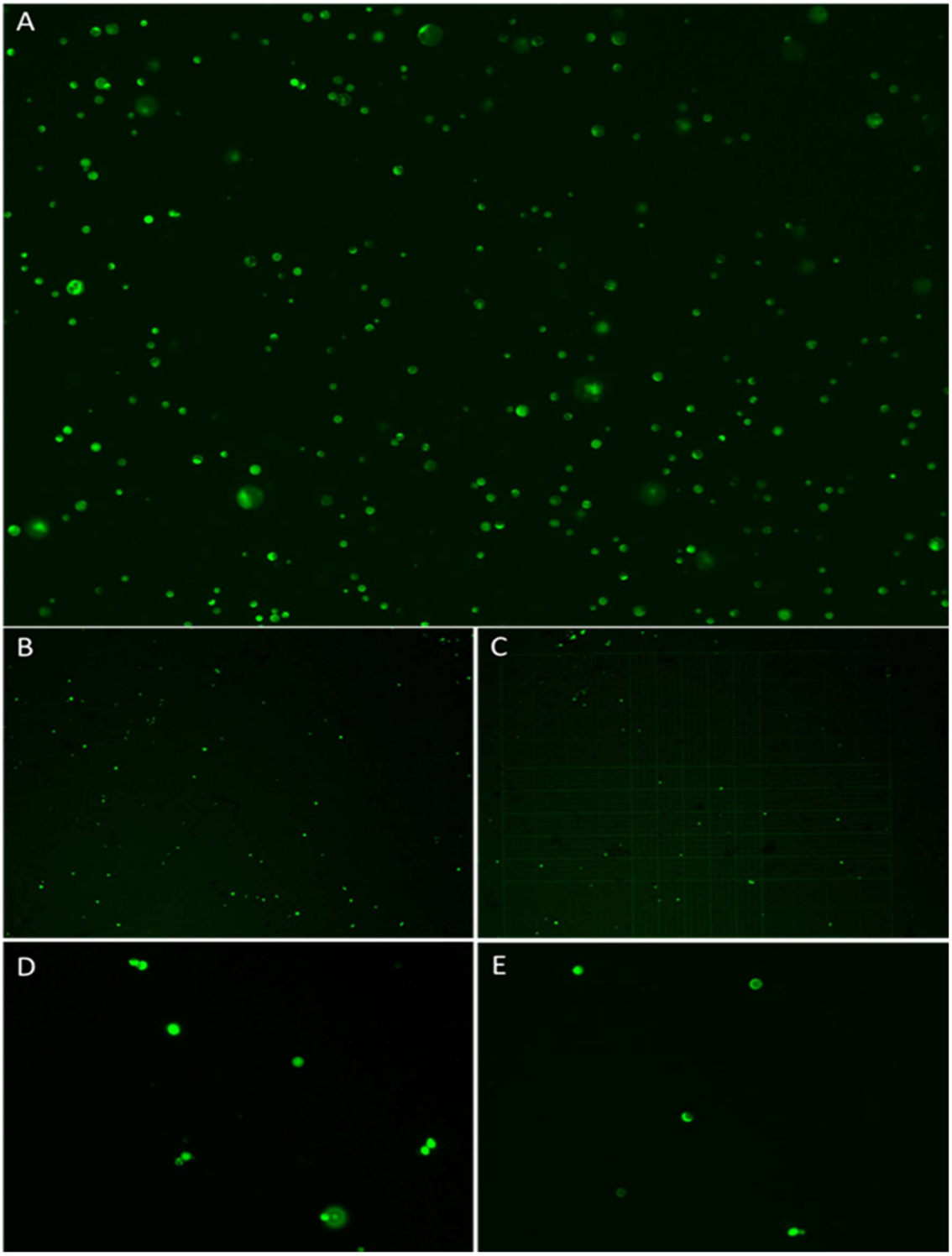

Figure 5.6 Protoplast viability (A) and transfection efficiency of the Green Fluorescent Protein (GFP) plasmid DNA [*pTRANS100* - pr35S::GFP] in the purple rice protoplast (B – E) after 48 hours of transfection.

6. CONCLUSIONS

6.1.Genome-wide association study for DDSR

Chromosome regions related to multiple key traits are attractive targets as they may well represent important breeding targets for crop improvement. These genomic regions may indicate the presence of a pleiotropic effect of a gene affecting multiple biological processes simultaneously (Pelgas *et al.*, 2011), or they could be several genes that affect different traits, which are just closely linked. We identified three QTL hotspots on chromosomes 2 (~24 Mb), 7 (~22Mb), and 12 (~19 Mb). QTL hotspots on chromosome 2, colocalized with *qHI-2, qSF-2-2, qPT-2, qSW-2 and qFe-2* may partly show the contribution of iron uptake during booting in increasing the number of filled grains, seed weight and plant biomass. This genomic region has already been reported to be a hotspot for multiple important traits, which directly or indirectly contribute to plant yield and biomass. Genomic regions detected on chromosome 2 is also colocalized genes/QTLs like *yld2.1* (Marri *et al.*, 2005), *GA2ox9* (Lo *et al.*, 2008), *OsMPS* (Schmidt *et al.*, 2013), tiller number (Miyamoto *et al.*, 2004), resistance to green rice leafhopper (Fujita *et al.*, 2006) and numerous flowering genes (*Oscry2*: Hirose *et al.*, 2006; *OsCOL4*: Lee *et al.*, 2010; *LTG1*: Lu *et al.*, 2014).

Another genomic region QTL hotspot has been detected on chromosome 7 for *qDTF-7, qGY-7,* and *qTN-7*. This genomic region has been reported to be significant for traits like grain size (*oswrky78*: Zhang *et al.*, 2011), flowering time (*OsUDT1*: Gong & He, 2014), leaf chlorophyll content (*qLCC-7*: Zuo *et al.*, 2007) and dwarfism (*OsBZR1*: Bai *et al.*, 2007) contributing positively towards the grain yield and flowering time in

rice. Interestingly, the QTL hotspot on chromosome 12 detected for *qTN-12, qPW-12* and *qGY-12* did not colocalize with any major genes or QTLs for grain yield, tiller number, or panicle weight. Further studying this novel genomic region may facilitate unraveling different molecular mechanisms underlying those key traits.

These genomic regions were identified across the dry and wet seasons, which illustrates the stability of these detected genes/QTLs. Such QTL clusters can be used as targets in marker-assisted breeding programs for DDSR varietal improvement and used for further molecular study, along with some other key QTL targets in different regions of the rice genomes.

6.2.Purple rice

The efficient transfection of purple rice protoplasts with the GFP plasmid DNA paves the path towards successful gene editing of purple rice using the PTG approach. The successful development and validation of the gRNAs for *OSB1* and *OSB2* will assist toward validation of these genes for purple leaf color in rice, which can be used as visual markers for plant transformation in the future.

7. REFERENCES

Azhiri-Sigari, TAHERE Gines, HERMENEGILDO Sebastian, LS, & Wade, L. (2005). Seedling vigor of rice cultivars in response to seeding depth and soil moisture. *Philippine Journal of Crop Science*, *30*(1), 53–58.

Bai, M.-Y., Zhang, L.-Y., Gampala, S. S., Zhu, S.-W., Song, W.-Y., Chong, K., & Wang, Z.-Y. (2007). Functions of OsBZR1 and 14-3-3 proteins in brassinosteroid signaling in rice. *Proceedings of the National Academy of Sciences*, *104*(34), 13839–13844.

Bao, Z., Watanabe, A., Sasaki, K., Okubo, T., Tokida, T., Liu, D., Ikeda, S., Imaizumi-Anraku, H., Asakawa, S., & Sato, T. (2014). A rice gene for microbial symbiosis, Oryza sativa CCaMK, reduces CH4 flux in a paddy field with low nitrogen input. *Applied and Environmental Microbiology*, *80*(6), 1995–2003.

Bates, T. R., & Lynch, J. P. (1996). Stimulation of root hair elongation in Arabidopsis thaliana by low phosphorus availability. *Plant, Cell & Environment*, *19*(5), 529–538.

Begum, H., Spindel, J. E., Lalusin, A., Borromeo, T., Gregorio, G., Hernandez, J., Virk, P., Collard, B., & McCouch, S. R. (2015). Genome-wide association mapping for yield and other agronomic traits in an elite breeding population of tropical rice (Oryza sativa). *PLoS One*, *10*(3), e0119873. https://doi.org/10.1371/journal.pone.0119873

Belhaj, K., Chaparro-Garcia, A., Kamoun, S., & Nekrasov, V. (2013). Plant genome editing made easy: targeted mutagenesis in model and crop plants using the CRISPR/Cas system. *Plant Methods*, *9*(1), 39. https://doi.org/10.1186/1746-4811-9-39

Belhaj, K., Chaparro-Garcia, A., Kamoun, S., Patron, N. J., & Nekrasov, V. (2015). Editing plant genomes with CRISPR/Cas9. *Current Opinion in Biotechnology*, *32*, 76–84. https://doi.org/https://doi.org/10.1016/j.copbio.2014.11.007

Bernier, J., Serraj, R., Kumar, A., Venuprasad, R., Impa, S., RP, V. G., Oane, R., Spaner, D., & Atlin, G. (2009). The large-effect drought-resistance QTL qtl12. 1 increases water uptake in upland rice. *Field Crops Research*, *110*(2), 139–146.

Bouman, B. A. M., & Tuong, T. P. (2001). Field water management to save water and increase its productivity in irrigated lowland rice. *Agricultural Water Management*, *49*(1), 11–30.

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., & Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics*, *23*(19), 2633–2635. https://doi.org/10.1093/bioinformatics/btm308

Bradshaw, A. D. (1965). Evolutionary significance of phenotypic plasticity in plants. In *Advances in genetics* (Vol. 13, pp. 115–155). Elsevier.

Carriger, S., & Vallée, D. (2007). More crop per drop. *Rice Today*, *6*(2), 10–13.

Caton, B. P., Cope, A. E., & Mortimer, M. (2003). Growth traits of diverse rice cultivars under severe competition: implications for screening for competitiveness. *Field Crops Research*, *83*(2), 157–172.

Chen, X., Shi, J., Hao, X., Liu, H., Shi, J., Wu, Y., Wu, Z., Chen, M., Wu, P., & Mao, C. (2013). OsORC 3 is required for lateral root development in rice. *The Plant Journal*, *74*(2), 339–350.

Cui, K.-H., Peng, S.-B., Xing, Y.-Z., Yu, S.-B., & Xu, C.-G. (2002). Genetic analysis of the panicle traits related to yield sink size of rice. *Yi Chuan Xue Bao= Acta Genetica Sinica*, *29*(2), 144–152.

Cui, K., Peng, S., Ying, Y., Yu, S., & Xu, C. (2004). Molecular dissection of the relationships among tiller number, plant height and heading date in rice. *Plant Production Science*, *7*(3), 309–318.

Cui, R., Han, J., Zhao, S., Su, K., Wu, F., Du, X., Xu, Q., Chong, K., Theißen, G., & Meng, Z. (2010). Functional conservation and diversification of class E floral homeotic genes in rice (Oryza sativa). *The Plant Journal*, *61*(5), 767–781.

Deng, J., Chen, S., Yin, X., Wang, K., Liu, Y., Li, S., & Yang, P. (2013). Systematic qualitative and quantitative assessment of anthocyanins, flavones and flavonols in the petals of 108 lotus (Nelumbo nucifera) cultivars. *Food Chemistry*, *139*(1–4), 307–312.

Dingkuhn, M, Schnier, H. F., De Datta, S. K., Dörffling, K., Javellana, C., & Pamplona, R. (1990). Nitrogen Fertilization of Direct-Seeded Flooded vs. Transplanted Rice: II. Interactions among Canopy Properties. *Crop Science*, *30*(6), 1284–1292.

Dingkuhn, Michael, Johnson, D. E., Sow, A., & Audebert, A. Y. (1999). Relationships between upland rice canopy characteristics and weed competitiveness. *Field Crops Research*, *61*(1), 79–95.

Domingo, C., Andrés, F., Tharreau, D., Iglesias, D. J., & Talón, M. (2009). Constitutive expression of OsGH3. 1 reduces auxin content and enhances defense response and resistance to a fungal pathogen in rice. *Molecular Plant-Microbe Interactions*, *22*(2), 201–210.

FAO. (2018). Food and Agriculture Organization of the United Nations (2018) Food outlook: biannual report on global food markets. *Rome, Italy*.

Farooq, M., Basra, S. M. A., Tabassum, R., & Afzal, I. (2015). Enhancing the Performance of Direct Seeded Fine Rice by Seed Priming. *Plant Production Science*, *9*(4), 446–456. https://doi.org/10.1626/pps.9.446

Fischer, A., Ramírez, H. V, & Lozano, J. (1997). Suppression of junglerice [Echinochloa

colona (L.) Link] by irrigated rice cultivars in Latin America. *Agronomy Journal*, *89*(3), 516–521.

Flint-Garcia, S. A., Thornsberry, J. M., & Buckler IV, E. S. (2003). Structure of linkage disequilibrium in plants. *Annual Review of Plant Biology*, *54*(1), 357–374.

Fornara, F., Pařenicová, L., Falasca, G., Pelucchi, N., Masiero, S., Ciannamea, S., Lopez-Dee, Z., Altamura, M. M., Colombo, L., & Kater, M. M. (2004). Functional characterization of OsMADS18, a member of the AP1/SQUA subfamily of MADS box genes. *Plant Physiology*, *135*(4), 2207–2219.

Fujita, D., Doi, K., Yoshimura, A., & Yasui, H. (2006). Molecular mapping of a novel gene, Grh5, conferring resistance to green rice leafhopper (Nephotettix cincticeps Uhler) in rice, Oryza sativa L. *Theoretical and Applied Genetics*, *113*(4), 567–573.

Fukai, S., & Inthapan, P. (1988). Growth and yield of rice cultivars under sprinkler irrigation in south-eastern Queensland. 3. Water extraction and plant water relations dash comparison with maize and grain sorghum. *Australian Journal of Experimental Agriculture*, *28*(2), 249–252.

Gamuyao, R., Chin, J. H., Pariasca-Tanaka, J., Pesaresi, P., Catausan, S., Dalid, C., Slamet-Loedin, I., Tecson-Mendoza, E. M., Wissuwa, M., & Heuer, S. (2012). The protein kinase Pstol1 from traditional rice confers tolerance of phosphorus deficiency. *Nature*, *488*(7412), 535–539.

Garrity, D. P., Movillon, M., & Moody, K. (1992). Differential weed suppression ability in upland rice cultivars. *Agronomy Journal*, *84*(4), 586–591.

Gleick, P. H. (1993). Water and conflict: Fresh water resources and international security. *International Security*, *18*(1), 79–112.

Gong, P., & He, C. (2014). Uncovering divergence of rice exon junction complex core heterodimer gene duplication reveals their essential role in growth, development, and reproduction. *Plant Physiology*, *165*(3), 1047–1061.

Hamada, H., Linghu, Q., Nagira, Y., Miki, R., Taoka, N., & Imai, R. (2017). An in planta biolistic method for stable wheat transformation. *Scientific Reports*, *7*(1), 1–8.

Hanzawa, E., Sasaki, K., Nagai, S., Obara, M., Fukuta, Y., Uga, Y., Miyao, A., Hirochika, H., Higashitani, A., & Maekawa, M. (2013). Isolation of a novel mutant gene for soil-surface rooting in rice (Oryza sativa L.). *Rice*, *6*(1), 1–11.

Hirose, F., Shinomura, T., Tanabata, T., Shimada, H., & Takano, M. (2006). Involvement of rice cryptochromes in de-etiolation responses and flowering. *Plant and Cell Physiology*, *47*(7), 915–925.

Huang, J., Kim, C. M., Xuan, Y., Park, S. J., Piao, H. L., Je, B. Il, Liu, J., Kim, T. H., Kim, B.-K., & Han, C.-D. (2013). OsSNDP1, a Sec14-nodulin domain-containing protein, plays a critical role in root hair elongation in rice. *Plant Molecular Biology*,

82(1), 39–50.

Hudson, D., Guevara, D. R., Hand, A. J., Xu, Z., Hao, L., Chen, X., Zhu, T., Bi, Y.-M., & Rothstein, S. J. (2013). Rice cytokinin GATA transcription Factor1 regulates chloroplast development and plant architecture. *Plant Physiology*, *162*(1), 132–144.

Inoue, H., Kobayashi, T., Nozoye, T., Takahashi, M., Kakei, Y., Suzuki, K., Nakazono, M., Nakanishi, H., Mori, S., & Nishizawa, N. K. (2009). Rice OsYSL15 is an iron-regulated iron (III)-deoxymugineic acid transporter expressed in the roots and is essential for iron uptake in early growth of the seedlings. *Journal of Biological Chemistry*, *284*(6), 3470–3479.

Ishimaru, Y., Masuda, H., Bashir, K., Inoue, H., Tsukamoto, T., Takahashi, M., Nakanishi, H., Aoki, N., Hirose, T., & Ohsugi, R. (2010). Rice metal-nicotianamine transporter, OsYSL2, is required for the long-distance transport of iron and manganese. *The Plant Journal*, *62*(3), 379–390.

Ithal, N., & Reddy, A. R. (2004). Rice flavonoid pathway genes, OsDfr and OsAns, are induced by dehydration, high salt and ABA, and contain stress responsive promoter elements that interact with the transcription activator, OsC1-MYB. *Plant Science*, *166*(6), 1505–1513.

Johnson, D. E., Mortimer, A. M., Singh, Y., Singh, V. P., Chauhan, B., Orr, A., & Hardy, B. (2008). Issues for weed management in direct-seeded rice and the development of decision-support frameworks. *Direct Seeding of Rice and Weed Management in the Irrigated Rice–Wheat Cropping System of the Indo-Gangetic Plains*, 223–228.

Kabir, M. K., Ullah, S. M., Jahan, S., Ullah, M. B., & Kamal, A. (2008). Influence of sewage sludge and nitrogen fertilization on growth, nutrient content and heavy metal uptake by rice straw. *Bangladesh Journal of Scientific and Industrial Research*, *43*(4), 571–580.

Kang, H.-G., Jeon, J.-S., Lee, S., & An, G. (1998). Identification of class B and class C floral organ identity genes from rice plants. *Plant Molecular Biology*, *38*(6), 1021–1029.

Kang, J.-S., Choi, S., & Rhew, H.-Y. (1998). Handbook of Formulas and Software for Plant Geneticists and Breeders. In *Journal of the Korean Continence Society* (Vol. 2, Issue 2). https://doi.org/10.5213/jkcs.1998.2.2.70

Katsura, K., Okami, M., Mizunuma, H., & Kato, Y. (2010). Radiation use efficiency, N accumulation and biomass production of high-yielding rice in aerobic culture. *Field Crops Research*, *117*(1), 81–89.

Kawakatsu, T., Taramino, G., Itoh, J., Allen, J., Sato, Y., Hong, S., Yule, R., Nagasawa, N., Kojima, M., & Kusaba, M. (2009). PLASTOCHRON3/GOLIATH encodes a glutamate carboxypeptidase required for proper development in rice. *The Plant*

Journal, *58*(6), 1028–1040.

Kitomi, Y., Ogawa, A., Kitano, H., & Inukai, Y. (2008). *CRL4* regulates crown root formation through auxin transport in rice. *Plant Root*, *2*, 19–28. https://doi.org/10.3117/plantroot.2.19

Kondo, M., Murty, M. V. R., Aragones, D. V, Okada, K., Winn, T., & Kwak, K. S. (1999). Characteristics of the root system and water uptake in upland rice. *Genetic Improvement of Rice for Water-Limited Environments*, 117–134.

Kumar, V., & Ladha, J. K. (2011). *Direct Seeding of Rice* (pp. 297–413). https://doi.org/10.1016/b978-0-12-387689-8.00001-1

Lal, R., Follett, R. F., Stewart, B. A., & Kimble, J. M. (2007). Soil carbon sequestration to mitigate climate change and advance food security. *Soil Science*, *172*(12), 943–956.

Lee, S., Kim, S. A., Lee, J., Guerinot, M. Lou, & An, G. (2010). Zinc deficiency-inducible OsZIP8 encodes a plasma membrane-localized zinc transporter in rice. *Molecules and Cells*, *29*(6), 551–558.

Lee, Y., Jeong, D., Lee, D., Yi, J., Ryu, C., Kim, S. L., Jeong, H. J., Choi, S. C., Jin, P., & Yang, J. (2010). OsCOL4 is a constitutive flowering repressor upstream of Ehd1 and downstream of OsphyB. *The Plant Journal*, *63*(1), 18–30.

Li, M., Tang, D., Wang, K., Wu, X., Lu, L., Yu, H., Gu, M., Yan, C., & Cheng, Z. (2011). Mutations in the F-box gene LARGER PANICLE improve the panicle architecture and enhance the grain yield in rice. *Plant Biotechnology Journal*, *9*(9), 1002–1013.

Liang, C., Wang, Y., Zhu, Y., Tang, J., Hu, B., Liu, L., Ou, S., Wu, H., Sun, X., & Chu, J. (2014). OsNAP connects abscisic acid and leaf senescence by fine-tuning abscisic acid biosynthesis and directly targeting senescence-associated genes in rice. *Proceedings of the National Academy of Sciences*, *111*(27), 10013–10018.

Liang, Q., Cheng, X., Mei, M., Yan, X., & Liao, H. (2010). QTL analysis of root traits as related to phosphorus efficiency in soybean. *Annals of Botany*, *106*(1), 223–234.

Lipka, A. E., Tian, F., Wang, Q., Peiffer, J., Li, M., Bradbury, P. J., Gore, M. A., Buckler, E. S., & Zhang, Z. (2012). GAPIT: genome association and prediction integrated tool. *Bioinformatics*, *28*(18), 2397–2399. https://doi.org/10.1093/bioinformatics/bts444

Liu, S., Wang, J., Wang, L., Wang, X., Xue, Y., Wu, P., & Shou, H. (2009). Adventitious root formation in rice requires OsGNOM1 and is mediated by the OsPINs family. *Cell Research*, *19*(9), 1110–1119.

Lo, S.-F., Yang, S.-Y., Chen, K.-T., Hsing, Y.-I., Zeevaart, J. A. D., Chen, L.-J., & Yu, S.-M. (2008). A novel class of gibberellin 2-oxidases control semidwarfism, tillering, and root development in rice. *The Plant Cell*, *20*(10), 2603–2618.

Lu, G., Wu, F., Wu, W., Wang, H., Zheng, X., Zhang, Y., Chen, X., Zhou, K., Jin, M., & Cheng, Z. (2014). Rice LTG 1 is involved in adaptive growth and fitness under low ambient temperature. *The Plant Journal*, *78*(3), 468–480.

Lu, Y., Song, Z., Lü, K., Lian, X., & Cai, H. (2012). Molecular characterization, expression and functional analysis of the amino acid transporter gene family (OsAATs) in rice. *Acta Physiologiae Plantarum*, *34*(5), 1943–1962.

Lynch, J., & van Beem, J. J. (1993). Growth and architecture of seedling roots of common bean genotypes. *Crop Science*, *33*(6), 1253–1257.

MacMillan, K., Emrich, K., Piepho, H.-P., Mullins, C. E., & Price, A. H. (2006). Assessing the importance of genotype× environment interaction for root traits in rice using a mapping population II: conventional QTL analysis. *Theoretical and Applied Genetics*, *113*(5), 953–964.

Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., Norville, J. E., & Church, G. M. (2013). RNA-guided human genome engineering via Cas9. *Science*, *339*(6121), 823–826.

Marri, P. R., Sarla, N., Reddy, L. V, & Siddiq, E. A. (2005). Identification and mapping of yield and yield related QTLs from an Indian accession of Oryza rufipogon. *BMC Genetics*, *6*(1), 1–14.

Miyamoto, N., Goto, Y., Matsui, M., Ukai, Y., Morita, M., & Nemoto, K. (2004). Quantitative trait loci for phyllochron and tillering in rice. *Theoretical and Applied Genetics*, *109*(4), 700–706.

Miyoshi, K., Ito, Y., Serizawa, A., & Kurata, N. (2003). OsHAP3 genes regulate chloroplast biogenesis in rice. *The Plant Journal*, *36*(4), 532–540.

Morita, Y., & Kyozuka, J. (2007). Characterization of OsPID, the rice ortholog of PINOID, and its possible involvement in the control of polar auxin transport. *Plant and Cell Physiology*, *48*(3), 540–549.

Nekrasov, V., Wang, C., Win, J., Lanz, C., Weigel, D., & Kamoun, S. (2017). Rapid generation of a transgene-free powdery mildew resistant tomato by genome deletion. *Scientific Reports*, *7*(1), 1–6.

Nguyen, H. T., Babu, R. C., & Blum, A. (1997). Breeding for drought resistance in rice: physiology and molecular genetics considerations. *Crop Science*, *37*(5), 1426–1434.

Oikawa, T., Maeda, H., Oguchi, T., Yamaguchi, T., Tanabe, N., Ebana, K., Yano, M., Ebitani, T., & Izawa, T. (2015). The birth of a black rice gene and its local spread by introgression. *The Plant Cell*, *27*(9), 2401–2414.

Okami, M., Kato, Y., & Yamagishi, J. (2011). Role of early vigor in adaptation of rice to water-saving aerobic culture: effects of nitrogen utilization and leaf growth. *Field Crops Research*, *124*(1), 124–131.

Pathak, H., Tewari, A. N., Sankhyan, S., Dubey, D. S., Mina, U., Singh, V. K., Jain, N., & Bhatia, A. (2011). Direct-seeded rice: potential, performance and problems–A review. *Current Advances in Agricultural Sciences*, *3*(2), 77–88.

Pelgas, B., Bousquet, J., Meirmans, P. G., Ritland, K., & Isabel, N. (2011). QTL mapping in white spruce: gene maps and genomic regions underlying adaptive traits across pedigrees, years and environments. *BMC Genomics*, *12*(1), 1–23.

Piao, R., Jiang, W., Ham, T.-H., Choi, M.-S., Qiao, Y., Chu, S.-H., Park, J.-H., Woo, M.-O., Jin, Z., & An, G. (2009). Map-based cloning of the ERECT PANICLE 3 gene in rice. *Theoretical and Applied Genetics*, *119*(8), 1497.

Pigliucci, M. (2001). *Phenotypic plasticity: beyond nature and nurture*. JHU Press.

Pimentel, D., Berger, B., Filiberto, D., Newton, M., Wolfe, B., Karabinakis, E., Clark, S., Poon, E., Abbett, E., & Nandagopal, S. (2004). Water resources: agricultural and environmental issues. *Bioscience*, *54*(10), 909–918.

Poorter, H., & De Jong, R. O. B. (1999). A comparison of specific leaf area, chemical composition and leaf construction costs of field plants from 15 habitats differing in productivity. *New Phytologist*, *143*(1), 163–176.

Qiu, S., Ma, N., Che, S., Wang, Y., Peng, X., Zhang, G., Wang, G., & Huang, J. (2014). Repression of OsEXPA3 expression leads to root system growth suppression in rice. *Crop Science*, *54*(5), 2201–2213.

Rose, M. T., Rose, T. J., Pariasca-Tanaka, J., Yoshihashi, T., Neuweger, H., Goesmann, A., Frei, M., & Wissuwa, M. (2012). Root metabolic response of rice (Oryza sativa L.) genotypes with contrasting tolerance to zinc deficiency and bicarbonate excess. *Planta*, *236*(4), 959–973.

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.

Sakamoto, W., Ohmori, T., Kageyama, K., Miyazaki, C., Saito, A., Murata, M., Noda, K., & Maekawa, M. (2001). The Purple leaf (Pl) locus of rice: the Pl w allele has a complex organization and includes two genes encoding basic helix-loop-helix proteins involved in anthocyanin biosynthesis. *Plant and Cell Physiology*, *42*(9), 982–991.

SANCHEZ, P. A. (1973). Puddling tropical rice soils: 2. Effects of water losses. *Soil Science*, *115*(4), 303–308.

Sandhu, N., Singh, A., Dixit, S., Cruz, M. T. S., Maturan, P. C., Jain, R. K., & Kumar, A. (2014). Identification and mapping of stable QTL with main and epistasis effect on rice grain yield under upland drought stress. *BMC Genetics*, *15*(1), 1–15.

Sandhu, N., Subedi, S. R., Singh, V. K., Sinha, P., Kumar, S., Singh, S. P., Ghimire, S. K., Pandey, M., Yadaw, R. B., & Varshney, R. K. (2019). Deciphering the genetic

basis of root morphology, nutrient uptake, yield, and yield-related traits in rice under dry direct-seeded cultivation systems. *Scientific Reports*, *9*(1), 1–16.

Sandhu, N., Torres, R. O., Sta Cruz, M. T., Maturan, P. C., Jain, R., Kumar, A., & Henry, A. (2015). Traits and QTLs for development of dry direct-seeded rainfed rice varieties. *Journal of Experimental Botany*, *66*(1), 225–244.

Sasaki, A., Ashikari, M., Ueguchi-Tanaka, M., Itoh, H., Nishimura, A., Swapan, D., Ishiyama, K., Saito, T., Kobayashi, M., & Khush, G. S. (2002). A mutant gibberellin-synthesis gene in rice. *Nature*, *416*(6882), 701–702.

Schmidt, R., Schippers, J. H. M., Mieulet, D., Obata, T., Fernie, A. R., Guiderdoni, E., & Mueller-Roeber, B. (2013). MULTIPASS, a rice R2R3-type MYB transcription factor, regulates adaptive growth by integrating multiple hormonal pathways. *The Plant Journal*, *76*(2), 258–273.

Septiningsih, E. M., Prasetiyono, J., Lubis, E., Tai, T. H., Tjubaryat, T., Moeljopawiro, S., & McCouch, S. R. (2003). Identification of quantitative trait loci for yield and yield components in an advanced backcross population derived from the Oryza sativa variety IR64 and the wild relative O. rufipogon. *Theoretical and Applied Genetics*, *107*(8), 1419–1432.

Shan, Q., Wang, Y., Li, J., & Gao, C. (2014). Genome editing in rice and wheat using the CRISPR/Cas system. *Nature Protocols*, *9*(10), 2395–2410.

Sharma, P K, Bhushan, L., Ladha, J. K., Naresh, R. K., Gupta, R. K., Balasubramanian, B. V, & Bouman, B. A. M. (2002). Crop-water relations in rice-wheat cropping under different tillage systems and water-management practices in a marginally sodic, medium-textured soil. *Water-Wise Rice Production. International Rice Research Institute, Los Baños, Philippines*, *8*, 223–235.

Sharma, Pradeep K, Ladha, J. K., & Bhushan, L. (2003). Soil physical effects of puddling in rice–wheat cropping systems. *Improving the Productivity and Sustainability of Rice-Wheat Systems: Issues and Impacts*, *65*, 97–113.

Sharma, S., Xu, S., Ehdaie, B., Hoops, A., Close, T. J., Lukaszewski, A. J., & Waines, J. G. (2011). Dissection of QTL effects for root traits using a chromosome arm-specific mapping population in bread wheat. *Theoretical and Applied Genetics*, *122*(4), 759–769.

Sheng, T., Qian, Q., Da Li, Z., Yasufumi, K., Da Nian, H., & Huang, Z. H. U. L. (2002). QTL analysis of rice peduncle vascular bundle system and panicle traits. *Journal of Integrative Plant Biology*, *44*(3), 301–306.

Shipley, B. (2006). Net assimilation rate, specific leaf area and leaf mass ratio: which is most closely correlated with relative growth rate? A meta-analysis. *Functional Ecology*, *20*(4), 565–574.

Singh, A., Choudhury, B., & Bouman, B. (2002). Effects of rice establishment methods

on crop performance, water use, and mineral nitrogen. *Water-Wise Rice Production. Los Baños (Philippines): International Rice Research Institute. P*, 237–246.

Singh, U. M., Yadav, S., Dixit, S., Ramayya, P. J., Devi, M. N., Raman, K. A., & Kumar, A. (2017). QTL hotspots for early vigor and related traits under dry direct-seeded system in rice (Oryza sativa L.). *Frontiers in Plant Science*, *8*, 286. https://doi.org/10.3389/fpls.2017.00286

Singh, U., Yadav, S., Dixit, S., Ramayya, P. J., Devi, M. N., Raman, K. A., & Kumar, A. (2017a). QTL hotspots for early vigor and related traits under dry direct-seeded system in rice (Oryza sativa L.). *Frontiers in Plant Science*, *8*, 286. https://doi.org/10.3389/fpls.2017.00286

Singh, U., Yadav, S., Dixit, S., Ramayya, P. J., Devi, M. N., Raman, K. A., & Kumar, A. (2017b). QTL hotspots for early vigor and related traits under dry direct-seeded system in rice (Oryza sativa L.). *Frontiers in Plant Science*, *8*, 286.

Song, G., Jia, M., Chen, K., Kong, X., Khattak, B., Xie, C., Li, A., & Mao, L. (2016). CRISPR/Cas9: a powerful tool for crop genome editing. *The Crop Journal*, *4*(2), 75–82.

Su, J.-Y., Zheng, Q., Li, H.-W., Li, B., Jing, R.-L., Tong, Y.-P., & Li, Z.-S. (2009). Detection of QTLs for phosphorus use efficiency in relation to agronomic performance of wheat grown under phosphorus sufficient and limited conditions. *Plant Science*, *176*(6), 824–836.

Subedi, S. R., Sandhu, N., Singh, V. K., Sinha, P., Kumar, S., Singh, S. P., Ghimire, S. K., Pandey, M., Yadaw, R. B., & Varshney, R. K. (2019). Genome-wide association study reveals significant genomic regions for improving yield, adaptability of rice under dry direct seeded cultivation condition. *BMC Genomics*, *20*(1), 1–20.

Sultan, S. E. (2000). Phenotypic plasticity for plant development, function and life history. *Trends in Plant Science*, *5*(12), 537–542.

Suzaki, T., Toriba, T., Fujimoto, M., Tsutsumi, N., Kitano, H., & Hirano, H.-Y. (2006). Conservation and diversification of meristem maintenance mechanism in Oryza sativa: function of the FLORAL ORGAN NUMBER2 gene. *Plant and Cell Physiology*, *47*(12), 1591–1602.

Swaminathan, M. S., & Jana, S. (1992). *Biodiversity implications for global food security*.

Thomson, M. J., Tai, T. H., McClung, A. M., Lai, X. H., Hinga, M. E., Lobos, K. B., Xu, Y., Martinez, C. P., & McCouch, S. R. (2003). Mapping quantitative trait loci for yield, yield components and morphological traits in an advanced backcross population between Oryza rufipogon and the Oryza sativa cultivar Jefferson.

Theoretical and Applied Genetics, *107*(3), 479–493.

Tripathi, R. P., Sharma, P., & Singh, S. (2007). Influence of tillage and crop residue on soil physical properties and yields of rice and wheat under shallow water table conditions. *Soil and Tillage Research*, *92*(1–2), 221–226.

Tseng, I.-C., Hong, C.-Y., Yu, S.-M., & Ho, T.-H. D. (2013). Abscisic acid-and stress-induced highly proline-rich glycoproteins regulate root growth in rice. *Plant Physiology*, *163*(1), 118–134.

Tuong, T. P., & Bouman, B. A. M. (2003). Rice production in water-scarce environments. *Water Productivity in Agriculture: Limits and Opportunities for Improvement*, *1*, 13–42.

Wang, H., Inukai, Y., & Yamauchi, A. (2006). Root development and nutrient uptake. *Critical Reviews in Plant Sciences*, *25*(3), 279–301.

Wang, W., Mauleon, R., Hu, Z., Chebotarov, D., Tai, S., Wu, Z., Li, M., Zheng, T., Fuentes, R. R., & Zhang, F. (2018). Genomic variation in 3,010 diverse accessions of Asian cultivated rice. *Nature*, *557*(7703), 43–49.

Xiao, J., Li, J., Yuan, L., & Tanksley, S. D. (1996). Identification of QTLs affecting traits of agronomic importance in a recombinant inbred population derived from a subspecific rice cross. *Theoretical and Applied Genetics*, *92*(2), 230–244.

Yadav, S., Singh, U. M., Naik, S. M., Venkateshwarlu, C., Ramayya, P. J., Raman, K. A., Sandhu, N., & Kumar, A. (2017a). Molecular mapping of QTLs associated with lodging resistance in dry direct-seeded rice (Oryza sativa L.). *Frontiers in Plant Science*, *8*, 1431.

Yadav, S., Singh, U. M., Naik, S. M., Venkateshwarlu, C., Ramayya, P. J., Raman, K. A., Sandhu, N., & Kumar, A. (2017b). Molecular Mapping of QTLs Associated with Lodging Resistance in Dry Direct-Seeded Rice (Oryza sativa L.). *Frontiers in Plant Science*, *8*(August), 1–10. https://doi.org/10.3389/fpls.2017.01431

Yamamoto, Y., Kamiya, N., Morinaka, Y., Matsuoka, M., & Sazuka, T. (2007). Auxin biosynthesis by the YUCCA genes in rice. *Plant Physiology*, *143*(3), 1362–1371.

Yan, C.-J., Liang, G.-H., Chen, F., Li, X., Tang, S.-Z., Yi, C.-D., Tian, S., Lu, J.-F., & Gu, M.-H. (2003). Mapping quantitative trait loci associated with rice grain shape based on an indica/japonica backcross population. *Yi Chuan Xue Bao= Acta Genetica Sinica*, *30*(8), 711–716.

Yonemaru, J., Yamamoto, T., Fukuoka, S., Uga, Y., Hori, K., & Yano, M. (2010). Q-TARO: QTL annotation rice online database. *Rice*, *3*(2–3), 194–203.

Yoo, S.-C., Cho, S.-H., Zhang, H., Paik, H.-C., Lee, C.-H., Li, J., Yoo, J.-H., Koh, H.-J., Seo, H. S., & Paek, N.-C. (2007). Quantitative trait loci associated with functional stay-green SNU-SG1 in rice. *Molecules & Cells (Springer Science & Business Media BV)*, *24*(1).

Yoshida, S., & Hasegawa, S. (1982). The rice root system: its development and function. *Drought Resistance in Crops with Emphasis on Rice*, *10*, 97–134.

Yu, J., & Buckler, E. S. (2006). Genetic association mapping and genome organization of maize. *Current Opinion in Biotechnology*, *17*(2), 155–160.

Yuan, L., & Grotewold, E. (2015). Metabolic engineering to enhance the value of plants as green factories. *Metabolic Engineering*, *27*, 83–91.

Zhang, C.-Q., Xu, Y., Lu, Y., Yu, H.-X., Gu, M.-H., & Liu, Q.-Q. (2011). The WRKY transcription factor OsWRKY78 regulates stem elongation and seed development in rice. *Planta*, *234*(3), 541–554.

Zhang, Q., Li, J., Zhang, W., Yan, S., Wang, R., Zhao, J., Li, Y., Qi, Z., Sun, Z., & Zhu, Z. (2012). The putative auxin efflux carrier OsPIN3t is involved in the drought stress response and drought tolerance. *The Plant Journal*, *72*(5), 805–816.

Zhang, Yingying, Zhu, Y., Peng, Y., Yan, D., Li, Q., Wang, J., Wang, L., & He, Z. (2008). Gibberellin homeostasis and plant height control by EUI and a role for gibberellin in root gravity responses in rice. *Cell Research*, *18*(3), 412–421.

Zhang, Yuan-Ji, Lynch, J. P., & Brown, K. M. (2003). Ethylene and phosphorus availability have interacting yet distinct effects on root hair development. *Journal of Experimental Botany*, *54*(391), 2351–2361.

Zhang, Z.-H., Qu, X.-S., Wan, S., Chen, L.-H., & Zhu, Y.-G. (2005). Comparison of QTL controlling seedling vigour under different temperature conditions using recombinant inbred lines in rice (Oryza sativa). *Annals of Botany*, *95*(3), 423–429.

Zhang, Z., Ersoz, E., Lai, C. Q., Todhunter, R. J., Tiwari, H. K., Gore, M. A., Bradbury, P. J., Yu, J., Arnett, D. K., Ordovas, J. M., & Buckler, E. S. (2010). Mixed linear model approach adapted for genome-wide association studies. *Nat Genet*, *42*(4), 355–360. https://doi.org/10.1038/ng.546

Zhao, D. L., Atlin, G. N., Bastiaans, L., & Spiertz, J. H. J. (2006). Comparing rice germplasm groups for growth, grain yield and weed-suppressive ability under aerobic soil conditions. *Weed Research*, *46*(6), 444–452.

Zhao, K., Tung, C. W., Eizenga, G. C., Wright, M. H., Ali, M. L., Price, A. H., Norton, G. J., Islam, M. R., Reynolds, A., Mezey, J., McClung, A. M., Bustamante, C. D., & McCouch, S. R. (2011). Genome-wide association mapping reveals a rich genetic architecture of complex traits in Oryza sativa. *Nature Communications*, *2*(1), 1–10. https://doi.org/10.1038/ncomms1467

ZhiMing, Y., Bo, K., XiaoWei, H., ShaoLei, L., YouHuang, B., WoNa, D., Ming, C., Hyung-Taeg, C., & Ping, W. (2011). Root hair-specific expansins modulate root hair elongation in rice. *The Plant Journal*, *66*(5), 725–734.

Zhu, J., Kaeppler, S. M., & Lynch, J. P. (2005). Mapping of QTL controlling root hair length in maize (Zea mays L.) under phosphorus deficiency. *Plant and Soil*, *270*(1),

299–310.
Zhuang, J.-Y., Fan, Y.-Y., Wu, J.-L., Xia, Y.-W., & Zheng, K.-L. (2001). Comparison of the detection of QTL for yield traits in different generations of a rice cross using two mapping approaches. *Yi Chuan Xue Bao= Acta Genetica Sinica*, *28*(5), 458–464.
Zou, Y., Liu, X., Wang, Q., Chen, Y., Liu, C., Qiu, Y., & Zhang, W. (2014). OsRPK1, a novel leucine-rich repeat receptor-like kinase, negatively regulates polar auxin transport and root development in rice. *Biochimica et Biophysica Acta (BBA)-General Subjects*, *1840*(6), 1676–1685.
Zuo, H., Xiao, K., Dong, Y., Xu, J., Li, Z., Luo, L., & Mei, H. (2007). Molecular detection of quantitative trait loci for leaf chlorophyll content at different growth-stages of rice (Oryza sativa L.). *Asian Journal of Plant Sciences*.

APPENDIX A
LIST OF ACCESSIONS USED IN THIS STUDY

IRGC	DESIGNATION	ORIGIN	Variety	Subpop
125751	GENIT::IRGC 3272-1	Argentina	*indica*	*admix*
126170	105::IRGC 40896-1	Sri Lanka	*indica*	*admix*
127065	PDR 34-2-1-2::IRGC 117020-1	Pakistan	*indica*	*admix*
117498	JC 157::IRGC 9074-1	India	*AusB, arom*	*aro*
125858	NS 1576::IRGC 68951-1	Madagascar	*indica*	*aro*
120876	AUS 439::IRGC 29221-1	Bangladesh	*AusB*	*aus*
120887	BATHURI::IRGC 25838-1	Bangladesh	*AusB*	*aus*
120894	BHADOIA 303::IRGC 6588-1	Bangladesh	*AusB*	*aus*
120915	CHUNGUR BALI::IRGC 25855-1	Bangladesh	*AusB*	*aus*
120927	DANGAR::IRGC 76296-1	India	*AusB*	*aus*
120968	HODARAWALA::IRGC 67631-1	Sri Lanka	*ind, AusB*	*aus*
120969	HOLOI BASH (SOLOI BASH)::IRGC 64778-1	Bangladesh	*AusB*	*aus*
120996	JAMBALI::IRGC 73101-1	Pakistan	*AusB*	*aus*
121005	KALIA::IRGC 34699-1	Bangladesh	*AusB*	*aus*
121022	KHARSU 80::IRGC 28016-1	Pakistan	*AusB*	*aus*
121026	KOYRA::IRGC 77267-1	Bangladesh	*AusB*	*aus*
121029	KURULU WEE (WHITE)::IRGC 66518-1	Sri Lanka	*AusB*	*aus*
121036	LALSAITA::IRGC 43915-1	Bangladesh	*AusB*	*aus*
121038	LENJA MURALI::IRGC 66815-1	Bangladesh	*AusB*	*aus*
121069	NOROI::IRGC 31611-1	Bangladesh	*ind, AusB*	*aus*
121120	SIMUL KHURI::IRGC 35154-1	India	*AusB*	*aus*
121128	SUFAID 246::IRGC 28303-1	Pakistan	*AusB*	*aus*
121134	TAK SUFAID::IRGC 73127-1	Pakistan	*AusB*	*aus*
121185	ARC 10100::IRGC 20709-1	India	*AusB*	*aus*
121233	JAGLI BORO::IRGC 27516-2	Bangladesh	*AusB*	*aus*
121473	PODI HEENATI::IRGC 36345-1	Sri Lanka	*AusB*	*aus*
121582	TAK::IRGC 73124-1	Pakistan	*AusB*	*aus*
121605	CHANDARHAT::IRGC 25845-1	Bangladesh	*AusB*	*aus*
121618	JABOR SAIL::IRGC 66831-1	Bangladesh	*AusB*	*aus*
125813	KURULUTUDU::IRGC 36304-1	Sri Lanka	*AusB*	*aus*
126249	N 22::IRGC 46459-1	India	*indica*	*aus*
127178	AUS 171::IRGC 29004-1	Bangladesh	*AusB*	*aus*
127179	AUS 219::IRGC 29031-1	Bangladesh	*AusB*	*aus*
127180	AUS 233::IRGC 29036-1	Bangladesh	*AusB*	*aus*
127182	AUS 278::IRGC 29068-1	Bangladesh	*AusB*	*aus*
127184	AUS 295::IRGC 29083-1	Bangladesh	*AusB*	*aus*
127186	AUS 301::IRGC 29089-1	Bangladesh	*AusB*	*aus*

127187	AUS 308::IRGC 29096-1	Bangladesh	*AusB*	*aus*
127188	AUS 329::IRGC 29116-1	Bangladesh	*AusB*	*aus*
127189	AUS 344::IRGC 29131-1	Bangladesh	*AusB*	*aus*
127192	AUS PADDY (RED)::IRGC 44978-1	India	*AusB*	*aus*
127205	BAK TULSI::IRGC 34831-1	India	*AusB*	*aus*
125618	LAI YIP ZIM::IRGC 4955-1	Taiwan	*indica*	*ind1A*
125627	UPRH 233::IRGC 61667-1	India	*indica*	*ind1A*
125669	BAI HE::IRGC 76438-1	China	*indica*	*ind1A*
125675	BA SHI ZAO::IRGC 67903-1	China	*indica*	*ind1A*
125702	CHIH SHEN LI::IRGC 1306-1	China	*indica*	*ind1A*
125704	CHI SHENG TAO::IRGC 4606-1	China	*indica*	*ind1A*
125716	CRILLO LA FRIA::IRGC 10793-1	Venezuela	*indica*	*ind1A*
125723	DA NUO (ZHAN)::IRGC 72025-1	China	*indica*	*ind1A*
125744	FU ZAO XIAN::IRGC 63619-1	China	*indica*	*ind1A*
125748	GAO JIAO BAI::IRGC 68047-1	China	*indica*	*ind1A*
125760	HE GU TSAO::IRGC 51302-1	China	*indica*	*ind1A*
125766	HUA LI ZAO::IRGC 80950-1	China	*indica*	*ind1A*
125770	I KUNG PAO::IRGC 114-1	Taiwan	*AusB*	*ind1A*
125809	KORASISI::IRGC 5285-1	Philippines	*indica*	*ind1A*
125841	MIN KE ZHAN::IRGC 72230-1	China	*indica*	*ind1A*
125842	MIN ZAO 6::IRGC 63772-1	China	*indica*	*ind1A*
125852	NCS 194::IRGC 51932-1	India	*indica*	*ind1A*
125865	PAI CHUEH CHIU LIU::IRGC 34259-1	China	*indica*	*ind1A*
125897	SAN SHIH TSI::IRGC 1038-1	China	*indica*	*ind1A*
125906	SSANGDUJO::IRGC 55632-1	South Korea	*indica*	*ind1A*
125913	TAIPEI WOO CO::IRGC 112-1	Taiwan	*indica*	*ind1A*
125914	TAITUNG WOO LI::IRGC 111-1	Taiwan	*indica*	*ind1A*
125928	TSAI YUAN CHON::IRGC 126-1	Taiwan	*indica*	*ind1A*
125929	TSAO SHENG LI 1::IRGC 1309-1	China	*indica*	*ind1A*
125937	WI BIR SHUN::IRGC 4602-1	China	*indica*	*ind1A*
125940	XIA ZHI BAI::IRGC 53437-1	China	*indica*	*ind1A*
125944	YA NONG ZAO 4::IRGC 63908-1	China	*indica*	*ind1A*
125946	YONG JIN ZAO 3::IRGC 70441-1	China	*indica*	*ind1A*
126115	CHANG LE SAN SHU ZAO::IRGC 63561-1	China	*indica*	*ind1A*
126122	PAI YI PING::IRGC 1368-1	China	*indica*	*ind1A*
126123	SAN CHIAO TSWEN::IRGC 1565-1	China	*AusB*	*ind1A*
127249	CE IN TSAN::IRGC 4362-1	China	*indica*	*ind1A*
127273	CHIAYI WU-K'O::IRGC 64974-1	Taiwan	*indica*	*ind1A*
120991	IRGA 318-11-6-9-2B::IRGC 117339-1	Colombia	*indica*	*ind1B*
125659	AUS 177::IRGC 29009-1	Bangladesh	*indica*	*ind1B*

125691	BR IRGA 409::IRGC 55915-1	Brazil	*indica*	*ind1B*
125692	C 662083::IRGC 62101-1	Taiwan	*indica*	*ind1B*
125699	CHANDINA::IRGC 36420-1	Sri Lanka	*indica*	*ind1B*
125713	CICA 9::IRGC 53079-1	Colombia	*indica*	*ind1B*
125730	E 2024::IRGC 67958-1	China	*indica*	*ind1B*
125778	IRI 339::IRGC 46956-1	South Korea	*indica*	*ind1B*
125800	KHAO GRADOOK CHAHNG::IRGC 17111-1	Thailand	*indica*	*ind1B*
125839	MILYANG 30::IRGC 46977-1	South Korea	*indica*	*ind1B*
125840	MILYANG 77::IRGC 69340-1	South Korea	*indica*	*ind1B*
125844	MUKKALA BAZAL::IRGC 77279-1	Bangladesh	*indica*	*ind1B*
125873	PSBRC 50::IRGC 99706-1	Philippines	*indica*	*ind1B*
125876	PSBRC 88::IRGC 99717-1	Philippines	*indica*	*ind1B*
125894	SADA RUPA::IRGC 77299-1	Bangladesh	*indica*	*ind1B*
125907	SUWEON 311::IRGC 61890-1	South Korea	*indica*	*ind1B*
125923	TIKAL 3::IRGC 50649-1	Guatemala	*indica*	*ind1B*
125950	ARAURE 1::IRGC 116956-1	Venezuela	*indica*	*ind1B*
125951	B 6136-3-TB-0-1-5::IRGC 117312-1	Indonesia	*indica*	*ind1B*
125952	B 6136 E-3-TB-0-1-5::IRGC 117311-1	Indonesia	*indica*	*ind1B*
125955	BW 295-5::IRGC 63098-1	Sri Lanka	*indica*	*ind1B*
125958	CHAMA (DWARF)::IRGC 69487-1	Zambia	*indica*	*ind1B*
125963	ELONI::IRGC 116980-1	Suriname	*indica*	*ind1B*
125967	ICTA CRISPO 38::IRGC 116994-1	Guatemala	*indica*	*ind1B*
125988	IRGA 370-38-1-1F-C4-2::IRGC 117342-1	Colombia	*indica*	*ind1B*
125989	IRGA 370-42-1-1F-C-1::IRGC 117343-1	Colombia	*indica*	*ind1B*
126013	WP 65::IRGC 36526-1	Thailand	*indica*	*ind1B*
126062	IR 75870-5-8-5-B-1::IRGC 117297-1	Philippines	*indica*	*ind1B*
126088	SIGARDIS::IRGC 15555-1	Sri Lanka	*indica*	*ind1B*
126092	WP 36::IRGC 55278-1	Thailand	*indica*	*ind1B*
126196	DAA MANSA::IRGC 67559-1	Ghana	*indica*	*ind1B*
127030	3210::IRGC 116950-1	Sri Lanka	*indica*	*ind1B*
127068	UQUIHUA::IRGC 117037-1	Peru	*indica*	*ind1B*
127075	IR 77390-1-6-4-19-1-B::IRGC 117303-1	Philippines	*indica*	*ind1B*
121441	MUTTU SAMBA::IRGC 36333-1	Sri Lanka	*indica*	*ind2*
122181	NONA BOKRA::IRGC 22710-C1	India	*indica*	*ind2*
124431	DA 11::IRGC 6046-1	Bangladesh	*indica*	*ind2*
125607	DUDH KADAR::IRGC 67707-1	India	*indica*	*ind2*
125615	KEERIPALA CHILL PADDY::IRGC 49790-1	India	*indica*	*ind2*
125616	KOTTEYARAN::IRGC 47383-1	Sri Lanka	*indica*	*ind2*
125619	LARHA MUGAD::IRGC 52339-1	India	*indica*	*ind2*
125621	PERUNEL::IRGC 63113-1	India	*indica*	*ind2*

125648	ARC 14060::IRGC 41374-1	India	*indica*	*ind2*
125650	ARC 14654::IRGC 41663-1	India	*indica*	*ind2*
125657	ASHMBER::IRGC 27522-1	Bangladesh	*indica*	*ind2*
125664	BADAL 1163::IRGC 32796-1	Bangladesh	*indica*	*ind2*
125666	BADUIE::IRGC 53715-1	India	*indica*	*ind2*
125674	BARIK KUDI::IRGC 52807-1	India	*indica*	*ind2*
125677	BENGALY MORIMO::IRGC 10976-1	Madagascar	*indica*	*ind2*
125684	BK 26::IRGC 45197-1	India	*indica*	*ind2*
125696	CHAKOL::IRGC 77226-1	Bangladesh	*indica*	*ind2*
125754	GOJOL GORIA::IRGC 26629-1	Bangladesh	*indica*	*ind2*
125759	HD 10::IRGC 6638-1	Australia	*indica*	*ind2*
125785	JHODI BIRUN::IRGC 31812-1	Bangladesh	*indica*	*ind2*
125789	KALABAIL::IRGC 25877-1	Bangladesh	*indica*	*ind2*
125792	KALU ILANKALAYAN::IRGC 36270-1	Sri Lanka	*indica*	*ind2*
125804	KITRANA 1007::IRGC 68517-1	Madagascar	*indica*	*ind2*
125814	KUSHIARA::IRGC 34709-1	Bangladesh	*indica*	*ind2*
125815	KUTTA::IRGC 52184-1	India	*indica*	*ind2*
125817	LABRA::IRGC 74757-1	India	*indica*	*ind2*
125818	LALKA (LAL DHAN)::IRGC 64946-1	Fiji	*indica*	*ind2*
125832	MAKALIOKA::IRGC 376-1	Madagascar	*indica*	*ind2*
125833	MAMORIAKA::IRGC 68672-1	Madagascar	*indica*	*ind2*
125845	MULLIKURUVA::IRGC 77529-1	India	*indica*	*ind2*
125847	MUTA GANJE::IRGC 26744-1	Bangladesh	*indica*	*ind2*
125853	NCS 237::IRGC 62202-1	India	*indica*	*ind2*
125868	PARA NELLU::IRGC 50009-1	India	*AusB*	*ind2*
125878	PURA BINNI::IRGC 26772-1	Bangladesh	*indica*	*ind2*
125881	RACE PERUMAL::IRGC 55347-1	Sri Lanka	*indica*	*ind2*
125887	RIZ TYPE SORGHO::IRGC 69015-1	Madagascar	*indica*	*ind2*
125935	WANGA BARUGULU::IRGC 52261-1	India	*indica*	*ind2*
126041	498-2A BR 8::IRGC 5891-1	India	*indica*	*ind2*
126083	ROJOFOTSY::IRGC 69402-1	Madagascar	*indica*	*ind2*
126119	NIBARI::IRGC 67742-1	India	*indica*	*ind2*
126136	KALO CHAKOL::IRGC 77258-1	Bangladesh	*indica*	*ind2*
126139	MODDAI KARUPPAN::IRGC 15465-1	Sri Lanka	*indica*	*ind2*
126143	SITHAIYAN KOTTAI SAMBA::IRGC50155-1	Sri Lanka	*indica*	*ind2*
126150	G 25::IRGC 45733-1	India	*indica*	*ind2*
126151	GOKULGANJA::IRGC 45701-1	India	*indica*	*ind2*
126158	MAKRO::IRGC 74763-1	India	*indica*	*ind2*
126161	SONAMUKHI::IRGC 46693-1	India	*indica*	*ind2*
126167	MEKENZIE SMALL::IRGC 49895-1	Guyana	*indica*	*ind2*

126199	DHANE BURWA::IRGC 10105-1	India	*indica*	*ind2*
126251	NCS 840::IRGC 62530-1	India	*AusB*	*ind2*
126280	T 315::IRGC 54792-1	India	*indica*	*ind2*
126289	URAIBOOL::IRGC 52785-1	India	*indica*	*ind2*
126294	XITTO::IRGC 6671-1	India	*indica*	*ind2*
127096	19::IRGC 70786-1	India	*indica*	*ind2*
127107	ADT 12::IRGC 6254-1	India	*indica*	*ind2*
127321	DISSI::IRGC 101346-1	Cameroon	*indica*	*ind2*
120902	BPI 76 NON SENSITIVE (GREEN)::IRGC 9790-1	Philippines	*indica*	*ind3*
121235	KHAO DAWK MALI 105::IRGC 27748-2	Thailand	*indica*	*ind3*
125628	YEBAWYIN::IRGC 33885-1	Myanmar	*indica*	*ind3*
125799	KETAN SERANG::IRGC 14615-1	Indonesia	*indica*	*ind3*
125838	MELEKE::IRGC 56823-1	Côte D'Ivoire	*indica*	*ind3*
125880	QUERO ASSAN::IRGC 28860-1	Portugal	*indica*	*ind3*
125882	RADEN KARAMUNTING::IRGC 20098-1	Indonesia	*indica*	*ind3*
125901	SIPULUT HITAM PENDEK::IRGC 20154-1	Indonesia	*indica*	*ind3*
125965	HAWM KRUA::IRGC 64333-1	Thailand	*indica*	*ind3*
126079	PATISAIL::IRGC 37562-1	Bangladesh	*indica*	*ind3*
126081	PLI KHAO::IRGC 64596-1	Thailand	*indica*	*ind3*
126132	E DAW HAWM::IRGC 47938-1	Thailand	*indica*	*ind3*
126142	SIMET 2::IRGC 25734-1	Indonesia	*indica*	*ind3*
126152	KAM PAI::IRGC 78245-1	Thailand	*indica*	*ind3*
126153	KHAO' HAWM::IRGC 78257-1	Thailand	*indica*	*ind3*
126157	LEUANG YAI 29-12-2::IRGC 881-1	Thailand	*indica*	*ind3*
126160	RELLY::IRGC 14623-1	Indonesia	*AusB*	*ind3*
126169	SRAU SENG::IRGC 30290-1	Vietnam	*indica*	*ind3*
126204	ES 21::IRGC 56171-1	Tanzania	*indica*	*ind3*
126216	JAO LEUANG::IRGC 65866-1	Thailand	*indica*	*ind3*
126223	KHAO THI RATE::IRGC 58041-1	Myanmar	*indica*	*ind3*
126226	KUNENG::IRGC 71545-1	Malaysia	*indica*	*ind3*
126236	LUA CHAN HUONG::IRGC 16800-1	Vietnam	*indica*	*ind3*
126250	Na souan::IRGC 11889-1	Lao Pdr	*indica*	*ind3*
126258	PAH WEAN::IRGC 78276-1	Thailand	*indica*	*ind3*
126262	PULUT BARAYA::IRGC 27393-1	Indonesia	*indica*	*ind3*
127207	BANDI::IRGC 17214-1	Indonesia	*indica*	*ind3*
127208	BANGKOUY::IRGC 94037-1	Cambodia	*indica*	*ind3*
127235	BONG SEN::IRGC 7011-1	Vietnam	*indica*	*ind3*
127244	C 166-135::IRGC 50633-1	Philippines	*indica*	*ind3*
127250	CEMPO MANGGAR::IRGC 27107-1	India	*indica*	*ind3*
127268	CHAO PEUAK DENG::IRGC 11602-1	Lao Pdr	*indica*	*ind3*

117454	CO 39::IRGC 51231-1	India	*indica*	*indx*
120914	CHUA DAU::IRGC 4785-1	China	*indica*	*indx*
120921	CSR-90 IR-2::IRGC 117327-1	India	*indica*	*indx*
120922	CT 9737-6-1-1-2-2P-M::IRGC 117330-1	Colombia	*indica*	*indx*
120993	IRGA 659-1-2-2-2::IRGC 117345-1	Colombia	*indica*	*indx*
121103	RR 272-17-829::IRGC 117354-1	Indonesia	*indica*	*indx*
121154	WAR 72-2-1-1::IRGC 117361-1	Sierra Leone	*indica*	*indx*
121167	BR 5230-46-4::IRGC 117318-1	Bangladesh	*indica*	*indx*
121599	BAT DO::IRGC 7014-1	Vietnam	*indica*	*indx*
122093	IR 2344-P1 PB-9-3-2B::IRGC 39317-C1	Philippines	*indica*	*indx*
122255	SONA::IRGC 26971-C1	India	*indica*	*indx*
124442	SINNA SITHIRA KALI::IRGC 51064-1	Sri Lanka	*indica*	*indx*
125609	FEI GAI 122::IRGC 63599-1	China	*indica*	*indx*
125613	JIN JUN DAO::IRGC 59710-1	China	*indica*	*indx*
125636	ARC 10594::IRGC 12524-1	India	*indica*	*indx*
125637	ARC 10754::IRGC 12603-1	India	*indica*	*indx*
125641	ARC 11524::IRGC 42672-1	India	*indica*	*indx*
125643	ARC 11857::IRGC 40972-1	India	*indica*	*indx*
125645	ARC 12576::IRGC 22163-1	India	*indica*	*indx*
125647	ARC 13778::IRGC 41216-1	India	*indica*	*indx*
125649	ARC 14064::IRGC 41377-1	India	*indica*	*indx*
125653	ARC 15873::IRGC 43250-1	India	*indica*	*indx*
125654	ARC 18092::IRGC 42256-1	India	*indica*	*indx*
125655	ARC 18112::IRGC 42274-1	India	*indica*	*indx*
125658	ASU::IRGC 62154-1	Bhutan	*indica*	*indx*
125663	BA BAI GU::IRGC 79580-1	China	*indica*	*indx*
125668	BAIANG 6::IRGC 6129-1	Indonesia	*indica*	*indx*
125671	BAMOA A 75::IRGC 51101-1	Mexico	*indica*	*indx*
125695	CAUVERY::IRGC 45255-1	India	*indica*	*indx*
125706	CHNNOR::IRGC 67485-1	India	*indica*	*indx*
125715	CR 60-10::IRGC 15777-1	India	*indica*	*indx*
125719	DA GANG ZHAN::IRGC 67103-1	China	*indica*	*indx*
125731	E 2040::IRGC 67968-1	China	*indica*	*indx*
125736	EX FOILAEIN (NAPUTO)::IRGC 81675-1	Mozambique	*indica*	*indx*
125755	GUI HUA ZAO::IRGC 68060-1	China	*indica*	*indx*
125756	H 6::IRGC 157-1	Sri Lanka	*indica*	*indx*
125765	HTA 22::IRGC 45827-1	Thailand	*indica*	*indx*
125773	IR 13429-109-2-2-1::IRGC 63491-1	Philippines	*indica*	*indx*
125805	KN 1 B 361-1-8-6-9::IRGC 46974-1	South Korea	*indica*	*indx*
125810	KULA KARUPPAN::IRGC 55328-1	Sri Lanka	*indica*	*indx*

125821	LEAD::IRGC 5805-1	Malawi	*indica*	*indx*
125829	LU MAO ZHAN::IRGC 68159-1	China	*indica*	*indx*
125836	MEI FENG 9::IRGC 63735-1	China	*indica*	*indx*
125850	NAZIRA SAIL::IRGC 77284-1	Bangladesh	*indica*	*indx*
125854	NCS 964 C::IRGC 62604-1	India	*indica*	*indx*
125859	NX 3533::IRGC 63796-1	China	*indica*	*indx*
125863	O. SATIVA::IRGC 17083-1	Taiwan	*indica*	*indx*
125866	PALEPYU::IRGC 33549-1	Myanmar	*indica*	*indx*
125874	PSBRC 68::IRGC 99711-1	Philippines	*indica*	*indx*
125890	RPA 5929 (K 45)::IRGC 33963-1	India	*indica*	*indx*
125904	SML AWINI::IRGC 13391-1	Suriname	*indica*	*indx*
125924	TOC 5430::IRGC 70487-1	Panama	*indica*	*indx*
125925	TONG GU HONG::IRGC 81026-1	China	*indica*	*indx*
125933	VEN THAP::IRGC 56138-1	Viet Nam	*indica*	*indx*
125953	B 6149 F-MR-7::IRGC 117314-1	Indonesia	*indica*	*indx*
125956	CAMPONI::IRGC 116963-1	Suriname	*indica*	*indx*
125957	CEA 3::IRGC 116965-1	Paraguay	*indica*	*indx*
125961	CUYAMEL 3820::IRGC 116975-1	Mexico	*indica*	*indx*
125966	IA CUBA 17::IRGC 116990-1	Cuba	*indica*	*indx*
125968	ICTA MOTAGUA::IRGC 116995-1	Guatemala	*indica*	*indx*
125972	IR 21015-72-3-3-3-1::IRGC 117004-1	Philippines	*indica*	*indx*
125979	IR 69502-6-SRN-3-UBN-1-B::IRGC 117290-1	Philippines	*indica*	*indx*
125986	IR 80310-12-B-1-3-B::IRGC 117307-1	Philippines	*indica*	*indx*
125987	IR 80340-23-B-12-6-B::IRGC 117309-1	Philippines	*indica*	*indx*
126014	XI GAN JING REN::IRGC 60035-1	China	*indica*	*indx*
126042	ARC 12884::IRGC 22417-1	India	*indica*	*indx*
126043	ARC 18597::IRGC 43299-1	India	*indica*	*indx*
126044	B 4414 F-MR-6-3::IRGC 117310-1	Indonesia	*indica*	*indx*
126064	IRGA 959-1-2-2F-4-1-4A-6-CA-6X::IRGC 117006-1	Brazil	*indica*	*indx*
126066	JUMA 62::IRGC 117011-1	Dominican Republic	*indica*	*indx*
126071	MENTIK TJERE BELUT::IRGC 18254-1	Indonesia	*indica*	*indx*
126075	NIAO YAO::IRGC 5496-1	Taiwan	*indica*	*indx*
126080	PICO NEGRO::IRGC 55849-1	Ecuador	*indica*	*indx*
126084	RPW 9-4 (SS 1)::IRGC 50690-1	India	*indica*	*indx*
126085	RUSTIC::IRGC 117026-1	Guyana	*indica*	*indx*
126131	ASHI BINNI::IRGC 77216-1	Bangladesh	*indica*	*indx*
126173	ARC 10812::IRGC 21074-1	India	*indica*	*indx*
126175	ARC 18202::IRGC 42328-1	India	*indica*	*indx*
126178	BALASURIYA A::IRGC 66509-1	Sri Lanka	*indica*	*indx*
126184	BAZAIL::IRGC 27526-1	Bangladesh	*indica*	*indx*

126203	EPEAL 102::IRGC 78698-1	Brazil	*indica*	*indx*
126215	IRRIBINI::IRGC 49094-1	Bangladesh	*indica*	*indx*
126309	RTS 16::IRGC 8235-1	Viet Nam	*indica*	*indx*
126312	SOLOMON RED RICE::IRGC 50950-1	Solomon Islands	*indica*	*indx*
127031	ALTAMIRA 9::IRGC 116953-1	Nicaragua	*indica*	*indx*
127033	ARC 11901::IRGC 21727-1	India	*indica*	*indx*
127050	INIAP 6::IRGC 117002-1	Ecuador	*indica*	*indx*
127053	IR 63295-AC 209-7::IRGC 117365-1	Philippines	*indica*	*indx*
127072	FONAIAP 2::IRGC 116985-1	Venezuela	*indica*	*indx*
127110	AE NOUA::IRGC 89308-1	Lao Pdr	*indica*	*indx*
127204	BAKASI::IRGC 27074-1	Indonesia	*indica*	*indx*
127231	BKN BR 1031-78-5-4::IRGC 55927-1	--	*indica*	*indx*
127237	BR 51-115-4::IRGC 43999-1	Bangladesh	*indica*	*indx*
127243	C 1016-1::IRGC 50368-1	Philippines	*indica*	*indx*
127255	CHAM LEK::IRGC 89387-1	Lao Pdr	*indica*	*indx*
127286	CN 44-40-7::IRGC 45368-1	India	*indica*	*indx*
127288	CUN GU NUO::IRGC 63576-1	China	*indica*	*indx*
125757	HAN NUO::IRGC 59591-1	China	*AusB*	*temp*
125879	PUTTIGE::IRGC 52588-1	India	*indica*	*temp*
117597	IR 62266-42-6-2::IRGC 117397-1	Philippines	*indica*	*trop*
126047	BEUREUM MEULIT::IRGC 35563-1	Indonesia	*indica*	*trop1*
126049	BOTOHAVANA MENA::IRGC 69349-1	Madagascar	*indica*	*trop1*
126146	ARC 6044::IRGC 12190-1	India	*indica*	*trop1*

IRGC: International Rice Germplasm Center Accession number, Species: *Oryza sativa* (Linnaeus), ind: *indica*, *Aus/boro: AusB*, Subpop: Subpopulation.

APPENDIX B

LIST OF PRIMERS USED IN PCR AMPLIFICATION OF PURPLE RICE GENES

OSB1 AND *OSB2*

Gene	Direction	Sequence
OSB1	Forward	GCATGACACGCCTTAATTTCC
OSB1	Reverse	ACCAGGATAGGACATCCCCC
OSB2	Forward	AGCTATGGTGCTCTTCCTCC
OSB2	Reverse	CCCAAGGCTCGTCTTCTTCT
OSB2	Forward	TTCCCGTATTACGTAGGACACTATC
OSB2	Reverse	TGGAGGAATAACTAGAAAACAAACGTGC

APPENDIX C

LIST OF gRNAs USED IN VALIDATION OF PURPLE RICE GENES *OSB1* AND *OSB2*

gRNAs	Exon	Gene	Sequence
gRNA1	2	*OSB1*	ACAACGGCGAGATAAAGACGAGG
gRNA2	2	*OSB1*	AAGATCTCGGGGACACGGAATGG
gRNA3	3	*OSB2*	TCTACGTGCGCTCTTAGCGAAGG
gRNA4	5	*OSB2*	CCCCTTCATGAGTGGCGTGCTTG