DISSERTATION

IDENTIFICATION OF A GENE THAT *MAKES-PLANTS-GIGANTIC-1*: CHARACTERIZATION OF *mpg1*, A NOVEL MUTANT OF RICE

Submitted by

Michael Maurice Friedman

Department of Biology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Spring 2020

Doctoral Committee:

Advisor: Daniel R. Bush

Cristiana T. Argueso
Taiowa A. Montgomery
A.S.N. Reddy

ABSTRACT

IDENTIFICATION OF A GENE THAT *MAKES-PLANTS-GIGANTIC-1*:

CHARACTERIZATION OF

mpg1, A NOVEL MUTANT OF RICE

The growing world population has been putting considerable strain on energy and food demands. To address the ever-growing need to meet these demands and service the global populous, emphasis has been placed on developing new methods to generate additional fuel and food. Plants play a unique role in this challenge, as they offer a means to create sustainable sources of energy as well as provide a source of food. It is important to investigate ways to increase plant productivity so that society can dually and effectively address these needs.

Plant material can be converted into numerous combustible fuel sources. High lignin content in plant secondary cell walls is desirable for thermochemical conversion, while low lignin content is more advantageous for enzymatic approaches targeting cellulose embedded in the cell wall matrix - both are substrates for converting biomass to energy. As essential building blocks of plant cell walls, increasing plant biomass fundamentally increases the energy stored in plant tissues.

Our original approach to increase biomass focused on manipulation of the source-to-sink transport of carbon. More specifically, we aimed to increase plant biomass by engineering a transfer DNA (T-DNA) expression cassette that drives the overexpression of sucrose transport in rice phloem. The hypothesis simply suggested transporting more carbon (sucrose) from leaf tissue to heterotrophic tissues in the plant would increase the biomass of individual plants. Rice

was used in these experiments because it is a model grass and staple crop, and is also a useful source for translational biology because it is closely related to bioenergy feedstocks such as sorghum, switchgrass, and *Miscanthus*.

Through the screening of a large population of transgenic rice plants, we discovered a single plant that was significantly larger than the wild-type control. Further investigation revealed that this transgenic line not only showed an increase in biomass but also exhibited an increase in seed yield as well. Additionally, the extent of growth enhancement varied in the presence of stress, where this plant yielded higher biomass than wild-type plants under various stressors. This rice was more robust during optimal conditions, and even more so during stressed conditions compared to wild-type plants.

Sequencing the DNA region around and including the T-DNA insertion event in this plant revealed that only a portion of the expression cassette was successfully inserted. Remarkably, the insertion did not contain the sucrose transporter gene that had been engineered into the cassette. Thus, the phenotype of the transgenic plant is not the result of the expressed transgene. Because integration of T-DNA into a chromosome can be a mutagenic event, we hypothesized that the insertion might have altered the expression of a nearby gene(s) that is responsible for the increased biomass, seed yield, and stress tolerance phenotype.

In support of this hypothesis, we showed that the expression cassette that inserted in the genome (monitored through molecular analysis of the insertion site) segregated with the phenotype across multiple generations. Due to the increase in biomass we refer to this mutant as *mpg1* (*makes plants gigantic-1*). Examining the expression of neighboring genes via semi-quantitative RT-PCR we discovered that one gene, a transcription factor from the APETALA 2-Ethylene-Responsive Element-Binding Protein (AP2/EREBP) transcription factor superfamily,

has markedly increased expression in *mpg1* compared to wild-type plants. This transcription factor belongs specifically to the AP2/ERF subfamily, which members have been shown to play a role in growth, development, and stress response. The *mpg1* plants exhibit a pleiotropic phenotype consisting of greater plant height, larger stems, larger leaves, increased seed yield, delay in flowering, enhanced ratooned growth, and degrees of stress tolerance compared to wild-type plants.

Transcriptomic analysis of homozygous *mpg1* and wild-type null segregants taken during the vegetative growth period prior to and during our ability to measure the biomass increases revealed a large-scale difference in gene expression from numerous genes that play roles in transcription factor activity, flower development, response to stress, DNA metabolism, cell cycle, defoliation response, cell wall metabolism, and hormone regulation. Identification of the mechanism(s) responsible for the increased biomass, seed yield, and degrees of stress tolerance may lead to strategies that could be applied to other plants to aid in both energy and food security alike.

ACKNOWLEDGMENTS

Here I would like to acknowledge the many people and programs that had a profound impact on the making of this PhD.

I would like to thank the members of my committee. First, to my advisor, Dr. Daniel Bush, thank you for allowing me to work in your lab. From you I have learned to better question and analyze information, as well as how to more effectively communicate and delineate data and its meaning. Dr. Cristiana Argueso, thank you for your continued understanding and support. Although I am not a member of your lab, on numerous occasions for several years, you have given me a significant level of guidance and counseling on not only science but life in general. Your professional demeanor and critical insights have fostered my upbringing in the scientific community. Dr. Taiowa Montgomery, thank you for providing me with knowledge regarding computational sciences and future career avenues. Your work and support allowed me to appropriately gauge and understand aspects of my research. Dr. A.S.N. Reddy, thank you for your guidance and flourishing outlook on science. Your mentality and curiosity provided me with a sense of devoted discovery and an unwavering need to persist in my studies.

Next, I would like to thank several research scientists I was fortunate enough to brush shoulders with. Dr. Bettina Broeckling has been there since the beginning of my research endeavors. You provided me with the basic skills and understanding necessary to participate in a research laboratory. You were a perpetual source of wisdom on my research, and always provided a helping hand throughout the process. From the bottom of my heart, thank you. Dr. Amanda Broz was another initial contact during my early years of research. You took me under your wing and educated me on a multitude of techniques and scientific approaches. I consider

you equally a friend and a mentor. Your invitation to participate in community programs involving sustainability made me feel more at home in Fort Collins. Your continued guidance and criticism has made me a better scientist. Without your initial research in rice my own would not have been possible, thank you.

I would also like to thank Dr. Jan Leach and members of her lab for sharing their high-caliber knowledge of rice research. They not only offered up members of their lab to help with some of my harvests, but also provided insight and speculation to my research. In particular, I would like to thank Dr. Jillian Lang for her assistance and successful co-authorship on the Student Collaborative Interdisciplinary Research Grant provided by the Cell and Molecular Biology Program at CSU, which funded aspects of my research.

I would also like to thank the following programs and affiliates: the NSF IGERT MAS BioEnergy (Multidisciplinary Approaches to Sustainable BioEnergy) program, and the NSF GAUSSI (Generating, Analyzing, and Understanding Sensory and Sequencing Information) program. My participation along-side these programs provided me with a better understanding of how research functions by collaboration, and how the influence of society and government directly impacts the adoption of innovation and discovery.

The technology transfer office, CSU – Ventures, deserves recognition for their years of mentorship and education. I enjoyed every moment during my time as an Ambassador with this amazing organization. My experience there helped illuminate the complex aspects involved in ushering innovation from conception through to commercialization. Thank you for giving me the opportunity and for connecting me with players involved in this field.

Lastly, I would like to thank the remaining sources of awards and grants that allowed me to further my research, and participate in scientific community engagements. The CSU

Department of Biology and the Colorado Center for Biorefining and Biofuels provided funding for me to travel to numerous conferences to present my research. The Illumina RNA-seq Library Prep and Grant Program Award enhanced my understanding of transcriptomics and provided funding to perform experimentation. This research was also supported by the USDA National Institute of Food and Agriculture (Award# 2008-35504-04852) and the Office of Science (BER) and US Department of Energy (DE-FG02- 08ER64629).

DEDICATION

In dedication.

My loving and supportive family and friends,

for without them

I would have not made it

through graduate school.

TABLE OF CONTENTS

DEFINITION OF TERMS

tWT:

True wild-type - plants whose predecessors didn't undergo any agrobacterium-mediated transformation.

WT-ns:

Wild-type null segregant - plants whose predecessors underwent agrobacterium-mediated transformation that, through segregation, resulted in the absence of the bi-laterally truncated T-DNA insertion indicative of the *mpg1* mutant.

HT-*mpg1*:

Heterozygous *mpg1* - plants whose predecessors underwent agrobacterium-mediated transformation that, through segregation, resulted in being heterozygous for the bi-laterally truncated T-DNA insertion indicative of the *mpg1* mutant.

HM-*mpg1*:

Homozygous *mpg1* - plants whose predecessors underwent agrobacterium-mediated transformation that, through segregation, resulted in being homozygous for the bi-laterally truncated T-DNA insertion indicative of the *mpg1* mutant.

DEG:

Differentially Expressed Genes – genes that were deemed significantly differentially expressed between treatment groups through transcriptomic analysis following RNA-seq experimentation.

GENE DESIGNATION (ID):

All genes named in this manuscript utilize the MSU identification naming style (LOC_OsXXgXXXXX) unless stated otherwise. Additionally, initial 'LOC_' is removed for most genes in the body of the text for sake of space and ease of reading

CHAPTER 1: Introduction – Rationale for experimentation: Global energy and food demands, hinderances, and current prospects in plant biotechnology

The rising population challenges resource needs, production, and stability

Elevated stress regarding energy and food have developed as a result of the steadily increasing global population, especially with respect to swelling demands. Means of generating food are insufficient to meet rising needs, and predominant sources of energy generation are unsustainable. Additionally, modern means of both resource production and use influence climate change, which adds to the complexity of this problem (Fedoroff et al., 2010).

The global population as of 2017 approached nearly 7.6 billion, of which, roughly 1 billion arose from the previous twelve years. The current trends in population growth indicate that it is rising at a rate of roughly 1.10% per year, meaning that the population is projected to reach 8.6 billion by 2030, 9.8 billion by 2050, and 11.2 billion by 2100 (UN, 2017). Global demand for food and energy are predicted to continue to rise for at least the next four decades (Godfray et al., 2010a).

Outlook on global energy demand

It is predicted that there will be an annual 1.3% rate increase in energy consumption from the year 2000 to 2050, suggesting that annual energy consumption levels will nearly double in that same period of time (Darmstadter, 2004). Although, there are many methods for generating sources of energy, historically, global energy generation has come primarily from fossil-fuel derivatives. These sources of energy include oil, coal, and natural gas. In 2015, global energy consumption could be comparatively broken down from the following sources: 32.94% oil, 29.2% coal, 23.85% gas, 6.79% hydro, 4.44% nuclear, 1.44% wind, 0.45% solar, and 0.89% other renewables (WEC, 2016).

Even with increased demands for greater diversification in alternative energy sources, the world still relies almost entirely on fossil-fuels. This raises concern as the use of this type of energy produces substantial levels of greenhouse gas emissions, largely contributing to climate change. The most abundant source of greenhouse gas that results in the 'greenhouse effect' is carbon dioxide – accumulating through human activities such as respiration, deforestation and land use change, and burning fossil-fuels (IPCC, 2014, Karl et al., 2009, Oreskes, 2004).

Due to the growing population, overwhelming use of fossil-fuels, and constantly rising demands for energy, greenhouse emissions continue to accumulate. Continuation of modern practices and refusal to transition to renewable, sustainable, and clean methods of energy generation could give rise to long-term or even irreversible climate change with perpetually stifling resource production. The Intergovernmental Panel on Climate Change (IPCC) is a global consortium charged with assessment and predictive analysis relevant to Article 2 of the United Nations Framework Convention on Climate Change (UNFCCC). They have developed Representative Concentration Pathways (RCPs) that gauge the socio-environmental impacts of varying degrees of continued atmospheric carbon accumulation (IPCC, 2014). Their evaluations conclude that it is in the best interest of the world to diminish current world practices for energy procurement and consumption to avoid devastating outcomes. Even in a 'best-case' scenario it will take significant global contribution and participation in a variety of aspects to try to stop, or even slow, current trends in (and consequences associated with) greenhouse gas accumulation while sustainably generating necessary resources.

Outlook on global food demand

Global food demand roughly tripled between the 1960s to the late 2000s (FAOSTAT, 2011), and is predicted to continue rising (Bodirsky et al., 2015). It is estimated that global

agricultural production will have to increase by 60-110% by 2050 to sufficiently meet expected demands (Tilman et al., 2011, FAO, 2009a, OECD/FAO, 2012). Beyond sheer demand for volume, as recent as the late 2000s it was determined that roughly 14% of people around the world still lacked access to the appropriate protein, energy, and nutrient content in their food (World-Bank, 2008, FAO, 2009b, FAOSTAT, 2009). Therefore, it is not only important to increase total amount of food, but also quality and type of foods.

Additional factors such as climate change, urbanization, and alternative economic focus areas challenge food production and influence demand even more so (WRI, 2005, Satterthwaite et al., 2010, Kearney, 2010, Godfray et al., 2010b). Regions specific to agricultural production are currently seeing and expected to undergo further reduction in crop yields due to climate change (Kurukulasuriya and Rosenthal, 2003). Continued trends in climate change predict a possible 11% decrease in global crop yields and a resulting increase in crop prices by 20% by 2050 (Wiebe et al., 2015).

Alternatively, some regions further from the equator are expected, at least initially, to experience beneficial attributes from climate change. Northern hemispheric areas such as China, Canada, and Russia are predicted to experience prolonged and warmer growing seasons, and enhanced carbon sequestration (Friedman et al., 2013, Myneni et al., 1997, Cox et al., 2000, Dufresne et al., 2002, Lobell and Gourdji, 2012). However, capitalizing on this could be difficult and require significant economic inputs and infrastructural reformation. With the increasing population, higher densities of people are accumulating in urban areas (Satterthwaite et al., 2010). This negatively influences the socio-economic drive on price and production of food with values of industry and services exceeding the primary sector (food production, forestry, and mining) (Satterthwaite et al., 2010, Godfray et al., 2010b).

Unfortunately, increasing the amount of land used to farm has proven difficult as there are a number of trade-offs associated with clearing more land for agricultural use - one of which being negative ecological impacts. Deforestation and land conversion gives rise to irreversible loss of biodiversity, and is the largest contributor to carbon emissions following fossil-fuel use (Balmford et al., 2005). If technology and rates of population growth are maintained, it is predicted that agricultural land will need to see a 50% increase in land area by 2030, and 66% increase by 2050 (compared to that of the late 2000s) to maintain current food consumption levels per capita (Schneider et al., 2011). Even with accepting any negative costs associated with increasing land for agricultural use this seems unlikely to happen; from the 1960s to the late 2000s the amount of land devoted to cultivated agriculture only increased by roughly 9% globally (Pretty, 2008).

Current crop yields, as measured by crops harvested per unit of land cultivated, are not growing at rate to match the projected demand for food, resulting in a yield gap (van Ittersum et al., 2013). This yield gap is a metric constrained by current technologies and management. Therefore, discovery of novel technologies and/or management strategies that increase yield per unit area are needed to help address the yield gap challenge.

Plants not only provide a way to generate food, but they also offer a way to more sustainably generate fuel. Developing strategies to improve plant growth and crop production represents a potential way to address both energy and food demands.

Plants offer a promising source of energy generation through lignocellulosic-based bioenergy

One area of noteworthy interest is bioenergy. The major energetic inputs for bioenergy arise from solar radiation, which is a source of substantial, clean, continuous, and sustainable

energy. The earth's surface receives roughly 2.5 x 10^{21} BTU/year of energy from the sun, which is around 12,000 times more than the current global energy need of 3.0 x 10^{17} BTU/year, and around 4,000 times more than predicted energy requirements of 2050 (Demain et al., 2005). Harvesting solar energy can be performed by a variety of sources, but only biological systems offer a sufficient way to utilize and subsequently convert solar energy (while capturing atmospheric carbon and reducing the carbon footprint) into forms of combustible fuel. Solar radiation utilized by these organisms is stabilized, and converts atmospheric carbon into chemical bonds in the form of sugars which are used to support autotrophic cellular metabolism resulting in the organisms ability to live, grow, and develop.

In 2005, it was estimated that photosynthetic organisms hold a conceivable energy level equivalent to ten times the world's energy usage (Demain et al., 2005). It was also projected that terrestrial plants alone produce around 1.3 x 10^{10} metric tons of dry weight biomass per year, which has potential energetic equivalencies to account for roughly two-thirds of the global energy demands (Demain et al., 2005).

The major constituent of plant biomass is lignocellulose. Lignocellulose is considered to be the most abundant organic material on earth and accounts for 50% of all biomass (Detroy, 1981). It is made up of three major components: cellulose, hemicellulose, and lignin, which act as a means of infrastructure and are distributed as a complex matrix throughout vascular plants. The quantity and distribution of lignocellulosic content varies by plant species, age, and plant part. Components of lignocellulose are comprised of a high density of energy rich structures (monolignols and carbohydrate biopolymers) (Northcote, 1972, Betts et al., 1991). Because this material is the primary component of plant biomass, and is rich in chemical energy, it is considered a desirable material for conversion to fuel-based energy.

The individual components of lignocellulose can be the substrate for various methods of degradation and chemical conversion to generate unique platform fuel sources. Cellulosic ethanol can be generated through acid or enzyme catalyzed hydrolysis of cellulose then fermented. Longer chain hydrocarbon-based fuels can be produced through gasification of total lignocellulose followed by steps such as Fischer-Tropsch, methanol, or dimethyl ether syntheses.

Biomass-to-liquid conversion is especially beneficial compared to grain/oil-based conversion as it is inedible and lowers the competition between food and fuel production. Crops are usually grown for food (grains, fruits, tubers, roots, and other structures or tissues with easily accessible sugar or nutrient content). However, agricultural cultivation results in large levels of waste residues that could be utilized for the generation of energy. The residual biomass (non-food) material produced during crop production is often considered a waste by-product that is regularly under-utilized. It is estimated that one-third of crop remains could be used to produce energy, while the other two-thirds could be used for additional resources (USAID-EIA, 2018). Rather than accumulating waste residues from agricultural practices, diligent and critical processes could be applied to salvage this material for use. Moreover, optimization of our agricultural systems, especially through the enhancement and intensification of crops, could play a significant role in successful implementation and scaling of bioenergy.

Bioenergy continues to be a desirable prospect for a more sustainable source of energy generation, however there are caveats associated with achieving a systematically desirable life cycle. There are numerous costs with implementation and scalability accompanied with widely transferring society to newly developing technologies that do not fall in line with the currently dominant platforms. Due to swelling energy demands and difficulty associated with increasing or altering land use for agricultural undertakings, generating greater plant biomass per unit area is

particularly desirable in fundamentally increasing feedstocks for bioenergy. Additional research and continued improvements are needed before bioenergy actually becomes widely adopted - nevertheless it remains a promising avenue for energy development.

Plant productivity limiting factors

Uncovering strategies to improve plants, in the form of both biomass production and seed yield, remains a primary goal to bolster energy and food security. Although effective approaches based solely on the enhancement of baseline plant productivity will be key to generating greater levels of biomass and seed yield, crops will still be exposed to environmentally limiting constraints. Crop cultivation is not a guaranteed process, farmers are subject to numerous variables that drastically influence their ability to produce healthy and successful plants from year-to-year. Crops are frequently grown in suboptimal conditions, limiting yields based on their full genetic potential (Rockström and Falkenmark, 2000, Tollenaar and Lee, 2002, Cassman et al., 2010). Therefore, it is important to define the major factors currently driving loss of crop productivity as well as predicted mechanisms of major hinderance in the future. Factors such as: poor environment and weather conditions, pollution, pest and pathogen infestation, and weed competition significantly impede crop production. Because of the more-or-less unpredictable nature of these stressors, it will be critical to understand and alleviate the risk driven by their inherent presence.

Abiotic stress is the leading cause of crop loss worldwide, accounting for major crop yield losses in excess of 50% annually (Boyer, 1982, Bray, 2000). Abiotic stressors include, but are not limited to, exposure to inappropriate levels (high and/or low) of: light, radiation, temperature, water, salinity, heavy metals, nutrients, and pollution. To complicate matters more, plants are often simultaneously exposed to multiple stressors in the field, resulting in complex

and unique responses not understood by independent evaluation of a single stress (Suzuki et al., 2014). These constraints are predicted to further compound in the future due to continued climate change and pollution alongside reduction of arable land and water resources (Lobell et al., 2011). It is predicted that global temperatures will continue to rise an average of 2-3° C within the next 30-50 years (IPCC, 2007), with occurrences of more intense and prolonged heatwaves expected as well (Meehl et al., 2007). Resulting shifts in climate are broadly increasing the total amount of arid and semi-arid landscape and pronounced desertification (Loehman, 2010, Brown et al., 1997), with some regions expected to encounter greater episodes of flooding (Hirabayashi and Kanae, 2009, Alfieri et al., 2017). Irrigation practices have not seen notable increases in recent years due to secondary salinization land degradation, to the point where cultivation is not considered economically viable (Hillel, 2000, Qadir et al., 2014, FAO, 2013a). Increases in pollutants (oxides of nitrogen, and volatile organic compounds) generated by combustion vehicles, power plants, refineries, and other industrial operations chemically react with heat and sunlight producing damaging levels of surface (tropospheric) ozone (O_3) (Avnery et al., 2011). Consequently, a myriad of abiotic stressors are likely to pose a significant threat to future crop yields.

Biotic stress factors also provide substantial levels of stress on plants. These include a variety of pathogens, pests, and competition from weeds that are estimated to negatively affect yield potentials of major crops by as much as 26-40% globally (Oerke, 2006). These factors not only pose a current risk to plant productivity, but might become more problematic in years to come. Similar to abiotic stress, continued human activities and changing environmental conditions are expected to play a role in increasing the precedence, breadth, and impact that biotic stress factors have on crop production. Global warming has been increasing temperatures

latitudinally away from the equator, giving rise to a larger geographic distribution of these stressors, potentially providing greater severity associated with their outbreaks on crop systems (Chakraborty and Newton, 2011, Lamichhane et al., 2015). This rouses a level of uncertainty pertaining to the migration of invasive species and the evolution of native populations. Environmental circumstances can influence the success of these factors (Shaw and Osborne, 2011, Gautam et al., 2013, Luck et al., 2011). For example, increased humidity within plant canopies is expected to provide a more amenable environment for some pathogens survival (Pangga et al., 2011), while cool wet environments, from flooding or irrigation, are ideal for a variety of soil-borne pathogens (Hong and Moorman, 2005). Moreover, abiotic stress has been shown to increase plant host susceptibility to a variety of biotic stress factors (Bostock et al., 2014, Pandey et al., 2017). For instance, elevated temperatures can provoke heat induced susceptibility, which can make plants more prone to the effects of a biological attack or infestation by bacterial or fungal pathogens (Zhu et al., 2010).

Competition from weeds is also becoming more prevalent not only due to changes in climate, but from repercussions associated with management practices. Currently used methods for managing weeds, and other biotic stressors, has largely relied on chemical methods (pesticides). Use, or overuse, of these chemical methods has led to tolerance amongst competing plants and undesired pests (Green, 2014, Yu and Powles, 2014, Shaner, 2017, Dhaliwal et al., 2010). Furthermore, chemical control agents are seeing greater regulation as a result of their negative impacts on human health and the ecosystem (Birch et al., 2011).

Accounting for and circumventing the effects of stress, especially those considered to become increasingly prevalent, is highly desired. Addressing environmental risk factors will not only help mitigate hazard but act as a synergistic mechanism for enhancing overall plant

productivity by reducing the gap between yield potential and actual yield.

Improvements through biotechnology

There are a few pathways for enhancing plant biomass and seed yield: increasing overall plant productivity, increasing the allocation and efficient use of resources, and/or overcoming any negative consequences of stress on growth and development.

Improving desirable agronomic traits has traditionally been accomplished through breeding techniques. One notable example is the green revolution in the 1960s, where breeding focused on improvement of architecture (developing semi-dwarf varieties), which resulted in substantial increases in wheat and rice crop yields (Khush, 1999). Although significant advancements have been attained by breeders, their achievements are subject to time consuming and laborious efforts. Additionally, for numerous crops, yield enhancements realized by breeding have grown notably faint, as much of the genetic potential pertaining to increases have already been taken advantage of. Consequentially, grain yield improvements by means of conventional breeding have practically reached a standstill in a variety of crops (Grassini et al., 2013, Ray et al., 2013, Ray et al., 2012).

Alternatively, biotechnological developments have provided hope by offering a way to enhance crop yields through the exploitation of novel genes that is not possible by breeding. The ability to engineer and tailor particular sequences provides a greater degree of control and specificity, while also saving time, to produce desired phenotypes. Access to modern tools, such as next-generation sequencing and genome editing technologies has made experimentation, validation, and implementation of biotechnology more timely for crop improvement.

Numerous genes have been found to play roles in either enhancing biomass accumulation, increasing seed yield, or generating degrees of tolerance and defense against

abiotic and biotic stressors. There is an ongoing need to develop ways in which novel molecular and biochemical mechanisms can be used to generate plants with desirable characteristics.

Enhancing baseline plant biomass

The genes underpinning the extent of biomass regulation are not well understood. Determinants and limitations of plant growth can often be the result of mechanisms surrounding resource production, resource utilization, and drive or stimulation to grow. Analysis of numerous species and mutants, their phenotypes, and underlying causal molecular components has provided insight into potential biomass regulatory factors. Notable mechanisms influencing plant productivity have been reviewed (Simkin et al., 2019, Yadav et al., 2015, Wilkinson et al., 2012, Mathan et al., 2016). In particular, components involved with photosynthesis, metabolism and nutrient use, hormone regulation, and plant morphology and architecture have been shown to impact biomass accumulation or biomass characteristics. Examples of genes that influence biomass characteristics and successful strategies used to enhance biomass through biotechnological means can be seen in (Table 1.1).

Table 1.1: Examples of biotechnological methods found to enhance biomass accumulation or biomass characteristics within plants.

Mechanism	Gene	Species	Format	Gene product	Functional mechanisms	Reference
Photosynthesis						
	PyCyt $_{c6}$	Arabidopsis	OX	cytochrome c6	Higher photosynthetic metabolite accumulation and increased leaf and root growth	(Chida et al., 2007)
	EcGCL, EcTSR, EcglcD/E/F	Arabidopsis	OX	Glycolate dehydrogenase/glyoxylate carboligase/tartronic semialdehyde reductase	Chloroplastic photorespiratory bypass increases photosynthesis and biomass production	(Kebeish et al., 2007)
	EcGlcDH (DEFp)	Potato	P/A	Polyprotein glycolate dehydrogenase complex	Improved CO2 uptake and higher levels of photosynthetic metabolites resulting in greater shoot and root biomass	(Nolke et al., 2014)
	OsHYR	Rice	OX	AP2/EREBP TF	Activation of photosynthetic genes and greater levels of growth and grain yield	(Ambavaram et al., 2014)
	OsPETC	Rice	KD	Rieske FeS	Higher rates of CO2 assimilation and enhanced biomass production and grain yield	(Yamori et al., 2016)
	AtZEP, AtVDE, AtPsbS	Tobacco	TSE	Zeaxanthin epoxidase, violaxanthin de-epoxidase, photosystem II subunit S	Elevated expression in leaves accelerated recovery from photoprotection producing plants with increased leaf area, plant height, and dry weight	(Willcox et al., 2016)
Metabolism						
	ZmSBEI, ZmSBEIIb	Arabidopsis	OX	Starch branching enzyme	Altered starch metabolism resulted in enhanced biomass and oilseed production	(Liu et al., 2016)
	AtNLP7	Arabidopsis and Tobacco	OX	NIN-like TF	Higher levels of nitrogen and carbon assimilation producing enhanced growth	(Yu et al., 2016)
	AtSPS, AtSPP	Arabidopsis and Poplar	OX	Sucrose phosphate synthase, sucrose phosphate phosphatase	Modified carbohydrate metabolism resulted in enhanced biomass accumulation	(Maloney et al., 2015)
	GhSusA1	Cotton	OX	Sucrose synthase	Increase carbon metabolism resulting in increased biomass and fiber yield	(Jiang et al., 2012b)
	Pine GS1	Poplar	OX	Glutamine synthetase	Altered nitrogen metabolism resulting in enhanced vegetative growth	(Jing et al., 2004)
	PsGPT with AtNTT1 or EcPPase with StAGPase	Potato	OX,OX/ OX,KD	Glucose 6-phosphate/phosphate translocator, adenylate translocator, pyrophosphatase, ADP-glucose	Modifying assimilate partitioning by increasing source and sink capacities resulting in increased tuber yield	(Jonik et al., 2012)

				pyrophosphorylase		
	PvSUS1	Switchgrass	OX	Sucrose synthase	Modified carbohydrate metabolism resulting in greater biomass	(Poovaiah et al., 2015)
Hormones & Development						
	AtDWF4	Arabidopsis	OX	Cytochrome P450	Modification of the brassinosteroid biosynthetic pathway increases biomass and seed yield	(Choe et al., 2001)
	AtARGOS	Arabidopsis	OX	Predicted integral membrane protein auxin-regulated gene involved in organ size	Modifies auxin signaling pathway resulting in plants with larger organs through stimulating cell proliferation	(Hu et al., 2003)
	AtNAC1	Arabidopsis	OX	NAC TF	Modulation of the auxin-responsive pathway producing plants with more roots, larger leaves, and thicker stems	(Xie et al., 2000)
	AtATAF2	Arabidopsis	OX	NAC TF	Repression of pathogenesis-related genes through modulated hormone-responsive pathways resulting in bigger leaves, and increased biomass	(Delessert et al., 2005)
	Bacterial IPT	Arabidopsis	OX	Isopentenyl transferase	Alters levels of cytokinin influencing shoot apical meristem activity	(Rupp et al., 1999)
	AtAVP1	Arabidopsis	OX	H^+-PPase	Hyperplasia through altered regulation of auxin transport	(Li et al., 2005b)
	PdGA20ox1	Arabidopsis and Poplar	OX	Gibberellin 20-oxidase	Modification of the gibberellin biosynthetic pathway increases biomass production	(Jeon et al., 2016)
	AtHOG1	Arabidopsis	KD	Cytokinin binding protein S-adenosyl-L-homocysteine hydrolase	Delayed flowering, increased biomass, and greater seed yield through manipulation of cytokinin regulation	(Godge et al., 2008)
	AtARF2	Arabidopsis	KD	ARF TF	Modified auxin responsive pathway regulation resultant in longer inflorescence, stems, and larger leaves	(Okushima et al., 2005b)
	AtGA20-OX1	Maize	OX	Gibberellin 20-oxidase	Modification of the gibberellin biosynthetic pathway increased vegetative biomass, and cellulose and lignin content	(Voorend et al., 2016)
	AtGA20ox	Poplar	OX	Gibberellin 20-oxidase	Modification of the gibberellin biosynthetic pathway produces plants with faster growth rate, increased height, girth, and leaf size	(Eriksson et al., 2000)
	PagBEE3L	Poplar	OX	bHLH TF	Modified brassinosteroid induced response pathway enhances vegetative growth and accumulation of xylem cells in stems	(Noh et al., 2015)
	AtDWF4	Rapeseed	OX	Cytochrome P450	Modification of the brassinosteroid biosynthetic	(Sahni et al., 2016)

					pathway increased root biomass and seed yield	
	ZmGA20ox	Switchgrass	OX	Gibberellin 20-oxidase	Modification of the gibberellin biosynthetic pathway generates longer leaves, internodes, and greater biomass	(Do et al., 2016)
	LlNAC	Tobacco	OX	NAC TF	Modulated hormone-responsive pathways generating greater biomass accumulation and tolerance to abiotic stressors	(Grover et al., 2014)
	NtGA2-ox	Tobacco	KD	Gibberellin 20-oxidase	Silencing the gibberellin deactivating enzyme influences the biosynthetic pathway increasing growth and fiber production	(Dayan et al., 2010)
	TaNAC69-1	Wheat	OX	NAC TF	Modified auxin regulation results in enhanced root length, above ground biomass, and grain yield	(Chen et al., 2016b)
Cell cycle & Architecture						
	LaAP2L1	Arabidopsis	OX	AP2/EREBP TF	Increased size of aerial organ, growth and final biomass through modulation of cell proliferation	(Li et al., 2013a)
	AtAPC10	Arabidopsis	OX	Anaphase-promoting complex/cyclosome subunit	Modulates cell division resulting in increased leaf size	(Eloy et al., 2011)
	AtGRF1,3,5	Arabidopsis	OX	GRF TF	Greater leaf and cotyledon promotion through greater cell proliferation.	(Kim et al., 2003, Horiguchi et al., 2005)
	AtAN3	Arabidopsis	OX	Transcription coactivator	Modification of cell proliferation increased leaf size	(Horiguchi et al., 2005, Kim et al., 2002)
	AtANT	Arabidopsis and Tobacco	OX	AP2-like TF	Enlarged embryonic and shoot organs through increased cell number	(Mizukami and Fischer, 2000)
	AtICK1/2/5/6/7	Arabidopsis	KO	Cyclin-dependent kinase inhibitors	Modulate cell cycle and endocycle, resulting in enhanced growth and larger seeds	(Cheng et al., 2013)
	AtBB	Arabidopsis	KO	RING-finger E3 ubiquitin-protein ligase	Reduced function of organ growth repressor generates plants with larger petals and sepals	(Disch et al., 2006)
	AtSAMBA	Arabidopsis	KO	Anaphase-promoting complex/cyclosome regulator	Regulates APC/C producing larger seeds, leaves, and roots	(Eloy et al., 2012)
	AtPPD1, AtPPD2	Arabidopsis	KO	Plant-specific putative DNA-binding proteins	Modified cell proliferation and differentiation resulting in larger leaves and cotyledon laminae	(White, 2006)
	AtABAP1	Arabidopsis	KO	Armadillo BTB pre-replication complex-interacting protein	Modification of DNA replication and transcription controls cell proliferation in leaves results in increases leaf size and cell number	(Masuda et al., 2008)

	PtSHR1	Poplar	KD	GRAS TF	Modulates coordination of acceleration of plant growth through cell division and fate resulting in increased biomass	(Wang et al., 2011b)
	AtCDC27a	Tobacco	OX	Anaphase-promoting complex subunit	Increased growth rate and increased organ size through cell cycle regulation	(Rojas et al., 2009)
	AtAPC10	Tobacco	OX	Anaphase-promoting complex/cyclosome subunit	Faster growth rate and enhanced biomass accumulation through cell cycle regulation	(de Freitas Lima et al., 2013)
	AtE2FB	Tomato	OX	Cell cycle TF	Accelerated rates of development through modulation of cell cycle regulation	(Abraham and del Pozo, 2011)
MicroRNA Regulation						
	msmiRNA156d	Alfalfa	OX	microRNA precursor	Downregulation of SPL genes increased root length and enhanced biomass production	(Aung et al., 2015)
	miRNA156b	Arabidopsis	OX	microRNA	Downregulation of SPL genes results in delay in flowering and increased biomass	(Schwab et al., 2005)
	miRNA397b	Arabidopsis	OX	microRNA	Modification of laccase results in reduced lignin content and improved biomass and grain yield	(Wang et al., 2014)
	miRNA858a	Arabidopsis	OX	microRNA	Down regulation of MYB transcription factors and altered phenylpropanoid pathway enhanced vegetative growth	(Sharma et al., 2016)
	miRNA156	Red Clover	OX	microRNA	Downregulation of SPL genes modifying vegetative phase change delayed flowering, and increased biomass production	(Zheng et al., 2016)
	miRNA156	Switchgrass	OX	microRNA	Downregulation of SPL genes results in increased tiller number and overall biomass	(Fu et al., 2012)

Table 1.2: List of genes and their effects in various species. Abbreviations include (P/A) – presence/absence, (OX) – overexpression, (KD) – knockdown, (KO) – knockout, (MOD) – modified sequence, (TSE) – tissue specific expression, (TF) – transcription factor.

The manipulation of specific genes has resulted in various physiological characteristics such as: increased photosynthetic rates and efficiency, enhanced carbon and nitrogen metabolism, stimulation of growth hormones, adjusted cell proliferation and differentiation, or altered development, all of which have shown promise in generating greater biomass. Of these

techniques, a number of similar strategies have proven useful in enhancing basal biomass accumulation across several species, suggesting the potential for further translational development and application.

Continued research will be needed to more comprehensively understand genes and pathways involved in growth regulation. Upstream functional mechanisms are the primary determinate driving downstream operations, therefore a systematic approach focusing on resource allocation and use followed by the stimulation of growth and formation of particular architecture might be a successful strategy to continue exploration. Optimization and evaluation of these biotechnological strategies should be investigated within different crops, under field conditions, especially those intended as biomass feedstocks to functionally aid in bioenergy capabilities.

Enhancing baseline seed yield

Proposed mechanisms to increase seed yield have been similar to strategies investigated for enhancing biomass production. Manipulation of various regulatory elements have been shown to improve resource development, use, and plant architecture resulting in increased seed yield in a variety of species. Successful mechanisms often includes the manipulation of genes involved in photosynthesis, metabolism and nutrient use, hormone regulation, and the cell cycle and architecture. Examples of molecular and biotechnological methods that have been shown to increase plant seed yield or seed yield traits can be seen in (Table 1.2).

Table 1.2: Examples of biotechnological methods found to enhance seed yield or seed yield characteristics within plants.

Mechanism	Gene	Species	Format	Gene product	Functional mechanisms	Reference
Photosynthesis						
	AtSBPase, AtFBPA, FpGCD-H	Arabidopsis	OX	Sedoheptulose-1,7-bisphosphatase, fructose 1,6-bisphosphate aldolase, glycine decarboxylase-H protein	Altered photorespiratory pathway lead to increased CO_2 assimilation and efficiency of PSII, resulting in greater biomass and seed yield	(Simkin et al., 2017a)
	NtRieskeFeS	Arabidopsis	OX	Rieske FeS (PetC)	Increase in components of the cytochrome b_6f complex resulted in elevated levels of other b_6f core complex proteins in PSI and PSII allowing for increased photosynthetic efficiency, and electron transport generating greater biomass and seed yield	(Simkin et al., 2017b)
	EcGDH, EcGCL, EcTSR	Camelina	OX	Glycolate dehydrogenase, glycolate carboxyligase, tartronic semialdehyde reductase	Introduction of photorespiratory bypass resulted in reduced photorespiration and increased photosynthesis increasing seed yield	(Dalal et al., 2015)
	OsPETC	Rice	KD	Rieske FeS	Higher rates of CO_2 assimilation and enhanced biomass production and grain yield	(Yamori et al., 2016)
	OsHYR	Rice	OX	AP2/EREBP TF	Activation of photosynthetic genes and greater levels of growth and grain yield	(Ambavaram et al., 2014)
	BdSBPase	Rice	TSE	Sedoheptulose-1,7-biphosphatase	Increased SBPase activity and increased CO_2 assimilation resulted in increased grain yield	(Driever et al., 2017)
Metabolism						
	AtPAP2	Arabidopsis	OX	Purple acid phosphatase	Expression alters flowering time, photosynthetic and growth rate, and seed yield by modulating carbon metabolism	(Zhang et al., 2012b)
	ScSuc2	Arabidopsis	TSE	Apoplastic invertase	Expression in meristematic tissues alters metabolism resulted in improved growth and seed yield	(Heyer et al., 2004)
	ZmGln1-3	Maize	OX	Glutamine synthetase	Altered nitrogen assimilation and recycling increasing grain yield through	(Martin et al., 2006)

					increased number of kernels	
	PsAAP1	Pea	OX	Amino acid permease	Increased phloem loading of amino acids resulting in increased sink development and seed yield by altering source and sink metabolism	(Zhang et al., 2015)
	ZmSh2r6hs	Rice, Wheat	TSE	ADP-glucose pyrophosphorylase subunit	Altered starch biosynthesis pathway in endosperm tissue increases efficiency producing greater seed yield	(Smidansky et al., 2002, Smidansky et al., 2003)
	ZmPRms	Tobacco	OX	Pathogenesis-related protein	An increase in symplastic sucrose transport results in increased seed yield	(Murillo et al., 2003)
Hormones & Development						
	AtHOG1	Arabidopsis	KD	Cytokinin binding protein S-adenosyl-L-homocysteine hydrolase	Delayed flowering, increased biomass, and greater seed yield through manipulation of cytokinin regulation	(Godge et al., 2008)
	AtDHS	Arabidopsis	KD	Deoxyhypusine synthase	Delayed leaf senescence delayed flowering resulting in increased reosette leaf and root biomass, and enhanced seed yield	(Wang et al., 2003)
	AtDWF4	Arabidopsis and Poplar	OX	Cytochrome P450	Modification of the brassinosteroid biosynthetic pathway results in an increase in biomass and seed yield	(Choe et al., 2001)
	AtDWF4	Rapeseed	OX	Cytochrome P450	Modification of the brassinosteroid biosynthetic pathway results in increased root biomass and seed yield	(Sahni et al., 2016)
	OsCPB1	Rice	OX	Cytochrome P450	Modified brassinosterioid biosynthesis pathway results in increase leaf angle and seed yield	(Wu et al., 2016)
	TaNAC69-1	Wheat	OX	NAC TF	Modified auxin regulation results in enhanced root length, above ground biomass, and grain yield	(Chen et al., 2016b)
Cell cycle & Architecture						
	AtRAN1	Arabidopsis	OX	Small GTP-binding protein Ran	Modulation of the cell cycle progression promoted vegetative growth and greater seed yield, as well as abiotic stress tolerance	(Xu et al., 2016b)
	Atda1-1	Arabidopsis	OX	Predicted ubiquitin receptor	Modulation of activity determining final seed and organ size results in	(Li et al., 2008)

					increased seed and organ size	
	AtSAMBA	Arabidopsis	KO	Anaphase-promoting complex/cyclosome regulator	Regulates APC/C producing larger seeds, leaves, and roots	(Eloy et al., 2012)
	ZmPLA1	Maize	OX, TSE	Cytochrome P450	Altered period of cell division duration results in enhanced biomass and seed yield	(Sun et al., 2017)
	AtPHYA	Rice	OX	Phytochrome A	Altered regulation of development and metabolism in response to changes in light influences cell elongation and cell number resulting in decreased plant size and increased grain yield by panicle number	(Garg et al., 2006)
	OsSPL14^{WFP}	Rice	P/A	Squamosa promoter binding protein-like 14 TF	Altered regulation of shoot and panicle branching results in higher grain yield	(Miura et al., 2010)
MicroRNA regulation						
	miRNA408	Arabidopsis, tobacco, rice	OX	microRNA	Enhanced photosynthesis through improved efficiency of irradiation utilization and carbon fixation, resulting in greater growth and seed yield	(Pan et al., 2018)
	OsNAC2 (OErN)	Rice	OX	microRNA resistant NAC TF	Development of miRNA164b-resistant *OsNAC2* remained unaffected by target miRNA which resulted in altered plant architecture and increased seed yield	(Jiang et al., 2018)
	miRNA397	Rice	OX	microRNA	Downregulating laccase results in larger grains and increased panicle branching	(Zhang et al., 2013)
	miRNA398a	Rice	OX	microRNA	Silencing of specific miRNAs alters regulation of grain formation number, resulting in increased panicle length, grain number and size producing higher grain yield	(Zhang et al., 2017b)
	miRNA156	Rice	TM	microRNA	Downregulation of *OsSPL14* regulator results in reduced number of tillers, and marginal increase in grain length and 1,000-grain weight	(Zhang et al., 2017b)
	miRNA396	Rice	TM	microRNA	Downregulation of growth regulating factor	(Chandran et al., 2018)

		repressor gene results in increased grain yield and greater resistance to *Magnaporthe oryzae*

Table 1.2: List of genes and their effects in various species. Abbreviations include (P/A) – presence/absence, (OX) – overexpression, (KD) – knockdown, (KO) – knockout, (MOD) – modified sequence, (TSE) – tissue specific expression, (TF) – transcription factor, (TM) – target mimicry.

Developing crops with greater seed yield is crucial to satisfactorily meet the food demands of future. In addition to the genes listed above, others have targeted genes involved with abiotic stress and biotic stresses as mechanisms to increase crop yields. Biotechnological interventions offer a meaningful way to potentially impact the production of grain and generation of biomass for food and energy security.

Conclusion

The rising global population is resulting in elevated demands for energy and food. Projected means of generating and use of these resources, by modern methods, are anticipated to be insufficient, unsustainable, and become increasingly more difficult with changing environmental conditions. It is imperative that society develops and implements the use of novel technological methods to sustainably and sufficiently meet needs. Thus, improving crop productivity has become a priority to help tackle the challenge of global energy and food security. The use of biotechnological strategies has become particularly enticing, as it offers a realistic and timely way to enhance plants by developing greater plant biomass and increase crop yields per unit area under either optimal or sub-optimal conditions.

In the results presented here, a biotechnological approach was used to improve plant growth by manipulating assimilate partitioning. Specifically, transgenic rice plants were generated containing a gene that functions in phloem loading of sucrose in aim to stimulate photosynthetic activity and increase flux of carbon to sink tissues. Collectively, this was

predicted to impact sink tissues, potentially resulting in greater biomass, and seed yield. While the hypothesis driving my experimental approach did not result in higher yields, a serendipitous insertion mutation occurred in one of the transgenic plants altering the expression pattern of a previously undescribed transcription factor which dramatically increased biomass and yield, and also had a positive impact under stress. Thus, the main focus of this dissertation research is describing this novel mutant.

BIBLIOGRAPHY

ABRAHAM, Z. & DEL POZO, J. C. 2011. Ectopic Expression of E2FB, a Cell Cycle Transcription Factor, Accelerates Flowering and Increases Fruit Yield in Tomato. *J Plant Growth Regul,* 31**,** 11-24.

ALFIERI, L., BISSELINK, B., DOTTORI, F., NAUMANN, G., DE ROO, A., SALAMON, P., WYSER, K. & FEYEN, L. 2017. Global projections of river flood risk in a warmer world. *Earth's Future,* 5**,** 171-182.

AMBAVARAM, M. M., BASU, S., KRISHNAN, A., RAMEGOWDA, V., BATLANG, U., RAHMAN, L., BAISAKH, N. & PEREIRA, A. 2014. Coordinated regulation of photosynthesis in rice increases yield and tolerance to environmental stress. *Nat Commun,* 5**,** 5302.

AUNG, B., GRUBER, M. Y., AMYOT, L., OMARI, K., BERTRAND, A. & HANNOUFA, A. 2015. MicroRNA156 as a promising tool for alfalfa improvement. *Plant Biotechnol J,* 13**,** 779-90.

AVNERY, S., MAUZERALL, D. L., LIU, J. & HOROWITZ, L. W. 2011. Global crop yield reductions due to surface ozone exposure: 2. Year 2030 potential crop production losses and economic damage under two scenarios of O3 pollution. *Atmos Environ*.

BALMFORD, A., GREEN, R. E. & SCHARLEMANN, J. P. W. 2005. Sparing land for nature: exploring the potential impact of changes in agricultural yield on the area needed for crop production. *Glob Change Biol,* 11.

BETTS, W. B., DART, R. K., BALL, A. S. & PEDLAR, S. L. 1991. Chapter 7: Biosynthesis and Structure of Lignocellulose. *In:* BETTS, W. B. (ed.) *Biodegredation: Natural and Synthetic Materials* Springer-Verlag.

BIRCH, A. N. E., BEGG, G. S. & SQUIRE, G. R. 2011. How agro-ecological research helps to address food security issues under new IPM and pesticide reduction policies for global crop production systems. *J Exp Bot,* 62**,** 3251-61.

BODIRSKY, B. L., ROLINSKI, S., BIEWALD, A., WEINDL, I., POPP, A. & LOTZE-CAMPEN, H. 2015. Global Food Demand Scenarios for the 21st Century. *PLoS One,* 10**,** e0139201.

BOSTOCK, R. M., PYE, M. F. & ROUBTSOVA, T. V. 2014. Predisposition in plant disease: exploiting the nexus in abiotic and biotic stress perception and response. *Annu Rev Phytopathol,* 52**,** 517-49.

BOYER, J. S. 1982. Plant productivity and environment. *Science,* 218**,** 444-448.

BRAY, E. A. 2000. *In Biochemistry and molecular biology of plants,* Rockville, MD, American Society of Plant Biologists.

BROWN, J. H., VALONE, T. J. & CURTIN, C. G. 1997. Reorganization of an arid ecosystem in response to recent climate change. *Proc Natl Acad Sci U S A* 94**,** 9729-9733.

CASSMAN, K. G., GRASSINI, P. & VAN WART, J. 2010. Crop Yield Potential, Yield Trends, and Global Food Security in a Changing Climate. *Handbook of Climate Change and Agroecosystems*.

CHAKRABORTY, S. & NEWTON, C. 2011. Climate change, plant diseases and food security: an overview. *Plant Pathol,* 60**,** 2-14.

CHANDRAN, V., WANG, H., GAO, F., CAO, X. L., CHEN, Y. P., LI, G. B., ZHU, Y., YANG, X. M., ZHANG, L. L., ZHAO, Z. X., ZHAO, J. H., WANG, Y. G., LI, S., FAN, J., LI, Y., ZHAO, J. Q., LI, S. Q. & WANG, W. M. 2018. miR396-OsGRFs Module Balances Growth and Rice Blast Disease-Resistance. *Front Plant Sci,* 9**,** 1999.

CHEN, D., RICHARDSON, T., CHAI, S., LYNNE MCINTYRE, C., RAE, A. L. & XUE, G. P. 2016. Drought-Up-Regulated TaNAC69-1 is a Transcriptional Repressor of TaSHY2 and TaIAA7, and Enhances Root Length and Biomass in Wheat. *Plant Cell Physiol,* 57**,** 2076-2090.

CHENG, Y., CAO, L., WANG, S., LI, Y., SHI, X., LIU, H., LI, L., ZHANG, Z., FOWKE, L. C., WANG, H. & ZHOU, Y. 2013. Downregulation of multiple CDK inhibitor ICK/KRP genes upregulates the E2F pathway and increases cell proliferation, and organ and seed sizes in Arabidopsis. *Plant J,* 75**,** 642-55.

CHIDA, H., NAKAZAWA, A., AKAZAKI, H., HIRANO, T., SURUGA, K., OGAWA, M., SATOH, T., KADOKURA, K., YAMADA, S., HAKAMATA, W., ISOBE, K., ITO, T., ISHII, R., NISHIO, T., SONOIKE, K. & OKU, T. 2007. Expression of the algal cytochrome c6 gene in Arabidopsis enhances photosynthesis and growth. *Plant Cell Physiol,* 48**,** 948-57.

CHOE, S., FUJIOKA, S., NOGUCHI, T., TAKATSUTO, S., YOSHIDA, S. & FELDMANN, K. A. 2001. Overexpression of DWARF4 in the brassinosteroid biosynthetic pathway results in increased vegetative growth and seed yield in Arabidopsis. *Plant J,* 26**,** 573-82.

COX, P. M., BETTS, R. A., JONES, C. D., SPALL, S. A. & TOTTERDELL, I. J. 2000. Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model. *Nature,* 408**,** 184-187.

DALAL, J., LOPEZ, H., VASANI, N. B., HU, Z., SWIFT, J. E., YALAMANCHILI, R., DVORA, M., LIN, X., XIE, D., QU, R. & SEDEROFF, H. W. 2015. A photorespiratory bypass increases plant growth and seed yield in biofuel crop Camelina sativa. *Biotechnol Biofuels,* 8**,** 175.

DARMSTADTER, J. 2004. Encyclopedia of Population. *In:* DEMENY, P. & MCNICOLL, G. (eds.) *Energy and Population.* Washington, D.C., USA: Resources for the Future.

DAYAN, J., SCHWARZKOPF, M., AVNI, A. & ALONI, R. 2010. Enhancing plant growth and fiber production by silencing GA 2-oxidase. *Plant Biotechnol J,* 8**,** 425-35.

DE FREITAS LIMA, M., ELOY, N. B., BOTTINO, M. C., HEMERLY, A. S. & FERREIRA, P. C. 2013. Overexpression of the anaphase-promoting complex (APC) genes in Nicotiana tabacum promotes increasing biomass accumulation. *Mol Biol Rep,* 40**,** 7093-102.

DELESSERT, C., KAZAN, K., WILSON, I. W., VAN DER STRAETEN, D., MANNERS, J., DENNIS, E. S. & DOLFERUS, R. 2005. The transcription factor ATAF2 represses the expression of pathogenesis-related genes in Arabidopsis. *Plant J,* 43**,** 745-57.

DEMAIN, A. L., NEWCOMB, M. & WU, J. H. 2005. Cellulase, clostridia, and ethanol. *Microbiol Mol Biol Rev,* 69**,** 124-54.

DETROY, R. W. 1981. Chapter 3: Bioconversion of Agricultural Biomass to Organic Chemicals *In:* GOLDSTEIN, I. S. (ed.) *Organic Chemicals from Biomass.* Boca Raton, Florida, USA: CRC Press, Inc.

DHALIWAL, G. S., JINDAL, V. & DHAWAN, A. K. 2010. Insect Pest Problems and Crop Losses: Changing Trends. *Indian J Ecol,* 37**,** 1-7.

DISCH, S., ANASTASIOU, E., SHARMA, V. K., LAUX, T., FLETCHER, J. C. & LENHARD, M. 2006. The E3 ubiquitin ligase BIG BROTHER controls arabidopsis organ size in a dosage-dependent manner. *Curr Biol,* 16**,** 272-9.

DO, P. T., DE TAR, J. R., LEE, H., FOLTA, M. K. & ZHANG, Z. J. 2016. Expression of ZmGA20ox cDNA alters plant morphology and increases biomass production of switchgrass (Panicum virgatum L.). *Plant Biotechnol J,* 14**,** 1532-40.

DRIEVER, S. M., SIMKIN, A. J., ALOTAIBI, S., FISK, S. J., MADGWICK, P. J., SPARKS, C. A., JONES, H. D., LAWSON, T., PARRY, M. A. J. & RAINES, C. A. 2017. Increased SBPase activity improves photosynthesis and grain yield in wheat grown in greenhouse conditions. *Philos Trans R Soc Lond B Biol Sci,* 372.

DUFRESNE, J. L., FAIRHEAD, L., LE TREUT, H., BERTHELOT, M., BOPP, L., CIAIS, P., FRIEDLINGSTEIN, P. & MONFRAY, P. 2002. On the magnitude of positive feedback between future climate change and the carbon cycle. *Geophys Res,* 29**,** 43-1-43-4.

ELOY, N. B., DE FREITAS LIMA, M., VAN DAMME, D., VANHAEREN, H., GONZALEZ, N., DE MILDE, L., HEMERLY, A. S., BEEMSTER, G. T., INZE, D. & FERREIRA, P. C. 2011. The APC/C subunit 10 plays an essential role in cell proliferation during leaf development. *Plant J,* 68**,** 351-63.

ELOY, N. B., GONZALEZ, N., VAN LEENE, J., MALEUX, K., VANHAEREN, H., DE MILDE, L., DHONDT, S., VERCRUYSSE, L., WITTERS, E., MERCIER, R., CROMER, L., BEEMSTER, G. T., REMAUT, H., VAN MONTAGU, M. C., DE JAEGER, G., FERREIRA, P. C. & INZE, D. 2012. SAMBA, a plant-specific anaphase-promoting complex/cyclosome regulator is involved in early development and A-type cyclin stabilization. *Proc Natl Acad Sci U S A,* 109**,** 13853-8.

ERIKSSON, M. E., ISRAELSSON, M., OLSSON, O. & MORITZ, T. 2000. Increased gibberellin biosynthesis in transgenic trees promotes growth, biomass production and xylem fiber length. *Nat Biotechnol,* 18**,** 784-8.

FAO 2009a. Global agriculture towards 2050. *How to feed the world 2050 High-Level Expert Forum.* Rome: FAO.

FAO 2009b. State of Food Insecurity in the World 2009. Rome: Food and Agriculture Organization of the United Nations (FAO).

FAO 2013. Area equipped for irrigation and percentage of cultivated land.: Available at http://www.fao.org/nr/water/aquastat/globalmaps/index.stm.

FAOSTAT. 2009. *http://www.fao.org/faostat/en/#home* [Online]. [Accessed].

FAOSTAT. 2011. *http://www.fao.org/faostat/en/#home* [Online]. [Accessed].

FEDOROFF, N. V., BATTISTI, D. S., BEACHY, R. N., COOPER, P. J., FISCHHOFF, D. A., HODGES, C. N., KNAUF, V. C., LOBELL, D., MAZUR, B. J., MOLDEN, D., REYNOLDS, M. P., RONALD, P. C., ROSEGRANT, M. W., SANCHEZ, P. A., VONSHAK, A. & ZHU, J. K. 2010. Radically rethinking agriculture for the 21st century. *Science,* 327**,** 833-4.

FRIEDMAN, A. R., HWANG, Y.-T., CHIANG, J. C. H. & FRIERSON, D. M. W. 2013. Interhemispheric Temperature Asymmetry over the Twentieth Century and in Future Projections. *J Climate,* 26**,** 5419-5433.

FU, C., SUNKAR, R., ZHOU, C., SHEN, H., ZHANG, J. Y., MATTS, J., WOLF, J., MANN, D. G., STEWART, C. N., JR., TANG, Y. & WANG, Z. Y. 2012. Overexpression of miR156 in switchgrass (Panicum virgatum L.) results in various morphological alterations and leads to improved biomass production. *Plant Biotechnol J,* 10**,** 443-52.

GARG, A. K., SAWERS, R. J., WANG, H., KIM, J. K., WALKER, J. M., BRUTNELL, T. P., PARTHASARATHY, M. V., VIERSTRA, R. D. & WU, R. J. 2006. Light-regulated overexpression of an Arabidopsis phytochrome A gene in rice alters plant architecture and increases grain yield. *Planta,* 223**,** 627-36.

GAUTAM, H. R., BHARDWAJ, M. L. & KUMAR, R. 2013. Climate change and its impact on plant diseases. *Curr Sci,* 105**,** 1685-1691.

GODFRAY, H. C., BEDDINGTON, J. R., CRUTE, I. R., HADDAD, L., LAWRENCE, D., MUIR, J. F., PRETTY, J., ROBINSON, S., THOMAS, S. M. & TOULMIN, C. 2010a. Food security: the challenge of feeding 9 billion people. *Science,* 327**,** 812-8.

GODFRAY, H. C., CRUTE, I. R., HADDAD, L., LAWRENCE, D., MUIR, J. F., NISBETT, N., PRETTY, J., ROBINSON, S., TOULMIN, C. & WHITELEY, R. 2010b. The future of the global food system. *Philos Trans R Soc Lond B Biol Sci,* 365**,** 2769-77.

GODGE, M. R., KUMAR, D. & KUMAR, P. P. 2008. Arabidopsis HOG1 gene and its petunia homolog PETCBP act as key regulators of yield parameters. *Plant Cell Rep,* 27**,** 1497-1507.

GRASSINI, P., ESKRIDGE, K. M. & CASSMAN, K. G. 2013. Distinguishing between yield advances and yield plateaus in historical crop production trends. *Nat Commun,* 4**,** 2918.

GREEN, J. M. 2014. Current state of herbicides in herbicide-resistant crops. *Pest Manag Sci*.

GROVER, A., SINGH, S., PANDEY, P., PATADE, V. Y., GUPTA, S. M. & NASIM, M. 2014. Overexpression of NAC gene from Lepidium latifolium L. enhances biomass, shortens life cycle and induces cold stress tolerance in tobacco: potential for engineering fourth generation biofuel crops. *Mol Biol Rep,* 41**,** 7479-89.

HEYER, A. G., RAAP, M., SCHROEER, B., MARTY, B. & WILLMITZER, L. 2004. Cell wall invertase expression at the apical meristem alters floral, architectural, and reproductive traits in Arabidopsis thaliana. *Plant J,* 39**,** 161-9.

HILLEL, D. 2000. Salinity Mangement for Sustainable Irrigation. Washington, D.C.: The World Bank.

HIRABAYASHI, Y. & KANAE, S. 2009. First estimate of the future global population at risk of flooding. *Hydrol Res Lett,* 3**,** 6-9.

HONG, C. X. & MOORMAN, G. W. 2005. Plant pathogens in irrigation water: challenges and opportunities. *Crit Rev in Plant Sci,* 24**,** 189-208.

HORIGUCHI, G., KIM, G. T. & TSUKAYA, H. 2005. The transcription factor AtGRF5 and the transcription coactivator AN3 regulate cell proliferation in leaf primordia of Arabidopsis thaliana. *Plant J,* 43**,** 68-78.

HU, Y., XIE, Q. & CHUA, N. H. 2003. The Arabidopsis auxin-inducible gene ARGOS controls lateral organ size. *Plant Cell,* 15**,** 1951-61.

IPCC 2007. Climate change 2007: impacts, adaptation and vulnerability. *In:* PARRY, M. L., CANZIANI, O. F., PALUTIKOF, J. P., VAN DER LINDEN, P. J. & HANSON, C. E. (eds.) *Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge, UK: Intergovernmental Panel on Climate Change.

IPCC 2014. Climate Change 2014: Synthesis Report - Summary for Policymakers. United Nations Intergovernmental Panel on Climate Change.

JEON, H. W., CHO, J. S., PARK, E. J., HAN, K. H., CHOI, Y. I. & KO, J. H. 2016. Developing xylem-preferential expression of PdGA20ox1, a gibberellin 20-oxidase 1 from Pinus

densiflora, improves woody biomass production in a hybrid poplar. *Plant Biotechnol J,* 14**,** 1161-70.
JIANG, D., CHEN, W., DONG, J., LI, J., YANG, F., WU, Z., ZHOU, H., WANG, W. & ZHUANG, C. 2018. Overexpression of miR164b-resistant OsNAC2 improves plant architecture and grain yield in rice. *J Exp Bot,* 69**,** 1533-1543.
JIANG, Y., GUO, W., ZHU, H., RUAN, Y. L. & ZHANG, T. 2012. Overexpression of GhSusA1 increases plant biomass and improves cotton fiber yield and quality. *Plant Biotechnol J,* 10**,** 301-12.
JING, Z. P., GALLARDO, F., PASCUAL, M. B., SAMPALO, R., ROMERO, J., DE NAVARRA, A. T. & CÁNOVAS, F. M. 2004. Improved growth in a field trial of transgenic hybrid poplar overexpressing glutamine synthetase. *New Phytologist,* 164**,** 137-145.
JONIK, C., SONNEWALD, U., HAJIREZAEI, M. R., FLUGGE, U. I. & LUDEWIG, F. 2012. Simultaneous boosting of source and sink capacities doubles tuber starch yield of potato plants. *Plant Biotechnol J,* 10**,** 1088-98.
KARL, T. R., MELILLO, J. M. & PETERSON, T. C. 2009. Global climate change impacts in the United States. U.S. Global Change Research Program.
KEARNEY, J. 2010. Food consumption trends and drivers. *Philos Trans R Soc Lond B Biol Sci,* 365**,** 2793-807.
KEBEISH, R., NIESSEN, M., THIRUVEEDHI, K., BARI, R., HIRSCH, H. J., ROSENKRANZ, R., STABLER, N., SCHONFELD, B., KREUZALER, F. & PETERHANSEL, C. 2007. Chloroplastic photorespiratory bypass increases photosynthesis and biomass production in Arabidopsis thaliana. *Nat Biotechnol,* 25**,** 593-9.
KHUSH, G. S. 1999. Green revolution: preparing for the 21st century. *Genome,* 42**,** 646-55.
KIM, G. T., SHODA, K., TSUGE, T., CHO, K. H., UCHIMIYA, H., YOKOYAMA, R., NISHITANI, K. & TSUKAYA, H. 2002. The ANGUSTIFOLIA gene of Arabidopsis, a plant CtBP gene, regulates leaf-cell expansion, the arrangement of cortical microtubules in leaf cells and expression of a gene involved in cell-wall formation. *EMBO J,* 21**,** 1267-79.
KIM, J. H., CHOI, D. & KENDE, H. 2003. The AtGRF family of putative transcription factors is involved in leaf and cotyledon growth inArabidopsis. *Plant J,* 36**,** 94-104.
KURUKULASURIYA, P. & ROSENTHAL, S. 2003. Climate Change and Agriculture A Review of Impacts and Adaptations. *Climate Change Series.* The World Bank, Washington, D.C.
LAMICHHANE, J. R., BARZMAN, M., BOOIJ, K., BOONEKAMP, P., DESNEUX, N., HUBER, L., KUDSK, P., LANGRELL, S. R. H., RATNADASS, A., RICCI, P., SARAH, J. P. & MESSEAN, A. 2015. Robust cropping systems to tackle pests under climate change. A review. *Agron Sustain Dev,* 35**,** 443-459.
LI, A., ZHOU, Y., JIN, C., SONG, W., CHEN, C. & WANG, C. 2013. LaAP2L1, a heterosis-associated AP2/EREBP transcription factor of Larix, increases organ size and final biomass by affecting cell proliferation in Arabidopsis. *Plant Cell Physiol,* 54**,** 1822-36.
LI, J., YANG, H., PEER, W. A., RICHTER, G., BLAKESLEE, J., BANDYOPADHYAY, A., TITAPIWANTAKUN, B., UNDURRAGA, S., KHODAKOVSKAYA, M., RICHARDS, E. L., KRIZEK, B., MURPHY, A. S., GILROY, S. & GAXIOLA, R. 2005. Arabidopsis H+-PPase AVP1 regulates auxin-mediated organ development. *Science,* 310**,** 121-5.

LI, Y., ZHENG, L., CORKE, F., SMITH, C. & BEVAN, M. W. 2008. Control of final seed and organ size by the DA1 gene family in Arabidopsis thaliana. *Genes Dev,* 22**,** 1331-6.
LIU, F., ZHAO, Q., MANO, N., AHMED, Z., NITSCHKE, F., CAI, Y., CHAPMAN, K. D., STEUP, M., TETLOW, I. J. & EMES, M. J. 2016. Modification of starch metabolism in transgenic Arabidopsis thaliana increases plant biomass and triples oilseed production. *Plant Biotechnol J,* 14**,** 976-85.
LOBELL, D. B. & GOURDJI, S. M. 2012. The influence of climate change on global crop productivity. *Plant Physiol,* 160**,** 1686-97.
LOBELL, D. B., SCHLENKER, W. & COSTA-ROBERTS, J. 2011. Climate trends and global crop production since 1980. *Science,* 333**,** 616-20.
LOEHMAN, R. 2010. Understanding the Science of Climate Change. *Talking Points - Impacts to Arid Lands.* Fort Collins, CO: National Park Service, Natural Resource Report NPS/NRPC/NRR-2010/209.
LUCK, J., SPACKMAN, M., FREEMAN, A., TREBICKI, P., GRIFFITHS, W., FINLAY, K. & CHAKRABORTY, S. 2011. Climate change and diseases of food crops. *Plant Pathol,* 60**,** 113-121.
MALONEY, V. J., PARK, J. Y., UNDA, F. & MANSFIELD, S. D. 2015. Sucrose phosphate synthase and sucrose phosphate phosphatase interact in planta and promote plant growth and biomass accumulation. *J Exp Bot,* 66**,** 4383-94.
MARTIN, A., LEE, J., KICHEY, T., GERENTES, D., ZIVY, M., TATOUT, C., DUBOIS, F., BALLIAU, T., VALOT, B., DAVANTURE, M., TERCE-LAFORGUE, T., QUILLERE, I., COQUE, M., GALLAIS, A., GONZALEZ-MORO, M. B., BETHENCOURT, L., HABASH, D. Z., LEA, P. J., CHARCOSSET, A., PEREZ, P., MURIGNEUX, A., SAKAKIBARA, H., EDWARDS, K. J. & HIREL, B. 2006. Two cytosolic glutamine synthetase isoforms of maize are specifically involved in the control of grain production. *Plant Cell,* 18**,** 3252-74.
MASUDA, H. P., CABRAL, L. M., DE VEYLDER, L., TANURDZIC, M., DE ALMEIDA ENGLER, J., GEELEN, D., INZE, D., MARTIENSSEN, R. A., FERREIRA, P. C. & HEMERLY, A. S. 2008. ABAP1 is a novel plant Armadillo BTB protein involved in DNA replication and transcription. *EMBO J,* 27**,** 2746-2756.
MATHAN, J., BHATTACHARYA, J. & RANJAN, A. 2016. Enhancing crop yield by optimizing plant developmental features. *Development,* 143**,** 3283-94.
MEEHL, G. A., STOCKER, T. F., COLLINS, W. D., GAYE, A. J., GREGORY, J. M., KITOH, A., KNUTTI, R., MURPHY, J. M., NODA, A., RAPER, S. C. B., WATTERSON, J. G., WEAVER, A. J. & ZHAO, Z. 2007. Global Climate Projections. *In:* CLIMATE CHANGE 2007: THE PHYSICAL SCIENCE BASIS. CONTRIBUTION OF WORKING GROUP I TO THE FOURTH ASSESSMENT REPORT OF THE INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE [SOLOMON, S., D. QIN, M. MANNING, Z. CHEN, M. MARQUIS, K.B. AVERYT, M. TIGNOR AND H.L. MILLER (EDS.)] (ed.). Cambridge, UK: IPCC.
MIURA, K., IKEDA, M., MATSUBARA, A., SONG, X. J., ITO, M., ASANO, K., MATSUOKA, M., KITANO, H. & ASHIKARI, M. 2010. OsSPL14 promotes panicle branching and higher grain productivity in rice. *Nat Genet,* 42**,** 545-9.
MIZUKAMI, Y. & FISCHER, R. L. 2000. Plant organ size control: AINTEGUMENTA regulates growth and cell numbers during organogenesis. *Proc Natl Acad Sci U S A,* 97**,** 942-7.

MURILLO, I., ROCA, R., BORTOLOTTI, C. & SEGUNDO, B. S. 2003. Engineering photoassimilate partitioning in tobacco plants improves growth and productivity and provides pathogen resistance. *Plant J,* 36**,** 330-341.
MYNENI, R. B., KEELING, C. D., TUCKER, C. J., ASRAR, G. & NEMANI, R. R. 1997. Increased plant growth in the northern high latitudes from 1981 to 1991. *Nature,* 386**,** 698-702.
NOH, S. A., CHOI, Y. I., CHO, J. S. & LEE, H. 2015. The poplar basic helix-loop-helix transcription factor BEE3 - Like gene affects biomass production by enhancing proliferation of xylem cells in poplar. *Biochem Biophys Res Commun,* 462**,** 64-70.
NOLKE, G., HOUDELET, M., KREUZALER, F., PETERHANSEL, C. & SCHILLBERG, S. 2014. The expression of a recombinant glycolate dehydrogenase polyprotein in potato (Solanum tuberosum) plastids strongly enhances photosynthesis and tuber yield. *Plant Biotechnol J,* 12**,** 734-42.
NORTHCOTE, D. H. 1972. Chemistry of the plant cell wall. *Ann Rev Plant Physiol,* 23**,** 113-132.
OECD/FAO 2012. OECD-FAO Agricultural Outlook 2012-2021. Available: http://dx.doi.org/10.1787/agr_outlook-2012-en.
OERKE, E. C. 2006. Crop losses to pests. *J Agr Sci,* 144**,** 31-43.
OKUSHIMA, Y., OVERVOORDE, P. J., ARIMA, K., ALONSO, J. M., CHAN, A., CHANG, C., ECKER, J. R., HUGHES, B., LUI, A., NGUYEN, D., ONODERA, C., QUACH, H., SMITH, A., YU, G. & THEOLOGIS, A. 2005. Functional genomic analysis of the AUXIN RESPONSE FACTOR gene family members in Arabidopsis thaliana: unique and overlapping functions of ARF7 and ARF19. *Plant Cell,* 17**,** 444-63.
ORESKES, N. 2004. Beyond the ivory tower. The scientific consensus on climate change. *Science,* 306**,** 1686.
PAN, J., HUANG, D., GUO, Z., KUANG, Z., ZHANG, H., XIE, X., MA, Z., GAO, S., LERDAU, M. T., CHU, C. & LI, L. 2018. Overexpression of microRNA408 enhances photosynthesis, growth, and seed yield in diverse plants. *J Integr Plant Biol,* 60**,** 323-340.
PANDEY, P., IRULAPPAN, V., BAGAVATHIANNAN, M. V. & SENTHIL-KUMAR, M. 2017. Impact of Combined Abiotic and Biotic Stresses on Plant Growth and Avenues for Crop Improvement by Exploiting Physio-morphological Traits. *Front Plant Sci,* 8**,** 537.
PANGGA, I. B., HANAN, J. & CHAKRABORTY, S. 2011. Pathogen dynamics in a crop canopy and their evolution under chaning climate. *Plant Pathol,* 60**,** 70-81.
POOVAIAH, C. R., MAZAREI, M., DECKER, S. R., TURNER, G. B., SYKES, R. W., DAVIS, M. F. & STEWART, C. N., JR. 2015. Transgenic switchgrass (Panicum virgatum L.) biomass is increased by overexpression of switchgrass sucrose synthase (PvSUS1). *Biotechnol J,* 10**,** 552-63.
PRETTY, J. 2008. Agricultural sustainability: concepts, principles and evidence. *Philos Trans R Soc Lond B Biol Sci,* 363**,** 447-65.
QADIR, M., QUILLEROU, E., NAGIA, V., MURTAZA, G., SINGH, M., THOMAS, R. J. & NOBEL, A. D. 2014. Economics of salt-induced land degradation and restoration. *Natural Resources Forum, A United Nations Sustainable Development Journal***,** 1-27.
RAY, D. K., MUELLER, N. D., WEST, P. C. & FOLEY, J. A. 2013. Yield Trends Are Insufficient to Double Global Crop Production by 2050. *PLoS One,* 8**,** e66428.
RAY, D. K., RAMANKUTTY, N., MUELLER, N. D., WEST, P. C. & FOLEY, J. A. 2012. Recent patterns of crop yield growth and stagnation. *Nat Commun,* 3**,** 1293.

ROCKSTRÖM, J. & FALKENMARK, M. 2000. Semiarid Crop Production from a Hydrological Perspective: Gap between Potential and Actual Yields. *Crit Rev Plant Sci,* 19**,** 319-346.
ROJAS, C. A., ELOY, N. B., LIMA MDE, F., RODRIGUES, R. L., FRANCO, L. O., HIMANEN, K., BEEMSTER, G. T., HEMERLY, A. S. & FERREIRA, P. C. 2009. Overexpression of the Arabidopsis anaphase promoting complex subunit CDC27a increases growth rate and organ size. *Plant Mol Biol,* 71**,** 307-18.
RUPP, H. M., FRANK, M., WERNER, T., STRNAD, M. & SCHMULLING, T. 1999. Increased steady state mRNA levels of the STM and KNAT1 homeobox genes in cytokinin overproducing Arabidopsis thaliana indicate a role for cytokinins in the shoot apical meristem. *Plant J,* 18**,** 557-63.
SAHNI, S., PRASAD, B. D., LIU, Q., GRBIC, V., SHARPE, A., SINGH, S. P. & KRISHNA, P. 2016. Overexpression of the brassinosteroid biosynthetic gene DWF4 in Brassica napus simultaneously increases seed yield and stress tolerance. *Sci Rep,* 6**,** 28298.
SATTERTHWAITE, D., MCGRANAHAN, G. & TACOLI, C. 2010. Urbanization and its implications for food and farming. *Philos Trans R Soc Lond B Biol Sci,* 365**,** 2809-20.
SCHNEIDER, U. A., HAVLÍK, P., SCHMID, E., VALIN, H., MOSNIER, A., OBERSTEINER, M., BÖTTCHER, H., SKALSKÝ, R., BALKOVIČ, J., SAUER, T. & FRITZ, S. 2011. Impacts of population growth, economic development, and technical change on global food production and consumption. *Agr Syst,* 104**,** 204-215.
SCHWAB, R., PALATNIK, J. F., RIESTER, M., SCHOMMER, C., SCHMID, M. & WEIGEL, D. 2005. Specific effects of microRNAs on the plant transcriptome. *Dev Cell,* 8**,** 517-27.
SHANER, D. L. 2017. Lessons Learned From the History of Herbicide Resistance. *Weed Sci,* 62**,** 427-431.
SHARMA, D., TIWARI, M., PANDEY, A., BHATIA, C., SHARMA, A. & TRIVEDI, P. K. 2016. MicroRNA858 Is a Potential Regulator of Phenylpropanoid Pathway and Plant Development. *Plant Physiol,* 171**,** 944-59.
SHAW, M. W. & OSBORNE, T. M. 2011. Geographic distribution of plant pathogens in response to climate change. *Plant Pathol,* 60**,** 31-43.
SIMKIN, A. J., LOPEZ-CALCAGNO, P. E., DAVEY, P. A., HEADLAND, L. R., LAWSON, T., TIMM, S., BAUWE, H. & RAINES, C. A. 2017a. Simultaneous stimulation of sedoheptulose 1,7-bisphosphatase, fructose 1,6-bisphophate aldolase and the photorespiratory glycine decarboxylase-H protein increases CO2 assimilation, vegetative biomass and seed yield in Arabidopsis. *Plant Biotechnol J,* 15**,** 805-816.
SIMKIN, A. J., LOPEZ-CALCAGNO, P. E. & RAINES, C. A. 2019. Feeding the world: improving photosynthetic efficiency for sustainable crop production. *J Exp Bot,* 70**,** 1119-1140.
SIMKIN, A. J., MCAUSLAND, L., LAWSON, T. & RAINES, C. A. 2017b. Overexpression of the RieskeFeS Protein Increases Electron Transport Rates and Biomass Yield. *Plant Physiol,* 175**,** 134-145.
SMIDANSKY, E. D., CLANCY, M., MEYER, F. D., LANNING, S. P., BLAKE, N. K., TALBERT, L. E. & GIROUX, M. J. 2002. Enhanced ADP-glucose pyrophosphorylase activity in wheat endosperm increases seed yield. *Proc Natl Acad Sci U S A,* 99**,** 1724-9.
SMIDANSKY, E. D., MARTIN, J. M., HANNAH, L. C., FISCHER, A. M. & GIROUX, M. J. 2003. Seed yield and plant biomass increases in rice are conferred by deregulation of endosperm ADP-glucose pyrophosphorylase. *Planta,* 216**,** 656-64.

SUN, X., CAHILL, J., VAN HAUTEGEM, T., FEYS, K., WHIPPLE, C., NOVAK, O., DELBARE, S., VERSTEELE, C., DEMUYNCK, K., DE BLOCK, J., STORME, V., CLAEYS, H., VAN LIJSEBETTENS, M., COUSSENS, G., LJUNG, K., DE VLIEGHER, A., MUSZYNSKI, M., INZE, D. & NELISSEN, H. 2017. Altered expression of maize PLASTOCHRON1 enhances biomass and seed yield by extending cell division duration. *Nat Commun,* 8**,** 14752.

SUZUKI, N., RIVERO, R. M., SHULAEV, V., BLUMWALD, E. & MITTLER, R. 2014. Abiotic and biotic stress combinations. *New Phytol,* 203**,** 32-43.

TILMAN, D., BALZER, C., HILL, J. & BEFORT, B. L. 2011. Global food demand and the sustainable intensification of agriculture. *Proc Natl Acad Sci U S A,* 108**,** 20260-4.

TOLLENAAR, M. & LEE, E. A. 2002. Yield potential, yield stability and stress tolerance in maize. *Field Crop Res,* 75**,** 161-169.

UN 2017. World Population Prospects - The 2017 Revision. *Key findings & advanced tables* New York: United Nations.

USAID-EIA. 2018. *Biomass in agriculture* [Online]. USAID Energy Investment Activity. Available: https://www.usaideia.ba/en/activities/biomass/what-we-do/biomass-in-agriculture/ [Accessed].

VAN ITTERSUM, M. K., CASSMAN, K. G., GRASSINI, P., WOLF, J., TITTONELL, P. & HOCHMAN, Z. 2013. Yield gap analysis with local to global relevance—A review. *Field Crop Res,* 143**,** 4-17.

VOOREND, W., NELISSEN, H., VANHOLME, R., DE VLIEGHER, A., VAN BREUSEGEM, F., BOERJAN, W., ROLDAN-RUIZ, I., MUYLLE, H. & INZE, D. 2016. Overexpression of GA20-OXIDASE1 impacts plant height, biomass allocation and saccharification efficiency in maize. *Plant Biotechnol J,* 14**,** 997-1007.

WANG, C. Y., ZHANG, S., YU, Y., LUO, Y. C., LIU, Q., JU, C., ZHANG, Y. C., QU, L. H., LUCAS, W. J., WANG, X. & CHEN, Y. Q. 2014. MiR397b regulates both lignin content and seed number in Arabidopsis via modulating a laccase involved in lignin biosynthesis. *Plant Biotechnol J,* 12**,** 1132-42.

WANG, J., ANDERSSON-GUNNERAS, S., GABOREANU, I., HERTZBERG, M., TUCKER, M. R., ZHENG, B., LESNIEWSKA, J., MELLEROWICZ, E. J., LAUX, T., SANDBERG, G. & JONES, B. 2011. Reduced expression of the SHORT-ROOT gene increases the rates of growth and development in hybrid poplar and Arabidopsis. *PLoS One,* 6**,** e28878.

WANG, T. W., LU, L., ZHANG, C. G., TAYLOR, C. & THOMPSON, J. E. 2003. Pleiotropic effects of suppressing deoxyhypusine synthase expression in Arabidopsis thaliana. *Plant Mol Biol,* 52**,** 1223-35.

WEC 2016. World Energy: Resources | 2016. London, UK: World Energy Council.

WHITE, D. W. 2006. PEAPOD regulates lamina size and curvature in Arabidopsis. *Proc Natl Acad Sci U S A,* 103**,** 13238-43.

WIEBE, K., LOTZE-CAMPEN, H., SANDS, R., TABEAU, A., VAN DER MENSBRUGGHE, D., BIEWALD, A., BODIRSKY, B., ISLAM, S., KAVALLARI, A., MASON-D'CROZ, D., MÜLLER, C., POPP, A., ROBERTSON, R., ROBINSON, S., VAN MEIJL, H. & WILLENBOCKEL, D. 2015. Climate change impacts on agriculture in 2050 under a range of plausible socioeconomic and emissions scenarios. *Environ Res Lett,* 10.

WILKINSON, S., KUDOYAROVA, G. R., VESELOV, D. S., ARKHIPOVA, T. N. & DAVIES, W. J. 2012. Plant hormone interactions: innovative targets for crop breeding and management. *J Exp Bot,* 63**,** 3499-509.

WILLCOX, D., CHAPPELL, B. G., HOGG, K. F., CALLEJA, J., SMALLEY, A. P. & GAUNT, M. J. 2016. A general catalytic beta-C-H carbonylation of aliphatic amines to beta-lactams. *Science,* 354**,** 851-857.

WORLD-BANK 2008. World Development Report 2008: Argiculture for Development. Washington, DC: World Bank.

WRI 2005. Millenium Ecosystem Assessment, Ecosystems and Human Well-Being. Washington, DC: World Resources Institute.

WU, Y., FU, Y., ZHAO, S., GU, P., ZHU, Z., SUN, C. & TAN, L. 2016. CLUSTERED PRIMARY BRANCH 1, a new allele of DWARF11, controls panicle architecture and seed size in rice. *Plant Biotechnol J,* 14**,** 377-86.

XIE, Q., FRUGIS, G., COLGAN, D. & CHUA, N. H. 2000. Arabidopsis NAC1 transduces auxin signal downstream of TIR1 to promote lateral root development. *Genes Dev,* 14**,** 3024-36.

XU, P., ZANG, A., CHEN, H. & CAI, W. 2016. The Small G Protein AtRAN1 Regulates Vegetative Growth and Stress Tolerance in Arabidopsis thaliana. *PLoS One,* 11**,** e0154787.

YADAV, U. P., AYRE, B. G. & BUSH, D. R. 2015. Transgenic approaches to altering carbon and nitrogen partitioning in whole plants: assessing the potential to improve crop yields and nutritional quality. *Front Plant Sci,* 6**,** 275.

YAMORI, W., KONDO, E., SUGIURA, D., TERASHIMA, I., SUZUKI, Y. & MAKINO, A. 2016. Enhanced leaf photosynthesis as a target to increase grain yield: insights from transgenic rice lines with variable Rieske FeS protein content in the cytochrome b6 /f complex. *Plant Cell Environ,* 39**,** 80-7.

YU, L. H., WU, J., TANG, H., YUAN, Y., WANG, S. M., WANG, Y. P., ZHU, Q. S., LI, S. G. & XIANG, C. B. 2016. Overexpression of Arabidopsis NLP7 improves plant growth under both nitrogen-limiting and -sufficient conditions by enhancing nitrogen and carbon assimilation. *Sci Rep,* 6**,** 27795.

YU, Q. & POWLES, S. 2014. Metabolism-based herbicide resistance and cross-resistance in crop weeds: a threat to herbicide sustainability and global crop production. *Plant Physiol,* 166**,** 1106-18.

ZHANG, H., ZHANG, J., YAN, J., GOU, F., MAO, Y., TANG, G., BOTELLA, J. R. & ZHU, J. K. 2017. Short tandem target mimic rice lines uncover functions of miRNAs in regulating important agronomic traits. *Proc Natl Acad Sci U S A,* 114**,** 5277-5282.

ZHANG, L., GARNEAU, M. G., MAJUMDAR, R., GRANT, J. & TEGEDER, M. 2015. Improvement of pea biomass and seed productivity by simultaneous increase of phloem and embryo loading with amino acids. *Plant J,* 81**,** 134-46.

ZHANG, Y., YU, L., YUNG, K. F., LEUNG, D. Y., SUN, F. & LIM, B. L. 2012. Over-expression of AtPAP2 in Camelina sativa leads to faster plant growth and higher seed yield. *Biotechnol Biofuels,* 5**,** 19.

ZHANG, Y. C., YU, Y., WANG, C. Y., LI, Z. Y., LIU, Q., XU, J., LIAO, J. Y., WANG, X. J., QU, L. H., CHEN, F., XIN, P., YAN, C., CHU, J., LI, H. Q. & CHEN, Y. Q. 2013. Overexpression of microRNA OsmiR397 improves rice yield by increasing grain size and promoting panicle branching. *Nat Biotechnol,* 31**,** 848-52.

ZHENG, Q., LIU, J., GOFF, B. M., DINKINS, R. D. & ZHU, H. 2016. Genetic Manipulation of miR156 for Improvement of Biomass Production and Forage Quality in Red Clover. *Crop Sci,* 56.
ZHU, Y., QIAN, W. & HUA, J. 2010. Temperature Modulates Plant Defense Responses through NB-LRR Proteins. *PLOS Pathog,* 6.

CHAPTER 2: Developing transgenic *Oryza sativa* containing a T-DNA expression cassette designed to enhance phloem loading, and discovery and initial characterization of mutant plant *mpg1*

Introduction

The increasing global population puts tension on vital supplies, particularly food and fuel. This, along with the dynamic changes in economic development, shifts in cultural lifestyles, and increasing rates of inclement environmental trends continues to exacerbate this situation (Edgerton, 2009, Schmidhuber and Tubiello, 2007). These issues will not likely be solved by a single solution, but rather by the implementation of numerous changes in the system at large. In an effort to aid in this large-scale problem, we have explored a novel way to increase plant productivity to help meet increased demands for food and fuel through the generation of greater seed yield and cellulosic material. Our particular strategy investigated altering carbon allocation as a way to enhance photosynthesis and stimulate greater growth in plants by capturing more fixed carbon. Specifically, we tried to enhance photosynthesis by modulating a sucrose regulatory network responsible for assimilate partitioning in plants.

Plants offer a way to service and possibly sustain an expanding world population by offering a renewable source of fuels, which make them an important alternative to petroleum products (Koonin, 2006). One of the key factors determining crop growth and yield is photosynthetic carbon metabolism (Gifford and Evans, 1981). Crop yields could be enhanced by altering photosynthetic carbon metabolism to improve photoassimilate production, transport, and usage in the sink tissues. We predicted that enhancing a key step in the transport of photoassimilates from leaf tissue would result in a larger supply of carbon in sink tissues and a greater overall level of throughput thereby increasing biomass and yield. We hypothesized

enhancing flux of photoassimilates out of source tissues would eliminate or reduce carbohydrate-mediated repression of photosynthesis and delay the induction of senescence, thus increasing overall plant productivity. Many studies have investigated ways in which to enhance photosynthetic capacity by strategies other than transport of photoassimilates, but these approaches experienced relatively limited success (Galtier et al., 1993, Miyagawa et al., 2001, Paul and Foyer, 2001, Lieman-Hurwitz et al., 2003, Nunes-Nesi et al., 2005). However, a recent study engineering more efficient photorespiratory pathways in tobacco (a C3 plant) increased photosynthetic efficiency and vegetative biomass significantly (South et al., 2019).

The partitioning of resources from sites of acquisition to sites of usage and storage is a fundamental regulatory feature of organisms. In plants, carbohydrates generated by photosynthesis in the leaf (source) are transported to heterotrophic (sink) tissues where they are used. Sucrose is the major product of photosynthesis and is transported from leaves (sources) to heterotrophic tissues (sinks) (Geiger, 1975). Photosynthesis occurs within mesophyll cells of the leaf. Once synthesized, sucrose diffuses to the phloem where it is actively loaded into phloem cells and transported throughout the plant. Sucrose moves from source tissue to the phloem via two non-exclusive pathways (Giaquinta, 1983, Turgeon, 1996, Van Bel, 1993). In the symplasmic transport pathway, sucrose moves from cell to cell through the plasmodesmata, and then moves directly into the companion cell/sieve tube complex, where it can be polymerized forming a raffinose family oligosaccharide. High concentrations of these sugars result in the influx of water via osmosis. Because the phloem cells are surrounded by cell walls they cannot expand thus creating high hydrostatic pressure that drives long-distance transport. (Turgeon, 1996, Turgeon and Medville, 2004). In the apoplasmic transport pathway, sucrose is released from the photosynthetic cell into the apoplast followed by active uptake into the phloem

companion cell for transport to heterotrophic tissues. The sucrose is loaded into the vasculature against a concentration gradient (often on the order of 50-fold), by a proton-coupled sucrose symporter (sucrose transporters, SUTs) in a 1:1 stoichiometry (located in companion cells) via proton-motive force (Bush, 1993). The proton motive-force is generated by the proton-pumping ATPase by hydrolysis of ATP. Sucrose is then transported by the sucrose symporter into the phloem. The concentration of sucrose in the phloem (often approaching 1 molar) causes water influx. The resulting high hydrostatic pressure causes mass flow to sink tissues where sucrose is released, lowering the hydrostatic pressure in the phloem (Giaquinta, 1983). Sucrose can then be stored or utilized by cells for growth and development.

Photosynthesis is impacted by translocation of photoassimilates. One early example of this showed a positive correlation between translocation efficiency and CO_2 uptake in *Phaseolus vulgaris* (Liu et al., 1973). Other studies have uncovered a positive relationship between photosynthetic rates and assimilate translocation (Geiger, 1975, Grodzinski et al., 1998). Sucrose translocation however has limitations. It has been proposed that sucrose concentration in leaves is maintained within a fixed range dependent on the rate of sucrose export, and the excess sucrose in the system is diverted to starch (Komor, 2000). This idea is supported by a study that investigated tomato plants overexpressing sucrose phosphate synthase. No changes in carbon export rates were observed in these plants, however they accumulated 2-3-fold higher concentrations of carbohydrates in source leaves (Galtier et al., 1995). Additionally, sucrose stimulates transcription of starch biosynthetic enzyme ADP-glucose pyrophosphyorylase (Muller-Rober et al., 1990, Sokolov et al., 1998). This suggests that starch biosynthesis increases because of an inability to export all newly fixed carbon. Indeed, starch accumulation during the day is almost as high as carbon exported from the leaf. That starch is then broken down at night

to drive night-time sucrose export (Fondy and Geiger, 1985). Higher accumulation of soluble sugars has also been seen in plants unable to synthesize starch, which is consistent with limitations in export capability (Caspar et al., 1985, Kofler et al., 2000).

Insufficient transport of photoassimilates results in down-regulation of photosynthesis. For example, wheat grown under elevated levels of CO_2 show greater accumulation of carbohydrates while experiencing decreased photosynthetic gene transcript levels (Nie et al., 1995). Photosynthetic rates are associated with metabolism of hexoses derived from sucrose. Inhibition of photosynthesis correlates with activities of soluble acid invertase, an enzyme that converts sucrose into the hexoses glucose and fructose (Goldschmidt and Huber, 1992, Moore et al., 1998). Hydrolysis of sucrose coupled with hexose phosphorylation triggers repression of photosynthetic gene expression (Krapp et al., 1993, Jang and Sheen, 1994, Moore et al., 1998). Studies have also indicated that the decline in photosynthetic activity that occurs during senescence could be due to hexokinase-mediated signaling (Pourtau et al., 2006). This suggests that enhancing sucrose transport to decrease sugar levels in leaf tissues, could delay senescence, thereby increasing efficiency and productivity of the plant over its lifetime. The overall flux of assimilates, both from storage and draw from sinks is a crucial element of assimilate partitioning of sucrose.

An increase in sink demand, leading to increased transport of photoassimilates has been shown to stimulate greater photosynthetic activity. Single rooted sweet potato leaves with increased sink demand have shown up-regulation of photosynthetic activity, suggesting that a feed-forward mechanism of sink activity on photosynthesis might exist (Sawada et al., 2003). A number of other plant species have shown a similar feed-forward relationship between production and transport of photoassimilates, and photosynthetic activity (Grodzinski et al.,

1998). Thus, photosynthetic rates of carbon assimilation are influenced by both source export capability and sink capacity, which is significantly impacted by the rate of sucrose loading into the phloem by proton-sucrose symporters (SUT's).

The prevalence of SUT proteins correlates with sucrose translocation ability. Sucrose has been shown to act as a signaling molecule influencing transcription levels of some SUT proteins. Transcript levels of *BvSUT1* have been shown to be inversely correlated to phloem sucrose concentration in *Beta vulgaris* (Vaughn et al., 2002, Chiou and Bush, 1998). Therefore sucrose abundance can regulate symporter abundance acting as a global control mechanism of assimilate partitioning.

*SUT*s play a significant role in sucrose transport to heterotrophic tissues. *SUT*s participate in sucrose retrieval, unloading, and loading to serve multiple physiological roles (Kuhn and Grof, 2010). *SUT*s are phylogenetically categorized into separate groups some of which are exclusive to either monocot or dicot species (Braun and Slewinski, 2009). For example, five monocot *SUT*s have been identified within the rice genome (Aoki et al., 2003). Of those *OsSUT1* has been localized to companion cells and plays a role in phloem loading. (Ishimaru et al., 2001, Scofield et al., 2002, Matsukura et al., 2000). Knock-downs of *OsSUT1* result in reduced grain filling and germination rates but maintained vegetative growth compared to wild-type (Matsukura et al., 2000, Ishimaru et al., 2001, Scofield et al., 2002). Alternatively, *SUT1* in other species have been shown to be critical in assimilate partitioning and resulting physiological impact. Inhibition of *StSUT1* in potato results in increased amounts of soluble and insoluble carbohydrates in leaves, inhibition of photosynthesis, and reduced tuber yield (Riesmeier et al., 1994, Kuhn et al., 1996). Inhibition of *LeSUT1* in tomato leads to inhibition of phloem loading (Hackel et al., 2006). Tobacco with inhibited *NtSUT1* contains higher amounts of soluble carbohydrates and exhibited

reduced rates of photosynthesis (Burkle et al., 1998). Thus, it is clear that the reduction of phloem loading by impaired functionality of *SUT1* negatively impacts photosynthesis, growth, and yield.

The effect of enhancing phloem loading via manipulation of *SUT1* has yet to be fully elucidated. Ectopic expression of *StSUT1* in storage parenchyma cells of developing pea seeds resulted in greater growth rate and enhanced sucrose influx into cotyledons (Rosche et al., 2002). Further, ectopic expression of *SoSUT1* driven by the constitutive promoter *CaMV35S* in potato resulted in some plants with enhanced rates of sucrose uptake into leaf plasma membrane vesicles but no increase in radio labeled CO_2 assimilation or tuber yield compared to wild-type (Leggewie et al., 2003). It is important to clarify that in that study plants were constitutively expressing a *SUT1* gene and in doing so likely resulted in mesophyll cells competing for sucrose released into the apoplast. Because the transporter wasn't expressed in a tissue-specific manner, but rather in all tissue types, it may have resulted in a futile cycle of sucrose partitioning (Leggewie et al., 2003). Specifically, this might have caused a continuous loop of releasing and re-capturing sucrose by the mesophyll cells, which would not allow for a net flux of sucrose to be loaded into the phloem. Appropriate evaluation of the effects of increasing *SUT1* activity would require investigation of enhancement specifically at the site of phloem loading - the companion cells.

We wanted to determine how companion cell specific over-expression of a *SUT* would alter assimilate partitioning. Rice was selected to perform our experiments for several reasons. Although modern breeding efforts have resulted in yield enhancements, increased yields still remain a priority to address future demands (Hickley et al., 2017). Rice is a model crop and a global staple grain with translational capability stemming to valued bioenergy crops such as

Sorghum and *Miscanthus*. In addition, this plant has its genome sequenced, and can be genetically transformed.

Rice contains both thin and thick-walled sieve tubes, which respectively function via the apoplasmic and the symplasmic transport pathways (Botha et al., 2008). Whether or not rice favors a symplasmic or apoplasmic mechanism of phloem loading is not clear (Braun and Slewinski, 2009). Rice plants with reduced transcript levels of *OsSUT1* have similar amounts of growth to wild-type plants suggesting that rice doesn't rely exclusively on apoplasstic phloem loading (Ishimaru et al., 2001, Scofield et al., 2002).

We originally hypothesized that if we could increase the activity of the proton-sucrose symporter activity within companion cells, more sucrose would be loaded into the phloem stimulating increased photosynthetic activity by decreasing sugar content in mesophyll cells and simultaneously providing greater overall carbon flux to sink tissues, thus potentially providing for increasing biomass and yield. To test this hypothesis we generated plants expressing a non-native hyperactive proton-sucrose symporter in the companion cells of rice. Transgenic rice plants were produced by inserting a T-DNA expression cassette engineered with a companion-cell specific promoter from the *GALACTINOL SYNTHASE 1* gene in *Cucumis melo* (*CmGAS1pro*) (Volk et al., 2003, Haritatos et al., 2000a) driving tissue-specific over-expression of the hyperactive sucrose symporter from *Arabidopsis thaliana* ($AtSUC1_{H65K}$) (Haritatos et al., 2000a, Lu and Bush, 1998). Expression analyses of the *CmGAS1* promoter have revealed that it is active in source leaves and is expressed in phloem companion cells (Haritatos et al., 2000b). Additionally, expression localization correlates with known phloem loading sites identified by expression localization of phloem-loading *SUT*s (Truernit and Sauer, 1995, Stadler and Sauer, 1996, Haritatos et al., 2000b). A secondary construct was created to test our hypothesis utilizing

a different companion cell-specific promoter from the *THEIRODOXIN H* gene in *Oryza sativa* (*OsTRXhpro*) (Ishiwatari et al., 2000, Ishiwatari et al., 1998). This protein was originally found in high abundance in rice phloem through analysis of phloem sap (Ishiwatari et al., 1995). Expression is tissue-specific and exhibits a consistent level of expression, with no evidence of negative regulation by sucrose abundance. These characteristics make it a suitable promoter element to try and increase phloem loading of sucrose. A non-native heterologous *SUT* ($AtSUC1_{H65K}$) was chosen instead of the native transporter in an effort to combat any negative feedback mechanisms that might result in co-suppression thereby limiting expression of the native transporter. The particular *SUC1* gene was selected because of its enhanced capability to transport sucrose. In yeast, $AtSUC1_{H65K}$ showed increased levels of sucrose transport activity in excess of 14-fold compared to the native wild-type transporter (Lu and Bush, 1998). Thus, the particular elements engineered here should directly assess the effects of increasing *SUT1* activity in the companion cell *in planta*.

Numerous transgenic plants (created via agrobacterium-mediated transformation) carrying our designed T-DNA expression cassette were generated. Transgenic plants were screened for biomass and yield-related traits (dry weight, plant height, tiller number, and seed yield). Surprisingly, plants with successful integration of the T-DNA expression cassette resulted in decreased levels of biomass and yield compared to wild-type plants. Serendipitously, however, a single plant (independent insertion event), resulted in a plant larger in size with greater yield than wild-type plants. This particular T-DNA insertion event contained a bi-laterally truncated version of our designed T-DNA expression cassette, lacking the entire transgene of interest ($AtSUC1_{H65K}$) and a significant portion of the companion-cell specific promoter sequence (*CmGAS1pro*). Because our original aim was to investigate ways in which to

increase plant productivity, we delved further into uncovering the reasons behind why this particular plant, now representing an insertion mutation, had increased biomass and yield compared to wild-type plants.

Materials and Methods

Plant materials

Rice (*Oryza sativa* L. spp. *japonica* cv. Kitaake background), including wild-type, transgenic *CmGAS1pro*::$AtSUC1_{H65K}$ generated by T-DNA expression cassette insertion, transgenic *OsTRXhpro*::$AtSUC1_{H65K}$ generated by T-DNA expression cassette insertion, transgenic *CmGAS1pro*::*GUS* generated by T-DNA expression cassette insertion, partial T-DNA insertion of expression cassette *CmGAS1pro*::$AtSUC1_{H65K}$ - mutant (*mpg1*), and non-transgenic tissue culture regenerated wild-type lines were used in this study to evaluate growth metrics, expression analyses, and localization of gene promoter function. The T-DNA expression cassettes were engineered in the binary vector pCambia1390. Transgenic plants were generated by induction of the T-DNA expression cassettes via *Agrobacterium tumefaciens*-mediated tissue culture transformation.

Transformation of (*Oryza sativa*)

The transformation procedure closely followed the version of (Nishimura et al., 2006).

Callus induction: Rice seeds (*Oryza sativa* L. spp. *japonica* cv. Kitaake) were sterilized for calli induction. Seeds husks were removed and submerged in 40% bleach with 1 drop of Tween20 per 50 mL, in sterile water. The solution was shaken at 150 rpm for 30 min. The seeds were then washed in sterile water three times. Seeds were transferred to callus induction media (MSD: 4.4g/L MS, 30 g/L sucrose, 4.0g/L phytagel, 2 mL of 1mg/mL 2-4-D, pH 5.8). MSD

plates were sealed with vent tape and incubated at 28-30º C under continuous light for 2 weeks. Resulting calli were transferred to new MSD for 4-7 days followed by co-cultivation.

Co-cultivation with agrobacterium: *Agrobacterium tumafaciens* strain LBA4404 containing the desired transformation construct were grown in 1 mL LB supplemented with 50 µg/mL hygromycin and agitated at 28-30º C at 250 rpm for a period of 24 hrs. Then 150 µL of the *Agrobacterium* culture were added to 5 mL of TY+ 100 mg/L acetosyringone and agitated at 25º C at 250 rpm for 2-4 hrs until the OD_{600} =0.1-0.2. Two-week-old calli were added to the *Agrobacturium* culture and agitated gently for 30 minutes. *Agrobacterium* were removed from the solution, and calli were transferred to co-cultivation media (MSD+AS; 4.4 g/L of MS, 30 g/L sucrose, 50 g/L sorbitol, 6.0 g/L phytagel; pH 5.8 using KOH. After sterilization 1 mL of 200 mM acetosyringone (3', 5'-Dimethoxy-4'-hydroxyacetophenone) and 2.0 mL of 1 mg/mL 2-4-D were added to the solution. Plates were kept in the dark at room temperature (21-25º C) for 2-3 days.

Callus selection: In the occurrence of *Agrobacterium* overgrowth Calli were washed gently with sterile MS+carb solution until solution was clear. The calli were then transferred to selection media (MSD+carb+drug: 4.4 g/L MS, 30g/L sucrose, 4.0g/L phytagel, 2 mL of 1mg/mL 2-4-D: pH of 5.8 using KOH. After sterilization 1.25 mL (per L) of 200mg/mL carbenicillin, 1.0 mL (per L) PPM, and 1.0 mL (per L) of 50 mg/mL hygromycin was added to the solution. Calli were incubated on selection media under continuous light at 30º C for 2-8 weeks and were transferred to new selection media as needed.

Regeneration of transgenic calli: Resistant microcalli were regenerated by transferring them to regeneration media (4.4 g/L MS, 30 g/L sucrose, 50 g/L sorbitol, 3 mL of 1 mg/mL BAP, 0.5 mL of 1 mg/mL NAA, 6.0g/L phytagel, and set to a pH of 5.8 using KOH and

autoclaved; once cooled to 50º C 625 µL (per L) of 200 mg/mL carbenicillin, and 1 mL (per L) of 50 mg/mL hygromycin was added to the solution). Calli were incubated on media at 30º C under continuous light for 2-3 weeks. Resistant calli were transferred to new regeneration media as needed until leaf primordia emerged (4 weeks – 2 months).

When calli developed primordial leaf tissue they were transferred to rooting medium in magenta boxes. Rooting medium was made up of 4.4 g/L of MS, 30 g/L sucrose, 2.0 g/L phytagel and set to a pH of 5.8 using KOH and autoclaved; once cooled to 50º C, 625 µL (per L) of 200 mg/mL carbenicillin and 1 mL (per L) of 50 mg/mL hygromycin were added to the solution and mixed. The calli were then incubated at 30º C under continuous light for 2-3 weeks until plantlets developed. Upon further development of individual root and shoot systems, individual plantlets were transferred to planting media (see below) and placed in the growth chamber or greenhouse until plants reached maturity and T_1 seed were harvested.

Plant growth conditions

Seeds were placed on germination paper and partially submerged in a 1:1000 dilution of MAXIM XL dual action fungicide (Syngenta) and sealed with parafilm. Seeds were incubated at 30º C under 12 h light cycles, until primary shoot and root development occurred (usually 5-7 days). Seedlings were then transferred to planting medium in the greenhouse. Planting medium (non-optimal) consisted of: 1 part play sand, 4 parts Canadian sphagnum peat moss, 4 parts (Promix) BX, mixed to homogeneity. Pots were organized in random fashion in a flat or tub with water covered with black plastic and watered until media was fully saturated and pots remained in roughly 3” of standing water. Greenhouse conditions were maintained at 30˚C and 70% RH with a 16h light cycle. Plant chlorosis was monitored and preemptively treated around the 3- to 4-leaf stage using Sprint 330 Iron Chelate at 0.3g/L and top-watered. At the same developmental

stage plants were fertilized using Scotts Peters Excel 15-5-15 Cal-Mag granular fertilizer at 24.22 g/1.0 L water and top-watered. Fertilizer treatment occurred twice weekly until harvest.

The optimal planting media consisted of: 1 part (Profile) Greens Grade porous ceramic particulate, and 1 part (Promix) BX, soil. The contents were mixed to homogeneity, and transferred to pots. Plant chlorosis was monitored and preemptively treated around the 3- to 4-leaf stage using (Sprint) 330 Iron Chelate at 0.3g/L and top-watered. At the same developmental stage plants were fertilized using granulized (Technigro) 15-5-15 Plus Cal-Mag at 48.87 g/1.0 L and top-watered. Fertilizer treatment occurred twice weekly through maturity until harvest.

DNA extraction and genotyping

Young, fresh plant tissue (3-leaf stage) was sampled for DNA extraction and analysis (2-5cm of leaf-tip). DNA was obtained via mechanical disruption of tissue and Shorty-Buffer extraction. Tissue was flash frozen in liquid nitrogen and disrupted using the Qiagen TissueLyser at 30 rps for a period of 1 minute. Five hundred μL of freshly prepared shorty buffer (0.2 M Tris HCl pH 9.0, 0.4 M LiCl, 25 mM EDTA, and 1.0% SDS) was added to each tissue sample, vortexed and centrifuged at max speed (13k rpm) for 5 min. Then 350 μL of supernatant was transferred to a new tube containing 400 μL isopropanol, mixed by inverting and centrifuged at max speed for 10 min. The supernatant was discarded and 1mL of 70% ethanol was added to each sample to wash the DNA pellet. Samples were then centrifuged at max speed for 10 minutes, the supernatant was discarded and the tubes were inverted for 30 minutes. The DNA pellet was resuspended in 100 μL of 10mM Tris-HCl pH 8 and stored short term at 4º C until use.

To verify insertion of transgenic cassettes plants were genotyped for the presence of various portions of the engineered T-DNA (*HYG*^{r} and *AtSUC1*$_{H65K}$) for *CmGAS1pro::*

AtSUC1$_{H65K}$ and *OsTRXhpro*::*AtSUC1*$_{H65K}$ constructs and (*CmGAS1pro* and *GUS*) for *CmGAS1pro*::*GUS* constructs. 20 µL PCR reactions with 2 µL of genomic DNA template and 2 µL of primers at 5 mM concentration were used with 10X EconoTaq Master Mix (Lucigen) under a normal 30-cycle amplification protocol. The primers for genotyping the presence of the *HYG*^{r} gene are HYGpc forward: 5'-TCCACTATCGGCGAGTACTTCTACACA-3', and HYGpc reverse: 5'-CACTGGCAAACTGTGATGGACGAC-3. The primers for genotyping the presence of the *AtSUC1*$_{H65K}$ transgene are AtSUC1H65K forward: 5'-AACTCCTCGGGATCCCTAAG-3', and AtSUC1r reverse: 5'-ACCGGAGCAGCTTGAGAATA – 3'. The primers for genotyping the presence of the *CmGAS1pro* are GASp forward: 5'-TAATGCCCTCCATCTCACCT-3', and GASp reverse: 5'-TGGAGTAAAATTTGGCAAACA-3'. The primers for genotyping the presence of GUS are GUSF forward: 5'-ACCGTTTGTGTGAACAACGA-3', and GUSR reverse: 5'-GGCACAGCACATCAAAGAGA-3'. PCR products were visualized on a 1% gel-agarose containing ethidium bromide.

GUS staining

Plants containing the *CmGAS*pro::*GUS* T-DNA construct were grown to various stages of maturity and sampled for GUS protein expression. Samples were put on ice in the dark and exposed to staining buffer (0.2% Triton X-100, 50 mM $NaHPO_4$ pH 7.2, 2 mM potassium Ferrocyanide, 2 mM potassium Ferricyanide) containing X-Gluc at a final concentration of 2 mM. Samples were then infiltrated under a vacuum for no more than 15 minutes and then incubated at 37ºC until blue staining was seen (14hrs). Staining buffer was removed and samples were submerged in ethanol to remove chlorophyll from tissues. Samples imaged using either a compound light microscope or a dissecting microscope.

Phenotypic analysis

Plants were grown to maturity and several metrics were recorded. Plant height was measured as the length from the planting media to the tip of the tallest leaf. Tiller number was recorded by counting number of stems with true leaves present. Seed yield was determined by harvesting panicles inferior to the panicle neck, drying for 7 days at 45º C and weighing samples. Leaf and stem biomass was measured by cutting plants with panicles removed 5 cm above soil line, drying at 45º C for 7-14 days and weighing samples.

Statistical analyses

Statistics were calculated for growth metrics using a one-way ANOVA and Games-Howell post-hoc multiple comparison test at 95% family-wise confidence level comparing all treatment groups in R using the 'userfriendlyscience' package. Graphical models were generated using the boxplot function in R and Microsoft Powerpoint.

Localization and sequencing of the bi-laterally truncated T-DNA insertion in *mpg1*

Thermal asymmetric interlaced polymerase chain reaction (hiTAIL-PCR) was used to locate the T-DNA insertion site that occurred in the *mpg1* mutant. Nested primers were designed to flank regions of known DNA (in the cassette) and paired with degenerate primers in PCR. The degenerate primers and protocol used followed the protocol by (Liu and Chen, 2007) for rice specifically. The nested primers were then designed specifically from the LB of the T-DNA expression cassette. Nested primers are as follows, primer LB-0A: 5'-GATCTGTCGATCGACAAGCTCGAGTTT, primer LB-1A: 5'-ACGATGGACTCCAGTCCGGCCGGGTT-3', primer LB-2A: 5'-CCAGTACTAAAATCCAGATCCCCCGA-3', primer LB-0B: 5'-GTTCCTATAGGGTTTCGCTCATGTGTT-3', primer LB-1B: 5'-

ACGATGGACTCCAGTCCGGCCCCCCC-3'. The products were visualized using gel-electrophoresis in a 1.0% agarose gel with ethidium bromide. Amplicons were excised and purified using the (Qiagen) Gel Extraction kit. The purified DNA was combined with the LB reverse nested primer and sent for Sanger sequencing. NCBI BLAST of sequences against the rice genome was used to verify the genomic location of the insertion (NCBI IRGSP-1.0 genome build).

The T-DNA insertion present in *mpg1* (a partial T-DNA insertion) was amplified using PCR and sequenced. Genomic DNA extractions were performed on 3-leaf stage leaf tissue of *mpg1* plants and collected using (Qiagen) DNeasy plant mini kit. PCR amplification and sequencing was performed by walking from the known (positively genotyped) insertion sequence and genomic sequence location derived from hiTAIL-PCR to reveal total T-DNA sequence and surrounding genomic DNA sequence.

Haplotyping mutants

To identify *mpg1* plants that were homozygous, heterozygous, or null segregants for the bi-laterally truncated T-DNA insertion, primers were designed in regions directly flanking the site of insertion, as well as primers spanning the integration site into the T-DNA, and used in PCR. The primers flanking the T-DNA insertion were wFLA forward: 5'-GGAAGTTGGAGATGGGAAACA-3', and wFLA reverse: 5'-GGCCTCGTGTGTCAGTAATAA-3'. The primers spanning the genomic region and the T-DNA insertion were wIN forward: 5'-ACACCGGAAGCATAGTCATTT-3', and wIN reverse: 5'-GGTCGCCAACATCTTCTTCT-3'.

RNA-extraction and gene expression analysis

Desired tissue (not more than 100mg) was sampled for RNA extraction and analysis. Tissue was placed in individual 2 mL tubes and flash frozen in liquid nitrogen. Tissue was ground using the TissueLyser (Qiagen) at 30 rps for 1 min, and RNA was extracted using (Qiagen) Plant RNeasy mini-kit. RNA was treated with DNase and purified using the Turbo DNase kit (Invitrogen). cDNA was synthesized from 1 μg of RNA using SuperScript (Invitrogen).

Primers were designed to gauge the expression of *AtSUC1*$_{H65K}$ in transgenic populations, housekeeping primers used in these expression analyses came from elongation factor 1 (EF1) and actin. The primer sequences for EF1 are EF1 forward: 5'-AACATCCGTAATATGTCCGTCATTGC-3', and EF1 reverse: 5'-ATGCAGACCAGCAGAGAAAGC-3'. The primer sequences for actin are, actin forward: 5'-GAGTATGATGAGTCGGGTCCA-3', and actin reverse primer: 5'-ACACCAACAATCCCAAACAGA-3'. Primers designed for expression of *AtSUC1*$_{H65K}$ are AtSUC1f forward: 5'-GACGACGACGAGAAGACCTC-3', and AtSUC1r reverse: 5'-ACCGGAGCAGCTTGAGAATA – 3'.

To determine whether the insertion cassette influenced expression of surrounding genes, primers were designed to several genes directly neighboring the site of the T-DNA insertion. Specific primers used for expression analysis via RT-PCR for the individual candidate gene of *LOC_Os08g41030* (*MPG1*) are 41030 forward: 5'-TCGCCATTGTTCAGCAAGAAGGA-3', and 41030 reverse: 5'-AAGTGCATGACCAAGTACAGA-3'. Housekeeping control primers were designed around actin, more specifically the sequences are, actin forward: 5'-GAGTATGATGAGTCGGGTCCA-3', and actin reverse primer: 5'-ACACCAACAATCCCAAACAGA-3'. PCR was performed in 20 μL reactions using Econo Taq

polymerase (Lucigen), 2 μL of cDNA and desired primers under a normal 30-cycle amplification protocol. PCR products were analyzed by electrophoresis on 1% gel-agarose containing ethidium bromide.

To further assess the expression level of candidate gene *MPG1* ddPCR was performed. 20 μL reactions containing 10 ng of cDNA were used in conjunction with EvaGreen fluorescence dye protocol (BioRad). Analysis was performed using (QuantaSoft) ddPCR software.

Experimental populations

Plants undergoing phenotypic assessment at the T_1 and T_2 generations were grown in 3.5" pots in non-optimal soil and fertilizer media and genotyped for T-DNA insertion. Plants were genotyped at the 3-leaf stage (leaf tissue). After genotyping, plants were transferred to 1.0 Ga pots and organized in random fashion including a wild-type brush-border to account for any edge effect. Plants were grown to maturity. Treatment groups WT (wild-type) n = 13, tcWT (plants that underwent tissue culture mediated Agrobacterium transformation that resulted in unsuccessful integration of T-DNA expression cassette) n = 5, *GS* (individual lines of *CmGAS1pro*:: $AtSUC1_{H65K}$ transformants labeled as independent insertion events, lines – 8, 12, 14, 21, 23) n = 22, 18, 16, 18, 23 respectively, and *mpg1* (a plant containing a partial T-DNA expression cassette insertion, not containing $AtSUC1_{H65K}$) n = 5.

T_3 segregating plants were utilized for expression analysis of *MPG1.* Plants were grown in 4.0" in pots in optimal soil and fertilizer media, where various tissues were collected for gene expression analysis.

Results

Cucumis melo GALACTINOL SYNTHASE 1 promoter shows companion cell specificity in *Oryza sativa*

Although the promoter for galactinol synthase has been shown to be companion cell specific in certain dicotyledonous species such as *Arabidopsis thaliana* and *Nicotiana tobacum* (Haritatos et al., 2000a) it was important to verify that it would behave similarly in rice, a monocotyledonous species. Here, rice calli were transformed with a T-DNA expression cassette containing the *Cucumis melo GALACTINOL SYNTHASE I* (*CmGAS1*) promoter driving -glucuronidase (GUS). Numerous T_0 lines were generated and several T_1 and T_2 plants from independent insertion events were grown to varying levels of maturity and stained for GUS activity (Figure 2.1). GUS was visualized in the vasculature of the plants, and closer inspection of leaf cross sections revealed that GUS was primarily located in the companion cells of vascular bundles. In vascular tissue, GUS appeared to be relatively uniform across leaf and stem tissues and also appeared in the lemma of developing spikelets and in panicle branches. GUS was intermittently seen in root tissues, present in lateral roots and sporadically through the primary root cortex.

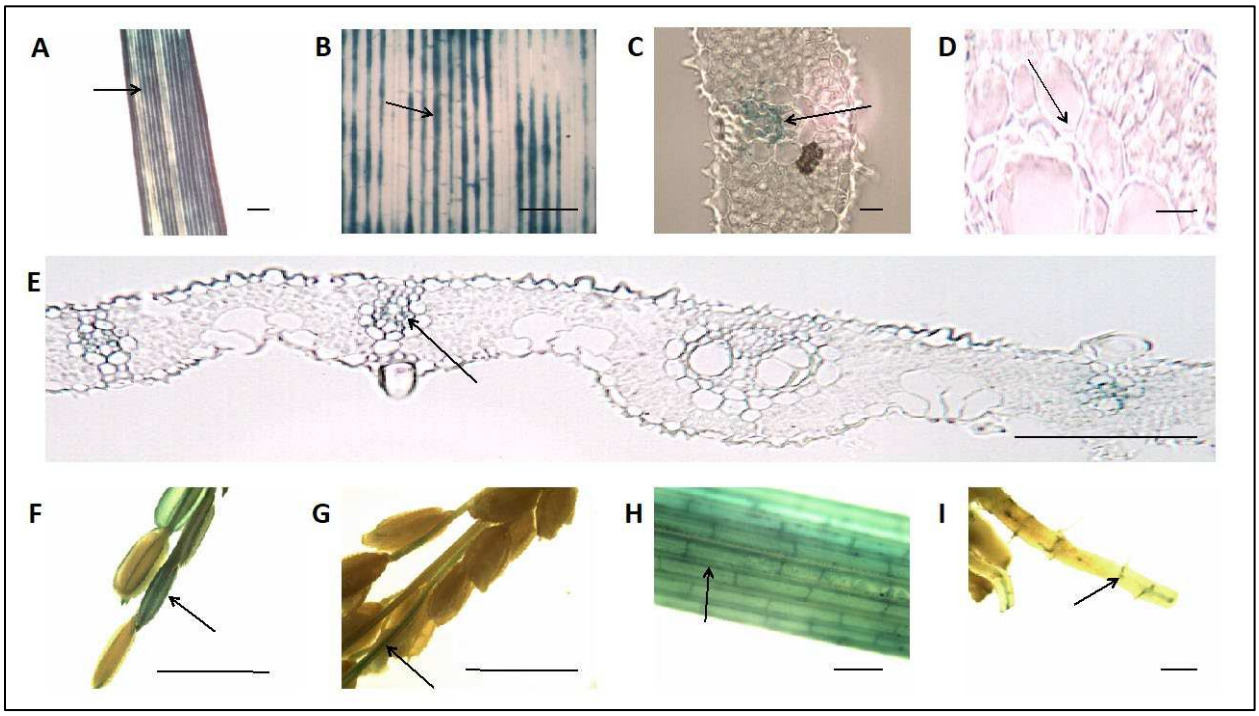


Figure 2.1: Localization of GUS in *CmGAS1pro*::*GUS* transgenic plants.
Images of *CmGAS1pro*::*GUS* were taken after performing GUS staining on various plant tissues. (A) Leaf tissue showing GUS in vascular tissue, scale = 1.0 mm. (B) Close-up of A, scale = 1.0 mm. (C) Cross section of leaf tissue showing GUS localized to companion cells in vascular bundles, scale = 10 μm. (D) Close-up of C showing GUS localized to companion cells proximal to phloem in vascular bundle (arrow), scale = 10 μm. (E) Broad image of leaf cross-section with GUS present in vascular tissues (arrow), scale = 100 μm. (F) Developing panicle showing GUS in panicle branches and developing spikelets, scale = 1.0 cm. (G) Mature panicle showing GUS in panicle branches, scale = 1.0 cm. (H) Mature stem tissue showing GUS in vascular tissue, scale = 1.0 mm. (I) Root tissue showing intermittent GUS staining in lateral roots and primary root cortex, scale = 1.0 mm.

Transgenic *Oryza sativa* containing a hyperactive proton-sucrose symporter driven by the *CmGAS1pro* have less biomass and seed yield compared to wild-type plants

To test our hypothesis on whether or not modulating source-sink interaction through increasing phloem loading of sucrose would impact biomass accumulation, transgenic plants were generated using a T-DNA expression cassette designed to drive companion cell-specific expression of a hyperactive proton-sucrose symporter (Figure 2.2). Several independent insertion lines of T_1 transgenic plants containing the *CmGAS1pro*::*AtSUC1*$_{H65K}$ T-DNA expression cassette were grown to maturity and biomass metrics were taken. All of the *CmGAS1pro*::

AtSUC1$_{H65K}$ lines exhibited lower total dry weight, height, tiller number and seed yield compared to wild-type plants (Figure 2.3).

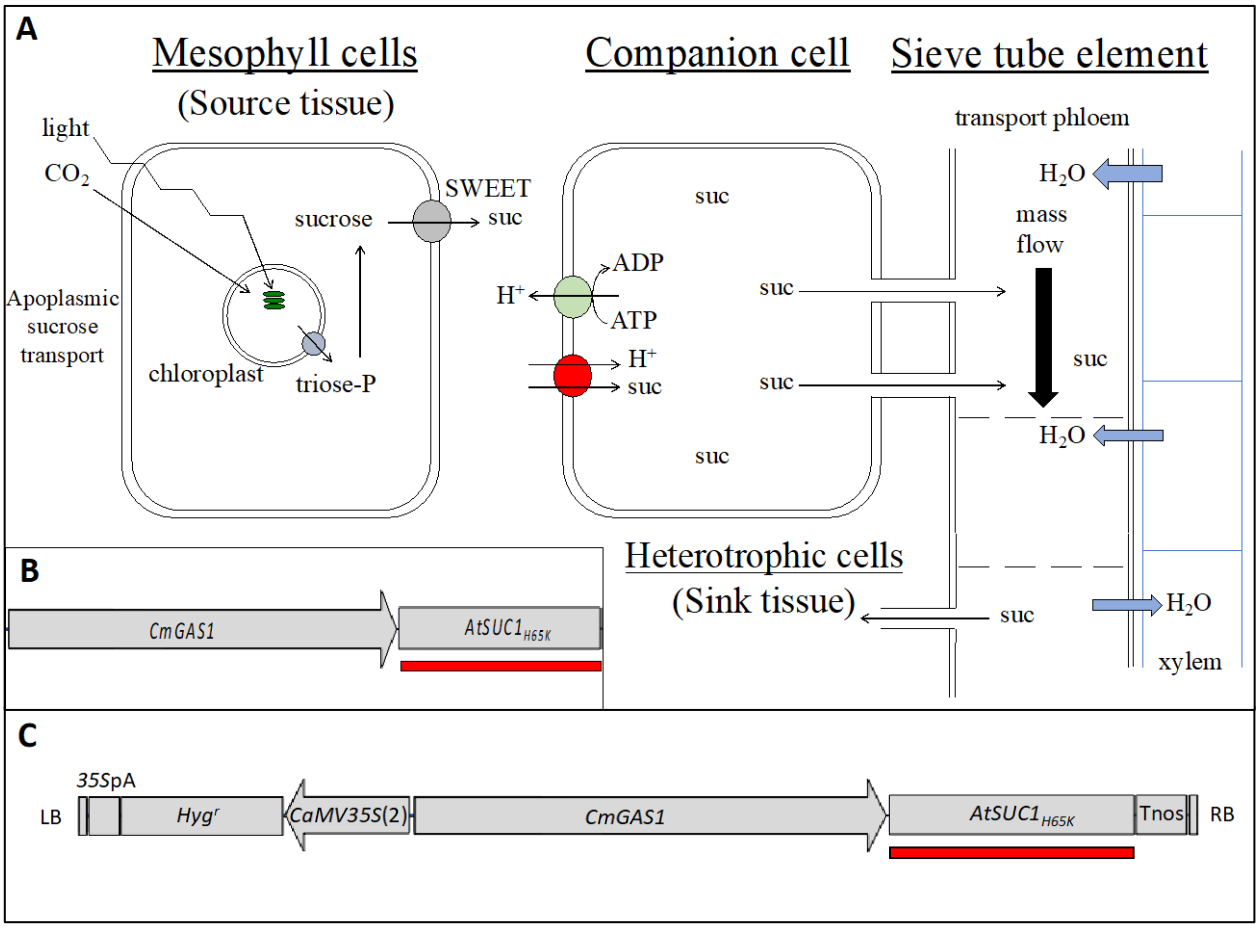


Figure 2.2: Schematic of apoplasmic sucrose transport in plant tissues and proposed design for enhancing sucrose loading into phloem of *Oryza sativa*.

(A) Visual diagram of the source-to-sink synthesis and transport of sucrose from mesophyll cells to heterotrophic tissues through the use of the proton-sucrose symporter in the apoplasmic sucrose transport pathway. (B) Display of a proposed gene construct to modify sucrose loading by utilizing a gene coding for a hyper-active sucrose symporter (*AtSUC1*$_{H65K}$) from *Arabidopsis thaliana* being driven by the companion-cell specific promoter, from the gene coding for *GALACTINOL SYNTHASE 1* (*CmGAS1*), from *Cucumis melo*. (C) Engineered T-DNA insertion cassette to test the proposed gene construct to be generated in binary vector pCambia1390. Elements of the T-DNA insertion construct contain sequences pertaining to the following: LB – left border, *35S*pA – *CaMV35S* poly-A signal, *Hyg*^{r} – the gene coding for hygromicin resistance, *CaMV35S*(2) – duplicated *CaMV35S* promoter sequence, *CmGAS1* – promoter sequence from *C. melo*, *AtSUC1*$_{H65K}$ – the gene coding for the hyperactive sucrose symporter from *A. thaliana*, Tnos – nos terminator, RB- right border.

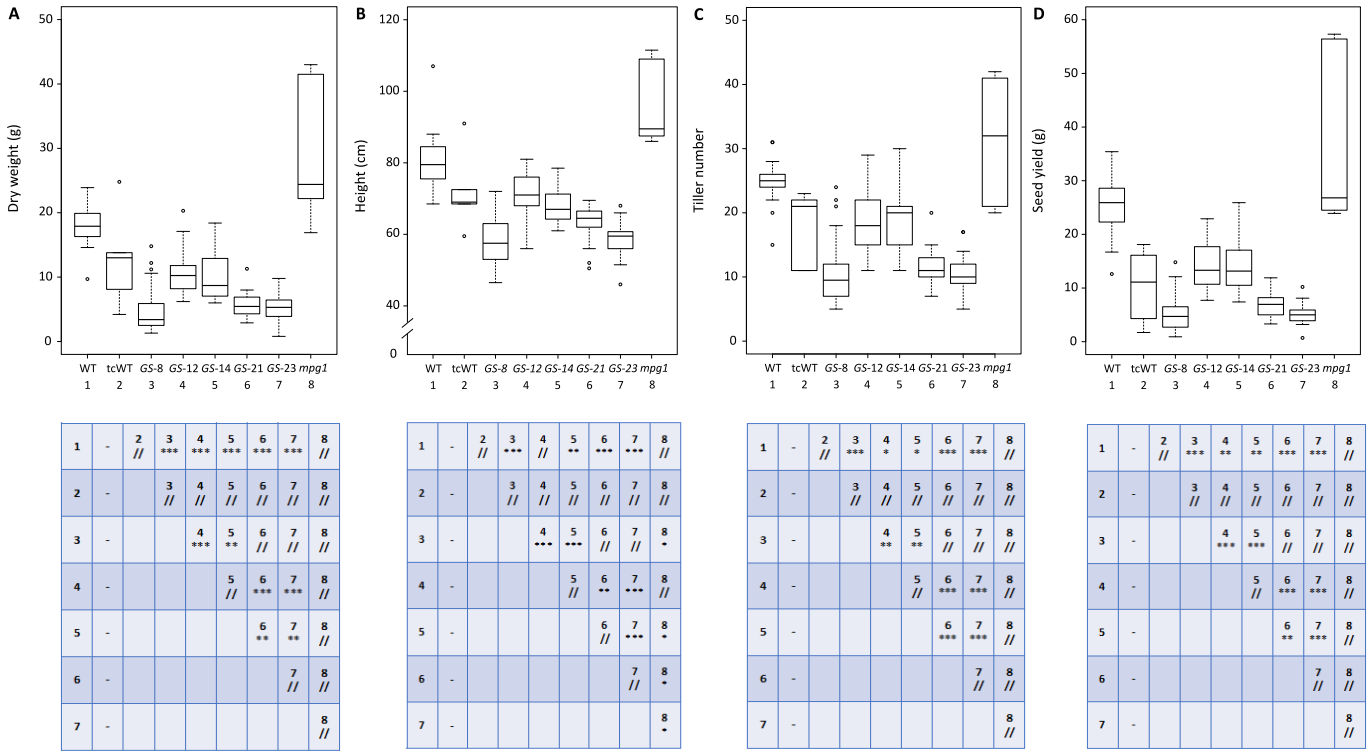


Figure 2.3: Phenotypic analysis of T₁ *CmGAS1pro*::*AtSUC1*$_{H65K}$ transformants and partial T-DNA insertion line.
(A) dry weight, (B) height, (C) tiller number, (D) seed yield. Treatment groups WT (wild-type) n = 13, tcWT (plants that underwent tissue culture mediated *Agrobacterium* transformation but did not contain T-DNA expression cassette) n = 5, *GS* (individual lines of *CmGAS1pro*::*AtSUC1*$_{H65K}$ transformants labeled as independent insertion events, lines – 8, 12, 14, 21, 23) n = 22, 18, 16, 18, 23 respectively, and *mpg1* (*makes-plants-gigantic-1*, a plant containing a partial T-DNA expression cassette insertion, not containing *AtSUC1*$_{H65K}$) n = 5. Analysis was conducted using a one-way ANOVA and Games Howell post-hoc multiple comparison test at 95% family-wise confidence level. Numeric value underneath treatment group abbreviation signifies its designation for comparative analysis in the tables (below). '//' indicates no change, '*' indicates p<0.05, '**' indicates p<0.01, '***' indicates p<0.001 assessing all treatment groups.

Transgenic *Oryza sativa* containing a hyperactive proton-sucrose symporter being driven by companion cell-specific promoters have intermittent expression of transgene

Semi-quantitative RT-PCR was performed to assess the expression of the transgene *AtSUC1*$_{H65K}$ to ensure that our designed T-DNA expression cassettes were functioning as intended in the T_1 generation grown for biomass observations. Leaf tissue from 3- to 4-leaf stage plants was sampled for expression analysis. As expected, transgenic *CmGAS1pro*::*AtSUC1*$_{H65K}$ plants showed expression of the transgene *AtSUC1*$_{H65K}$ while wild-type plants did not (Figure 2.4). We developed an alternate companion cell-specific T-DNA expression construct utilizing the companion cell-specific promoter from the gene coding for *THEIRODOXIN H* in *Oryza sativa* (*OsTRXhpro*::*AtSUC1*$_{H65K}$) (Ishiwatari et al., 2000, Fukuda et al., 2005, Ishiwatari et al., 1998). These plants were also grown in a growth trial experiment alongside *CmGAS1pro*::*AtSUC1*$_{H65K}$ plants. Both transgenic plant lines still resulted in smaller plants relative to wild-type (data not shown). Expression analysis from this growth trial also showed different rates of expression, not untypical of transgenic plants, of the transgene *AtSUC1*$_{H65K}$ between and within individual lines of both *CmGAS1pro* and *OsTRXhpro* (Figure 2.4B).

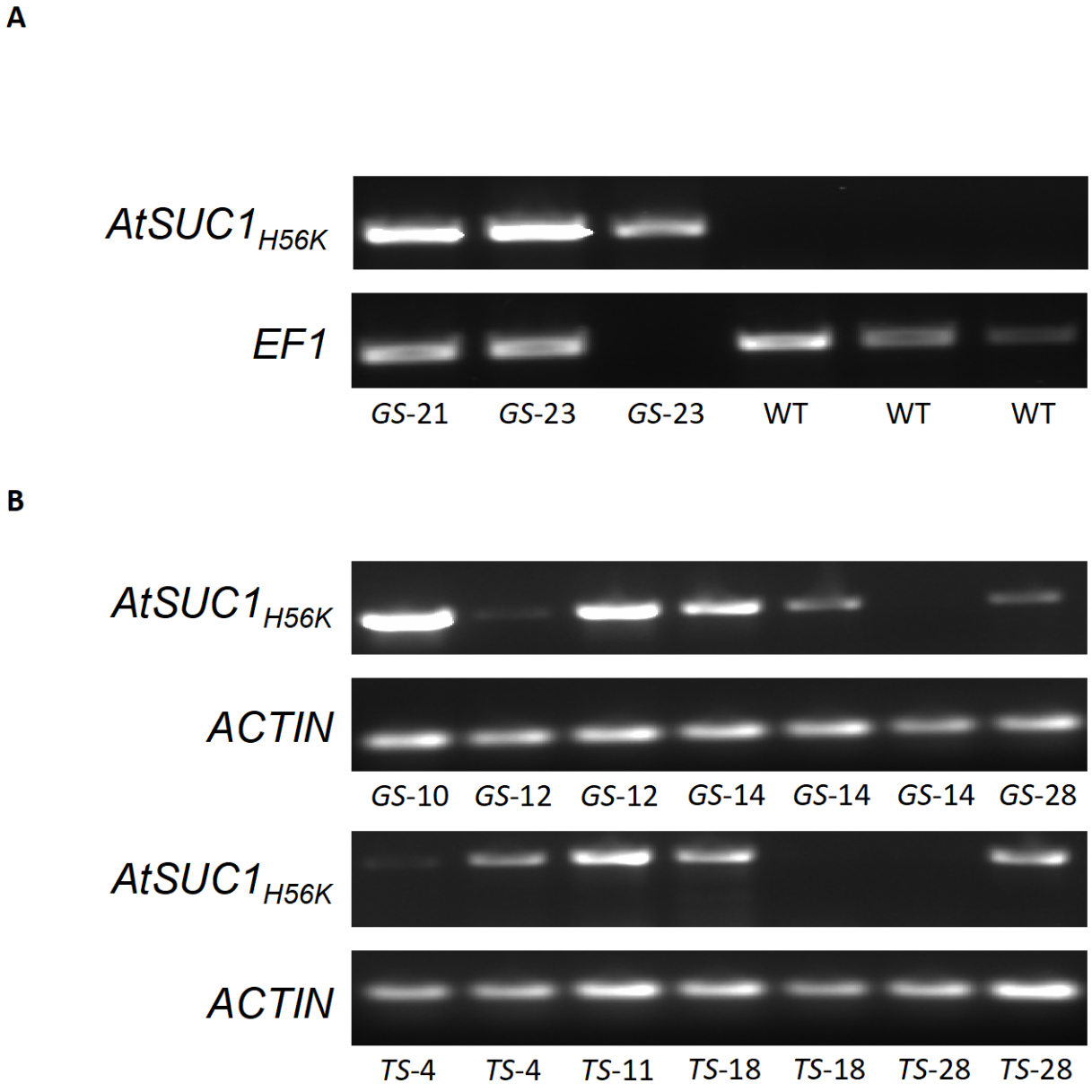


Figure 2.4: *AtSUC1$_{H65K}$* expression in transgenic Oryza sativa.
(A) RT-PCR gel-electrophoresis image from leaf tissue at 3- to 4-leaf stage of the representative T_1 growth assessment of *CmGAS1pro*::*AtSUC1$_{H65K}$* transgenics. Positively genotyped transgenic plants show expression of *AtSUC1$_{H65K}$*, while wild-type plants do not. Elongation factor 1 (EF1) was used as a control, and each lane represents an individual plant. *GS-# = CmGAS1pro*:: *AtSUC1$_{H65K}$* – individual insertion event (line). (B) RT-PCR gel-electrophoresis image from leaf tissue at 3- to 4-leaf stage T_1 *OsTRXhpro*::*AtSUC1$_{H65K}$* and *CmGAS1pro*:: *AtSUC1$_{H65K}$* as well as T_2 *CmGAS1pro*::*AtSUC1$_{H65K}$* transgenics. Positively genotyped transgenic plants show intermittent expression of *AtSUC1$_{H65K}$*. Actin was used as a control, and each lane represents an individual plant. *GS-(#) = CmGAS1pro*::*AtSUC1$_{H65K}$* – individual insertion event (line), *TS-(#) = OsTRXhpro*::*AtSUC1$_{H65K}$* – individual insertion event (line).

Plant containing a partial T-DNA insertion exhibits increased biomass and seed yield compared to wild-type

Initial assessment of T_0 generation transgenic plants containing either the *CmGAS1*pro::*AtSUC1*$_{H65K}$ or *OsTRXhpro*::*AtSUC1*$_{H65K}$ revealed a single plant that was visibly larger in size than all other plants (a representative photograph from the T_3 generation of this line is shown in Figure 2.5). After initial genotyping, this plant (originally transformed with the *CmGAS1pro*::*AtSUC1*$_{H65K}$ T-DNA expression cassette) was found to contain the gene coding for hygromycin resistance (*HYG*^{r}) but not the transgene *AtSUC1*$_{H65K}$ suggesting that a truncated version of the intended expression cassette was inserted. Because this plant was larger in size we continued to grow it alongside plants containing the full T-DNA expression cassette to further monitor its growth and behavior. In the T_1 generation grown to characterize the transgenic plants expressing *AtSUC1*$_{H65K}$ in companion cells, the partial T-DNA insertion plants exhibited greater biomass and seed yield compared to other treatments (Figure 2.3, Figure 2.5). Although not significant, likely due to sample size, *mpg1* plants were larger and produced more seed yield than wild-type. Because of the greater biomass observed in these partial T-DNA insertion plants, the line was termed *makes-plants-gigantic-1* (*mpg1*).

Figure 2.5: Comparison of *mpg1* to wild-type plant.
(Left) wild-type plant (right) *mpg1* T_3 plant, scale = 1.0 m.

mpg1 plants contain a bi-laterally truncated T-DNA expression cassette integrated in an intergenic region on chromosome 8

The *mpg1* line was further investigated to identify why it exhibited enhanced biomass and seed yield compared to wild-type plants. Primer walking from the known portion of the T-DNA insertion (gene coding for hygromycin resistance) revealed that the insertion is bi-laterally truncated (Figure 2.6). Forty base-pairs comprising the end left border are absent, as well as sequence comprising the end right border through *CmGAS1pro* (first 1037 bp) sequence are absent as well. The total size of the insertion is 3190 base-pairs in length.

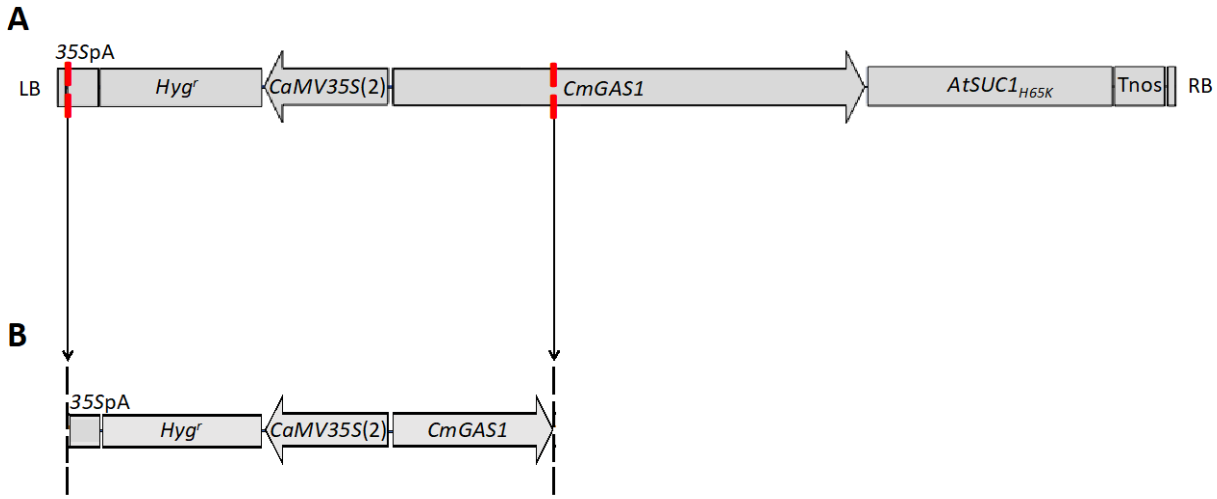


Figure 2.6: Schematic representing the extent of bi-lateral truncation in the T-DNA expression cassette integrated in *mpg1*.
(A) Visual representation of the complete T-DNA insertion cassette designed to generate intended transformants. LB = left border, 35SpA = *CaMV35S* poly A sequence, *Hyg*^{r} = gene coding for hygromycin resistance, *CaMV35S(2)* = duplicated *CaMV35S* promoter sequence, *CmGAS1* = *CmGAS1* promoter sequence, *AtSUC1*$_{H65K}$ = gene coding for hyperactive sucrose symporter, Tnos = nos terminator, RB = right border. (B) The remnants of the intended T-DNA expression cassette present in *mpg1*. The insertion is bi-laterally truncated. 40 base-pairs are absent from the left border, and all sequence from the right border through to the first 1037 base-pairs of the *CmGAS1* promoter sequence are absent compared to the complete T-DNA sequence.

We hypothesized that the integration of the truncated T-DNA was somehow responsible for generating the phenotype of increased biomass and seed yield. Thermal asymmetric interlaced PCR (TAIL-PCR) revealed that the insertion event occurred on chromosome 8 between *Os08g41010* and *Os08g41020* (NCBI IRGSP-1.0 genome build, just after locus Chr8:25,942,615), not disrupting any known or annotated functional elements (Figure 2.7). Additional sequencing at the location of the insertion revealed a 20 base-pair genomic DNA (gDNA) deletion occurred (deleting Chr8:25,942,616-25,942,635) at the site of integration.

Because no coding sequences were disrupted, we predicted that the T-DNA insertion might be altering expression of a gene or genes neighboring the site of integration.

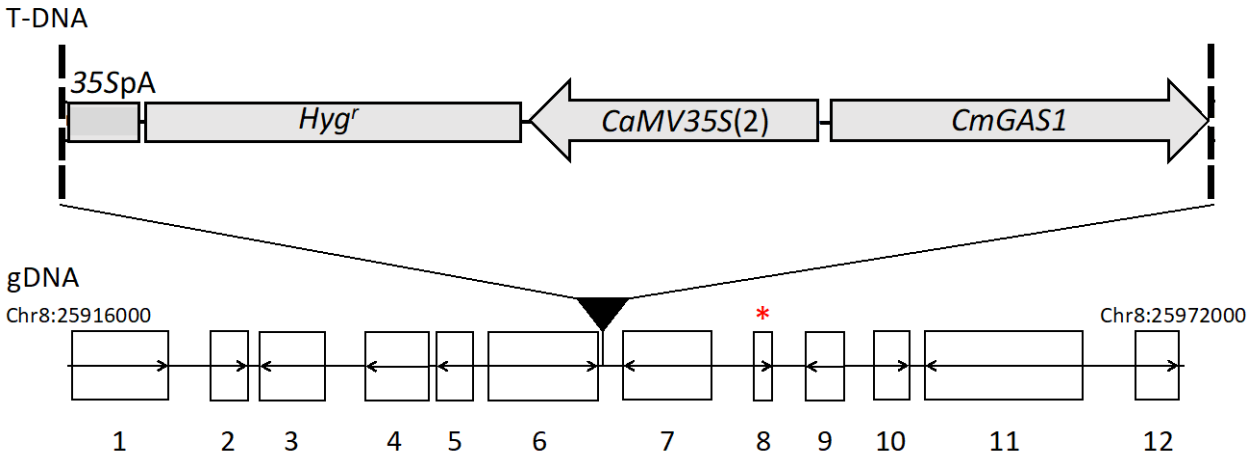


Figure 2.7: Genomic localization of the bi-laterally truncated T-DNA insertion in *mpg1* and results of neighboring differential gene expression between wild-type and *mpg1*.
A visual schematic showing the location of the bi-laterally truncated T-DNA insertion present in *mpg1* (above) and its location of integration in the rice genome, gDNA (below). HI-TAIL-PCR revealed that the insertion event occurred on chromosome 8 between Os08g41010 and Os08g41020. Further sequencing of the region concluded that the T-DNA integrated just after loci Chr8:25,942,615 (NCBI IRGSP-1.0 genome build). At the location of the insertion a 20 base-pair genomic genomic DNA (gDNA) deletion occurred (deleting Chr8:25,942,616-25,942,635). The total size of the insertion is 3190 base-pairs in length. The genomic region displayed spans from Chr8:25916000-Chr8:2597200. Expression analysis was performed via semi-quantitative RT-PCR on neighboring genes flanking the insertion site for both wild-type and *mpg1* plants. The red asterisk represents the only gene experiencing notable differential expression, LOC_Os08g41030 (8). Genes listed: LOC_Os08g40960 – retrotransposon protein (1), LOC_Os08g40970 – retrotransposon protein (2), LOC_Os08g40980 – retrotransposon protein (3), LOC_Os08g40990 – receptor-like protein kinase 1 (4), LOC_Os08g41000 – extracellular ligand-gated ion channel (5), LOC_Os08g41010 – zinc finger family protein (6), LOC_Os08g41020 – retrotransposon protein (7), LOC_Os08g41030 – AP2 domain containing protein (8), LOC_Os08g41040 – expressed protein (9), LOC_Os08g41054 – hypothetical protein (10), LOC_Os08g41070 – retrotransposon protein (11), LOC_Os08g41080 – expressed protein (12).

The T-DNA insertion in *mpg1* results in increased expression of a neighboring transcription factor

Expression of genes neighboring the T-DNA insertion in *mpg1* plants was compared to expression in wild-type plants in 4-leaf stage leaf tissue. Several genes directly upstream and downstream of the bi-laterally truncated T-DNA insertion site were investigated. Semi-quantitative RT-PCR showed that only one gene had a noticeably different pattern of expression in *mpg1* compared to wild-type plants, *Os08g41030* (Figure 2.7, red star). *Os08g41030* encodes a transcription factor belonging to the APETALA 2/Ethylene-Responsive Element-Binding Protein (AP2/EREBP) transcription factor superfamily. In *mpg1* plants *Os08g41030* has much higher expression than in wild-type plants, which had little to no expression (Figure 2.8). In segregating *mpg1* lines *Os08g41030* expression was high in plants heterozygous or homozygous for the T-DNA insertion, while null segregants and wild-type plants lacked expression (Figure 2.8). This pattern of expression had a 100% correlation with the presence of the T-DNA insertion and enhanced biomass. This is further discussed in Chapter 3, where comprehensive phenotyping is discussed in more detail. Because of the differential regulation of this gene in the mutant *mpg1,* we termed LOC_Os08g41030 (*MPG1*). To more accurately compare expression levels of *MPG1* between *mpg1* and wild-type plants, ddPCR was performed on homozygous *mpg1* plants and wild-type null segregants. *mpg1* plants homozygous for the T-DNA insertion showed a 190-fold increase in expression of *MPG1* compared to wild-type null segregants, which still reported little to no expression (Figure 2.8).

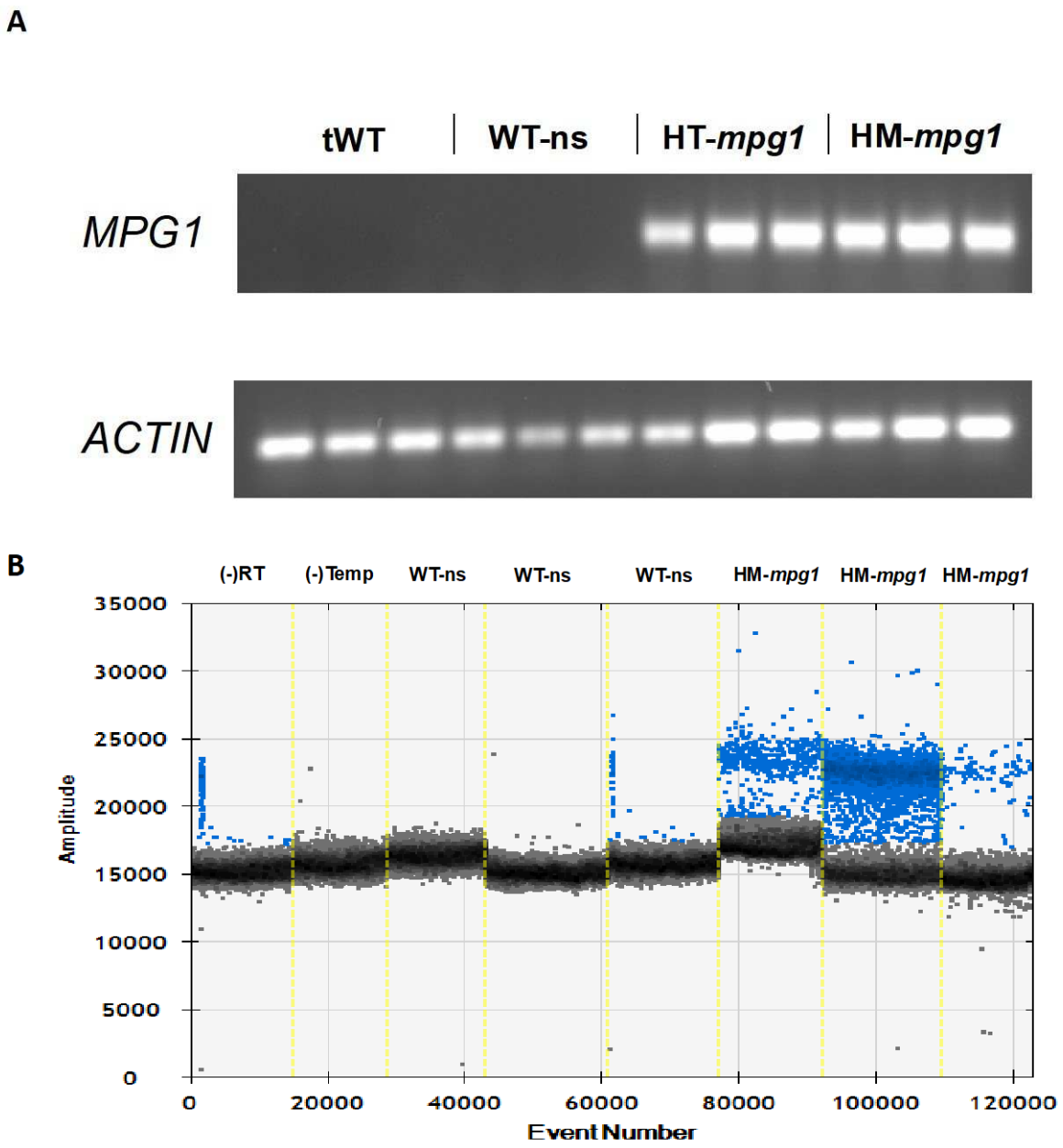


Figure 2.8: Expression analysis of AP2-EREBP (*MPG1*).
(A) RT-PCR gel-electrophoresis image from leaf tissue at 21 days post-planting. There is greater expression of (*MPG1*) in plants that are heterozygous and homozygous for the truncated T-DNA insertion present in *mpg1* compared to wild-type plants which show no expression. Each lane represents a single biological replicate, and the gene coding for actin was used as a control. (B) dd-PCR report from leaf tissue at 42 days post-planting showing greater expression of (*MPG1*) in plants that are homozygous for the truncated T-DNA insertion present in *mpg1* compared to wild-type segregant plants which show little to no expression. Homozygous plants show a 190-fold increase in *MPG1* expression compared to WT-ns. -RT represents a negative reverse transcriptase control and -Temp represents a negative template control. Each lane represents a single biological replicate.

Discussion

Transgenic *Oryza sativa* plants engineered to enhance phloem loading of sucrose in the companion cell resulted in stunted plants

Taking the approach to increase plant productivity by modulation of sucrose loading in phloem, in contrast to expectations, resulted in smaller plants. Rather than generating plants with increased sink organs, plants expressing a hyperactive sucrose symporter in companion cells were stunted and produced reduced yield compared to wild-type plants (Figure 2.3). Using a GUS reporter system, we validated that the dicot companion cell specific promoter (*CmGAS1pro*) used in the generation of the T-DNA expression cassette was targeted to companion cells in rice, a monocot (Figure 2.1). Further we utilized a separate companion cell specific promoter from rice (*OsTRXhpro*) to drive expression of our transgene and found similar phenotypic results. Therefore, it is unlikely that mis-targeting of the $AtSUC1_{H65K}$ transgene caused the reductions in growth and yield.

We found that transgene expression in leaves of both constructs (*CmGAS1pro::*$AtSUC1_{H65K}$ and *OsTRXhpro::*$AtSUC1_{H65K}$) varied. Expression between and within individuals lines was inconsistent, with some lines showing little to no expression. The observed variation in expression suggests there might be some factor limiting the transcriptional message of this transgene. Assuming productivity to be relatively similar between independent insertion events, tissues for transcription profiling were all taken at the same time of development (4-leaf stage), at the same time of day (mid-day), with the same tissue (fully expanded leaf). A previous mechanism that has been shown to limit the transcript abundance for *BvSUC1* in *Beta vulgaris* was the concentration of sucrose in source tissues (Vaughn et al., 2002). Perhaps initial excess loading of sucrose resulted in a period of reduced phloem loading

through transient sucrose accumulation and lack of export from source tissue resulting in decreased photosynthetic activity and transcription of $AtSUC1_{H65K}$.

However, these transgenic plants were utilizing expression cassettes driving a modified non-native *SUC1* under control of companion cell specific promoters not specific to *SUC1*. The native promoter elements driving companion cell-specific phloem loading through *SUC1* has been shown to be the site for negatively regulated transcription under presence of heightened levels of sucrose in the source tissue (Chiou and Bush, 1998, Vaughn et al., 2002). It isn't clear if a regulatory mechanism is impacting the activity of the companion cell-specific galactinol synthase promoter and thioredoxin h promoter elements driving expression of the proton-sucrose symporter involved with phloem loading. The *CmGAS1* gene, like other galactinol synthase genes, is involved with the formation of raffinose family oligosaccharides which provide for various plant functions (Sengupta et al., 2015). Examination of this gene class has revealed that they are associated with several functions including: protection of embryos from maturation associated desiccation (Downie et al., 2003), transport and storage of carbon - including signal transduction (Stevenson et al., 2000, Xue et al., 2007), membrane trafficking (Thole and Nielsen, 2008), mRNA export (Okada and Ye, 2009), response to pathogen attack (Couee et al., 2006, Kim et al., 2008), and response to abiotic stressors (Zuther et al., 2004, Zuther et al., 2012, Peters and Keller, 2009). Specifically, *CmGAS1*'s expression has been localized to companion cells, not changing expression levels under induced stress or diurnal cycles, suggesting its function might pertain specifically to phloem loading of RFO's as translocation sugars (Volk et al., 2003). Further analysis of the *CmGAS1* promoter has revealed similar *cis*-elements to that of the *AtSUC2* promoter – transcriptional activity has been shown to be regulated by two putative binding sites, one for homeodomain leucine zipper and another for a DNA-binding-with-one-

finger transcription factor sites (Schneidereit et al., 2008, Ayre et al., 2003, Tsuwamoto and Harada, 2010). Again, decreased transcriptional activity of native *BvSUT1* was seen under increased levels of sucrose (Vaughn et al., 2002). Additionally, an *Arabidopsis* mutant, *pho3*, which accumulates greater levels of sucrose and other carbohydrates, resulted in a 50% reduction in transcript abundance of *AtSUC2* compared to wild-type (Lloyd and Zakhleniuk, 2004). Similar observations have also been visualized in assessment of transcript responses to supplemental sucrose exposure (Osuna et al., 2007). These together might suggest that because of the similar sequence elements present in these promoters, that *CmGAS1pro* might be transcriptionally limited by similar constraints of the native coupled regulation of sucrose concentration in the source pool and phloem loading regulation by proton-sucrose symporter genes via sucrose signaling. Alternatively, the other selected companion cell-specific promoter used in our study was from thioredoxin h in rice. Assessment of this gene uncovered that the promoter from rice thioredoxin h was companion cell-specific (Ishiwatari et al., 2000, Ishiwatari et al., 1998). Characterization of this gene has revealed that, natively, it is ubiquitously expressed and further induced by salt exposure and abscisic acid treatments, and is predicted to function as a means to regulate the redox state of the apoplast and influence plant development and response to stressors (Zhang et al., 2011). The nature of this promoters transcriptional activity has not been shown to be linked to sucrose abundance, so it should be continuously functioning.

Because both the *CmGAS1pro*::*AtSUC1*$_{H65K}$ and *OsTRXhpro*::*AtSUC1*$_{H65K}$ transgenic plants showed transient and inconsistent abundance of *AtSUC1*$_{H65K}$ message, then perhaps there is an alternative regulatory mechanism resulting in reduced message abundance. It would be interesting to gain a better understanding of other companion cell-specific promoters, their regulatory and expression patterns to further assess the capabilities of enhanced phloem loading

on plant productivity. In particular through the use of continually high expressive companion cell-specific promoters.

Sucrose transport genes have been shown to also be regulated post-transcriptionally (Kuhn and Grof, 2010). In *Arabidopsis* there are nine known sucrose transporters, seven of them including *AtSUC1* have been shown to be targeted by microRNAs (Lu et al., 2005). Perhaps the reduced message of $AtSUC1_{H65K}$ was subject to post-transcriptional silencing. This regulatory mechanism might occur from modulating the presence of additional sucrose transport proteins, or perhaps abundance of sucrose in source tissues as a result of the expression of the transgene. Regardless of the reason behind the varying levels of expression within independent transgenic lines, plants that exhibited expression of the $AtSUC1_{H65K}$ still resulted in smaller plants than wild-type.

Since performing this study, other researchers have investigated the effect of proposed enhancement of phloem loading of sucrose through manipulation of sucrose transporters in companion cells. As discussed, sucrose transporter expression through the control of native promoter elements is highly regulated (Vaughn et al., 2002, Chiou and Bush, 1998). Therefore it has been proposed that non-native or exotic transcriptional elements could persist in enhancing phloem loading of sucrose under higher concentrations of sucrose in source tissues. One particular study used a companion cell-specific promoter not subject to feedback inhibition by sucrose (Dasgupta et al., 2014). The promoter from *Commelina yellow mottle virus* (*CoYMVpro*) (Medberry et al., 1992, Matsuda et al., 2002) was used to drive expression of several different *SUT* genes in *Arabidopsis* plants homozygous for the *Atsuc2-4* mutant, which are stunted due to the inability to load and transport sucrose (Dasgupta et al., 2014, Srivastava et al., 2008). *AtSUC1* and *AtSUC2* genes were able to rescue the phenotype of the mutant. Assessment of

CoYMVpro::*SUT* in wild-type backgrounds showed enhanced loading and transport of sucrose, yet resulted in stunted plants (Dasgupta et al., 2014). These stunted plants showed increased expression of phosphate starvation-induced (PSI) genes, and upon application of supplemental phosphate to the growth medium were no longer stunted. The growth results and perceived phosphate deficiency mimic closely that of what is observed in plants constitutively overexpressing *AtSUC2* in *Arabidopsis* (Lei et al., 2011). It is important to note that even after the addition of the supplemental phosphate these plants were unable to exceed the growth and yield of wild-type plants. This suggests that these plants perceived a phosphate deficiency. The perception on phosphate limitations could be the cause of unbalanced carbon-to-phosphate ratio giving way to insufficient biochemical and signaling cascades resulting in the inability to use carbon in sink tissues for growth and development (Dasgupta et al., 2014).

It is likely that stimulating plant growth by means of enhanced photosynthate partitioning is more complex than initially predicted. Assimilate partitioning of sucrose in higher plants is a complex system with numerous regulators. It is likely that other nutrients, or abundance of available elements participate in multifaceted physiological crosstalk influencing productivity and functionality, as has been in other studies (Kellermeier et al., 2014). Further evaluation of the affects elicited by enhancing photosynthate partitioning will be necessary to successfully engineer plants capable of greater plant productivity.

***mpg1* is a T-DNA mutant with enhanced plant productivity compared to wild-type plants**

During the generation of transgenic plants used to investigate the manipulation of phloem loading by expressing a hyperactive sucrose symporter in companion cells, a mutant was found. This mutant, named *makes-plants-gigantic-1* (*mpg1*) was discovered because it was larger than any of the other transgenic plants generated or wild-type controls. Initial investigation of this

plant revealed, through genotyping, that it contained a portion of the originally designed *CmGAS1pro::AtSUC1$_{H65K}$* T-DNA expression cassette.

Sequencing the remnant contents of the T-DNA expression cassette showed that it was bi-laterally truncated and missing the transgene, *AtSUC1$_{H65K}$*. The insertion contained the gene coding for hygromycin resistance driven by a duplicated *CaMV35S* promoter, and the first third of the *CmGAS1* promoter sequence (companion cell specific promoter intended to drive expression of *AtSUC1$_{H65K}$* in the complete T-DNA expression construct) (Figure 2.6). We found that the insertion integrated into an inter-genic region on chromosome 8 (Figure 2.7).

Because the insertion didn't disrupt any known or annotated functional elements, we hypothesized that the presence of the insertion resulted in a mutagenic event altering the expression of a gene or genes neighboring the site of insertion. We found altered expression of a single neighboring gene, *LOC_Os08g41030*, an APETALA 2/ Ethylene-Responsive Element-Binding Protein (AP2/EREBP) transcription factor from the ERF subfamily, which we termed *MPG1*. In the mutant *mpg1,* both plants that are heterozygous and homozygous for the T-DNA insertion showed markedly higher expression of *MPG1* than in true wild-type plants as well as wild-type null segregants that arose as progeny from segregating *mpg1* plants. This suggests that the presence of the T-DNA is a dominant mutation. Further backcrosses will aid in validation of this speculation.

MPG1 encodes a transcription factor belonging to the (AP2/EREBP) gene superfamily. This gene family is comprised of transcription factors shown to be involved with regulation of growth and development, as well as response to stresses (Nakano et al., 2006). Interestingly, a phylogenetic assessment investigating this particular AP2/EREBP gene with others from rice and *Arabidopsis* showed limited homology to other genes in this family suggesting that it is exclusive

to rice (Nakano et al., 2006). The rice genome contains >1,300 transcriptional regulators and over 40% of those are believed to be from plant-specific families (Rice Full-Length c et al., 2003). Because *MPG1* is specific to rice the question regarding its functionality in other plant species remains, however if mechanisms of action regarding this gene remain well conserved, translational application could exist.

The native functionality of *MPG1* has not been well characterized, however gene expression databases and previous gene expression studies have provided a degree of information on its native expression. MSU Rice Genome Annotation Database's RNA-seq coverage data shows expression in numerous tissue types (flowering, embryo, shoot, and leaf tissues) (Ouyang et al., 2007). Specifically, expression of *MPG1* was reported to be within pre- and post- emergence inflorescence, seed 5 days post-pollination, embryo 25 days post-pollination, and 20 days post-planting shoot and leaf tissues. Of these tissues, greater levels of expression were visualized in flowering tissues over vegetative tissues. *MPG1*'s expression has also been seen in response to stress. Increased transcription of this gene has been reported, via micro array and RT-PCR analysis under exogeneous salicylic acid, submergence, laid-down submergence, and cold (Sharoni et al., 2011).

A more recent study has demonstrated a native function of *MPG1*. *MPG1*, also known as *OsERF#115*, was found to play a significant role in endosperm development showing expression in aleurone tissue of developing seeds (Xu et al., 2016a). It acts as the DNA recognition protein necessary to operate a transcriptional complex involving several other transcription factors (nuclear factor transcription factors) regulating endosperm development and grain filling. RNA-seq analysis investigating the global effects elicited by the suppression of one of the proteins in this transcriptional complex (knockdown of *OsNF-YB1* by RNAi) generated a list of co-regulated

(down-regulated) genes and associated ontological enrichments. The co-regulated genes were enriched for transport, ATP synthesis, protein folding, response to stimuli, and metabolic processes (Xu et al., 2016a). A number of these genes and associated ontological functions might provide critical insight into the role that the *MPG1* transcription factor plays natively, as well as what might be affected in the mutant *mpg1*.

The presence of the T-DNA insertion, correlated strongly with the elevated level of expression of this AP2/EREBP transcription factor *MPG1*. Although proximity and correlation are not concrete examples of evidence of causation, it seemed advantageous to investigate this gene and further comprehensively phenotype *mpg1* in effort to uncover the molecular mechanisms resulting in increased plant productivity. Continued assessment of the presence of the T-DNA insertion, correlating with continued elevated expression of *MPG1*, and deeper analysis of the phenotypic characteristics present in this mutant through multiple generations and backcrosses should help solidify that this caused the phenotype in *mpg1* and further uncover the functionality of this candidate gene (assessed in Chapter 3).

Additionally, recapitulation of the observed phenotype in *mpg1* through independent genetic confirmation would validate the case that *MPG1* plays a role in generating the phenotype in the mutant. Replicating the phenotype however could prove difficult, as there is a level of uncertainty involved with what caused the elevated expression of *MPG1* in the mutant lines containing the bi-laterally truncated T-DNA insertion. As of this point we have observed what appears to be respectively uniform levels of heightened expression of *MPG1* in mutant plants both heterozygous and homozygous for the T-DNA insertion in both vegetative and reproductive stages across varying tissue types. This suggests that *mpg1* plants are constitutively overexpressing *MPG1*, as we have not been able to find tissues with less or differing expression.

It is possible that promoter sequences remaining in the truncated T-DNA insertion are directly influencing the expression of *MPG1*. The truncated T-DNA contains a duplicated *CaMV35S* promoter sequence, as well as roughly the first third of the *CmGAS1* promoter sequence. It is likely promoter elements from the expression cassette are driving expression of the candidate gene, which would mean that it could be constitutively overexpressed through the influence of the duplicated *CaMV35S* promoter sequence, or it could be expressed exclusively within the companion cells through the influence of the remaining *CmGAS1* promoter sequence, assuming the remaining sequence is sufficient to mediate tissue specific gene expression. Alternatively, expression could be driven by a complex interplay between these two promoters, or surrounding genomic elements.

Other factors that might play a role in the expression pattern of *MPG1* observed in the mutant could be a result of disruption by the T-DNA insertion itself. A number of genes flanking the T-DNA insertion site are genes coding for retrotransposon proteins. Often, genes coding for retrotransposon proteins become immobilized through transcriptional limitation via methylation (Kato et al., 2003). It is possible that the particular region that the T-DNA integrated is naturally highly condensed, excluding transcriptional machinery from physically accessing this area under native conditions. In support of this the Plant Methylome Database shows regions flanking the insertion site are highly methylated (Niederhuth et al., 2016). The T-DNA insertion, although bi-laterally truncated, is still quite large (3190 base-pairs). Although none of the other genes surrounding the insertion site experienced noticeably different expression in *mpg1* compared to wild-type, it is possible that the T-DNA allowed the region of gDNA to be more presentable to transcriptional machinery inducing greater transcription of the candidate gene. (Gelvin and Kim, 2007).

Again, there was a 20 base-pair genomic DNA deletion that occurred at the site of the T-DNA integration. It is possible that either the T-DNA or the deletion interfered with a functional element not yet described. For example, it is possible that the site of integration, or immediately proximal, natively contains a sequence utilized by a repressor that acts on *MPG1* that was removed or disrupted when the T-DNA integrated leading to the heightened levels of expression in *mpg1*.

Any number of these reasons, or combination, could explain the altered expression of *MPG1* in the mutant. Regardless of why the gene of interest is experiencing elevated expression compared to wild-type plants, additional investigation into *MPG1*'s global effects could provide insight into its functionality. Specifically evaluating transcriptome differences as well as uncovering the specific target sites of this transcription factor could reveal other genes and pathways that are affected resulting in the specific characteristics observed in *mpg1*.

BIBLIOGRAPHY

AOKI, N., HILROSE, T., SCOFIELD, G. N., WHITFELD, P. R. & FURBANK, R. T. 2003. The sucrose transporter gene family in rice. *Plant Cell Physiol,* 44**,** 223-232.

AYRE, B. G., BLAIR, J. E. & TURGEON, R. 2003. Functional and phylogenetic analyses of a conserved regulatory program in the phloem of minor veins. *Plant Physiol,* 133**,** 1229-39.

BRAUN, D. M. & SLEWINSKI, T. L. 2009. Genetic control of carbon partitioning in grasses: roles of sucrose transporters and tie-dyed loci in phloem loading. *Plant Physiol,* 149**,** 71-81.

BURKLE, L., HIBBERD, J. M., QUICK, W. P., KUHN, C., HIRNER, B. & FROMMER, W. B. 1998. The H+-sucrose cotransporter NtSUT1 is essential for sugar export from tobacco leaves. *Plant Physiol,* 118**,** 59-68.

BUSH, D. R. 1993. Proton-coupled sugar and amino acid transporters in plants. *Ann Rev Plant Physiol Plant Mol Bio,* 44**,** 513-542.

CASPAR, T., HUBER, S. C. & SOMERVILLE, C. 1985. Alterations in Growth, Photosynthesis, and Respiration in a Starchless Mutant of Arabidopsis thaliana (L.) Deficient in Chloroplast Phosphoglucomutase Activity. *Plant Physiol,* 79**,** 11-7.

CHIOU, T. & BUSH, D. R. 1998. Sucrose is a signal molecule in assimilate partitioning. *Proc Natl Acad Sci U S A,* 95**,** 4784-4788.

COUEE, I., SULMON, C., GOUESBET, G. & EL AMRANI, A. 2006. Involvement of soluble sugars in reactive oxygen species balance and responses to oxidative stress in plants. *J Exp Bot,* 57**,** 449-59.

DASGUPTA, K., KHADILKAR, A. S., SULPICE, R., PANT, B., SCHEIBLE, W. R., FISAHN, J., STITT, M. & AYRE, B. G. 2014. Expression of Sucrose Transporter cDNAs Specifically in Companion Cells Enhances Phloem Loading and Long-Distance Transport of Sucrose but Leads to an Inhibition of Growth and the Perception of a Phosphate Limitation. *Plant Physiol,* 165**,** 715-731.

DOWNIE, B., GURUSINGHE, S., DAHAL, P., THACKER, R. R., SNYDER, J. C., NONOGAKI, H., YIM, K., FUKANAGA, K., ALVARADO, V. & BRADFORD, K. J. 2003. Expression of a GALACTINOL SYNTHASE gene in tomato seeds is up-regulated before maturation desiccation and again after imbibition whenever radicle protrusion is prevented. *Plant Physiol,* 131**,** 1347-59.

EDGERTON, M. D. 2009. Increasing crop productivity to meet global needs for feed, food, and fuel. *Plant Physiol,* 149**,** 7-13.

FONDY, B. R. & GEIGER, D. R. 1985. Diurnal changes in allocation of newly fixed carbon in exporting sugar beet leaves. *Plant Physiol,* 78**,** 753-7.

FUKUDA, A., FUJIMAKI, S., MORI, T., SUZUI, N., ISHIYAMA, K., HAYAKAWA, T., YAMAYA, T., FUJIWARA, T., YONEYAMA, T. & HAYASHI, H. 2005. Differential distribution of proteins expressed in companion cells in the sieve element-companion cell complex of rice plants. *Plant Cell Physiol,* 46**,** 1779-86.

GALTIER, N., FOYER, C. H., HUBER, J., VOELKER, T. A. & HUBER, S. C. 1993. Effects of Elevated Sucrose-Phosphate Synthase Activity on Photosynthesis, Assimilate Partitioning, and Growth in Tomato (Lycopersicon esculentum var UC82B). *Plant Physiol,* 101**,** 535-543.

GALTIER, N., FOYER, C. H., MURCHIE, E., ALRED, R., QUICK, P., VOELKER, T. A., THEPENIER, C., LASCEVE, G. & BETSCHE, T. 1995. Effects of light and atmospheric carbon dioxide enrichment on photosynthesis and carbon partitioning in the leaves of tomato (Lycopersicon esculentum L.) plants over-expressing sucrose phosphate synthase. *J Exp Bot,* 46**,** 1335-1344.

GEIGER, D. R. 1975. Phloem Loading. *In:* ZIMMERMANN, M. H. & MILBURN, J. A. (eds.) *Transport in Plants I. Encyclopedia of Plant Physiology (New Series).* Berlin, Heidelberg: Springer.

GELVIN, S. B. & KIM, S. I. 2007. Effect of chromatin upon Agrobacterium T-DNA integration and transgene expression. *Biochim Biophys Acta,* 1769**,** 410-21.

GIAQUINTA, R. T. 1983. Phloem loading of sucrose. *Ann Rev Plant Physiol,* 34**,** 347-387.

GIFFORD, R. M. & EVANS, L. T. 1981. Photosynthesis, carbon partitioning, and yield. *Ann Rev Plant Physiol,* 32**,** 485-509.

GOLDSCHMIDT, E. E. & HUBER, S. C. 1992. Regulation of photosynthesis by end-product accumulation in leaves of plants storing starch, sucrose, and hexose sugars. *Plant Physiol,* 99**,** 1443-8.

GRODZINSKI, B., JIAO, J. & LEONARDOS, E. D. 1998. Estimating photosynthesis and concurrent export rates in C3 and C4 species at ambient and elevated CO2. *Plant Physiol,* 117**,** 207-215.

HACKEL, A., SCHAUER, N., CARRARI, F., FERNIE, A. R., GRIMM, B. & KUHN, C. 2006. Sucrose transporter LeSUT1 and LeSUT2 inhibition affects tomato fruit development in different ways. *Plant J,* 45**,** 180-92.

HARITATOS, E., AYRE, B. G. & TURGEON, R. 2000a. Identification of phloem involved in assimilate loading in leaves by the activity of the galactinol synthase promoter. *Plant Physiol,* 123**,** 929-37.

HARITATOS, E., MEDVILLE, R. & TURGEON, R. 2000b. Minor vein structure and sugar transport in Arabidopsis thaliana. *Planta,* 211**,** 105-11.

HICKLEY, J. M., CHIURUGWI, T., MACKAY, I., POWELL, W. & CGIAR 2017. Genomic prediction unifies animal and plant breeding programs to form platforms for biological discovery. *Nature Genetics,* 49**,** 1297-1303.

ISHIMARU, K., HIROSE, T., AOKI, N., TAKAHASHI, S., ONO, K., YAMAMOTO, S., WU, J., SAJI, S., BABA, T., UGAKI, M., MATSUMOTO, T. & OHSUGI, R. 2001. Antisense expression of a rice sucrose transporter OsSUT1 in rice (Oryza sativa L.). *Plant Cell Physiol,* 42**,** 1181-5.

ISHIWATARI, Y., FUJIWARA, T., MCFARLAND, K. C., NEMOTO, K., HAYASHI, H., CHINO, M. & LUCAS, W. J. 1998. Rice phloem thioredoxin h has the capacity to mediate its own cell-to-cell transport through plasmodesmata. *Planta,* 205**,** 12-22.

ISHIWATARI, Y., HONDA, C., KAWASHIMA, I., NAKAMURA, S., HIRANO, H., MORI, S., FUJIWARA, T., HAYASHI, H. & CHINO, M. 1995. Thioredoxin h is one of the major proteins in rice phloem sap. *Planta,* 195**,** 456-463.

ISHIWATARI, Y., NEMOTO, K., FUJIWARA, T., CHINO, M. & HAYASHI, H. 2000. In situ hybridization study of the rice phloem thioredoxin h mRNA accumulation - possible involvement in the differentiation of vascular tissues. *Physiol Plantarum,* 109**,** 90-96.

JANG, J. C. & SHEEN, J. 1994. Sugar sensing in higher plants. *Plant Cell,* 6**,** 1665-79.

KATO, M., MIURA, A., BENDER, J., JACOBSEN, S. E. & KAKUTANI, T. 2003. Role of CG and non-CG methylation in immobilization of transposons in Arabidopsis. *Curr Biol,* 13**,** 421-6.

KELLERMEIER, F., ARMENGAUD, P., SEDITAS, T. J., DANKU, J., SALT, D. E. & AMTMANN, A. 2014. Analysis of the Root System Architecture of Arabidopsis Provides a Quantitative Readout of Crosstalk between Nutritional Signals. *Plant Cell,* 26**,** 1480-1496.

KIM, M. S., CHO, S. M., KANG, E. Y., IM, Y. J., HWANGBO, H., KIM, Y. C., RYU, C., YANG, K. Y., CHUNG, G. C. & CHO, B. H. 2008. Galactinol is a signaling component of the induced systemic resistance caused by Pseudomonas chlororaphis O6 root colonization. *Mol Plant Microbe In,* 21**,** 1653-1653.

KOFLER, H., HAUSLER, R. E., SCHULZ, B., GRONER, F., FLUGGE, U. I. & WEBER, A. 2000. Molecular characterisation of a new mutant allele of the plastid phosphoglucomutase in Arabidopsis, and complementation of the mutant with the wild-type cDNA. *Mol Gen Genet,* 263**,** 978-86.

KOMOR, E. 2000. Source physiology and assimilate transport: the interaction of sucrose metabolism, starch storage and phloem export in source leaves and the effects on sugar status in phloem. *Aust J Plant Physiol,* 27**,** 497-505.

KOONIN, S. E. 2006. Getting serious about biofuels. *Science,* 311**,** 435.

KRAPP, A., HOFMANN, B., SCHAFER, C. & STITT, M. 1993. Regulation of the expression of rbcS and other photosynthetic genes by carbohydrates: a mechanism for the 'sink regulation' of photosynthesis? *Plant J,* 3**,** 817-828.

KUHN, C. & GROF, C. P. 2010. Sucrose transporters of higher plants. *Curr Opin Plant Biol,* 13**,** 288-98.

KUHN, C., QUICK, W. P., SCHULZ, A., RIESMEIER, J. W., SONNEWALD, U. & FROMMER, W. B. 1996. Companion cell-specific inhibition of the potato sucrose transporter SUT1. *Plant Cell Environ,* 19**,** 1115-1123.

LEGGEWIE, G., KOLBE, A., LEMOINE, R., ROESSNER, U., LYTOVCHENKO, A., ZUTHER, E., KEHR, J., FROMMER, W. B., RIESMEIER, J. W., WILLMITZER, L. & FERNIE, A. R. 2003. Overexpression of the sucrose transporter SoSUT1 in potato results in alterations in leaf carbon partitioning and in tuber metabolism but has little impact on tuber morphology. *Planta,* 217**,** 158-67.

LEI, M., LIU, Y., ZHANG, B., ZHAO, Y., WANG, X., ZHOU, Y., RAGHOTHAMA, K. G. & LIU, D. 2011. Genetic and genomic evidence that sucrose is a global regulator of plant responses to phosphate starvation in Arabidopsis. *Plant Physiol,* 156**,** 1116-30.

LIEMAN-HURWITZ, J., RACHMILEVITCH, S., MITTLER, R., MARCUS, Y. & KAPLAN, A. 2003. Enhanced photosynthesis and growth of transgenic plants that express ictB, a gene involved in HCO3- accumulation in cyanobacteria. *Plant Biotechnol J,* 1**,** 43-50.

LIU, P., WALLACE, D. H. & OZBUN, J. L. 1973. Influence of Translocation of Photosynthetic Efficiency of Phaseolus vulgaris L. *Plant Physiol,* 52**,** 412-5.

LIU, Y. G. & CHEN, Y. 2007. High-efficiency thermal asymmetric interlaced PCR for amplification of unknown flanking sequences. *Biotechniques,* 43**,** 649-50, 652, 654 passim.

LLOYD, J. C. & ZAKHLENIUK, O. V. 2004. Responses of primary and secondary metabolism to sugar accumulation revealed by microarray expression analysis of the Arabidopsis mutant, pho3. *J Exp Bot,* 55**,** 1221-30.

LU, C., TEJ, S. S., LUO, S., HAUDENSCHILD, C. D., MEYERS, B. C. & GREEN, P. J. 2005. Elucidation of the small RNA component of the transcriptome. *Science,* 309**,** 1567-9.
LU, J. M. & BUSH, D. R. 1998. His-65 in the proton-sucrose symporter is an essential amino acid whose modification with site-directed mutagenesis increases transport activity. *Proc Natl Acad Sci U S A,* 95**,** 9025-9030.
MATSUDA, Y., LIANG, G., ZHU, Y., MA, F., NELSON, R. S. & DING, B. 2002. The Commelina Yellow Mottle Virus promoter drives companion-cell-specific gene expression in multiple organs of tansgenic tobacco. *Protoplasma,* 220**,** 0051-0058.
MATSUKURA, C., SAITOH, T., HIROSE, T., OHSUGI, R., PERATA, P. & YAMAGUCHI, J. 2000. Sugar uptake and transport in rice embryo. Expression of companion cell-specific sucrose transporter (OsSUT1) induced by sugar and light. *Plant Physiol,* 124**,** 85-93.
MEDBERRY, S. L., LOCKHART, B. E. L. & OLSZEWSKI, N. E. 1992. The Commelina Yellow Mottle Virus promoter is a strong promoter in vascular and reproductive tissues. *the Plant Cell,* 4**,** 185-192.
MIYAGAWA, Y., TAMOI, M. & SHIGEOKA, S. 2001. Overexpression of a cyanobacterial fructose-1,6-/sedoheptulose-1,7-bisphosphatase in tobacco enhances photosynthesis and growth. *Nat Biotechnol,* 19**,** 965-969.
MOORE, B. D., CHENG, S. H., RICE, J. & SEEMANN, J. R. 1998. Sucrose cycling, Rubisco expression, and prediction of photosynthetic acclimation to elevated atmospheric CO2. *Plant Cell Environ,* 21**,** 905-915.
MULLER-ROBER, B. T., KOSSMANN, J., HANNAH, L. C., WILLMITZER, L. & SONNEWALD, U. 1990. One of two different ADP-glucose pyrophosphorylase genes from potato responds strongly to elevated levels of sucrose. *Mol Gen Genet,* 224**,** 136-46.
NAKANO, T., SUZUKI, K., FUJIMURA, T. & SHINSHI, H. 2006. Genome-wide analysis of the ERF gene family in Arabidopsis and rice. *Plant Physiol,* 140**,** 411-32.
NIE, G., HENDRIX, D. L., WEBBER, A. N., KIMBALL, B. A. & LONG, S. P. 1995. Increased accumulation of carbohydrates and decreased photosynthetic gene transcript levels in wheat grown at an elevated CO2 concentration in the field. *Plant Physiol,* 108**,** 975-983.
NIEDERHUTH, C. E., BEWICK, A. J., JI, L., ALABADY, M. S., KIM, K. D., LI, Q., ROHR, N. A., RAMBANI, A., BURKE, J. M., UDALL, J. A., EGESI, C., SCHMUTZ, J., GRIMWOOD, J., JACKSON, S. A., SPRINGER, N. M. & SCHMITZ, R. J. 2016. Widespread natural variation of DNA methylation within angiosperms. *Genome Biol,* 17**,** 194.
NUNES-NESI, A., CARRARI, F., LYTOVCHENKO, A., SMITH, A. M., LOUREIRO, M. E., RATCLIFFE, R. G., SWEETLOVE, L. J. & FERNIE, A. R. 2005. Enhanced photosynthetic performance and growth as a consequence of decreasing mitochondrial malate dehydrogenase activity in transgenic tomato plants. *Plant Physiol,* 137**,** 611-22.
OKADA, M. & YE, K. 2009. Nuclear phosphoinositide signaling regulates messenger RNA export. *RNA Biol,* 6**,** 12-6.
OSUNA, D., USADEL, B., MORCUENDE, R., GIBON, Y., BLASING, O. E., HOHNE, M., GUNTER, M., KAMLAGE, B., TRETHEWEY, R., SCHEIBLE, W. R. & STITT, M. 2007. Temporal responses of transcripts, enzyme activities and metabolites after adding sucrose to carbon-deprived Arabidopsis seedlings. *Plant J,* 49**,** 463-91.
OUYANG, S., ZHU, W., HAMILTON, J., LIN, H., CAMPBELL, M., CHILDS, K., THIBAUD-NISSEN, F., MALEK, R. L., LEE, Y., ZHENG, L., ORVIS, J., HAAS, B., WORTMAN,

J. & BUELL, C. R. 2007. The TIGR Rice Genome Annotation Resource: improvements and new features. *Nucleic Acids Res,* 35**,** D883-7.

PAUL, M. J. & FOYER, C. H. 2001. Sink regulation of photosynthesis. *J Exp Bot,* 52**,** 1383-400.

PETERS, S. & KELLER, F. 2009. Frost tolerance in excised leaves of the common bugle (Ajuga reptans L.) correlates positively with the concentrations of raffinose family oligosaccharides (RFOs). *Plant Cell Environ,* 32**,** 1099-107.

POURTAU, N., JENNINGS, R., PELZER, E., PALLAS, J. & WINGLER, A. 2006. Effect of sugar-induced senescence on gene expression and implications for the regulation of senescence in Arabidopsis. *Planta,* 224**,** 556-68.

RICE FULL-LENGTH C, D. N. A. C., NATIONAL INSTITUTE OF AGROBIOLOGICAL SCIENCES RICE FULL-LENGTH C, D. N. A. P. T., KIKUCHI, S., SATOH, K., NAGATA, T., KAWAGASHIRA, N., DOI, K., KISHIMOTO, N., YAZAKI, J., ISHIKAWA, M., YAMADA, H., OOKA, H., HOTTA, I., KOJIMA, K., NAMIKI, T., OHNEDA, E., YAHAGI, W., SUZUKI, K., LI, C. J., OHTSUKI, K., SHISHIKI, T., FOUNDATION OF ADVANCEMENT OF INTERNATIONAL SCIENCE GENOME, S., ANALYSIS, G., OTOMO, Y., MURAKAMI, K., IIDA, Y., SUGANO, S., FUJIMURA, T., SUZUKI, Y., TSUNODA, Y., KUROSAKI, T., KODAMA, T., MASUDA, H., KOBAYASHI, M., XIE, Q., LU, M., NARIKAWA, R., SUGIYAMA, A., MIZUNO, K., YOKOMIZO, S., NIIKURA, J., IKEDA, R., ISHIBIKI, J., KAWAMATA, M., YOSHIMURA, A., MIURA, J., KUSUMEGI, T., OKA, M., RYU, R., UEDA, M., MATSUBARA, K., RIKEN, KAWAI, J., CARNINCI, P., ADACHI, J., AIZAWA, K., ARAKAWA, T., FUKUDA, S., HARA, A., HASHIZUME, W., HAYATSU, N., IMOTANI, K., ISHII, Y., ITOH, M., KAGAWA, I., KONDO, S., KONNO, H., MIYAZAKI, A., OSATO, N., OTA, Y., SAITO, R., SASAKI, D., SATO, K., SHIBATA, K., SHINAGAWA, A., SHIRAKI, T., YOSHINO, M., HAYASHIZAKI, Y. & YASUNISHI, A. 2003. Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice. *Science,* 301**,** 376-9.

RIESMEIER, J. W., WILLMITZER, L. & FROMMER, W. B. 1994. Evidence for an essential role of the sucrose transporter in phloem loading and assimilate partitioning. *EMBO J,* 13**,** 1-7.

ROSCHE, E., BLACKMORE, D., TEGEDER, M., RICHARDSON, T., SCHROEDER, H., HIGGINS, T. J., FROMMER, W. B., OFFLER, C. E. & PATRICK, J. W. 2002. Seed-specific overexpression of a potato sucrose transporter increases sucrose uptake and growth rates of developing pea cotyledons. *Plant J,* 30**,** 165-75.

SAWADA, S., SATO, M., KASAI, A., YAOCHI, D., KAMEYA, Y., MATSUMOTO, I. & KASAI, M. 2003. Analysis of the Feed-Forward Effects of Sink Activity on the Photosynthetic Source-Sink Balance in Single-Rooted Sweet Potato Leaves. I. Activation of RuBPcase through the Development of Sinks. *Plant Cell Physiol,* 44**,** 190-197.

SCHMIDHUBER, J. & TUBIELLO, F. N. 2007. Global food security under climate change. *Proc Natl Acad Sci U S A,* 104**,** 19703-8.

SCHNEIDEREIT, A., IMLAU, A. & SAUER, N. 2008. Conserved cis-regulatory elements for DNA-binding-with-one-finger and homeo-domain-leucine-zipper transcription factors regulate companion cell-specific expression of the Arabidopsis thaliana SUCROSE TRANSPORTER 2 gene. *Planta,* 228**,** 651-62.

SCOFIELD, G. N., HAIROSE, T., GAUDRON, J. A., FURBANK, R. T., UPADHYAYA, N. M. & OHSUGI, R. 2002. Antisense suppression of the rice transporter gene, OsSUT1,

leads to impaired grain filling and germination but does not affect photosynthesis. *Funct Plant Biol,* 29**,** 815-826.
SENGUPTA, S., MUKHERJEE, S., BASAK, P. & MAJUMDER, A. L. 2015. Significance of galactinol and raffinose family oligosaccharide synthesis in plants. *Front Plant Sci,* 6**,** 656.
SHARONI, A. M., NURUZZAMAN, M., SATOH, K., SHIMIZU, T., KONDOH, H., SASAYA, T., CHOI, I. R., OMURA, T. & KIKUCHI, S. 2011. Gene structures, classification and expression models of the AP2/EREBP transcription factor family in rice. *Plant Cell Physiol,* 52**,** 344-60.
SOKOLOV, L. N., DÉJARDIN, A. & KLECZKOWSKI, L. A. 1998. Sugars and light/dark exposure trigger differential regulation of ADP-glucose pyrophosphorylase genes in Arabidopsis thaliana (thale cress). *Biochem J,* 336**,** 681-687.
SOUTH, P. F., CAVANAGH, A. P., LIU, H. W. & ORT, D. R. 2019. Synthetic glycolate metabolism pathways stimulate crop growth and productivity in the field. *Science,* 363.
SRIVASTAVA, A. C., GANESAN, S., ISMAIL, I. O. & AYRE, B. G. 2008. Functional characterization of the Arabidopsis AtSUC2 Sucrose/H+ symporter by tissue-specific complementation reveals an essential role in phloem loading but not in long-distance transport. *Plant Physiol,* 148**,** 200-11.
STADLER, R. & SAUER, N. 1996. The Arabidopsis thaliana AtSUC2 Gene is Specifically Expressed in Companion Cells. *Bot Acta***,** 299-306.
STEVENSON, J. M., PERERA, I. M., HEILMANN, I., PERSSON, S. & BOSS, W. F. 2000. Inositol signaling and plant growth. *Trends Plant Sci,* 5**,** 252-258.
THOLE, J. M. & NIELSEN, E. 2008. Phosphoinositides in plants: novel functions in membrante trafficking. *Curr Opin Plant Biol,* 11**,** 620-631.
TRUERNIT, E. & SAUER, N. 1995. The promoter of the Arabidopsis thaliana SUC2 sucrose-H+ symporter gene directs expression of B-glucuronidase to the phloem: Evidence for phloem loading and unloading by SUC2. *Planta,* 196**,** 564-570.
TSUWAMOTO, R. & HARADA, T. 2010. Identification of a cis-regulatory element that acts in companion cell-specific expression of AtMT2B promoter through the use of Brassica vasculature and gene-gun-mediated transient assay. *Plant Cell Physiol,* 51**,** 80-90.
TURGEON, R. 1996. Phloem loading and plasmodesmata. *Trends Plant Sci,* 1**,** 418-423.
TURGEON, R. & MEDVILLE, R. 2004. Phloem loading. A reevaluation of the relationship between plasmodesmatal frequencies and loading strategies. *Plant Physiol,* 136**,** 3795-803.
VAN BEL, A. J. E. 1993. Strategies of phloem loading. *Ann Rev Plant Physiol Plant Mol Biol,* 44**,** 253-281.
VAUGHN, M. W., HARRINGTON, G. N. & BUSH, D. R. 2002. Sucrose-mediated transcriptional regulation of sucrose symporter activity in the phloem. *Proc Natl Acad Sci U S A,* 99**,** 10876-80.
VOLK, G. M., HARITATOS, E. E. & TURGEON, R. 2003. Galactinol synthase gene expression in melon. *J Amer Soc Hort Sci* 128**,** 8-15.
XU, J. J., ZHANG, X. F. & XUE, H. W. 2016. Rice aleurone layer specific OsNF-YB1 regulates grain filling and endosperm development by interacting with an ERF transcription factor. *J Exp Bot,* 67**,** 6399-6411.
XUE, H., CHEN, X. & LI, G. 2007. Involvement of phospholipid signaling in plant growth and hormone effects. *Curr Opin Plant Biol,* 10**,** 483-9.

ZHANG, C. J., ZHAO, B. C., GE, W. N., ZHANG, Y. F., SONG, Y., SUN, D. Y. & GUO, Y. 2011. An apoplastic h-type thioredoxin is involved in the stress response through regulation of the apoplastic reactive oxygen species in rice. *Plant Physiol,* 157**,** 1884-99.

ZUTHER, E., BUCHEL, K., HUNDERTMARK, M., STITT, M., HINCHA, D. K. & HEYER, A. G. 2004. The role of raffinose in the cold acclimation response of Arabidopsis thaliana. *FEBS Lett,* 576**,** 169-73.

ZUTHER, E., SCHULZ, E., CHILDS, L. H. & HINCHA, D. K. 2012. Clinal variation in the non-acclimated and cold-acclimated freezing tolerance of Arabidopsis thaliana accessions. *Plant Cell Environ,* 35**,** 1860-1878.

CHAPTER 3: Characterization and comprehensive phenotyping of mutant rice plant *mpg1*

Introduction

Numerous transgenic plants containing a T-DNA expression cassette engineered to modulate a pivotal point in carbon assimilation were generated. The expression cassette contained a hyperactive sucrose symporter from *Arabidopsis* (*AtSUC1$_{H65K}$*) involved in phloem loading of sucrose (Lu and Bush, 1998), directed by a companion-cell specific promoter from *Cucumis melo* (*CmGAS1pro*) (Haritatos et al., 2000a, Volk et al., 2003). This design was intended to increase phloem loading of sucrose, thereby enhancing export of sucrose from source tissues, increasing photosynthetic activity and reducing effects of carbohydrate-mediated repression, resulting in greater pools of carbon in sink tissues, potentially raising overall plant productivity. These transgenic plants were screened for enhanced biomass and yield characteristics.

Surprisingly, none of the plants containing the full transgene expression cassette (incorporating *AtSUC1$_{H65K}$*) showed an increase in biomass or yield. In fact, these plants were significantly smaller compared to wild-type (Chapter 2). Confirmation of companion cell specific targeting of the *CmGAS1* promoter in rice was validated using a promoter-GUS fusion construct. Thus, the lack of biomass phenotype observed in the over-expression *AtSUC1$_{H65K}$* rice plants is not likely due to mis-targeting, as the *AtSUC1$_{H65K}$* should be expressed in a companion cell specific fashion. Expression analysis of *AtSUC1$_{H65K}$* in plants containing the full transgene expression cassette was verified using RT-PCR. Plants expressing the sucrose transporter transgene were stunted compared to wild-type plants. There could be numerous reasons why the *CmGAS1*pro::*AtSUC1$_{H65K}$* transgenic plants didn't result in increased biomass. Other studies

investigating the function of enhancing phloem loading by manipulation of sucrose transporters in companion cells observed similar phenotypic outcomes to these transgenic plants (Dasgupta et al., 2014). Their studies uncovered that *Arabidopsis* driving expression of sucrose transporters under the direction of a virus based companion cell-specific promoter (*CoYMVpro*) (Matsuda et al., 2002, Medberry et al., 1992) were able to enhance phloem loading and long distance transport of carbon, however gave way to the activation of phosphate starvation genes, and reduced overall biomass (Dasgupta et al., 2014). Their study suggests that the phosphate to carbon ratio must be properly balanced for plants to appropriately partition and metabolize assimilates. A shift in this ratio likely results from altered signaling and inability for the plant to use sink carbon sources for enhanced growth and productivity. Assimilate partitioning is a relatively complex pathway with many dynamic characteristics. Further investigation in the crosstalk and balance of nutrients will be needed to better understand the effects of enhanced photoassimilate partitioning and mechanisms by which to enhance plant productivity.

However, out of the numerous transgenics generated to assess the effects of enhanced phloem loading, there was a single plant that was significantly larger than its counterparts and wild-type plants. Inspection of this particular plant, showing enhanced biomass accumulation, revealed that it only contained a portion of the full T-DNA expression cassette intended for integration (Figure 3.1). Sequencing of the insertion showed that the T-DNA expression cassette was bi-laterally truncated and resulted in a 20 base-pair genomic DNA deletion at the site of integration. This type of disrupted T-DNA insertion is not uncommon (Castle et al., 1993, Gheysen et al., 1991, Mayerhofer et al., 1991). The remnants of the insertion mainly contained a duplicated *CaMV35S* promoter driving a gene coding for hygromycin resistance, as well as roughly one-third of the total *CmGAS1* promoter sequence (Figure 3.1). These partial insertions

have been shown to integrate based on microhomologies consistent between the T-DNA and genomic DNA sequence at the site of insertion (Pelczar et al, 2004). This could potentially explain the reason behind the truncated T-DNA and site of insertion in *mpg1*.

The insertion did not contain the gene of interest pertaining to sucrose transport. Thus, the phenotype of the large transgenic plant is not the result of $AtSUC1_{H65K}$ expression. We gathered seed from this plant and confirmed the high biomass phenotype in the progeny, which segregated with the presence of the T-DNA expression cassette.

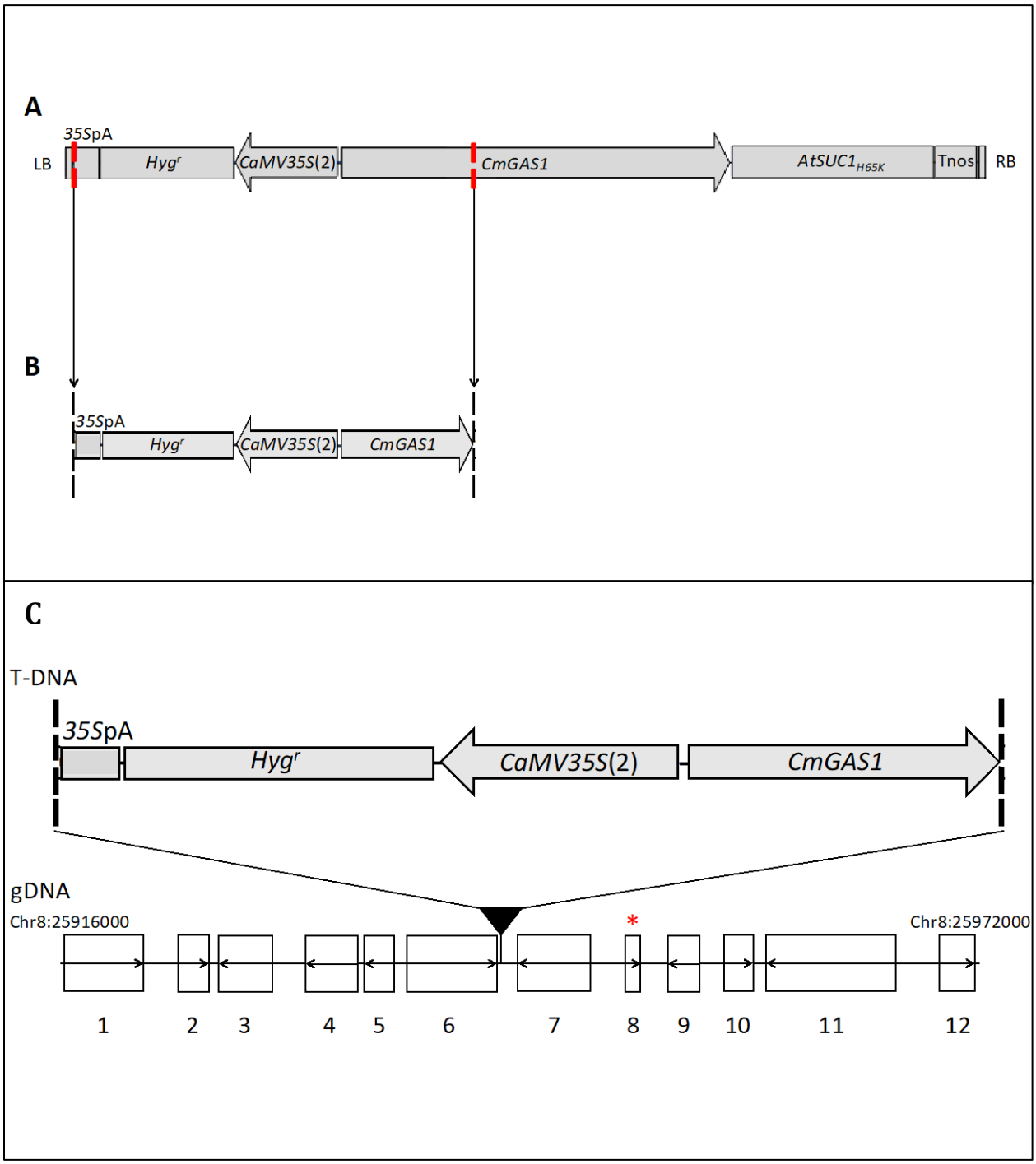


Figure 3.1: Schematic representing the extent and orientation of sequence elements of the T-DNA insertion, site of genomic integration, and proximity and location of differentially expressed AP2/EREBP (*MPG1*).

(A) Visual representation of the complete T-DNA insertion cassette designed to generate intended transformants. LB = left border, 35SpA = *CaMV35S* poly A sequence, *Hygr* = gene coding for hygromycin resistance, *CaMV35S(2)* = duplicated *CaMV35S* promoter sequence,

CmGAS1 = *CmGAS1* promoter sequence, $AtSUC1_{H65K}$ = gene coding for hyperactive sucrose symporter, Tnos = nos terminator, RB = right border. (B) The remnants of the intended T-DNA expression cassette present in *mpg1*. The insertion is bi-laterally truncated, with 40 base-pairs absent from the left border, and everything from the right border up through the first 1037 base-pairs of the *CmGAS1* promoter sequence, compared to the complete T-DNA sequence. (C) A visual schematic showing the location of the bi-laterally truncated T-DNA insertion present in *mpg1* (above) and its location of integration in the rice genome, gDNA (below). HI-TAIL-PCR revealed that the insertion event occurred on chromosome 8 between Os08g41010 and Os08g41020. Further sequencing of the region concluded that the T-DNA integrated just after loci Chr8:25,942,615 (NCBI IRGSP-1.0 genome build). At the location of the insertion a 20 base-pair genomic genomic DNA (gDNA) deletion occurred (deleting Chr8:25,942,616-25,942,635). The total size of the insertion is 3190 base-pairs in length. The genomic region displayed spans from Chr8:25916000-Chr8:2597200. Expression analysis was performed via semi-quantitative RT-PCR on neighboring genes flanking the insertion site for both wild-type and *mpg1* plants. The red asterisk represents the only gene experiencing notable differential expression, LOC_Os08g41030 (8). Genes listed: LOC_Os08g40960 – retrotransposon protein (1), LOC_Os08g40970 – retrotransposon protein (2), LOC_Os08g40980 – retrotransposon protein (3), LOC_Os08g40990 – receptor-like protein kinase 1 (4), LOC_Os08g41000 – extracellular ligand-gated ion channel (5), LOC_Os08g41010 – zinc finger family protein (6), LOC_Os08g41020 – retrotransposon protein (7), LOC_Os08g41030 – AP2 domain containing protein (8), LOC_Os08g41040 – expressed protein (9), LOC_Os08g41054 – hypothetical protein (10), LOC_Os08g41070 – retrotransposon protein (11), LOC_Os08g41080 – expressed protein (12).

Genetic engineering and integration of T-DNA into plant cells was accomplished by *Agrobacterium* mediated transformation. *Agrobacterium*, a soil-based bacterial plant pathogen, works by utilizing its native tumor-inducing (Ti) plasmid which transfers specific genes to host cells, which natively, results in crown-gall disease (Smith and Townsend, 1907), a neoplastic growth. This system can be hijacked for genetic engineering or biotechnological purposes. The genes resulting in pathogenesis can be replaced by exogenous or modified DNA, which is transferred into host cells and stably incorporated into the host's genomic DNA. *Agrobacterium* mediated transformation relies on certain cues (from pathogen and host) that result in T-DNA transport into the host cells. Important genetic components present in the bacterium chromosomal DNA and T-plasmid include chromosomal virulence (*chv*) genes, T-plasmid virulence (*vir*) genes, and T-DNA border sequences, reviewed in (Gelvin, 2000, Gelvin, 2003,

Tzfira and Citovsky, 2000, Tzfira and Citovsky, 2002, Zupan et al., 2000, Zambryski, 1992, Pitzschke, 2013, Dafny-Yelin et al., 2008, Pitzschke and Hirt, 2010). During the transformation process, the *Agrobacterium vir* region is induced by host phenolic signals as a result of wounding (Stachel et al., 1985), leading to insertion of bacterial substrates (T-DNA and virulence proteins) via a type IV secretion system (Christie, 2004). The induction of two virulence proteins, VirD1 and VirD2, act as site-specific nucleases cleaving the bottom strand at the T-DNA borders giving rise to a single-stranded T-DNA molecule (T-strand) (Scheiffele et al., 1995). This T-strand joins together with several Vir proteins and is exported from the *Agrobacterium* to the host cell through a channel created by VirB and VirD4 (Christie and Vogel, 2000). By the time the T-strand makes its way to the host cytoplasm, it forms a nucleoprotein complex (T-complex), where the VirD2 molecule is covalently attached to the 5'-end of the T-DNA and the remainder of the structure surrounded by VirE2 molecules, which interact with VirE2-interacting protein 1 (VIP1), amongst other potential cellular interactions, to guide the T-complex to the nucleus of the host cell via the importin α-mediated nuclear import pathway (Loyter et al., 2005, Tzfira et al., 2001, Tzfira et al., 2002). Once the T-complex has entered the host's nucleus, VIP1 mediates its association with chromatin through interaction with core histones (Li et al., 2005a, Lacroix et al., 2008). The masking of the T-DNA-protein-complex is degraded by Skp1-Cul1-F-box protein (SCF) host ubiquitin/proteasome machinery (Zaltsman et al., 2010a, Zaltsman et al., 2013, Tzfira et al., 2004b, Zaltsman et al., 2010b) exposing the T-DNA to chromatin. Incorporation of the T-DNA into the host genome is not well understood, however it is suggested that foreign DNA integrates through illegitimate recombination or non-homologous end joining (Paszkowski et al., 1988, Gheysen et al., 1991, Mayerhofer et al., 1991). It is suggested that VirD2 might play a role in integration as it possesses site-specific cleavage and reversal capabilities (Pansegrau et al.,

1993). *In vitro* models including VirD2 and plant extracts were able to mediate T-DNA ligation to plant DNA, while the use of VirD2 exclusively could not (Ziemienowicz et al., 2000). This supports a claim that host proteins are necessary to aid in the ligation of T-DNA integrating into a host genome. Direct mechanisms of integration have been proposed as either a double-strand-break repair (DSBR), or a single-strand-gap repair (SSGR) (Tzfira et al., 2004a).

Because we were ultimately interested in uncovering mechanisms that could lead to strategies to enhance plant productivity, we focused our studies on this truncated T-DNA mutant line exhibiting increased biomass and yield. Due to the increase in biomass and seed yield we refer to this mutant as *mpg1* (*makes plants gigantic-1*). We hypothesized that the bi-laterally truncated T-DNA expression cassette's presence might have been disrupting a functional element in the genome. However, through TAIL-PCR, we localized the T-DNA expression cassette to an intergenic region on chromosome 8. T-DNA's have a bias propensity to integrate into intergenic regions (particularly within promoter, or even the 5'-UTR, or 3'-UTR elements), presumably because they are more transcriptionally active (Alonso et al., 2003, Rosso et al., 2003, Tzfira et al., 2004a). The particular locus of its integration didn't contain any known or annotated functional elements. Further, the integration occurred downstream of the predicted 3'UTR sequences of the two adjacent genes. We hypothesized that the bi-laterally truncated T-DNA insertion resulted in a mutagenic event altering the expression of a gene(s) neighboring the T-DNA insertion site responsible for the *mpg1* phenotype.

In support of our hypothesis, expression analysis of neighboring genes revealed a single gene with noticeably different expression in *mpg1* compared to wild-type plants in leaf tissue. The particular gene with differential expression showed significantly greater expression in *mpg1* than wild-type plants, which showed little to no expression. The gene with heightened expression

in *mpg1* (referred to here as *MPG1*) encodes a transcription factor belonging to the APETALA 2/Ethylene-Responsive Element-Binding Protein (AP2/EREBP) gene superfamily, under the ERF gene family, within the ERF subfamily (Nakano et al., 2006). This gene family is made up of transcription factors that have been shown to be involved with regulation of growth and development, as well as response to stresses (Nakano et al., 2006). There are currently 163 AP2/ERF genes identified in the rice genome (Rashid et al., 2012). A conserved AP2/ERF DNA binding domain, a unique sequence 60-70 residues in size, characterizes the gene superfamily (Riechmann and Meyerowitz, 1998, Sakuma et al., 2002). The AP2/EREBP superfamily is divided into three major families of genes (AP2, RAV, and ERF) (Rashid et al., 2012). The genes are further segregated into four subfamilies based on number of AP2/ERF domains and general functionality (AP2, RAV, DREB, and ERF) (Nakano et al., 2006). The ERF family members characteristically contain a single AP2/ERF domain (Rashid et al., 2012). The ERF subfamily is grouped based upon their encoded transcription factor's expected DNA binding motif. The ERF subfamily binds specifically to the GCC-box *cis*-acting element (GCCGCC). These particular motifs have been found in promoter regions of ethylene-inducible related genes (Ohme-Takagi and Shinshi, 1995).

We have been unable to detect expression of *MPG1* in wild-type plants, however analysis of gene expression databases and previous gene expression studies has provided insight into native expression. MSU Rice Genome Annotation Database's RNA-seq coverage data show expression in numerous tissue types (flowering, embryo, shoot, and leaf tissues) (Ouyang et al., 2007). Specifically, expression of *MPG1* was reported to be within pre- and post- emergence inflorescence, seed 5 days post-pollination, embryo 25 days post-pollination, and 20 days post-planting shoot and leaves. Of these tissues, greater levels of expression were measured in

flowering tissues over vegetative tissues. *MPG1*'s expression has also been seen in response to stress. Increased transcription of this gene has been reported, via micro array and RT-PCR analysis under exogeneous salicylic acid, submergence, laid-down submergence, and cold (Sharoni et al., 2011). Perhaps, *MPG1* plays a role in stress response.

Recent studies have provided further insight into the native function of *MPG1*. *MPG1*, also known as *OsERF#115*, has been found to interact, through yeast-two hybrid analysis, with other transcriptional regulators (*NUCLEAR FACTOR – YB1* (*OsNF-YB1*), and *NUCLEAR FACTOR – YC* (*OsNF- YC11/12*)) to aid in rice endosperm development and grain filling (Xu et al., 2016a). These nuclear factor transcription factors are specifically expressed within the aleurone layer of developing seeds and directly regulate endosperm development and grain filling. RNA-seq was conducted focusing on the effects of *OsNF-YB1* by assessment of *OsNF-YB1*RNAi lines. This generated a list of coregulated genes, which through gene ontology enrichment analysis suggest involvement with transport, ATP synthesis, protein folding, response to stimuli, and metabolic processes. *OsNF-YB1* and *OsNF-YC11/12* form a transcriptional complex with *MPG1*, which is responsible for binding to DNA at appropriate target sequences. Yeast-one hybrid analysis validated the GCC-box (GCCGCC) as the DNA recognition sequence for *MPG1* (Xu et al., 2016a). This study directly shows support that *MPG1* plays a role in development.

To ensure that the increased level of expression of *MPG1* was caused by the bi-laterally truncated T-DNA insertion resulting in the phenotype observed in *mpg1*, and not by a footprinting event generated by agrobacterium-mediated transformation (Castle et al., 1993, Gheysen et al., 1991, Mayerhofer et al., 1991) we continued to track the presence of the bi-laterally truncated T-DNA insertion, elevated expression of *MPG1*, and continued

comprehensive phenotyping of *mpg1* across multiple selfed segregating generations, and several lines of F_2BC_1 backcrosses. We grew plants under optimal, non-optimal, field, and specific stress conditions to further characterize this mutant, and residually better understand the effects of *MPG1*'s expression in *mpg1*.

Our studies revealed that *MPG1*'s expression pattern in *mpg1* continues to correlate with the presence of the bi-laterally truncated T-DNA insertion and phenotype (delayed flowering, increased biomass, greater seed yield, and possibly a degree of stress tolerance) across several generations. It is likely that the over-expression of this specific AP2/EREBP (*MPG1*) is involved with both stress response, growth, and development.

Materials and Methods

Plant materials

Rice (*Oryza sativa* L. spp. *japonica* cv. Kitaake), including wild-type, and segregating lines of partial T-DNA insertion of expression cassette *CmGAS1pro*::*AtSUC1*$_{H65K}$ - mutant (*mpg1*), were used to assess phenotypic metrics, and expression analyses.

Plant growth conditions

Seeds were placed on germination paper and partially submerged in a 1:1000 dilution of MAXIM XL dual action fungicide (Syngenta) and sealed with parafilm. Seeds were incubated at 30º C under 12 h light cycles, until primary shoot and root development occurred (usually 5-7 days). Seedlings were then transferred to planting medium in the greenhouse. Planting medium (non-optimal) consisted of: 1 part play sand, 4 parts Canadian sphagnum peat moss, 4 parts (Promix) BX, mixed to homogeneity. Plants were either transferred to 3.5" pots until 3-leaf stage where they were genotyped then transplanted into 1.0 gallon pots, or directly into 1.0 gallon experimental pots. Pots were organized in random fashion in a flat or tub with water covered

with black plastic and watered until media was fully saturated and pots remained in roughly 3” of standing water. Greenhouse conditions were maintained at 30˚C and 70% RH with a 16h light cycle. Plant chlorosis was monitored and preemptively treated around the 3- to 4-leaf stage using Sprint 330 Iron Chelate at 0.3 g/L water and top-watered. At the same developmental stage plants were fertilized using Scotts Peters Excel 15-5-15 Cal-Mag granular fertilizer at 24.22 g/L water and top-watered. Fertilizer treatment occurred twice weekly until harvest.

The optimal planting media consisted of: 1 part (Profile) Greens Grade porous ceramic particulate, and 1 part (Promix) BX, soil. The contents were mixed to homogeneity, and transferred to 1.0 gallon experimental pots. Pots were organized in random fashion in a flat or tub with water covered with black plastic and watered until media was fully saturated and pots remained in roughly 3” of standing water. Greenhouse conditions were maintained at 30˚C and 70% RH with a 16h light cycle. Plant chlorosis was monitored and preemptively treated around the 3- to 4-leaf stage using (Sprint) 330 Iron Chelate at 0.3 g/L water and top-watered. At the same developmental stage plants were fertilized using granulized (Technigro) 15-5-15 Plus Cal-Mag at 48.87 g/L and top-watered. Fertilizer treatment occurred twice weekly through maturity until harvest.

Experimental growth replicates

Four independent rounds of experimentation were performed using both non-optimal and optimal conditions. Plant growth experiments were performed based on the availability of greenhouse space, thus plants were grown in differing seasons and were subject to different levels of variability concerning pest outbreaks and/or greenhouse regulatory stability. The experimental trials performed are labeled (1-4) for both non-optimal and optimal conditions.

Several stunted variants (not linked to our study, present even in tWT plants) were removed from our analyses as outliers.

The layout for non-optimal trials are as follows. Trial 1 consisted of T_2 plants grown during winter of 2013, and comprised of 72 total plants - 22 tWT, 10 WT-ns, 16 HT-*mpg1*, and 24 HM-*mpg1*. Trial 2 consisted of T_3 and T_4 plants grown during summer of 2013, and comprised of 174 total plants – 0 tWT, 43 WT-ns, 61 HT-*mpg1*, and 70 HM-*mpg1*. Trial 3 consisted of T_4 and T_5 plants grown during spring of 2014, and comprised of 103 total plants – 12 tWT, 11 WT-ns, 24 HT-*mpg1*, and 56 HM-*mpg1*. Trial 4 consisted of T_3 plants grown during fall of 2015, and comprised of 76 total plants – 20 tWT, 20 WT-ns, 17 HT-*mpg1*, and 19 HM-*mpg1*. Cumulative analysis of these trials can be found in (APPENDIX). Due to potential extraneous variability between trials, trial 4 was chosen as a representative population for analysis and description because of its healthy performance, lack of environmental variability, experimental sample size, and strong similarity between other trials. Similarly trial 3 was selected as a representative for analysis of leaf length and width, as this measurement was not evaluated during trial 4.

The layout for optimal trials are as follows. Trial 1 consisted of T_3 and T_4 plants grown during fall of 2014, and comprised of 86 plants – 12 tWT, 24 WT-ns, 38 HT-*mpg1*, and 22 HM-*mpg1*. Trial 2 consisted of T_3 plants grown during summer of 2016, and comprised of 72 plants – 0 tWT, 11 WT-ns, 35 HT-*mpg1*, and 26 HM-*mpg1*. Trial 3 consisted of T_4 plants grown in summer of 2017, and comprised of 80 plants – 20 tWT, 19 WT-ns, 29 HT-*mpg1*, and 12 HM-*mpg1*. Trial 4 consisted of T_4 plants grown during fall of 2017, and comprised of 75 plants – 20 tWT, 12 WT-ns, 25 HT-*mpg1*, and 18 HM-*mpg1*. Cumulative analysis of these trials can be found in (APPENDIX). Due to potential extraneous variability between trials, trial 3 was chosen

as a representative population for analysis and description because of its healthy performance, lack of environmental variability, experimental sample size, and strong similarity between other trials. Similarly trial 4 was selected as a representative for analysis of leaf length and leaf width, as this measurement was not evaluated during trial 3.

DNA extraction and genotyping

Young, fresh plant tissue (3-leaf stage) was sampled for DNA extraction and analysis (2-5cm of leaf-tip). DNA was obtained via mechanical disruption of tissue and Shorty-Buffer extraction. Tissue was flash frozen in liquid nitrogen and disrupted using the Qiagen TissueLyser at 30 rps for a period of 1 minute. Five hundred μL of freshly prepared shorty buffer (0.2 M Tris HCl pH 9.0, 0.4 M LiCl, 25 mM EDTA, and 1.0% SDS) was added to each tissue sample, vortexed and centrifuged at max speed (13k rpm) for 5 min. Then 350 μL of supernatant was transferred to a new tube containing 400 μL isopropanol, mixed by inverting and centrifuged at max speed for 10 min. The supernatant was discarded and 1mL of 70% ethanol was added to each sample to wash the DNA pellet. Samples were then centrifuged at max speed for 10 minutes, the supernatant was discarded and the tubes were inverted for 30 minutes. The DNA pellet was resuspended in 100 μL of 10mM Tris-HCl pH 8 and stored short term at 4º C until use.

Haplotyping mutants

To identify *mpg1* plants that were homozygous, heterozygous, or null segregants for the bi-laterally truncated T-DNA insertion, primers were designed in regions directly flanking the site of insertion, as well as primers spanning the integration site into the T-DNA, and used in PCR. The primers flanking the T-DNA insertion were wFLA forward: 5'-GGAAGTTGGAGATGGGAAACA-3', and wFLA reverse: 5'-

GGCCTCGTGTGTCAGTAATAA-3'. The primers spanning the genomic region and the T-DNA insertion were wIN forward: 5'-ACACCGGAAGCATAGTCATTT-3', and wIN reverse: 5'-GGTCGCCAACATCTTCTTCT-3'.

RNA-extraction and gene expression analysis

Desired tissue from both stem, leaf, and root across development from various selfed and backcrossed populations (not more than 100 mg) was sampled for RNA extraction and analysis. Tissue was placed in individual 2 mL tubes and flash frozen in liquid nitrogen. Tissue was ground using the TissueLyser (Qiagen) at 30 rps for 1 min, and RNA was extracted using (Qiagen) Plant RNeasy mini-kit. RNA was treated with DNase and purified using the Turbo DNase kit (Invitrogen). cDNA was synthesized from 1 μg of RNA using SuperScript (Invitrogen).

Specific primers used for expression analysis via RT-PCR for the individual candidate gene of LOC_Os08g41030 (*MPG1*) are 41030 forward: 5'-TCGCCATTGTTCAGCAAGAAGGA-3', and 41030 reverse: 5'-AAGTGCATGACCAAGTACAGA-3'. Housekeeping control primers were designed around actin, more specifically the sequences are, actin forward: 5'-GAGTATGATGAGTCGGGTCCA-3', and actin reverse primer: 5'-ACACCAACAATCCCAAACAGA-3'. PCR was performed in 20 μL reactions using Econo Taq polymerase (Lucigen), 2 μL of cDNA and desired primers under a normal 30-cycle amplification protocol. PCR products were analyzed by electrophoresis on 1% gel-agarose containing ethidium bromide.

Field trial

Both WT-ns and HM-*mpg1* seeds were germinated and plants were grown under optimal conditions in a local greenhouse (Fort Collins, CO) where they were subjected to a period of

'hardening' by exposure to fans, which were rotated periodically for several days. In June, when plants were around the 4-leaf stage they were transferred to a field plot at Colorado State University's Agricultural Development & Education Center (ARDEC) in Fort Collins, CO. The soil was pre-fertilized and plants were grown to maturity. Plants were irrigated weekly with supplemental hand-watering as necessary. It is important to note that plants were harvested prior to *mpg1*'s ability to fully progress through panicle development and seed filling due to inclement cold weather at the end of the growth experiment.

Backcross population

Rice panicles were assessed during the emerging stage of flowering. To generate backcross lines, green-seed spikelets from one tWT individual were cut in half prior to the milk stage, and de-masculinated by removing all interior pollen. Pollen was then taken from HM-*mpg1* plants and applied via shaking onto the de-masculinated spikelets. Pollinations were bagged for the remainder of panicle development and seed filling stages. Three independent crosses were successful. These three F_1BC_1 plants were genotyped and grown to maturity and seed was collected. The subsequent segregating F_2BC_1 population was then grown under optimal conditions and phenotypically characterized.

Phenotypic analyses

Plants were grown to maturity and several metrics were recorded. During plant growth and development height, tiller number, and girth were measured several times weekly from shortly after planting through to end of vegetative growth. Plant height was measured as the length from the planting media to the tip of the tallest leaf. Tiller number was recorded by counting number of stems with true leaves present. Girth was measured as length of circumference at roughly 5 cm above soil line. Time to flowering was measured by the number

of days post-planting (dpp) until panicle emergence (or heading). At harvest height, tiller number, and girth were assessed again. In addition, leaf length, and leaf width were measured. These characteristics were taken from the leaf that provided the highest point of the plant and measured its length from tip to culm, and the width at its widest point. Additionally, panicles inferior to the panicle neck were harvested, dried for 7 days in a 45º C drying oven and weighed to determine seed yield. Plants with panicles removed were cut roughly 5 cm above soil line, dried at 45º C for 7-14 days and weighed for a measure of total biomass. Total seed yield with panicle tissue was recorded, as well as number of panicles. Panicles were also monitored for the presence and length of awn development. After harvest residual plant matter remained in growing conditions and ratooning was assessed. Measurements were taken at 41 dpp, with preliminary assessment at a second harvest.

Various measurements were recorded for preliminary analysis of panicle architecture (APPENDIX). During individual experiments, a 1000 grain seed count was performed by removing the hull and weighing 1000 grains; an assessment of spikelet number per panicle, panicle length, and panicle branch number were measured; and a qualitative observation of panicle length and density were also noted via pictorial assessment.

Stress trials

Plants were grown under optimal growth conditions in 3.5” pots randomly assortment in flats until the 3-leaf stage of development. The plants undergoing salinity stress at 3-leaf stage were then watered with 100mM NaCl via removal of previous water and re-watering with NaCl solution. These flats were continually re-filled with 100mM NaCl and the plants were grown to maturity under these salinity conditions. Plants undergoing the drought stress were grown to the 3-leaf stage under optimal conditions at which point water was removed from the flat until plants

experienced leaf curling. Upon visualization of leaf curling water was re-introduced to the plants until leaf curling subsided, then water removal was reintroduced. The process of water removal and rehydration continued through plant maturity.

***Xanthomonas* pathogen assay (APPENDIX)**

Xanthomonas oryzae pv. *oryzae* strain PXO86 was grown on PSA plates at 28 C for 2 days. A new PSA plate was then inoculated from the first to make a lawn. From the lawn plate a loop-full of PXO86 was transferred in to ~6.0mL of DI H_2O and was mixed via inversion. The resulting suspension concentration was determined and manipulated using a spectrophotometer till the OD at 600nm = 0.2. Scissors were dipped in the suspension and used to clip inoculate 2 leaves on a single plant at 21 days post planting. Eight of each type of plant (tWT, WT-ns, HT-*mpg1*, and HM-*mpg1*) were inoculated. Half of the plants were used to gauge gene expression of the candidate gene LOC_Os08g41030 (*MPG1*) by RT-PCR while the other half of the plants were used to assess the PXO86 stress challenge. Samples for assessment of gene expression were taken at time-point 0 (pre- inoculation), 6 hours post-inoculation, and 12 hours post-inoculation. Samples were taken from one of the two inoculated leaves on each plant. At the 6 hour post-inoculation time-point, wild-type plants showed increased expression of candidate gene LOC_Os08g41030 (*MPG1*). The plants that weren't sampled for gene expression were left for a period of 12 days. The degree of the PXO86 infection was determined by the length of the lesion formed from the cut inoculation site.

Statistical analyses

Statistics were calculated for growth metrics using a one-way ANOVA and Games-Howell post-hoc multiple comparison test at 95% family-wise confidence level comparing all treatment groups in R using the 'userfriendlyscience' package. The field trial and ratooned

backcrossed populations were evaluated using a Student's t-test. Graphical models were generated using boxplot in R, Microsoft Excel, and Powerpoint.

Results

mpg1 plants grown under non-optimal conditions accumulate more biomass and seed yield compared to wild-type plants

Selfed segregating populations of *mpg1* plants were grown to continue assessment of the phenotype and its correlation to the presence of the bi-laterally truncated T-DNA insertion and elevated expression of *MPG1*. During initial experiments growing these plants for a more comprehensive phenotypic assessment surrounding enhanced plant productivity we observed a state of stress on the experimental population. Specifically, wild-type Kitaake were experiencing symptoms of stress (smaller plant size and paler complexion). Although *mpg1* plants in this population didn't present with these symptomatic indicators we investigated potential sources of stress. We concluded that potential sources of stress might have arisen from our growth media and fertilizer treatments, as we visibly noticed salt accumulation in the growing tubs. The growth media we were using contained 1 part (Quikrete) Play Sand, 4 parts (Sungro) Canadian Sphagnum peat moss, 4 parts (Promix) BX soil. The fertilizer used was (Scott's) 15-5-15 CalMag. Because we visibly noticed salt accumulation in the plant basin/tubs we hypothesized that this combination of growth media and nutrients was somehow resulting in stressful conditions (salt and pH). Assessments of the element blend in the fertilizer suggested that the nitrate-nitrogen concentration might result in an acidic reaction in growth media (Pierre, 1928, Hedley and Bolan, 2003) potentially resulting in lower than optimal pH levels for rice growth. pH analysis of the fertilizer prior to application (100 ppm N using a Dosatron) registered around a pH of 3.0. Measurements of the growth media after fertilizer treatment revealed a pH range of

5.3-5.6, and an electrical conductivity (EC) range of 1.6-3.23 mS/cm during late vegetative growth. The optimal threshold rice is a pH of 5.5-6.5 (Yu, 1991, Mosaic, 2018) and EC of 1.7-2.1 (USDA, 2011, Mosaic, 2018). This suggests that the plants might have been under stress as these conditions exceed their optimum parameters, especially in terms of salt concentration. The cause of the stressful conditions is likely multi-factorial, consisting of higher than optimal concentrations of fertilizer, soil-media accumulating nutrients from lack of leaching, and plants continually maturing and using less fertilizer. The exact cause, and nature of the stress observed in these growth conditions remains unknown.

Comprehensive phenotyping of *mpg1* plants began by growing plants from seedling to maturity in a controlled greenhouse and measuring fully matured plants at harvest. As previously stated, during several rounds of this experiment some of the wild-type plants appeared to be stressed (general unhealthy appearance – plants were smaller and more pale compared to optimally grown WT-ns and tWT Kitaake). Further evaluation of our planting media and fertilizer combination revealed that their use potentially resulted in salt accumulation and acidic conditions, which could have been impacting the growth and development of these plants. It is important to note that these specific conditions could be sufficient in generating a stressful environment, or it could be that this nutrient combination creates conditions resulting in a stressful or sub-optimal environment for wild-type plants.

Interestingly, *mpg1* plants didn't experience any of these visually symptomatic characteristics. These non-optimal conditions were repeated several times with some variation, but saw the same overall trend between phenotypes in all of the replications (APPENDIX). Here we selected one of the replicate trials as a representative population. Under these non-optimal conditions *mpg1* plants accumulate more biomass and seed yield compared to wild-type plants

(Figure 3.2, 3.3). Total above ground biomass (dry weight) was significantly greater in *mpg1* plants compared to wild-type plants (~5.7-fold HM-*mpg1* vs. tWT). Biomass metrics were broken down into several components. Of these, *mpg1* plants were significantly taller (~1.3-fold HM-*mpg1* vs. tWT), had more tillers (~1.9-fold HM-*mpg1* vs. tWT), and had greater total plant girth (~1.9-fold HM-*mpg1* vs. tWT) compared to wild-type plants (Figure 3.2). Assessment of time until flowering was performed as well. *mpg1* exhibited a delay in flowering compared to wild-type plants (~23 days HM-*mpg1* vs. tWT) (Figure 3.3). Panicle number (~1.7-fold HM-*mpg1* vs. tWT) and overall seed yield (~3.0-fold HM-*mpg1* vs. tWT) in *mpg1* was also significantly greater compared to wild-type plants (Figure 3.3).

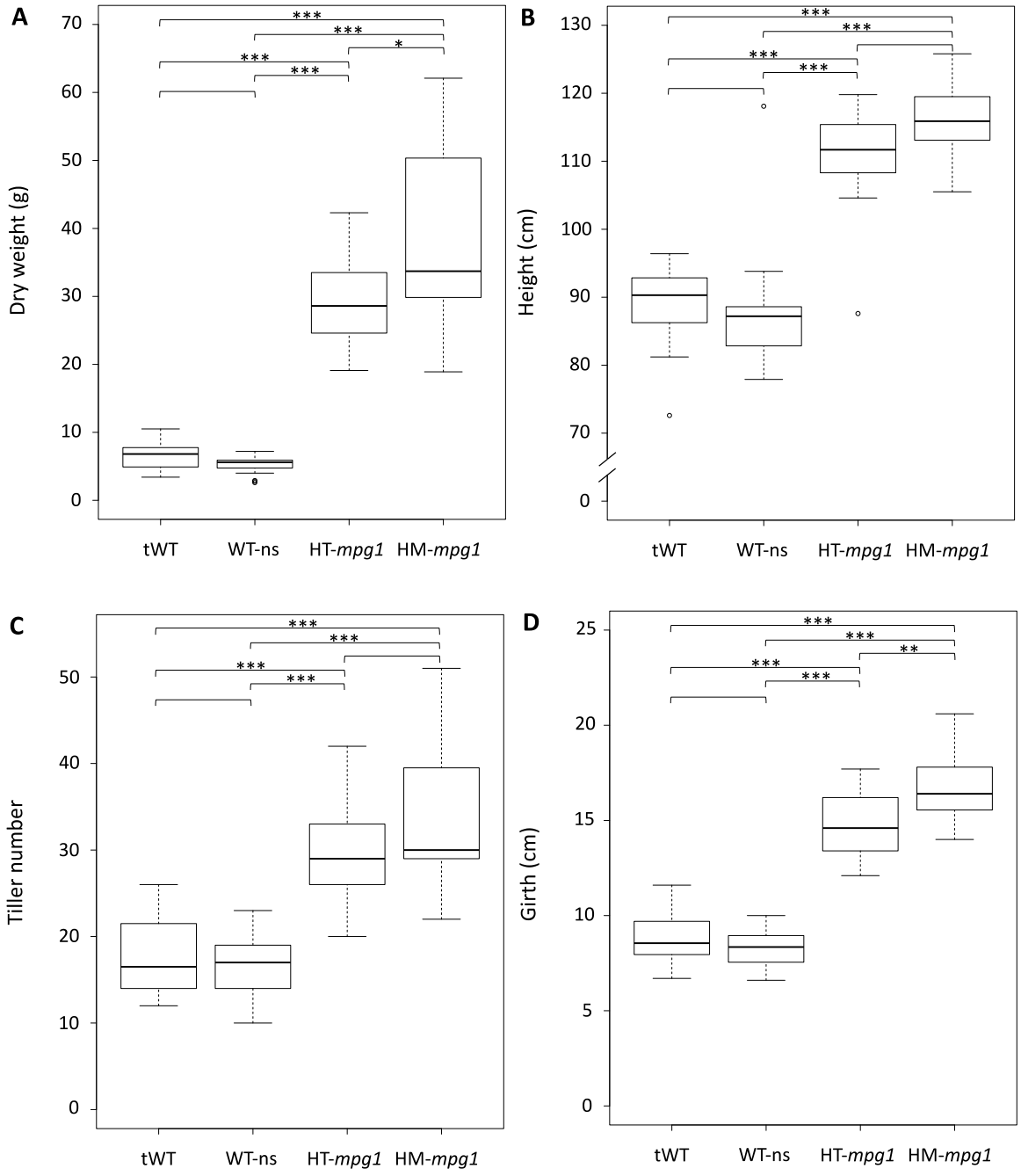


Figure 3.2: Analysis of biomass-related characteristics within a representative segregating population of T_3 *mpg1* grown under non-optimal conditions.
(A) dry weight, (B) height, (C) tiller number, and (D) girth. tWT (n = 20, 19 for height), WTns (n = 20), HT-*mpg1* (n = 17), HM-mpg1 (n = 19, 18 for height). Analysis was conducted using a one-way ANOVA and Games-Howell post-hoc multiple comparison test at 95% family-wise confidence level. '*' indicates $p<0.05$, '**' indicates $p<0.01$, '***' indicates $p<0.001$.

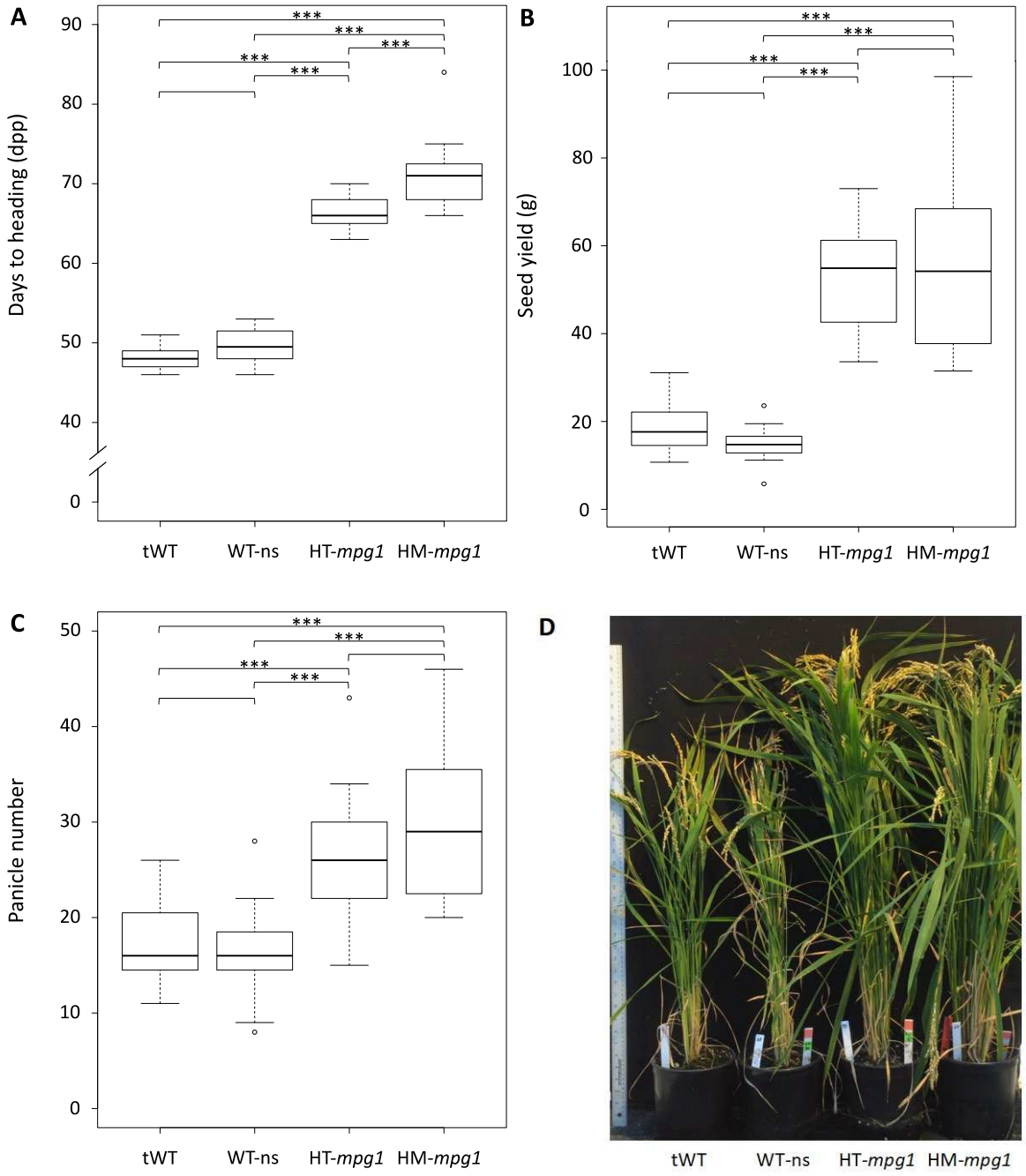


Figure 3.3: Analysis of seed yield-related characteristics within a representative segregating population of T_3 *mpg1* grown under non-optimal conditions.
(A) days to heading measured by days post-planting, (B) seed yield, and (C) panicle number. (D) a photo of each of the treatment groups side-by-side, scale = 1m. tWT (n = 20, 19 for panicle number), WTns (n = 20, 19 for panicle number), HT-*mpg1* (n = 17), HM-mpg1 (n = 19). Analysis was conducted using a one-way ANOVA and Games-Howell post-hoc multiple comparison test at 95% family-wise confidence level. '***' indicates $p<0.001$.

In other replicates of comprehensive phenotyping of plants grown under non-optimal conditions, leaf characteristics were also measured. *mpg1* plants exhibited longer leaves (~1.3-fold HM-*mpg1* vs. tWT), and wider leaves (~1.2-fold HM-*mpg1* vs. tWT) compared to wild-type plants (Figure 3.4). HM-*mpg1* plants usually measured slightly greater across all metrics compared to HT-*mpg1*, implying that a dosage effect might exist. Some traits were greater than others, however some differences were not significant suggesting that the effect of overexpression reaches a saturation limit in plants that are homozygous for the T-DNA insertion for some phenotypic characteristics. The insertion segregates 3:1 phenotypically as a dominant mutation. The phenotyping results indicate that *mpg1* plants experience a delay in flowering, have an increase in in biomass and seed yield, and possibly have a degree of abiotic stress tolerance.

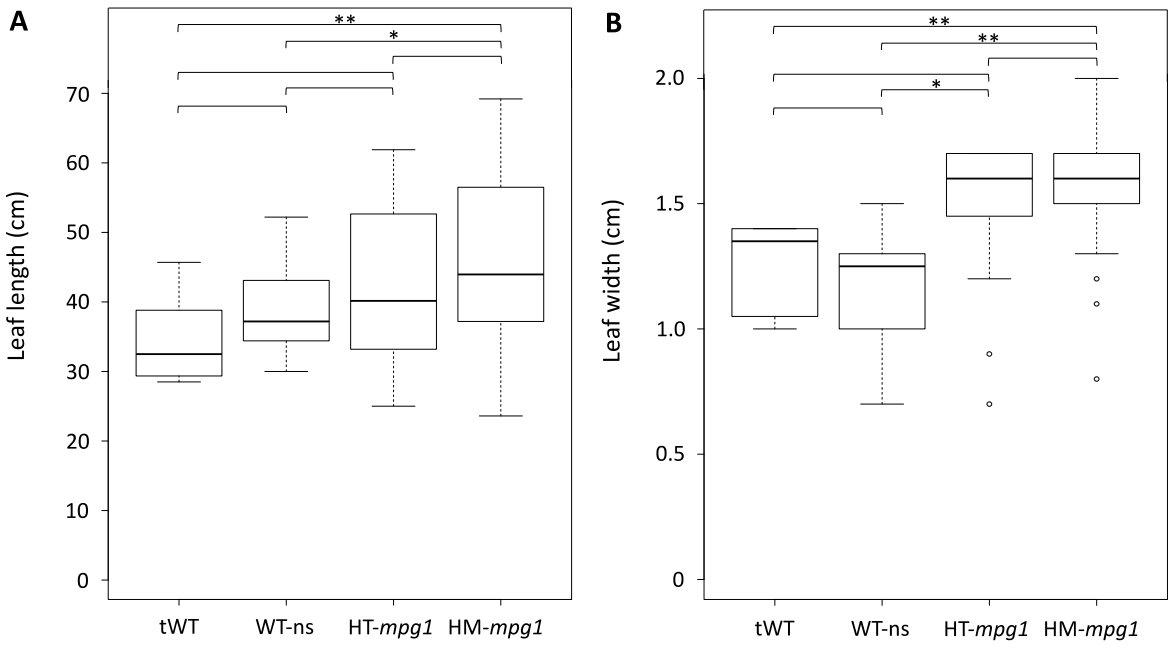


Figure 3.4: Analysis of leaf characteristics within an alternate representative segregating population of T_4 and T_5 *mpg1* grown under non-optimal conditions (trial 3).
(A) Leaf length measurements, (B) leaf width measurements taken at harvest. tWT (n = 22), WT-ns (n = 10), HT-*mpg1* (n = 16), HM-mpg1 (n = 24). Analysis was conducted using a one-way ANOVA and Games-Howell post-hoc multiple comparison test at 95% family-wise confidence level. '*' indicates $p<0.05$, '**' indicates $p<0.01$.

mpg1 plants grown under optimal conditions accumulate more biomass and seed yield compared to wild-type plants

A new fertilizer and media regiment was adopted that yielded optimum growth for wild-type rice. After trying different combinations of fertilizer and growth media (data not shown) we settled on using a soil medium consisting of 1 part (Profile) Greens Grade, and 1 part (Promix) BX, and (Technigro) 15-5-15 CalMag+. The use of this combination removed any symptoms of stress observed under the non-optimal growth conditions. Both wild-type and *mpg1* segregants grown under these optimum conditions were able to achieve greater levels of biomass accumulation, and seed yield compared to the non-optimal growth conditions. This indicated that the plants grown under the non-optimal conditions were indeed undergoing stress induced decreases in growth.

Comprehensive phenotyping of *mpg1* plants continued after developing optimal growth conditions. We changed the planting media and fertilizer to try and remove the deleterious factors seen in our first planting media and fertilizer combination (possibly salt accumulation and acidic conditions). These plants were again grown from seedling to maturity in a controlled greenhouse taking measurements during growth, at harvest, and after ratooning. All of the plants in these trials appeared healthy throughout the experiment. These optimal conditions were repeated several times (APPENDIX) with some variation, but showed the same overall trend between phenotypes in all of the replications. Here we selected one of the replicate trials as a representative population. Under these optimal conditions *mpg1* plants continued to accumulate more biomass and seed yield compared to wild-type plants (Figure 3.5, 3.6). Total above ground biomass (dry weight) was significantly higher in *mpg1* plants compared to wild-type plants (~2.7-fold HM-*mpg1* vs. tWT).

Biomass metrics were broken down into several components. Of these, *mpg1* plants were significantly taller (~1.1-fold HM-*mpg1* vs. tWT), and had greater total plant girth (~1.9-fold HM-*mpg1* vs. tWT) compared to wild-type plants (Figure 3.5, 3.9). Tiller number was unaffected between *mpg1* plants and wild-type under optimum conditions. *mpg1* showed a delay in flowering compared to wild-type plants (~14 days later HM-*mpg1* vs. tWT) (Figure 3.6), however analysis of multiple replicates suggests the delay in flowering is closer to roughly 20 days (APPENDIX). Overall seed yield in *mpg1* was also significantly greater compared to wild-type plants (~1.6-fold HM-*mpg1* vs. tWT), however panicle number was not greater in *mpg1* under optimal conditions. This might suggest that there is a seed yield increase due to increased spikelet number per panicle.

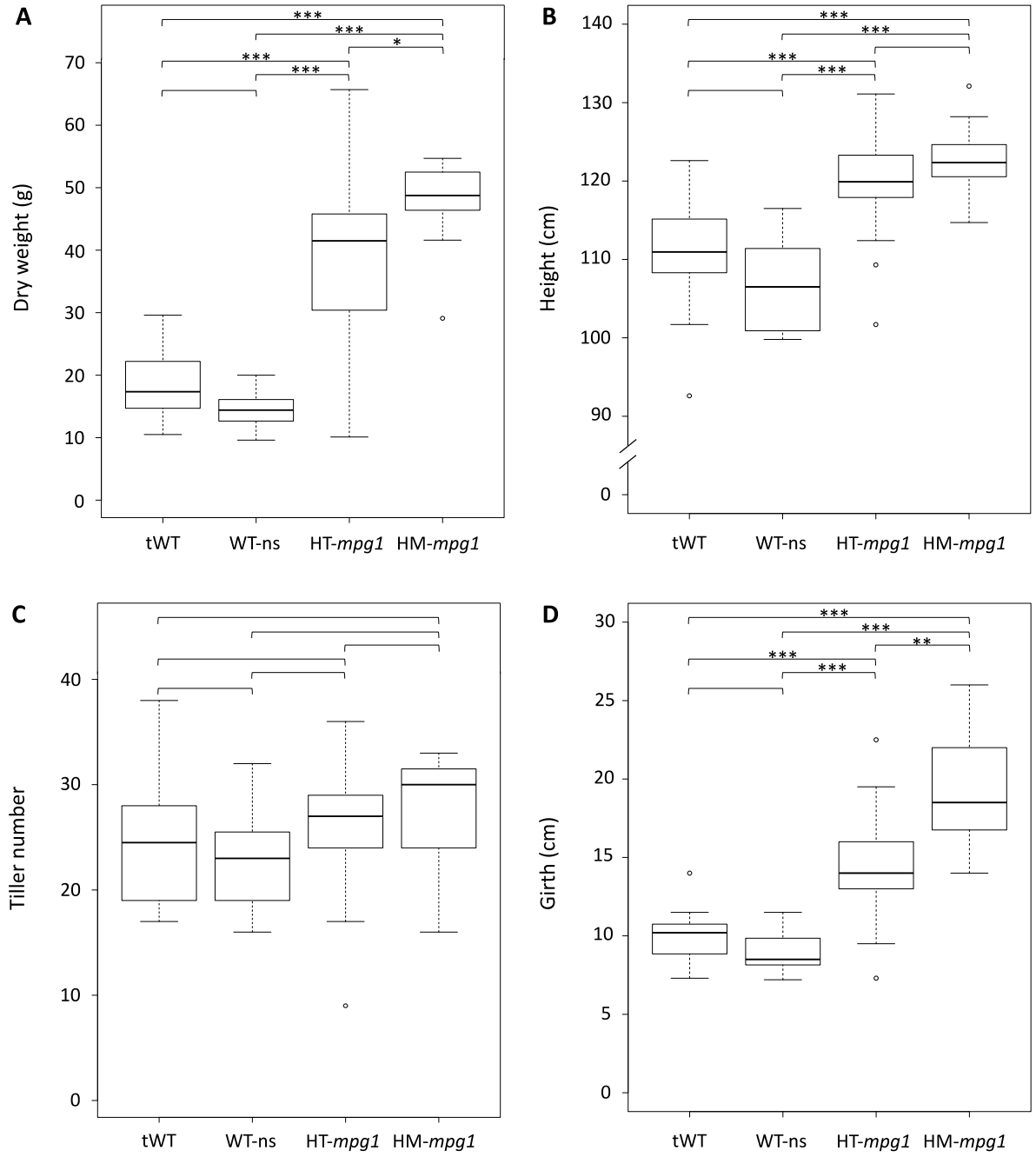


Figure 3.5: Analysis of biomass-related characteristics within a representative segregating population of T_4 *mpg1* grown under optimal conditions.
(A) dry weight, (B) height, (C) tiller number, and (D) girth. tWT (n = 20), WTns (n = 19), HT-*mpg1* (n = 29), HM-mpg1 (n = 12). Analysis was conducted using a one-way ANOVA and Games-Howell post-hoc multiple comparison test at 95% family-wise confidence level. '*' indicates $p<0.05$, '**' indicates $p<0.01$, '***' indicates $p<0.001$.

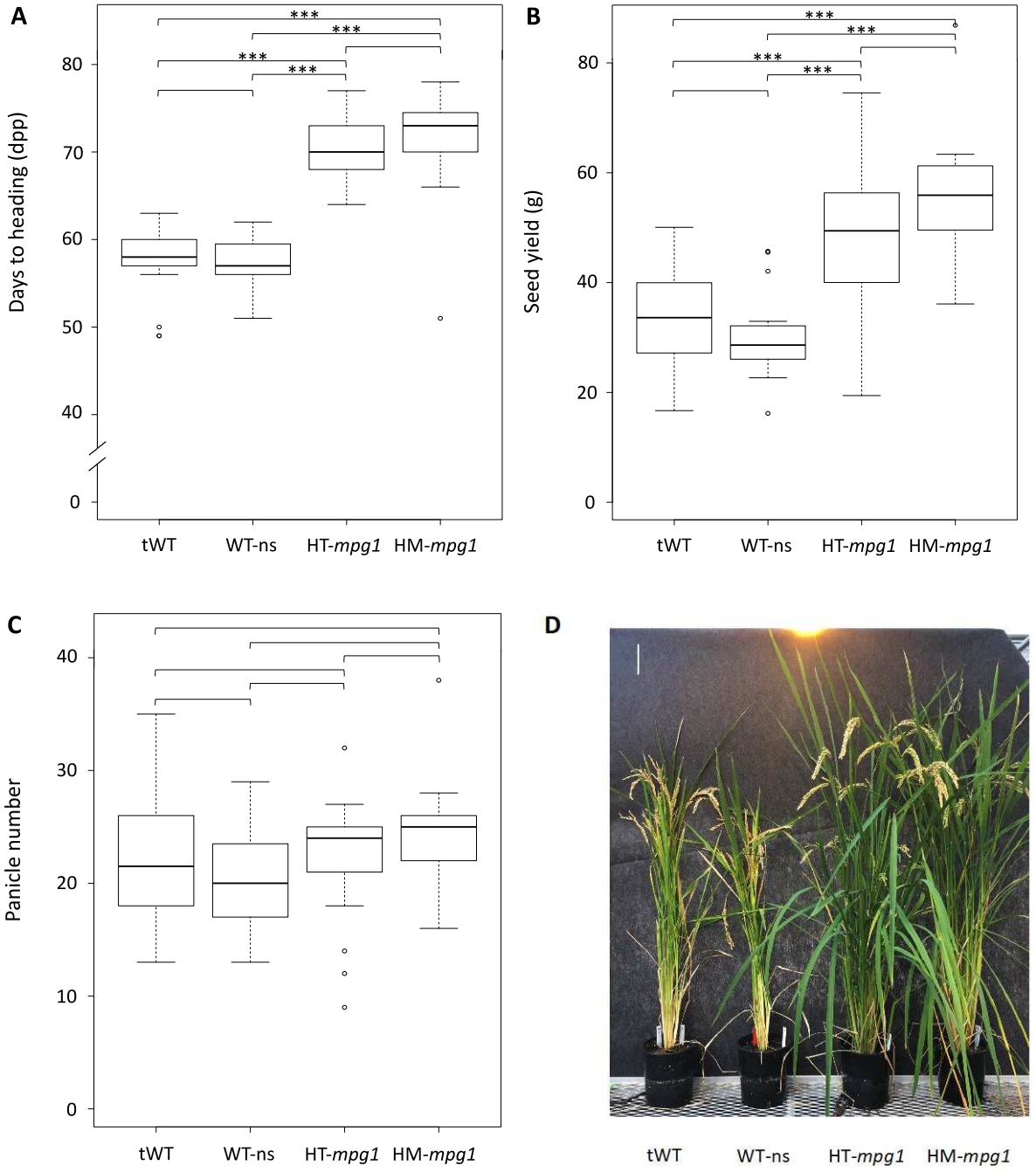


Figure 3.6: Analysis of seed yield-related characteristics within a representative segregating population of T_4 *mpg1*, and photo of plants grown under optimal conditions. (A) days to heading measured by days post-planting, (B) seed yield, and (C) panicle number. (D) a photo of each of the treatment groups side-by-side, scale = 10 cm. tWT (n = 20), WTns (n = 19), HT-*mpg1* (n = 29), HM-mpg1 (n = 12). Analysis was conducted using a one-way ANOVA and Games-Howell post-hoc multiple comparison test at 95% family-wise confidence level. '***' indicates $p<0.001$.

In other replicates of comprehensive phenotyping of plants grown under optimal conditions, leaf characteristics were also measured. *mpg1* plants exhibit wider leaves (~1.2-fold HM-*mpg1* vs. tWT) compared to wild-type plants (Figure 3.7). In all of the biomass and seed measurements recorded, *mpg1* and wild-type plants both saw better growth and higher values within these metrics under the optimum conditions compared to the non-optimal conditions (Figure 3.2-7). The degree of difference between *mpg1* and wild-type measurements was greater under non-optimal conditions compared to optimal conditions, again suggesting that *mpg1*'s phenotype might also be linked to stress.

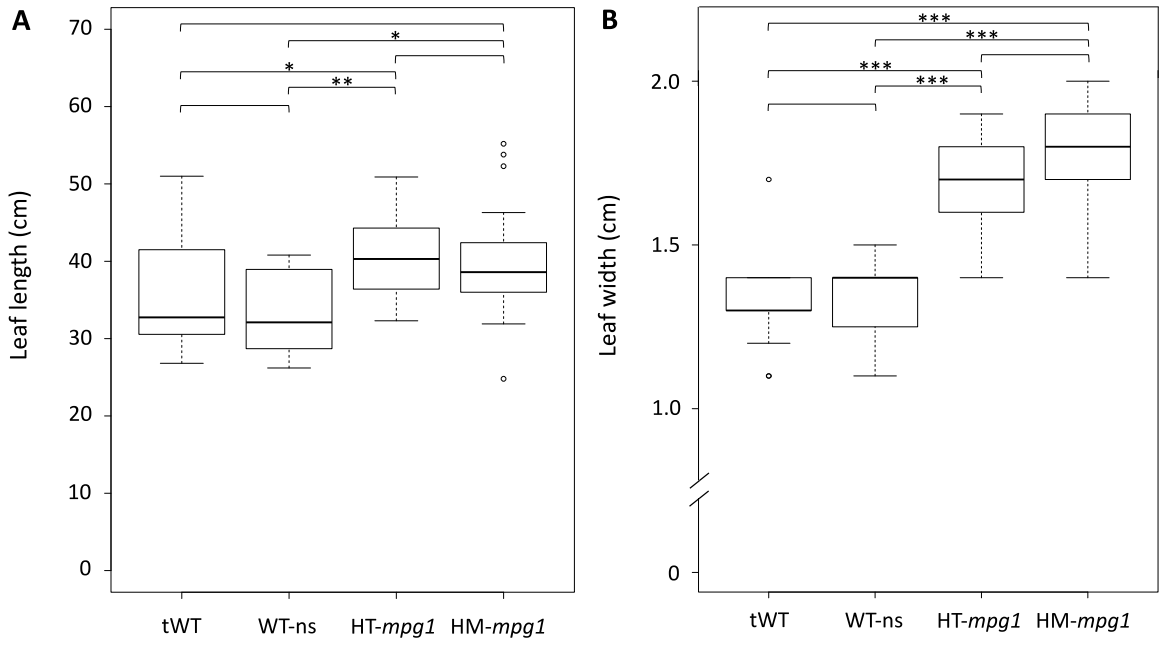


Figure 3.7: Analysis of leaf characteristics within an alternate segregating population of T_4 *mpg1* plants grown under optimal conditions (trial 4).
(A) Leaf length measurements, (B) leaf width measurements taken at harvest. tWT (n = 20), WTns (n = 12), HT-*mpg1* (n = 25), HM-mpg1 (n = 18). Analysis was conducted using a one-way ANOVA and Games-Howell post-hoc multiple comparison test at 95% family-wise confidence level. '*' indicates $p<0.05$, '**' indicates $p<0.01$, '***' indicates $p<0.001$.

Measurements were also taken over time to assess the growth rate of *mpg1* relative to wild-type plants. Height, tiller number, and girth were recorded several times a week during vegetative development and again during harvest. These measurements were observed alongside flowering time to gauge growth respective to total vegetative growth time (Figure 3.8). Height in both *mpg1* and wild-type plants remains relatively similar until about 63 days post-planting (roughly 6 days after panicle heading of wild-type plants), at which time *mpg1* superseded height and continued to grow slightly through seed maturation. Tiller number in both *mpg1* and wild-type plants remains roughly the same until about 46 days post-planting (roughly 11 days prior to panicle heading of wild-type plants), at which point *mpg1* superseded wild-type and continued to grow till about 56 days post-planting (roughly 16 days prior to panicle heading of *mpg1* plants). Girth in both *mpg1* and wild-type plants remained similar until about 42 days post-planting (roughly 15 days prior to panicle heading of wild-type plants), when *mpg1* superseded wild-type plants and continued to grow till about 63 days post planting (roughly 9 days prior to panicle heading of *mpg1* plants). The total accumulation of increased biomass results from an increase in multiple traits: height, girth, and leaf width. Out of all of the biomass characteristics measured, girth has the greatest difference compared to the other traits. Biomass accumulation in relation to flowering time reveals that *mpg1* is able to accumulate more growth prior to panicle heading of wild-type plants. This suggests that *mpg1* is able to generate more tissue than wild-type plants early in vegetative growth while also having a more prolonged period of growth time dually adding to its enhanced growth compared to wild-type. Further investigation into the true time of vegetative to reproductive transition for both *mpg1* and wild-type plants will need to be determined to assess whether or not *mpg1*'s biomass phenotype is solely a result of increased vegetative growth time alone.

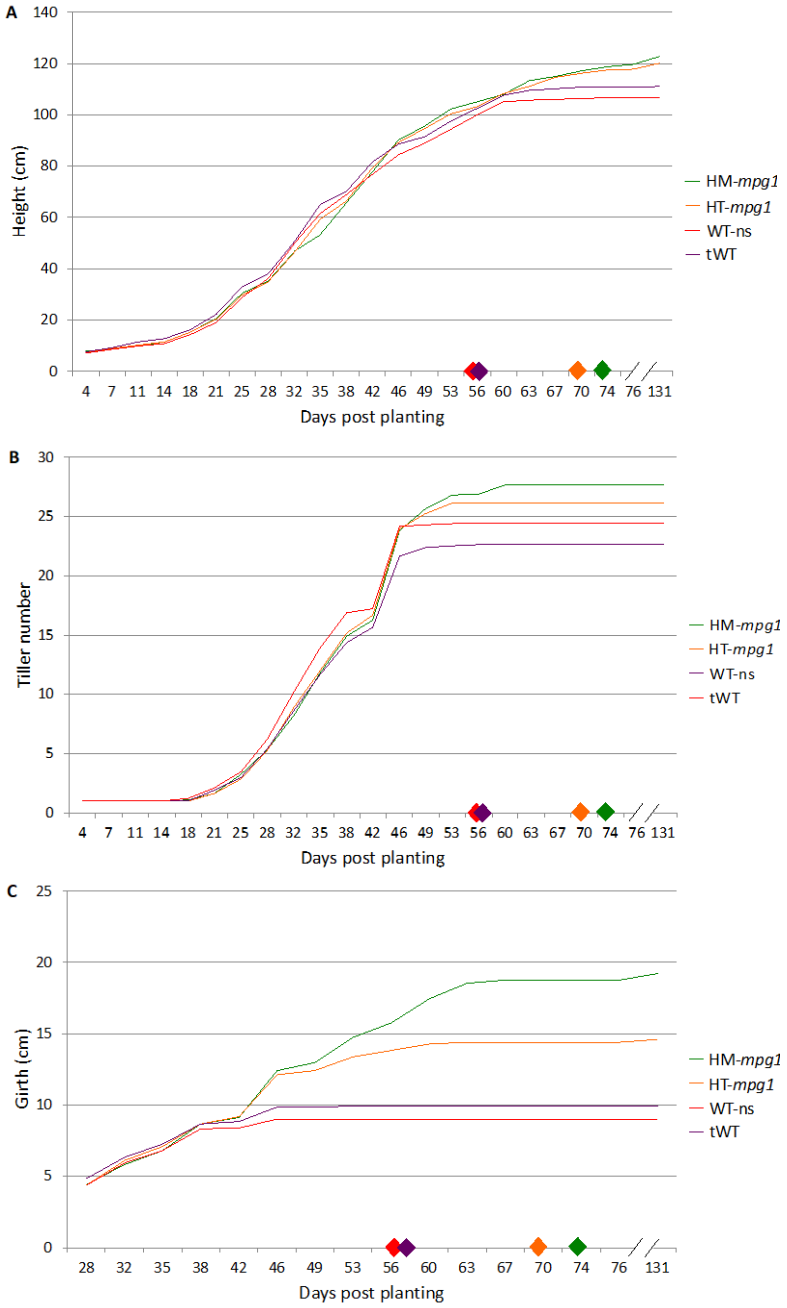


Figure 3.8: Analysis of growth metrics over time within a representative segregating population of T_4 *mpg1* grown under optimal conditions.
Growth parameters pertaining to biomass accumulation were monitored throughout the growth of plants. Specific measurements consisted of (A) height, (B) tiller number, and (C) girth. tWT (n = 20), WTns (n = 19), HT-*mpg1* (n = 29), HM-mpg1 (n = 12). Diamonds correlating with their line color represent the time at which heading occurred for that treatment relative to growth.

After harvesting *mpg1* we observed that it was able to regenerate more tissue and at a quicker rate than wild-type plants after harvesting mature plants (Figure 3.9). This is a process known as ratooning. Ratooning involves the cutting/removal of stem tissue from grasses, leaving residual stubble and subterranean buds, with the intent that the residual plant matter will regenerate seed-bearing tissues for an additional harvest. Even as short as 2 days post-harvest *mpg1* plants were able to regenerate more tissue than wild-type plants (Figure 3.9).

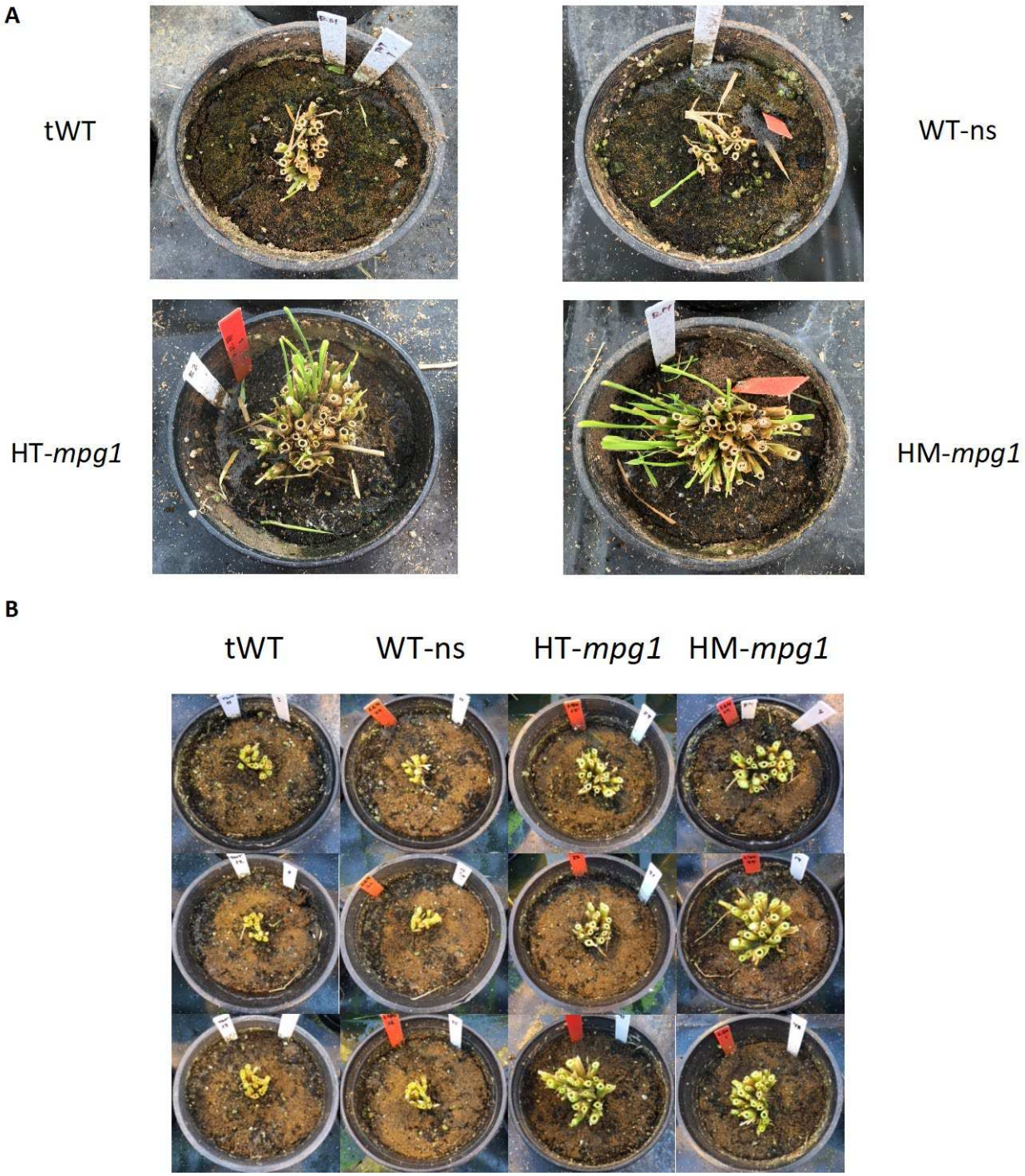


Figure 3.9: Analysis of initial ratooning and final girth characteristics within a representative and additional segregating populations of T_4 *mpg1* grown under optimal conditions.

(A) Photographs of plants 2 days post-harvest. The beginning of ratooned growth can be visualized. (B) Photographs of plants immediately after harvesting, differences in girth and tiller size can be visualized.

The same optimally grown representative population was harvested and left to ratoon. At 41 days post-planting the population was measured again for height, tiller number, girth, and panicle number (Figure 3.10, 3.11). The survivability rate of *mpg1* were much higher than wild-type plants. All of the *mpg1* plants survived while only 65% of tWT, and 42% of WT-ns survived. Of the surviving plants *mpg1* plants were significantly taller (~1.4-fold HM-*mpg1* vs. tWT), have a greater number of tillers (~4.3-fold HM-*mpg1* vs. tWT), have greater girth (~3.3-fold HM-*mpg1* vs. tWT), and more panicles (~7.2-fold HM-*mpg1* vs. tWT) than wild-type plants. Again, these results similarly follow the phenotypic pattern seen in other trials of experimentation assessing ratooning (APPENDIX).

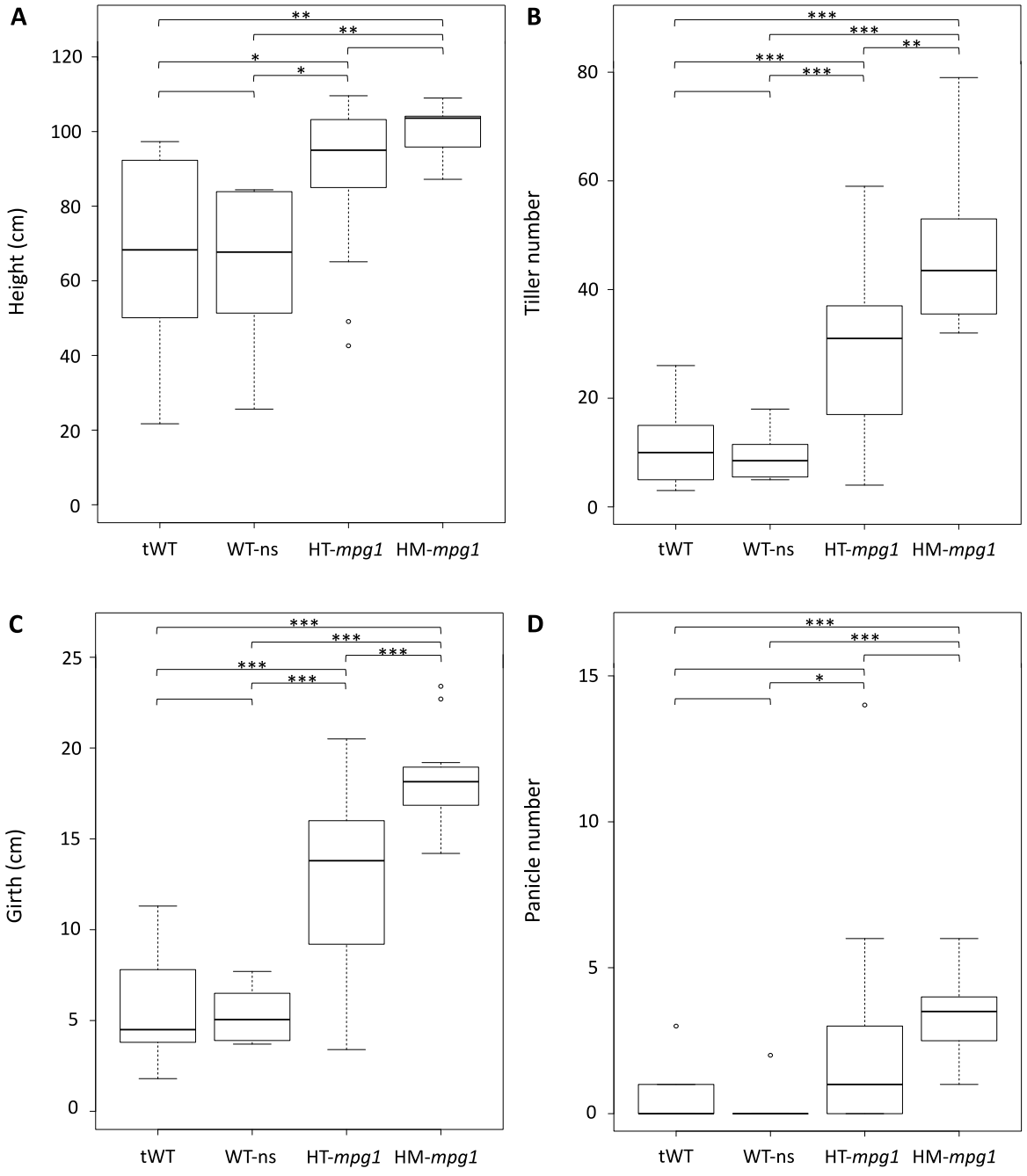


Figure 3.10: Analysis of ratooning at 41 days post-harvest within a representative segregating population of T_4 *mpg1* grown under optimal conditions.
Biomass and seed-related traits were assessed at 41 days post-harvest. All of the *mpg1* plants survived while only 65% of tWT, and 42% of WT-ns survived. Of the surviving plants measurements of (A) height, (B) tiller number, (C) girth, and (D) panicle number were assessed. tWT (n = 13), WTns (n = 8), HT-*mpg1* (n = 29), HM-mpg1 (n = 12). Analysis was conducted using a one-way ANOVA and Games-Howell post-hoc multiple comparison test at 95% family-wise confidence level. '*' indicates $p<0.05$, '**' indicates $p<0.01$, '***' indicates $p<0.001$.

Figure 3.11: Photograph of ratooning plants at 41 days post-harvest within a representative segregating population of T_4 *mpg1* grown under optimal conditions. Scale = 10 cm

Ratooned *mpg1* plants flowered earlier than wild-type plants, and also accumulated more biomass by 41 days post-planting. These same plants remained growing until maturity and were measured again at second harvest (Figure 3.12). Of the biomass metrics, *mpg1* had significantly greater dry biomass (~6.0-fold HM-*mpg1* vs. tWT), height (~1.5-fold HM-*mpg1* vs. tWT), tiller number (~4.0-fold HM-*mpg1* vs. tWT), girth (~2.4-fold HM-*mpg1* vs. tWT), leaf length (~1.4-fold HM-*mpg1* vs. tWT), and leaf width (~1.1-fold HM-*mpg1* vs. tWT). Out of all the biomass characteristics measured, height and girth showed the greatest difference between wild-type and *mpg1* compared to the other traits.

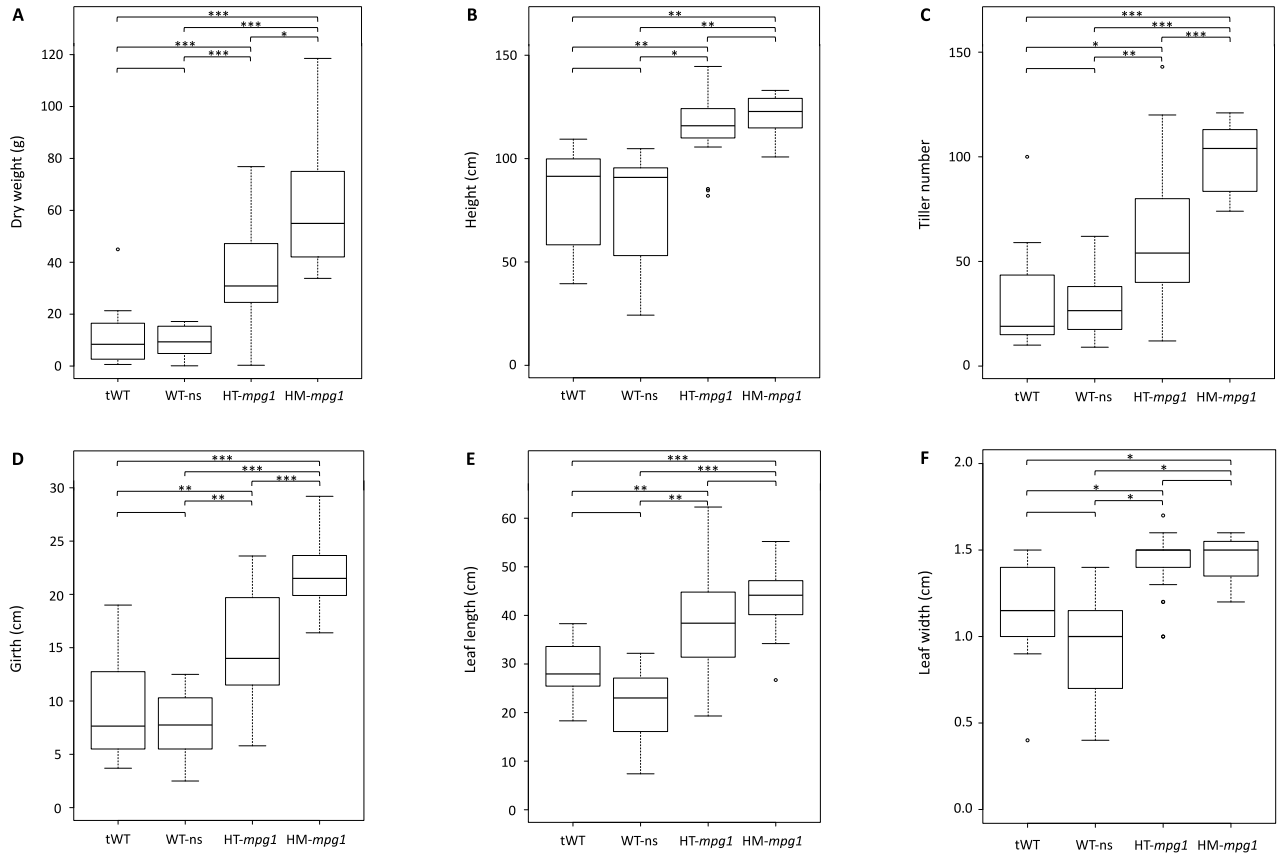


Figure 3.12: Preliminary analysis of biomass metrics of ratooning at maturity within a representative segregating population of T$_4$ *mpg1* grown under optimal conditions.

Biomass and seed-yield metrics were measured at plant maturity after ratooning (second-harvest). All of the *mpg1* plants survived while only 65% of tWT, and 42% of WT-ns survived. Of the surviving plants measurements of (A) dry weight, (B) height, (C) tiller number, (D) girth, (E) leaf length, and (F) leaf width were assessed. tWT (n = 13), WTns (n = 8), HT-*mpg1* (n = 29), HM-mpg1 (n = 12). Analysis was conducted using a one-way ANOVA and Games-Howell post-hoc multiple comparison test at 95% family-wise confidence level. '*' indicates $p<0.05$, '**' indicates $p<0.01$, '***' indicates $p<0.001$ assessing tWT plants against all other treatment groups.

Of the seed yield metrics (Figure 3.13), *mpg1* plants were able to accumulate a higher grain yield (~3.7-fold HM-*mpg1* vs. tWT) and a greater number of panicles than wild-type (~3.5-fold HM-*mpg1* vs. tWT). Many of the wild-type plant's panicles were unable to complete seed filling and mature fully by the time *mpg1* plants matured and began senescence. *mpg1* plants accumulated greater biomass during primary growth and after rationing, suggesting that the mutant has altered growth and development compared to wild-type.

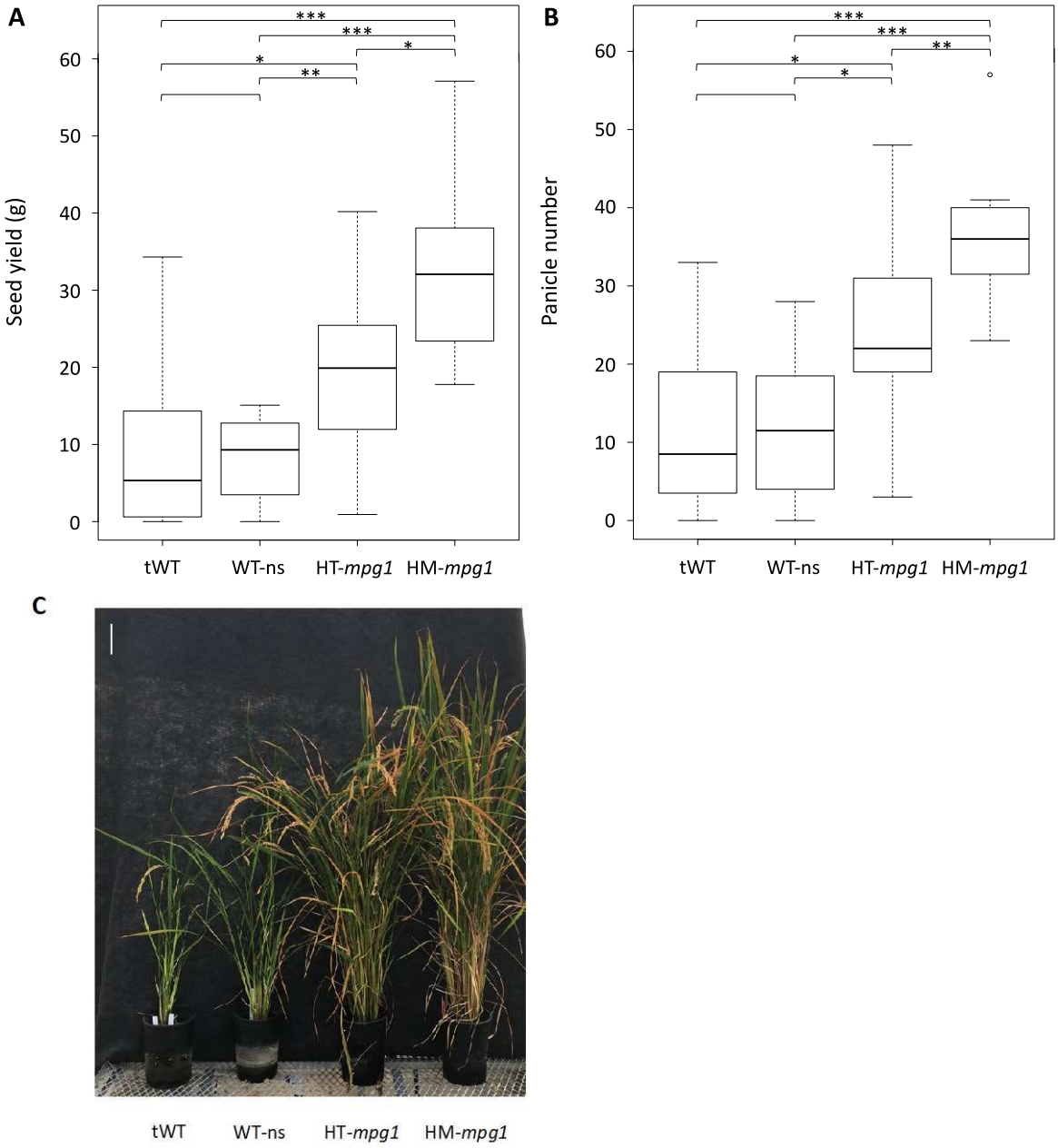


Figure 3.13: Preliminary analysis of seed-yield metrics of ratooning at maturity within a representative segregating population of T_4 *mpg1* grown under optimal conditions.
Biomass and seed-yield metrics were observed at plant maturity after ratooning (second-harvest). All of the *mpg1* plants survived while only 65% of tWT, and 42% of WT-ns survived. Of the surviving plants measurements of (A) seed yield, and (B) panicle number were observed. (C) A photograph of plants at maturity post-ratooning, scale = 10 cm. tWT (n = 13), WTns (n = 8), HT-*mpg1* (n = 29), HM-mpg1 (n = 12). Analysis was conducted using a one-way ANOVA and Games-Howell post-hoc multiple comparison test at 95% family-wise confidence level. '*' indicates $p<0.05$, '**' indicates $p<0.01$, '***' indicates $p<0.001$.

Again, HM-*mpg1* plants exhibited slightly greater traits across several measurements compared to HT-*mpg1*, some significantly, implying that a dosage effect exists. This suggests that the effect of elevated *MPG1* expression reaches a saturation limit for some of the characteristics observed in *mpg1*.

mpg1 plants grown under field conditions accumulate more biomass compared to wild-type plants

Greenhouse environments provide a degree of control, reducing variable environmental interactions with plant growth and development. The environment can interact with the genotype changing the phenotype. To better understand *mpg1*'s phenotype more thoroughly, *mpg1* plants homozygous for the bi-laterally truncated T-DNA insertion, and wild-type null segregants were grown under field conditions in Fort Collins, CO. Plants were started in the greenhouse under local environmental conditions and underwent a period of hardening (~4 weeks), through implementation of artificial wind and were then transplanted outdoors, grown to maturity, and measured (Figure 3.14). It is important to note that the field conditions weren't optimal for rice. Rice is usually grown in areas of high relative humidity (60%-90% relative humidity) (Hirai et al., 2000) in flooded conditions in combinations of silty clay loam (Klotzbücher et al., 2015). The soil content in Colorado consists mainly of hard clay, and the atmospheric conditions contain very low relative humidity. The general per cent humidity in Fort Collins, CO during June to September doesn't exceed 20% (weatherspark.com, 2019). This implies that these plants were likely grown under a degree of stress. *mpg1* plants accumulated significantly greater biomass (~2.1-fold HM-*mpg1* vs. WT-ns), height (~1.1-fold HM-*mpg1* vs. WT-ns), and girth (~1.4-fold HM-*mpg1* vs. WT-ns) compared to wild-type null segregant plants.

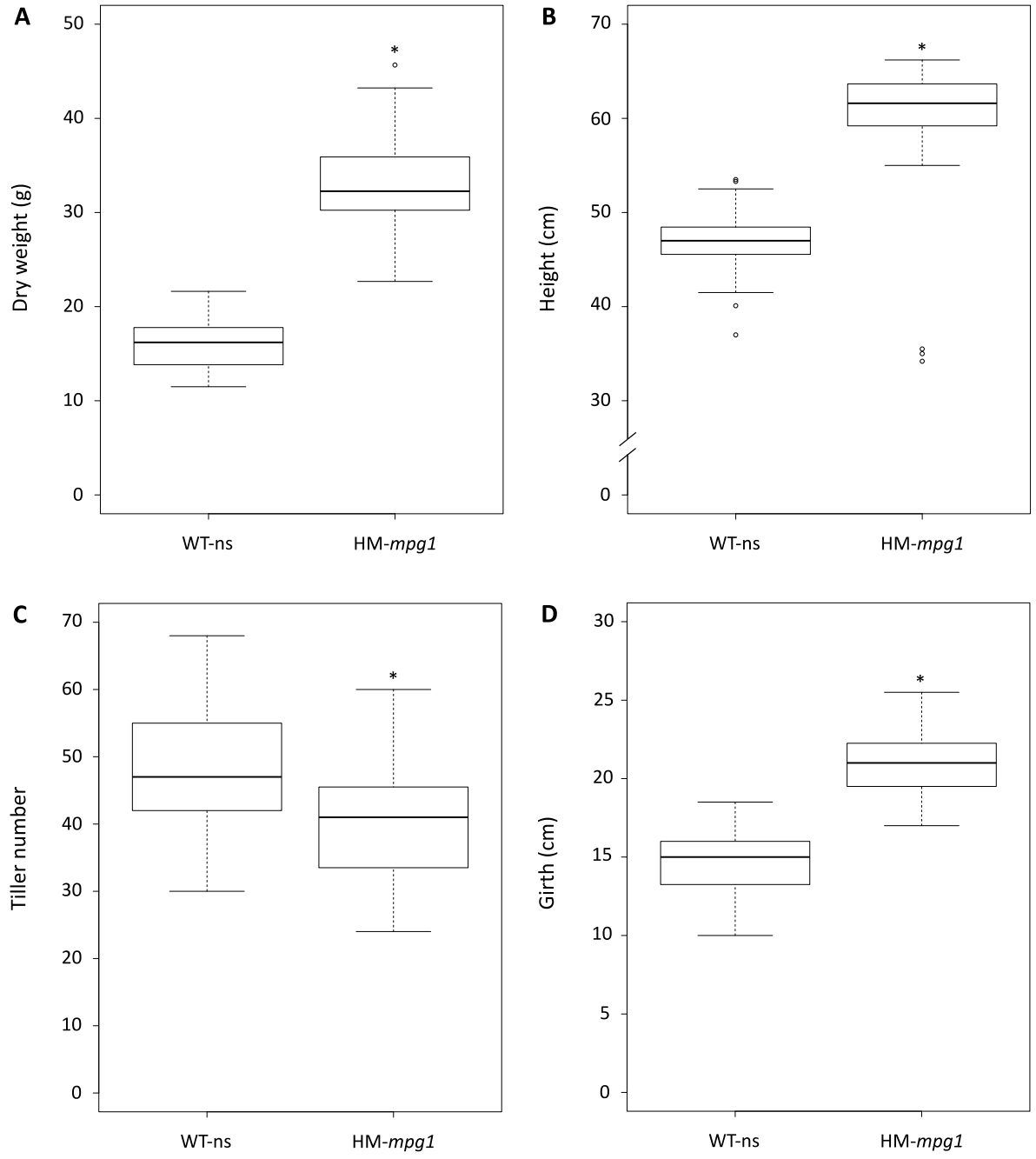


Figure 3.14: Analysis of biomass metrics within a population of T$_4$ *mpg1* grown under field conditions.

To measure the effects of real-world conditions plants were grown in the field. Several measurements were taken to help evaluate phenotypic outcomes related to biomass accumulation from *mpg1* plants, (A) dry weight, (B) height, (C) tiller number, and (D) girth. WT-ns (n = 60), HM-mpg1 (n = 60). Analysis was conducted using a Student's t-test. '*' indicates $p<0.05$ assessing WT-ns plants against HM-*mpg1*.

Interestingly, *mpg1* plants had significantly less tillers compared to wild-type plants. During the tillering stage of development, top-down visualization of the field showed that *mpg1* plants had greater girth and canopy compared to wild-type plants (Figure 3.15, A). At harvest, pictures of the field also show greater plant height of *mpg1* plants compared to wild-type plants (Figure 3.15).

Figure 3.15: Photographs of a population of T_4 *mpg1* grown under field conditions. (A) Field trial layout showing an aerial view of plants during tillering stage of development. Canopy differences can be noted between WT-ns and *mpg1*. (B) Aerial view of plants prior to harvest. (C) Additional aerial view of plants prior to harvest. (Black = WT-ns) (Red = HM-*mpg1*)

The delay in flowering phenotype observed in *mpg1* plants in the greenhouse was also present in the field, however to a greater extent. In the field, *mpg1* plants reached heading of panicles ~30 days after WT-ns, rather than ~14 day delay in heading time seen in the greenhouse within the representative population under optimum conditions. Measurements of seed yield were not taken because plants were harvested prior to seed filling and full reproductive development due to inclement cold weather in the region. The number of spikelets on the three eldest panicles were counted as a surrogate measure of seed yield potential. *mpg1* plants had a greater number of spikelets per panicle compared to wild-type plants (APPENDIX).

F_2BC_1 *mpg1* plants grown under optimal conditions accumulate more biomass and seed yield compared to wild-type plants

The process of tissue-culture *Agrobacterium*-mediated plant transformation can result in T-DNA footprinting, emerging in potentially numerous insertions, deletions, or mutations throughout the genome. To 'clean up' the genetic background of *mpg1* from subsequent potential mutations, the mutant was backcrossed with wild-type. Continued presentation of the *mpg1* phenotype in backcrossed lines will better assure that the bi-laterally truncated T-DNA insertion and increased expression of *MPG1* resulted in the pleiotropic phenotype and not just one of the many characteristics visualized in *mpg1*. In rice it has been suggested to perform at least six if not more backcrosses to mitigate linkage drag of genes tightly linked to the target locus (Hasan et al., 2015). Three independent (BC_1) crosses were generated and the following segregating F_2BC_1 generations were grown and measured. The three independent crossed lines were assessed together against wild-type plants. Backcrossed *mpg1* plants were still able to accumulate significantly greater overall biomass (~4.7-fold HM-*mpg1* vs. tWT), height (~1.2-fold HM-*mpg1*

vs. tWT), girth (~2.0-fold HM-*mpg1* vs. tWT), leaf length (~1.3-fold HM-*mpg1* vs. tWT), and leaf width (~1.3-fold HM-*mpg1* vs. tWT) compared to wild-type plants (Figure 3.16).

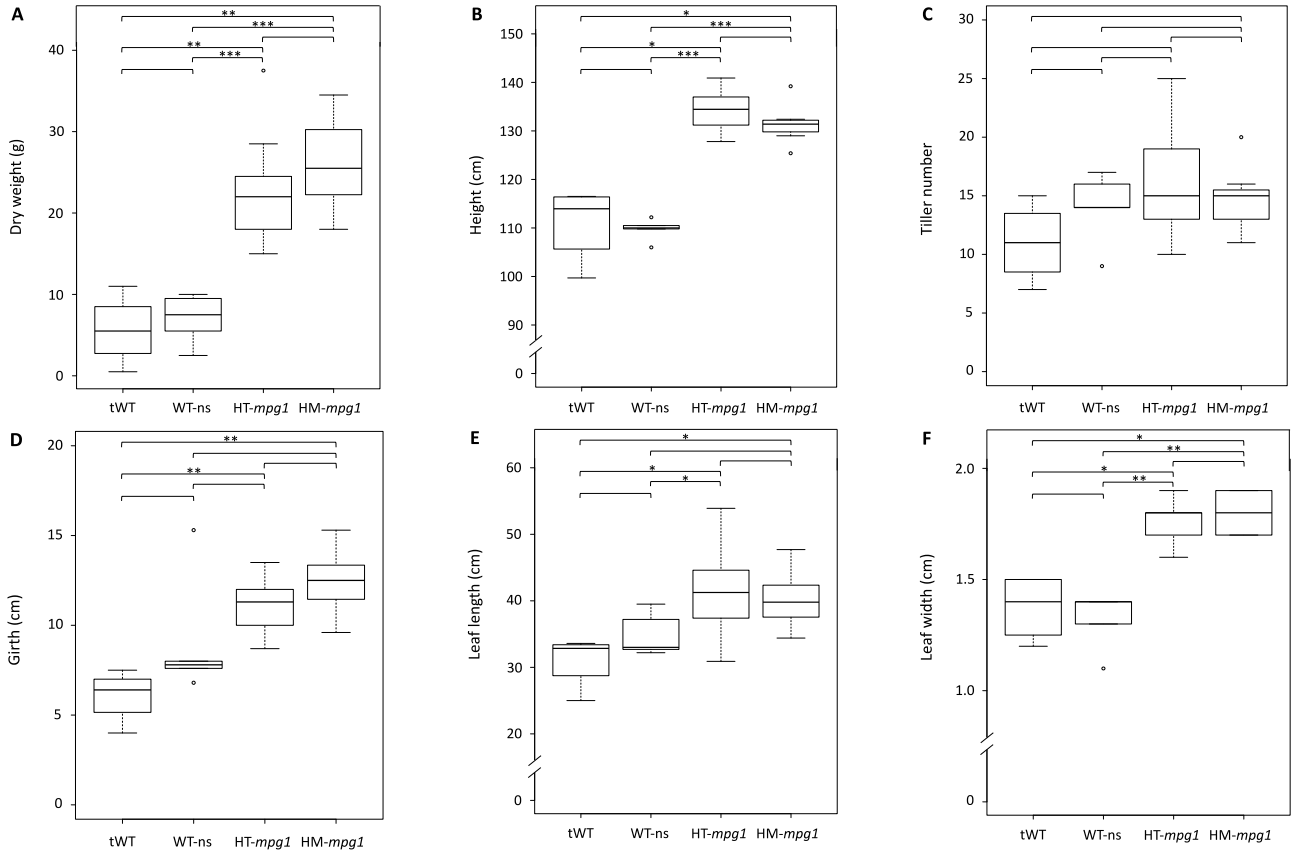


Figure 3.16: Preliminary analysis of biomass metrics within a segregating backcross population of F_2BC_1 *mpg1* grown under optimal conditions.

To clean up the genetic background of *mpg1* and assess the resulting phenotype, backcrosses with tWT plants were grown. The segregating treatments consist of 3 independent crossing events pooled together. Several measurements were taken to help evaluate phenotypic outcomes related to biomass accumulation from *mpg1* plants. (A) dry weight, (B) height, (C) tiller number, (D) girth, (E) leaf length, and (F) leaf width were assessed. tWT (n = 4), WTns (n = 5), HT-*mpg1* (n = 18), HM-*mpg1* (n = 7). Analysis was conducted using a one-way ANOVA and Games-Howell post-hoc multiple comparison test at 95% family-wise confidence level. '*' indicates $p<0.05$, '**' indicates $p<0.01$, '***' indicates $p<0.001$.

Backcrossed *mpg1* also experienced a delay in flowering measured by time to panicle heading compared to wild-type plants (~17 days HM-*mpg1* vs. tWT) (Figure 3.17). *mpg1* also have significantly greater seed yield (~2.2-fold HM-*mpg1* vs. tWT), but not a greater number of panicles compared to wild-type plants (Figure 3.17). This would again suggest that the increase in seed yield in *mpg1* can likely result by number of seeds per panicle and not from an increased number of panicles.

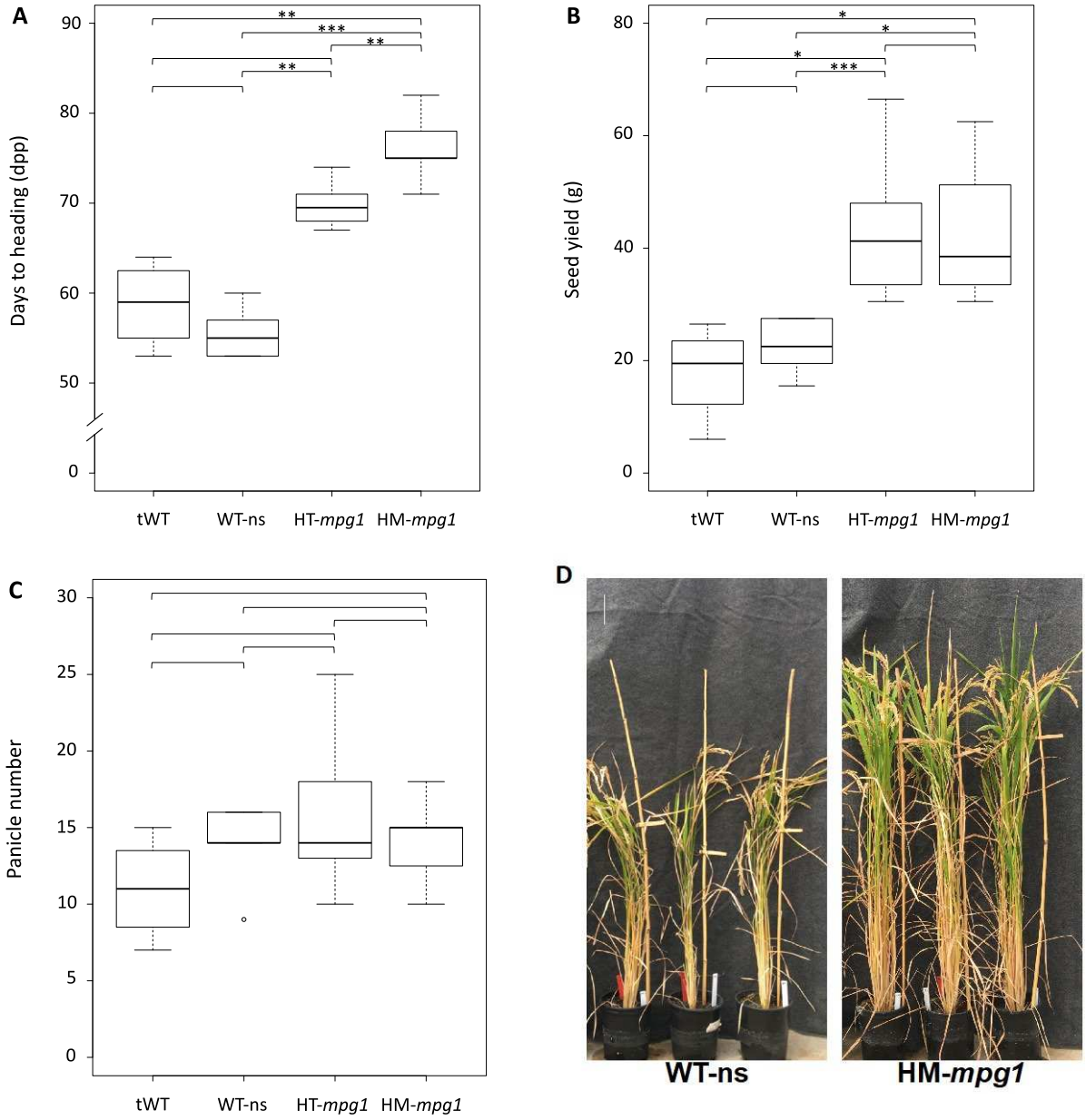


Figure 3.17: Preliminary analysis of seed-yield metrics within a segregating backcross population of F_2BC_1 *mpg1* grown under optimum conditions.
To clean up the genetic background of *mpg1* and validate portions of its pleiotropic phenotype, backcrosses with tWT plants were grown. The segregating treatments consist of 3 independent crossing events pooled together. Several measurements were taken to help evaluate phenotypic outcomes related to seed yield from *mpg1* plants, (A) days to heading measured by days post-planting, (B) seed yield, and (C) panicle number. (D) a photo of three WT-ns and HM-*mpg1* side-by-side, scale = 10 cm. tWT (n = 4), WTns (n = 5), HT-*mpg1* (n = 18), HM-*mpg1* (n = 7). Analysis was conducted using a one-way ANOVA and Games-Howell post-hoc multiple comparison test at 95% family-wise confidence level. '*' indicates $p<0.05$, '**' indicates $p<0.01$, '***' indicates $p<0.001$.

These plants were left to ratoon and biomass and seed-yield metrics were observed at plant maturity (Figure 3.18, 3.19). 25% of tWT, 20% of WT-ns, 50% of HT-*mpg1,* and 85% of HM-*mpg1* survived. The survivability rate of *mpg1* were much higher than wild-type plants post-harvest (ratooning). Although we are unable to report statistics on the surviving plants (tWT and WT-ns only having one surviving plant each), of the surviving plants, *mpg1* plants had noticeably greater biomass characteristics. Additionally, ratooned plants didn't exhibit the same extent of re-growth observed during assessment of non-backcrossed populations suggesting that these plants might have been harvested further into senescence, possibly as a result of insuring maximum seed filling and maturation.

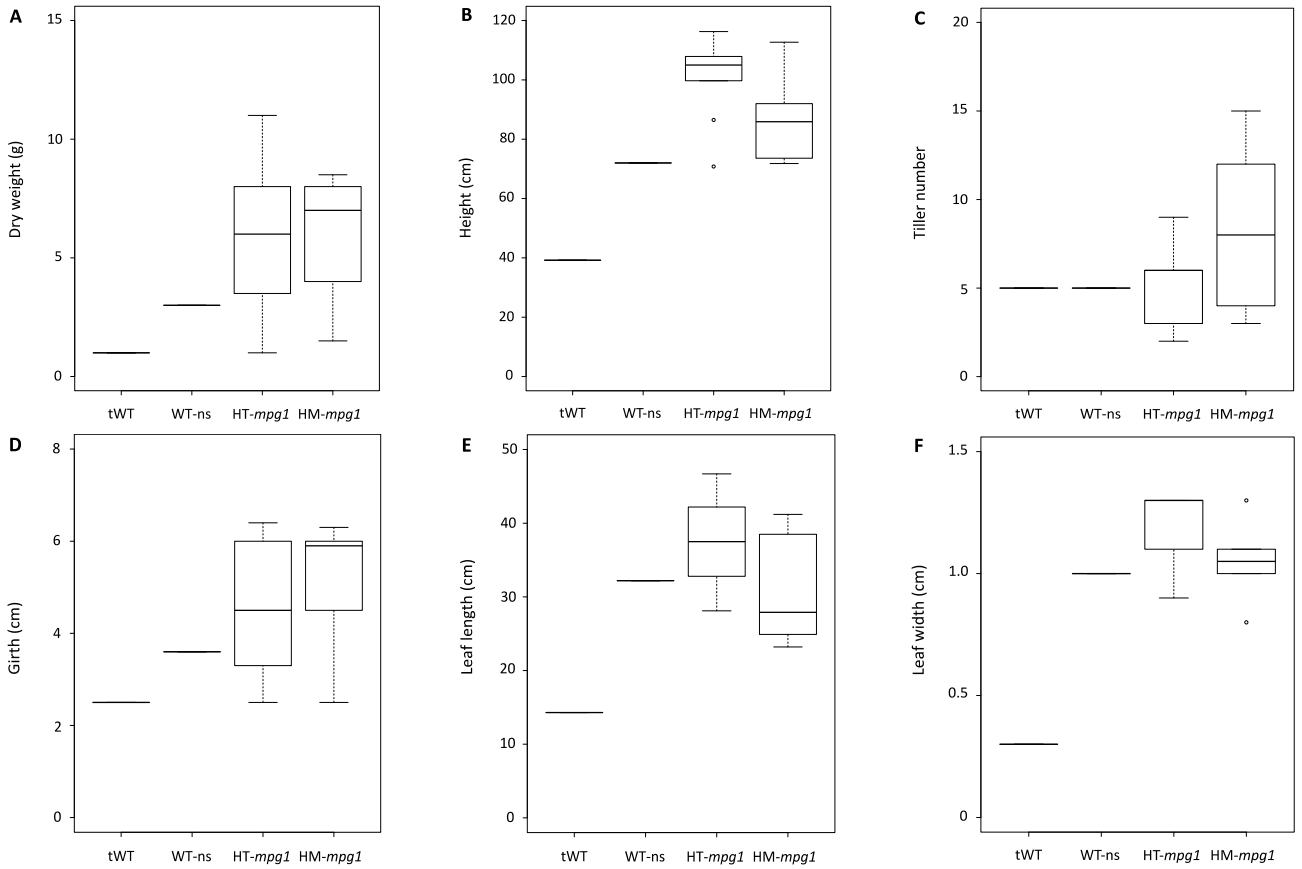


Figure 3.18: Preliminary analysis of biomass metrics within a segregating backcross population of F_2BC_1 *mpg1* at maturity post-ratooning, grown under optimum conditions.

To clean up the genetic background of *mpg1* and validate portions of its pleiotropic phenotype, backcrosses with tWT plants were grown. The segregating treatments consist of 3 independent crossing events pooled together. To better understand the difference between *mpg1* and wild-type plants during ratooning, biomass metrics were observed at plant maturity after ratooning (second-harvest). 25% of tWT, 20% of WT-ns, 50% of HT-*mpg1,* and 85% of HM-*mpg1* survived. Of the surviving plants measurements of (A) dry weight, (B) height, (C) tiller number, (D) girth, (E) leaf length, and (F) leaf width were assessed. tWT (n = 1), WTns (n = 1), HT-*mpg1* (n = 9), HM-*mpg1* (n = 6). Analysis was conducted using a Student's t-test between HT-*mpg1* and HM-*mpg1*.

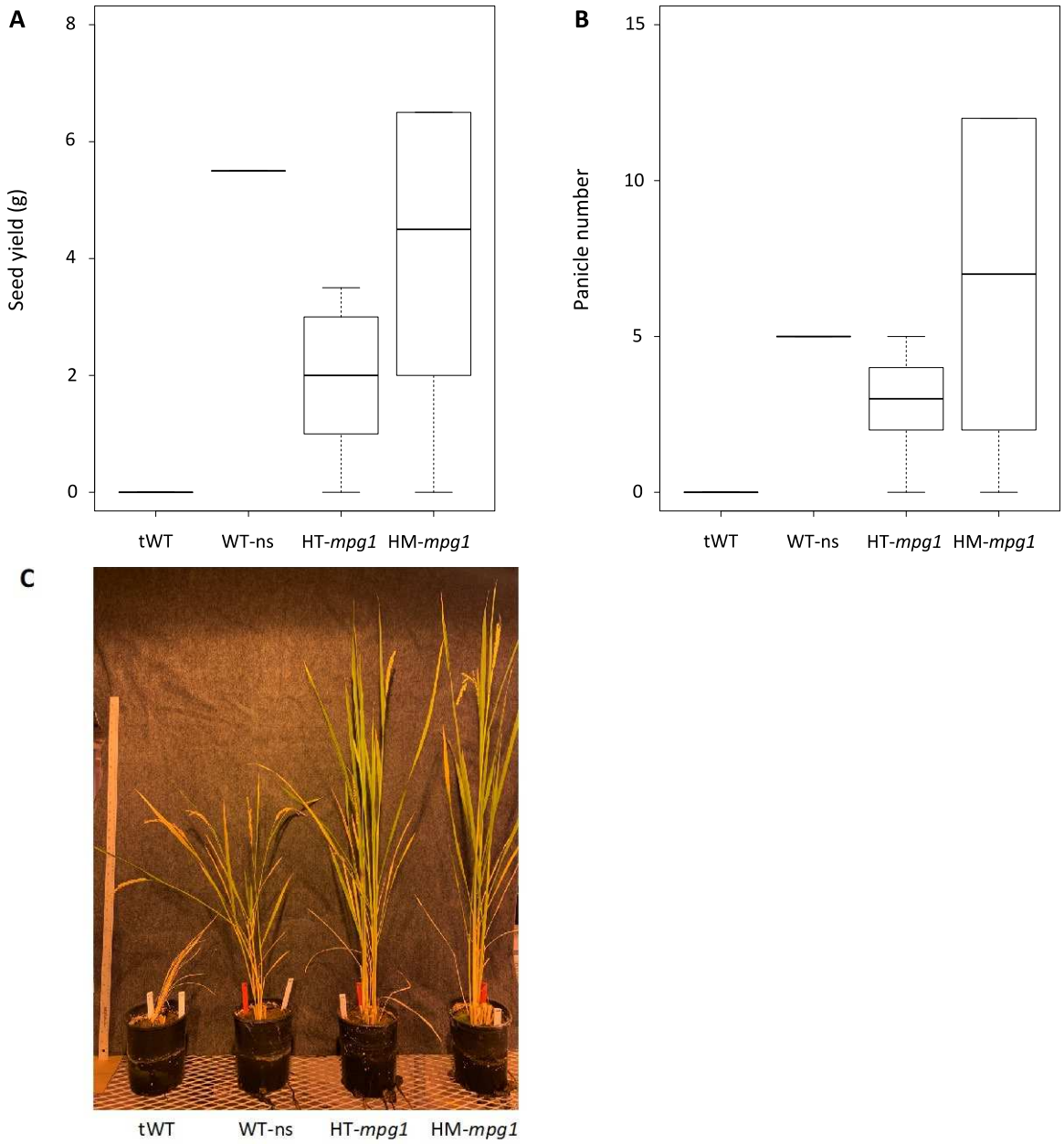


Figure 3.19: Preliminary analysis of seed-yield metrics within a segregating backcross population of F_2BC_1 *mpg1* at maturity post-ratooning, grown under optimum conditions. To clean up the genetic background of *mpg1* and validate portions of its pleiotropic phenotype, backcrosses with tWT plants were grown. The segregating treatments consist of 3 independent crossing events pooled together. To better understand the difference between *mpg1* and wild-type plants during ratooning, seed-yield metrics were observed at plant maturity after ratooning (second-harvest). 25% of tWT, 20% of WT-ns, 50% of HT-*mpg1,* and 85% of HM-*mpg1* survived. Of the surviving plants (A) seed-yield, and (B) panicle number were assessed. (C) photograph of plants. tWT (n = 1), WTns (n = 1), HT-*mpg1* (n = 9), HM-*mpg1* (n = 6). Analysis was conducted using a Student's t-test between HT-*mpg1* and HM-*mpg1*.

***mpg1* plants develop spikelets with awns while wild-type plants do not**

Although, not a very consistent trait from generation and replication, a high number of *mpg1* plants tended to develop awns on a majority of their spikelets (Figure 3.20). Sometimes awns were more or less pronounced, but usually present on the majority of *mpg1* spikelets. Wild-type spikelets never showed any signs of awn development on the majority of their spikelets.

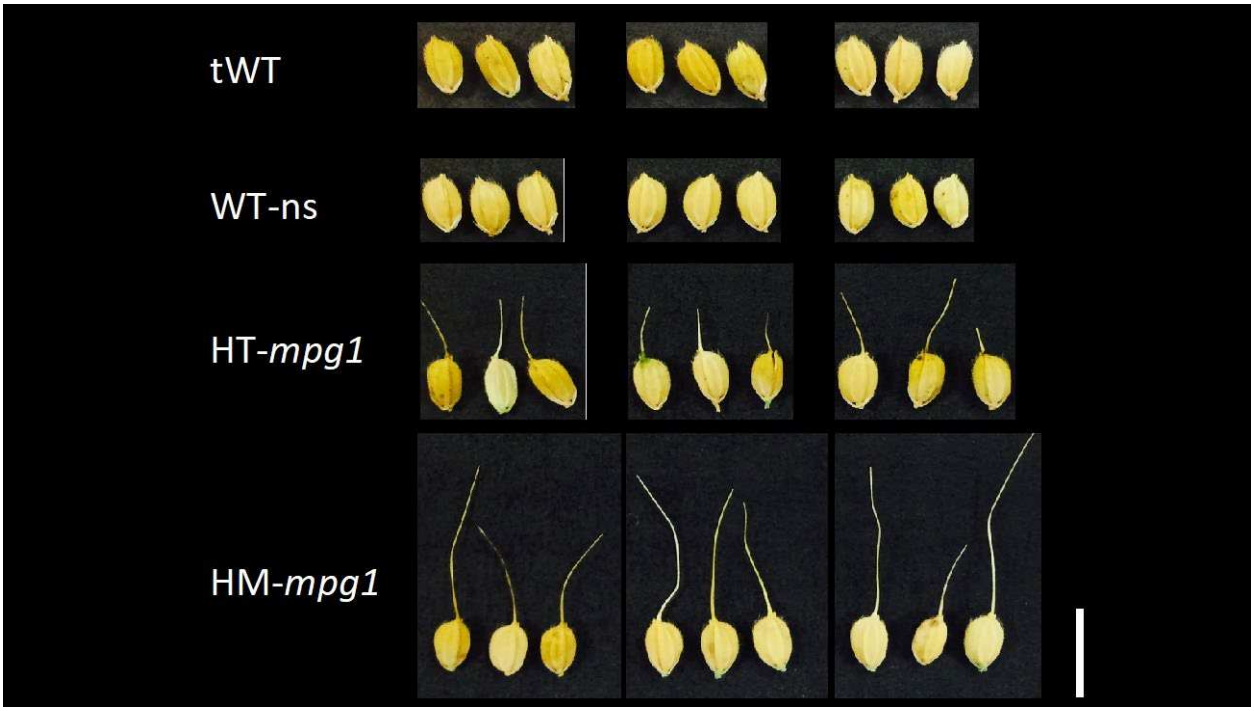


Figure 3.20: Analysis of seed characteristics within an example population of T_3 *mpg1* at maturity.
The presence and length of awns is present in *mpg1* plants. This seems to vary from experimental replication, however is usually present to some degree. Scale = 1.0 cm

***mpg1* plants accumulate greater biomass under drought and salt stress compared to wild-type plants**

The initial growth experiments characterizing the *mpg1* mutant were later determined to be under stressful conditions due to appearance and growth of wild-type plants. Although we were unable to determine the exact cause of the stress we hypothesized that the non-optimally grown plants were exposed to potentially stressful levels of salt and/or acidic conditions due to

visible salt accumulation in watering tubs and higher than optimal EC measurements as a result of the particular growth medium and fertilizer regiments used. To better evaluate *mpg1*'s response to stress, segregating populations of *mpg1* were exposed to specific osmotic stressors (drought and salt). Plants were exposed to these stressful conditions from seedling stage through maturity and measured (Figure 3.21-23). *mpg1* plants accumulated significantly greater overall biomass (~2.2-fold HM-*mpg1* vs. tWT), number of tillers (~1.6-fold HM-*mpg1* vs. tWT), and girth (~1.2-fold HM-*mpg1* vs. tWT) compared to wild-type plants exposed to drought stress. In optimal growth conditions plants homozygous and heterozygous for the bi-laterally truncated T-DNA insertion (*mpg1*) were significantly larger than wild-type, however a number of the measured metrics under defined stress conditions showed a significant difference only between HT-*mpg1* and wild-type plants (not HM-*mpg1* vs. tWT), especially under salt stress (Figure 3.21, 3.22). Interestingly, *mpg1* plants grown under drought and salt stress were not largely different from wild-type plants grown under optimal conditions (control) for overall biomass accumulation (Figure 3.21). This suggests that *mpg1* plants and underlying molecular mechanisms responsible for its phenotype could prove useful in plant growth under stressful conditions without loss of biomass production. The delay in flowering phenotype of *mpg1* compared to wild-type was still present under both drought and salt stress conditions, however both wild-type plants and *mpg1* under both of these stressors saw a delay in flowering beyond what is seen in optimally grown conditions, with drought (~24 day delay HM-*mpg1* vs. tWT) and salt (~25 day delay HM-*mpg1* vs. tWT) (Figure 3.22).

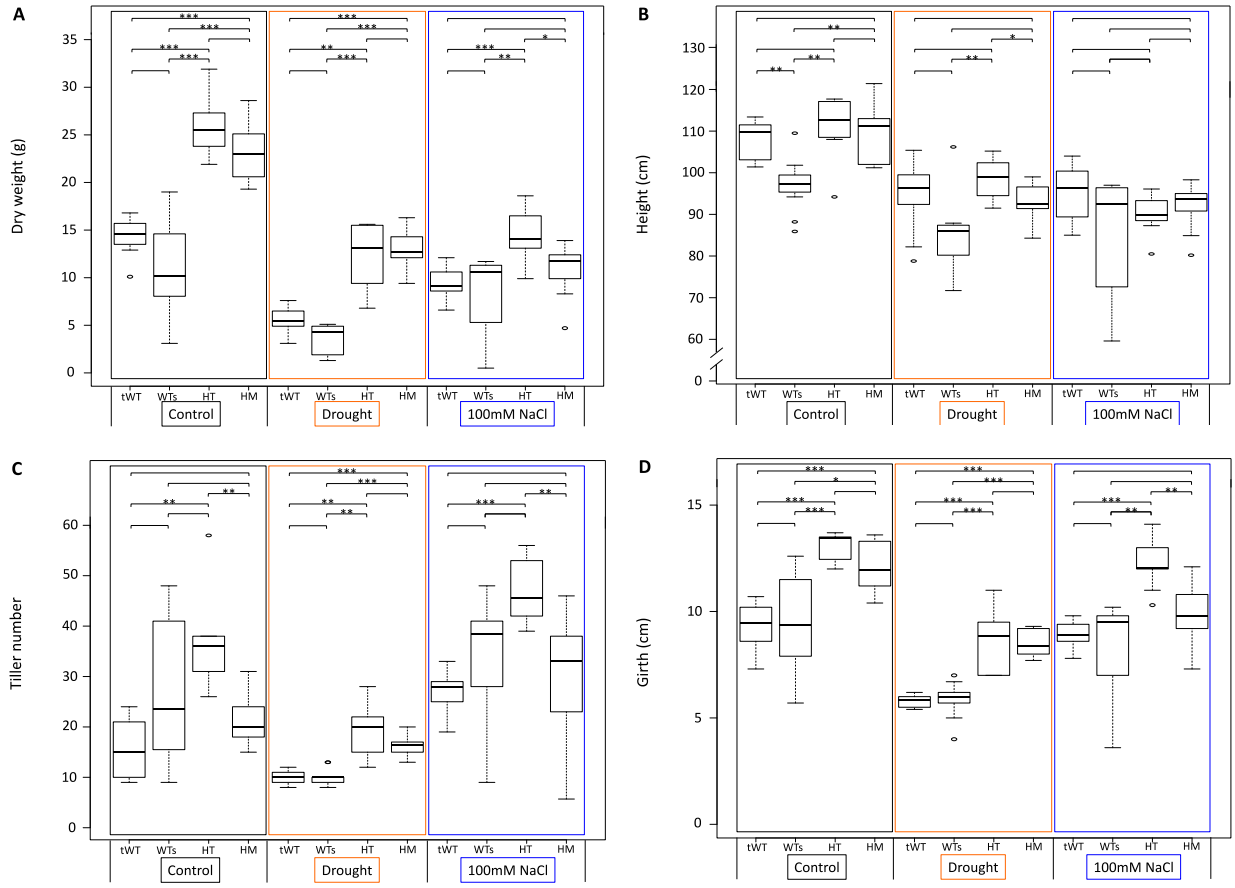


Figure 3.21: Analysis of biomass metrics within an example population of T_3 *mpg1* grown under specific stress treatments.
To better understand the effect of stress on *mpg1*, plants were exposed to drought and salt stresses from 4-leaf stage through to maturity. Measurements were taken at maturity. (A) dry weight, (B) height, (C) tiller number, and (D) girth were assessed. tWT (n = 10), WT-ns (WTs) (n = 10), HT-*mpg1* (HT) (n = 10), HM-*mpg1* (HM) (n = 10). Analysis was conducted using a one-way ANOVA and Games-Howell post-hoc multiple comparison test at 95% family-wise confidence level within experimental groups. '*' indicates $p<0.05$, '**' indicates $p<0.01$, '***' indicates $p<0.001$.

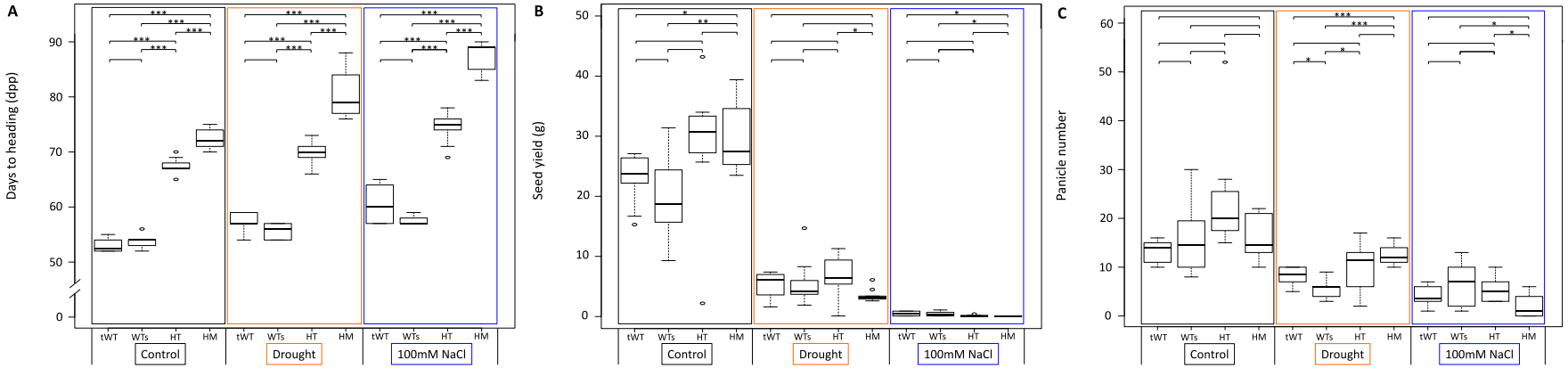


Figure 3.22: Analysis of seed-yield metrics within an example population of T_3 *mpg1* grown under specific stress treatments. To better understand the effect of stress on *mpg1*, plants were exposed to drought and salt stresses from 4-leaf stage through to maturity. Measurements were taken at maturity. (A) days to heading, (B) seed yield, and (C) panicle number were assessed. tWT (n = 10), WT-ns (WTs) (n = 10), HT-*mpg1* (HT) (n = 10), HM-*mpg1* (HM) (n = 10). Analysis was conducted using a one-way ANOVA and Games-Howell post-hoc multiple comparison test at 95% family-wise confidence level within experimental groups. '*' indicates $p<0.05$, '**' indicates $p<0.01$, '***' indicates $p<0.001$

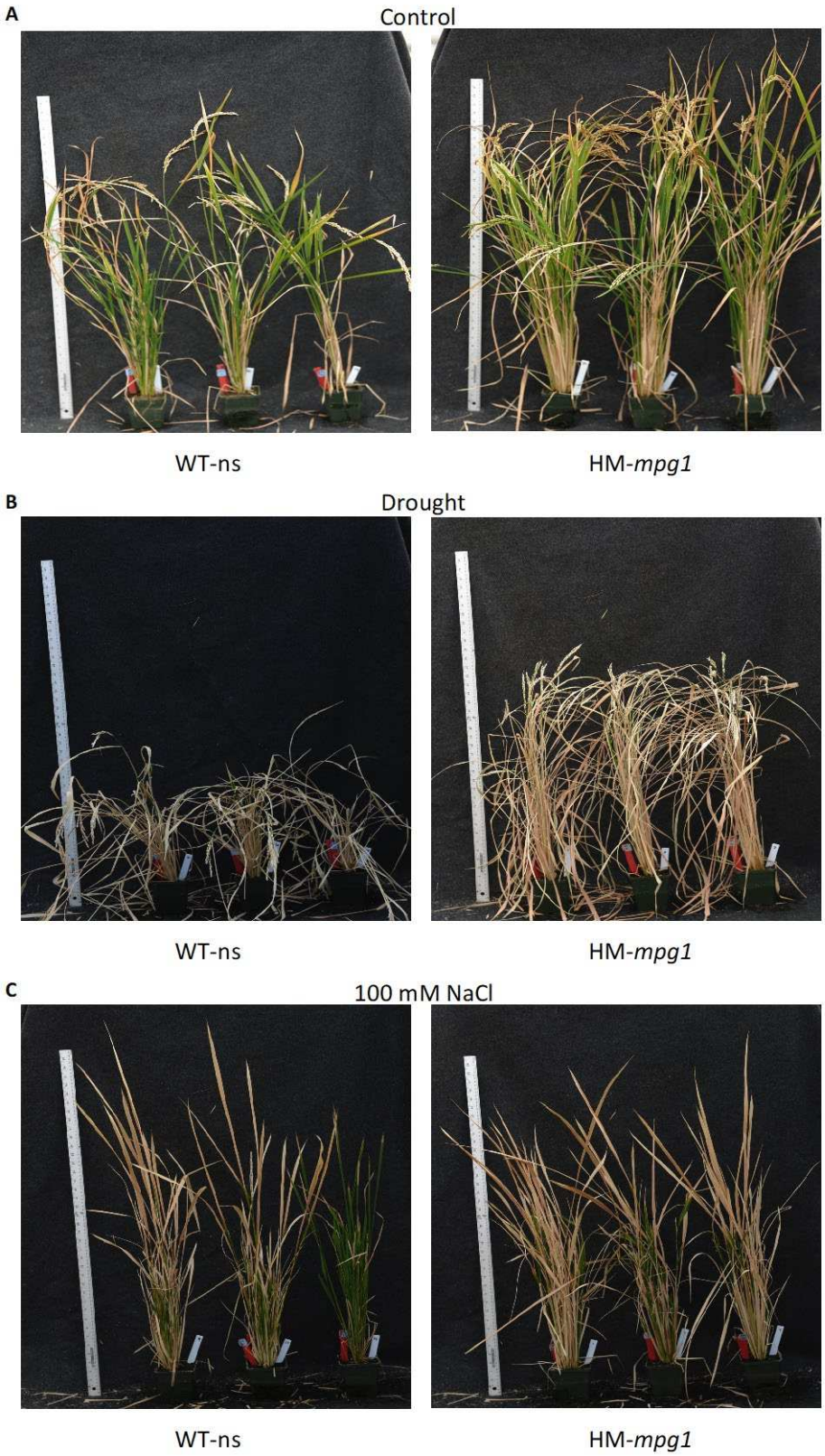


Figure 3.23: Photographs of an example population of T_3 *mpg1* grown under specific stress treatments. Scale = 1.0 m.

Seed yield and panicle number characteristics between *mpg1* and wild-type were difficult to ascertain under the stress conditions because plants were unable to complete seed filling and maturation. *mpg1* plants did however exhibit greater dry weight and girth compared to wild-type, under drought conditions (Figure 3.21A). Interestingly, the control group didn't present with the same degree of phenotypic difference that was observed during assessment of the optimal and non-optimal condition growth trials. This might suggest that pot size influences the degree of the phenotype, as pot size was the only considerably different variable between these experiments.

The *MPG1* gene is not constitutively overexpressed in *mpg1* plants

The presence of the bi-laterally truncated T-DNA insertion in *mpg1* was tracked alongside expression of *MPG1* in wild-type and segregating *mpg1* plants within above ground tissues for multiple generations. The presence of the truncated T-DNA insertion and increased expression of *MPG1* compared to wild-type plants (which showed little to no expression of *MPG1*) directly correlated with the greater biomass characteristics, increased seed yield, delayed flowering, and performance under sub-optimal conditions observed in *mpg1*. Thus, expression of *MPG1* likely plays a role in the phenotype of *mpg1*. We initially predicted that the presence of the bi-laterally truncated T-DNA expression cassette resulted in constitutive overexpression of *MPG1* because the various tissue types and time-points initially assessed (leaf, stem, flowering) showed continually, and similarly, high levels of expression. However, in root tissue of *mpg1* there was little to no expression of *MPG1* (Figure 3.24). This demonstrates that the expression of *MPG1* appears in specific tissue types and that the gene is not constitutively overexpressed in the mutant.

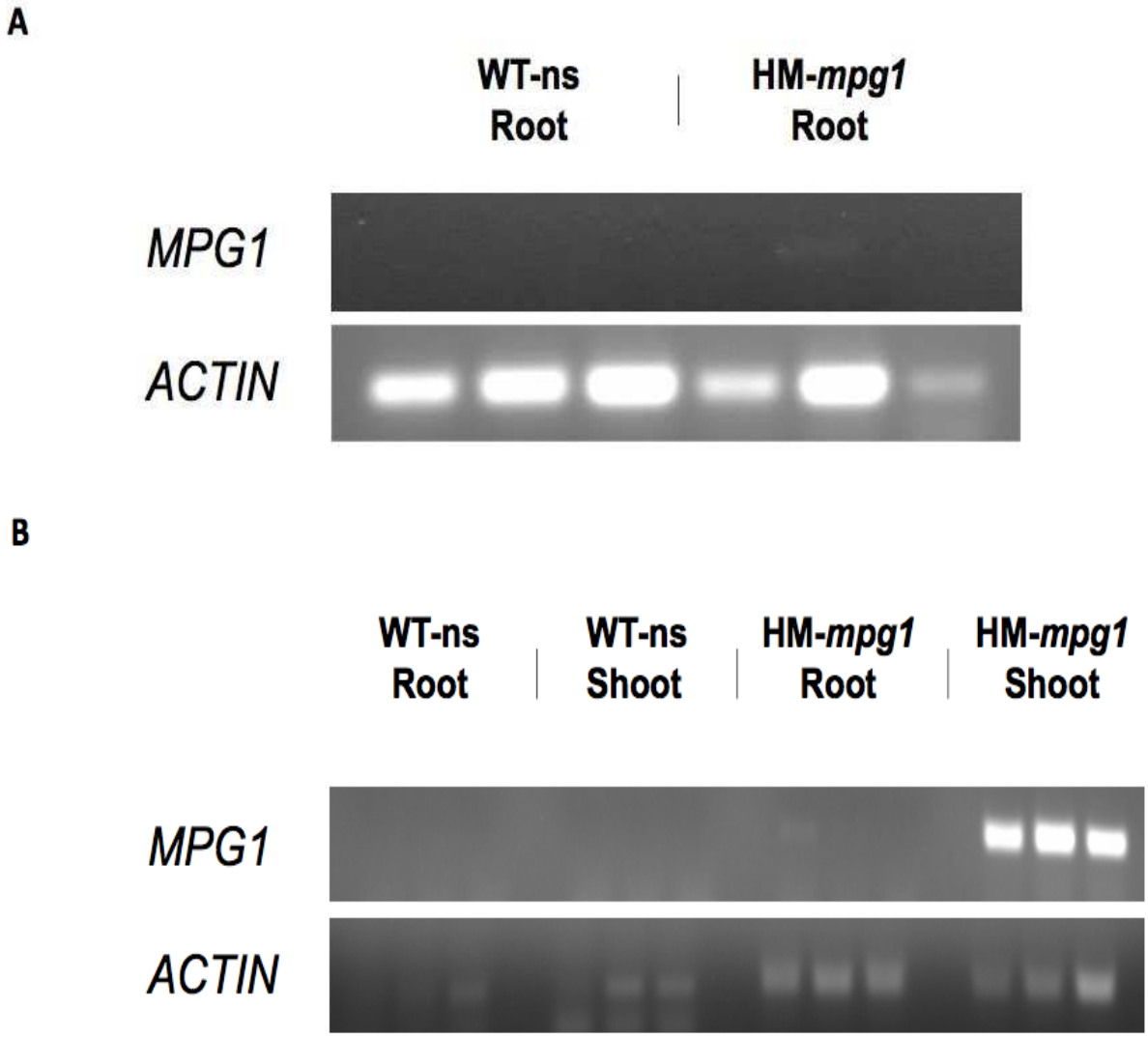


Figure 3.24: Tissue specific expression analysis of *MPG1*.
(A) RT-PCR gel-electrophoresis image from whole root tissue at 32 days post-planting from plants grown in soil under optimum conditions. There is little to no expression of *MPG1* in plants that are homozygous for the T-DNA insertion present in *mpg1,* and no visible expression present in wild-type null segregant plants. Each lane represents a single biological replicate, and the gene coding for actin was used as a control. (B) RT-PCR gel-electrophoresis image from whole root tissue and whole shoot tissue at 21 days post-planting from plants grown in hydroponic conditions (Yoshida Solution). There is little to no expression of *MPG1* root tissue in plants that are homozygous for the T-DNA insertion present in *mpg1,* and no visible expression present in wild-type null segregant plants. Expression can still be seen however in shoot tissue of HM-*mpg1* and not in wild-type null segregants. Each lane represents a single biological replicate, and the gene coding for actin was used as a control.

Native *MPG1*'s expression is activated under *Xoo* challenge

Ectopic regulation of AP2/ERF transcription factors has led to enhanced tolerances of abiotic stress, but also resistance to multiple diseases (Xu et al., 2011). Because of this, *mpg1* was further assessed by means of a pathogen challenge assay. *Xanthomonas oryzae pv. oryzae* (*Xoo*), a bacterial pathogen that results in bacterial blight in rice, was selected. *mpg1* and wild-type plants were exposed to *Xoo* by means of cutting leaf tissue dipped in a pathogen culture (clipping inoculation method), and tested for gene expression of *MPG1* hours following inoculation, as well as legion length days later. Interestingly, there wasn't a significant difference in legion length between *mpg1* and wild-type (data not shown), however at 6 hours post-inoculation *MPG1* expression was visualized in wild-type plants (APPENDIX). This suggests that *MPG1* might natively function in defense response, and that its ectopic expression in *mpg1* might impact defense response.

Discussion

***mpg1* plants are more productive than wild-type plants under numerous conditions**

Segregating populations of *mpg1* plants showed increased productivity (particularly in biomass accumulation) under 'non-ideal', optimum, specific stress, and field conditions. Under optimal conditions *mpg1* plants accumulated greater biomass than wild-type in the forms of overall dry-weight, height, girth, and leaf size at maturity. Under the non-optimal conditions *mpg1* plants also had greater tiller number compared to wild-type suggesting that stress impacts tiller formation. Tiller number in rice is regulated by the phytohormone strigolactone (Arite et al., 2009, Lin et al., 2009, Minakuchi et al., 2010). Seed yield characteristics also varied from non-optimal conditions to optimal ones. The number of panicles present in *mpg1* was greater than wild-type under non-optimal conditions however this was not observed in *mpg1* plants

grown under optimal conditions. This is likely a result of not having greater tiller numbers, as panicles arise from individual tillers. However, under both growth conditions *mpg1* accumulated greater seed yield than wild-type plants. This indicates that *mpg1*'s enhanced seed yield is likely a result of developing a greater number of seeds per panicle.

The degree of phenotypic differences observed in *mpg1* was greater under non-optimal conditions than under optimal ones. However, there was still a significant difference in biomass and seed yield under optimal conditions. This suggests that stress plays an important role in the mutant phenotype. Although we were able to link stress response to the phenotype in *mpg1* it was unclear from our initial experiments what type of stress it is associated with. Our observations suggested that the plants were possibly under salt and/or pH stress. Additionally, expression analyses of plants undergoing arrays of different treatments and stressors (exogeneous salicylic acid, submergence, laid-down submergence, and cold) revealed the activation of *MPG1* (Sharoni et al., 2011). Several of these particular conditions stimulate osmotic stress; provoking further investigation into gene expression databases specific to transcription factors in response to stress (RiceSRTFDB) and found that *MPG1* is activated under drought and salt conditions as well (Priya and Jain, 2013). Analysis of *MPG1*'s promoter region also revealed a high number of *cis*-elements specific to 'drought stress' and 'abiotic stress', with a smaller number pertaining to 'biotic stress' and 'salinity stress' present as well (Priya and Jain, 2013). Thus, we opted to evaluate *mpg1* under controlled salt and drought stress independently.

Because the original observation of reduced growth due to an unknown stress, under non-optimal conditions, was likely a prolonged exposure, we grew plants under extended periods of salt or drought. Under extreme salt and drought stress, *mpg1* plants accumulated greater biomass,

however we were unable to evaluate seed yield, as both wild-type and mutant populations were unable to successfully generate flowering tissue and complete seed filling. Both *mpg1* and wild-type plants exhibited symptomatic indicators of stress (bleaching, streaking, and burning) at similar time of onset and degree. Prolonged exposure to stress can have cumulative effects potentially becoming excessive and difficult-to-gauge. The level of stress (particularly salt) used in our examination was below commonly used levels for acute evaluations of stress response (Sengupta and Majumder, 2009, Hasegawa et al., 2000, Saijo et al., 2000, Ohta et al., 2002). Induction of drought and monitoring of leaf curling is also a good metric to insure drought stress is occurring (Wopereis et al., 1996, O'Toole and Cruz, 1980). However, it is likely that over time, the accumulation of salt or drought exceeded the tolerance levels of both plant types over the course of stress treatment. Evaluating and insuring prolonged stress could have provoked a compounding level of cell death, further than that of an acute exposure to stress, and is probably unlike what would be seen in field conditions (Hsiao et al., 1984, Reddy et al., 2017). Because of this it is difficult to specifically evaluate if *mpg1* has a degree of salt or drought tolerance. The increase in biomass exhibited under these specific stressors could just be an effect of enhanced growth prior to reaching sufficient levels of stress. Further experimentation, possibly evaluating the effects of acute stress during different points of development might serve better in understanding the effect of particular stress on the *mpg1* phenotype.

mpg1 plants were grown in the field to evaluate the persistence of their phenotype under exposure to environmental conditions. Often phenotypes observed in growth chamber or greenhouse settings do not translate to the field because conditions are generally optimal in controlled conditions versus the wide range of conditions seen in nature, which can affect phenotypes. *mpg1* accumulated a greater amount of biomass compared to null segregants. The

particular metrics that contributed to increased biomass were height, girth, and leaf width. Although we were unable to measure seed yield in the field, we did count the number of seeds on the oldest 3 panicles as a preliminary means of assessment. *mpg1* had a greater number of seeds per panicle compared to wild-type null segregents (APPENDIX). This correlated with what was observed in plants grown in the greenhouse under optimal conditions, as greater seed yield was produced in *mpg1* without an increase in panicle number compared to wild-type.

There isn't a noticeable difference in *mpg1* and wild-type seeds. A 1000 grain weight was assessed on a segregating T_2 population grown in the greenhouse, showing no significant difference between *mpg1* and wild-type (APPENDIX). In support of this, *mpg1* seedlings monitored throughout various populations did not exhibit any noticeable differences in shape and size compared to wild-type. An interesting observation was that the delay in flowering phenotype in the field was roughly double of that observed under greenhouse conditions. *Oryza sativa* L. is a short-day photoperiodic flowering plant regulated by several genes including *EARLY HEADING DATE 1* (*Ehd1*), *HEADNIG DATE 1* (*Hd1*), and MADS box transcription factors that signal for expression of *FLOWERING LOCUS T* (*FT*) group of floral inducers (Doi et al., 2004). It is likely that the delay in flowering retarded the onset and transition from vegetative development to reproductive development by interacting with a gene or genes in this pathway. It would be interesting to look at the expression of these genes in *mpg1* relative to wild-type. Although longer daylength usually inhibits flowering of rice (Lee and Gynheung, 2015), the extended delay in flowering time seen in field conditions could potentially be an effect of the difference in day length compared to the controlled light cycle (16 hour per day year round) in greenhouse conditions.

Biomass accumulation, seed yield, and stress tolerance are complex traits that can involve

a multitude of genes (Rojas et al., 2010, Van Camp, 2005, Gao et al., 2007, Ellis et al., 2000, Qiu et al., 2007). A deeper investigation, specifically assessing, genes involved with plant height, girth, leaf size, panicle length, seed number, and abiotic and biotic stress response could help to determine the underlying molecular mechanisms involved with formation of these traits in *mpg1*.

MPG1 expression correlates well with the increased biomass phenotype of _mpg1_

During the comprehensive phenotyping of *mpg1* we periodically sampled various tissues across multiple generations of selfed segregating plants. *mpg1* plants, both heterozygous and homozygous for the bi-laterally truncated T-DNA insertion, showed high levels of expression of *MPG1* compared to null segregants and true wild-type plants (which showed no measurable level of report). Tissue types consisted of leaf and stem tissues from the vegetative phase of development, as well as, whole panicles from the reproductive phase of development. However, later experiments revealed that expression was low to non-existent for *MPG1* in root tissues of the mutant. The dramatic change in this genes expression and its proximity to the T-DNA insertion increases our confidence that *MPG1* is responsible for the characteristics observed in *mpg1* (delayed flowering, increased biomass accumulation, increased seed yield, and possibly a degree of stress tolerance). Additionally, phenotypic measurements between tWT and WT-ns plants did not usually differ, suggesting that the presence and influence of any unlinked T-DNA footprinting events are unlikely. An evaluation of several independent segregating (BC_1) backcrossed lines resulted in similar phenotypic observations. Although further backcrosses are needed, presence of the T-DNA insertion and ectopic expression of *MPG1* continues to result in the observed mutant phenotype. Continued backcrosses will help ensure that the phenotypes associated with *mpg1* are not the result of footbrinting from *Agrobacterium*-mediated tissue culture transformation that might be tightly linked and co-segregates with the T-DNA insertion.

***mpg1's* delay in development might be effecting our understanding of the phenotype**

Under all the comprehensive phenotypic growth assessments, *mpg1* had a significant delay in flowering compared to wild-type plants. We took measurements of this delay in flowering by the number of days post-planting that it took plants to reach the panicle heading stage of floral development. Although this is a sufficient metric of gauging temporal rates of flowering time, it is insufficient in uncovering the timepoint in which the vegetative phase of development transitions to a reproductive phase of development (molecularly and morphologically). There is a wide variety of transcriptional and physiological changes that occur when a plant undergoes this transition process.

mpg1 plants accumulate greater biomass prior the time point where wild-type plants begin heading. However, date to heading isn't the actual marker of reproductive transition. Therefore, this could indicate that the enhanced growth metrics present in *mpg1* occur from either a continuously greater growth rate along with an extended period of growth due to a prolonged vegetative phase, or that *mpg1* plants are greater in size solely because of an extended amount of time to grow resulting in the accumulation of greater biomass compared to wild-type. We noticed that in most metrics we begin to visualize *mpg1* superseding growth of wild-type plants around the 42 days post-planting time point. Based on observations of growth over time, *mpg1* appears to experience roughly two weeks of extended growth compared to wild-type based on when growth appears to plateau (Figure 3.8), which closely matches the extent of time observed in the delay of flowering. Dissecting tiller tissue and discerning when panicle initiation (the true marker for developmental transition) occurs in wild-type plants relative to *mpg1* will allow for a better understanding of when vegetative phase ends and how the extent and rate of biomass accumulation transpires as it relates to periods of development.

Additionally, the assessment we performed to address ratooning differences between *mpg1* and wild-type might be effected by this delay in flowering as well. *mpg1* grown in the greenhouse experienced about a 14 day delay in flowering from wild-type. This effect might be even more exacerbated by time of seed filling and senescence. Because wild-type plants flower earlier compared to *mpg1* they complete seed maturation and filling earlier as well. These plants might undergo a higher degree of senescence compared to when *mpg1* completes seed filling, which could directly affect its ratooning capability as both populations were ratooned at the same time. Future experimentation should include harvesting wild-type and *mpg1* at similar stage of development after flowering is complete, likely similar to the 2 week delay observed in flowering.

The shift in developmental progression between *mpg1* and wild-type might also affect our ability to evaluate prolonged stress. Stress effects tissues differently throughout development (Cooper et al., 2003). *mpg1* plants were likely exposed to an additional 2 weeks of stress during vegetative development, while wild-type plants began their reproductive phase, complicating our ability to appropriately observe effects of stress side-by-side. Treatments should likely be evaluated with respect to the delay in development, or exclusively within similar temporal ranges of development likely through exposure to an acute stress.

mpg1* plants do not constitutively overexpress *MPG1

Early evaluation of *MPG1* expression showed that leaf tissue, stem tissue, and panicle tissue, all exhibited continually high levels of expression in *mpg1*. Due to this, we originally anticipated that it was being constitutively overexpressed, and hypothesized that the ectopic expression of *MPG1* in mutants was a result of transcriptional elements present in the T-DNA insertion. Two promoter element sequences remained present in this insertion, one being the

constitutive overexpression promoter from the *CaMV35S* gene (duplicated promoter), and the other being the companion cell-specific promoter (National Center for Bio-technology Information [NCBI] accession no. AF249912) from the *CmGAS1* gene (only the first third of the promoter region is still present in truncated T-DNA). If one of these elements were driving the expression of *MPG1* we would either see continually high expression in all tissue types (assuming the remnant *CaMV35S* promoter sequence is driving expression), or expression specifically in the vasculature (assuming the remnant *CmGAS1* promoter sequence is driving expression). The insertion itself occurred roughly 7kb upstream of *MPG1*, with one gene between the T-DNA and *MPG1*. Additional things to consider is the orientation of these promoter sequences relative to *MPG1*. The duplicated *CaMV35S* promoter sequence is oriented inversely to the direction of *MPG1* with the presence of a terminator poly-adenylation sequence still remaining, while the *CmGAS1* promoter sequence is oriented in the same direction. Although the *CmGAS1* promoter sequence is oriented in the same direction upstream of *MPG1*, it is missing the latter two-thirds of its sequence so it is questionable whether or not it is functional, much less able to direct tissue-specific expression.

Although we observed high levels of expression of *MPG1* in above ground tissues of mpg1 plants, we saw little to no expression in the roots of *mpg1* suggesting that the duplicated *CaMV35S* promoter is not driving expression of *MPG1* in *mpg1* mutants. The expression patterns of *CaMV35S* promoter have been shown to act in a constitutive manor in rice using a promoter-*GUS* fusion, examples (Chen et al., 2017b, Chen et al., 2017a) showing consistent high levels of GUS in whole root tissue. This expression pattern does however closely mimic the expression pattern observed in rice containing *CmGAS1* promoter driving the reporter gene *GUS*. Utilization of the *CmGAS1* promoter fused to GUS in rice has shown expression primarily in the

vasculature of plants - specifically in leaf, stem, and panicle tissue, with sporadic and limited expression in root tissue (Chapter 2). This result is somewhat surprising because there is only a small portion of the *CmGAS1* promoter remaining in the integrated bi-laterally truncated T-DNA. Functional analyses of the *CmGAS1* promoter in *Arabidopsis* uncovered the necessary and sufficient sequence elements to drive tissue-specific expression (minor vein companion cells) in this species (Ayre et al., 2003). The portion of the promoter sequence present in *mpg1 (*-3081 to -2045 from the open reading frame of *CmGAS1*) does not contain the sequences determined to be necessary and sufficient to drive tissue specific expression in minor veins (-1333 to -983 from the open reading frame of *CmGAS1*) (Ayre et al., 2003). The region present in *mpg1* is -3081 to -2045 from the open reading frame of *CmGAS1*, suggesting that the remaining sequence of the promoter should not be sufficient to drive expression in a companion cell-specific manner. The authors didn't however investigate any of the sequence upstream of -1800. Functional elements might remain in this region as there are conserved sequences of unknown function that are found across several Cucurbitaceae members (Ayre et al., 2003). For example, the shortened *CmGAS1* promoter sequence in *mpg1* includes a G-Box motif (CACGTG), at -2885 bases before the putative *CmGAS1* translational start site. (Volk et al., 2003). G-Box motifs are highly conserved sequences that have been shown to be associated with genes regulated by environmental signals and physiological cues (Menkens et al., 1995).

Although *CaMV35S* promoter is generally considered a unidirectional promoter, recent studies evaluating the nature of T-DNA and resulting expression cassettes utilizing this promoter sequence have uncovered a lack in engineered transgene expressing specificity. Regardless of vector orientation (promoters, genes, and terminator sequences) when using a *CaMV35S* promoter to drive expression of selectable markers, ectopic and non-specific expression of genes

within or neighboring the T-DNA insertion site can occur (Zhou et al., 2014b). It has been shown that the activity of the *CaMV35S* promoter can cause interference in the strength and specificity of adjacent promoters, especially ones that are spatiotemporally specific (Zheng et al., 2007, Yoo et al., 2005). Further, *CaMV35S* promoter can function as an enhancer in either orientation to increase the transcription of genes close in proximity (Benfey et al., 1989, Benfey et al., 1990b, Benfey et al., 1990a). The 35S enhancer sequences (present in the duplicated *CaMV35S* promoter within *mpg1*) *trans* activate and/or override the regulation of transgenes or neighboring genes increasing their expression (Weigel et al., 2000, Kay et al., 1987). Interaction of the *CaMV35S* promoter on adjacent promoters may be even greater if the neighboring promoter is truncated as it lacks potentially natural insulating sequences (Beilmann et al., 1992, Ohtsuki et al., 1998). The bi-directional capability of this promoter/enhancer on non-targeted elements is referred to as enhancer-promoter interference and has been reviewed for its abilities and mitigation (Gudynaite-Savitch et al., 2009, Singer et al., 2011). Based on our results, it is possible that enhancer elements present in the duplicated *CaMV35S* promoter of the T-DNA insertion in *mpg1* are influencing the activity of the native promoter of *MPG1* or the remnants of the *CmGAS1* promoter, or both, driving ectopic and tissue specific expression in *mpg1* plants.

Future prospective to validate phenotype

Recapitulation of the *mpg1* phenotype through independent molecular analysis via newly generated transgenic plants remains difficult due to a lack of understanding of *MPG1* expression in *mpg1*. Because we originally hypothesized that the mutant's phenotype might have resulted from constitutive overexpression of *MPG1*, transgenic plants were generated containing *MPG1* fused to the maize ubiquitin constitutive overexpression promoter (*ZmUBI*). In the T_0 generation we noticed some phenotypic similarity to the *mpg1*, mainly larger plants and delayed flowering

relative to plants not overexpressing *MPG1*. However, the progeny of these plants containing the *ZmUBIpro*::*MPG1* construct were unable to germinate, as only T_1 wild-type null segregants were recovered and able to produce seedlings (data not shown). These findings along with the lack of expression in root tissue further support that *MPG1* is not constitutively expressed in *mpg1*, and show that constitutive expression renders the plant incapable of producing viable seed. Dissection of this seed could potentially reveal consequences involved in seed development as one of the known native functions of *MPG1* involves the appropriate regulation of embryo development and grain filling. Therefore, altered expression by constitutive expression of *MPG1* might somehow effect these structures. This remains to be an intriguing discovery because *mpg1* exhibits expression of *MPG1* in panicle tissue, suggesting that the manner of expression present in the mutant is necessary to achieve satisfactory seed development while overrepresentation results in deleterious consequences. Implementation of inducible promoter systems driving expression of *MPG1* might be potentially useful in generating plants capable of not only exhibiting the *mpg1* phenotype but also produce viable seed. Decisive control through temporal regulation of *MPG1* might allow for fine-tuning of the phenotype and a better understanding of when expression is important for the generation of particular features and characteristics.

Northern blot tissue printing or *in-situ* hybridization of *MPG1* in the mutant could reveal the tissue-specific localization of expression in *mpg1*. Performing this technique along different developmental stages within varying tissue types could better allow us to understand exactly when and where *MPG1*'s expression occurs in the mutant. However, even if tissue- and/or temporal-specific expression of *MPG1* is revealed, it will still be difficult to ascertain the molecular nature of how it is being directed. The remaining duplicated *CaMV35S* promoter/enhancer could be acting on the remaining *CmGAS1* promoter or on *MPG1*'s native

promoter deriving altered expression. Independent assessment of both of these promoters with the influence of the 35S enhancer and/or insertion of the entire truncated T-DNA event and surrounding genomic location could be used to try and replicate the phenotype. Alternatively, portions of the remaining genomic sequence between or around the T-DNA insertion site and *MPG1*, could be influencing the manner of expression in *mpg1*, which would be difficult to test using a recombinant approach.

Because the nature of expression of *MPG1* remains unclear in *mpg1*, an alternate method to assess functional elements responsible could be achieved through modifying the mutant line itself. More specifically, CRISPR/Cas9 technology could be used to delete specific regions of interest within *mpg1*. Subsequently the deleted regions could be re-inserted to validate their effect. CRISPR/Cas9-mediated deletion has been successful in deleting large genomic fragments in plants such as soybean and rice (Cai et al., 2018, Zhou et al., 2014a). Through walking and deleting portions of the T-DNA insertion and/or proximal genomic DNA, and characterization of the effects resultant could be used to determine necessary and sufficient sequences for generating the phenotype of *mpg1*. This approach could not only be used to increase our confidence in *mpg1*'s elevated expression of *MPG1* resulting in the phenotype but be used to complement other verification methods and better understand sequences and molecular mechanisms deriving the nature of expression.

Phenotypic assessment of *mpg1* in direct comparison to wild-type has proven difficult because of its developmental delay. Evaluating *MPG1*'s effects within a loss-of-flowering mutant population could lead to a better understanding of biomass accumulation with respect to vegetative growth period. This type of comparison could also be useful in assessing the effects of specific stressors, as we would be able to insure that both plants modulating *MPG1* and wild-

type would be in similar stages of development. Double knockdown lines for *RICE FLOWERING LOCUS T 1* (*RFT1*) and *Heading date 3a* (*Hd3a*) by RNAi have shown to be successful in generating Kitaake with a complete defect in floral transition (Komiya et al., 2008). Although using lines like this should be successful in evaluating *MPG1*'s influence on plants outside of delay in flowering, other Kitaake with delay in flowering have been shown to enhance yield metrics, similar to what is observed in *mpg1*. Plants overexpressing DHD1, a floral repressor upstream of *Ehd1, Hd3a,* and *RFT1* that directly acts on *OsHAP5C* and *OsHAP5D* results in a 2 to 3 week delay in flowering resulting in plants exhibiting greater height, panicle length, panicle branch number, and grains per panicle (Zhang et al., 2019). Although the authors didn't measure culm size or total biomass, pictorial representation of their data suggest that these metrics were increased as well. This provokes the idea that the pleiotropic phenotype observed in *mpg1* might be a direct result of the manner by which flowering is delayed by ectopic expression of *MPG1*.

Although it is important to understand how the T-DNA insertion led to the increased expression of *MPG1*, it is dually important to uncover what systematic influence its ectopic expression is having on the plant as a whole. Because *MPG1* is a transcription factor, it functions to target and influence the expression of other genes. As *MPG1* is non-constitutively up-regulated and likely influenced by stress and development, evaluation of global differences in gene expression under similar states of development (between *mpg1* and wild-type) might lead to the discovery of impacted genes resulting in the phenotype. Although *MPG1*'s expression might effect *mpg1* in complex and distinct ways throughout its cycle of development, both *mpg1* and wild-type appear to be under similar states of development early in vegetative growth. Assessment of differentially regulated genes at this timepoint, particularly within tissues that

greatly differ between the mutant and wild-type, could provide meaningful information in the form of transcriptomic differences caused by the mutation and ectopic expression of *MPG1*. Identification and evaluation of the effected genes and direct targets of *MPG1*, could allow for further hypothesis generation and tentative directions to take this study forward by uncovering molecular mechanism relevant to the enhancement of plant productivity.

The outlook on *mpg1* remains promising

We originally sought after a way to make plants larger, and have now identified a plant with increased biomass as a result of a T-DNA insertion mutation exhibited by ectopic expression of *MPG1* (an AP2/EREBP transcription factor). Regardless of pinpointing the individual molecular mechanisms pertinent to the *mpg1* phenotype, this mutant continues to be a much larger plant than wild-type. It has the additional benefits of generating increased seed yield, possesses enhanced ratooning capabilities, and potentially exhibits a degree of stress tolerance. The implementation of this discovery could aid in the generation of domestic and international bioenergy feedstock supplies and produce greater levels of food globally. Its use could not only increase basal levels of resources but also improve those in areas with inclement environmental and social conditions.

Rice is considered the most important human food crop in the world, as it directly feeds more people than any other crop (FAO, 2013b). More than half of the world's population uses rice on a daily basis, being particularly prevalent in areas of poor economic status such as Asia, Africa, and Latin America.

The only potentially negative factor in its use in the commercial sector is its

delay in flowering. The seasonal timing of planting, growth, and harvest can matter in the production of grain. If rice is planted too early, plants can experience slower emergence, poor growth, decreased seed vigor, increased seedling disease damage and increased bird predation. If rice is planted too late, plants can experience decreases in yield, reduced grain quality, panicle blight, increased incidences for disease and insect damage, and reduced success with ratooning (Lack et al., 2012, Slaton et al., 2003). Addressing whether or not the required additional time to grow *mpg1* plants will be suitable or detrimental in the process of cultivating rice is difficult to determine.

Modern rice production usually occurs as an annual plant, although in certain temperate regions can produce successful ratooned crops for decades and be treated as a perennial. A large number of cultivated rice varieties exist, and are regionally grown for their successful generation of grain. Inherently, a number of these varieties also have different lengths of total growth as a measurement from seed to seed. However, modern cultivation practices commonly account for roughly 30+ days for sowing and 30+ days for harvest (DRD, 2002). This allotted time, along with precise implementation of transplanting times, could be sufficient in accounting for the necessary growth period needed for the application of *mpg1* in a real-world setting. Strategic use of ratooned crops could also be managed to potentially increase yields and residues in appropriate environments by increasing the number of crops harvested per unit of time.

Perhaps crossing *mpg1* with other desirable varieties of rice and generating

recombinant inbred lines (RILs) would prove particularly valuable. Not only would this serve as a means to validate the *mpg1* region of interest with the phenotype, but also potentially tailor rice phenotypes for intensified cropping and increased yield with limited hindering factors.

BIBLIOGRAPHY

ALONSO, J. M., STEPANOVA, A. N., LEISSE, T. J., KIM, C. J., CHEN, H., SHINN, P., STEVENSON, D. K., ZIMMERMAN, J., BARAJAS, P., CHEUK, R., GADRINAB, C., HELLER, C., JESKE, A., KOESEMA, E., MEYERS, C. C., PARKER, H., PREDNIS, L., ANSARI, Y., CHOY, N., DEEN, H., GERALT, M., HAZARI, N., HOM, E., KARNES, M., MULHOLLAND, C., NDUBAKU, R., SCHMIDT, I., GUZMAN, P., AGUILAR-HENONIN, L., SCHMID, M., WEIGEL, D., CARTER, D. E., MARCHAND, T., RISSEEUW, E., BROGDEN, D., ZEKO, A., CROSBY, W. L., BERRY, C. C. & ECKER, J. R. 2003. Genome-wide insertional mutagenesis of Arabidopsis thaliana. *Science,* 301**,** 653-7.

ARITE, T., UMEHARA, M., ISHIKAWA, S., HANADA, A., MAEKAWA, M., YAMAGUCHI, S. & KYOZUKA, J. 2009. d14, a strigolactone-insensitive mutant of rice, shows an accelerated outgrowth of tillers. *Plant Cell Physiol,* 50**,** 1416-24.

AYRE, B. G., BLAIR, J. E. & TURGEON, R. 2003. Functional and phylogenetic analyses of a conserved regulatory program in the phloem of minor veins. *Plant Physiol,* 133**,** 1229-39.

BEILMANN, A., ALBRECHT, K., SCHULTZE, S., WANNER, G. & PFITZNER, U. M. 1992. Activation of a truncated PR-1 promoter by endogenous enhancers in transgenic plants. *Plant Mol Biol,* 18**,** 65-78.

BENFEY, P. N., REN, L. & CHUA, N. H. 1989. The CaMV 35S enhancer contains at least two domains which can confer different developmental and tissue-specific expression patterns. *EMBO J,* 8**,** 2195-2202.

BENFEY, P. N., REN, L. & CHUA, N. H. 1990a. Combinatorial and synergistic properties of CaMV 35S enhancer subdomains. *EMBO J,* 9**,** 1685-96.

BENFEY, P. N., REN, L. & CHUA, N. H. 1990b. Tissue-specific expression from CaMV 35S enhancer subdomains in early stages of plant development. *EMBO J,* 9**,** 1677-84.

CAI, Y., CHEN, L., SUN, S., WU, C., YAO, W., JIANG, B., HAN, T. & HOU, W. 2018. CRISPR/Cas9-Mediated Deletion of Large Genomic Fragments in Soybean. *Int J Mol Sci,* 19.

CASTLE, L. A., ERRAMPALLI, D., ATHERTON, T. L., FRANZMANN, L. H., YOON, E. S. & MEINKE, D. W. 1993. Genetic and molecular characterization of embryonic mutants identified following seed transformation in Arabidopsis. *Mol Gen Genet,* 241**,** 504-14.

CHEN, Z., CHENG, Q., HU, C., GUO, X., CHEN, Z., LIN, Y., HU, T., BELLIZZI, M., LU, G., WANG, G. L., WANG, Z., CHEN, S. & WANG, F. 2017a. A Chemical-Induced, Seed-Soaking Activation Procedure for Regulated Gene Expression in Rice. *Front Plant Sci,* 8**,** 1447.

CHEN, Z., KONG, L., ZHOU, Y., CHEN, Z., TIAN, D., LIN, Y., WANG, F. & CHEN, S. 2017b. Endosperm-specific OsPYL8 and OsPYL9 act as positive regulators of the ABA signaling pathway in rice seed germination. *Funct Plant Biol,* 44.

CHRISTIE, P. J. 2004. Type IV secretion: the Agrobacterium VirB/D4 and related conjugation systems. *Biochim Biophys Acta,* 1694**,** 219-34.

CHRISTIE, P. J. & VOGEL, J. P. 2000. Bacterial type IV secretion: conjugation systems adapted to deliver effector molecules to host cells. *Trends Microbiol,* 8**,** 354-60.

COOPER, B., CLARKE, J. D., BUDWORTH, P., KREPS, J., HUTCHISON, D., PARK, S., GUIMIL, S., DUNN, M., LUGINBUHL, P., ELLERO, C., GOFF, S. A. & GLAZEBROOK, J. 2003. A network of rice genes associated with stress response and seed development. *Proc Natl Acad Sci U S A,* 100**,** 4945-50.

DAFNY-YELIN, M., LEVY, A. & TZFIRA, T. 2008. The ongoing saga of Agrobacterium-host interactions. *Trends Plant Sci,* 13**,** 102-5.

DASGUPTA, K., KHADILKAR, A. S., SULPICE, R., PANT, B., SCHEIBLE, W. R., FISAHN, J., STITT, M. & AYRE, B. G. 2014. Expression of Sucrose Transporter cDNAs Specifically in Companion Cells Enhances Phloem Loading and Long-Distance Transport of Sucrose but Leads to an Inhibition of Growth and the Perception of a Phosphate Limitation. *Plant Physiol,* 165**,** 715-731.

DOI, K., IZAWA, T., FUSE, T., YAMANOUCHI, U., KUBO, T., SHIMATANI, Z., YANO, M. & YOSHIMURA, A. 2004. Ehd1, a B-type response regulator in rice, confers short-day promotion of flowering and controls FT-like gene expression independently of Hd1. *Genes Dev,* 18**,** 926-36.

DRD 2002. Rice in India: A Status Paper. Patna, Bihar: Government of India - Ministry of Agriculture & Farmer's Welfare.

ELLIS, J., DODDS, P. & PRYOR, T. 2000. Structure, function and evolution of plant disease resistance genes. *Curr Opin Plant Biol,* 3**,** 278-284.

FAO 2013. FAO Statistical Yearbook 2013: World food and agriculture. Rome: Food and Agriculture Organization of the United Nations.

GAO, J., CHAO, D. & LIN, H. 2007. Understanding abiotic stress tolerance mechanisms: recent studies on stress response in rice. *J Integr Plant Biol,* 49**,** 742-750.

GELVIN, S. B. 2000. Agrobacterium and Plant Genes Involved in T-DNA Transfer and Integration. *Annu Rev Plant Physiol Plant Mol Biol,* 51**,** 223-256.

GELVIN, S. B. 2003. Agrobacterium-Mediated Plant Transformation: the Biology behind the "Gene-Jockeying" Tool. *Microbiol Mol Biol R,* 67**,** 16-37.

GHEYSEN, G., VILLARROEL, R. & VAN MONTAGU, M. 1991. Illegitimate recombination in plants: a model for T-DNA integration. *Genes Dev,* 5**,** 287-97.

GUDYNAITE-SAVITCH, L., JOHNSON, D. A. & MIKI, B. L. 2009. Strategies to mitigate transgene-promoter interactions. *Plant Biotechnol J,* 7**,** 472-85.

HARITATOS, E., AYRE, B. G. & TURGEON, R. 2000. Identification of phloem involved in assimilate loading in leaves by the activity of the galactinol synthase promoter. *Plant Physiol,* 123**,** 929-37.

HASAN, M. M., RAFII, M. Y., ISMAIL, M. R., MAHMOOD, M., RAHIM, H. A., ALAM, M. A., ASHKANI, S., MALEK, M. A. & LATIF, M. A. 2015. Marker-assisted backcrossing: a useful method for rice improvement. *Biotechnol Biotechnol Equip,* 29**,** 237-254.

HASEGAWA, P. M., BRESSAN, R. A., ZHU, J. & BOHNERT, H. J. 2000. Plant cellular and molecular responses to high salinity. *Annu Rev Plant Physiol Plant Mol Biol,* 51**,** 463-99.

HEDLEY, M. J. & BOLAN, N. S. 2003. Role of carbon, nitrogen, and sulfur cycles in soil acidification. *In:* RENGEL, Z. (ed.) *Handbook of Soil Acidity.* CRC Press - Technology & Engineering.

HIRAI, G., OKUMURA, T., TAKEUCHI, S., TANAKA, O. & CHUJO, H. 2000. Studies on the effect of the relative humidity of the atmosphere on the growth and physiology of rice plants. *Plant Prod Sci,* 3**,** 129-133.

HSIAO, T. C., O'TOOLE, J. C., YAMBAO, E. B. & TURNER, N. C. 1984. Influence of osmotic adjusment of leaf roling and tissue death in rice (Oryza sativa L.). *Plant Physiol,* 75**,** 338-341.
KAY, R., CHAN, A., DALY, M. & MCPHERSON, J. 1987. Duplication of CaMV 35S promoter sequences creates a strong enhancer for plant genes. *Science,* 236**,** 1299-1302.
KLOTZBÜCHER, T., MARXEN, A., VETTERLEIN, D., SCHNEIKER, J., TÜRKE, M., VAN SINH, N., MANH, N. H., VAN CHIEN, H., MARQUEZ, L., VILLAREAL, S., BUSTAMANTE, J. V. & JAHN, R. 2015. Plant-available silicon in paddy soils as a key factor for sustainable rice production in Southeast Asia. *Basic Appl Ecol,* 16**,** 665-673.
KOMIYA, R., IKEGAMI, A., TAMAKI, S., YOKOI, S. & SHIMAMOTO, K. 2008. Hd3a and RFT1 are essential for flowering in rice. *Development,* 135**,** 767-74.
LACK, S., MARANI, N. M. & MOMBENI, M. 2012. The effects of planting date on grain yield and yield components of rice cultivars. *Adv Environ Biol,* 6**,** 406-413.
LACROIX, B., LOYTER, A. & CITOVSKY, V. 2008. Association of the Agrobacterium T-DNA-protein complex with plant nucleosomes. *Proc Natl Acad Sci U S A,* 105**,** 15429-34.
LEE, Y. S. & GYNHEUNG, A. 2015. Complex Regulatory Networks of Flowering Time in Rice. *Rice Research: Open Access,* 03.
LI, J., KRICHEVSKY, A., VAIDYA, M., TZFIRA, T. & CITOVSKY, V. 2005. Uncoupling of the functions of the Arabidopsis VIP1 protein in transient and stable plant genetic transformation by Agrobacterium. *Proc Natl Acad Sci U S A,* 102**,** 5733-8.
LIN, H., WANG, R., QIAN, Q., YAN, M., MENG, X., FU, Z., YAN, C., JIANG, B., SU, Z., LI, J. & WANG, Y. 2009. DWARF27, an iron-containing protein required for the biosynthesis of strigolactones, regulates rice tiller bud outgrowth. *Plant Cell,* 21**,** 1512-25.
LOYTER, A., ROSENBLUH, J., ZAKAI, N., LI, J., KOZLOVSKY, S. V., TZFIRA, T. & CITOVSKY, V. 2005. The plant VirE2 interacting protein 1. a molecular link between the Agrobacterium T-complex and the host cell chromatin? *Plant Physiol,* 138**,** 1318-21.
LU, J. M. & BUSH, D. R. 1998. His-65 in the proton-sucrose symporter is an essential amino acid whose modification with site-directed mutagenesis increases transport activity. *Proc Natl Acad Sci U S A,* 95**,** 9025-9030.
MATSUDA, Y., LIANG, G., ZHU, Y., MA, F., NELSON, R. S. & DING, B. 2002. The Commelina Yellow Mottle Virus promoter drives companion-cell-specific gene expression in multiple organs of tansgenic tobacco. *Protoplasma,* 220**,** 0051-0058.
MAYERHOFER, R., KONCZ-KALMAN, Z., NAWRATH, C., BAKKEREN, G., CRAMERI, A., ANGELIS, K., REDEI, G. P., SCHELL, J., HOHN, B. & KONCZ, C. 1991. T-DNA integration: a mode of illegitimate recombination in plants. *EMBO J,* 10**,** 697-704.
MEDBERRY, S. L., LOCKHART, B. E. L. & OLSZEWSKI, N. E. 1992. The Commelina Yellow Mottle Virus promoter is a strong promoter in vascular and reproductive tissues. *Plant Cell,* 4**,** 185-192.
MENKENS, A. E., SCHINDLER, U. & CASHMORE, A. R. 1995. The G-box: a ubiquitous regulatory DNA element in plants bound by the GBF family of bZIP proteins. *Trends Biochem Sci,* 20**,** 506-10.
MINAKUCHI, K., KAMEOKA, H., YASUNO, N., UMEHARA, M., LUO, L., KOBAYASHI, K., HANADA, A., UENO, K., ASAMI, T., YAMAGUCHI, S. & KYOZUKA, J. 2010.

FINE CULM1 (FC1) works downstream of strigolactones to inhibit the outgrowth of axillary buds in rice. *Plant Cell Physiol,* 51**,** 1127-35.
MOSAIC. 2018. *Soil pH* [Online]. Available: https://www.cropnutrition.com/efu-soil-ph [Accessed].
NAKANO, T., SUZUKI, K., FUJIMURA, T. & SHINSHI, H. 2006. Genome-wide analysis of the ERF gene family in Arabidopsis and rice. *Plant Physiol,* 140**,** 411-32.
O'TOOLE, J. C. & CRUZ, R. T. 1980. Response of leaf water potential, stomatal resistance, and leaf rolling to water stress. *Plant Physiol,* 65**,** 428-432.
OHME-TAKAGI, M. & SHINSHI, H. 1995. Ethylene-inducible DNA binding proteins that interact with an ethylene responsive element. *Plant Cell,* 7**,** 173-182.
OHTA, M., HAYASHI, Y., NAKASHIMA, A., HAMADA, A., TANAKA, A., NAKAMURA, T. & HAYAKAWA, T. 2002. Introduction of a Na+/H+ antiporter gene from Atriplex gmelini confers salt tolerance to rice. *FEBS Lett,* 532**,** 279-282.
OHTSUKI, S., LEVINE, M. & CAI, H. N. 1998. Different core promoters possess distinct regulatory activities in the Drosophila embryo. *Genes Dev,* 12**,** 547-56.
OUYANG, S., ZHU, W., HAMILTON, J., LIN, H., CAMPBELL, M., CHILDS, K., THIBAUD-NISSEN, F., MALEK, R. L., LEE, Y., ZHENG, L., ORVIS, J., HAAS, B., WORTMAN, J. & BUELL, C. R. 2007. The TIGR Rice Genome Annotation Resource: improvements and new features. *Nucleic Acids Res,* 35**,** D883-7.
PANSEGRAU, W., SCHOUMACHER, F., HOHN, B. & LANKA, E. 1993. Site-specific cleavage and joining of single-stranded DNA by VirD2 protein of Agrobacterium tumefaciens Ti plasmids: analogy to bacterial conjugation. *Proc Natl Acad Sci U S A,* 90**,** 11538-42.
PASZKOWSKI, J., BAUR, M., BOGUCKI, A. & POTRYKUS, I. 1988. Gene targeting in plants. *EMBO J,* 7**,** 4021-6.
PIERRE, W. H. 1928. Nitrogenous fertilizers and soil acidity. I. Effect of various nitrogenous fertilizers on soil reaction. *Food and Agriculture Organization of the United States*
PITZSCHKE, A. 2013. Agrobacterium infection and plant defense-transformation success hangs by a thread. *Front Plant Sci,* 4**,** 519.
PITZSCHKE, A. & HIRT, H. 2010. New insights into an old story: Agrobacterium-induced tumour formation in plants by plant transformation. *EMBO J,* 29**,** 1021-32.
PRIYA, P. & JAIN, M. 2013. RiceSRTFDB: a database of rice transcription factors containing comprehensive expression, cis-regulatory element and mutant information to facilitate gene function analysis. *Database (Oxford),* 2013**,** bat027.
QIU, D., XIAO, J., DING, X., XIONG, M., CAI, M., CAO, Y., LI, X., XU, C. & WANG, S. 2007. OsWRKY13 mediates rice disease resistance by regulating defense-related genes in salicylate- and jasmonate-dependent signaling. *Mol Plant Microbe In,* 20**,** 492-499.
RASHID, M., GUANGYUAN, H., GUANGXIAO, Y., HUSSAIN, J. & XU, Y. 2012. AP2/ERF Transcription Factor in Rice: Genome-Wide Canvas and Syntenic Relationships between Monocots and Eudicots. *Evol Bioinform Online,* 8**,** 321-55.
REDDY, I. N. B. L., KIM, B.-K., YOON, I.-S., KIM, K.-H. & KWON, T.-R. 2017. Salt Tolerance in Rice: Focus on Mechanisms and Approaches. *Rice Science,* 24**,** 123-144.
RIECHMANN, J. L. & MEYEROWITZ, E. M. 1998. The AP2/EREBP family of plant transcription factors. *Biol Chem,* 379**,** 633-46.
ROJAS, C. A., HEMERLY, A. S. & FERREIRA, P. C. 2010. Genetically modified crops for biomass increase. Genes and strategies. *GM Crops,* 1**,** 137-42.

ROSSO, M. G., LI, Y., STRIZHOV, N., REISS, B., DEKKER, K. & WEISSHAAR, B. 2003. An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for flankin sequence tag-based reverse genetics. *Plant Mol Biol,* 53**,** 347-259.

SAIJO, Y., HATA, S., KYOZUKA, J., SHIMAMOTO, K. & IZUI, K. 2000. Over-expression of a single Ca2+-dependent protein kinase confers both cold and salt/drought tolerance on rice plants. *Plant J,* 23**,** 319-327.

SAKUMA, Y., LIU, Q., DUBOUZET, J. G., ABE, H., SHINOZAKI, K. & YAMAGUCHI-SHINOZAKI, K. 2002. DNA-binding specificity of the ERF/AP2 domain of Arabidopsis DREBs, transcription factors involved in dehydration- and cold-inducible gene expression. *Biochem Biophys Res Commun,* 290**,** 998-1009.

SCHEIFFELE, P., PANSEGRAU, W. & LANKA, E. 1995. Initiation of Agrobacterium tumefaciens T-DNA processing. Purified proteins VirD1 and VirD2 catalyze site- and strand-specific cleavage of superhelical T-border DNA in vitro. *J Biol Chem,* 270**,** 1269-76.

SENGUPTA, S. & MAJUMDER, A. L. 2009. Insight into the salt tolerance factors of a wild halophytic rice, Porteresia coarctata: a physiological and proteomic approach. *Planta,* 229**,** 911-29.

SHARONI, A. M., NURUZZAMAN, M., SATOH, K., SHIMIZU, T., KONDOH, H., SASAYA, T., CHOI, I. R., OMURA, T. & KIKUCHI, S. 2011. Gene structures, classification and expression models of the AP2/EREBP transcription factor family in rice. *Plant Cell Physiol,* 52**,** 344-60.

SINGER, S. D., COX, K. D. & LIU, Z. 2011. Enhancer-promoter interference and its prevention in transgenic plants. *Plant Cell Rep,* 30**,** 723-31.

SLATON, N. A., LINSCOMBE, S. D., NORMAN, R. J. & GBUR, E. E. 2003. Seeding date effect on rice grain yields in Arkansas and Louisiana. *Agron J,* 95**,** 218-223.

SMITH, E. F. & TOWNSEND, C. O. 1907. A Plant-Tumor of Bacterial Origin. *Science,* 25**,** 671-3.

STACHEL, S. E., MESSENS, E., VAN MONTAGU, M. & ZAMBRYSKI, P. 1985. Identification of the signal molecules produced by wounded plant cells that activate T-DNA transfer in Agrobacterium tumefaciens. *Nature***,** 624-629.

TZFIRA, T. & CITOVSKY, V. 2000. From host recognition to T-DNA integration: the function of bacterial and plant genes in the Agrobacterium-plant cell interaction. *Mol Plant Pathol,* 1**,** 201-12.

TZFIRA, T. & CITOVSKY, V. 2002. Partners-in-infection: host proteins involved in the transformation of plant cells by Agrobacterium. *Trends Cell Biol,* 12**,** 121-9.

TZFIRA, T., LI, J., LACROIX, B. & CITOVSKY, V. 2004a. Agrobacterium T-DNA integration: molecules and models. *Trends Genet,* 20**,** 375-83.

TZFIRA, T., VAIDYA, M. & CITOVSKY, V. 2001. VIP1, an Arabidopsis protein that interacts with Agrobacterium VirE2, is involved in VirE2 nuclear import and Agrobacterium infectivity. *EMBO J,* 20**,** 3596-3607.

TZFIRA, T., VAIDYA, M. & CITOVSKY, V. 2002. Increasing plant susceptibility to Agrobacterium infection by overexpression of the Arabidopsis nuclear protein VIP1. *Proc Natl Acad Sci U S A,* 99**,** 10435-40.

TZFIRA, T., VAIDYA, M. & CITOVSKY, V. 2004b. Involvement of targeted proteolysis in plant genetic transformation by Agrobacterium. *Nature,* 431**,** 87-92.

USDA 2011. Soil Quality Indicators. *In:* SERVICE, N. R. C. (ed.). https://www.nrcs.usda.gov/wps/PA_NRCSConsumption/download?cid=nrcs142p2_053136&ext=pdf: USDA.

VAN CAMP, W. 2005. Yield enhancement genes: seeds for growth. *Curr Opin Biotechnol,* 16**,** 147-53.

VOLK, G. M., HARITATOS, E. E. & TURGEON, R. 2003. Galactinol synthase gene expression in melon. *J Amer Soc Hort Sci* 128**,** 8-15.

WEATHERSPARK.COM. 2019. *Average Weather in Fort Collins Colorado, United States. Accessed at: https://weatherspark.com/y/3565/Average-Weather-in-Fort-Collins-Colorado-United-States-Year-Round* [Online]. [Accessed].

WEIGEL, D., AHN, J. H., BLAZQUEZ, M. A., BOREVITZ, J. O., CHRISTENSEN, S. K., FANKHAUSER, C., FERRANDIZ, C., KARDAILSKY, I., MALANCHARUVIL, E. J., NEFF, M. M., NGUYEN, J. T., SATO, S., WANG, Z. Y., XIA, Y., DIXON, R. A., HARRISON, M. J., LAMB, C. J., YANOFSKY, M. F. & CHORY, J. 2000. Activation tagging in Arabidopsis. *Plant Physiol,* 122**,** 1003-13.

WOPEREIS, M. C. S., KROPFF, M. J., MALIGAYA, A. R. & TUONG, T. P. 1996. Drought-stress responses of two lowland rice cultivars to soil water status. *Field Crop Res,* 46**,** 21-39.

XU, J. J., ZHANG, X. F. & XUE, H. W. 2016. Rice aleurone layer specific OsNF-YB1 regulates grain filling and endosperm development by interacting with an ERF transcription factor. *J Exp Bot,* 67**,** 6399-6411.

XU, Z. S., CHEN, M., LI, L. C. & MA, Y. Z. 2011. Functions and application of the AP2/ERF transcription factor family in crop improvement. *J Integr Plant Biol,* 53**,** 570-85.

YOO, S. Y., BOMBLIES, K., YOO, S. K., YANG, J. W., CHOI, M. S., LEE, J. S., WEIGEL, D. & AHN, J. H. 2005. The 35S promoter used in a selectable marker gene of a plant transformation vector affects the expression of the transgene. *Planta,* 221**,** 523-30.

YU, T. R. 1991. Characteristics of soil acidity of paddy soils in relation to rice growth. *In:* WRIGHT, R. J., BALIGAR, V. C. & MURRMANN, R. P. (eds.) *Plant-Soil Interactions at Low pH. Dev Plant Soil Sci.* Dordrecht: Springer.

ZALTSMAN, A., KRICHEVSKY, A., KOZLOVSKY, S. V., YASMIN, F. & CITOVSKY, V. 2010a. Plant defense pathways subverted by Agrobacterium for genetic transformation. *Plant Signal Behav,* 5**,** 1245-8.

ZALTSMAN, A., KRICHEVSKY, A., LOYTER, A. & CITOVSKY, V. 2010b. Agrobacterium induces expression of a host F-box protein required for tumorigenicity. *Cell Host Microbe,* 7**,** 197-209.

ZALTSMAN, A., LACROIX, B., GAFNI, Y. & CITOVSKY, V. 2013. Disassembly of synthetic Agrobacterium T-DNA-protein complexes via the host SCF(VBF) ubiquitin-ligase complex pathway. *Proc Natl Acad Sci U S A,* 110**,** 169-74.

ZAMBRYSKI, P. C. 1992. Chronicles from the Agrobacterium-plant cell DNA transfer story. *Ann Rev Plant Physiol and Plant Mol Biol,* 43**,** 456-490.

ZHANG, H., ZHU, S., LIU, T., WANG, C., CHENG, Z., ZHANG, X., CHEN, L., SHENG, P., CAI, M., LI, C., WANG, J., ZHANG, Z., CHAI, J., ZHOU, L., LEI, C., GUO, X., WANG, J., WANG, J., JIANG, L., WU, C. & WAN, J. 2019. DELAYED HEADING DATE1 interacts with OsHAP5C/D, delays flowering time and enhances yield in rice. *Plant Biotechnol J,* 17**,** 531-539.

ZHENG, X., DENG, W., LUO, K., DUAN, H., CHEN, Y., MCAVOY, R., SONG, S., PEI, Y. & LI, Y. 2007. The cauliflower mosaic virus (CaMV) 35S promoter sequence alters the level and patterns of activity of adjacent tissue- and organ-specific gene promoters. *Plant Cell Rep,* 26**,** 1195-203.

ZHOU, H., LIU, B., WEEKS, D. P., SPALDING, M. H. & YANG, B. 2014a. Large chromosomal deletions and heritable small genetic changes induced by CRISPR/Cas9 in rice. *Nucleic Acids Res,* 42**,** 10903-14.

ZHOU, J., YU, F., WANG, X., YANG, Y., YU, C., LIU, H., CHENG, Y., YAN, C. & CHEN, J. 2014b. Specific expression of DR5 promoter in rice roots using a tCUP derived promoter-reporter system. *PLoS One,* 9**,** e87008.

ZIEMIENOWICZ, A., TINLAND, B., BRYANT, J., GLOECKLER, V. & HOHN, B. 2000. Plant enzymes but not Agrobacterium VirD2 mediate T-DNA ligation in vitro. *Mol Cell Biol,* 20**,** 6317-6322.

ZUPAN, J., MUTH, T. R., DRAPER, O. & ZAMBRYSKI, P. 2000. The transfer of DNA from Agrobacterium tumefaciens into plants: a feast of fundamental insights. *Plant J,* 23**,** 11-28.

CHAPTER 4: Comparative transcriptomic assessment of *mpg1* mutant plants vs. wild-type null segregant *Oryza sativa* in whole tiller tissue

Introduction

Transcription factors are important regulatory elements that function to control the expression of other genes. More specifically they are *trans*-acting proteins that acts on *cis*-acting elements (genomic sequences) which serve to modulate another gene(s) transcription. Transcription factors are crucial for a variety of reasons; they participate in growth and development, as well as respond to stimuli. One transcription factor can target numerous other genes (including other transcription factors), which can result in cascading effects within biochemical processes and physiological functions. Investigation and characterization of transcription factors could play a significant role in understanding key regulators and potentially their use in translational biology.

A particular transcription factor presented itself of noteworthy interest. *MPG1*, also known as *OsERF#115*, is a transcription factor belonging to the APETALA 2/Ethylene-Responsive Element-Binding Protein (AP2/EREBP) gene superfamily, under the ERF gene family, within the ERF subfamily in rice (Nakano et al., 2006). Characterization of this gene originally began through the analysis of a mutant rice plant, *mpg1* (*makes plants gigantic-1*), generated from a bi-laterally truncated T-DNA expression cassette that exhibits an interesting pleiotropic phenotype. This plant accumulates greater biomass, seed yield, and experiences a delay in flowering compared to wild-type plants. Ratooning capabilities are also higher in *mpg1* plants. Additionally, *mpg1* plants exhibit a greater delta in phenotype relative to wild-type under non-optimal conditions, compared to plants grown under optimum conditions, and are able to accumulate greater biomass under prolonged salt and drought stress.

The T-DNA insertion present in *mpg1* correlated strongly with the elevated expression of *MPG1*, a neighboring gene to the T-DNA insertion site. The reason for enhanced expression by means of the T-DNA insertion has yet to be fully uncovered as the T-DNA integrated in an intergenic region not disrupting any known or annotated functional elements. Expression of *MPG1* in this mutant is a dominant mutation, with elevated expression and presence of the mutant phenotype in plants that are both heterozygous and homozygous for the T-DNA insertion (monitored across several selfed generations and within BC_1F_2 populations). Ectopic expression of this gene in *mpg1* was found in leaf, and stem tissue throughout its life cycle, and in panicle tissue during reproductive development. Wild-type plants and wild-type null segregants had little to no message within these tissues using semi-quantitative RT-PCR and ddPCR methods respectively. Although expression of *MPG1* in the mutant *mpg1* led us to originally believe that it was constitutively overexpressed, this was refuted after transcript abundance in root tissues was found to be little to none.

The ERF subfamily is grouped based upon their encoded transcription factor's expected DNA binding motif. The ERF subfamily binds specifically to the GCC-box, characterized by the motif (GCCGCC). These particular *cis*-acting elements have been found in promoter regions of ethylene-inducible related genes (Ohme-Takagi and Shinshi, 1995). These genes have been shown to play a role in both growth and development, and responses to stress. The broad functions associated with this subfamily of genes aligns well with the phenotype observed in *mpg1*.

Ethylene is a volatile phytohormone that influences plant growth and development (Burg, 1973). This hormone plays roles in flowering (Achard et al., 2007), fruit ripening (Barry and Giovannoni, 2007), senescence, abscission (Burg, 1968), and seed germination (Matilla, 2000).

Ethylene also mediates stress-related responses from pathogens and wounding (van Loon et al., 2006, Leon et al., 2001).

Ethylene is synthesized and regulated under specific conditions, reviewed in (Johnson and Ecker, 1998). The biosynthesis of ethylene begins from the conversion of methionine to *S*-adenosylmethionine (*S*-AdoMet) via SAM synthetase, catalyzed by ATP (Ravanel et al., 1998). The *S*-AdoMet can then be used as a methyl group donor for numerous molecules or act as a precursor to the ethylene biosynthesis pathway (Ravanel et al., 1998). Under the ethylene biosynthesis pathway *S*-AdoMet is converted to ACC (Adams and Yang, 1979) via the ACC synthase (*S*-adenosyl-L-methionine methylthioadenosine-lyase) (Sato and Theologies, 1989). ACC synthase can also produce 5'-methylthioadenosine (MTA), which acts as an intermediate in a salvage pathway to preserve the methyl group for additional cycles of ethylene production without increasing need for a greater pool of methionine (Yang Cycle) (Miyazaki and Yang, 1987). ACC oxidase (ACO) finalizes the generation of ethylene synthesis using ACC as a substrate (Hamilton et al., 1991, Spanu et al., 1991). ACC synthase is an important enzyme in the production of ethylene, whose transcriptional induction is stimulated via several abiotic and biotic cues, and developmental stages (Argueso et al., 2007). ACC oxidase, is the final step in ethylene biosynthesis, and is often considered the rate limiting step especially under times of increased ethylene production. Both ACC synthase and ACC oxidase belong to multigene families (Kende, 1993, Zarembinski and Theologis, 1994, Fluhr et al., 1996), and have been shown to be differentially expressed under different stimulus or developmental periods (Fluhr et al., 1996).

A classic observation of the effect of ethylene is the triple response. The triple response is a physiological observation of plants exposed to ethylene that result in inhibition of elongation of

the internode, increasing thickness of a stem, and loss of gravitropism from an imbalance of auxin (Guzman and Ecker, 1990, Bleecker et al., 1988, Merchante and Stepanova, 2017). These responses have been observed in developing seedlings that are under mechanical stress (direct growth block by physical barrier), as a mechanism to circumvent these physical growth barriers. Additionally etiolated plants subject to ectopic saturation of ethylene or the ethylene precursor ACC result in the triple response (Merchante and Stepanova, 2017), as well as mutant plants with constitutive triple response in absence of ethylene (Kieber et al., 1993, Roman and Ecker, 1995). Characterization of Arabidopsis mutant lines with impaired triple response have provided evidence pertaining to the ethylene signaling pathway.

Ethylene signaling networks continue to be reviewed (Wang et al., 2002, Gallie, 2015). Ethylene is perceived via ethylene receptors on the membrane of the endoplasmic reticulum (Schaller and Bleecker, 1995, Schaller et al., 1995, Rodriguez et al., 1999, Ju et al., 2012). Absence of the hormone results in CONSTITUTIVE TRIPLE RESPOSNE 1 (CTR1) associating with receiver domains of the receptor (Kieber et al., 1993, Clark et al., 1998), maintaining an active state responsible for phosphorylation of ETHYLENE INSENSITIVE 2 (EIN2), which also interacts with the ethylene receptors, inhibiting the induction of ethylene response by EIN2 (Alonso et al., 1999, Ju et al., 2012, Qiao et al., 2012, Wen et al., 2012). A copper transporter, RESPONSIVE-TO-ANTAGONIST1 (RAN1), is responsible for delivering copper, a co-factor necessary for perceiving ethylene (Binder et al., 2010). When ethylene is present, it binds to copper, bound in the receptors N-terminal transmembrane domain, inactivating the ethylene receptor-CTR complex. Ethylene binding inactivates the receptors, deactivating CTR1, a Raf-like kinase, responsible for the repression of EIN2 (Bisson et al., 2009, Bisson and Groth, 2010). The loss of EIN2 phosphorylation results in proteolytic release of its C-terminal domain, which

targets the nucleus (Ju et al., 2012). The EIN2 C-terminal domain once present in the nucleus prevents degradation of the ethylene response transcription factors, ETHYLENE INSENSITIVE 3, and ETHYLENE INSENSITIVE LIKE-1 (EIN3/EIL1) by two F-box proteins (EBF1, EBF2) and results in the induction of their degradation (Ju et al., 2012, Qiao et al., 2012, Wen et al., 2012). Dimerization of EIN3/EIL1 is followed by their binding to the promoter region of ethylene response factor (ERF) transcription factor genes. EIN3 binds directly to the promoter of ETHYLENE RESPONSE FACTOR 1 (ERF1) and targets genes in an ethylene dependent style (Ju et al., 2012, Qiao et al., 2012, Wen et al., 2012). This results in the modulated transcription of various downstream ethylene responsive genes (Gallie, 2015). ERF1 and other EREBP transcription factors have been shown to target GCC-box motifs in the promoter elements of genes in response to ethylene.

Interestingly, many EREBP transcription factors have been found in plant species (Riechmann and Meyerowitz, 1998) but only a small portion of them have been shown to be regulated by ethylene (Thara et al., 1999, Yamamoto et al., 1999). Rather, other things have been shown to regulate the expression of these transcription factors. Stressors, such as salt and drought, as well as the introduction of SA and JA, are stimuli that have been shown to regulate these genes (Ohme-Takagi and Shinshi, 1995, Buttner and Singh, 1997, Suzuki et al., 1998, Thara et al., 1999, Fujimoto et al., 2000, Gu et al., 2000). Therefore it has been suggested that the GCC-box motif functions as a general transcriptional regulatory element for genes with the EREBP transcription factor superfamily and is not exclusively specific to ethylene response (Wang et al., 2002).

The most notable characteristic present in *mpg1* is an increase in overall biomass (dry weight) compared to wild-type. The enhanced biomass observed in *mpg1* comes from a

combination of increased height, increased girth, and increased leaf size. The most significant biomass contributor is from increased girth. A potential mechanism of enhanced terminal biomass accumulation might be the result of an extended vegetative growth phase in *mpg1*, as it experiences a significant delay in flowering compared to wild-type (~14 days). However, *mpg1* plants accumulate greater metrics of growth prior to wild-type's stage of panicle immergence indicating that the increase in biomass might not be solely due to a prolonged vegetative growth phase. The *mpg1* mutant is also capable of generating greater biomass than wild-type post ratooning as well as when exposed to prolonged salt and drought stress. Interestingly, original observations of this mutant also revealed that it was able to accumulate a greater difference in biomass than wild-type under sub-optimal conditions compared to optimal conditions, further suggesting an interaction with stress. The pleiotropic phenotype visualized in *mpg1* aligns similarly with the functionality of genes from the AP2/ERF transcription factors. It is likely that *MPG1* acts as a transcriptional regulator for growth and development, as well as response to stimuli/stress.

It is important to recognizes that the nature of the phenotype in *mpg1* is pleiotropic. Regulation of the 'regulators' can widely and diversely influence plant phenotypes, as modulation of transcription factors often results in pleiotropic phenotypes by influencing a number of downstream genes (Wang et al., 2005, Okushima et al., 2005a, Doebley and Lukens, 1998, Mawlong et al., 2015, Yaish et al., 2010).

One of the clear phenotypes observed in *mpg1* is a delay in flowering. Genes involved in transition to flowering and floral initiation have been somewhat well analyzed (Tsuji et al., 2011, Tsuji et al., 2013, Endo-Higashi and Izawa, 2011, Sui et al., 2013, Sun et al., 2014a, Kobayashi et al., 2012, Tamaki et al., 2015, Cho et al., 2017, Lee and An, 2015). Flowering in rice functions

by environmental cues by day length, generating florigens. Rice, in general, is recognized as a short-day flowering plant, however variability exists between cultivars (Poonyarit et al., 1988, Li et al., 1995). Kitaake, the cultivar used as the background of *mpg1*, is an early-flowering rice, that has reduced transcript of flowering repressor *Ghd7*, and increased transcript of downstream genes *Ehd1, Hd3a,* and *RFT1*, responsible for flowering progression (Kim et al., 2013). Perception of florigens induces several transcription factors transitioning the shoot apical meristem (SAM) to a reproductive state generating panicle and later inflorescence tissues. One or more of the genes involved in this transition or generation of flowering tissues may be differentially regulated in *mpg1*.

mpg1 performs better under sub-optimal conditions (potentially salt and pH). Additionally, *mpg1* plants were able to accumulate greater levels of biomass under prolonged exposure to salt or drought stress. This suggests that *mpg1* plants might have differentially regulated genes that function in stress response, particularly abiotic stress. There are an abundance of genes that play roles in stress response in plants. Reactive oxygen species and antioxidants (Gill and Tuteja, 2010), chaperones (Wang et al., 2004b), accumulation of compatible solutes (Chen and Murata, 2004, Garg et al., 2002), and hormones (Peleg and Blumwald, 2011) play a crucial role in abiotic stress response and tolerance. Many response mechanisms for stress are driven by transcription factors. Several transcription factor gene families that have been shown to respond to abiotic stress include: AP2/ERF (Mizoi et al., 2012), NAC (Nakashima et al., 2012), WRKY (Chen et al., 2010), MYB (Agarwal and Jha, 2010), bHLH (Pireyre and Burow, 2015), and bZIP (Liao et al., 2008).

Genetic mechanisms associated with biomass accumulation are rather complex. Numerous factors could be contributing to the overall increase in size in *mpg1*. Energy and

nutrients are required to perform metabolism necessary for growth and development. Photosynthesis is relatively inefficient, when it comes to converting available radiation energy into new plant growth (Zhu et al., 2008). Alterations in photosynthesis could potentially influence plant productivity and biomass by fixing greater amounts of carbon and utilizing it downstream for growth and biomass accumulation. Alterations of photosynthetic machinery for thermostability in *Arabidopsis* provided enhanced growth under heat stress (Kurek et al., 2007). The use of glycolate catabolic pathway genes from *E. coli* in *Arabidopsis* resulted in enhanced photosynthesis and reduced photorespiration resulting in enhanced biomass production (Kebeish et al., 2007). Altered synchronization of circadian light cycles with photosynthesis result in enhanced photosynthesis and growth in *Arabidopsis* (Dodd et al., 2005). Constitutive overexpression of the transcription factor *HYR* in rice led to enhanced photosynthesis and primary carbon metabolism, resulting in enhanced biomass and yield (Ambavaram et al., 2014). Adjustments in photosynthetic activity and resulting biomass production seems to need equal adjustments in factors influencing metabolic balance. Metabolism often requires input between nutrients and other molecular factors that function in physiological cross-talk (Kellermeier et al., 2014). Due to this, alterations in nutrient assimilation and use efficiency might also influence biomass accumulation (Xu et al., 2012, Lopez-Arredondo et al., 2014, Fageria et al., 2008). With this in mind, amino acid biosynthesis (Zhou et al., 2009, Kishor et al., 2005, Carrari et al., 2005), carbon metabolism (Coleman et al., 2009, Sturm and Tang, 1999, Huber and Huber, 1996, Fredeen, 1988), and transport (Julius et al., 2017, Ortiz-Lopez et al., 2000) could all impact plant growth as well.

If pools of resources necessary to support growth and development are present, additional factors can influence growth. Hormones are key growth regulators (Santner et al., 2009). Auxin

is responsible for cell enlargement and plant growth (Teale et al., 2006), cytokinin has been shown to be involved in cell division and shoot initiation (Werner et al., 2001), gibberellin functions in stem growth (Richards et al., 2001), and brassinosteroids play a role in cell elongation and division (Clouse, 1996). Hormones influence different downstream genes that directly impact growth. One particular downstream mechanism influencing growth are meristematic tissues (Somssich et al., 2016, Perilli et al., 2012). Differential expression and altered maintenance of meristematic tissues can result in differential plant growth and architecture (Reinhardt and Kuhlemeier, 2002, Liu et al., 2009, Clark et al., 1997, Mizukami and Fischer, 2000). The interaction of hormones on meristematic tissue often influences genes pertaining to cell cycle and differentiation (Dello Ioio et al., 2007). The cell cycle functions to generate new cells during growth and development through mitosis, therefore these genes directly influence cell number and cell size (A et al., 2017, Sablowski and Carnier Dornelas, 2014). Alterations in these genes can directly result in differential plant sizes and morphology (Qi and John, 2007, Wang et al., 2000, Dewitte et al., 2003). Cell cycle genes also regulate endoreduplication, which can significantly increase cell sizes (Sugimoto-Shirasu and Roberts, 2003, Park et al., 2005). *mpg1* possibly experiences altered hormone and/or cell cycle regulation generating plants that are larger due to differences in cell size or number compared to wild-type.

Considering the numerous ways that growth can be influenced in plants, investigation of distinct characteristics present in *mpg1* could potentially help narrow our focus. Specific growth metrics observed in *mpg1* contribute to its increase in size. Specifically plant height, girth, and wider leaves resulted in larger plants compared to wild-type. Additionally, *mpg1* plants were able to accumulate greater biomass post-ratooning.

Plant height in rice has been found to be mostly controlled by the synthesis and regulation of phytohormones GA and BR (Ashikari et al., 1999, Yamamuro et al., 2000, Hong et al., 2003, Sasaki et al., 2002, Sasaki et al., 2003, Itoh et al., 2004, Tanabe et al., 2005). Genes influencing either of these hormones are likely differentially regulated in *mpg1*.

Culm size in rice has been found to corollate with numerous physiological properties within rice. Larger culm varieties have greater rates of photosynthesis during heading and grain filling, apoplasmic transport ability and gas exchange, and higher yield production through greater number of grains per panicle (Wu et al., 2011). Many of these physiological capabilities are expected to be a result of increased vascular bundle area compared to common small culm plants (Wu et al., 2011). Large culm plants also exhibited a higher potential for lodging resistance as they possessed fewer tillers; they were taller, had thicker and wider culms, and longer leaves - which also directly resulted in greater biomass (Wu et al., 2011). Several genes have been described pertaining to culm thickness. Cinnamyl-alcohol dehydrogenase (CAD) mutants, which play a role in cell wall synthesis, experience differential culm thickness (Ookawa et al., 2014, Li et al., 2009). Other genes involved in cell wall synthesis have also been found to play a role in culm size as well (Hirano et al., 2010, Aohara et al., 2009). Other genes, including *PLANT ARCHITECTURE AND YIELD 1,* (*PAY1*) (*Os08g31470*), encoding a peptidase, trypsin-like serine and cystine protease, affects plant architecture in the form of greater culm diameter, height, and grains per panicle, possibly through reduced polar auxin transport activity and increased cell size in rice (Zhao et al., 2015). Taken together, this again suggests that greater culm sizes can result from genes functioning in cell wall biosynthesis and/or cell number/size regulation.

Increased leaf width was also observed in *mpg1*. Several genes controlling leaf architecture in rice have been found. *NARROW LEAF 1* (*NAL1*) (*Os04g52479*), encodes a putative trypsin-like serine/cystine protease which regulates leaf growth and shape, vascular orientation, root architecture, and polar auxin transport (Qi et al., 2008). *NARROW LEAF 2* (*NAL2*) (*Os11g01130*) and *NARROW LEAF 3* (*NAL3*) (*Os12g01120*) are two duplicated WUSCHEL-related homeobox genes (*WOX3A*) that are involved in leaf margin, and vascular development potentially through auxin distribution by PIN proteins influencing leaf width, grain shape, tiller, and lateral root abundance (Ishiwata et al., 2013, Cho et al., 2013). *NARROW LEAF 7* (*NAL7*) (*Os03g06654*), encodes for a flavin-containing monooxygenase, that regulates leaf width mediated by auxin biosynthesis (Fujino et al., 2008). *NARROW AND ROLED LEAF 1* (*NRL1*) (*Os12g36890*), also known as *OsCslD4*, encodes a cellulose synthase-like protein D4, which regulates leaf morphogenesis and vegetative development (Hu et al., 2010). Meristem maintenance regulates plant architecture including leaf morphology. It is likely that genes involved in meristematic function and/or auxin transport are differentially regulated in *mpg1*.

mpg1 accumulates greater seed yield by generating panicles with increased grain number. Greater number of grain per panicle usually arises from increase panicle branching and/or panicle length. Several genes influencing panicle morphology and increased grain yield in rice have been uncovered. *IDEAL PLANT ARCHITECTURE 1, IPA1/OsSPL14* (*Os08g39890*), encodes a squamosa promoter binding protein-like 14, protein which has been shown to stimulate panicle branching leading to higher grain yield through unknown mechanisms (Miura et al., 2010). A dominant gain-of-function allele of *DEP1* (*Os09g26999*), which encodes a phosphatidylethanolamine-binding protein (PEBP), has been shown to increase the number of grains per panicle and overall grain yield through enhanced meristematic activity (Huang et al.,

2009b). Genes relevant to hormone regulation alter panicle morphology. A semi-dominant allele of DST (*Os03g57240*), a rice zinc finger protein that perturbs its interaction with its target *Gn1a/OsCKX2* and in doing so regulates cytokinin oxidase/dehydrogenase activity resulting in increased CK activity in the shoot apical meristem, generates an increasing number of reproductive organs and panicle branches (Li et al., 2013b, Ashikari et al., 2005). *OsLAC* (*Os05g38420*), encodes a laccase-like protein involving sensitivity of plants to BR, which when repressed via miR397 produces long panicles with greater branches, and when overexpressed results in small grains and panicles (Zhang et al., 2013). Additional factors, such as pleiotropic characteristics surrounding heading date, have correlated with panicle morphology and grain number. *HEADING DATE 7* (*Ghd7*) (*Os07g49460*), a flowering repressor encoding for a CCT domain protein, when expressed produces greater seed yield through panicle branching, and increased plant height (Xue et al., 2008, Liu et al., 2013b). *HEADING DATE* (*Ghd8/LHD1*) (*Os08g07740*), encodes a HAP3 subunit of a CCAAT binding protein, when expressed correlates with heading date, seeds per panicle, and plant height as well (Zhang et al., 2006, Yan et al., 2011, Dai et al., 2012). Potentially, *mpg1* experiences an increase in seeds per panicle through differential regulation of genes pertaining to hormonal regulation, meristematic activity, and/or complex pleiotropic traits involved in flowering control.

Besides having an increased number of spikelets, *mpg1*'s seeds developed awns, where wild-type Kitaake do not. Awns are long needle-like projections forming from the top of the lemma in the floret. Domesticated variants within *japonica* have had awn characteristics bred away, as awnless seeds are more favorable during harvest and storage. Several genes and underlying mechanisms have been identified for the development of awns in rice. Single mutations in several genes in *japonica* have given rise to awn formation including:

SHOOTLESS2 (*SHL2*) (*Os01g34350*) , *SHOOT ORGANIZATION1* (*SHO1*) (*Os04g43050*), *SHOOT ORGANIZATION2* (*SHO2*) (*Os03g33650*), *WAVY LEAF1* (*WAF1*) (*Os07g06970*), and *TONGARI-BOUSHI1* (*TOB1*) (*Os04g45330*) (Itoh et al., 2000, Itoh et al., 2008, Liu et al., 2007a, Toriba et al., 2010, Song et al., 2012, Tanaka et al., 2012). These genes are associated with meristem maintenance and spatial orientation establishment. These mutants, capable of awn development, have partial defects in maintenance and adaxial-abaxial polarity. Several QTLs have been associated with awn development in rice, two of which are *An-1* and *An-2*. *An-1* was confirmed as a bHLH transcription factor (*Os04g28280*) that has pleiotropic effects on awn initiation and length, grain length, and grain number per panicle by promoting cell division, and down-regulating meristematic activity (Luo et al., 2013). *An-2*, also known as *LONELY GUY LIKE PROTEIN 6* (*LOGL6*) (*Os04g43840*), encodes a protein that catalyzes the final step of cytokinin synthesis in ancestral wild rice *Oryza rufipogon*, that influences awn elongation by enhanced cell division and decreased grain production (Gu et al., 2015). *GRAIN NUMBER, GRAIN LENGTH AND AWN DEVELOPMENT1* (*GAD1*) (*Os08g37890*), predicted to encode a small secretory signal peptide from the *EPIDERMAL PATTERNING FACTOR-LIKE* (*EPFL*) family, has been shown to regulate awn development, grain number, and grain length potentially through cytokinin influence and synergistic interaction with *An-1* and *An-2* (Jin et al., 2016). The genes *DROOPING LEAF* (*DL*) (*Os03g11600*), a YABBY gene required for carpel specification and midrib formation (Yamaguchi et al., 2004), and *OsETTIN2* (*OsETT2*) (*Os01g48060*), a gene orthologous to *Arabidopsis ETTIN* (*ETT*)/ *AUXIN RESPONSIVE FACTOR3* (*ARF3*) responsible for abaxial cell fate (Toriba et al., 2010), together are sufficient for the formation of awns in *indica* varieties and relay dormant in *japonica* (Toriba and Hirano, 2014). These genes have suggested that awn primordium functions with meristematic tissues. The molecular mechanisms

determining awn formation remain complex and elusive, however it has been hypothesized that particular tissues might function as a 'quasi-meristem' in that they give rise to specialized and determinate tissues through active cell proliferation (Girin et al., 2009), and that awn primordium might function in this way (Toriba and Hirano, 2014). Because *mpg1* is associated with ectopic formation of awns, and some similar pleiotropic characteristics as the genes mentioned above, it is likely that genes and pathways that associate with meristematic tissue are being differentially regulated.

Ratooning is the agricultural practice of cutting tiller/culm tissue at harvest to stimulate the regeneration of panicle bearing culms in order to render a secondary harvest. This type of cultivation is not broadly practiced, but shows some promise for increasing yields today (Faruq et al., 2014). Recent use of ratooning has been implemented in China to increase rice yields, suggesting its potential economic importance (Hart, 2018). Mutant *mpg1* is able to regenerate more tissue at a quicker rate with greater yields after ratooning compared to wild-type. Genes involved in ratooning, however, have not been well characterized. Studies investigating ratooning efficiency indicate that successful ratooning seems to rely on appropriate agricultural practices, by means of tiller cutting time and height, as well as application of sufficient nitrogen and subsequent efficiency for its translocation to induce successful post-cutting regrowth (IRRI, 1988, Jones, 1993, Oad et al., 2002, Daliri et al., 2009, Kailou, 2012). *mpg1* might have differentially expressed genes pertaining to nutrient use efficiency or regulation of axillary bud activity, or even general growth factors.

When assessing the function of *MPG1* in *mpg1*, it is important to note that its expression is modified from its native pattern. Recent studies have begun to characterize *MPG1*'s native function and expression patterns. MSU Rice Genome Annotation Database's RNA-seq coverage

data show expression in numerous tissue types specifically within pre- and post- emergence inflorescence, seed 5 days post-pollination, embryo 25 days post-pollination, and 20 days post-planting shoot and leaves (Ouyang et al., 2007). Of these tissues, greater levels of expression were measured in flowering and reproductive tissues. *MPG1* has also been shown to play a role in endosperm development and grain filling (Xu et al., 2016a). It acts in association with two nuclear factor transcription factors (*OsNF-YB1* and *OsNF-YC*) to form a transcriptional complex within the aleurone layer of developing seed tissue, where it acts as the DNA binding protein (Xu et al., 2016a). Genes in this transcriptional complex, through transcriptome analysis of *OsNF-YB1* RNAi lines, correlate strongly with genes responsible for transport, ATP synthesis, protein folding, response to stimuli, and metabolic processes (Xu et al., 2016a). ChIP-seq and expression analyses found genes functioning in sugar and amino acid transport to be targets of *OsNF-YB1* (Bai et al., 2016, Xu et al., 2016a). Additionally, this nuclear factor transcription factor is suspected to regulate cell cycle genes, further influencing the development and proliferation of endosperm tissue (Sun et al., 2014b).

Expression of *MPG1* has also been observed in response to stress. Studies evaluating several stressors noted an increased expression via micro array and RT-PCR analyses. In particular, exogeneous salicylic acid, submergence, laid-down submergence, and cold revealed the activation of *MPG1* (Sharoni et al., 2011).

The ectopic expression of *MPG1* is likely modulating multiple pathways involved with the phenotypes observed in *mpg1*. The nature of why the T-DNA insertion resulted in elevated levels of expression of *MPG1* has yet to be elucidated. Expression of *MPG1* in *mpg1* was found in shoot, leaf, and reproductive tissues over the course of development, however little to no report was found in root tissues during vegetative growth. Perhaps the expression of *MPG1* in

mpg1 is spatiotemporally driven. Additional factors contributing to the nature of *MPG1* expression in the mutant might have resulted from disrupting local regulatory mechanisms not yet characterized. There was a 20 base-pair deletion of genomic DNA at the site of the T-DNA integration. The region surrounding its incorporation was also highly methylated, possibly allowing for enhanced physical interaction to native or ectopic transcriptional regulators (present in the T-DNA). In addition to the transcriptional regulatory elements in the T-DNA, local elements proximal to the site of the T-DNA integration could also be driving, directing, or influencing expression of *MPG1* in a complex fashion. Indeed, the absence of *MPG1* expression in the root of the mutant containing the constitutive 35S promoter in the T-DNA expression cassette shows additional regulatory elements are impacting *MPG1* expression. The temporal/developmental stage and tissue type where this up-regulation of *MPG1* is present could be playing a role in generating individual characteristics observed in the mutant. Regardless, further evaluation of how *MPG1* is expressed will be crucial in determining the systematic function associated with the phenotype visualized in *mpg1*.

Because *MPG1* is operating outside of its normal pattern of expression in *mpg1*, understanding the impact of this ectopic expression is of crucial importance. The altered expression of *MPG1* likely caused numerous other genes to be differentially regulated, any of which potentially play role in the phenotype observed in the mutant. Broadly assessing global changes in gene expression in the mutant compared to wild-type could help reveal mechanisms that are differentially regulated and important to plant productivity and allow for the generation of further hypotheses surrounding the formation of the specific characteristics present in *mpg1*.

A global evaluation of gene expression in *mpg1* was performed utilizing RNA-sequencing to assess transcriptomic differences between the mutant and wild-type. Out of all the

physical characteristics that differ between *mpg1* and wild-type, plant girth was the most drastic, so we opted to sample tiller tissue at the points around where the phenotype became notably different. Two separate rounds of RNA-seq were performed to try and best capture which genes were differentially expressed just prior to and during our ability to measure the tiller girth phenotype in *mpg1*.

In one round of RNA-seq we chose to sample whole tiller tissue prior to any physical evidence of the girth phenotype. At 42 days post-planting *mpg1* plants begin to noticeably outpace wild-type plants in term of the size of tillers by girth. Therefore we selected to profile tissue 10 days prior to this observation in hopes of capturing differences in transcript abundance before the generation of the phenotype. We created libraries from individual plants (4 different HM-*mpg1* plants, and 4 different wild-type null segregants (WT-ns)). The libraries were run on the Illumina HiSeq2500 instrument utilizing a 125bp paired-end reads format. Differential gene expression was determined and the genes were further assessed by ontological function and presence of the known DNA binding motif of *MPG1* to more clearly partition these gene into possible genes of interest pertaining to characteristics present in the *mpg1* phenotype. This analysis resulted in a number of significantly differentially regulated genes.

During a second round of RNA-seq we chose to use whole tiller tissue at 42 days post-planting, as it was at this timepoint where the *mpg1* plants exceeded wild-type plants in term of the size of tillers by girth. By selecting this timepoint we were hoping to capture transcriptional differences occurring when the phenotype of the mutant was present. We generated libraries from individual plants (3 individual homozygous (HM-*mpg1*) plants, and 3 individual wild-type null segregants (WT-ns)). The libraries were run on the Illumina NextSeq instrument utilizing a 75bp paired-end reads format. Again, differential gene expression was performed to find genes

that were significantly differentially expressed in HM-*mpg1* plants compared to wild-type null segregants.

Better understanding of the development of rice through dissection of crown tissue during different points of development revealed the temporal shift from vegetative phase of development to the reproductive phase of development by determining panicle initiation. The timepoint at which this occurs lines up well with when *mpg1* plants began to accumulate greater girth, so it likely that our profiling of the 42 days post-planting tissue captured *mpg1* and wild-type null segregants under different developmental stages, rendering the resulting data from 42 days post-planting particularly complex because of the developmental shift. This being said however, when assessing the analysis of both RNA-seq experiments collectively there are several genes that overlap between the two experiments. These genes and associated predicted ontological function have allowed us to develop further hypotheses that might help further our understanding of the phenotype observed in *mpg1*. There were a myriad of differentially expressed genes in *mpg1*. Differentially regulated genes were enriched for ontological functions associated with transcription factor activity, flower development, stress response, DNA metabolism, and the cell cycle. Other genes of interest were found in high correlation with defoliation response, cell wall formation, and hormone regulation.

Materials and Methods

Plant materials

Rice (*Oryza sativa* L. spp. *japonica* cv. Kitaake), including wild-type, and segregating lines of partial T-DNA insertion of expression cassette *CmGAS1pro*::*AtSUC1*$_{H65K}$ - mutant (*mpg1*), were used to assess expression analyses.

Growth conditions

Seeds were placed on germination paper and partially submerged in a 1:1000 dilution of MAXIM XL dual action fungicide (Syngenta) and sealed with parafilm. Seeds were incubated at 30º C under 12 h light cycles, until primary shoot and root development occurred (usually 5-7 days). Seedlings were thin transferred to planting medium in the greenhouse or growthchamber. Pots were organized in random fashion in a flat or tub with water covered with black plastic and watered until media was fully saturated and pots remained in roughly 3" of standing water. Plant chlorosis was monitored and preemptively treated around the 3- to 4-leaf stage using Sprint 330 Iron Chelate at 0.3 g/L water and top-watered. At the same developmental stage plants were fertilized using granulized (Technigro) 15-5-15 Plus Cal-Mag at 48.87 g/L and top-watered. Fertilizer treatment occurred twice weekly through maturity until harvest.

For plants assessed at the 32 day post-planting timepoint, seedlings were then transferred to planting medium in the greenhouse. The contents were mixed to homogeneity, and transferred to pots. Plants were grown in 1.0 gallon size pots in greenhouse conditions maintained at roughly 26˚C and 75% RH with a 16h light cycle. For plants assessed at the 42 day post-planting timepoint, seedlings were grown in 0.75 gallon size pots in growthchamber conditions maintained at roughly 26˚C and 80% RH with a 13h light cycle.

DNA extraction

Young, fresh plant tissue (3-leaf stage) was sampled for DNA extraction and analysis (2-5 cm of leaf-tip). DNA was obtained via mechanical disruption of tissue and Shorty-Buffer extraction. Tissue was flash frozen in liquid nitrogen and disrupted using the Qiagen TissueLyser at 30 rps for a period of 1 minute. Five hundred μL of freshly prepared shorty buffer (0.2 M Tris HCl pH 9.0, 0.4 M LiCl, 25 mM EDTA, and 1.0% SDS) was added to each tissue sample, vortexed and centrifuged at max speed (13k rpm) for 5 min. Then 350 μL of supernatant was

transferred to a new tube containing 400 μL isopropanol, mixed by inverting and centrifuged at max speed for 10 min. The supernatant was discarded and 1mL of 70% ethanol was added to each sample to wash the DNA pellet. Samples were then centrifuged at max speed for 10 minutes, the supernatant was discarded and the tubes were inverted for 30 minutes. The DNA pellet was resuspended in 100 μL of 10mM Tris-HCl pH 8 and stored short term at 4º C until use.

Haplotyping mutants

To identify *mpg1* plants that were homozygous, heterozygous, or null segregants for the bi-laterally truncated T-DNA insertion, primers were designed in regions directly flanking the site of insertion, as well as primers spanning the integration site into the T-DNA, and used in PCR. The primers flanking the T-DNA insertion were wFLA forward: 5'-GGAAGTTGGAGATGGGAAACA-3', and wFLA reverse: 5'-GGCCTCGTGTGTCAGTAATAA-3'. The primers spanning the genomic region and the T-DNA insertion were wIN forward: 5'-ACACCGGAAGCATAGTCATTT-3', and wIN reverse: 5'-GGTCGCCAACATCTTCTTCT-3'.

RNA-seq

***RNA-extraction and Library Preparation*:** Two RNA-seq experiments were performed, one on 32 days post-planting material (4 individual plants each from two genotypes grown in greenhouse conditions), and another on 42 days post-planting material (3 individual plants each from two genotypes grown in growth chamber conditions). Whole tiller tissue was selected for analysis, from roughly 5 cm above soil line and up (collected mid-day) of two genotypes (HM-*mpg1* and WT-ns. Tissue was placed in individual 50 mL conical tubes and flash frozen in liquid nitrogen. Tissue was ground using the TissueLyser (Qiagen) at 30 rps for 1 min, and RNA was

extracted using (Qiagen) Plant RNeasy mini-kit. Desired tissue (not more than 100 mg) was sampled for RNA extraction and analysis. RNA was treated with DNase and purified using the Turbo DNase kit (Invitrogen). RNA quality control was verified using a Bioanalyzer (Agilent) and TapeStation (Agilent) for the 32 and 42 days post-planting respectively. Libraries were generated using the TruSeq RNA-seq kit (Illumina) as per manufacture instructions, 125bp paired-end sequencing of the library for 32 days post-planting material was done at the RTSF GENOMICS Core facility at Michigan State University using Illumina Hi seq 2500 system. 75bp paired-end sequencing of the library for 42 days post-planting material was done at the Self-Service Next Generation Sequencing Core facility at Colorado State University using Illumina NextSeq system.

***Mapping of Reads and Identification of the DEG*:** Resulting reads were assessed for quality control using (FastQC) (Andrews, 2010), where results fell within acceptable parameters. This was followed with FASTQ Toolkit (Illumina-BaseSpace-Labs, 2018) performing quality trimming of anything less than 20 phred score. Due to the inherent nature of RNA-seq pipelines, several methods were originally evaluated (APPENDIX). The analysis pipeline selected utilized STAR (2.5.3a) (Dobin et al., 2013), Htseq-count-merge (0.6.1) (Anders et al., 2015), and edgeR (3.0) (Robinson et al., 2010) through the Cyverse workflow interface (Merchant et al., 2016). Reads were aligned to the Ensembl MSU6.0 version of the rice genome available through (http://rice.plantbiology.msu.edu). Total mapped read count ranged between 40-70, and 50-220 million reads per sample for 32 and 42 days post-planting RNA-seq experiments respectively.

Hierarchical clustering and heatmaps were created using the R package heatmap.2 with hclust and dendrogram formation functions. Gene co-expression was evaluated using Genevestigatior software (Hruz et al., 2008). Venn diagrams were generated using BioVenn

(Hulsen et al., 2008).

Promoter analysis

Assessment of the number of GCC-box (GCCGCC) motif elements present in promoter regions was conducted using R and searching the Ensemble MSU6.0 genome build and extracting number of motif elements present in the range of -2000 bp from the transcription start site (TSS) for every gene. These numbers were later cross referenced against our generated list of DEG from RNA-seq.

GO enrichment analysis

GO analysis was performed for term enrichment using RiceNetDB (Chen, M. 2013). Single enrichment analysis with GO annotations was performed with a corrected P-value <= 0.01. The genes that are up- or down-regulated for each data set were analyzed separately. To identify various genes that correlate with the pleiotropic phenotype of *mpg1*, several GO terms and their correlating genes were placed in separate tables for partitioning and assessment.

Results

Mutant plant *mpg1* experiences differential expression of 297 genes in 32 days post-planting stem tissue

Differential gene expression was analyzed using the STAR-HTSeq-edgeR alignment, mapping, and assessment pipeline (APPENDIX). Libraries were constructed from individual plant's whole stem tissue, four HM-*mpg1* and four WT-ns plants. Results indicate that there are 297 differentially expressed genes (Figure 4.1), of which, 106 were up-regulated, and 191 were down regulated in HM-*mpg1* compared to WT-ns. Because MPG1 targets the GCC-box motif directly, we assessed the differentially expressed genes promoter regions (-2000 from transcription start site (TSS)) for the presence of this motif. Of the differentially expressed genes,

56 up-regulated genes, and 107 down-regulated genes contained at least one GCC-box in their respective promoter regions (-2000 from TSS) (Figure 4.1, C). Although, we cannot claim that these genes are being targeted by MPG1 without independent verifying them, they are good to note for future investigation.

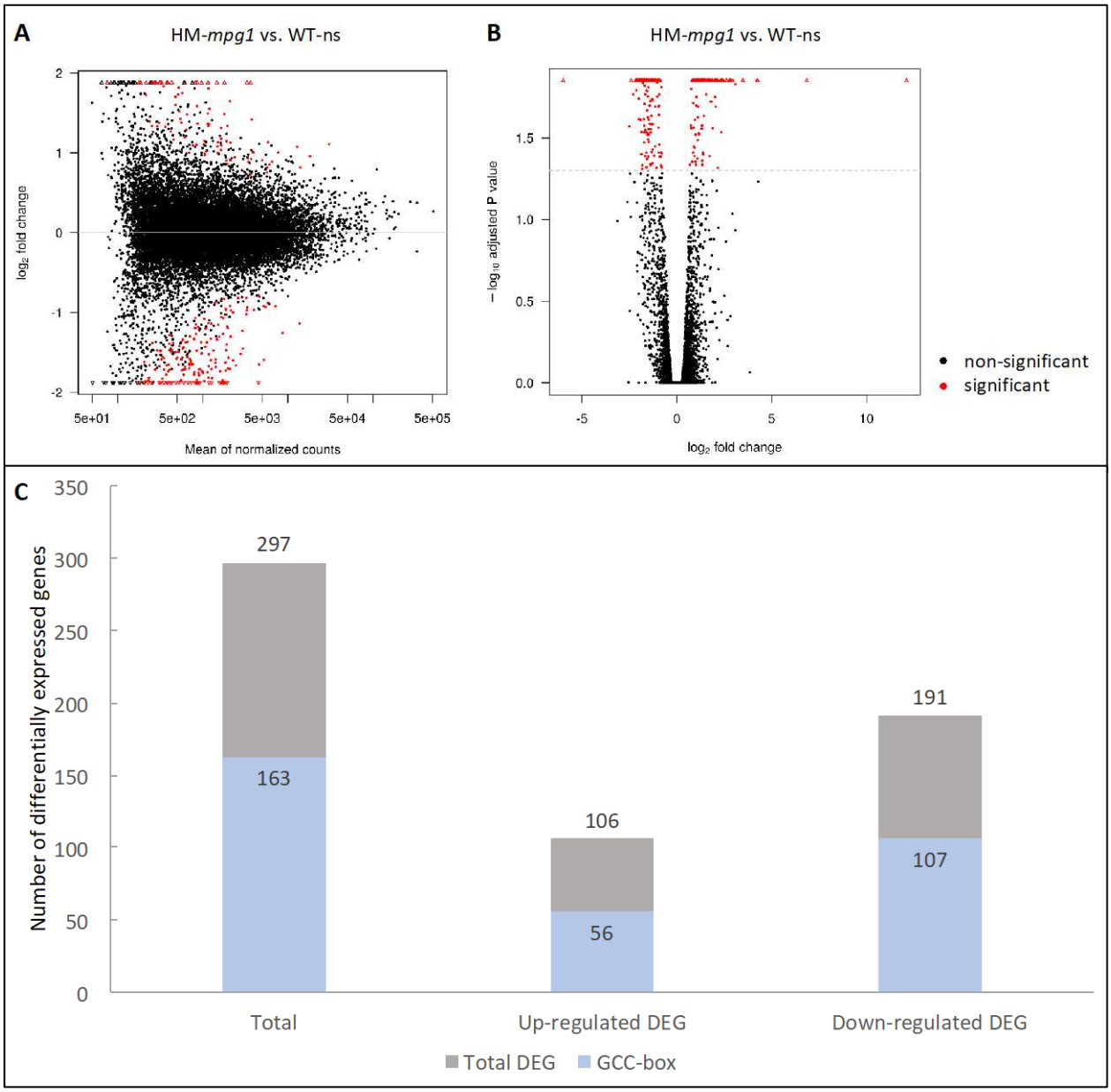


Figure 4.1: Differentially expressed genes found from 32 days post-planting stem tissue. (A) MA-plot, (B) volcano plot, and (C) bar graph with number of genes up- and down-regulated. The bar graph also contains the number of genes differentially expressed that contain a GCC-box motif in their promoter region (-2000 from TSS).

MPG1 is the only gene with noticeably different expression within the region of the T-DNA insertion in 32 days post-planting in stem tissue

Using the Integrated Genome Viewer (IGV) software, reads were mapped from each of the samples (individual plants) to the genome (MSU6.0 build) to visually assess the region of the T-DNA integration. Visualization of seven surrounding genes show that *MPG1* is up-regulated in HM-*mpg1* compared to WT-ns which have little to no expression (Figure 4.2). Additionally, the other genes in the region surrounding the T-DNA insertion appear to have little to no noticeable difference in expression between HM-*mpg1* and WT-ns.

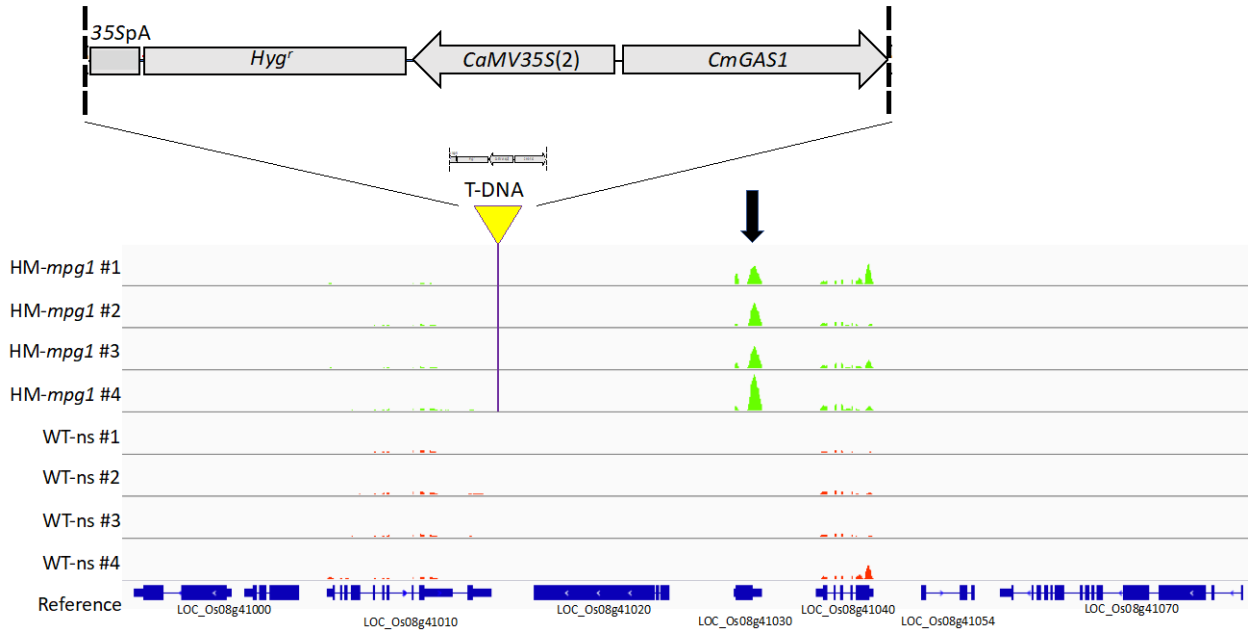


Figure 4.2: Gene expression surrounding the site of the T-DNA insertion.
IGV visualization of read abundance of genes proximal to the T-DNA insertion in tissue collected from 32 days post-planting whole tiller tissue. The schematic shows the orientation, location, and size of the insertion (purple bar) in *mpg1* plants. Each row represents a single biological replicate. *MPG1* (*Os08g41030*) is notated by the black arrow. Scale: 0-4129 read abundance.

mpg1 plants have similar expression patterning between samples

Using normalized counts, a hierarchical clustering of genes was plotted as a heatmap

using each sample. Plants within each sample group (HM-*mpg1* and WT-ns) show a degree of variability amongst one and other, which is easily expected from individual to individual, however show a decent correlation of gene expression within each genotype (Figure 4.3). Segments of differential gene expression clearly differ between HM-*mpg1* and WT-ns suggesting similar and consistent behavior of gene expression within each genotype.

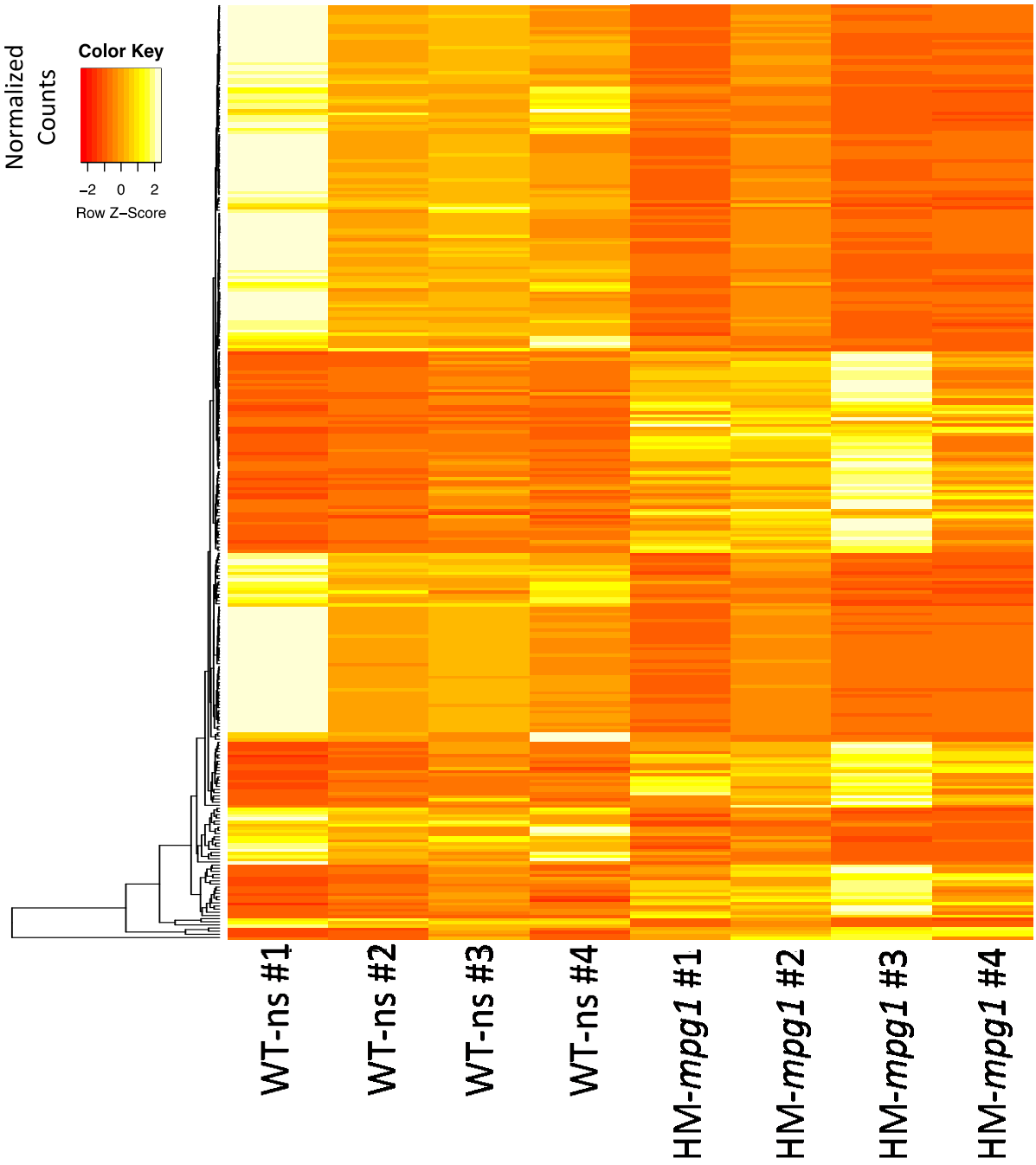


Figure 4.3: Heatmap and hierarchical clustering analysis.
Each lane represents an individual biological replicate from 32 days post-planting tissue by normalized counts of DEG.

Surprisingly, clusters that differed between *HM-mpg1* and WT-ns didn't reveal any particular type or category of genes sufficient for interpretation or analysis regarding the phenotype. Many of the genes categorized together pertained to several and/or different

functions. Further assessment via expression levels and ontology enrichment are necessary to help understand the types of differentially regulated genes in *mpg1*.

MPG1 is the most significantly differentially regulated gene in 32 days post-planting tiller tissue

Assessment of the genes determined to be significantly differentially expressed between HM-*mpg1* and WT-ns provided further insight into the behavior of the transcriptome and individual genes. *MPG1* (*Os08g41030*) was the most significantly differentially expressed gene experiencing a 12.1 log2 fold-change between HM-*mpg1* and WT-ns (Table 4.1). This further suggests its importance in influencing the phenotype observed in *mpg1*.

Table 4.1: DEG list from RNA-seq analysis of HM-*mpg1* compared to WT-ns within 32 days post-planting tiller tissue.

MSU ID	RAP ID	log2 Fold Change	Corrected p-value	GCC-box	Description
Up-regulated DEG					
LOC_Os08g41030	Os08g0521600	12.102	3.1769E-167	2	AP2 domain containing protein, expressed
LOC_Os07g40380	Os07g0594400	6.841	2.52939E-07	2	expressed protein
LOC_Os11g05380	Os11g0151400	4.254	0.012423981	0	cytochrome P450, putative, expressed
LOC_Os06g29260	None	4.228	3.76167E-16	0	retrotransposon protein, putative, unclassified, expressed
LOC_Os01g41420	Os01g0597600	3.491	0.006146145	1	transmembrane amino acid transporter protein, putative, expressed
LOC_Os03g08310	Os03g0180800	3.476	0.000727219	0	ZIM domain containing protein, putative, expressed
LOC_Os12g08850	Os12g0190400	3.083	0.014783458	0	expressed protein
LOC_Os12g14440	Os12g0247700	2.935	0.000727219	3	Jacalin-like lectin domain containing protein, putative, expressed
LOC_Os01g60600	Os01g0821300	2.828	9.56729E-05	1	WRKY108, expressed
LOC_Os04g23550	Os04g0301500	2.828	0.000165919	0	basic helix-loop-helix family protein, putative, expressed
LOC_Os03g53020	Os03g0741100	2.774	0.002370743	1	helix-loop-helix DNA-binding domain containing protein, expressed
LOC_Os02g11070	Os02g0205500	2.632	6.9291E-05	0	3-ketoacyl-CoA synthase, putative, expressed
LOC_Os08g28710	Os08g0374600	2.629	0.001720238	2	receptor protein kinase CRINKLY4 precursor, putative, expressed
LOC_Os03g05880	Os03g0153500	2.622	7.66277E-07	0	monooxygenase, putative, expressed
LOC_Os03g08520	Os03g0183500	2.436	0.001582899	0	DUF581 domain containing protein, expressed
LOC_Os01g64360	Os01g0863300	2.409	0.002423154	6	MYB family transcription factor, putative, expressed
LOC_Os10g25290	Os10g0392400	2.357	0.003615628	1	ZIM domain containing protein, putative, expressed
LOC_Os04g27340	Os04g0341500	2.353	4.97594E-05	0	terpene synthase, putative, expressed

LOC_Os01g18120	Os01g0283700	2.35	0.000126019	1	cinnamoyl CoA reductase, putative, expressed
LOC_Os06g03670	Os06g0127100	2.341	0.029199294	5	dehydration-responsive element-binding protein, putative, expressed
LOC_Os02g08440	Os02g0181300	2.31	0.000314578	0	WRKY71, expressed
LOC_Os02g52210	Os02g0759400	2.305	0.002206617	0	zinc finger, C3HC4 type domain containing protein, expressed
LOC_Os01g09220	Os01g0186900	2.177	0.048166977	0	transposon protein, putative, CACTA, En/Spm sub-class, expressed
LOC_Os05g44060	Os05g0516700	2.151	0.006885263	1	expressed protein
LOC_Os02g11859	Os02g0209300	2.113	0.000707832	1	expressed protein
LOC_Os05g12040	Os05g0211100	2.092	0.000165919	1	cytochrome P450 51, putative, expressed
LOC_Os02g08270	Os02g0179200	2.074	0.001252941	1	class I glutamine amidotransferase, putative, expressed
LOC_Os02g45420	Os02g0676800	2.051	0.020342379	3	AP2 domain containing protein, expressed
LOC_Os11g02520	Os11g0117400	2.042	2.21885E-06	0	WRKY104, expressed
LOC_Os03g08330	Os03g0181100	2.022	0.017725838	0	ZIM domain containing protein, putative, expressed
LOC_Os05g46830	Os05g0546300	1.993	0.000884752	0	proline-rich protein, putative, expressed
LOC_Os04g41960	Os04g0497000	1.895	0.027454263	8	NADP-dependent oxidoreductase, putative, expressed
LOC_Os08g07100	Os08g0168000	1.879	0.000223937	0	terpene synthase, putative, expressed
LOC_Os10g02840	Os10g0118000	1.833	0.000165919	0	O-methyltransferase, putative, expressed
LOC_Os03g08320	Os03g0180900	1.818	0.007078376	1	ZIM domain containing protein, putative, expressed
LOC_Os10g30790	Os10g0444700	1.808	0.018925408	0	inorganic phosphate transporter, putative, expressed
LOC_Os02g32580	Os02g0527200	1.75	0.007814787	6	expressed protein
LOC_Os08g34580	Os08g0445700	1.733	0.010741315	0	trehalose-6-phosphate synthase, putative, expressed
LOC_Os07g46920	None	1.727	0.040742897	0	sex determination protein tasselseed-2, putative, expressed
LOC_Os04g33640	Os04g0412300	1.705	0.009452213	4	glycosyl hydrolases family 17, putative, expressed
LOC_Os07g34260	Os07g0526400	1.648	1.10124E-06	0	chalcone and stilbene synthases, putative, expressed
LOC_Os10g37760	Os10g0521900	1.6	0.005745346	0	OsRhmbd17 - Putative Rhomboid homologue, expressed
LOC_Os03g17700	Os03g0285800	1.587	0.003068219	1	CGMC_MAPKCGMC_2_ERK.2 - CGMC includes CDA, MAPK, GSK3, and CLKC kinases, expressed
LOC_Os07g09190	Os07g0190000	1.554	0.017089824	0	transketolase, putative, expressed
LOC_Os08g13440	Os08g0231400	1.523	0.000884752	4	cupin domain containing protein, expressed
LOC_Os11g45740	Os11g0684000	1.495	0.000241253	1	MYB family transcription factor, putative, expressed
LOC_Os05g27730	Os05g0343400	1.486	0.011176548	2	WRKY53, expressed
LOC_Os02g02930	Os02g0121700	1.482	0.008390101	0	terpene synthase, putative, expressed
LOC_Os08g10500	Os08g0205800	1.427	0.023171071	0	expressed protein
LOC_Os02g48770	Os02g0719600	1.417	2.12788E-05	4	SAM dependent carboxyl methyltransferase, putative, expressed
LOC_Os03g03700	Os03g0129100	1.396	0.039896743	1	MLO domain containing protein, putative, expressed
LOC_Os12g36110	Os12g0547600	1.38	0.005968712	0	calmodulin binding protein, putative, expressed
LOC_Os01g56810	Os01g0775400	1.379	0.038921032	3	cytokinin dehydrogenase precursor, putative, expressed

LOC_Os12g26290	Os12g0448900	1.367	0.001582899	0	alpha-DOX2, putative, expressed
LOC_Os06g10210	Os06g0203600	1.366	0.000528889	0	expressed protein
LOC_Os05g10310	Os05g0191500	1.36	0.013188333	1	acid phosphatase, putative, expressed
LOC_Os01g66860	Os01g0892800	1.338	9.56729E-05	0	serine/threonine protein kinase, putative, expressed
LOC_Os09g17560	Os09g0344500	1.337	0.011795236	0	O-methyltransferase, putative, expressed
LOC_Os08g10150	Os08g0201700	1.335	0.046043213	9	SHR5-receptor-like kinase, putative, expressed
LOC_Os03g19990	Os03g0314500	1.333	0.043428005	1	WD40-like Beta Propeller Repeat family protein, expressed
LOC_Os12g17430	None	1.325	0.026505655	1	NBS-LRR disease resistance protein, putative, expressed
LOC_Os11g13750	Os11g0241700	1.313	0.027683623	1	expressed protein
LOC_Os02g06930	Os02g0165100	1.289	0.011137219	2	protein kinase, putative, expressed
LOC_Os04g52440	Os04g0614500	1.269	0.012593484	10	aminotransferase, putative, expressed
LOC_Os04g05650	Os04g0142400	1.196	0.011452357	0	expressed protein
LOC_Os04g57200	Os04g0667600	1.177	0.029553586	3	heavy metal transport/detoxification protein, putative, expressed
LOC_Os06g20920	Os06g0314600	1.167	0.030040896	0	SAM dependent carboxyl methyltransferase, putative, expressed
LOC_Os04g33390	Os04g0406600	1.148	0.013188333	1	prephenate dehydratase domain containing protein, expressed
LOC_Os08g03350	Os08g0127100	1.144	0.002356747	1	amino acid transporter, putative, expressed
LOC_Os12g36910	Os12g0556200	1.134	0.001582899	5	calmodulin binding protein, putative, expressed
LOC_Os03g28940	Os03g0402800	1.134	0.004321813	1	ZIM domain containing protein, putative, expressed
LOC_Os11g41710	Os11g0635500	1.13	0.000962048	3	cytochrome P450, putative, expressed
LOC_Os09g36680	Os09g0537700	1.111	0.020342379	0	ribonuclease T2 family domain containing protein, expressed
LOC_Os03g56250	Os03g0773300	1.105	0.008691891	1	LRR receptor-like protein kinase, putative, expressed
LOC_Os04g15920	Os04g0229100	1.104	0.023566018	0	dehydrogenase, putative, expressed
LOC_Os08g08970	Os08g0189200	1.096	0.008242338	2	Cupin domain containing protein, expressed
LOC_Os07g41060	Os07g0601900	1.095	0.032512432	0	dihydroflavonol-4-reductase, putative, expressed
LOC_Os03g58290	Os03g0797300	1.082	0.042166915	0	indole-3-glycerol phosphate lyase, chloroplast precursor, putative, expressed
LOC_Os06g16640	Os06g0278000	1.078	0.004980284	1	carboxyl-terminal peptidase, putative, expressed
LOC_Os05g05680	Os05g0149400	1.035	0.017725838	0	1-aminocyclopropane-1-carboxylate oxidase, putative, expressed
LOC_Os06g29730	Os06g0493100	1.019	0.019983155	4	RALFL28 - Rapid ALkalinization Factor RALF family protein precursor, expressed
LOC_Os02g03410	Os02g0126400	1.014	0.004788142	0	CAMK_CAMK_like.12 - CAMK includes calcium/calmodulin depedent protein kinases, expressed
LOC_Os01g22900	Os01g0332100	1.007	0.012423981	1	neutral/alkaline invertase, putative, expressed
LOC_Os03g06520	Os03g0161200	1.004	0.004225916	1	sulfate transporter, putative, expressed
LOC_Os02g55970	Os02g0803300	1.002	0.046043213	0	ANTH, putative, expressed
LOC_Os09g38130	Os09g0554300	1	0.010479699	0	auxin efflux carrier component, putative, expressed
LOC_Os04g33660	Os04g0412500	0.988	0.030148233	2	bifunctional monodehydroascorbate reductase and carbonic anhydrasenectarin-3 precursor, putative, expressed

LOC_Os10g35460	Os10g0497700	0.972	0.044079374	3	COBRA, putative, expressed
LOC_Os07g31720	Os07g0500300	0.961	0.007814787	0	GTPase activating protein, putative, expressed
LOC_Os04g27670	Os04g0344100	0.945	0.024458713	0	terpene synthase family, metal binding domain containing protein, expressed
LOC_Os03g13840	Os03g0241900	0.933	0.036316675	2	senescence-associated protein, putative, expressed
LOC_Os11g42220	Os11g0641800	0.927	0.040742897	0	laccase precursor protein, putative, expressed
LOC_Os10g38540	Os10g0528900	0.908	0.035580782	0	glutathione S-transferase, putative, expressed
LOC_Os01g61850	NONE	0.905	0.003541198	0	NONE
LOC_Os10g02880	Os10g0118200	0.902	0.009951258	0	O-methyltransferase, putative, expressed
LOC_Os12g02320	Os12g0115100	0.891	0.006890709	4	LTPL12 - Protease inhibitor/seed storage/LTP family protein precursor, expressed
LOC_Os07g36170	Os07g0545800	0.887	0.027295537	1	chitin-inducible gibberellin-responsive protein, putative, expressed
LOC_Os04g10350	Os04g0182200	0.862	0.01931512	0	1-aminocyclopropane-1-carboxylate oxidase homolog 2, putative, expressed
LOC_Os03g03510	Os03g0126800	0.854	0.005968712	1	CAMK_KIN1/SNF1/Nim1_like.15 - CAMK includes calcium/calmodulin depedent protein kinases, expressed
LOC_Os03g29190	Os03g0405500	0.853	0.038921032	2	PDI, putative, expressed
LOC_Os07g45570	Os07g0650600	0.818	0.011795236	0	expressed protein
LOC_Os04g28620	Os04g0354600	0.817	0.044215314	0	male sterility protein, putative, expressed
LOC_Os04g50216	Os04g0592600	0.764	0.027454263	0	SNARE associated Golgi protein, putative, expressed
LOC_Os05g33630	Os05g0406100	0.754	0.015769534	1	inosine-uridine preferring nucleoside hydrolase family protein, putative, expressed
LOC_Os04g54830	Os04g0640850	0.737	0.026505655	3	expressed protein
LOC_Os05g48760	Os05g0561600	0.695	0.046043213	1	protein of unknown function DUF1421 domain containing protein, expressed
Down-regulated DEG					
LOC_Os03g54160	Os03g0752800	-5.988	4.1366E-29	1	OsMADS14 - MADS-box family gene with MIKCc type-box, expressed
LOC_Os10g26280	Os10g0402200	-2.49	0.026776109	3	ORC3 - Putative origin recognition complex subunit 3, expressed
LOC_Os03g51230	Os03g0722400	-2.408	0.002389921	0	SNF2 family N-terminal domain containing protein, expressed
LOC_Os09g37920	Os09g0551800	-2.305	0.017072543	0	helicase, putative, expressed
LOC_Os08g03560	Os08g0129600	-2.165	0.00852403	2	chloroplast unusual positioning protein, putative, expressed
LOC_Os03g50420	Os03g0712100	-2.155	0.013188333	0	expressed protein
LOC_Os01g67100	Os01g0896300	-2.144	0.00393184	3	expressed protein
LOC_Os06g14460	Os06g0256000	-2.128	0.011389258	0	chromosome condensation protein like, putative, expressed
LOC_Os03g48490	Os03g0691500	-2.121	0.016154124	3	centromere protein, putative, expressed
LOC_Os07g22580	Os07g0408500	-2.117	0.001653115	0	rhoGAP domain containing protein, expressed
LOC_Os03g11540	Os03g0214100	-2.086	0.0046764	1	RPA1B - Putative single-stranded DNA binding complex subunit 1, expressed
LOC_Os03g46920	Os03g0672400	-2.065	0.000165919	0	expressed protein
LOC_Os09g30070	Os09g0477700	-2.057	0.011710029	12	expressed protein
LOC_Os05g39850	Os05g0476200	-2.024	0.003803193	0	MCM3 - Putative minichromosome maintenance MCM complex subunit 3, expressed

LOC_Os01g56020	Os01g0765500	-2.022	0.003541198	1	expressed protein
LOC_Os02g56540	Os02g0810200	-2.02	0.000323953	3	kinesin motor domain containing protein, putative, expressed
LOC_Os04g48760	Os04g0576900	-2.02	0.001838904	0	leucine-rich repeat family protein, putative, expressed
LOC_Os12g39980	Os12g0590500	-2.005	0.006826929	0	kinesin motor domain containing protein, putative, expressed
LOC_Os11g29380	Os11g0484300	-2.004	0.000977446	12	MCM2 - Putative minichromosome maintenance MCM complex subunit 2, expressed
LOC_Os10g01570	Os10g0104900	-1.998	0.006680249	0	C-5 cytosine-specific DNA methylase, putative, expressed
LOC_Os05g41750	Os05g0497150	-1.998	0.018096953	0	RecF/RecN/SMC N terminal domain containing protein, expressed
LOC_Os05g14590	Os05g0235800	-1.993	0.011573367	3	MCM6 - Putative minichromosome maintenance MCM complex subunit 6, expressed
LOC_Os01g49200	Os01g0685900	-1.989	0.000165919	0	microtubule associated protein, putative, expressed
LOC_Os05g11980	Os05g0210500	-1.989	0.024218537	7	timeless protein, expressed
LOC_Os06g02530	Os06g0115700	-1.982	0.008756356	0	expressed protein
LOC_Os06g03710	Os06g0127800	-1.978	0.001128387	1	DELLA protein SLR1, putative, expressed
LOC_Os02g47150	Os02g0699700	-1.971	0.001582899	0	DNA topoisomerase 2, putative, expressed
LOC_Os06g08790	Os06g0187000	-1.966	0.011176548	1	ORC1 - Putative origin recognition complex subunit 1, expressed
LOC_Os01g67740	Os01g0904400	-1.943	0.006826929	0	chromosome segregation protein, putative, expressed
LOC_Os03g41100	Os03g0607600	-1.942	0.006826929	2	cyclin, putative, expressed
LOC_Os03g45760	Os03g0659800	-1.936	0.004335784	2	expressed protein
LOC_Os11g03430	Os11g0128400	-1.924	0.006826929	2	CDC45B - Putative DNA replication initiation protein, expressed
LOC_Os01g55560	Os01g0760900	-1.914	0.024458713	0	ABIL3, putative, expressed
LOC_Os12g13950	Os12g0242900	-1.899	0.007078376	1	POLA2 - Putative DNA polymerase alpha complex subunit, expressed
LOC_Os02g42560	Os02g0638200	-1.897	0.00447532	1	expressed protein
LOC_Os10g25450	Os10g0394200	-1.885	0.004938225	1	OsSub60 - Putative Subtilisin homologue, expressed
LOC_Os08g01100	Os08g0101100	-1.882	0.001061004	1	HMG1/2, putative, expressed
LOC_Os01g33040	Os01g0513900	-1.876	0.002370743	0	kinesin motor domain containing protein, expressed
LOC_Os12g06980	Os12g0167700	-1.875	0.009509699	3	SAP domain containing protein, expressed
LOC_Os03g18630	Os03g0297800	-1.871	0.012309165	3	receptor-like kinase RHG1, putative, expressed
LOC_Os04g39670	Os04g0472700	-1.862	0.039726956	1	expressed protein
LOC_Os03g11400	Os03g0212600	-1.861	0.007938142	6	targeting protein-related, putative, expressed
LOC_Os01g08150	Os01g0176500	-1.86	0.002566861	2	expressed protein
LOC_Os01g01170	Os01g0101800	-1.854	0.02911742	1	expressed protein
LOC_Os03g53920	Os03g0750300	-1.845	0.008691891	3	kinesin motor domain containing protein, putative, expressed
LOC_Os02g28850	Os02g0489800	-1.845	0.0124456	2	Kinesin motor domain domain containing protein, expressed
LOC_Os02g56520	Os02g0809900	-1.843	0.003093143	1	expressed protein
LOC_Os04g31050	Os04g0379800	-1.828	0.048977376	1	expressed protein

LOC_Os07g46540	Os07g0659500	-1.821	0.005111201	0	condensin complex subunit 1, putative, expressed
LOC_Os03g38010	Os03g0577100	-1.818	0.046388349	0	nuf2 family protein, expressed
LOC_Os02g55410	Os02g0797400	-1.814	0.007421892	1	MCM5 - Putative minichromosome maintenance MCM complex subunit 5, expressed
LOC_Os01g34870	Os01g0532800	-1.785	0.010067904	3	expressed protein
LOC_Os03g56070	Os03g0770900	-1.781	0.006135912	3	expressed protein
LOC_Os03g12140	Os03g0221500	-1.779	0.002356747	0	glucan endo-1,3-beta-glucosidase precursor, putative, expressed
LOC_Os09g38710	Os09g0560000	-1.779	0.014547125	0	HEAT repeat family protein, putative, expressed
LOC_Os01g15480	Os01g0259400	-1.759	0.013182537	16	EMB3013, putative, expressed
LOC_Os02g50910	Os02g0742800	-1.756	0.005998005	0	expressed protein
LOC_Os12g42160	Os12g0616000	-1.752	0.004789434	1	kinesin motor domain containing protein, putative, expressed
LOC_Os01g09580	Os01g0191800	-1.75	0.011864396	0	CAMK_CAMK_like_Aur_like.1 - CAMK includes calcium/calmodulin depedent protein kinases, expressed
LOC_Os03g02290	Os03g0114000	-1.741	0.010206293	1	kinesin motor domain containing protein, putative, expressed
LOC_Os02g10020	Os02g0193600	-1.74	0.024218537	0	Mad3/BUB1 homology region 1 domain containing protein, expressed
LOC_Os03g19080	Os03g0302900	-1.735	0.001020368	0	expressed protein
LOC_Os01g64820	Os01g0868300	-1.731	0.033631101	0	POLA1 - Putative DNA polymerase alpha catalytic subunit, expressed
LOC_Os05g36280	Os05g0438700	-1.727	0.000648202	1	histone H3, putative, expressed
LOC_Os12g31810	Os12g0502300	-1.725	0.026505655	0	cyclin, putative, expressed
LOC_Os12g01700	Os12g0107700	-1.72	0.000948597	2	inactive receptor kinase At2g26730 precursor, putative, expressed
LOC_Os04g35420	Os04g0433800	-1.718	0.029243297	0	helicase conserved C-terminal domain containing protein, expressed
LOC_Os01g70560	Os01g0931200	-1.716	0.004938225	0	expressed protein
LOC_Os03g51870	Os03g0728500	-1.716	0.009349775	0	FHA domain containing protein, putative, expressed
LOC_Os10g28230	Os10g0418000	-1.712	0.000271188	3	Core histone H2A/H2B/H3/H4 domain containing protein, putative, expressed
LOC_Os12g13570	Os12g0238000	-1.712	0.019537207	6	MYB family transcription factor, putative, expressed
LOC_Os06g51110	Os06g0726800	-1.709	0.011710029	0	cyclin, putative, expressed
LOC_Os05g38480	Os05g0459400	-1.708	0.009509699	2	kinesin motor domain containing protein, putative, expressed
LOC_Os07g41370	Os07g0605200	-1.707	1.56687E-08	5	OsMADS18 - MADS-box family gene with MIKCc type-box, expressed
LOC_Os08g44420	Os08g0558400	-1.699	0.007510897	0	kinesin-related protein, putative, expressed
LOC_Os01g42070	Os01g0605500	-1.697	0.015117733	0	kinesin motor domain containing protein, putative, expressed
LOC_Os01g36390	Os01g0544450	-1.682	0.011176548	3	MCM4 - Putative minichromosome maintenance MCM complex subunit 4, expressed
LOC_Os05g33890	Os05g0409400	-1.68	0.029106256	0	microtubule associated protein, putative, expressed
LOC_Os11g05730	Os11g0155900	-1.671	0.040742897	4	histone H3, putative, expressed
LOC_Os05g33030	Os05g0397900	-1.668	0.011452357	0	kinesin motor domain containing protein, putative, expressed
LOC_Os03g39020	Os03g0587200	-1.662	0.002423154	7	Kinesin motor domain domain containing protein, expressed

LOC_Os10g42490	Os10g0575600	-1.662	0.028197883	0	homeobox and START domains containing protein, putative, expressed
LOC_Os09g38450	Os09g0556750	-1.661	0.017089824	0	expressed protein
LOC_Os03g52650	Os03g0736500	-1.66	0.001582899	0	syntaxin-related protein, putative, expressed
LOC_Os03g04550	Os03g0138500	-1.657	0.001039621	0	expressed protein
LOC_Os01g31800	Os01g0502700	-1.655	0.01040715	4	Core histone H2A/H2B/H3/H4 domain containing protein, putative, expressed
LOC_Os03g17960	Os03g0288900	-1.655	0.014052588	1	expressed protein
LOC_Os10g35580	Os10g0498900	-1.654	0.0046764	7	ATEB1A-like microtubule associated protein, putative, expressed
LOC_Os03g17100	Os03g0279200	-1.651	0.000973277	0	Core histone H2A/H2B/H3/H4 domain containing protein, putative, expressed
LOC_Os08g32600	Os08g0421800	-1.645	0.013146309	0	STE_MEKK_ste11_MAP3K.21 - STE kinases include homologs to sterile 7, sterile 11 and sterile 20 from yeast, expressed
LOC_Os05g41390	Os05g0493500	-1.641	0.003068219	4	cyclin, putative, expressed
LOC_Os04g50960	Os04g0597900	-1.635	0.004466763	1	expressed protein
LOC_Os02g15810	Os02g0258200	-1.631	0.000165919	0	HMG1/2, putative, expressed
LOC_Os07g32390	Os07g0507200	-1.629	0.008084822	1	targeting protein-related, putative, expressed
LOC_Os04g47890	Os04g0566600	-1.62	0.047682262	0	MYB family transcription factor, putative, expressed
LOC_Os03g17164	Os03g0279816	-1.605	0.038921032	6	kinesin-related protein, putative, expressed
LOC_Os12g41230	Os12g0605500	-1.603	0.016154124	7	tesmin/TSO1-like CXC domain containing protein, expressed
LOC_Os01g59120	Os01g0805600	-1.599	0.001797948	0	cyclin, putative, expressed
LOC_Os08g38300	Os08g0490900	-1.594	2.52939E-07	0	Core histone H2A/H2B/H3/H4 domain containing protein, putative, expressed
LOC_Os03g06670	Os03g0162200	-1.587	7.23581E-05	1	Core histone H2A/H2B/H3/H4 domain containing protein, putative, expressed
LOC_Os09g36790	Os09g0539000	-1.587	0.047605144	3	expressed protein
LOC_Os04g15800	Os04g0228100	-1.585	0.008936865	0	expressed protein
LOC_Os07g08500	Os07g0182900	-1.58	0.04248804	0	C-5 cytosine-specific DNA methylase, putative, expressed
LOC_Os08g42600	Os08g0538700	-1.579	0.029243297	3	retinoblastoma-related protein-like, putative, expressed
LOC_Os01g74146	Os01g0972900	-1.572	0.021393347	0	WD repeat-containing protein, putative, expressed
LOC_Os02g56130	Os02g0805200	-1.568	0.002538809	1	PCNA - Putative DNA replicative polymerase clamp, expressed
LOC_Os11g37100	Os11g0580000	-1.565	0.016297586	2	expressed protein
LOC_Os04g41900	Os04g0496300	-1.564	0.012830513	1	expressed protein
LOC_Os07g33360	Os07g0517300	-1.555	0.030148233	8	expressed protein
LOC_Os01g01890	Os01g0108800	-1.549	0.007078376	0	expressed protein
LOC_Os04g28260	Os04g0350300	-1.521	0.023582979	2	Kinesin motor domain domain containing protein, expressed
LOC_Os01g64640	Os01g0866200	-1.516	0.001582899	1	histone H3, putative, expressed
LOC_Os12g37400	Os12g0560700	-1.514	0.017587983	0	MCM7 - Putative minichromosome maintenance MCM complex subunit 7, expressed
LOC_Os02g39390	Os02g0606700	-1.514	0.032236886	2	expressed protein
LOC_Os10g42230	Os10g0572900	-1.509	0.014124235	2	AT hook motif domain containing protein, expressed

LOC_Os09g37510	Os09g0547200	-1.496	0.041075289	0	DUF292 domain containing protein, expressed
LOC_Os03g44890	Os03g0652000	-1.49	0.011795236	2	anthranilate phosphoribosyltransferase, putative, expressed
LOC_Os12g40860	Os12g0601000	-1.489	0.009452213	0	Leucine Rich Repeat family protein, expressed
LOC_Os04g40940	Os04g0486500	-1.487	0.015419278	1	mitotic spindle checkpoint protein MAD2, putative, expressed
LOC_Os10g42660	Os10g0577400	-1.481	0.027850644	1	expressed protein
LOC_Os07g01490	Os07g0105700	-1.479	0.027850644	1	kinesin motor domain containing protein, putative, expressed
LOC_Os05g02300	Os05g0113900	-1.477	0.012423981	7	Core histone H2A/H2B/H3/H4 domain containing protein, putative, expressed
LOC_Os06g42020	Os06g0625700	-1.475	0.024416414	4	CSLA9 - cellulose synthase-like family A, expressed
LOC_Os02g45940	Os02g0684500	-1.473	0.003068219	1	Core histone H2A/H2B/H3/H4 domain containing protein, putative, expressed
LOC_Os09g36900	Os09g0540600	-1.472	0.033339486	0	WD domain, G-beta repeat domain containing protein, expressed
LOC_Os12g17530	Os12g0273700	-1.467	0.030148233	2	expressed protein
LOC_Os05g49860	Os05g0574300	-1.454	0.000555408	0	Core histone H2A/H2B/H3/H4 domain containing protein, putative, expressed
LOC_Os08g22604	Os08g0316900	-1.453	0.020071293	0	expressed protein
LOC_Os11g01620	Os11g0107700	-1.449	0.003120053	0	inactive receptor kinase At2g26730 precursor, putative, expressed
LOC_Os12g36890	Os12g0555600	-1.438	0.006327022	5	CSLD4 - cellulose synthase-like family D, expressed
LOC_Os05g01230	Os05g0102600	-1.429	0.019983155	4	zinc finger, C3HC4 type domain containing protein, expressed
LOC_Os02g19990	Os02g0302900	-1.421	0.025871527	0	reticulon domain containing protein, putative, expressed
LOC_Os02g47180	Os02g0700100	-1.398	0.018770734	1	WD repeat-containing protein, putative, expressed
LOC_Os11g35090	Os11g0552600	-1.392	0.045669572	0	kinesin motor domain containing protein, putative, expressed
LOC_Os01g55220	Os01g0756900	-1.382	0.013322879	1	expressed protein
LOC_Os03g16709	Os03g0274400	-1.381	0.046043213	5	expressed protein
LOC_Os01g12860	Os01g0229000	-1.376	0.027850644	2	MYB family transcription factor, putative, expressed
LOC_Os01g13260	Os01g0233500	-1.374	0.01140625	0	cyclin-A1, putative, expressed
LOC_Os04g43300	Os04g0512400	-1.373	0.029243297	0	BRCA1 C Terminus domain containing protein, expressed
LOC_Os07g23660	Os07g0418700	-1.371	0.001582899	0	retrotransposon protein, putative, unclassified, expressed
LOC_Os01g16650	Os01g0273100	-1.367	0.012006746	4	ubiquitin-conjugating enzyme, putative, expressed
LOC_Os03g41060	Os03g0607200	-1.365	0.002423154	2	GASR2 - Gibberellin-regulated GASA/GAST/Snakin family protein precursor, putative, expressed
LOC_Os02g06340	Os02g0158100	-1.339	0.023171071	4	EH domain-containing protein 1, putative, expressed
LOC_Os12g24550	Os12g0433500	-1.337	0.025871895	11	expressed protein
LOC_Os08g40170	Os08g0512600	-1.333	0.00852403	0	cyclin-dependent kinase B2-1, putative, expressed
LOC_Os12g42700	Os12g0621700	-1.329	0.038504571	1	expressed protein
LOC_Os02g47130	Os02g0699433	-1.305	0.018253415	2	expressed protein
LOC_Os01g11550	Os01g0213800	-1.303	0.000165919	0	TCP family transcription factor, putative, expressed
LOC_Os12g34510	Os12g0530000	-1.299	0.012830513	1	Core histone H2A/H2B/H3/H4 domain containing protein, putative, expressed

LOC_Os03g49750	Os03g0704400	-1.282	0.045151628	4	protein kinase family protein, putative, expressed
LOC_Os06g07210	Os06g0168600	-1.269	0.018609635	2	ribonucleoside-diphosphate reductase large subunit, putative, expressed
LOC_Os04g47580	Os04g0563700	-1.269	0.037174644	1	cyclin, putative, expressed
LOC_Os03g21160	Os03g0329200	-1.258	0.025993405	0	RNA-binding zinc finger protein, putative, expressed
LOC_Os07g23640	Os07g0418600	-1.254	3.9648E-05	0	retrotransposon protein, putative, Ty3-gypsy subclass, expressed
LOC_Os09g38400	Os09g0556300	-1.229	0.001039621	1	prostatic spermine-binding protein precursor, putative, expressed
LOC_Os12g44090	Os12g0638100	-1.2	0.01236593	1	leucine-rich repeat family protein, putative, expressed
LOC_Os09g27060	Os09g0442700	-1.194	0.040613484	0	SNF2 family N-terminal domain containing protein, expressed
LOC_Os02g41904	Os02g0629800	-1.193	0.004938225	0	DEF7 - Defensin and Defensin-like DEFL family, expressed
LOC_Os12g39830	Os12g0588800	-1.193	0.013188333	5	cyclin, putative, expressed
LOC_Os12g38140	Os12g0569200	-1.183	0.010741315	0	expressed protein
LOC_Os11g32100	Os11g0523700	-1.175	0.00393184	0	inducer of CBF expression 1, putative, expressed
LOC_Os07g42860	Os07g0620800	-1.156	0.007078376	0	cyclin, putative, expressed
LOC_Os10g05600	Os10g0146200	-1.149	0.013182537	0	thaumatin-like protein 1 precursor, putative, expressed
LOC_Os08g44360	Os08g0557800	-1.137	0.035580782	0	male sterility protein 2, putative, expressed
LOC_Os09g37910	Os09g0551600	-1.136	0.002538809	1	HMG1/2, putative, expressed
LOC_Os06g03380	Os06g0124300	-1.132	0.020310352	4	expressed protein
LOC_Os05g44400	Os05g0520300	-1.079	0.012830513	0	GATA zinc finger domain containing protein, expressed
LOC_Os05g32760	Os05g0394200	-1.067	0.008744052	4	BZIP protein, putative, expressed
LOC_Os04g38570	Os04g0459000	-1.067	0.048059378	0	multidrug resistance protein, putative, expressed
LOC_Os01g05610	Os01g0149400	-1.059	0.002370743	0	Core histone H2A/H2B/H3/H4 domain containing protein, putative, expressed
LOC_Os05g37160	Os05g0443800	-1.027	0.014412512	3	tubulin/FtsZ domain containing protein, putative, expressed
LOC_Os07g05190	Os07g0145400	-1.023	0.017614214	2	leucine-rich repeat family protein, putative, expressed
LOC_Os01g56320	Os01g0769200	-1	0.006894113	0	OsSub4 - Putative Subtilisin homologue, expressed
LOC_Os05g34700	Os05g0419800	-0.975	0.008242338	1	GDSL-like lipase/acylhydrolase, putative, expressed
LOC_Os09g37600	Os09g0548200	-0.973	0.006325692	2	lysM domain-containing GPI-anchored protein precursor, putative, expressed
LOC_Os03g05520	Os03g0149200	-0.969	0.020374462	0	nicotiana lesion-inducing like, putative, expressed
LOC_Os07g40480	Os07g0596000	-0.965	0.027295537	0	zinc finger family protein, putative, expressed
LOC_Os06g43600	Os06g0643500	-0.953	0.038921032	1	LTPL129 - Protease inhibitor/seed storage/LTP family protein precursor, expressed
LOC_Os01g06010	Os01g0153300	-0.936	0.022751123	0	Core histone H2A/H2B/H3/H4 domain containing protein, putative, expressed
LOC_Os01g05970	Os01g0152900	-0.93	0.020046687	1	OsFBO1 - F-box and other domain containing protein, expressed
LOC_Os01g14850	Os01g0251400	-0.918	0.007078376	7	MFS18 protein precursor, putative, expressed
LOC_Os03g04240	Os03g0135100	-0.917	0.017072543	0	glutathione S-transferase, putative, expressed
LOC_Os05g04530	Os05g0135900	-0.916	0.034554753	4	IF, putative, expressed
LOC_Os01g62230	Os01g0839500	-0.882	0.011176548	3	Core histone H2A/H2B/H3/H4 domain containing protein, putative, expressed

LOC_Os10g40810	Os10g0557600	-0.878	0.041075289	0	GATA zinc finger domain containing protein, expressed
LOC_Os07g35520	Os07g0539400	-0.84	0.041075289	0	glucan endo-1,3-beta-glucosidase precursor, putative, expressed
LOC_Os06g45990	Os06g0671800	-0.83	0.046441077	0	patellin-5, putative, expressed
LOC_Os07g23540	Os07g0418100	-0.819	0.038921032	0	retrotransposon protein, putative, unclassified, expressed
LOC_Os03g01530	Os03g0105600	-0.816	0.015380312	2	tubulin/FtsZ domain containing protein, putative, expressed
LOC_Os01g60740	Os01g0822900	-0.8	0.047687794	1	LTPL16 - Protease inhibitor/seed storage/LTP family protein precursor, expressed
LOC_Os05g30750	Os05g0370600	-0.786	0.039206768	1	anthranilate phosphoribosyltransferase, putative, expressed

GO terms enriched from DEG

Using (RiceNETDB) (Liu et al., 2013a) differentially expressed genes were assessed for gene ontology enrichment between HM-*mpg1* and WT-ns for biological processes (Figure 4.4). Of the differentially up-regulated genes, some of the highest enrichment categories belonged to physiological process, metabolism, cellular process, and response to stimulus. Of the differentially down-regulated genes, some of the highest enrichment categories were DNA metabolism, cell cycle, metabolism, and development.

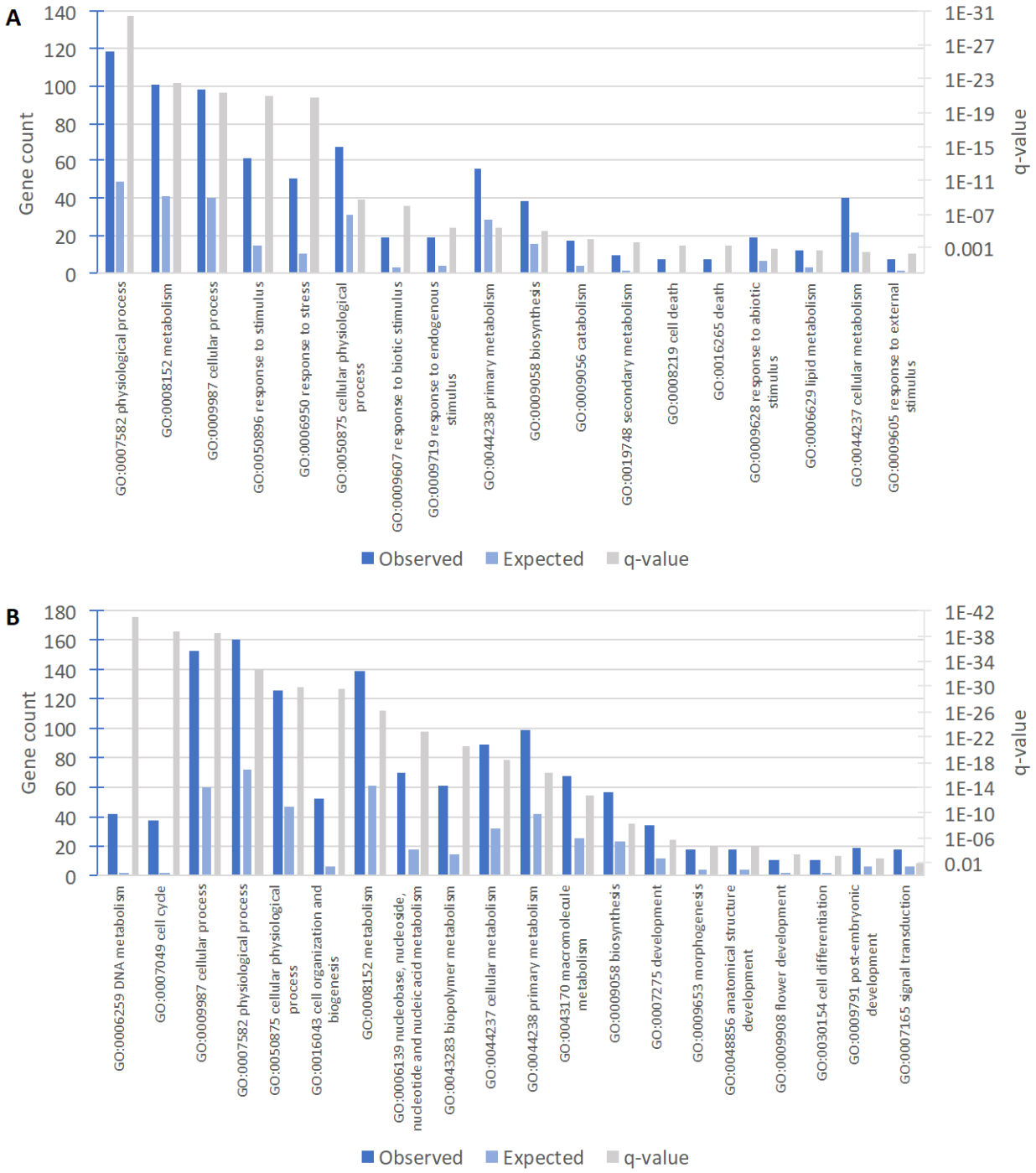


Figure 4.4: RiceNETDB biological process gene ontology enrichment analysis from DEG from 32 days post-planting tiller tissue.
(A) up-regulated DEG, (B) down-regulated DEG. RiceNETDB accounts for genes including their differential splice isoforms.

Again, using (RiceNETDB) (Liu et al., 2013a) gene ontology enrichment for differentially expressed genes pertaining to cellular localization was assessed (Figure 4.5). Many of the enrichment categories were rather non-descript. Of the differentially up-regulated genes, the highest enrichment categories belonged to extracellular region, cell part, cell wall, and external encapsulating structure. Of the differentially down-regulated genes the highest enrichment categories were cell part, intracellular part, organelle, nucleus, cytoskeleton, and cell wall to name a few.

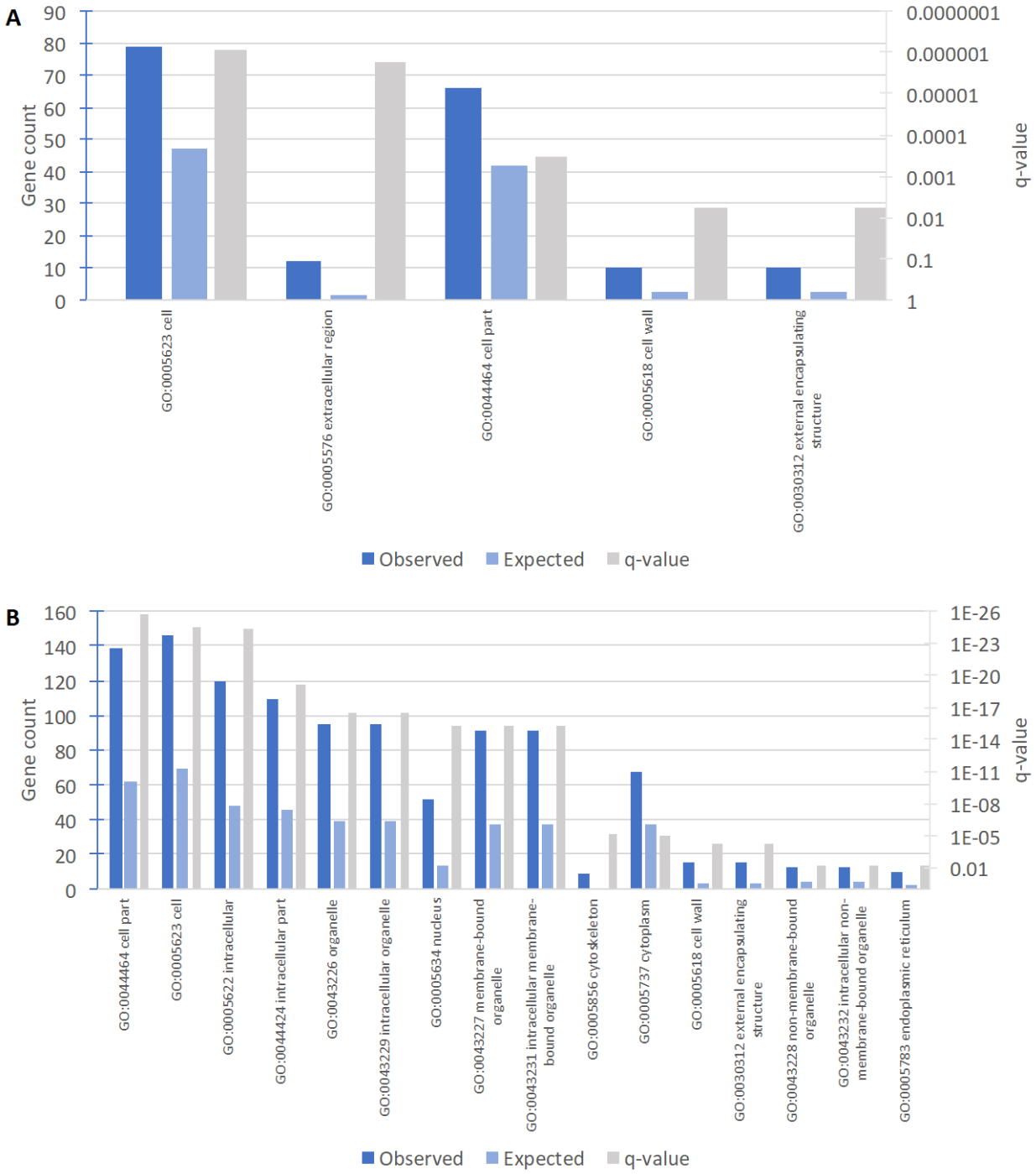


Figure 4.5: RiceNETDB cellular localization gene ontology enrichment analysis from DEG from 32 days post-planting tiller tissue.
(A) up-regulated DEG, (B) down-regulated DEG. RiceNETDB accounts for genes including their differential splice isoforms.

(RiceNETDB) (Liu et al., 2013a) was used again to assess genes pertaining to molecular functionality (Figure 4.6). Of the up-regulated genes, gene enrichment categories pertained to catalytic activity, binding, transferase activity, transcription factor activity, and transcription regulator activity. The down-regulated genes were enriched for DNA binding, motor activity, nucleic acid binding, enzyme regulator activity, protein binding, and transcription factor activity. These categories are not necessarily telling and will require investigation into individual categories genes to help hypothesize their potential role in *mpg1*'s phenotype.

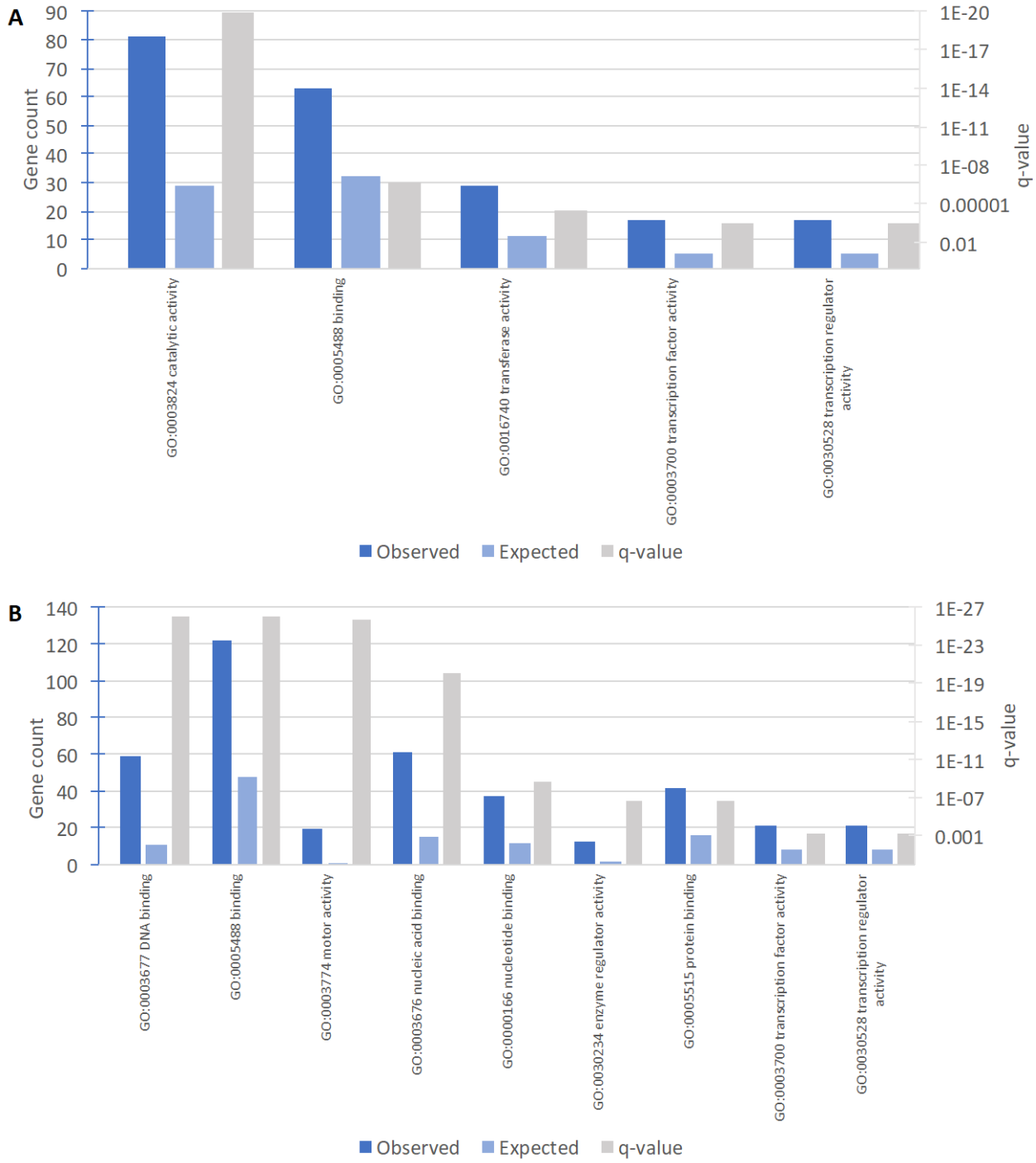


Figure 4.6: RiceNETDB molecular function gene ontology enrichment analysis from DEG from 32 days post-planting tiller tissue.
(A) up-regulated DEG, (B) down-regulated DEG. RiceNETDB accounts for genes including their differential splice isoforms.

GO term enrichment of AP2/ERF binding motif-containing genes

Although we are unaware of the specific functional GCC-box motifs that are targeted by MPG1 within these differentially regulated genes, we are aware of the genes that at least contain the GCC-box motif sequence in their promoter region. Assessing the genes containing the (GCCGCC) sequence within their promoter region along with their ontological function could reveal categories of genes that *MPG1* might directly regulate. Using only the DEG that had at least one GCC-box motif in the possible promoter region (-2000 from TSS), we assessed for gene ontology enrichment using (RiceNETDB) (Figure 4.7, 4.8).

Of the up-regulated genes that contain a GCC-box motif, genes were enriched for several ontological categories. Under biological process, some of the categories were response to stress, physiological process, and metabolism that were enriched. Some of the ontological categories pertaining to cellular component consisted of extracellular region, cell wall, and external encapsulating structure. A few of the molecular function ontological categories included catalytic activity, binding, and transcription factor activity.

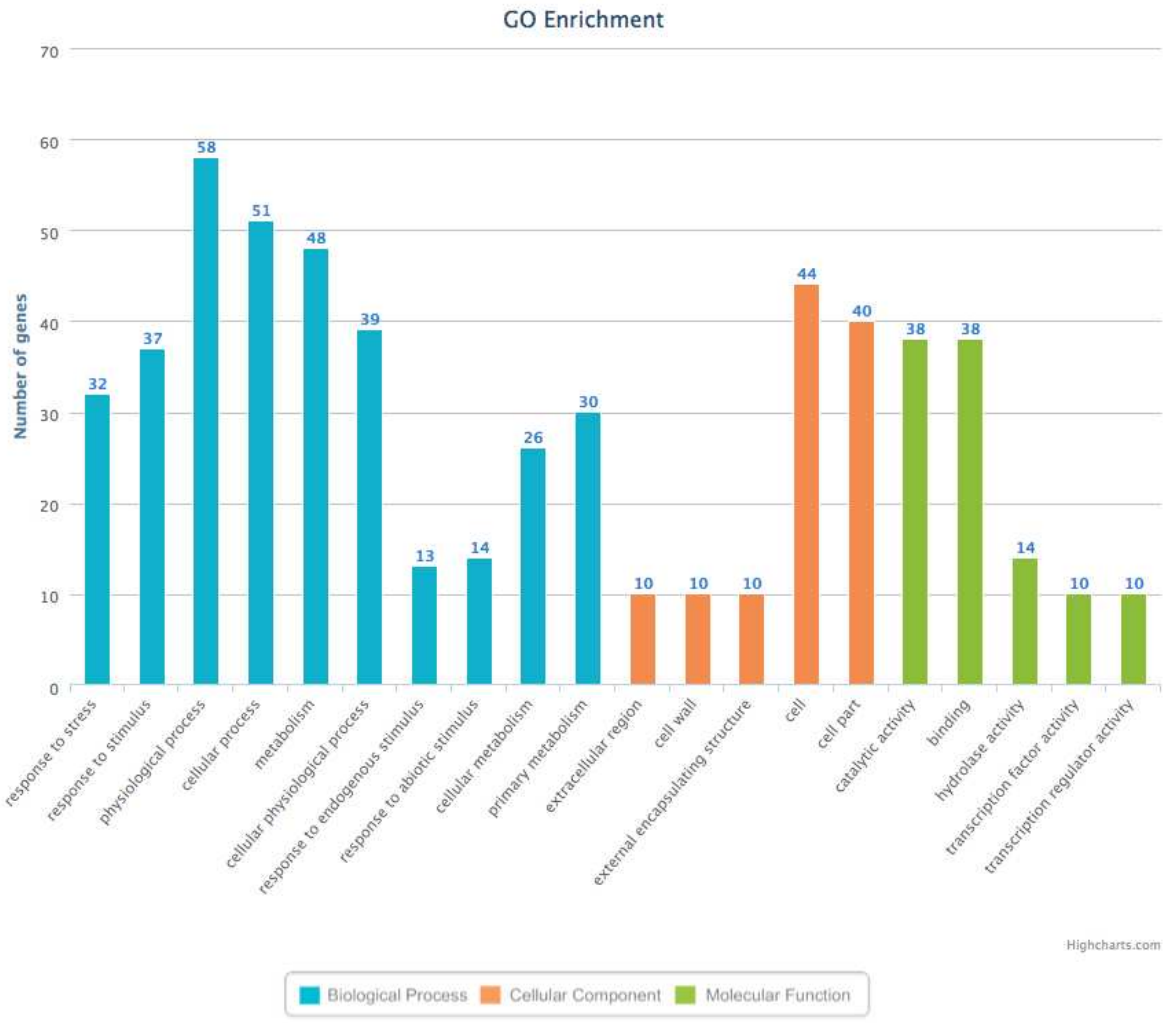


Figure 4.7: RiceNETDB ontological enrichment categories pertaining to up-regulated DEG containing at least one GCC-box motif in -2000 from TSS.
RiceNETDB accounts for genes including their differential splice isoforms.

Of the down-regulated genes that contain a GCC-box motif, some of the biological process categories consisted of DNA metabolism, cellular process, cell cycle, cell organization and biogenesis, and cell differentiation. Some of the ontological categories pertaining to cellular component were membrane bound-organelle, nucleus, cell wall, cytoplasm, and cytoskeleton. And some of the molecular function ontological categories that were enriched were DNA binding, motor activity, and transcription factor activity.

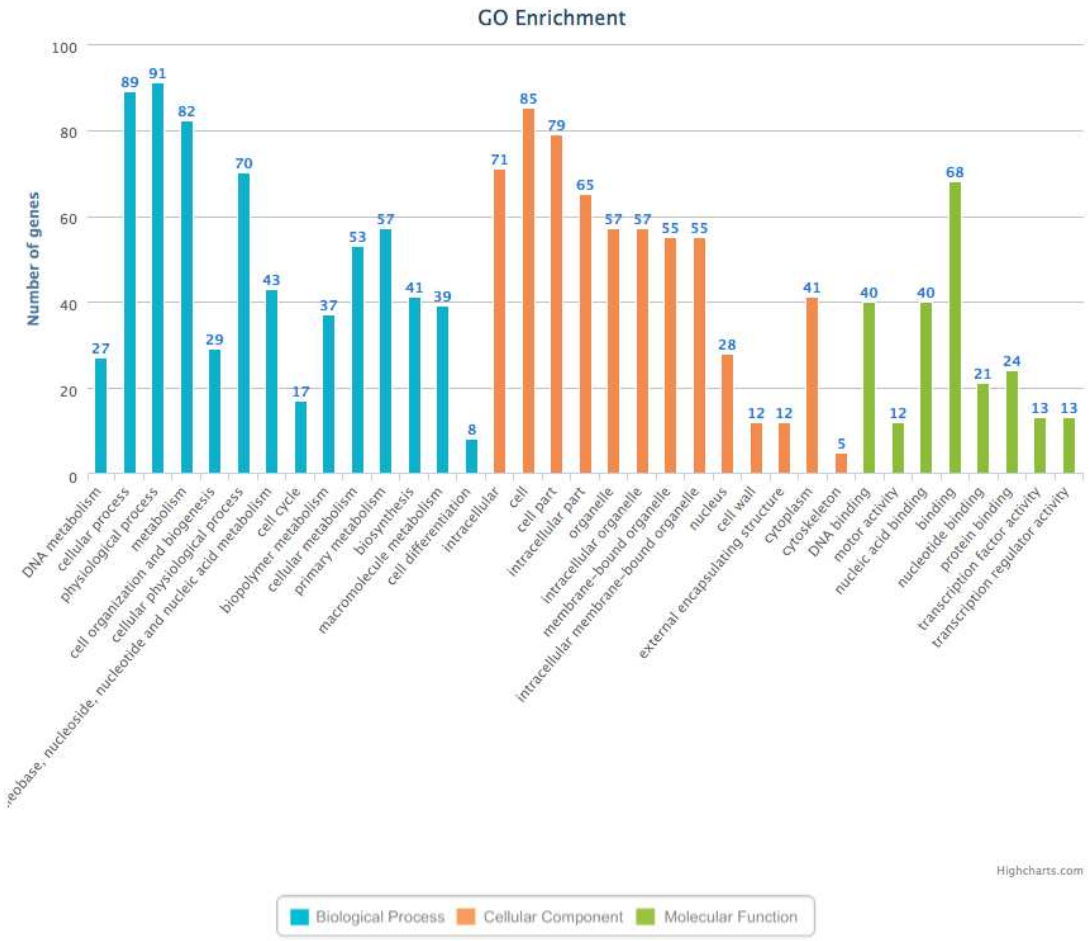


Figure 4.8: RiceNETDB ontological enrichment categories pertaining to down-regulated DEG containing at least one GCC-box motif in -2000 from TSS.
RiceNETDB accounts for genes including their differential splice isoforms.

Use of these categories and genes might aid future molecular analyses (transactivation assays, Y1H, EMSA, or ChIP-seq) to verify MPG1 targets. ChIP-seq, in particular, would prove useful to globally validate and verify target sites and influenced genes.

The MSU6.0 genome build contains 55,956 genes. Of those 19,273 contain at least one GCC-box motif in their promoter regions within -2000 from TSS. Using a 95 percent confidence interval on all genes, we uncovered genes that are enriched for this motif (containing 6 or more GCC-box motifs). There are 3,364 gens in rice that contain 6 or more of these motifs within their promoters. Of those genes that are enriched for this motif, we cross referenced them against our

DEG (Table 4.2). There were a total of five up-regulated and fourteen down-regulated DEG that contain an enriched number of GCC-box motifs.

Table 4.2: DEG list from RNA-seq analysis of HM-*mpg1* compared to WT-ns within 32 days post-planting tiller tissue containing an enriched number of GCC-box motifs in -2000 from TSS.

MSU ID	RAP ID	log2 Fold Change	GCC-box motifs	Description
Up-regulated DEG				
LOC_Os01g64360	Os01g0863300	2.409	6	MYB family transcription factor, putative, expressed
LOC_Os04g41960	Os04g0497000	1.895	8	NADP-dependent oxidoreductase, putative, expressed
LOC_Os02g32580	Os02g0527200	1.75	6	expressed protein
LOC_Os08g10150	Os08g0201700	1.335	9	SHR5-receptor-like kinase, putative, expressed
LOC_Os04g52440	Os04g0614500	1.269	10	aminotransferase, putative, expressed
Down-regulated DEG				
LOC_Os09g30070	Os09g0477700	-2.057	12	expressed protein
LOC_Os11g29380	Os11g0484300	-2.004	12	MCM2 - Putative minichromosome maintenance MCM complex subunit 2, expressed
LOC_Os05g11980	Os05g0210500	-1.989	7	timeless protein, expressed
LOC_Os03g11400	Os03g0212600	-1.861	6	targeting protein-related, putative, expressed
LOC_Os01g15480	Os01g0259400	-1.759	16	EMB3013, putative, expressed
LOC_Os12g13570	Os12g0238000	-1.712	6	MYB family transcription factor, putative, expressed
LOC_Os03g39020	Os03g0587200	-1.662	7	Kinesin motor domain domain containing protein, expressed
LOC_Os10g35580	Os10g0498900	-1.654	7	ATEB1A-like microtubule associated protein, putative, expressed
LOC_Os03g17164	Os03g0279816	-1.605	6	kinesin-related protein, putative, expressed
LOC_Os12g41230	Os12g0605500	-1.603	7	tesmin/TSO1-like CXC domain containing protein, expressed
LOC_Os07g33360	Os07g0517300	-1.555	8	expressed protein
LOC_Os05g02300	Os05g0113900	-1.477	7	Core histone H2A/H2B/H3/H4 domain containing protein, putative, expressed
LOC_Os12g24550	Os12g0433500	-1.337	11	expressed protein
LOC_Os01g14850	Os01g0251400	-0.918	7	MFS18 protein precursor, putative, expressed

Genes that were differentially regulated and enriched for the GCC-box motifs were assessed by gene ontology enrichment using (RiceNetDB) (Liu et al., 2013a). The up-regulated genes as a whole were not enriched for any particular ontological category, while the down-regulated genes were enriched for DNA binding and motor activity under molecular function ontological enrichment categories.

Mutant plant *mpg1* experiences differential expression of 1249 genes in 42 days post-planting tiller tissue

Differential gene expression was analyzed using the STAR-HTSeq-edgeR alignment, mapping, and assessment pipeline (APPENDIX). Libraries were constructed from individual plant's whole stem tissue, three HM-*mpg1* and three WT-ns plants. Results indicate that there are 1249 differentially expressed genes (Figure 4.9), of which, 265 were up-regulated, and 984 were down regulated in HM-*mpg1* compared to WT-ns. Because MPG1 targets the GCC-box motif directly, we again assessed the differentially expressed genes promoter regions (-2000 from TSS) for the presence of this motif. Of the differentially expressed genes, 127 up-regulated genes and 438 down-regulated genes contained at least one GCC-box (Figure 4.9, C). Although, we cannot verify that these genes are being targeted by MPG1 they are good to note for further investigation.

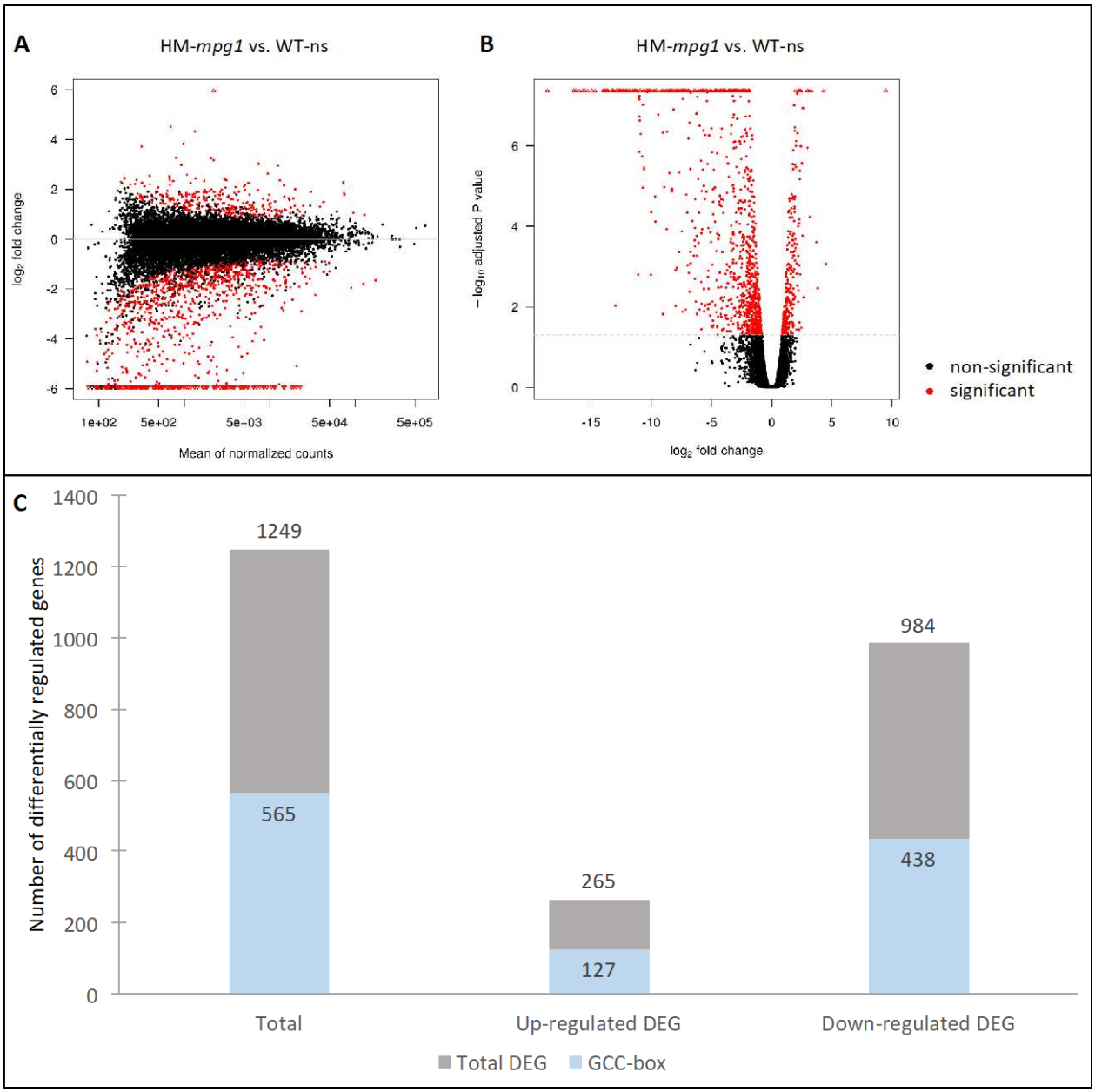


Figure 4.9: Differentially expressed genes found from 42 days post-planting stem tissue. (A) MA-plot, (B) volcano plot, and (C) bar graph with number of genes up- and down-regulated. The bar graph also contains the number of genes differentially expressed that contain a GCC-box motif in their promoter region.

MPG1 is the most noticeably differentially expressed gene within the neighboring region of the T-DNA insertion at 42 days post-planting in stem tissue

Using the Integrated Genome Viewer (IGV) software, we mapped reads from each of the samples (individual plants) to the genome to visually assess the region of the T-DNA integration.

Assessment of nine surrounding genes shows that *MPG1* is up-regulated in HM-*mpg1* compared to WT-ns which have little to no expression, and is the most noticeably differentially regulated gene between assessed genotypes (Figure 4.10).

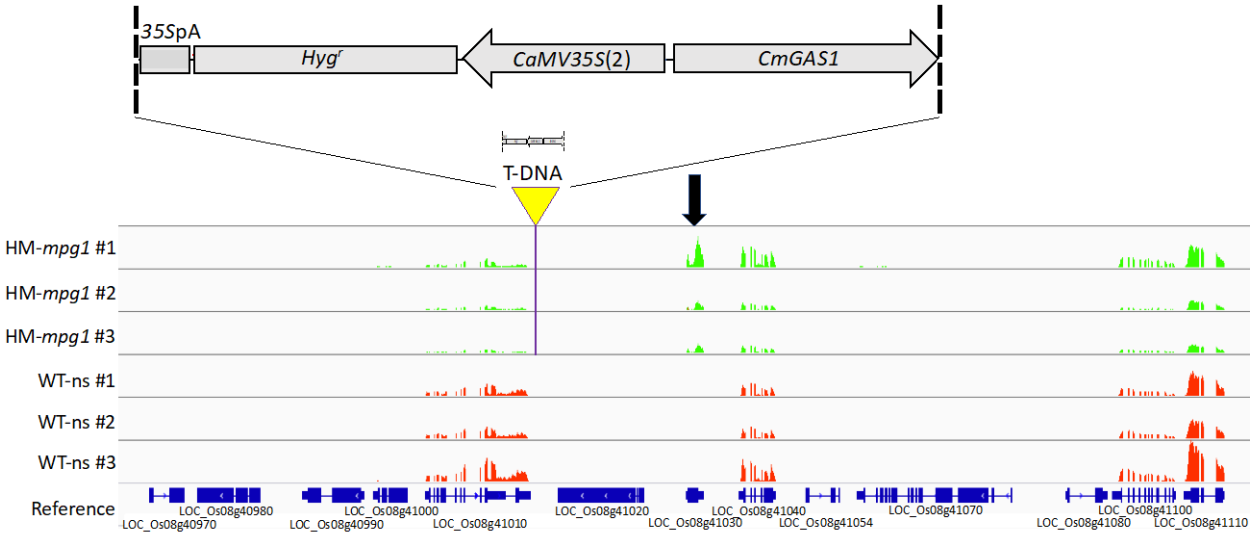


Figure 4.10: Gene expression surrounding the site of the T-DNA insertion.
IGV visualization of read abundance of genes proximal to the T-DNA insertion in tissue collected from 42 days post-planting whole tiller tissue. The schematic shows the orientation, location, and size of the insertion (purple bar) in *mpg1* plants. Each row represents a single biological replicate. *MPG1* (*Os08g41030*) is notated by the black arrow. Scale: 0-1123 read abundance.

Shift from vegetative development to reproductive development in wild-type plants occurs around 42 days post-planting

mpg1 plants experience a delay in flowering compared to wild-type plants (~14 days). Prior to extensively evaluating the RNA-seq experiments we wanted to insure that we were capturing transcriptome differences between HM-*mpg1* and WT-ns at similar developmental stages. Although we have observed a two week delay in heading, we were unaware at which timepoint the actual shift from vegetative phase to reproductive phase of development occurred

in Kitaake. The actual transition to reproductive development occurs with panicle initiation at the shoot apical meristem located in the crown of the plant. To better ascertain this timepoint we grew plants and dissected the crown at various timepoints post-planting and searched for panicle imitation. Panicle initiation is characterized by the presence and formation of 'green-ring' tissue at the basal internode, accompanied by basal internode elongation proceeding to panicle tissue formation. In line with our RNA-seq experiments, we evaluated tissues respective of the 32 days post-planting and 42 days post-planting timepoints for their developmental progress (Figure 4.11). Both HM-*mpg1* and WT-ns plants at 32 days post-planting were absent of any signs of reproductive transition, however at the 42 days post-planting timepoint WT-ns presented with green-ring formation, internode elongation, and the beginning of panicle tissue formation.

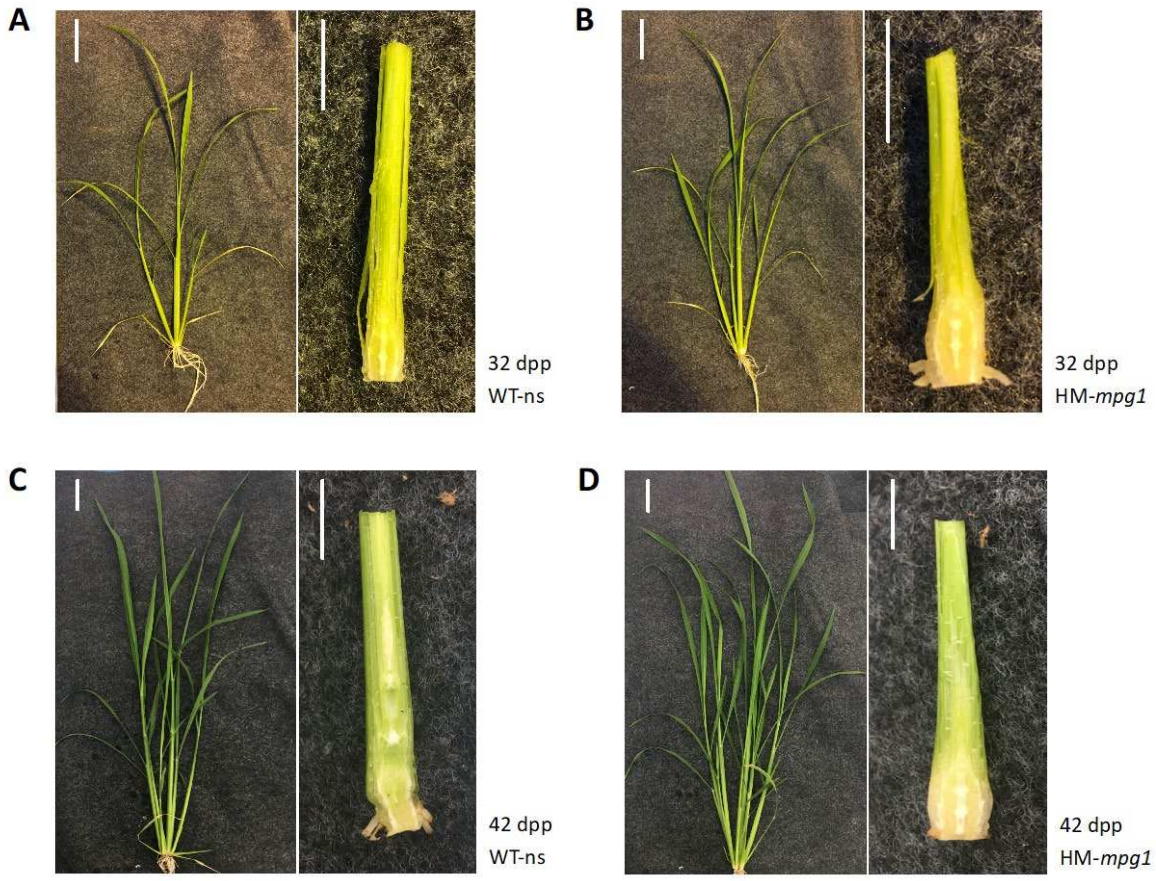


Figure 4.11: Assessment of temporal transition from vegetative to reproductive phase of development.
(A,B) pictures of whole plant at 32 days post-planting along with same plant cross section of crown tissue. (C,D) pictures of whole plant at 42 days post-planting along with same plant cross section of crown tissue. Green-ring formation and internode elongation and panicle development can be visualized in WT-ns plants at 42 days post-planting. Whole plant scale = 5.0 cm, cross section scale = 1.0 cm.

Evaluation of panicle initiation suggests that WT-ns plants are in a different phase of development than HM-*mpg1* plants at 42 days post-planting. Assessment of their transcriptomes at this timepoint would prove difficult because it would no longer be assessing just the effects of the T-DNA insertion and increased expression of *MPG1*, but also the differences regarding different stages of development. Because of this, and that both experiments were conducted in differing growth conditions, we opted to primarily investigate differences in just the 32 days post-planting timepoint to more appropriately gauge differential gene expression between HM-

mpg1 and WT-ns, as they appear to be in similar stages of development. While this being observed, it is still not well understood when gene expression is temporally shifted driving the transition from vegetative to reproductive development, as it is probably prior to the visualization of flowering tissue in the crown. Because there is a 10 day difference between our selected tissue and noting the beginning of morphological transition to a reproductive phase in WT-ns, we expect that we have captured the transcriptome prior to molecular states of reproductive transition. All differentially regulated genes from the RNA-seq experiment at 42 days post-planting are provided in the (APPENDIX).

Shared DEG from 32 and 42 days post-planting stem tissues

Although the RNA-seq experiment assessing stem tissue at 42 days post-planting was determined to be unacceptable as a means to assess transcriptomic differences between HM-*mpg1* and WT-ns at large, we did use it to observe if there were any shared genes that were differentially express between both the 32 and 42 days post-planting timepoints as a potential insight into strongly co-regulated genes present in *mpg1*. There were in fact, 14 up-regulated genes, and 4 down-regulated genes shared between the tissues sampled at both timepoints (Table 4.3).

Table 4.3: Shared DEG from RNA-seq analysis of HM-*mpg1* compared to WT-ns within 32 and 42 days post-planting tiller tissue.

MSU ID	RAP ID	log2 Fold Change (32dpp, 42dpp)	Description
Up-regulated DEG			
LOC_Os08g41030	Os08g0521600	12.102, 9.478	AP2 domain containing protein ***(MPG1)***
LOC_Os01g09220	Os01g0186900	2.177, 1.636	transposon protein, putative, CACTA, En/Spm sub-class, expressed
LOC_Os10g37760	Os10g0521900	1.6, 1.368	OsRhmbd17 - Putative Rhomboid homologue, expressed
LOC_Os12g36110	Os12g0547600	1.38, 1.194	calmodulin binding protein, putative, expressed
LOC_Os09g17560	Os09g0344500	1.337, 2.973	O-methyltransferase, putative, expressed
LOC_Os08g10150	Os08g0201700	1.335, 1.682	SHR5-receptor-like kinase, putative, expressed
LOC_Os06g20920	Os06g0314600	1.167, 2.194	SAM dependent carboxyl methyltransferase, putative, expressed

LOC_Os11g41710	Os11g0635500	1.13, 1.688	cytochrome P450, putative, expressed
LOC_Os05g05680	Os05g0149400	1.035, 1.189	1-aminocyclopropane-1-carboxylate oxidase, putative, expressed
LOC_Os07g31720	Os07g0500300	0.961, 1.054	GTPase activating protein, putative, expressed
LOC_Os04g27670	Os04g0344100	0.945, 1.511	terpene synthase family, metal binding domain containing protein, expressed
LOC_Os12g02320	Os12g0115100	0.891, 1.44	LTPL12 - Protease inhibitor/seed storage/LTP family protein precursor, expressed
LOC_Os07g45570	Os07g0650600	0.818, 1.156	expressed protein
LOC_Os04g28620	Os04g0354600	0.817, 2.634	male sterility protein, putative, expressed
Down-regulated DEG			
LOC_Os03g54160	Os03g0752800	-5.988, -1.956	OsMADS14 - MADS-box family gene with MIKCc type-box, expressed
LOC_Os03g41060	Os03g0607200	-1.365, -1.446	GASR2 - Gibberellin-regulated GASA/GAST/Snakin family protein precursor, putative, expressed
LOC_Os08g44360	Os08g0557800	-1.137, -1.653	male sterility protein 2, putative, expressed
LOC_Os06g43600	Os06g0643500	-0.953, -2.04	LTPL129 - Protease inhibitor/seed storage/LTP family protein precursor, expressed

Of this list of co-regulated genes, *MPG1* (*Os08g41030*) remained the highest differentially up-regulated gene in HM-*mpg1* compared to WT-ns. The other co-up-regulated genes include a variety of genes pertaining to various functions. (*Os01g09220*), codes for a transposon protein/putative nuclease HARBI1, found to be highly expressed in pith parenchyma (Nakano et al., 2013) and in response to glutamate (Kan et al., 2017). *OsRhmbd17* (*Os10g37760*), codes for a rhomboid-like protease found to be expressed in response to glutamate (Kan et al., 2017), and highly in cold-tolerant cultivars (da Maia et al., 2017). Glutamate is an active amino acid involved in protein synthesis and possibly signaling (Kan et al., 2017). *CaMBP* (*Os12g36110*), codes for a calmodulin binding protein, which is a target of OsWRKY47 a positive regulator of water deficit stress (Raineri et al., 2015), as well as observed in panicle tissue during heat stress (Zhang et al., 2012a). (*Os09g17560*), codes for an o-methyltransferase, suggested to be associated with drought tolerance (Silveira et al., 2015). (*Os08g10150*), codes for a SHR5-receptor-like kinase, found to be up-regulated in rice overexpressing *OsNAC9,* which results in drought resistance and altered root architecture (Redillas et al., 2012). (*Os06g20920*), codes for a SAM dependent carboxyl methyltransferase,

shown to be up-regulated in disease resistance varieties (Jain et al., 2017), and is responsible for converting SA into methyl salicylate which acts as a signal for pathogen attack (Shulaev et al., 1997). (*Os11g41710*), codes for a cytochrome P450, shown to be a drought responsive gene (Xia et al., 2019). *OsACO5* (*Os05g05680*), codes for a 1-aminocyclopropane-1-carboxylate oxidase, involved in ethylene biosynthesis (Hamilton et al., 1991, Spanu et al., 1991). (*Os07g31720*), codes for a GTPase activating protein, likely involved with signaling (Ma, 1994). (*Os04g27670*), codes for a terpene synthase family, metal binding domain containing protein, involved with defense (Espinas, 2018). *LTPL12/OsLTP1* (*Os12g02320*), codes for a protease inhibitor/seed storage/LTP family protein (Zi et al., 2013). *BLE2* (*Os07g45570*), codes for an expressed protein, a brassinosteroid signaling gene involved in root development (Krishna et al., 2017, Qin et al., 2018). (*Os04g28620*), codes for a male sterility protein homologous to *Arabidopsis MALE STERIL2*, which is an acyl CoA reductase that has been found to play a role in anther and microspore development (Chen et al., 2011, Shi et al., 2011), and fatty acid synthesis (Mao et al., 2012). Four of these genes contain at least one GCC-box motif in their promoter region (-2000 from TSS) - *MPG1 (Os08g41030)*, (*Os08g10150*), (*Os11g41710*), and *LTPL12/OsLTP1* (*Os12g02320*).

Of the 4 co-down-regulated genes observed: *OsMADS14* (*Os03g54160*), encodes a transcription factor involved in flowering activation (Jeon et al., 2000, Kim et al., 2007). *GASR2* (*Os03g41060*), codes for a transcription factor involved in photosynthetic carbon metabolism (Ambavaram et al., 2014, Satoh et al., 2008), also found in high levels in apical meristematic tissue proposed to play a role in cell proliferation and panicle development (Furukawa et al., 2006). (*Os08g44360*), encodes a male sterility protein 2, a protein involved with anther development and cell cycle regulation (Hobo et al., 2008, Huang et al., 2009a). *LTPL129*

(*Os06g43600*), codes for a protease inhibitor/seed storage/LTP family protein precursor, that is not well characterized. All of these genes except for male sterility protein 2, (*Os08g44360*), contain at least one GCC-box motif in their promoter region (-2000 from TSS).

Some of these genes might be directly influencing a degree of the phenotype observed in *mpg1*. *MADS* genes regulate transition to reproductive phases (Jeon et al., 2000, Kim et al., 2007) and *GASR2* acts as a mechanism to repress starch phosphorylase and GDSL-like lipases responsible for phosphorylation of starch intermediates and accumulation of starch (Ambavaram et al., 2014, Satoh et al., 2008), potentially influencing carbon metabolism in *mpg1*. *CaMBP* (*Os12g36110*), (*Os09g17560*), (*Os08g10150*), and (*Os11g41710*) have all been associated with abiotic stress response, and are potentially influencing the stress-like responses seen in *mpg1* (Raineri et al., 2015), Redillas et al., 2012, Silveira et al., 2015, H. Xia et al., 2019, (Zhang et al., 2012a).

DEG from 32 days post-planting stem tissue are involved in transcriptional regulation

There are several phenotypic characteristics observed in *mpg1*. This is not surprising as the candidate gene predicted to be the cause of the phenotype visualized in the mutant is a transcription factor. Ectopic expression of this gene likely results in a multitude of downstream changes in gene expression, potentially influencing numerous additional genes, including other transcription factors. It will remain unclear which of the differentially expressed downstream genes are directly or indirectly affected by the overexpression of *MPG1* until direct target sites are more thoroughly evaluated. Transcriptome assessment has however revealed several genes involving transcriptional regulation to be differentially expressed in *mpg1*. The altered expression of a single transcription factor can lead to a domino effect resulting in cascading

differences in gene expression. Many transcription factor families have been studied, and have provided further evidence to the nature of their expression and resultant functionality.

Using (RiceNETDB) (Liu et al., 2013a) gene ontology enrichment with our list of DEG, showed enrichment for the category GO:0003700 transcription factor activity. Specifically there were twelve up-regulated genes, and sixteen down-regulated genes in HM-*mpg1* compared to WT-ns (Table 4.4). This could potentially provide insight and rationale to individual characteristics observed in the *mpg1* pleiotropic phenotype.

Table 4.4: DEG list from RNA-seq analysis of HM-*mpg1* compared to WT-ns within 32 days post-planting tiller tissue pertaining to the gene ontology category GO:00003700 – transcription factor activity.

MSU ID	RAP ID	log2 Fold Change	Description
Up-regulated DEG			
LOC_Os08g41030	Os08g0521600	12.102	AP2 domain containing protein, expressed
LOC_Os01g60600	Os01g0821300	2.828	WRKY108, expressed
LOC_Os04g23550	Os04g0301500	2.828	basic helix-loop-helix family protein, putative, expressed
LOC_Os03g53020	Os03g0741100	2.774	helix-loop-helix DNA-binding domain containing protein, expressed
LOC_Os01g64360	Os01g0863300	2.774	MYB family transcription factor, putative, expressed
LOC_Os06g03670	Os06g0127100	2.409	dehydration-responsive element-binding protein, putative, expressed
LOC_Os02g08440	Os02g0181300	2.341	WRKY71, expressed
LOC_Os02g45420	Os02g0676800	2.31	AP2 domain containing protein, expressed
LOC_Os11g02520	Os11g0117400	2.051	WRKY104, expressed
LOC_Os11g45740	Os11g0684000	2.042	MYB family transcription factor, putative, expressed
LOC_Os05g27730	Os05g0343400	1.495	WRKY53, expressed
LOC_Os07g36170	Os07g0545800	0.887	chitin-inducible gibberellin-responsive protein, putative, expressed
Down-regulated DEG			
LOC_Os03g54160	Os03g0752800	-5.988	OsMADS14 - MADS-box family gene with MIKCc type-box, expressed
LOC_Os03g48490	Os03g0691500	-2.121	centromere protein, putative, expressed
LOC_Os06g03710	Os06g0127800	-1.978	DELLA protein SLR1, putative, expressed
LOC_Os08g01100	Os08g0101100	-1.882	HMG1/2, putative, expressed
LOC_Os12g13570	Os12g0238000	-1.712	MYB family transcription factor, putative, expressed
LOC_Os07g41370	Os07g0605200	-1.707	OsMADS18 - MADS-box family gene with MIKCc type-box, expressed
LOC_Os10g42490	Os10g0575600	-1.662	homeobox and START domains containing protein, putative, expressed

LOC_Os02g15810	Os02g0258200	-1.631	HMG1/2, putative, expressed
LOC_Os04g47890	Os04g0566600	-1.62	MYB family transcription factor, putative, expressed
LOC_Os12g41230	Os12g0605500	-1.603	tesmin/TSO1-like CXC domain containing protein, expressed
LOC_Os01g12860	Os01g0229000	-1.376	MYB family transcription factor, putative, expressed
LOC_Os01g11550	Os01g0213800	-1.303	TCP family transcription factor, putative, expressed
LOC_Os11g32100	Os11g0523700	-1.175	inducer of CBF expression 1, putative, expressed
LOC_Os09g37910	Os09g0551600	-1.136	HMG1/2, putative, expressed
LOC_Os05g44400	Os05g0520300	-1.079	GATA zinc finger domain containing protein, expressed
LOC_Os10g40810	Os10g0557600	-0.878	GATA zinc finger domain containing protein, expressed

Of the up-regulated genes, there were three AP2 transcription factors (one being *MPG1*), four WRKY family transcription factors, two bHLH transcription factors, two MYB transcription factors, and one CIGR protein. Most of these genes showed at least 2 log2-fold change different in HM-*mpg1* compared to WT-ns. A greater look into previous assessments of these genes gave an insight to their potential function.

Of the AP2 domain containing proteins, *MPG1* (*Os08g41030*) has been shown to be associated with endosperm development and grain filling (Xu et al., 2016a) and response to stress (Sharoni et al., 2011). AP2 domain containing protein *OsDREB1E/OsERF20* (*Os02g45420*) has been shown to be up-regulated in host-pathogen interactions within rice-blast resistant lines (Jain et al., 2017), and water deficit (Auler et al., 2019). *OsDREB1C* (*Os06g03670*), has shown to be associated with response water deficit and drought response (Wang et al., 2011a, Auler et al., 2019).

Of the WRKY family transcription factors, *WRKY108* (*Os01g60600*), has been shown to be potentially involved in rice morphology and cytokinin metabolism (Hirose et al., 2007), host-pathogen interactions in rice-blast resistant lines (Jain et al., 2017), and drought tolerance (Sinha et al., 2017). *WRKY71* (*Os02g08440*), has been shown to be involved in rice defense response (Liu et al., 2007b), and cold-stress tolerance (Kim et al., 2016), and GA-related interactions in

plant height (Chen et al., 2015). *WRKY104/WRKY89* (*Os11g02520*) has induced expression under heat stress response (Zhang et al., 2012a), methyl jasmonate, and UV-B radiation (Wang et al., 2007).

WRKY53 (*Os05g27730*), has been shown to be involved in oxidative responses to biotic and abiotic stressors (Van Eck et al., 2014).

Of the helix-loop-helix family transcription factors *RERJ1* (*Os04g23550*), has been shown to be involved with JA response (Kiribuchi et al., 2004), and wounding and drought stress response (Kiribuchi et al., 2005). *OsbHLH148* (*Os03g53020*), has been shown to be associated with JA signaling and drought tolerance (Seo et al., 2011).

Of the MYB family transcription factors, (*Os01g64360*) may play a role in salt stress at the seedling stage (Kong et al., 2019), and anoxic stress (Mohanty et al., 2012). *OsJAmyb* (*Os11g45740*), has been shown to be involved in high-osmotic stress (Yokotani et al., 2013), and biotic stress response (Cao et al., 2015).

CIGR1 (*Os07g36170*), might play a role in response to elicitor-induced defense response (Day et al., 2003).

All but *RERJ1* (*Os04g23550*), *WRKY71* (*Os02g08440*), and *WRKY104* (*Os11g02520*) contain at least 1 GCC-box motif in the predicted promoter region (-2000 from TSS). In particular two have a higher number of motifs, (*Os01g64360*), contains 6 GCC-box motifs, and (*Os06g03670*) contains 5 GCC-box motifs.

Of the down-regulated genes, there were two MADS-box transcription factors, three high mobility group (HMG) chromosomal proteins, three MYB transcription factors, two GATA zinc finger domain containing proteins, a centromere related, a DELLA-SLR1-like GRAS, a homeobox and START domain containing, a tesmin/TSO1-like CXC domain containing, a TCP

family transcription factor, and a bHLH transcription factor proteins.

The MADS-box transcription factor family is classically involved in floral development. *OsMADS14* (*Os03g54160*) is involved in flowering activation (Jeon et al., 2000, Kim et al., 2007). *OsMADS18* (*Os07g41370*), is involved with specifying floral determinacy and organ identity (Fornara et al., 2004), seed maturation and germination, response to ABA signaling, IAA- and SL-regulated growth (Yin et al., 2019).

HMG chromosomal protein (*Os08g01100*) expression was found to correlate with abiotic stress-related genes (Yang et al., 2015b). (*Os02g15810*), expression was found to correlate with brown planthopper resistance (Wang et al., 2012), and developing zygotes (Abiko et al., 2013). (*Os09g37910*), was also down-regulated but has not been well characterized.

Of the MYB transcription factors, (*Os12g13570*), has been shown to be suppressed in developing seeds under heat stress (Chen et al., 2016a). (*Os04g47890*), remains uncharacterized however shares a high level of homology to *OsMPH1* (*Os06g45890*), which is involved with improved plant height and grain yield (Zhang et al., 2017c). (*Os01g12860*), has been shown to be suppressed in developing seeds under heat stress (Chen et al., 2016a), and expression correlates with stress response (Smita et al., 2015).

The GATA zinc-finger domain containing protein's (*Os05g44400*) expression is down regulated with *S. hermonthica* infected rice roots (Mutuku et al., 2015). *OsGATA7* (*Os10g40810*), is involved in brassinosteroid-mediated growth and regulation affecting architecture and grain shape (Zhang et al., 2018).

The centromere related protein (*Os03g48490*), has not been well characterized.

The DELLA-SLR1-like GRAS*, DLT/OsGRAS-32* (*Os06g03710*), has shown to be involved in BR signal regulation and GA metabolism (Li et al., 2010b) influencing height and

tillers (Tong et al., 2009), and negative regulation of grain size (Sun et al., 2013).

The homeobox and START domain containing protein, *ROC3* (*Os10g42490*), has been shown to be negatively regulated under nitrogen starvation (Yang et al., 2015a).

The tesmin/TSO1-like CXC domain containing protein (*Os12g41230*) gene expression has been shown to be co-expressed with genes pertaining to the cell cycle (Yu et al., 2017).

The TCP family transcription factors have historically played roles in cell proliferation in vegetative and reproductive structures and morphology, *OsPCF5* (*Os01g11550*), has been shown to play a role in greater sized leaf morphogenesis and improved cold tolerance when silenced (Yang et al., 2013).

The bHLH transcription factor *OsICE1* (*Os11g32100*) is involved with cold tolerance and acclimation (Nakamura et al., 2011, Deng et al., 2017) and stomata development (Wu et al., 2019).

Of the down-regulated genes, all have a GCC-box within their promoter regions (-2000 from TSS) except *ROC3* (*Os10g42490*), (*Os02g15810*), (*Os04g47890*), *OsPCF5* (*Os01g11550*), *OsICE1* (*Os11g32100*), (*Os05g44400*), and *OsGATA7* (*Os10g40810*). From the genes containing a GCC-box motif, several had a high number of these elements, *OsMADS18* (*Os07g41370*) containing 5, (*Os12g13570*) containing 6, and (*Os12g41230*) containing 7.

DEG from 32 days post-planting stem tissue are involved in flower development

Using (RiceNETDB) (Liu et al., 2013a) gene ontology enrichment, there were a number of differentially regulated genes within the category GO:0009908 flower development. There were zero up-regulated genes, and eight down-regulated in HM-*mpg1* compared to WT-ns (Table 4.5). This matched well with the phenotype of delayed flowering observed in *mpg1*.

Table 4.5: DEG list from RNA-seq analysis of HM-*mpg1* compared to WT-ns within 32 days post-planting tiller tissue pertaining to the gene ontology category GO:0009908 – flower development.

MSU ID	RAP ID	log2 Fold Change	Description
Up-regulated DEG			
-	-	-	-
Down-regulated DEG			
LOC_Os03g54160	Os03g0752800	-5.988	OsMADS14 - MADS-box family gene with MIKCc type-box, expressed
LOC_Os01g64820	Os01g0868300	-1.731	POLA1 - Putative DNA polymerase alpha catalytic subunit, expressed
LOC_Os10g28230	Os10g0418000	-1.712	Core histone H2A/H2B/H3/H4 domain containing protein, putative, expressed
LOC_Os07g41370	Os07g0605200	-1.707	OsMADS18 - MADS-box family gene with MIKCc type-box, expressed
LOC_Os08g32600	Os08g0421800	-1.645	STE_MEKK_ste11_MAP3K.21 - STE kinases include homologs to sterile 7, sterile 11 and sterile 20 from yeast, expressed
LOC_Os03g06670	Os03g0162200	-1.587	Core histone H2A/H2B/H3/H4 domain containing protein, putative, expressed
LOC_Os07g08500	Os07g0182900	-1.58	C-5 cytosine-specific DNA methylase, putative, expressed
LOC_Os08g42600	Os08g0538700	-1.579	retinoblastoma-related protein-like, putative, expressed

The genes down-regulated from this category included genes coding for two MADS transcription factors, a POLAI, two core histone domain containing proteins, a DNA methylase, a STE kinase, and a retinoblastoma-related protein-like proteins.

Of the MADS-box transcription factors, *OsMADS14* (*Os03g54160*) is involved in flowering activation (Jeon et al., 2000, Kim et al., 2007). *OsMADS18* (*Os07g41370*), is involved with specifying floral determinacy and organ identity (Fornara et al., 2004), seed maturation and germination, response to ABA signaling, IAA- and SL-regulated growth (Yin et al., 2019).

OsPOLAI (*Os01g64820*), is a gene coding for a putative DNA polymerase alpha catalytic subunit, which is involved with DNA replication, and found to be necessary for gamete formation in *Arabidopsis* (Barrero et al., 2007).

Histone modifications have been shown to promote flowering in rice by targeting sites relevant to genes responsible for reproductive development (Sun et al., 2012, Shi et al., 2014). Two core histone domain containing proteins were found in this category (*Os10g28230* and *Os03g06670*).

The DNA methytransferase, *OsMET1b/OsMET1-2* (*Os07g08500*), has been found in high expression in the SAM during panicle initiation and floral organ initiation, suggesting a role in flowering (Sharma et al., 2009).

The STE kinase, *OsMAPKKK15* (*Os08g32600*), also found in this category, belongs to a large family of kinases usually involved in signaling for plant growth and development (Rao et al., 2010) and the cell cycle (Takahashi et al., 2011).

The retinoblastoma-related protein-like found in this category, *OsRBR1* (*Os08g42600*), has been shown to play a role in cell division cycle regulation (Abraham et al., 2015).

All of these genes possess at least one GCC-box within their promoter regions (-2000 from the TSS) except *OsPOLA1* (*Os01g64820*), *OsMAPKKK15* (*Os08g32600*), and *OsMET1b/OsMET1-2* (*Os07g08500*). From the genes containing a GCC-box motif, one had a higher number of these elements, *OsMADS18* (*Os07g41370*) containing 5.

DEG from 32 days post-planting stem tissue are involved in stress response

Using (RiceNETDB) (Liu et al., 2013a) gene ontology enrichment, genes pertaining to the category GO:0006950 response to stress were assessed. In HM-*mpg1* thirty genes were found up-regulated in this category while none were found to be enriched from the down-regulated genes in *mpg1* (Table 4.6). This potentially correlates well with the observations of the possible stress response phenotype observed in *mpg1*.

Table 4.6: DEG list from RNA-seq analysis of HM-*mpg1* compared to WT-ns within 32 days post-planting tiller tissue pertaining to the gene ontology category GO:0006950 – response to stress.

MSU ID	RAP ID	log2 Fold Change	Description
Up-regulated DEG			
LOC_Os11g05380	Os11g0151400	4.254	cytochrome P450, putative, expressed
LOC_Os02g11070	Os02g0205500	2.632	3-ketoacyl-CoA synthase, putative, expressed
LOC_Os03g05880	Os03g0153500	2.622	monooxygenase, putative, expressed

LOC_Os10g25290	Os10g0392400	2.357	ZIM domain containing protein, putative, expressed
LOC_Os01g18120	Os01g0283700	2.35	cinnamoyl CoA reductase, putative, expressed
LOC_Os06g03670	Os06g0127100	2.341	dehydration-responsive element-binding protein, putative, expressed
LOC_Os02g08440	Os02g0181300	2.31	WRKY71, expressed
LOC_Os02g08270	Os02g0179200	2.074	class I glutamine amidotransferase, putative, expressed
LOC_Os02g45420	Os02g0676800	2.051	AP2 domain containing protein, expressed
LOC_Os03g08330	Os03g0181100	2.022	ZIM domain containing protein, putative, expressed
LOC_Os04g41960	Os04g0497000	1.895	NADP-dependent oxidoreductase, putative, expressed
LOC_Os03g08320	Os03g0180900	1.818	ZIM domain containing protein, putative, expressed
LOC_Os07g34260	Os07g0526400	1.648	chalcone and stilbene synthases, putative, expressed
LOC_Os03g17700	Os03g0285800	1.587	CGMC_MAPKCGMC_2_ERK.2 - CGMC includes CDA, MAPK, GSK3, and CLKC kinases, expressed
LOC_Os11g45740	Os11g0684000	1.495	MYB family transcription factor, putative, expressed
LOC_Os05g27730	Os05g0343400	1.486	WRKY53, expressed
LOC_Os02g48770	Os02g0719600	1.417	SAM dependent carboxyl methyltransferase, putative, expressed
LOC_Os03g03700	Os03g0129100	1.396	MLO domain containing protein, putative, expressed
LOC_Os12g26290	Os12g0448900	1.367	alpha-DOX2, putative, expressed
LOC_Os12g17430	None	1.325	NBS-LRR disease resistance protein, putative, expressed
LOC_Os04g52440	Os04g0614500	1.269	aminotransferase, putative, expressed
LOC_Os03g28940	Os03g0402800	1.134	ZIM domain containing protein, putative, expressed
LOC_Os11g41710	Os11g0635500	1.13	cytochrome P450, putative, expressed
LOC_Os03g58290	Os03g0797300	1.082	indole-3-glycerol phosphate lyase, chloroplast precursor, putative, expressed
LOC_Os01g22900	Os01g0332100	1.007	neutral/alkaline invertase, putative, expressed
LOC_Os10g35460	Os10g0497700	0.972	COBRA, putative, expressed
LOC_Os03g13840	Os03g0241900	0.933	senescence-associated protein, putative, expressed
LOC_Os03g03510	Os03g0126800	0.854	CAMK_KIN1/SNF1/Nim1_like.15 - CAMK includes calcium/calmodulin depedent protein kinases, expressed
LOC_Os04g28620	Os04g0354600	0.817	male sterility protein, putative, expressed
LOC_Os05g33630	Os05g0406100	0.754	inosine-uridine preferring nucleoside hydrolase family protein, putative, expressed
Down-regulated DEG			
-	-	-	-

The genes up-regulated in *mpg1* that pertain to this gene ontology category consist of genes coding for: two cytochrome P450 proteins, a 3-ketoacyl-CoA synthase protein, a monooxygenase protein, four TIFY ZIM domain containing transcription factor proteins, a cinnamoyl CoA reductase protein, two AP2/EREBP transcription factor proteins, two WRKY

transcription factor proteins, a class I glutamine amidotransferase protein, a NADP-dependent oxidoreductase protein, a chalcone and stilbene synthase protein, two kinase proteins, a MYB transcription factor protein, a SAM dependent carboxyl methyltransferase, a MLO domain containing protein, an alpha-DOX2 protein, a NBS-LLR disease resistance protein, an aminotransferase protein, an indole-3-glycerol phosphate lyase protein, a neutral/alkaline invertase protein, a COBRA protein, a senescence-associated protein, a male sterility protein, and an inosine-uridine preferring nucleoside hydrolase family protein.

Ten of these genes have at least a 2.0 log2 fold change in HM-*mpg1* compared to WT-ns. Of the genes from this list, the gene with greatest differential expression is (*Os11g05380*), which encodes for cytochrome P450, although not well characterized this gene class is involved in a variety of functions associated with metabolism, defense, and response to stress (Mizutani, 2012). The other cytochrome P450, (*Os11g41710*), has been shown to play a role in drought response (Xia et al., 2019).

(*Os02g11070*) encodes for a 3-ketoacyl-CoA synthase protein, a lipid metabolism protein that has been shown to be responsive to salt stress (Yuenyong et al., 2018).

AUMO1 (*Os03g05880*), encodes for a monooxygenase, that is a target of OsABF1 playing a role in drought tolerance (Zhang et al., 2017a).

Of the ZIM domain containing proteins, *OsJAZ12* (*Os10g25290*), has been shown to play a role in stress and JA response (Ye et al., 2009), and interact with *OsbHLH148*, which when overexpressed gives rise to drought tolerance (Seo et al., 2011). *OsJAZ10* (*Os03g08330*), is co-expressed in drought resistant lines overexpressing transcription factor *OsDRAP1* (Huang et al., 2018), and is co-expressed in blight resistant rice overexpressing *OsMYC2* (Uji et al., 2016). *OsJAZ11* (*Os03g08320*), has been shown to be down-regulated in *pi21* RNAi lines; *pi21*

functional lines confer partial resistance to *Magnaporthe oryzae* (Zhang et al., 2016). Additionally this gene is induced under overexpression of *OsbHLH148* that results in drought tolerance (Seo et al., 2011). And, *OsJAZ6* (*Os03g28940*), is shown to be up-regulated under *Xanthomonas oryzae* pv. *oryzae* infection (Wang et al., 2019).

(*Os01g18120*), encodes for a cinnamoyl-CoA reductase classically involved in lignin biosynthesis, has been found to be highly expressed in pith parenchyma cells of wild-type plants and not in F71 dwarf mutant pith parenchyma (Nakano et al., 2013). Additionally, greater expression through induction of this class of genes has been visualized under exposure to abiotic stressors (Srivastava et al., 2015).

From the AP2/EREBPs, expression of *OsDREB1C* (*Os06g03670*) and *OsERF20* (*Os02g45420*), have been shown to be associated with drought stress response (Sharoni et al., 2012)). Additionally, *OsERF20/OsDREB1E*, has been shown to be up-regulated in host-pathogen interactions in rice-blast resistant lines (Jain et al., 2017), and in response to phosphate starvation (Yu et al., 2018).

Of the WRKY transcription factors, *OsWRKY71* (*Os02g08440*), has been shown to be involved in rice defense response (Liu et al., 2007b), cold-stress tolerance (Kim et al., 2016), and GA-related interactions in plant height (Chen et al., 2015). *OsWRKY53* (*Os05g27730*), has been shown to be involved in oxidative responses to biotic and abiotic stressors (Van Eck et al., 2014).

(*Os02g08270*), encodes a class I glutamine amidotransferase protein, is not well characterized however this class of proteins have been shown to be associated with glutamate metabolism and plant defense (Seifi et al., 2013).

(*Os04g41960*), encodes a NADP-dependent oxidoreductase protein, which has been shown to be expressed in rice roots under exposure to exogenous glutamate (Kan et al., 2017).

OsPKS15 (*Os07g34260*), encodes a chalcone and stilbene synthase protein, whose gene family has been associated with stress response through the SA defense pathway resulting in accumulation of flavonoid and isoflavonoid phytoalexins (Dao et al., 2011). Additionally, this gene has been shown to be up-regulated under *Xanthomonas oryzae* pv. *oryzae* infection (Yu et al., 2015).

Of the kinase proteins, *OsMSRMK2/OsMAP1/OsMPK3/OsMPK5/OsMAPK2/OsMAPK5/OsBIMK1* (*Os03g17700*), encodes a mitogen-activated kinase that has been shown to be expressed and potentially play a role in numerous defense/stress response pathways including mechanical wounding, biotic, and abiotic stressors (Agrawal et al., 2002). One specific function of this gene is shown to play a pivotal role in submergence tolerance via interaction with *OsSUB1A1* (Singh and Sinha, 2016). Overexpression of this gene enhances low phosphate tolerance in both rice and *Arabidopsis* (Hur and Kim, 2014). *OsCIPK9* (*Os03g03510*), encodes a calcineurin B-like-interacting protein kinase which has been shown to be induced under various stimuli including drought stress, salt stress, PEG exposure, and ABA exposure (Xiang et al., 2007).

OsJAmyb (*Os11g45740*), encodes for a MYB transcription factor that has been shown to play a role in abiotic stress response (Yokotani et al., 2013). Overexpression of this gene confers blast resistance (Cao et al., 2015).

(*Os02g48770*), encodes a SAM dependent carboxyl methyltransferase, that acts as a jasmonate O-methyltransferase, that functions in the jasmonic acid synthesis and signaling pathway; this gene has been found to be highly expressed in *cea62* (a jasmonic acid overproduction mutant of rice), which is resistant to bacterial blight (Liu, et al. 2012).

OsMLO3 (*Os03g03700*), encodes a MILDEW LOCUS O (MLO) transcription factor plays a role in powdery mildew fungus defense response (Devoto et al., 2003, Win et al., 2018). Additionally, high levels of induced expression of this gene are visible in rice *bbr1* mutant that confers *Xanthomonas oryzae* pv. *oryzae* resistance (Yi et al., 2013).

PIOX (*Os12g26290*), encodes an alpha-DOX2 protein, expression has been found under alkaline stress-tolerant variety WD20340 (Li et al., 2018). Induction of expression for this gene is also seen during response to *Nilaparvata lugens* (Brown Planthopper) infestation (Wei et al., 2009).

(*Os12g17430*), encodes a NBS-LRR disease resistance protein, has not been well characterized, however genes from this family are involved as R-genes against herbivorous insects (Broekgaarden et al., 2011), and disease resistance (Bozkurt et al., 2007).

(*Os04g52440*) encodes an aminotransferase protein, which has been shown be highly induced during Asian rice gall midge infestation in resistance variety RP2068-18-3-5 (Agarrwal et al., 2016). Additionally, this gene has been potentially linked to source-sink and yield traits in rice (Xu et al., 2015).

(*Os03g58290*), encodes an indol-3-glycerol phosphate lyase protein, which has been shown to have induced expression under *Xanthomonas oryzae* pv. *oryzae* infection (Yu et al., 2015).

OsNIN2 (*Os01g22900*) encodes a neutral/alkaline invertase, is notably expressed in plants overexpressing *OsCaM1-1*, that exhibit salt stress tolerance (Yuenyong et al., 2018). Additionally, this gene is associated with carbon metabolism, and has shown enhanced expression during response to drought (Reguera et al., 2013), and during mechanical removal of the cell wall as a potential means of defense response (Sharma et al., 2011).

(*Os10g35460*), encodes a COBRA-like protein which has been shown to be expressed during glutamate-response in roots (Kan et al., 2017). COBRA family genes produce GPI-anchored proteins involved in cellulose synthesis (Maleki et al., 2016, Dai et al., 2011), as well as even oriented cell expansion in *Arabidopsis* (Doerks et al., 2002).

(*Os03g13840*), encodes a senescence-associated protein, which has been shown to be expressed during heat stress response (James et al., 2015).

OsCER4 (*Os04g28620*), encodes a male sterility protein, involved in fatty acid synthesis and cuticular wax formation (Mao et al., 2012).

(*Os05g33630*), encodes an inosine-uridine proffering nucleoside hydrolase, which has been shown to be up-regulated in *OsPsbS* knockout rice plants that present enhanced resistance to pathogens (Zulfugarov et al., 2016).

All but (*Os11g05380*), (*Os02g11070*), *AUMO1* (*Os03g05880*), *OsWRKY71* (*Os02g08440*), *OsJAZ10* (*Os03g08330*), *OsPKS15* (*Os07g34260*), *PIOX* (*Os12g26290*), (*Os03g58290*), *OsCER4* (*Os04g28620*) contain at least 1 GCC-box motif in the promoter region (-2000 from TSS). In particular three have a higher number of motifs, (*Os04g52440*) containing 10, (*Os04g41960*) containing 8, and *OsDREB1C* (*Os06g03670*) containing 5 motifs.

DEG from 32 days post-planting stem tissue are similar to genes that respond to defoliation and possibly biomass accumulation

Ratooning, is the practice of cutting back culm tissue of developed monocots in order to stimulate the generation of new panicle bearing tillers. This is a method that can be used to produce a secondary harvest for some monocots. *mpg1* presented with an enhanced ratooning ability by having a greater survivability rate, generating greater amount of biomass, flowering earlier, and generating greater seed yield compared to wild-type plants post ratooning. The

molecular mechanisms involved in ratooning remain unknown. To potentially identify genes involved in this phenotype we compared the differentially regulated genes in *mpg1* with an expression analysis of defoliated rice seedlings (Chen et al., 2009). Although this is not a perfect metric for comparison, it might provide insight into genes that play roles in tissue re-generation post-wounding (cutting/ratooning), or even genes involved with growth that could explain the general overall increase in biomass observed in *mpg1*.

(Chen et al., 2009) examined defoliation by assessing gene expression through microarray analysis on rice seedlings at least 18 cm in height (3- to 4- leaf stage), where two-thirds of their stem tissue length was mechanically removed and remaining tissue was sampled at 2, 6, and 24 h post-removal. Their results showed that defoliation resulted in altered expression of 466 genes. differentially regulated genes pertained to categories within carbohydrate metabolism, biosynthesis of secondary metabolism, amino acid metabolism, lipid metabolism, membrane transport, signal transduction, and cell growth and death (Chen et al., 2009).

A number of genes found in defoliation response were also found to be similarly co-regulated in *mpg1*. There were 26 up-regulated differentially expressed genes in *mpg1* that were also found to be up-regulated in response to defoliation in seedlings. Additionally, two genes from the *mpg1* DEG were inversely expressed compared to the gene expression found in the defoliated rice (Table 4.7).

It could be interesting to evaluate the similarities found in this study to our own to potentially understand genes that might play a role in tissue regeneration and biomass accumulation in *mpg1*.

Table 4.7: DEG list from RNA-seq analysis of HM-*mpg1* compared to WT-ns within 32 days post-planting tiller tissue that are similarly co-expressed under rice defoliation (Chen et al., 2009).

MSU ID	RAP ID	log2 Fold Change	Description

Up-regulated DEG			
LOC_Os12g14440	Os12g0247700	2.935	Jacalin-like lectin domain containing protein, putative, expressed
LOC_Os02g11070	Os02g0205500	2.632	3-ketoacyl-CoA synthase, putative, expressed
LOC_Os03g08520	Os03g0183500	2.436	DUF581 domain containing protein, expressed
LOC_Os10g25290	Os10g0392400	2.357	ZIM domain containing protein, putative, expressed
LOC_Os02g08440	Os02g0181300	2.31	WRKY71, expressed
LOC_Os05g44060	Os05g0516700	2.151	expressed protein
LOC_Os11g02520	Os11g0117400	2.042	WRKY104, expressed
LOC_Os04g41960	Os04g0497000	1.895	NADP-dependent oxidoreductase, putative, expressed
LOC_Os10g02840	Os10g0118000	1.833	O-methyltransferase, putative, expressed
LOC_Os03g08320	Os03g0180900	1.818	ZIM domain containing protein, putative, expressed
LOC_Os07g09190	Os07g0190000	1.554	transketolase, putative, expressed
LOC_Os11g45740	Os11g0684000	1.495	MYB family transcription factor, putative, expressed
LOC_Os02g02930	Os02g0121700	1.482	terpene synthase, putative, expressed
LOC_Os02g48770	Os02g0719600	1.417	SAM dependent carboxyl methyltransferase, putative, expressed
LOC_Os12g36110	Os12g0547600	1.38	calmodulin binding protein, putative, expressed
LOC_Os12g26290	Os12g0448900	1.367	alpha-DOX2, putative, expressed
LOC_Os05g10310	Os05g0191500	1.36	acid phosphatase, putative, expressed
LOC_Os03g19990	Os03g0314500	1.333	WD40-like Beta Propeller Repeat family protein, expressed
LOC_Os04g05650	Os04g0142400	1.196	expressed protein
LOC_Os04g15920	Os04g0229100	1.104	dehydrogenase, putative, expressed
LOC_Os07g41060	Os07g0601900	1.095	dihydroflavonol-4-reductase, putative, expressed
LOC_Os03g06520	Os03g0161200	1.004	sulfate transporter, putative, expressed
LOC_Os09g38130	Os09g0554300	1.000	auxin efflux carrier component, putative, expressed
LOC_Os10g35460	Os10g0497700	0.972	COBRA, putative, expressed
LOC_Os10g02880	Os10g0118200	0.902	O-methyltransferase, putative, expressed
LOC_Os04g10350	Os04g0182200	0.862	1-aminocyclopropane-1-carboxylate oxidase homolog 2, putative, expressed
Down-regulated DEG			
-	-	-	-
Opposite regulation			
Up-regulated that were down-regulated DEG in defoliation			
LOC_Os09g36680	Os09g0537700	1.111	ribonuclease T2 family domain containing protein, expressed
Down-regulated that were up-regulated DEG in defoliation			

LOC_Os04g38570	Os04g0459000	-1.067	multidrug resistance protein, putative, expressed

Using (RiceNETDB) (Liu et al., 2013a) gene ontology enrichment for the differentially regulated genes that shared similar expression between *mpg1* and defoliated seedlings was performed. Categories of enrichment spanned a variety of biological processes, some noteworthy ones including: response to stress, lipid metabolism, and biosynthesis (listed by order of significance). The only enrichment category pertaining to molecular function was catalytic activity.

The genes from the gene ontology enrichment category response to stress (GO: 0006950) encode for a 3-ketoacyl-CoA synthase, two ZIM domain containing proteins, one WRKY transcription factor, one NADP-dependent oxidoreductase, one MYB transcription factor, one SAM dependent carboxyl methyltransferase, one alpha-DOX2 protein, and a COBRA-like protein.

The genes from the gene ontology enrichment category lipid metabolism (GO: 0006629) encode for a transketolase, one terpene synthase, one SAM dependent carboxyl methyltransferase, an alpha-DOX2 protein, and a dihydroflavonol-4-reductase.

The genes from the gene ontology enrichment category biosynthesis (GO: 0009058) encode for two WRKY transcription factors, an O-methyltransferase, a transketolase, a terpene synthase, a SAM dependent carboxyl methyltransferase, a dehydrogenase, and a dihydroflavonol-4-reductase.

Although not enriched for any particular categories using (RiceNETDB), many of the other genes found to be up-regulated in both *mpg1* and defoliation response might directly be linked to biomass accumulation.

OsJAC1 (*Os12g14440*), encodes a jacalin-like lectin domain containing protein which has been shown to be involved in growth and development, and when overexpressed results in decreased coleoptile and stem elongation (Jiang et al., 2007) as well as bacterial and fungal pathogen resistance. Interestingly, this gene is co-up-regulated under rice plants overexpressing a heme activator protein (*OsHAP2E*), that confers abiotic and biotic resistance, increased photosynthesis and tiller number (Alam et al., 2015).

OsaFLZ24/DUF581 (*Os03g08520*), encodes a FCS-like-zinc-finger domain containing protein, this family of genes have been shown to play a role in sugar and energy signaling and plant development (Jamsheer et al., 2015). This gene in particular has been shown to be up-regulated under nitrogen starvation (Hsieh et al., 2018) suggesting it may function in nitrogen efficiency aiding in biomass production (Richard-Molard et al., 2008).

OsJAZ12 (*Os10g25290*), encodes a ZIM domain containing protein, which are classically recognized for their ability to target and repress JA signaling. This gene in particular was shown to have induced expression in plants exposed to MoHrip1, a protein elicitor isolated from *Magnaporthe oryzae* that induced blast-resistance enhanced growth (Lv et al., 2016).

OsWRKY71 (*Os02g08440*), encodes a transcription factor classically noted to play a role in a variety of stress responses, one in particular being induction after nitrogen starvation (Yang et al., 2015a), suggesting it might play a role in nitrogen efficiency.

(*Os05g44060*), is an unknown expressed protein, which has been shown to be up-regulated under exogenous glutamate, and may play a role in synergistic metabolism (Kan et al., 2017).

(*Os04g41960*), encodes a NADP-dependent oxidoreductase, which has also been shown to be induced under exogenous glutamate (Kan et al., 2017).

(*Os10g02840*), encodes an O-methyltransferase protein, which has been shown to be upregulated under nitrogen and phosphorous starvation (Cai et al., 2013).

OsDXS3 (*Os07g09190*), encodes a 1-Deoxy-d:-xylulose 5-phosphate synthase/transketolase, belonging to a family of proteins that function early in the isoprenoid 2C-methyl-D-erythritol 4-phosphase (MEP) synthesis pathway involved in diverse functions including photosynthesis, respiration, growth, cell cycle control, defense, and response to stimulus (Estevez et al., 2001). The MEP pathway is important because it is an upstream process that can be important to generate numerous biomass phytohormones such as gibberellic acid, abscisic acid, and strigolactones, as well as photosynthetically important molecules like plastoquinone and chlorophylls (Cordoba et al., 2009).

(*Os02g02930*), encodes a terpene synthase protein, which has been shown to be upregulated under nitrogen and phosphorus starvation (Cai et al., 2013).

(*Os02g48770*), encodes a SAM dependent carboxyl methyltransferase, which has also been shown to be upregulated under nitrogen and phosphorus starvation (Cai et al., 2013).

(*Os05g10310*), encodes an acid phosphatase, shown to play a role in carbon metabolism (Morita et al., 2015).

RIF3/OsCAD6 (*Os04g15920*), encodes a dehydrogenase protein, which has not only been shown to be upregulated under nitrogen and phosphorous starvation stress (Cai et al., 2013); but members of this gene family play a role in monolignol biosynthesis (Hirano et al., 2012) for lignification and stress response (Park et al., 2018).

(*Os07g41060*), encodes a dihydroflavonol-4-reductase, which has been shown to be induced under potassium starvation, and down-regulated in reintroduction of potassium (Shankar et al., 2013).

(*03g0161200*), encodes a sulfate transporter, which has not been well characterized. Sulfur transporters could play a role in biomass accumulation because sulfur is an essential macronutrient for plant growth (Gigolashvili and Kopriva, 2014) and has been linked to nitrogen and sulfur uptake and remobilization during vegetative growth and (Abdallah et al., 2010).

OsPILS7a (*Os09g38130*), encodes an auxin efflux carrier component, which functions to transport auxin during vegetative developmental phases (Mohanta and Mohanta, 2014). Auxin is an important phytohormone which operates in cell division, and elongation (Mohanta et al., 2018).

(*Os10g35460*), encodes a COBRA-like protein, which has up-regulated expression in *glup4* rice mutants that have increased carbohydrate metabolism (Doroshenk et al., 2010) and to exposure to glutamate (Kan et al., 2017). Additionally this gene family has been linked to cell expansion and cellulose crystallinity (Schindelman et al., 2001, Calderan-Rodrigues et al., 2019).

(*Os10g02880*), encodes an O-methyltransferase, which has been shown to have induced expression associated with root growth and development under root-specific expression of *OsNAC5* (Jeong et al., 2013) and *OsNAC9* (Redillas et al., 2012) giving way to drought tolerance and enhanced grain yield.

(*Os04g10350*), encodes a 1-aminocyclopropane-1-carboxylate oxidase homolog 2 (ACC-oxidase), which plays a role in ethylene biosynthesis (Hamilton et al., 1991, Spanu et al., 1991).

There were two genes that were inversely-regulated between defoliation response and the differentially expressed genes identified in *mpg1.* One that was down-regulated in defoliation response, yet up-regulated in *mpg1* was *OsRNS4/OsRRP* (*Os09g36680*), which encodes a ribonuclease T2 family domain containing protein. This gene has been shown to be expressed in rice stem tissue in wild-type conditions and reduced in dwarf mutant *ext37* (Wei et al., 2006).

This gene is also highly up-regulated in *OsHAP2E* overexpression lines that confer abiotic and biotic stress tolerance as well as an increase in photosynthesis and tiller number (Alam et al., 2015). The one gene that was up-regulated in defoliation response and down-regulated in *mpg1* was *MDR13* (*Os04g38570*), which encodes for a multidrug resistance protein. This gene has been shown to play a role in pathogen resistance pathways (Divya et al., 2018).

All but (*Os02g11070*), *OsaFLZ24/DUF581* (*Os03g08520*), *OsWRKY71* (*Os02g08440*), *OsWRKY104/OsWRKY89* (*Os11g02520*), (*Os10g02840*), *OsDXS3* (*Os07g09190*), (*Os02g02930*), *CaMBP* (*Os12g36110*), *PIOX* (*Os12g26290*), (*Os04g05650*), *RIF3/OsCAD6* (*Os04g15920*), (*Os07g41060*), *OsPILS7a* (*Os09g38130*), (*Os10g02880*), and (*Os04g10350*) contain at least 1 GCC-box motif in theit promoter region (-2000 from TSS). Four genes have a higher number of GCC-box motifs than the others, (*Os04g41960*) contains 8, (*Os02g48770*) contains 4, *OsJAC1* (*Os12g14440*) contains 3, and (*Os10g35460*) contains 3.

DEG from 32 days post-planting stem tissue possess genes involved in cell wall formation

The similarities in the comparative analysis between defoliation and *mpg1*, as well as previous studies suggesting altered biomass and plant morphology, provoked us to investigate genes within additional specific traits, in particular (cell wall biosynthesis, carbohydrate metabolism, and cell growth).

By scanning DEG in the enriched gene ontology term (metabolism) we looked for genes that might be associated with these traits. Indeed, there were several genes that potentially relate to biomass accumulation and/or seed yield.

There were a number of genes that were differentially regulated pertaining to cell wall proteins and cell wall biosynthesis (Table 4.8).

Table 4.8: DEG list from RNA-seq analysis of HM-*mpg1* compared to WT-ns within 32 days post-planting tiller tissue pertaining to cell wall related genes from the gene ontology category GO:0008152 – metabolism.

Cell Wall Function	MSU ID	RAP ID	log2 Fold Change	Description
Suberin biosynthesis	LOC_Os02g11070	Os02g0205500	2.632	3-ketoacyl-CoA synthase, putative, expressed
Lignin biosynthesis	LOC_Os01g18120	Os01g0283700	2.35	cinnamoyl CoA reductase, putative, expressed
	LOC_Os11g42220	Os11g0641800	0.927	laccase precursor protein, putative, expressed
	LOC_Os07g41060	Os07g0601900	1.095	dihydroflavonol-4-reductase, putative, expressed
Carbohydrate metabolism	LOC_Os08g34580	Os08g0445700	1.733	trehalose-6-phosphate synthase, putative, expressed
	LOC_Os01g22900	Os01g0332100	1.007	neutral/alkaline invertase, putative, expressed
Cell Wall Metabolism	LOC_Os04g33640	Os04g0412300	1.705	glycosyl hydrolases family 17, putative, expressed
	LOC_Os03g12140	Os03g0221500	-1.779	glucan endo-1,3-beta-glucosidase precursor, putative, expressed
Cellulose crystallization	LOC_Os10g35460	Os10g0497700	0.972	COBRA, putative, expressed
Cutin deposition	LOC_Os12g02320	Os12g0115100	0.891	LTPL12 - Protease inhibitor/seed storage/LTP family protein precursor, expressed
Cellulose synthesis	LOC_Os12g36890	Os12g0555600	-1.438	CSLD4 - cellulose synthase-like family D, expressed
	LOC_Os06g42020	Os06g0625700	-1.475	CSLA9 - cellulose synthase-like family A, expressed
Saccharification	LOC_Os12g01700	Os12g0107700	-1.72	inactive receptor kinase At2g26730 precursor, putative, expressed
	LOC_Os11g01620	Os11g0107700	-1.449	inactive receptor kinase At2g26730 precursor, putative, expressed

Several genes pertaining to cell wall formation were found to be differentially regulated between WT-ns and HM-*mpg1*. Genes were found through brute force analysis within the GO:0008152 metabolism enriched category from (RiceNETDB). Genes pertaining to the cell wall involved suberin biosynthesis, lignin biosynthesis, cell wall modification, cell wall metabolism, cellulose biosynthesis, cutin deposition, cell wall synthesis, cellulose synthesis, and saccharification.

One DEG, highly up-regulated, influences suberin biosynthesis. (*Os02g11070*) encodes a 3-ketoacyl-CoA synthase a key enzyme in biosynthesis of suberin (Franke et al., 2012). Suberin, along with cutin, and lignins form a complex matrix within higher plant epidermis and periderm cell-well tissues. A closely related 3-ketoacyl-CoA synthase gene in rice, *OsWSL1*, characterized under a loss of function mutant revealed that it affected plant morphology in terms of height and

leaf length, as well as root length and abundance, and panicle length (Yu et al., 2008), suggesting that 3-ketoacy-CoA's might play a role in growth and development processes.

Several DEG involved in lignin biosynthesis were up-regulated in *mpg1*. (*Os01g18120*), encodes a cinnamoyl CoA reductase, an enzyme classically involved in plant lignin biosynthesis (Lacombe et al., 1997). *OsLAC24* (*Os11g42220*), encodes a laccase precursor protein, which could play a role in plant lignification (Dean and Eriksson, 1994). (*Os07g41060*) encodes a dihydroflavinol-4-reductase, which family members have shown to play a role in formation of lignin (Sharma et al., 2011).

Genes involved in cell wall modification/metabolism were also up-regulated in *mpg1*. *OsTPS8* (*Os08g34580*), encodes for a trehalose-6-phosphate-synthase, which family members are involved in sucrose utilization through carbon metabolism, starch synthesis, cell wall structure, and stress response (Sharma et al., 2011, Ponnu et al., 2011, Li et al., 2011, Gomez et al., 2006). *OsTPS8* is a class II type trehalose-phosphate-synthase gene, which has been shown to play a role regulation of suberin deposition via ABA signaling, resulting in salinity tolerance without yield penalty (Vishal et al., 2019). *OsNIN2* (*Os01g22900*), encodes a neutral/alkaline invertase, which has been shown to play a role carbon metabolism (Sharma et al., 2011). Invertases function to hydrolyze sucrose into monosaccharides glucose and fructose commonly localized in cell walls and have been linked to cell growth and development (Cho et al., 2005, Kohorn et al., 2006).

Of the genes involved in cell wall metabolism, (*Os04g33640*) encodes a glycosyl hydrolase family 17 protein, which family members have been shown to play a role in cell wall regulation (Minic and Jouanin, 2006, Irshad et al., 2008, Jamet et al., 2006), plant development, and defense (Thomas et al., 2000) was up-regulated in *mpg1*. (*Os03g12140*), encodes for a

glucan endo-1,3-beta-glucosidase precursor, which have been shown to play a role in cell wall biogenesis (Bosch et al., 2011) is down-regulated in *mpg1*.

OsBC1L9 (*Os10g35460*), encodes a COBRA-like protein, which is suggested to be involved in secondary cell wall formation (Dai et al., 2009) and homologous members in *Arabidopsis* have also been shown to participate in cellulose biosynthesis and secondary cell wall formation (Brown et al., 2005, Persson et al., 2005).

OsLTP1.18/OsLTP1 (*Os12g02320*), encodes for LTPL12 - protease inhibitor/seed storage/LTP family protein precursor, has been shown to play a role in cutin deposition (Samuel et al., 2002), is up-regulated in *mpg1*.

There were two cellulose synthase genes that were down-regulated in *mpg1* compared to WT-ns. *OsCD1/OsCSLD4/OsNRL1/OsND1/Ossle1/OsDNL1* (*Os12g36890*), encodes a CSLD4 - cellulose synthase-like family D protein, which has been shown to play a role in architecture and growth. This can be seen in several studies including in *OsCD1* RNAi lines that culminate in dwarfed plants with curled thin leaves (Luan et al., 2011) and in mutant lines that result in expression of xylan synthesis related genes and alteration in cell cycle regulation (Li et al., 2010a). Its native role has been expected to play a role in M phase regulation and cell proliferation (Yoshikawa et al., 2013). *OsCSLA9* (*Os06g42020*), encodes a CSLA9 - cellulose synthase-like family A protein, which has been shown to participate in cellulose synthesis in secondary cell walls (Kotake et al., 2011).

There are two saccharification genes that were also down-regulated in *mpg1*. (*Os12g01700*) and (*Os11g01620*), which both encode an inactive receptor kinase At2g26730 precursor protein. This protein in *Arabidopsis*, an LRR kinase, has been shown to play a role in enzymatic saccharification within the cell wall (Ohtani et al., 2017).

Of the genes assessed under metabolism for additional growth metrics (*Os04g52440*), *OsBC1L9* (*Os10g35460*), and (*Os03g29190*) belong to the gene ontology term GO:0016049 cell growth. Additionally *OsTPS8* (*Os08g34580*), (*Os04g33640*), and *OsDXS3* (*Os07g09190*) belong to the gene ontology term GO: 0005975 carbohydrate metabolism.

All but (*Os02g11070*), *OsLAC24* (*Os11g42220*), (*Os07g41060*), *OsTPS8* (*Os08g34580*), (*Os03g12140*), and (*Os11g01620*) have at least one GCC-box within their promoter region (-2000 from TSS). Several of the genes had higher numbers of the GCC-box motifs, (*Os04g33640*) contains 4, *OsLTP1.18/OsLTP1* (*Os12g02320*) contains 4, *OsCD1/OsCSLD4/OsNRL1/OsND1/Ossle1/OsDNL1* (*Os12g36890*) contains 5, and *OsCSLA9* (*Os06g42020*) contains 4.

Other potentially interesting DEG that could relate to plant productivity

Other genes that potentially influence biomass accumulation and seed yield were found by brute force analysis of independent genes derived from our DEG list from RNA-seq. *OsCYP51G3* (*Os05g12040*), encodes a cytochrome P450 51 protein, which has been shown to be involved with steroid biosynthesis regulating plant height, seed setting rate (Xia et al., 2015) and grain yield (Ma et al., 2016) under osmotic stress. *OsRLCK253* (*Os08g28710*), encodes a receptor protein kinase CRINKLY4 precursor protein, which family members in Maize have been shown to be responsible for cell-autonomous differentiation response and cell size (Becraft et al., 2001). (*Os04g52440*), encodes for an aminotransferase, which has been shown to play a role in QTL SS1 pathway which interacts with QTL SS2. SS1 acts as a regulator controlling source leaf width and grain number per panicle (Xu et al., 2015). *OsRIF3/OsCAD6* (*Os04g15920*) was found to be regulated by expression of *OsMPH1*, which when overexpressed leads to plants with increased height and grain yield (Zhang et al., 2017c). *CIGR1* (*Os07g36170*),

encodes for a chitin-inducible gibberellin-responsive protein, which has been found to be associated with tillering (Park, 2014, Paul et al., 2012). *OsRhmbd17* (*Os10g37760*), codes for a putative Rhomboid homologue, which expression has been noted during starch biosynthesis and mobilization in *glup4 OsRAB5* loss of function mutant in developing seeds (Doroshenk et al., 2010). *OsWRKY53* (*Os05g27730*), encodes a WRKY family transcription factor, that has been shown to play a role in plant architecture and seed morphology through regulation of brassinosteroid signaling (Tian et al., 2017).

One particular gene of interest that was down regulated in *mpg1* was *OsDLT/OsGRAS-32/OsD62/OsGS6/OsSMOS2* (*Os06g03710*). This gene encodes a DELLA-like protein from the GRAS family, that plays a positive role in BR signaling (Tian et al., 2004, Tong et al., 2009). Loss of function mutants for this gene result in alterations in plant morphology, including height, tillering, leaf size, root abundance, panicle and seed morphology through regulation of GA biosynthesis and BR signaling (Li et al., 2010b).

A number of the genes differentially expressed in *mpg1* have also been shown to be expressed under nutrient limitation response, therefore it might be interesting to speculate that *mpg1* is more efficient in resource use. Some of the genes involving transport include: amino acid transporters, element transporters, and a hormone transporter. Of the amino acid transporters (*Os01g41420*), encodes a transmembrane amino acid transporter, is up-regulated in *mpg1* which could potentially allow for enhanced protein generation. Additionally, *OsHT* (*Os08g03350*), encodes an amino acid transporter for histidine that could also aid with this function (Liu et al., 2005). Three amino acid biosynthesis genes were also found to be up-regulated, (*Os04g33390*) involved in phenylalanine biosynthesis, (*Os03g58290*) involved in tryptophan biosynthesis, and *OsDXS3* (*Os07g09190*) involved in thiamine biosynthesis (Naithani et al., 2017, Gupta et al.,

2016). *OsPht1;8/OsPT8* (*Os10g30790*), encodes an inorganic phosphate transporter, which is involved in phosphate homeostasis in rice (Jia et al., 2011). (*Os03g06520*), encodes a sulfate transporter, which could potentially play roles in rice morphology and metabolism (Hirose et al., 2007). (*Os04g57200*), encodes a heavy metal transporter. *OsPILS7a* (*Os09g38130*) encodes for an auxin efflux carrier, which plays a role in growth and development (Mohanta et al., 2015). Perhaps hormone pathways are also influencing the phenotype observed in *mpg1*.

DEG from 32 days post-planting stem tissue are involved in hormone regulation

The DEG gene list was taken and assessed against the KEGG and Gramene plant reactome pathways to assess genes influencing hormone biosynthesis and signal transduction in order to evaluate any potential differences in *mpg1* (Table 4.9).

Table 4.9: DEG that play a role in regulation of hormones from KEGG (Kanehisa et al., 2019, Kanehisa et al., 2017, Kanehisa and Goto, 2000) and Gramene Plant Reactome (Naithani et al., 2017, Gupta et al., 2016).

Hormone	Function	MSU ID	RAP ID	log2 Fold Change	Description
Auxin	Transport	LOC_Os09g38130	Os09g0554300	1	auxin efflux carrier component, putative, expressed
		LOC_Os04g38570	Os04g0459000	-1.067	OsABCB14 - multidrug resistance protein, putative, expressed
Brassinosteroid	Signal transduction	LOC_Os06g03710	Os06g0127800	-1.978	DELLA protein SLR1, putative, expressed
Cytokinin	Biosynthesis	LOC_Os01g56810	Os01g0775400	1.379	cytokinin dehydrogenase precursor, putative, expressed
Ethylene	Biosynthesis	LOC_Os05g05680	Os05g0149400	1.035	1-aminocyclopropane-1-carboxylate oxidase, putative, expressed
		LOC_Os04g10350	Os04g0182200	0.862	1-aminocyclopropane-1-carboxylate oxidase homolog 2, putative, expressed
Gibberellin	Signal transduction	LOC_Os03g28940	Os03g0402800	1.134	ZIM domain containing protein, putative, expressed
Jasmonic acid	Biosynthesis	LOC_Os02g48770	Os02g0719600	1.417	SAM dependent carboxyl methyltransferase, putative, expressed
	Signal transduction	LOC_Os03g08310	Os03g0180800	3.476	ZIM domain containing protein, putative, expressed
		LOC_Os10g25290	Os10g0392400	2.357	ZIM domain containing protein, putative, expressed
		LOC_Os03g08330	Os03g0181100	2.022	ZIM domain containing protein, putative, expressed
		LOC_Os03g08320	Os03g0180900	1.818	ZIM domain containing protein, putative, expressed
		LOC_Os03g28940	Os03g0402800	1.134	ZIM domain containing protein, putative, expressed
		LOC_Os02g41904	Os02g0629800	-1.193	DEF7 - Defensin and Defensin-like DEFL family, expressed
Salicylic acid	Signal Transduction	LOC_Os10g25290	Os10g0392400	2.357	ZIM domain containing protein, putative, expressed

Hormones play a key role in growth and development and response to stimulus. There are a few key genes that are differentially regulated in *mpg1* compared to wild-type plants that involve either the biosynthesis, transport, or signal transduction of several hormones.

There are two auxin transport genes that are differentially regulated in *mpg1*. *OsPILS7a* (*Os09g38130*), encodes an auxin efflux carrier component, which is shown to have developmentally regulated expression where expression is noted during 4-leaf vegetative growth and under IAA and CK treatments (Mohanta et al., 2015), and is up-regulated in *mpg1*. *OsABCB14* (*Os04g38570*), encodes an auxin transporter, that has been shown to transport auxin and regulate iron homeostasis, shown by knockdown mutants that exhibited decreased levels of polar auxin transport and insensitivity to iron deficiency (Xu et al., 2014). Auxin is an important plant hormone that regulates cell enlargement and plant growth, and potentially nutrient homeostasis. Perhaps there is a altered flux of auxin resulting in greater growth of *mpg1*.

Brassinosteroids play a diverse role in plant growth and development, particularly through cell elongation and cell division. A gene involving brassinosteroid signal transduction is down-regulated in *mpg1*. (*Os06g03710*), encodes a DELLA-SLR1-like GRAS protein*, DLT/OsGRAS-32,* is a positive regulator of BR signal regulation and GA metabolism; *DLT* loss of function mutants had enhanced expression of GA biosynthesis genes *OsGA20ox2/SD1* and *OsGA20ox3* (Li et al., 2010b). This gene has been shown to affect height and tillers (Tong et al., 2009), and negatively regulates grain size (Sun et al., 2013). The decreased expression of *DLT* in *mpg1* might be affecting height and biomass production through decreased levels of BR signaling and increased GA biosynthesis. Although not gathered from KEGG analysis or Gramene Plant Reactome, *OsXIAO* (*Os04g48760*), encodes a leucine-rich repeat family protein, is down-regulated in *mpg1*. This gene has been shown to play a role in BR signaling and cell

division as well (Jiang et al., 2012a). Again, not found in KEGG or Gramene Plant reactiome, *BLE2* (*Os07g45570*), codes for an expressed protein, found to be a brassinosteroid signaling gene involved in root development (Krishna et al., 2017, Qin et al., 2018), which is up-regulated in *mpg1*.

Cytokinin, like other plant hormones has diverse functionality, but can function in cell division and shoot initiation (Ferguson and Beveridge, 2009). *mpg1* experiences an up-regulation of *OsCKX5* (*Os01g56810*), which encodes a cytokinin dehydrogenase precursor, that functions in the cytokinin biosynthesis pathway. This particular genes expression has been found in relation to analysis of *OsAAP3* and its effects on bud outgrowth (Lu et al., 2018).

Ethylene is another plant hormone that can impact a variety things including plant growth. *mpg1* experiences up-regulation in two genes involved in the ethylene biosynthesis pathway. Both *OsACO5/OsACO6* (*Os05g05680*) and (*Os04g10350*) encode 1-aminocyclopropane-1-carboxylate oxidase proteins (ACC oxidase genes – (ACO)). *OsACO5/OsACO6* has been shown to be induced under fungal pathogen infection (Iwai et al., 2006) and during Brown Planthopper feeding (Hu et al., 2011). Not only does this hormone play a role in response defense response, but also growth and development. Ethylene can affect the development of vegetative tissues and senescence in a concentration or species dependent manner (Iqbal et al., 2017). Perhaps *mpg1* is experiencing altered levels of auxin-to-ethylene ratios and is resulting in the generation of enhanced growth or prolonging senescence.

Genes involved in the jasmonic acid signal transduction pathway are differentially expressed in *mpg1*. Five genes encoding ZIM-domain containing proteins are up-regulated and one gene encoding a defensin-like protein is down-regulated. These genes are classically involved in defense and stress response, but also play a role in hormone crosstalk with

gibberellin (Hou et al., 2013) and salicylic acid (Wei et al., 2014). Gibberellins are hormones commonly associated with growth, specifically stem growth. Salicilic acid is a hormone known to play a role in disease response and resistance. ZIM-domain containing proteins, also known as JAZ proteins, act as transcriptional repressors of JA responses in plants (Chini et al., 2007, Thines et al., 2007, Yan et al., 2007). Perhaps there is heightened expression of several JAZ proteins resulting in the inhibition of JA signal transduction, allowing for growth and development free from retardation as a result of de-activated JA signaling pathway (Liu et al., 2015). Additionally, (*Os02g48770*), which encodes a SAM dependent carboxyl methyltransferase is also up-regulated. This gene functions as a jasmonate O-methyltransferase that catalyzes the generation of methyl jasmonate from jasmonic acid. Methyl jasmonate operates as a cellular regulator involved with development and defense responses (Seo, et al., 2001).

Many of these genes have GCC-box motifs in their promoter regions and therefore might be targets for AP2/EREBP transcription factors. One gene that plays a role in JA signal transduction, *OsDEF7/OsCAL1/OsAFP1* (*Os02g41904*), encodes a defensin-like protein, is down regulated in *mpg1*. This gene has been shown to play a role cadmium translocation (Luo et al., 2018), stimulation of antifungal activity (Sagehashi et al., 2017), and are naturally upregulated under exposure to plant pathogens (Tantong et al., 2016).

All of these genes with the exception of *OsPILS7a* (*Os09g38130*), *OsABCB14* (*Os04g38570*), *OsACO5/OsACO6* (*Os05g05680*), (*Os04g10350*), *OsJAZ9* (*Os03g08310*), *OsJAZ10* (*Os03g08330*), and *OsDEF7/OsCAL1/OsAFP1* (*Os02g41904*) contain at least one GCC-box within their promoter regions (-2000 from TSS).

DEG from 32 days post-planting stem tissue are involved in DNA metabolism and cell cycle

The two greatest down-regulated gene ontology enrichment categories in *mpg1* fall under GO:0006259 DNA metabolism (Table 4.10) and GO:0007049 cell cycle (Table 4.11). It was unexpected to see a high number of genes from these categories down-regulated in *mpg1* because of its larger overall size compared to wild-type plants.

Table 4.10: DEG list from RNA-seq analysis of HM-*mpg1* compared to WT-ns within 32 days post-planting tiller tissue pertaining to the gene ontology category GO:0006259 - DNA metabolism.

MSU ID	RAP ID	log2 Fold Change	Description
Up-regulated DEG			
-	-	-	-
Down-regulated DEG			
LOC_Os10g26280	Os10g0402200	-2.49	ORC3 - Putative origin recognition complex subunit 3, expressed
LOC_Os01g67100	Os01g0896300	-2.144	expressed protein
LOC_Os06g14460	Os06g0256000	-2.128	chromosome condensation protein like, putative, expressed
LOC_Os03g11540	Os03g0214100	-2.086	RPA1B - Putative single-stranded DNA binding complex subunit 1, expressed
LOC_Os05g39850	Os05g0476200	-2.024	MCM3 - Putative minichromosome maintenance MCM complex subunit 3, expressed
LOC_Os11g29380	Os11g0484300	-2.004	MCM2 - Putative minichromosome maintenance MCM complex subunit 2, expressed
LOC_Os10g01570	Os10g0104900	-1.998	C-5 cytosine-specific DNA methylase, putative, expressed
LOC_Os05g14590	Os05g0235800	-1.993	MCM6 - Putative minichromosome maintenance MCM complex subunit 6, expressed
LOC_Os02g47150	Os02g0699700	-1.971	DNA topoisomerase 2, putative, expressed
LOC_Os06g08790	Os06g0187000	-1.966	ORC1 - Putative origin recognition complex subunit 1, expressed
LOC_Os11g03430	Os11g0128400	-1.924	CDC45B - Putative DNA replication initiation protein, expressed
LOC_Os12g13950	Os12g0242900	-1.899	POLA2 - Putative DNA polymerase alpha complex subunit, expressed
LOC_Os04g39670	Os04g0472700	-1.862	expressed protein
LOC_Os07g46540	Os07g0659500	-1.821	condensin complex subunit 1, putative, expressed
LOC_Os02g55410	Os02g0797400	-1.814	MCM5 - Putative minichromosome maintenance MCM complex subunit 5, expressed
LOC_Os01g64820	Os01g0868300	-1.731	POLA1 - Putative DNA polymerase alpha catalytic subunit, expressed
LOC_Os05g36280	Os05g0438700	-1.727	histone H3, putative, expressed
LOC_Os05g38480	Os05g0459400	-1.708	kinesin motor domain containing protein, putative, expressed
LOC_Os01g36390	Os01g0544450	-1.682	MCM4 - Putative minichromosome maintenance MCM complex subunit 4, expressed
LOC_Os11g05730	Os11g0155900	-1.671	histone H3, putative, expressed
LOC_Os01g31800	Os01g0502700	-1.655	Core histone H2A/H2B/H3/H4 domain containing protein, putative, expressed
LOC_Os03g17100	Os03g0279200	-1.651	Core histone H2A/H2B/H3/H4 domain containing protein, putative, expressed
LOC_Os08g38300	Os08g0490900	-1.594	Core histone H2A/H2B/H3/H4 domain containing protein, putative, expressed

LOC_Os07g08500	Os07g0182900	-1.58	C-5 cytosine-specific DNA methylase, putative, expressed
LOC_Os08g42600	Os08g0538700	-1.579	retinoblastoma-related protein-like, putative, expressed
LOC_Os02g56130	Os02g0805200	-1.568	PCNA - Putative DNA replicative polymerase clamp, expressed
LOC_Os01g64640	Os01g0866200	-1.516	histone H3, putative, expressed
LOC_Os12g37400	Os12g0560700	-1.514	MCM7 - Putative minichromosome maintenance MCM complex subunit 7, expressed
LOC_Os05g02300	Os05g0113900	-1.477	Core histone H2A/H2B/H3/H4 domain containing protein, putative, expressed
LOC_Os02g45940	Os02g0684500	-1.473	Core histone H2A/H2B/H3/H4 domain containing protein, putative, expressed
LOC_Os05g49860	Os05g0574300	-1.454	Core histone H2A/H2B/H3/H4 domain containing protein, putative, expressed
LOC_Os05g01230	Os05g0102600	-1.429	zinc finger, C3HC4 type domain containing protein, expressed
LOC_Os04g43300	Os04g0512400	-1.373	BRCA1 C Terminus domain containing protein, expressed
LOC_Os12g34510	Os12g0530000	-1.299	Core histone H2A/H2B/H3/H4 domain containing protein, putative, expressed
LOC_Os06g07210	Os06g0168600	-1.269	ribonucleoside-diphosphate reductase large subunit, putative, expressed
LOC_Os01g05610	Os01g0149400	-1.059	Core histone H2A/H2B/H3/H4 domain containing protein, putative, expressed
LOC_Os01g06010	Os01g0153300	-0.936	Core histone H2A/H2B/H3/H4 domain containing protein, putative, expressed
LOC_Os01g05970	Os01g0152900	-0.93	OsFBO1 - F-box and other domain containing protein, expressed
LOC_Os01g62230	Os01g0839500	-0.882	Core histone H2A/H2B/H3/H4 domain containing protein, putative, expressed

The genes that were down-regulated pertaining to DNA metabolism consisted of: two ORC origin recognition subunits, two uncharacterized proteins, two chromosome condensation related proteins, one RPA proteins, six minichromosome maintenance proteins, two C-5 DNA methylase proteins, ten core histone H2A/H2B/H3/H4 domain containing proteins, one DNA topoisomerase, one CDC DNA replication initiation protein, two POLA putative DNA polymerase subunits, three histone H3 proteins, one kinesin motor domain containing protein, one retinoblastoma related proteins, one PCNA putative DNA replicative polymerase clamp, one zinc finger, C3HC4 type domain containing protein, one BRCA1 C terminus domain containing protein, one ribonucleoside diphosphate reductase subunit, and *OsFBO1* F-box and other domain containing protein.

ORC proteins along with MCM proteins help to make up the pre-replication complex which functions to replicate DNA, and is responsible for plants to transition from G1 to S in

during interphase. Loss of function, overexpression, and ectopic expression of these genes results in differential plant growth (Brasil et al., 2017). Additionally, these proteins have been found in response to abiotic stress (Tuteja et al., 2011). Retinoblastoma genes are also an important regulatory step in this process, acting as negative regulators for cell cycle progression commonly expressed during floral organ development (Almutairi and Sadder, 2014). *OsRBR1* (*Os08g42600*), encoding a retinoblastoma protein, was found to be down regulated in *mpg1*. Down regulation of RBR1 in *Arabidopsis* results in increased G_2-phase cells (Hirano et al., 2008), and in rice increases S-phase cell frequency and enhanced biomass production in callus (Dudits et al., 2011).

Additional structural regulation genes are down-regulated in *mpg1*. *OsCMT3a* (*Os10g01570*), encodes a C-5 cytosine-specific DNA methylase, which has shown to be expressed in high levels during panicle development (Sharma et al., 2009) and *OsMET1-2* (*Os07g08500*), shows high levels of expression in all tissue types during floral tissue development (Sharma et al., 2009).

A large number of core histone H2A/H2B/H3/H4 domain containing proteins were also down-regulated in *mpg1*. These proteins have been shown to play a role in chromatin structure. Chromatin structure regulation can be induced under response to environmental stimulus and stress (Pawlak and Deckert, 2007, Chinnusamy and Zhu, 2009, Feng et al., 2010), during growth (Sui et al., 2012), and development (Feng et al., 2010, Shi et al., 2014). *OsRFPHC-11* (*Os05g01230*) is a variation in methylation gene responsible for methylation maintenance (Hu et al., 2014, Law and Jacobsen, 2010).

Additional genes from this enrichment category not mentioned play obvious roles in DNA replication and DNA structural orientation.

All of the genes from the gene ontology enrichment category for DNA metabolism contained GCC-box motifs within their promoter region (-2000 from TSS) with the exception of (*Os06g14460*), *MCM3* (*Os05g39850*), *OsCMT3a* (*Os10g01570*), (*Os02g47150*), (*Os07g46540*), (*Os01g64820*), (*Os03g17100*), (*Os08g38300*), *OsMet1-2/OsMET1b* (*Os07g08500*), *MCM7* (*Os12g37400*), (*Os05g49860*), *OsRFPHC-12* (*Os04g43300*), (*Os01g05610*), and (*Os01g06010*). Several of the genes that did contain the GCC-box element in their promoter had a higher number of these elements. *MCM2* (*Os11g29380*) contains 12, (*Os11g05730*) contains 4, (*Os01g31800*) contains 4, and (*Os05g02300*) contains 7 GCC-box sequences.

These results could indicate that *mpg1* might be experiencing less DNA replication, possibly through impacted cell cycle regulation. Additionally, major portions of DNA might be structurally modified in comparison to wild-type. The function of these genes play a role in growth and development, as well as response to stimulus, possibly playing a role in the phenotype of *mpg1*.

Table 4.11: DEG list from RNA-seq analysis of HM-*mpg1* compared to WT-ns within 32 days post-planting tiller tissue pertaining to the gene ontology category GO:0007049- cell cycle.

MSU ID	RAP ID	log2 Fold Change	Description
Up-regulated DEG			
-	-	-	-
Down-regulated DEG			
LOC_Os06g14460	Os06g0256000	-2.128	chromosome condensation protein like, putative, expressed
LOC_Os01g67100	Os01g0896300	-2.144	expressed protein
LOC_Os10g01570	Os10g0104900	-1.998	C-5 cytosine-specific DNA methylase, putative, expressed
LOC_Os01g49200	Os01g0685900	-1.989	microtubule associated protein, putative, expressed
LOC_Os02g47150	Os02g0699700	-1.971	DNA topoisomerase 2, putative, expressed
LOC_Os03g41100	Os03g0607600	-1.942	cyclin, putative, expressed
LOC_Os01g33040	Os01g0513900	-1.876	kinesin motor domain containing protein, expressed
LOC_Os04g39670	Os04g0472700	-1.862	expressed protein
LOC_Os02g28850	Os02g0489800	-1.845	Kinesin motor domain domain containing protein, expressed

LOC_Os07g46540	Os07g0659500	-1.821	condensin complex subunit 1, putative, expressed
LOC_Os03g38010	Os03g0577100	-1.818	nuf2 family protein, expressed
LOC_Os01g15480	Os01g0259400	-1.759	EMB3013, putative, expressed
LOC_Os02g10020	Os02g0193600	-1.74	Mad3/BUB1 homology region 1 domain containing protein, expressed
LOC_Os01g64820	Os01g0868300	-1.731	POLA1 - Putative DNA polymerase alpha catalytic subunit, expressed
LOC_Os12g31810	Os12g0502300	-1.725	cyclin, putative, expressed
LOC_Os06g51110	Os06g0726800	-1.709	cyclin, putative, expressed
LOC_Os05g33890	Os05g0409400	-1.68	microtubule associated protein, putative, expressed
LOC_Os05g41390	Os05g0493500	-1.641	cyclin, putative, expressed
LOC_Os07g32390	Os07g0507200	-1.629	targeting protein-related, putative, expressed
LOC_Os01g59120	Os01g0805600	-1.599	cyclin, putative, expressed
LOC_Os07g08500	Os07g0182900	-1.58	C-5 cytosine-specific DNA methylase, putative, expressed
LOC_Os08g42600	Os08g0538700	-1.579	retinoblastoma-related protein-like, putative, expressed
LOC_Os04g28260	Os04g0350300	-1.521	Kinesin motor domain domain containing protein, expressed
LOC_Os12g40860	Os12g0601000	-1.489	Leucine Rich Repeat family protein, expressed
LOC_Os04g40940	Os04g0486500	-1.487	mitotic spindle checkpoint protein MAD2, putative, expressed
LOC_Os07g01490	Os07g0105700	-1.479	kinesin motor domain containing protein, putative, expressed
LOC_Os01g13260	Os01g0233500	-1.374	cyclin-A1, putative, expressed
LOC_Os12g24550	Os12g0433500	-1.337	expressed protein
LOC_Os08g40170	Os08g0512600	-1.333	cyclin-dependent kinase B2-1, putative, expressed
LOC_Os04g47580	Os04g0563700	-1.269	cyclin, putative, expressed
LOC_Os12g39830	Os12g0588800	-1.193	cyclin, putative, expressed
LOC_Os07g42860	Os07g0620800	-1.156	cyclin, putative, expressed
LOC_Os08g44360	Os08g0557800	-1.137	male sterility protein 2, putative, expressed

The genes pertaining to the cell cycle gene ontology category were: two chromosome condensation proteins, three uncharacterized proteins, two DNA methylase proteins, two microtubule associated proteins, one topoisomerase, nine cyclin proteins, four kinesin motor domain containing proteins, a nuf2 family protein, an EMB3013 protein, a Mad3/BUB1 homology region 1 domain containing protein, a DNA polymerase, a targeting protein-related protein, a retinoblastoma-related protein, a leucine rich repeat family protein, a MAD2 mitotic spindle checkpoint protein, a cyclin-dependent kinase, and a male sterility protein.

A number of genes coding for cyclins were down regulated in *mpg1*. Cyclins are proteins responsible for progression of the cell cycle. A total of nine cyclin genes were found to be differentially regulated. *Orysa;CYCA3;2* (*Os03g41100*), is a cell cycle gene whose class of cyclins are involved in cell division and differentiation with high levels of expression in meristematic tissue (Yu et al., 2003, Pettko-Szandtner et al., 2015). *OsCYCA2;1* (*Os12g31810*), has been shown to play a role in suppression of DNA endoduplication and cell enlargement (Qu et al., 2018). *Orysa;CYCB2;2* (*Os06g51110*), has been shown to relate to cell proliferation (Ni et al., 2014) and expressed in a G2/M phase-specific manner (Ma et al., 2009) and is recognized as a mitotic cyclin (Umeda et al., 1999). *Orysa;CYCB1;3* (Os05g41390), encodes a cyclin that coexpression network analyses have associated with growth characteristics (Ficklin et al., 2010). *Orysa;CYCB1;4* (*Os01g59120*), has been shown to play a role in cold stress response and embryo development (Ma et al., 2009, Guo et al., 2010); also referred to as *OsCYCB1;4*, this gene is found in high levels of expression of leaf meristematic tissue (Pettko-Szandtner et al., 2015). *OsCYCA2.3/ Orysa;CYCA1;3* (*Os01g13260*) is a G2 related cyclin which has been found in gene expression networks assessing cell cycle genes with growth, development, and survivability (Shi et al., 2019). *Orysa;CYCB2;1* (*Os04g47580*) is expressed during G2 and M phase and is recognized as a mitotic cyclin (Umeda et al., 1999, Lee et al., 2003). *Orysa;CYCD5;3* (*Os12g39830*) was found to be highly up-regulated in basal leaf segments (Pettko-Szandtner et al., 2015). *Orysa;CYCD4;1* (Os07g42860), has been found to be a potential target of AGL2 which interacts with MADS-box containing genes in *Arabidopsis* (Pinyopich et al., 2003, Ferrario et al., 2004). These genes are recognized as core cell cycle gene (Wang et al., 2004a, La et al., 2006, Guo et al., 2007, Pettko-Szandtner et al., 2015).

Many cyclin genes are transcriptionally regulated in response to stress as well (Zhao et al., 2014). Perhaps the transcriptional down-regulation that was noted in *mpg1* is merely due to a number of stress response genes that are up-regulated.

In leu of all of the differentially expressed cyclins, *CDKB2;1* (*Os08g40170*), which encodes a cyclin-dependent kinase was also down-regulated. This gene has been shown to be important for mitosis (Endo et al., 2012), with high activity in the meristem (Pettko-Szandtner et al., 2015).

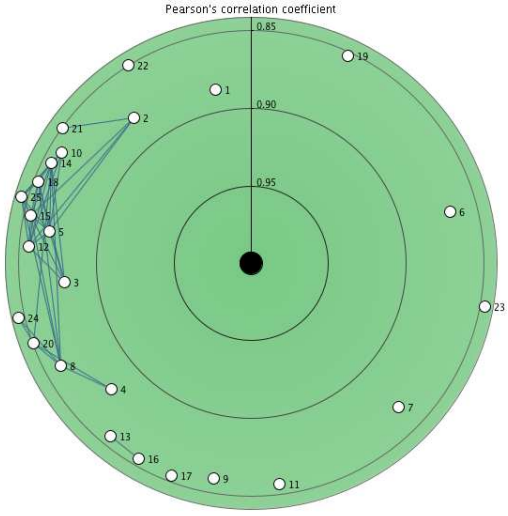
DBS1 (Os01g33040), encodes a kinesin motor domain containing protein, NACK-type kinesisn-like protein, which participates in initiating organ primordia in rice (Sazuka et al., 2005). (*Os02g28850*), another down-regulated kinesin motor domain containing protein, is closely related to other kinesins expressed exclusively in rapidly dividing cells towards the plus-end of phragmoplast microtubules (Guo et al., 2009). *OsKRP1* (*Os04g28260*), a kinesin motor domain containing protein, is involved with cell proliferation as well (Meguro and Sato, 2014). Overexpression of this gene resulted in inhibition of development and reproductive tissues (Barroco et al., 2006).

Of the genes involved with the cell cycle gene ontology category the following contained GCC-box motifs within their promoter region (-2000 from TSS): (*Os01g67100*), *Orysa;CYCA3;2* (*Os03g41100*), (*Os04g39670*), (*Os02g28850*), (*Os01g15480*), *Orysa;CYCB1;3* (*Os05g41390*), (*Os07g32390*), *OsRBR1* (*Os08g42600*), *OsKRP1* (*Os04g28260*), (*Os04g40940*), *OsDLK* (*Os07g01490*), (*Os12g24550*), *Orysa;CYCB2;1* (*Os04g47580*), and *Orysa;CYCD5;3* (*Os12g39830*). In particular there were several genes with a high number of GCC-box motifs in their promoter region. (*Os01g15480*) contains 16, (*Os12g24550*) contains 11, and

Orysa;CYCD5;3 (*Os12g39830*) contains 5 GCC-box sequences.

MPG1 co-expression network reveals potential insights into functionality

To see if *MPG1* is still regulating or potentially influencing similar things as it is natively, we evaluated transcriptome differences between the mutant *mpg1* and WT-ns, with native co-expressed genes of *MPG1*. Co-expression was found using affymatrix rice genome array using Genevestigator co-expression software. Analysis of *MPG1* (*Os08g41030*) and top 25 highest positive co-expressed genes (Figure 4.12).



	Score	Gene	Description
1	0.89	LOC_Os03g29570	mps one binder kinase activator-like 1A, putative, expressed
2	0.88	LOC_Os10g11580	histone-like transcription factor and archaeal histone, putative, expressed
3	0.88	LOC_Os07g47240	expressed protein
4	0.88	LOC_Os02g36680	expressed protein
5	0.87	LOC_Os01g29840	no apical meristem protein, putative, expressed
6	0.87	LOC_Os11g31380	no apical meristem protein, putative, expressed
7	0.87	LOC_Os04g44580	expressed protein
8	0.86	LOC_Os02g10800	mitochondrial carrier protein, putative, expressed
9	0.86	LOC_Os01g33350	zinc-binding protein, putative, expressed
10	0.86	LOC_Os08g16860	OsFBX282 - F-box domain containing protein, expressed
11	0.86	LOC_Os06g02460	expressed protein
12	0.86	LOC_Os02g15070	glutelin, putative, expressed
13	0.86	LOC_Os02g19420	expressed protein
14	0.86	LOC_Os12g42670	expressed protein
15	0.85	LOC_Os07g42390	DUF581 domain containing protein, expressed
16	0.85	LOC_Os10g04890	expressed protein
17	0.85	LOC_Os02g09270	cytochrome P450 71A1, putative, expressed
18	0.85	LOC_Os02g12310	no apical meristem protein, putative, expressed
19	0.85	LOC_Os02g25860	glutelin, putative, expressed
20	0.85	LOC_Os01g41370	FBD domain containing protein, putative, expressed
21	0.85	LOC_Os02g49410	histone-like transcription factor and archaeal histone, putative, expressed
22	0.85	LOC_Os01g24460	histone-like transcription factor and archaeal histone, putative, expressed
23	0.85	LOC_Os07g11510	RAL6 - Seed allergenic protein RA5/RA14/RA17 precursor, expressed
24	0.85	LOC_Os02g28580	expressed protein
25	0.85	LOC_Os03g55734	SSA3 - 2S albumin seed storage family protein precursor, expressed

Figure 4.12: Co-expression network analysis of *MPG1*.
Genevestigatior software analysis of genes with high co-expression with *MPG1* using a Pearson's correlation coefficient of 85% and higher with connected genes with mutual correlation of at least 97% and higher. Genes are listed numerically in order of correlation with associated gene ID and description.

Assessment of the genes found in positive co-expression with *MPG1* are of potential interest. Of the top 25 co-expressed genes there were: one binder kinase activator-like 1A protein, three histone-like transcription factor and archaeal histone proteins, eight uncharacterized proteins, three no apical meristem proteins, one mitochondrial carrier protein, one zinc-binding protein, one OsFBX282 F-Box domain containing protein, two glutelin proteins, one DUF581 domain containing protein, one cytochrome P450 71A1 protein, one FBD

domain containing protein, one RAL6 – seed allergenic protein RA5/RA14/RA17 precursor protein, and one SSA3 – 25 albumin seed storage family protein precursor protein.

Not surprisingly a number of these genes function in seed development, as *MPG1* is natively expressed in aleurone tissue and functions in endosperm development (Xu et al., 2016a). The histone-like transcription factor and archaeal histone proteins are also known as nuclear factor transcription factors. From the list of these proteins found to be co-expressed are (*NF-YC12*) found to play a direct role in grain quality (Bello et al., 2018), (*OsNF-YB1*) involved in grain filling and endosperm development (Xu et al., 2016a), and (*NF-YC10*) regulates grain width through affecting cell proliferation (Jia et al., 2019).

Another highly co-expressed gene family represented are glutelin proteins. These class of proteins act as seed storage proteins and make up a major component of endosperm (Okita et al., 1989).

The other largely represented gene class present amongst the co-expressed genes were no apical meristem transcription factors. Three were listed, (*ONACO26*) functions in grain size/weight (Mathew et al., 2016), (*ONAC129*) associated with seed specific expression (Mathew et al., 2016), (*HDA707*), actually a histone deacetylase, function to modulate gene expression under developmental regulation (Hu et al., 2009). The no apical meristem class of genes was named as such because mutations in this gene resulted in the loss of shoot apical meristem development, and are expected to play a role in determining positions of meristems and primordia (Souer et al., 1996). Perhaps this means *MPG1*, might play a role in the regulation of meristem activity, which could be responsible for a number of the characteristics observed in *mpg1* phenotype.

Although co-expression doesn't equate to co-regulation, we continued evaluation of co-expressed genes with the differentially expressed genes of *mpg1*. Surprisingly, cross-evaluating the greatest positively and negatively co-expressed 400 genes with *mpg1* with our list of differentially regulated genes revealed zero similarly co-expressed genes. However, there were two genes from the 800 co-expressed genes that were inversely regulated in *mpg1*. (*Os11g41710*), encoding a cytochrome P450, and (*Os05g33630*), encoding an inosine-uridine preferring nucleoside hydrolase, normally found in negative correlation co-expression with *MPG1* expression were up-regulated in *mpg1*.

Again, because *MPG1* is being expressed outside of its native function, it is difficult to determine the impact that its expression is having in tissues not normally expressing it. The lack of correlation between natively co-expressed genes and genes differentially expressed in *mpg1* makes it difficult to ascertain what *MPG1*'s expression affects. Evaluation of conserved domains by assessment of *MPG1*'s nucleic acid and protein sequences in NCBI showed only the AP2/EREBP ERF domain. Yet, there might be other functional domains present in *MPG1* that are not yet well characterized. The differentially expressed genes uncovered in our transcriptomic analysis of *mpg1* gives us insight into the impact ectopic expression of *MPG1* might be playing. Genes involving transcriptional regulation, flowering, stress response, and metabolism are affected.

Discussion

MPG1 is the most significantly differentially regulated gene in mpg1

By in large the gene with the greatest degree of differential expression between HM-*mpg1* and WT-ns at 32 days post-planting was *MPG1* (*Os08g41030*). *MPG1* was also the most significantly differentially up-regulated gene at 42 days post-planting. It experienced up-

regulation by roughly 12 $\log_2$ fold change and 9.5 $\log_2$ fold change in 32 and 42 days post-planting tiller tissue respectively (Table 4.1, 4.3). Additionally, it is the only gene directly neighboring the site of the T-DNA insertion with noticeably altered expression in *mpg1* (Figure 4.2, 4.10). This increases our confidence that *MPG1*, the AP2/EREBP transcription factor, is the gene responsible for generating the phenotype observed in *mpg1*. The expression of this gene correlates with several characteristics seen in *mpg1* compared to wild-type. The mutant experiences greater biomass accumulation through increased height, increased leaf size, and increased tiller girth. Ratooning capabilities by means of survivability, biomass regeneration, and seed yield are also higher in *mpg1*. *mpg1* plants also experience a delay in flowering. Additionally, *mpg1* produces greater seed yield, and spikelets with awns. Lastly, *mpg1* appears to possess a degree of stress tolerance. The expression of *MPG1* is likely influencing the expression of numerous genes that independently relate to growth, development, and stress response resulting in each of the traits seen in *mpg1*. *MPG1* also possesses two GCC-box motifs in its promoter region, perhaps it functions in a positive feedback fashion targeting itself for further expression, allowing for continuous expression.

Genes with GCC-box motifs belong to a variety of functional activities

MPG1 is a transcription factor that has been shown to target GCC-box motif elements (GCCGCC) (Xu et al., 2016a). This sequence is relatively frequent in the rice genome with at least one of these elements present in the -2000 from TSS of roughly 34% of all genes. Nearly half of the DEG generated from our RNA-seq experiment contained this motif in their assessed promoter regions (Figure 4.1). DEG containing at least one GCC-box motif were analyzed for ontological enrichment.

Up-regulated genes containing this *cis*-element pertained to a variety of ontological categories. Under biological process, genes were enriched for response to stress, physiological process, and metabolism. Under cellular component, ontological categories were enriched for extracellular region, cell wall, and external encapsulating structure. Under molecular function, categories were enriched for catalytic activity, binding, and transcription factor activity.

From the down-regulated genes that contain a GCC-box motif, genes were enriched for ontological categories pertaining to biological processes of DNA metabolism, cellular process, cell cycle, cell organization and biogenesis, and cell differentiation. Cellular component enriched categories were membrane bound-organelle, nucleus, cell wall, cytoplasm, and cytoskeleton. Under molecular function enriched categories were DNA binding, motor activity, and transcription factor activity.

Again, the presence of the this *cis*-motif in the promoter regions of these genes does not mean they are functional targets of MPG1 under the ectopic expression present in *mpg1*. However, the list of differentially regulated genes generated in this study could be used to investigate particular genes of interest and their association with MPG1.

Genes with a higher number of *cis*-elements tend to show greater functionality for their *trans*-acting factors. Of the DEG, promoter regions were assessed for genes containing a high number of the GCC-box *cis*-acting elements (Table 4.2). There were only a total of 19 genes that were differentially regulated, and contained an enriched number of GCC-box motifs in their promoter regions, 5 genes up-regulated and 14 genes down-regulated. Ontological assessment of these genes revealed no enrichment categories for the up-regulated genes but the down-regulated genes pertained to DNA binding and motor activity under molecular function. Perhaps these

genes are direct targets of MPG1.

Several genes are strongly co-regulated between 32 and 42 days post-planting

Although it was determined that the RNA-seq experiment performed at 42 days post-planting was not acceptable to evaluate transcriptional differences between *mpg1* and WT-ns because of a differential shift in development (Figure 4.11), we evaluated its results for similarly co-regulated genes that might be of significant influence in determining the phenotype of *mpg1*. A small number of genes, fourteen up-regulated and four down-regulated genes were found (Table 4.3). The greatest co-expressed up-regulated gene was *MPG1*. Of the other up-regulated genes most of the genes expression and predicted function were found in accordance with stress/stimulus response. In particular a high number of these genes were found during drought stress, pathogen response, and glutamate response. Further investigation is necessary surrounding *mpg1's* interaction with particular stressors and/or stimuli to fully evaluate its capabilities. Many of these genes elude to the possibility that ectopic expression of *MPG1* influences genes that are induced under osmotic stress, particularly drought, and pathogen attack. A number of genes were found that have also been shown to respond to glutamate – an active amino acid which can be involved in protein synthesis and metabolism. Maybe *mpg1* is able to optimize resources for growth, playing a part in its ability to generate greater biomass and seed yield.

Of the down-regulated genes co-expressed between 32 and 42 days post-planting pertain to flowering regulation and development, seed storage, and carbon metabolism. In particular, the greatest down-regulated gene was *OsMADS14* (Os03g54160), which codes for a transcription factor involved in flowering activation (Jeon et al., 2000, Kim et al., 2007). This gene is also the most significantly down-regulated gene in tiller tissue at 32 days post-planting between *mpg1*

and WT-ns. Perhaps these genes play a significant role in the delay in flowering phenotype, and enhanced biomass accumulation through increased vegetative development time.

Mutant *mpg1* experiences differential expression of a variety of transcription factors

Numerous transcription factors were differentially expressed in *mpg1* compared to WT-ns. The DEG pertained to a variety of functions including growth and development, and response to stress. Many transcriptional regulators were both up- and down-regulated in *mpg1,* possibly explaining the large and diverse phenotypic differences observed compared to WT-ns plants. A number of these genes also contained GCC-box *cis*-elements within their promoter regions (-2000 from TSS), possibly being directly regulated by the enhanced expression of *MPG1*. A number of these transcription factors provided further insight into potential causes of a few of the characteristics observed in *mpg1* (delayed flowering – MADS genes, stress response - AP2/EREBP, WRKY, bHLH, MYB, growth and development – various hormone, metabolism, and cell cycle regulatory genes).

***mpg1* experiences differential gene expression in genes regulating vegetative to reproductive transition**

The process of flowering in plants requires the transition between the vegetative and reproductive stage of development. Flowering can be triggered in different ways depending on the particular organism. In rice, once flowering is triggered the shoot apical meristem (SAM) converts into rachis meristem (RM), leading to the production of the bract primordium and primary branch meristem (PBM). Elongation of the PBM leads to production of secondary branch meristem (SBM), both which finalize in terminal spikelet meristem, culminating in individual floral meristems (FM) leading to the proliferation of floral organs (Itoh et al., 2005).

Flowering is triggered in rice by the accumulation of florigens in leaf phloem (Komiya et al., 2008, Tamaki et al., 2007). The particular induction florigens are *HEADING DATE 3a* (*Hd3a*) for short-day conditions (Komiya et al., 2008), and *RICE FLOWERING LOCUS T1* (*RFT1*) for long-day conditions (Tamaki et al., 2015). Sufficient environmental conditions drive expression of these florigens in phloem tissue to act as a mobile signaling molecule (Cho et al., 2017). Expression of florigen genes results from homodimerization of a type-B responsive regulatory element Ehd1, which is inhibited by cytokinin-inducible OsRR1 (Cho et al., 2016). Transcription of *Ehd1* can be induced or suppressed by numerous upstream regulatory elements. Elements that induce transcription include OsID1, Ehd4, Hd1, and OsMADS51, while elements that suppress transcription include AP2, Ghd7, Hd1, COL4, and OsLFL1 (Tsuji et al., 2011, Lee and An, 2015, Cho et al., 2017).

Hd3a and RFT1 proteins translocate to the SAM, where activation of downstream target genes responsible for initiation of reproductive tissues occurs. Hd3a interacts with a Gf14 family protein 14-3-3, allowing further interaction with OsFD1 (forming a florigen activation complex (FAC)) allowing nuclear targeting (Taoka et al., 2017, Tsuji et al., 2013). The presence of FAC induces expression of *OsMADS14, OsMADS15,* and *OsMADS18*, which are members of the *FRUITFUL* (*FUL*)-clade MADS box genes, and *OsMADS34*, a *SEPALLATA* (*SEP*)-clade gene, are activated (Kobayashi et al., 2012). Both *Hd3a* RNAi and *RFT1* RNAi lines have decreased expression of *OsMADS14* and *OsMADS15* in the SAM, indication that these genes act downstream of the florigens (Komiya et al., 2008, Komiya et al., 2009, Tsuji et al., 2011).

Individual suppression of the (*FUL*)-clade MADS box genes *OsMADS14, OsMADS15,* and *OsMADS18* via RNAi resulted in no noticeable differences in flowering time or development, whereas suppression of all three together resulted in delayed flowering (Kobayashi

et al., 2012). This suggests that *OsMADS14, OsMADS15,* and *OsMADS18* function redundantly. We see a notable decrease in expression of *OsMADS14,* and *OsMADS18*, perhaps reduction of expression of these genes results in the delay in flowering phenotype observed in *mpg1*.

Interestingly, a study investigating the effects of *OsMADS51* not only revealed that it functions upstream of *Ehd1, Hd3a, and OsMADS14*, but also that loss of function mutants for *OsMADS51* under short days flowered roughly 2 weeks later than wild-type segregants (Kim et al., 2007). Although we do not see a significant difference in *OsMADS51* expression in *mpg1*, it too experiences a 2 week delay in flowering compared to wild-type and WT-ns. This is interesting to note because of the concern that *mpg1*'s increased biomass might be solely from a 2 week increase in vegetative growth, however the *OsMADS51* null mutants at time of wild-type segregant heading showed no visible increase in height or girth (Kim et al., 2007), which did occur in the *mpg1* mutant. Although this study used a different cultivar of rice this suggests that *mpg1* might have increased biomass accumulation not only from increased vegetative phase, but from other means as well.

However, *OsDTH8*, a gene that encodes a putative HAP3 subunit of the CCAAT-box-binding transcription factor, when expressed resulted in a roughly 14 day delay in flowering, increase in plant height, and increase in number of spikelets per panicle (Wei et al., 2010). This flowering suppressor influences plant height and yield potential simultaneously, which aligns similarly to what is observed in *mpg1*. Therefore it is possible that the events of suppressed flowering in *mpg1* also result in its enhanced growth metrics.

OsMADS14 is possibly repressed by *MPG1*, as it contains a GCC-box motif in its promoter region (-2000 from TSS). The expression of *MPG1* and reduced expression of this florigen activated reproductive transition gene might be directly responsible for the delay in

flowering phenotype as well as add to growth due to an increase in vegetative growing period. Additionally, altered regulation of this gene might account for the even greater delay in flowering observed in *mpg1* under field conditions.

mpg1 plants experience differential regulation in numerous genes pertaining to stress response

A large number of genes were up-regulated in *mpg1* plants that pertained to the ontological category of stress response. This category contained a high number of the most differentially expressed genes between HM-*mpg1* and WT-ns. Many of these genes have been shown to play a role in abiotic and biotic stress response. Specifically, a number of these genes have been shown to be induced under drought stress. Other genes from this gene ontology category that were up-regulated in *mpg1* have been shown to play a role in defense response to a variety of different bacterial and fungal pathogens, as well as response to pest feeding.

Previous studies on *mpg1* have investigated the effects of prolonged salt and drought stress exposure. Under both of these stressors, *mpg1* was able to accumulate greater biomass but appeared to have to greater tolerance to salt or drought. It was difficult to determine if *mpg1* actually had any stress tolerance or if it just accumulated biomass at a greater rate throughout its life cycle, as both *mpg1* and WT-ns experienced a similar degree symptomatic response Additional studies evaluating the effects of these stressors on *mpg1* will be necessary. The delay in flowering phenotype increases the difficulty of evaluating these stressors as *mpg1* develops at a different rate compared to WT-ns plants. These stress treatments are important to perform during similar stages of development to appropriately evaluate their response.

A pathogen challenge utilizing *Xanthomonas oryzae* pv. *oryzae* was conducted on *mpg1* plants. Although there was no difference in necrotic legion sizes between the mutant and wild-

type plants, expression of *MPG1* was induced in wild-type plants post-inoculation (APPENDIX). This suggests that maybe this gene does play a role in defense response both natively and during ectopic expression. Further evaluation of different pathogen challenges will be necessary to evaluate this trait further.

DEG potentially involved with enhanced ratooning and biomass accumulation in *mpg1*

Biomass accumulation and growth remain particularly complex. Molecular mechanisms of ratooning have not been evaluated in detail, so a pseudo-analogous assessment of 're-growth' was performed by analyzing the transcriptome of defoliated seedlings and comparing differentially regulated genes to the ones affected in *mpg1* (Table 4.7). Numerous co-expressed genes were found and their ontological functions mainly pertain to stress response, lipid metabolism, and biosynthesis.

Beyond the genes assessed by ontological enrichment, through brute-force assessment, a number of other differentially regulated genes presented with potential interest as well. Although not enriched under a particular category, a number of genes were found that correlate to nutrient depletion response, hormone regulation, and carbon metabolism. Maybe *mpg1* is more proficient at performing resource allocation, and translocation providing greater biomass accumulation. This might also explain the greater phenotypic difference between *mpg1* and WT-ns under non-optimum conditions (expected to create a high salt, low pH environment. The fertilizer used during non-optimum conditions was made up of a different combination of elements than the optimum conditions fertilizer. It is possible that *mpg1* is able to acquire and/or use these nutrients in a more efficient manner.

Additionally, a number of genes involving cell wall biosynthesis were differentially expressed in *mpg1* (Table 4.8). These genes could be playing a role in greater tissue generation

and morphology. It would be enlightening to assess the structure and orientation of cells in *mpg1* throughout development compared to wild-type plants. Perhaps *mpg1* has an altered cell number/cell size ratio, which could also potentially explain the enhanced biomass accumulation of *mpg1*.

DEG indicate a potential for differential hormone regulation in *mpg1*

Hormones influence a variety of factors ranging from growth, and development, to response to stimulus, and stress. These hormones are used to signal the induction and suppression of genes to elicit physiological function. Numerous genes involved in hormone regulation were differentially expressed in *mpg1* (Table 4.9).

mpg1 has differential expression of genes regulating auxin transport, perhaps the flux of auxin in the mutant is resulting in greater growth, particularly greater height and girth. *mpg1* also experiences a decrease in genes involved in brassinosteroid signal transduction, which could affect cell elongation and cell division. This could be influencing the cell cycle affecting cell size or number. Genes involved in the biosynthesis of cytokinin and ethylene are also up-regulated. It could be possible that *mpg1* is producing greater amounts of these hormones leading to enhanced shoot growth and stress response. Signal transduction of gibberellins and salicylic acid are also up-regulated in *mpg1* affecting both plant defense and stem growth.

One major factor influencing grain per panicle is the accumulation of cytokinins (Ashikari et al., 2005, Li et al., 2013b). Similarly, *mpg1* experiences an up-regulation in a cytokinin biosynthesis gene, perhaps *mpg1* is accumulating a greater level of cytokinins giving rise to this phenotype.

Of the genes affecting hormone interaction, a large number of ZIM domain containing proteins or (JAZ) proteins were up-regulated. These genes largely participate in the signal

transduction pathway of jasmonic acid. They function to transcriptionally repress JA responses in plants (Chini et al., 2007, Thines et al., 2007, Yan et al., 2007). Perhaps the higher number of up-regulated JAZ proteins arises from a higher degree of bioactive jasmonic acid present in *mpg1*, and are present in effort to down regulate this response. Or, maybe *mpg1* influences the over-production of these genes and results in jasmonic acid insensitive plants. The JA pathway is commonly stimulated by defense responses, which influences growth-defense trade-offs, producing smaller plants (Guo et al., 2018). The up-regulation of these genes might be contributing to increased plant growth by reducing the jasmonic acid defense response pathway. Interestingly other defense and stress response genes are activated, possibly allowing for enhanced growth without a tradeoff with this hormone-directed defense mechanism.

If MPG1 is targeting JAZ proteins directly or even influencing their expression downstream, this might explain why plants overexpressing *MPG1* are incapable of germinating. Perhaps constitutive overexpression resulted in even greater stimulation of JAZ proteins creating a true jasmonic acid insensitive plant. Jasmonic acid insensitive mutants are often sterile because of defective anther and filament formation and pollen viability, as appropriate jasmonic acid signaling is necessary for the development of these structures (Browse, 2009). JAZ proteins have also been found to bind to targets other than jasmonic acid signaling, influencing other networks and elaborating on hormone cross talk (Pauwels and Goossens, 2011). Hormone cross-talk and interaction is very complex and these are merely speculations that could aid in the explanation of the *mpg1* phenotype.

Of the genes that were similarly differentially regulated under defoliation and *mpg1,* as well as genes regulating stress response, growth, nutrient regulation/response, and hormone biosynthesis, transport, and signaling could all potentially play roles in the increased plant

productivity observed in *mpg1*. Any one of these factors, or combination thereof, could aid in generating the phenotype of enhanced plant productivity in the form of biomass accumulation, seed yield, and stress response.

Many DEG pertain to DNA metabolism and cell cycle

The most significantly enriched down-regulated ontological categories consisted of DNA metabolism (Table 4.10) and cell cycle (Table 4.11). These categories often go hand in hand as they are necessary for continued proliferation of cells.

The DEG relevant to DNA metabolism consisted of a high number of genes coding for minichromosome maintainince (MCM) proteins, core histone domain containing proteins, and other chromatin modification genes. MCM proteins are involved in transitional regulation from G1 to S phase of interphase. Core histone domain containing proteins are found to be expressed during response to stress and stimulus.

The specific DEG pertaining to the cell cycle category were mainly cyclins. The MCM and chromatin modification genes along with cyclins indicate that *mpg1* plants are experiencing different patterns of growth compared to wild-type, as they are all heavily responsible for regulating progression of the cell cycle.

Because many of these genes were down regulated in *mpg1* it might suggest that the mutant doesn't proliferate cells as quickly as wild-type plants. This is counterintuitive to what would be expected, as *mpg1* clearly has more tissue than wild-type. Perhaps the cell division/expansion ratio resulted in altered plant morphology. It is possible that the increase in biomass was resultant of increased cell size in *mpg1* compared to wild-type plants. Previous studies have tried manipulating the expression of cell cycle genes (Czerednik et al., 2015, Czerednik et al., 2012, Gonzalez et al., 2007, Mathieu-Rivet et al., 2010, Nafati et al., 2011) to

alter cell number/cell size garnering significant phenotypic and morphological changes. Inhibition of cell division can cause a compensation effect of increasing cell size (Hemerly et al., 1995). Alterations in the cell cycle can by means of degradation of M-phase cyclins result in endoreduplication and increased cell size as well (Boudolf et al., 2004).

One of several scenarios might explain the differentially expressed genes involved in DNA metabolism and cell cycle in *mpg1*. One being that the rate of proliferation to growth was modulated generating larger plants, and/or possibly undergoing endoreduplication. Another being that wild-type plants began to transcriptionally transition to the reproductive phase of development and are beginning to generate flowering tissue. The higher expression rates of cell cycle regulatory genes might be explained in wild-type plants, as they might have begun the transition from vegetative to reproductive phases of development and in doing so have elevated proliferation of cells for the creation of flowering tissues. Lastly, alterations in DNA metabolism and cell cycle could have arose from the stimulation of stress response genes. These pathways have been shown to be modulated under stressful conditions.

Regardless, evaluation of cell size and cell number in tissues across development would be enlightening. Additionally, observation of meristematic tissues could also prove to be an important insight into understanding the nature of enhanced biomass visualized in *mpg1*. Altered plant morphology, and growth are very often the result of meristematic tissue influence. The presence of awns, also present in *mpg1*, has been reported in plants with altered or dysfunctional meristematic tissues. It is very likely that meristematic tissues are important to the *mpg1* phenotype.

It is difficult to speculate which genes are responsible for the phenotype of *mpg1* however we can use the information generated in this study to hypothesize what might be driving

individual characteristics observed in *mpg1*. Further independent validation of these genes will be necessary to gauge the effects elicited by these genes. Because *MPG1*'s nature of expression in the mutant is not well understood other criterion investigating affected genes will need to be taken into account. The expression of these genes might be dependent on their temporal expression, or within specific tissue types.

Future considerations for the continuation of this study

This transcriptomic analysis was conducted under a variety of RNA-seq pipeline platforms and one was selected for discussion in this manuscript. This single pipeline was chosen because it had good overlap between other pipelines for calling DEG, and also produced the smallest number of DEG's compared to the others. It was therefore selected to reduce the probability of false-positive or false-negative calls (Dobin et al., 2013, Conesa et al., 2016, Varet et al., 2016, Seyednasrollah et al., 2015). By using only this platform for analysis, other genes of interest that are differentially expressed, however that were not deemed significant might still be of interest. RNA-seq is a good tool for further hypothesis generation, however it has its limitations. Analysis of this particular mutant is difficult because of its altered rate of development compared to wild-type plants. Tissue selection and timepoint are critical factors to consider when analyzing RNA-seq experiments. Future assessment of other tissues or timepoints could further illuminate what is occurring in *mpg1*. Because notable changes are occurring in growth, evaluation of meristematic tissue would be of particular interest.

If indeed the process of flowering at the transcriptional level begins around or prior to the 32 days post-planting timepoint, than this RNA-seq evaluation might not be sufficient in assessing the global gene expression differences occurring between *mpg1* and wild-type plants. Earlier evaluation in development may be necessary to capture these differences. This being said,

genes affected down-stream in the developmental cycle might also be crucial in recognizing the molecular mechanisms at play and necessary to understand what is occurring in the *mpg1* mutant. Because the mutant lines are expressing this transcription factor out of its native context, other variables must be taken into consideration when assessing the function of overexpression of this gene. Many transcription factors bind different DNA-binding domains under different circumstances. Their function can largely depend on other transcription factors and molecular mechanisms that arise from response to stimulus, or activate for growth and development. *MPG1* might have a variety of different functional aspects spatiotemporally, as expression is not found constitutively in all tissue types.

Dissection and assessment of the entire developmental shift could better allow us to assess *mpg1* side-by-side with wild-type plants. Maybe sampling and assessment of both *mpg1* and WT-ns, will need to be done at different timepoints that represent the closest physiological developmental stage. Once a comprehensive understanding of developmental timepoints is made, we might be able to better assess these plants and what is occurring differently throughout the progression of growth.

Potentially, a better way to assess *mpg1* could be through the use of loss-of-flowering mutants and/or inducible promoters. Removing the complication of flowering transition in *mpg1* would allow for more appropriate assessment of the mutant side-by-side with wild-type. The use of these molecular tools could help tailor experiments to assess these plants at similar developmental and temporal windows as well as allow for a better understanding of what ectopic expression of *MPG1* is doing across the entirety of the plants life cycle. This could also allow for better conditions to assess the effects of stressors on *mpg1*, as plants would be in a more amenable and comparable stage of development.

Many flowering mutants mentioned above present with similar pleiotropic phenotypes to *mpg1*. So, perhaps the down-regulation the floral regulators seen in *mpg1* are actually the direct result of the phenotype, and that assessment of *mpg1* crossed with loss-of-flowering mutants will prove uneventful.

While assessment of meristematic tissues and cell number/size within various plant tissues is important, we remain unclear what phenotypic differences occur in below-ground tissues. During harvests, occasionally root tissue was sampled for gene expression analyses, however it was never measured or characterized in any further way. Root tissues in *mpg1* appeared to be in higher abundance than wild-type plants during harvesting. There is a possibility root architecture or root accumulation is also altered in *mpg1* plants. These characteristics must be explored to further rationalize and evaluate reasons for *mpg1*'s phenotype.

Recapitulation of the phenotype will be necessary to validate the functionality of ectopic expression of *MPG1* resulting in the phenotype. Constitutive overexpression of *MPG1* resulted in T_0 plants with some notable characteristics similar to *mpg1*, however also gave way to the inability of T_1 progeny containing T-DNA expression cassettes overexpressing *MPG1* to germinate. This increases our confidence that *MPG1* is indeed the cause of the phenotype in *mpg1*, however under constitutive overexpression results in a deleterious phenotype suggesting that a specific pattern and/or a certain degree of expression is necessary for generating functional plants presenting with the phenotype. As previously stated, this is complex in that numerous possibilities exist to explain the nature of *MPG1*'s expression in *mpg1*. Remnant sequence coding for transcriptional regulators remain in the T-DNA insertion proximal to *MPG1*. A genomic deletion event occurred at the site of the T-DNA integration. Roughly 7kb of genomic

sequence separates the T-DNA from *MPG1.* Any number of factors surrounding these sequences neighboring *MPG1* could be influencing the nature of *MPG1*'s expression in the mutant. Several techniques could be explored to clarify this matter. CRISPR technology could be implemented in the mutant by removing the T-DNA insertion to see if removal of the phenotype can be achieved. Additionally, different segments of the T-DNA could be removed to gauge if certain components are necessary for generating the phenotype. Similarly, CRISPR technology could also be used to delete proximal segments of genomic DNA to see if there are functional regions aiding the regulation of *MPG1* in the mutant. Further verification could also be implored by generating new transgenic plants containing altered versions of the surrounding sequences driving *MPG1* to see if the phenotype can be generated again. If sequence elements driving expression of *MPG1* in the mutant are better understood further complementation or application could be explored by gauging the effect of tissue- and/or temporal-specific expression one or many of the differentially expressed genes found in this study.

BIBLIOGRAPHY

ABDALLAH, M., DUBOUSSET, L., MEURIOT, F., ETIENNE, P., AVICE, J. C. & OURRY, A. 2010. Effect of mineral sulphur availability on nitrogen and sulphur uptake and remobilization during the vegetative growth of Brassica napus L. *J Exp Bot,* 61**,** 2635-46.

ABIKO, M., MAEDA, H., TAMURA, K., HARA-NISHIMURA, I. & OKAMOTO, T. 2013. Gene expression profiles in rice gametes and zygotes: identification of gamete-enriched genes and up- or down-regulated genes in zygotes after fertilization. *J Exp Bot,* 64**,** 1927-40.

ABRAHAM, E., FARKAS, I., DARULA, Z., VARGA, E., LUKACS, N., AYAYDIN, F., MEDZIHRADSZKY, K. F., DOMBRADI, V., DUDITS, D. & HORVATH, G. V. 2015. The B" regulatory subunit of protein phosphatase 2A mediates the dephosphorylation of rice retinoblastoma -related protein-1. *Plant Mol Biol,* 87**,** 125-141.

ACHARD, P., BAGHOUR, M., CHAPPLE, A., HEDDEN, P., VAN DER STRAETEN, D., GENSCHIK, P., MORITZ, T. & HARBERD, N. P. 2007. The plant stress hormone ethylene controls floral transition via DELLA-dependent regulation of floral meristem-identity genes. *Proc Natl Acad Sci U S A,* 104**,** 6484-9.

ADAMS, D. O. & YANG, S. F. 1979. Ethylene biosynthesis: Identification of 1-aminocyclopropane-1-carboxylic acid as an intermediate in the conversion of methionine to ethylene. *Proc Natl Acad Sci U S A,* 761**,** 170-174.

AGARRWAL, R., PADMAKUMARI, A. P., BENTUR, J. S. & NAIR, S. 2016. Metabolic and transcriptomic changes induced in host during hypersensitive response mediated resistance in rice against the Asian rice gall midge. *Rice (N Y),* 9**,** 5.

AGARWAL, P. K. & JHA, B. 2010. Transcription factors in plants and ABA dependent and independent abiotic stress signalling. *Biologia Plantarum,* 54**,** 201-212.

AGRAWAL, G. K., RAKWAL, R. & IWAHASHI, H. 2002. Isolation of novel rice (Oryza sativa L.) multiple stress responsive MAP kinase gene, OsMSRMK2, whose mRNA accumulates rapidly in response to environmental cues. *Biochem Biophys Res Commun,* 294**,** 1009-16.

ALAM, M. M., TANAKA, T., NAKAMURA, H., ICHIKAWA, H., KOBAYASHI, K., YAENO, T., YAMAOKA, N., SHIMOMOTO, K., TAKAYAMA, K., NISHINA, H. & NISHIGUCHI, M. 2015. Overexpression of a rice heme activator protein gene (OsHAP2E) confers resistance to pathogens, salinity and drought, and increases photosynthesis and tiller number. *Plant Biotechnol J,* 13**,** 85-96.

ALMUTAIRI, Z. M. & SADDER, M. T. 2014. Cloning and Expression Profiling of the Polycomb Gene, Retinoblastoma-related Protein from Tomato Solanum lycopersicum L. *Evol Bioinform Online,* 10**,** 177-85.

ALONSO, J. M., HIRAYAMA, T., ROMAN, G., NOURIZADEH, S. & ECKER, J. R. 1999. EIN2, a bifunctional transducer of ethylene and stress responses in Arabidopsis. *Science,* 284**,** 2148-52.

AMBAVARAM, M. M., BASU, S., KRISHNAN, A., RAMEGOWDA, V., BATLANG, U., RAHMAN, L., BAISAKH, N. & PEREIRA, A. 2014. Coordinated regulation of photosynthesis in rice increases yield and tolerance to environmental stress. *Nat Commun,* 5**,** 5302.

ANDERS, S., PYL, P. T. & HUBER, W. 2015. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics,* 31**,** 166-9.

ANDREWS, S. 2010. FastQC: a quality control tool for high throughput sequence data. *Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc*.

AOHARA, T., KOTAKE, T., KANEKO, Y., TAKATSUJI, H., TSUMURAYA, Y. & KAWASAKI, S. 2009. Rice BRITTLE CULM 5 (BRITTLE NODE) is involved in secondary cell wall formation in the sclerenchyma tissue of nodes. *Plant Cell Physiol,* 50**,** 1886-97.

ARGUESO, C. T., HANSEN, M. & KIEBER, J. J. 2007. Regulation of Ethylene Biosynthesis. *J Plant Growth Regul,* 26**,** 92-105.

ASHIKARI, M., SAKAKIBARA, H., LIN, S., YAMAMOTO, T., TAKASHI, T., NISHIMURA, A., ANGELES, E. R., QIAN, Q., KITANO, H. & MATSUOKA, M. 2005. Cytokinin oxidase regulates rice grain production. *Science,* 309**,** 741-5.

ASHIKARI, M., WU, J., YANO, M., SASAKI, T. & YOSHIMURA, A. 1999. Rice gibberellin-insensitive dwarf mutant gene Dwarf 1 encodes the alpha-subunit of GTP-binding protein. *Proc Natl Acad Sci U S A,* 96**,** 10284-9.

AULER, P. A., AMARAL, M. N. D., ROSSATTO, T., VIGHI, I. L., BENITEZ, L. C., DA MAIA, L. C. & BRAGA, E. J. B. 2019. Research Article Expression of transcription factors involved with dehydration in contrasting rice genotypes submitted to different levels of soil moisture. *Genet Mol Res,* 18.

BAI, A. N., LU, X. D., LI, D. Q., LIU, J. X. & LIU, C. M. 2016. NF-YB1-regulated expression of sucrose transporters in aleurone facilitates sugar loading to rice endosperm. *Cell Res,* 26**,** 384-8.

BARRERO, J. M., GONZALEZ-BAYON, R., DEL POZO, J. C., PONCE, M. R. & MICOL, J. L. 2007. INCURVATA2 encodes the catalytic subunit of DNA Polymerase alpha and interacts with genes involved in chromatin-mediated cellular memory in Arabidopsis thaliana. *Plant Cell,* 19**,** 2822-38.

BARROCO, R. M., PERES, A., DROUAL, A. M., DE VEYLDER, L., NGUYEN LE, S. L., DE WOLF, J., MIRONOV, V., PEERBOLTE, R., BEEMSTER, G. T., INZE, D., BROEKAERT, W. F. & FRANKARD, V. 2006. The cyclin-dependent kinase inhibitor Orysa;KRP1 plays an important role in seed development of rice. *Plant Physiol,* 142**,** 1053-64.

BARRY, C. S. & GIOVANNONI, J. J. 2007. Ethylene and Fruit Ripening. *J Plant Growth Regul,* 26**,** 143-159.

BECRAFT, P. W., KANG, S. H. & SUH, S. G. 2001. The Maize CRINKLY4 Receptor Kinase Controls a Cell-Autonomous Differentiation Response. *Plant Physiol,* 127**,** 486-496.

BELLO, B. K., HOU, Y., ZHAO, J., JIAO, G., WU, Y., LI, Z., WANG, Y., TONG, X., WANG, W., YUAN, W., WEI, X. & ZHANG, J. 2018. NF-YB1-YC12-bHLH144 complex directly activates Wx to regulate grain quality in rice (Oryza sativa L.). *Plant Biotechnol J.*

BINDER, B. M., RODRIGUEZ, F. I. & BLEECKER, A. B. 2010. The copper transporter RAN1 is essential for biogenesis of ethylene receptors in Arabidopsis. *J Biol Chem,* 285**,** 37263-70.

BISSON, M. M., BLECKMANN, A., ALLEKOTTE, S. & GROTH, G. 2009. EIN2, the central regulator of ethylene signalling, is localized at the ER membrane where it interacts with the ethylene receptor ETR1. *Biochem J,* 424**,** 1-6.

BISSON, M. M. & GROTH, G. 2010. New insight in ethylene signaling: autokinase activity of ETR1 modulates the interaction of receptors and EIN2. *Mol Plant,* 3**,** 882-9.
BLEECKER, A. B., ESTELLE, M. A., SOMERVILLE, C. & KENDE, H. 1988. Insensitivity to Ethylene Conferred by a Dominant Mutation in Arabidopsis thaliana. *Science,* 241**,** 1086-9.
BOSCH, M., MAYER, C. D., COOKSON, A. & DONNISON, I. S. 2011. Identification of genes involved in cell wall biogenesis in grasses by differential gene expression profiling of elongating and non-elongating maize internodes. *J Exp Bot,* 62**,** 3545-61.
BOUDOLF, V., VLIEGHE, K., BEEMSTER, G. T., MAGYAR, Z., TORRES ACOSTA, J. A., MAES, S., VAN DER SCHUEREN, E., INZE, D. & DE VEYLDER, L. 2004. The plant-specific cyclin-dependent kinase CDKB1;1 and transcription factor E2Fa-DPa control the balance of mitotically dividing and endoreduplicating cells in Arabidopsis. *Plant Cell,* 16**,** 2683-92.
BOZKURT, O., HAKKI, E. E. & AKKAYA, M. S. 2007. Isolation and sequence analysis of wheat NBS-LRR type disease resistance gene analogs using degenerate PCR primers. *Biochem Genet,* 45**,** 469-486.
BRASIL, J. N., COSTA, C. N. M., CABRAL, L. M., FERREIRA, P. C. G. & HEMERLY, A. S. 2017. The plant cell cycle: Pre-Replication complex formation and controls. *Genet Mol Biol,* 40**,** 276-291.
BROEKGAARDEN, C., SNOEREN, T. A., DICKE, M. & VOSMAN, B. 2011. Exploiting natural variation to identify insect-resistance genes. *Plant Biotechnol J,* 9**,** 819-25.
BROWN, D. M., ZEEF, L. A., ELLIS, J., GOODACRE, R. & TURNER, S. R. 2005. Identification of novel genes in Arabidopsis involved in secondary cell wall formation using expression profiling and reverse genetics. *Plant Cell,* 17**,** 2281-95.
BROWSE, J. 2009. The power of mutants for investigating jasmonate biosynthesis and signaling. *Phytochemistry,* 70**,** 1539-46.
BURG, S. P. 1968. Ethylene, plant senescence and abscission. *Plant Physiol,* 43**,** 1503-11.
BURG, S. P. 1973. Ethylene in plant growth. *Proc Natl Acad Sci U S A,* 70**,** 591-7.
BUTTNER, M. & SINGH, K. B. 1997. Arabidopsis thaliana ethylene-responsive element binding protein (AtEBP), an ethylene-inducible, GCC box DNA-binding protein interacts with an ocs element binding protein. *Proc Natl Acad Sci U S A,* 94**,** 5961-6.
CAI, H., XIE, W. & LIAN, X. 2013. Comparative analysis of differentially expresed genes in rice under nitrogen and phosphorus starvation stress conditions. *Plant Mol Biol,* 31**,** 160-173.
CALDERAN-RODRIGUES, M. J., GUIMARAES FONSECA, J., DE MORAES, F. E., VAZ SETEM, L., CARMANHANIS BEGOSSI, A. & LABATE, C. A. 2019. Plant Cell Wall Proteomics: A Focus on Monocot Species, Brachypodium distachyon, Saccharum spp. and Oryza sativa. *Int J Mol Sci,* 20.
CAO, W. L., CHU, R. Z., ZHANG, Y., LUO, J., SU, Y. Y., XIE, L. J., ZHANG, H. S., WANG, J. F. & BAO, Y. M. 2015. OsJAMyb, a R2R3-type MYB transcription factor, enhanced blast resistance in transgenic rice. *Physiol Mol Plant P,* 92**,** 154-160.
CARRARI, F., COLL-GARCIA, D., SCHAUER, N., LYTOVCHENKO, A., PALACIOS-ROJAS, N., BALBO, I., ROSSO, M. & FERNIE, A. R. 2005. Deficiency of a plastidial adenylate kinase in Arabidopsis results in elevated photosynthetic amino acid biosynthesis and enhanced growth. *Plant Physiol,* 137**,** 70-82.

CHEN, C., BEGCY, K., LIU, K., FOLSOM, J. J., WANG, Z., ZHANG, C. & WALIA, H. 2016. Heat stress yields a unique MADS box transcription factor in determining seed size and thermal sensitivity. *Plant Physiol,* 171**,** 606-22.

CHEN, H., LAI, Z., SHI, J., XIAO, Y., CHEN, Z. & XU, X. 2010. Roles of arabidopsis WRKY18, WRKY40 and WRKY60 transcription factors in plant responses to abscisic acid and abiotic stress. *BMC Plant Biol,* 10**,** 281.

CHEN, S., LI, X. Q., ZHAO, A., WANG, L., LI, X., SHI, Q., CHEN, M., GUO, J., ZHANG, J., QI, D. & LIU, G. 2009. Genes and pathways induced in early response to defoliation in rice seedlings. *Curr Issues Mol Biol,* 11**,** 81-100.

CHEN, T. H. H. & MURATA, N. 2004. Enhancement of tolerance of abiotic stress by metabolic engineering of betaines and other compatible solutes. *Curr Opin Plant Biol,* 5**,** 250-257.

CHEN, W., YU, X. H., ZHANG, K., SHI, J., DE OLIVEIRA, S., SCHREIBER, L., SHANKLIN, J. & ZHANG, D. 2011. Male Sterile2 encodes a plastid-localized fatty acyl carrier protein reductase required for pollen exine development in Arabidopsis. *Plant Physiol,* 157**,** 842-53.

CHEN, X., LU, S., WANG, Y., ZHANG, X., LV, B., LUO, L., XI, D., SHEN, J., MA, H. & MING, F. 2015. OsNAC2 encoding a NAC transcription factor that affects plant height through mediating the gibberellic acid pathway in rice. *Plant J,* 82**,** 302-14.

CHINI, A., FONSECA, S., FERNANDEZ, G., ADIE, B., CHICO, J. M., LORENZO, O., GARCIA-CASADO, G., LOPEZ-VIDRIERO, I., LOZANO, F. M., PONCE, M. R., MICOL, J. L. & SOLANO, R. 2007. The JAZ family of repressors is the missing link in jasmonate signalling. *Nature,* 448**,** 666-71.

CHINNUSAMY, V. & ZHU, J. K. 2009. Epigenetic regulation of stress responses in plants. *Curr Opin Plant Biol,* 12**,** 133-9.

CHO, J. I., LEE, S. K., KO, S., KIM, H. K., JUN, S. H., LEE, Y. H., BHOO, S. H., LEE, K. W., AN, G., HAHN, T. R. & JEON, J. S. 2005. Molecular cloning and expression analysis of the cell-wall invertase gene family in rice (Oryza sativa L.). *Plant Cell Rep,* 24**,** 225-36.

CHO, L. H., YOON, J. & AN, G. 2017. The control of flowering time by environmental factors. *Plant J,* 90**,** 708-719.

CHO, L. H., YOON, J., PASRIGA, R. & AN, G. 2016. Homodimerization of Ehd1 Is Required to Induce Flowering in Rice. *Plant Physiol,* 170**,** 2159-71.

CHO, S. H., YOO, S. C., ZHANG, H., PANDEYA, D., KOH, H. J., HWANG, J. Y., KIM, G. T. & PAEK, N. C. 2013. The rice narrow leaf2 and narrow leaf3 loci encode WUSCHEL-related homeobox 3A (OsWOX3A) and function in leaf, spikelet, tiller and lateral root development. *New Phytol,* 198**,** 1071-84.

CLARK, K. L., LARSEN, P. B., WANG, X. & CHANG, C. 1998. Association of the Arabidopsis CTR1 Raf-like kinase with the ETR1 and ERS ethylene receptors. *Proc Natl Acad Sci U S A,* 95**,** 5401-6.

CLARK, S. E., WILLIAMS, R. W. & MEYEROWITZ, E. M. 1997. The CLAVATA1 gene encodes a putative receptor kinase that controls shoot and floral meristem size in Arabidopsis. *Cell,* 89**,** 575-585.

CLOUSE, S. D. 1996. Molecular genetic studies confirm the role of brassinosteroids in plant growth and development. *Plant J,* 10**,** 1-8.

COLEMAN, H. D., YAN, J. & MANSFIELD, S. D. 2009. Sucrose synthase affects carbon partitioning to increase cellulose production and altered cell wall ultrastructure. *Proc Natl Acad Sci U S A,* 106**,** 13118-23.

CONESA, A., MADRIGAL, P., TARAZONA, S., GOMEZ-CABRERO, D., CERVERA, A., MCPHERSON, A., SZCZESNIAK, M. W., GAFFNEY, D. J., ELO, L. L., ZHANG, X. & MORTAZAVI, A. 2016. A survey of best practices for RNA-seq data analysis. *Genome Biol,* 17**,** 13.

CORDOBA, E., SALMI, M. & LEON, P. 2009. Unravelling the regulatory mechanisms that modulate the MEP pathway in higher plants. *J Exp Bot,* 60**,** 2933-43.

CZEREDNIK, A., BUSSCHER, M., ANGENENT, G. C. & DE MAAGD, R. A. 2015. The cell size distribution of tomato fruit can be changed by overexpression of CDKA1. *Plant Biotechnol J,* 13**,** 259-68.

CZEREDNIK, A., BUSSCHER, M., BIELEN, B. A., WOLTERS-ARTS, M., DE MAAGD, R. A. & ANGENENT, G. C. 2012. Regulation of tomato fruit pericarp development by an interplay between CDKB and CDKA1 cell cycle genes. *J Exp Bot,* 63**,** 2605-17.

DA MAIA, L. C., CADORE, P. R. B., BENITEZ, L. C., DANIELOWSKI, R., BRAGA, E. J. B., FAGUNDES, P. R. R., MAGALHÃES, A. M. & COSTA DE OLIVEIRA, A. 2017. Transcriptome profiling of rice seedlings under cold stress. *Funct Plant Biol,* 44.

DAI, X., DING, Y., TAN, L., FU, Y., LIU, F., ZHU, Z., SUN, X., SUN, X., GU, P., CAI, H. & SUN, C. 2012. LHD1, an allele of DTH8/Ghd8, controls late heading date in common wild rice (Oryza rufipogon). *J Integr Plant Biol,* 54**,** 790-9.

DAI, X., YOU, C., CHEN, G., LI, X., ZHANG, Q. & WU, C. 2011. OsBC1L4 encodes a COBRA-like protein that affects cellulose synthesis in rice. *Plant Mol Biol,* 75**,** 333-345.

DAI, X., YOU, C., WANG, L., CHEN, G., ZHANG, Q. & WU, C. 2009. Molecular characterization, expression pattern, and function analysis of the OsBC1L family in rice. *Plant Mol Biol,* 71**,** 469-81.

DALIRI, M. S., EFTEKHARI, A., MOBASSER, H. R., TARI, D. B. & PORKALHOR, H. 2009. Effect of cutting time and cutting height on yield and yield components of ratoon rice (Tarom Langrodi Variety). *Asian J Plant Sci,* 8**,** 89-91.

DAO, T. T., LINTHORST, H. J. & VERPOORTE, R. 2011. Chalcone synthase and its functions in plant resistance. *Phytochem Rev,* 10**,** 397-412.

DAY, R. B., SHIBUYA, N. & MINAMI, E. 2003. Identification and characterization of two new members of the GRAS gene family in rice responsive to N-acetylchitooligosaccharide elicitor. *Biochim Biophys Acta,* 1625**,** 261-8.

DEAN, J. F. D. & ERIKSSON, K.-E. L. 1994. Laccase and the Deposition of Lignin in Vascular Plants. *Holzforschung,* 48**,** 21-33.

DELLO IOIO, R., LINHARES, F. S., SCACCHI, E., CASAMITJANA-MARTINEZ, E., HEIDSTRA, R., COSTANTINO, P. & SABATINI, S. 2007. Cytokinins determine Arabidopsis root-meristem size by controlling cell differentiation. *Curr Biol,* 17**,** 678-82.

DENG, C., YE, H., FAN, M., PU, T. & YAN, J. 2017. The rice transcription factors OsICE confer enhanced cold tolerance in transgenic Arabidopsis. *Plant Signal Behav,* 12**,** e1316442.

DEVOTO, A., HARTMANN, H. A., PIFFANELLI, P., ELLIOTT, C., SIMMONS, C., TARAMINO, G., GOH, C. S., COHEN, F. E., EMERSON, B. C., SCHULZE-LEFERT, P. & PANSTRUGA, R. 2003. Molecular phylogeny and evolution of the plant-specific seven-transmembrane MLO family. *J Mol Evol,* 56**,** 77-88.

DEWITTE, W., RIOU-KHAMLICHI, C., SCOFIELD, S., HEALY, J. M., JACQMARD, A., KILBY, N. J. & MURRAY, J. A. 2003. Altered cell cycle distribution, hyperplasia, and

inhibited differentiation in Arabidopsis caused by the D-type cyclin CYCD3. *Plant Cell,* 15**,** 79-92.
DIVYA, D., NAIR, S. & BENTUR, J. S. 2018. Expression profiles of key genes involved in rice gall midge interactions reveal diversity in resistance pathways. *Curr Sci,* 115**,** 74-82.
DOBIN, A., DAVIS, C. A., SCHLESINGER, F., DRENKOW, J., ZALESKI, C., JHA, S., BATUT, P., CHAISSON, M. & GINGERAS, T. R. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics,* 29**,** 15-21.
DODD, A. N., SALATHIA, N., HALL, A., KEVEI, E., TOTH, R., NAGY, F., HIBBERD, J. M., MILLAR, A. J. & WEBB, A. A. 2005. Plant circadian clocks increase photosynthesis, growth, survival, and competitive advantage. *Science,* 309**,** 630-3.
DOEBLEY, J. & LUKENS, L. 1998. Transcriptional regulators and the evolution of plant form. *Plant Cell,* 10**,** 1075-82.
DOERKS, T., COPLEY, R. R., SCHULTZ, J., PONTING, C. P. & BORK, P. 2002. Systematic identification of novel protein domain families associated with nuclear functions. *Genome Res,* 12**,** 47-56.
DOROSHENK, K. A., CROFTS, A. J., WASHIDA, H., SATOH-CRUZ, M., CROFTS, N., SUGINO, A., OKITA, T. W., MORRIS, R. T., WYRICK, J. J., FUKUDA, M., KUMAMARU, T. & SATOH, H. 2010. Characterization of the rice glup4 mutant suggests a role for the small GTPase Rab5 in the biosynthesis of carbon and nitrogen storage reserves in developing endosperm. *Breeding Sci,* 60**,** 556-567.
DUDITS, D., ÁBRAHÁM, E., MISKOLCZI, P., AYAYDIN, F., BILGIN, M. & HORVÁTH, G. V. 2011. Cell-cycle control as a target for calcium, hormonal and developmental signals: the role of phosphorylation in the retinoblastoma-centred pathway. *Annals of Botany,* 107**,** 1193-1202.
ENDO, M., NAKAYAMA, S., UMEDA-HARA, C., OHTSUKI, N., SAIKA, H., UMEDA, M. & TOKI, S. 2012. CDKB2 is involved in mitosis and DNA damage response in rice. *Plant J,* 69**,** 967-77.
ENDO-HIGASHI, N. & IZAWA, T. 2011. Flowering time genes Heading date 1 and Early heading date 1 together control panicle development in rice. *Plant Cell Physiol,* 52**,** 1083-94.
ESPINAS, N. 2018. *rCBP-dependent regulation in rice innate immunity.* Doctor of Philosophy, Okinawa Institute of Science and Technology Graduate University.
ESTEVEZ, J. M., CANTERO, A., REINDL, A., REICHLER, S. & LEON, P. 2001. 1-Deoxy-D-xylulose-5-phosphate synthase, a limiting enzyme for plastidic isoprenoid biosynthesis in plants. *J Biol Chem,* 276**,** 22901-9.
FAGERIA, N. K., BALIGAR, V. C. & LI, Y. C. 2008. The Role of Nutrient Efficient Plants in Improving Crop Yields in the Twenty First Century. *J Plant Nutr,* 31**,** 1121-1157.
FARUQ, G., TAHA, R. & PRODHAN, Z. 2014. Rice Ratoon Crop: A Sustainable Rice Production System for Tropical Hill Agriculture. *Sustainability,* 6**,** 5785-5800.
FENG, S., JACOBSEN, S. E. & REIK, W. 2010. Epigenetic reprogramming in plant and animal development. *Science,* 330**,** 622-7.
FERGUSON, B. J. & BEVERIDGE, C. A. 2009. Roles for auxin, cytokinin, and strigolactone in regulating shoot branching. *Plant Physiol,* 149**,** 1929-44.
FERRARIO, S., IMMINK, R. G. & ANGENENT, G. C. 2004. Conservation and diversity in flower land. *Curr Opin Plant Biol,* 7**,** 84-91.

FICKLIN, S. P., LUO, F. & FELTUS, F. A. 2010. The association of multiple interacting genes with specific phenotypes in rice using gene coexpression networks. *Plant Physiol,* 154**,** 13-24.
FLUHR, R., MATTOO, A. K. & DILLEY, D. R. 1996. Ethylene - Biosynthesis and perception. *CRC Crit Rev Plant Sci,* 15**,** 479-523.
FORNARA, F., PARENICOVA, 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 Physiol,* 135**,** 2207-19.
FRANKE, R. B., DOMBRINK, I. & SCHREIBER, L. 2012. Suberin goes genomics: use of a short living plant to investigate a long lasting polymer. *Front Plant Sci,* 3**,** 4.
FREDEEN, A. L. 1988. Influence of phosphorus nutrition on growth and and carbon partitioning in Blycine max. *Plant Physiol*.
FUJIMOTO, S. Y., OHTA, M., USUI, A., SHINSHI, H. & OHME-TAKAGI, M. 2000. Arabidopsis ethylene-responsive element binding factors act as transcriptional activators or repressors of GCC box-mediated gene expression. *Plant Cell,* 12**,** 393-404.
FUJINO, K., MATSUDA, Y., OZAWA, K., NISHIMURA, T., KOSHIBA, T., FRAAIJE, M. W. & SEKIGUCHI, H. 2008. NARROW LEAF 7 controls leaf shape mediated by auxin in rice. *Mol Genet Genomics,* 279**,** 499-507.
FURUKAWA, T., SAKAGUCHI, N. & SHIMADA, H. 2006. Two OsGASR genes, rice GAST homologue genes that are abundant in proliferating tissues, show different expression patterns in developing panicles. *Genes Genet Syst,* 81**,** 171-180.
GALLIE, D. R. 2015. Ethylene receptors in plants - why so much complexity? *F1000Prime Rep,* 7**,** 39.
GARG, A. K., KIM, J. K., OWENS, T. G., RANWALA, A. P., CHOI, Y. D., KOCHIAN, L. V. & WU, R. J. 2002. Trehalose accumulation in rice plants confers high tolerance levels to different abiotic stresses. *Proc Natl Acad Sci U S A,* 99**,** 15898-903.
GIGOLASHVILI, T. & KOPRIVA, S. 2014. Transporters in plant sulfur metabolism. *Front Plant Sci,* 5**,** 442.
GILL, S. S. & TUTEJA, N. 2010. Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. *Plant Physiol Biochem,* 48**,** 909-30.
GIRIN, T., SOREFAN, K. & OSTERGAARD, L. 2009. Meristematic sculpting in fruit development. *J Exp Bot,* 60**,** 1493-502.
GOMEZ, L. D., BAUD, S., GILDAY, A., LI, Y. & GRAHAM, I. A. 2006. Delayed embryo development in the ARABIDOPSIS TREHALOSE-6-PHOSPHATE SYNTHASE 1 mutant is associated with altered cell wall structure, decreased cell division and starch accumulation. *Plant J,* 46**,** 69-84.
GONZALEZ, N., GEVAUDANT, F., HERNOULD, M., CHEVALIER, C. & MOURAS, A. 2007. The cell cycle-associated protein kinase WEE1 regulates cell size in relation to endoreduplication in developing tomato fruit. *Plant J,* 51**,** 642-55.
GU, B., ZHOU, T., LUO, J., LIU, H., WANG, Y., SHANGGUAN, Y., ZHU, J., LI, Y., SANG, T., WANG, Z. & HAN, B. 2015. An-2 Encodes a Cytokinin Synthesis Enzyme that Regulates Awn Length and Grain Production in Rice. *Mol Plant,* 8**,** 1635-50.
GU, Y. Q., YANG, C., THARA, V. K., ZHOU, J. & MARTIN, G. B. 2000. Pti4 is induced by ethylene and salicylic acid, and its product is phosphorylated by the Pto kinase. *Plant Cell,* 12**,** 771-785.

GUO, J., SONG, J., WANG, F. & ZHANG, X. S. 2007. Genome-wide identification and expression analysis of rice cell cycle genes. *Plant Mol Biol,* 64**,** 349-60.
GUO, J., WANG, F., SONG, J., SUN, W. & ZHANG, X. S. 2010. The expression of Orysa;CycB1;1 is essential for endosperm formation and causes embryo enlargement in rice. *Planta,* 231**,** 293-303.
GUO, L., HO, C. M., KONG, Z., LEE, Y. R., QIAN, Q. & LIU, B. 2009. Evaluating the microtubule cytoskeleton and its interacting proteins in monocots by mining the rice genome. *Ann Bot,* 103**,** 387-402.
GUO, Q., YOSHIDA, Y., MAJOR, I. T., WANG, K., SUGIMOTO, K., KAPALI, G., HAVKO, N. E., BENNING, C. & HOWE, G. A. 2018. JAZ repressors of metabolic defense promote growth and reproductive fitness in Arabidopsis. *Proc Natl Acad Sci U S A,* 115**,** E10768-E10777.
GUPTA, P., NAITHANI, S., TELLO-RUIZ, M. K., CHOUGULE, K., D'EUSTACHIO, P., FABREGAT, A., JIAO, Y., KEAYS, M., LEE, Y. K., KUMARI, S., MULVANEY, J., OLSON, A., PREECE, J., STEIN, J., WEI, S., WEISER, J., HUERTA, L., PETRYSZAK, R., KERSEY, P., STEIN, L. D., WARE, D. & JAISWAL, P. 2016. Gramene Database: Navigating Plant Comparative Genomics Resources. *Curr Plant Biol,* 7-8**,** 10-15.
GUZMAN, P. & ECKER, J. R. 1990. Exploiting the triple response of Arabidopsis to identify ethylene-related mutants. *Plant Cell,* 2**,** 513-23.
HAMILTON, A. J., BOUZAYEN, M. & GRIERSON, D. 1991. Identification of a tomato gene for the ethylene-forming enzyme by expression in yeast. *Proc Natl Acad Sci U S A,* 88**,** 7434-7.
HART, N. 2018. *Rice fields in China double yields by "ratooning"* [Online]. International Atomic Energy Agency. Available: https://www.iaea.org/newscenter/news/rice-fields-in-china-double-yields-by-ratooning [Accessed].
HEMERLY, A., ENGLER JDE, A., BERGOUNIOUX, C., VAN MONTAGU, M., ENGLER, G., INZE, D. & FERREIRA, P. 1995. Dominant negative mutants of the Cdc2 kinase uncouple cell division from iterative plant development. *EMBO J,* 14**,** 3925-36.
HIRANO, H., HARASHIMA, H., SHINMYO, A. & SEKINE, M. 2008. Arabidopsis RETINOBLASTOMA-RELATED PROTEIN 1 is involved in G1 phase cell cycle arrest caused by sucrose starvation. *Plant Mol Biol,* 66**,** 259-75.
HIRANO, K., AYA, K., KONDO, M., OKUNO, A., MORINAKA, Y. & MATSUOKA, M. 2012. OsCAD2 is the major CAD gene responsible for monolingol biosynthesis in rice culm. *Plant Cell Rep,* 31**,** 91-101.
HIRANO, K., KOTAKE, T., KAMIHARA, K., TSUNA, K., AOHARA, T., KANEKO, Y., TAKATSUJI, H., TSUMURAYA, Y. & KAWASAKI, S. 2010. Rice BRITTLE CULM 3 (BC3) encodes a classical dynamin OsDRP2B essential for proper secondary cell wall synthesis. *Planta,* 232**,** 95-108.
HIROSE, N., MAKITA, N., KOJIMA, M., KAMADA-NOBUSADA, T. & SAKAKIBARA, H. 2007. Overexpression of a type-A response regulator alters rice morphology and cytokinin metabolism. *Plant Cell Physiol,* 48**,** 523-39.
HOBO, T., SUWABE, K., AYA, K., SUZUKI, G., YANO, K., ISHIMIZU, T., FUJITA, M., KIKUCHI, S., HAMADA, K., MIYANO, M., FUJIOKA, T., KANEKO, F., KAZAMA, T., MIZUTA, Y., TAKAHASHI, H., SHIONO, K., NAKAZONO, M., TSUTSUMI, N., NAGAMURA, Y., KURATA, N., WATANABE, M. & MATSUOKA, M. 2008. Various

spatiotemporal expression profiles of anther-expressed genes in rice. *Plant Cell Physiol,* 49**,** 1417-28.
HONG, Z., UEGUCHI-TANAKA, M., UMEMURA, K., UOZU, S., FUJIOKA, S., TAKATSUTO, S., YOSHIDA, S., ASHIKARI, M., KITANO, H. & MATSUOKA, M. 2003. A rice brassinosteroid-deficient mutant, ebisu dwarf (d2), is caused by a loss of function of a new member of cytochrome P450. *Plant Cell,* 15**,** 2900-10.
HOU, X., DING, L. & YU, H. 2013. Crosstalk between GA and JA signaling mediates plant growth and defense. *Plant Cell Rep,* 32**,** 1067-1074.
HRUZ, T., LAULE, O., SZABO, G., WESSENDORP, F., BLEULER, S., OERTLE, L., WIDMAYER, P., GRUISSEM, W. & ZIMMERMANN, P. 2008. Genevestigator v3: a reference expression database for the meta-analysis of transcriptomes. *Adv Bioinformatics,* 2008**,** 420747.
HSIEH, P. H., KAN, C. C., WU, H. Y., YANG, H. C. & HSIEH, M. H. 2018. Early molecular events associated with nitrogen deficiency in rice seedling roots. *Sci Rep,* 8**,** 12207.
HU, J., ZHOU, J., PENG, X., XU, H., LIU, C., DU, B., YUAN, H., ZHU, L. & HE, G. 2011. The Bphi008a gene interacts with the ethylene pathway and transcriptionally regulates MAPK genes in the response of rice to brown planthopper feeding. *Plant Physiol,* 156**,** 856-72.
HU, J., ZHU, L., ZENG, D., GAO, Z., GUO, L., FANG, Y., ZHANG, G., DONG, G., YAN, M., LIU, J. & QIAN, Q. 2010. Identification and characterization of NARROW AND ROLLED LEAF 1, a novel gene regulating leaf morphology and plant architecture in rice. *Plant Mol Biol,* 73**,** 283-92.
HU, L., LI, N., XU, C., ZHONG, S., LIN, X., YANG, J., ZHOU, T., YULIANG, A., WU, Y., CHEN, Y. R., CAO, X., ZEMACH, A., RUSTGI, S., VON WETTSTEIN, D. & LIU, B. 2014. Mutation of a major CG methylase in rice causes genome-wide hypomethylation, dysregulated genome expression, and seedling lethality. *Proc Natl Acad Sci U S A,* 111**,** 10642-7.
HU, Y., QIN, F., HUANG, L., SUN, Q., LI, C., ZHAO, Y. & ZHOU, D. X. 2009. Rice histone deacetylase genes display specific expression patterns and developmental functions. *Biochem Biophys Res Commun,* 388**,** 266-71.
HUANG, L., WANG, Y., WANG, W., ZHAO, X., QIN, Q., SUN, F., HU, F., ZHAO, Y., LI, Z., FU, B. & LI, Z. 2018. Characterization of Transcription Factor Gene OsDRAP1 Conferring Drought Tolerance in Rice. *Front Plant Sci,* 9**,** 94.
HUANG, M. D., WEI, F. J., WU, C. C., HSING, Y. I. & HUANG, A. H. 2009a. Analyses of advanced rice anther transcriptomes reveal global tapetum secretory functions and potential proteins for lipid exine formation. *Plant Physiol,* 149**,** 694-707.
HUANG, X., QIAN, Q., LIU, Z., SUN, H., HE, S., LUO, D., XIA, G., CHU, C., LI, J. & FU, X. 2009b. Natural variation at the DEP1 locus enhances grain yield in rice. *Nat Genet,* 41**,** 494-7.
HUBER, S. C. & HUBER, J. L. 1996. Role and Regulation of Sucrose-Phosphate Synthase in Higher Plants. *Annu Rev Plant Physiol Plant Mol Biol,* 47**,** 431-444.
HULSEN, T., DE VLIEG, J. & ALKEMA, W. 2008. BioVenn - a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. *BMC Genomics,* 9**,** 488.
HUR, Y. J. & KIM, D. H. 2014. Overexpression of OsMAPK2 Enhances Low Phosphate Tolerance in Rice and Arabidopsis thaliana. *Am J Plant Sci,* 05**,** 452-462.

ILLUMINA-BASESPACE-LABS 2018. FASTQ Toolkit. *Available online at: https://www.illumina.com/products/by-type/informatics-products/basespace-sequence-hub/apps/fastq-toolkit.html*.

IQBAL, N., KHAN, N. A., FERRANTE, A., TRIVELLINI, A., FRANCINI, A. & KHAN, M. I. R. 2017. Ethylene Role in Plant Growth, Development and Senescence: Interaction with Other Phytohormones. *Front Plant Sci,* 8**,** 475.

IRRI 1988. *Rice Ratooning,* Los Banos, Laguna, Philippines, International Rice Research Institute.

IRSHAD, M., CANUT, H., BORDERIES, G., PONT-LEZICA, R. & JAMET, E. 2008. A new picture of cell wall protein dynamics in elongating cells of Arabidopsis thaliana: confirmed actors and newcomers. *BMC Plant Biol,* 8**,** 94.

ISHIWATA, A., OZAWA, M., NAGASAKI, H., KATO, M., NODA, Y., YAMAGUCHI, T., NOSAKA, M., SHIMIZU-SATO, S., NAGASAKI, A., MAEKAWA, M., HIRANO, H. Y. & SATO, Y. 2013. Two WUSCHEL-related homeobox genes, narrow leaf2 and narrow leaf3, control leaf width in rice. *Plant Cell Physiol,* 54**,** 779-92.

ITOH, H., TATSUMI, T., SAKAMOTO, T., OTOMO, K., TOYOMASU, T., KITANO, H., ASHIKARI, M., ICHIHARA, S. & MATSUOKA, M. 2004. A rice semi-dwarf gene, Tan-Ginbozu (D35), encodes the gibberellin biosynthesis enzyme, ent-kaurene oxidase. *Plant Mol Biol,* 54**,** 533-47.

ITOH, J., NONOMURA, K., IKEDA, K., YAMAKI, S., INUKAI, Y., YAMAGISHI, H., KITANO, H. & NAGATO, Y. 2005. Rice plant development: from zygote to spikelet. *Plant Cell Physiol,* 46**,** 23-47.

ITOH, J., SATO, Y. & NAGATO, Y. 2008. The SHOOT ORGANIZATION2 gene coordinates leaf domain development along the central-marginal axis in rice. *Plant Cell Physiol,* 49**,** 1226-36.

ITOH, J. I., KITANO, H., MATSUOKA, M. & NAGATO, Y. 2000. Shoot organization genes regulate shoot apical meristem organization and the pattern of leaf primordium initiation in rice. *Plant Cell,* 12**,** 2161-74.

IWAI, T., MIYASAKA, A., SEO, S. & OHASHI, Y. 2006. Contribution of ethylene biosynthesis for resistance to blast fungus infection in young rice plants. *Plant Physiol,* 142**,** 1202-15.

JAIN, P., SINGH, P. K., KAPOOR, R., KHANNA, A., SOLANKE, A. U., KRISHNAN, S. G., SINGH, A. K., SHARMA, V. & SHARMA, T. R. 2017. Understanding Host-Pathogen Interactions with Expression Profiling of NILs Carrying Rice-Blast Resistance Pi9 Gene. *Front Plant Sci,* 8**,** 93.

JAMES, D., TARAFDAR, A., BISWAS, K., SATHYAVATHI, T. C., PADARIA, J. C. & KUMAR, P. A. 2015. Development and characterization of a high temperature stress responsive subtractive cDNA library in Pearl Millet Pennisetum glaucum (L.) R.Br. *Indian J Exp Biol,* 53**,** 543-50.

JAMET, E., CANUT, H., BOUDART, G. & PONT-LEZICA, R. 2006. Cell wall proteins: a new insight through proteomics. *Trends Plant Sci,* 11**,** 33-39.

JAMSHEER, K. M., MANNULLY, C. T., GOPAN, N. & LAXMI, A. 2015. Comprehensive Evolutionary and Expression Analysis of FCS-Like Zinc finger Gene Family Yields Insights into Their Origin, Expansion and Divergence. *PLoS One,* 10**,** e0134328.

JEON, J. S., LEE, S., JUNG, K. H., JUN, S. H., JEONG, D. H., LEE, J., KIM, C., JANG, S., YANG, K., NAM, J., AN, K., HAN, M. J., SUNG, R. J., CHOI, H. S., YU, J. H., CHOI,

J. H., CHO, S. Y., CHA, S. S., KIM, S. I. & AN, G. 2000. T-DNA insertional mutagenesis for functional genomics in rice. *Plant J,* 22**,** 561-70.
JEONG, J. S., KIM, Y. S., REDILLAS, M. C., JANG, G., JUNG, H., BANG, S. W., CHOI, Y. D., HA, S. H., REUZEAU, C. & KIM, J. K. 2013. OsNAC5 overexpression enlarges root diameter in rice plants leading to enhanced drought tolerance and increased grain yield in the field. *Plant Biotechnol J,* 11**,** 101-14.
JIA, H., REN, H., GU, M., ZHAO, J., SUN, S., ZHANG, X., CHEN, J., WU, P. & XU, G. 2011. The phosphate transporter gene OsPht1;8 is involved in phosphate homeostasis in rice. *Plant Physiol,* 156**,** 1164-75.
JIA, S., XIONG, Y., XIAO, P., WANG, X. & YAO, J. 2019. OsNF-YC10, a seed preferentially expressed gene regulates grain width by afecting cell proliferation in rice. *Plant Sci,* 280**,** 219-227.
JIANG, J. F., XU, Y. Y. & CHONG, K. 2007. Overexpression of OsJAC1, a lectin gene, suppresses coleoptile and stem elongation in rice. *J Integr Plant Biol,* 49**,** 230-237.
JIANG, Y., BAO, L., JEONG, S. Y., KIM, S. K., XU, C., LI, X. & ZHANG, Q. 2012. XIAO is involved in the control of organ size by contributing to the regulation of signaling and homeostasis of brassinosteroids and cell cycling in rice. *Plant J,* 70**,** 398-408.
JIN, J., HUA, L., ZHU, Z., TAN, L., ZHAO, X., ZHANG, W., LIU, F., FU, Y., CAI, H., SUN, X., GU, P., XIE, D. & SUN, C. 2016. GAD1 Encodes a Secreted Peptide That Regulates Grain Number, Grain Length, and Awn Development in Rice Domestication. *Plant Cell,* 28**,** 2453-2463.
JOHNSON, P. R. & ECKER, J. R. 1998. The ethylene gas signal transduction pathway: a molecular perspective. *Annu Rev Genet,* 32**,** 227-54.
JONES, A.R., FORERO-VARGAS, M., WITHERS, S. P., SMITH, R. S., TRAAS, J., DEWITTE, W. & MURRAY, J. A. H. 2017. Cell-size dependent progression of the cell cycle creates homeostasis and flexibility of plant cell size. *Nat Commun,* 8**,** 15060.
JONES, D. B. 1993. Rice ratoon response to main crop harvest cutting height. *Agronomy,* 85**,** 1139-1142.
JU, C., YOON, G. M., SHEMANSKY, J. M., LIN, D. Y., YING, Z. I., CHANG, J., GARRETT, W. M., KESSENBROCK, M., GROTH, G., TUCKER, M. L., COOPER, B., KIEBER, J. J. & CHANG, C. 2012. CTR1 phosphorylates the central regulator EIN2 to control ethylene hormone signaling from the ER membrane to the nucleus in Arabidopsis. *Proc Natl Acad Sci U S A,* 109**,** 19486-91.
JULIUS, B. T., LEACH, K. A., TRAN, T. M., MERTZ, R. A. & BRAUN, D. M. 2017. Sugar Transporters in Plants: New Insights and Discoveries. *Plant Cell Physiol,* 58**,** 1442-1460.
KAILOU, L. 2012. Physiological traits, yields and nitrogen translocation of ratoon rice in response to different cultivations and planting periods. *Afr J Agr Res,* 7.
KAN, C. C., CHUNG, T. Y., WU, H. Y., JUO, Y. A. & HSIEH, M. H. 2017. Exogenous glutamate rapidly induces the expression of genes involved in metabolism and defense responses in rice roots. *BMC Genomics,* 18**,** 186.
KANEHISA, M., FURUMICHI, M., TANABE, M., SATO, Y. & MORISHIMA, K. 2017. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res,* 45**,** D353-D361.
KANEHISA, M. & GOTO, S. 2000. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res,* 28**,** 27-30.

KANEHISA, M., SATO, Y., FURUMICHI, M., MORISHIMA, K. & TANABE, M. 2019. New approach for understanding genome variations in KEGG. *Nucleic Acids Res,* 47**,** D590-D595.
KEBEISH, R., NIESSEN, M., THIRUVEEDHI, K., BARI, R., HIRSCH, H. J., ROSENKRANZ, R., STABLER, N., SCHONFELD, B., KREUZALER, F. & PETERHANSEL, C. 2007. Chloroplastic photorespiratory bypass increases photosynthesis and biomass production in Arabidopsis thaliana. *Nat Biotechnol,* 25**,** 593-9.
KELLERMEIER, F., ARMENGAUD, P., SEDITAS, T. J., DANKU, J., SALT, D. E. & AMTMANN, A. 2014. Analysis of the Root System Architecture of Arabidopsis Provides a Quantitative Readout of Crosstalk between Nutritional Signals. *Plant Cell,* 26**,** 1480-1496.
KENDE, H. 1993. Ethylene biosynthesis. *Annu Rev Plant Physiol Plant Mol Biol,* 44**,** 283-307.
KIEBER, J. J., ROTHENBERG, M., ROMAN, G., FELDMANN, K. A. & ECKER, J. R. 1993. CTR1, a negative regulator of the ethylene response pathway in Arabidopsis, encodes a member of the raf family of protein kinases. *Cell,* 72**,** 427-41.
KIM, C. Y., VO, K. T. X., NGUYEN, C. D., JEONG, D. H., LEE, S. K., KUMAR, M., KIM, S. R., PARK, S. H., KIM, J. K. & JEON, J. S. 2016. Functional analysis of a cold-responsive rice WRKY gene, OsWRKY71. *Plant Biotechnol Rep,* 10**,** 13-23.
KIM, S. L., CHOI, M., JUNG, K. H. & AN, G. 2013. Analysis of the early-flowering mechanisms and generation of T-DNA tagging lines in Kitaake, a model rice cultivar. *J Exp Bot,* 64**,** 4169-82.
KIM, S. L., LEE, S., KIM, H. J., NAM, H. G. & AN, G. 2007. OsMADS51 is a short-day flowering promoter that functions upstream of Ehd1, OsMADS14, and Hd3a. *Plant Physiol,* 145**,** 1484-94.
KIRIBUCHI, K., JIKUMARU, Y., KAKU, H., MINAMI, E., HASEGAWA, M., KODAMA, O., SETO, H., OKADA, K., NOJIRI, H. & YAMANE, H. 2005. Involvement of the basic helix-loop-helix transcription factor RERJ1 in wounding and drought stress responses in rice plants. *Biosci Biotechnol Biochem,* 69**,** 1042-4.
KIRIBUCHI, K., SUGIMORI, M., TAKEDA, M., OTANI, T., OKADA, K., ONODERA, H., UGAKI, M., TANAKA, Y., TOMIYAMA-AKIMOTO, C., YAMAGUCHI, T., MINAMI, E., SHIBUYA, N., OMORI, T., NISHIYAMA, M., NOJIRI, H. & YAMANE, H. 2004. RERJ1, a jasmonic acid-responsive gene from rice, encodes a basic helix-loop-helix protein. *Biochem Biophys Res Commun,* 325**,** 857-63.
KISHOR, P. B. K., SANGAM, S., AMRUTHA, R. N., LAXMI, P. S., NAIDU, K. R., RAO, K. R. S. S., RAO, S., REDDY, K. J., THERIAPPAN, P. & SREENIVASULU, N. 2005. Regulation of proline biosynthesis, degredation, uptake and transport in higher plants: Its implications in plant growth and abiotic stress tolerance. *Curr Sci,* 88**,** 424-438.
KOBAYASHI, K., YASUNO, N., SATO, Y., YODA, M., YAMAZAKI, R., KIMIZU, M., YOSHIDA, H., NAGAMURA, Y. & KYOZUKA, J. 2012. Inflorescence meristem identity in rice is specified by overlapping functions of three AP1/FUL-like MADS box genes and PAP2, a SEPALLATA MADS box gene. *Plant Cell,* 24**,** 1848-59.
KOHORN, B. D., KOBAYASHI, M., JOHANSEN, S., RIESE, J., HUANG, L. F., KOCH, K., FU, S., DOTSON, A. & BYERS, N. 2006. An Arabidopsis cell wall-associated kinase required for invertase activity and cell growth. *Plant J,* 46**,** 307-16.
KOMIYA, R., IKEGAMI, A., TAMAKI, S., YOKOI, S. & SHIMAMOTO, K. 2008. Hd3a and RFT1 are essential for flowering in rice. *Development,* 135**,** 767-74.

KOMIYA, R., YOKOI, S. & SHIMAMOTO, K. 2009. A gene network for long-day flowering activates RFT1 encoding a mobile flowering signal in rice. *Development,* 136**,** 3443-50.
KONG, W., ZHONG, H., GONG, Z., FANG, X., SUN, T., DENG, X. & LI, Y. 2019. Meta-Analysis of Salt Stress Transcriptome Responses in Different Rice Genotypes at the Seedling Stage. *Plants (Basel),* 8.
KOTAKE, T., AOHARA, T., HIRANO, K., SATO, A., KANEKO, Y., TSUMURAYA, Y., TAKATSUJI, H. & KAWASAKI, S. 2011. Rice Brittle culm 6 encodes a dominant-negative form of CesA protein that perturbs cellulose synthesis in secondary cell walls. *J Exp Bot,* 62**,** 2053-62.
KRISHNA, G. K., VISHWAKARMA, C., THOMAS, P., ARAVIND, J., KUSHWAHA, S. & CHINNUSAMY, V. 2017. Association between ABA-and drought-mediated regulation of root traits and identification of potential SNPs in genes for root development in rice. *Indian J Genet Pl Br,* 78.
KUREK, I., CHANG, T. K., BERTAIN, S. M., MADRIGAL, A., LIU, L., LASSNER, M. W. & ZHU, G. 2007. Enhanced Thermostability of Arabidopsis Rubisco activase improves photosynthesis and growth rates under moderate heat stress. *Plant Cell,* 19**,** 3230-41.
LA, H., LI, J., JI, Z., CHENG, Y., LI, X., JIANG, S., VENKATESH, P. N. & RAMACHANDRAN, S. 2006. Genome-wide analysis of cyclin family in rice (Oryza Sativa L.). *Mol Genet Genomics,* 275**,** 374-86.
LACOMBE, E., HAWKINS, S., VAN DOORSSELAERE, J., PIQUEMAL, J., GOFFNER, D., POEYDOMENGE, O., BOUDET, A. M. & GRIMA-PETTENATI, J. 1997. Cinnamoyl CoA reductase, the first committed enzyme of the lignin branch biosynthetic pathway: cloning, expression and phylogenetic relationships. *Plant J,* 11**,** 429-41.
LAW, J. A. & JACOBSEN, S. E. 2010. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat Rev Genet,* 11**,** 204-20.
LEE, J., DAS, A., YAMAGUCHI, M., HASHIMOTO, J., TSUTSUMI, N., UCHIMIYA, H. & UMEDA, M. 2003. Cell cycle function of a erice B2-type cyclin interacting with a B-type cyclin-dependent kinase. *Plant J,* 34**,** 417-425.
LEE, Y.-S. & AN, G. 2015. Regulation of flowering time in rice. *J Plant Biol,* 58**,** 353-360.
LEON, J., ROJO, E. & SANCHEZ-SERRANO, J. J. 2001. Wound signalling in plants. *J Exp Bot,* 52**,** 1-9.
LI, H. W., ZANG, B. S., DENG, X. W. & WANG, X. P. 2011. Overexpression of the trehalose-6-phosphate synthase gene OsTPS1 enhances abiotic stress tolerance in rice. *Planta,* 234**,** 1007-18.
LI, N., LIU, H., SUN, J., ZHENG, H., WANG, J., YANG, L., ZHAO, H. & ZOU, D. 2018. Transcriptome analysis of two contrasting rice cultivars during alkaline stress. *Sci Rep,* 8**,** 9586.
LI, R., XIONG, G., ZHANG, B. & ZHOU, Y. 2010a. Rice plants response to the disruption of OsCSLD4 gene. *Plant Signal Behav,* 5**,** 136-9.
LI, S., ZHAO, B., YUAN, D., DUAN, M., QIAN, Q., TANG, L., WANG, B., LIU, X., ZHANG, J., WANG, J., SUN, J., LIU, Z., FENG, Y. Q., YUAN, L. & LI, C. 2013. Rice zinc finger protein DST enhances grain production through controlling Gn1a/OsCKX2 expression. *Proc Natl Acad Sci U S A,* 110**,** 3167-72.
LI, W., WU, J., WENG, S., ZHANG, Y., ZHANG, D. & SHI, C. 2010b. Identification and characterization of dwarf 62, a loss-of-function mutation in DLT/OsGRAS-32 affecting gibberellin metabolism in rice. *Planta,* 232**,** 1383-96.

LI, X., YANG, Y., YAO, J., CHEN, G., LI, X., ZHANG, Q. & WU, C. 2009. FLEXIBLE CULM 1 encoding a cinnamyl-alcohol dehydrogenase controls culm mechanical strength in rice. *Plant Mol Biol,* 69**,** 685-97.

LI, Z., PINSON, S. R. M., STANSEL, J. W. & PARK, W. D. 1995. Identification of quantitative trait loci (QTLs) for heading date and plant height in cultivated rice (Oryza sativa L.). *Theor Appl Genet,* 91**,** 374-381.

LIAO, Y., ZOU, H. F., WEI, W., HAO, Y. J., TIAN, A. G., HUANG, J., LIU, Y. F., ZHANG, J. S. & CHEN, S. Y. 2008. Soybean GmbZIP44, GmbZIP62 and GmbZIP78 genes function as negative regulator of ABA signaling and confer salt and freezing tolerance in transgenic Arabidopsis. *Planta,* 228**,** 225-40.

LIU, B., CHEN, Z., SONG, X., LIU, C., CUI, X., ZHAO, X., FANG, J., XU, W., ZHANG, H., WANG, X., CHU, C., DENG, X., XUE, Y. & CAO, X. 2007a. Oryza sativa dicer-like4 reveals a key role for small interfering RNA silencing in plant development. *Plant Cell,* 19**,** 2705-18.

LIU, D., GONG, W., BAI, Y., LUO, J. & ZHU, Y. 2005. OsHT, a rice gene encoding for a plasma-membrane localized histidine transporter. *J Integr Plant Biol,* 47**,** 92-99.

LIU, L., MEI, Q., YU, Z., SUN, T., ZHANG, Z. & CHEN, M. 2013a. An integrative bioinformatics framework for genome-scale multiple level network reconstruction of rice. *J Integr Bioinform,* 10**,** 223.

LIU, Q., YAO, X., PI, L., WANG, H., CUI, X. & HUANG, H. 2009. The ARGONAUTE10 gene modulates shoot apical meristem maintenance and establishment of leaf polarity by repressing miR165/166 in Arabidopsis. *Plant J,* 58**,** 27-40.

LIU, T., LIU, H., ZHANG, H. & XING, Y. 2013b. Validation and characterization of Ghd7.1, a major quantitative trait locus with pleiotropic effects on spikelets per panicle, plant height, and heading date in rice (Oryza sativa L.). *J Integr Plant Biol,* 55**,** 917-27.

LIU, X., BAI, X., WANG, X. & CHU, C. 2007b. OsWRKY71, a rice transcription factor, is involved in rice defense response. *J Plant Physiol,* 164**,** 969-79.

LIU, X., LI, F., TANG, J., WANG, W., ZHANG, F., WANG, G., CHU, J., YAN, C., WANG, T., CHU, C., LI, C. 2012. Activation of the jasmonic acid pathway by depletion of the hydroperoxide lyase OsHPL3 reveals crosstalk between the HPL and AOS branches of the oxylipin pathway in rice. *PLoS One,* 7, e50089.

LIU, Z., ZHANG, S., SUN, N., LIU, H., ZHAO, Y., LIANG, Y., ZHANG, L. & HAN, Y. 2015. Functional diversity of jasmonates in rice. *Rice (N Y),* 8**,** 42.

LOPEZ-ARREDONDO, D. L., LEYVA-GONZALEZ, M. A., GONZALEZ-MORALES, S. I., LOPEZ-BUCIO, J. & HERRERA-ESTRELLA, L. 2014. Phosphate nutrition: improving low-phosphate tolerance in crops. *Annu Rev Plant Biol,* 65**,** 95-123.

LU, K., WU, B., WANG, J., ZHU, W., NIE, H., QIAN, J., HUANG, W. & FANG, Z. 2018. Blocking amino acid transporter OsAAP3 improves grain yield by promoting outgrowth buds and increasing tiller number in rice. *Plant Biotechnol J,* 16**,** 1710-1722.

LUAN, W., LIU, Y., ZHANG, F., SONG, Y., WANG, Z., PENG, Y. & SUN, Z. 2011. OsCD1 encodes a putative member of the cellulose synthase-like D sub-family and is essential for rice plant architecture and growth. *Plant Biotechnol J,* 9**,** 513-24.

LUO, J., LIU, H., ZHOU, T., GU, B., HUANG, X., SHANGGUAN, Y., ZHU, J., LI, Y., ZHAO, Y., WANG, Y., ZHAO, Q., WANG, A., WANG, Z., SANG, T., WANG, Z. & HAN, B. 2013. An-1 encodes a basic helix-loop-helix protein that regulates awn development, grain size, and grain number in rice. *Plant Cell,* 25**,** 3360-76.

LUO, J. S., HUANG, J., ZENG, D. L., PENG, J. S., ZHANG, G. B., MA, H. L., GUAN, Y., YI, H. Y., FU, Y. L., HAN, B., LIN, H. X., QIAN, Q. & GONG, J. M. 2018. A defensin-like protein drives cadmium efflux and allocation in rice. *Nat Commun,* 9**,** 645.

LV, S., WANG, Z., YANG, X., GUO, L., QIU, D. & ZENG, H. 2016. Transcriptional Profiling of Rice Treated with MoHrip1 Reveal the Function of Protein Elicitor in Enhancement of Disease Resistance and Plant Growth. *Front Plant Sci,* 7**,** 1818.

MA, H. 1994. GTP-binding proteins in plants: new members of an old family. *Plant Mol Biol,* 26**,** 1611-36.

MA, Q., DAI, X., XU, Y., GUO, J., LIU, Y., CHEN, N., XIAO, J., ZHANG, D., XU, Z., ZHANG, X. & CHONG, K. 2009. Enhanced tolerance to chilling stress in OsMYB3R-2 transgenic rice is mediated by alteration in cell cycle and ectopic expression of stress genes. *Plant Physiol,* 150**,** 244-56.

MA, X., FENG, F., WEI, H., MEI, H., XU, K., CHEN, S., LI, T., LIANG, X., LIU, H. & LUO, L. 2016. Genome-Wide Association Study for Plant Height and Grain Yield in Rice under Contrasting Moisture Regimes. *Front Plant Sci,* 7**,** 1801.

MALEKI, S. S., MOHAMMADI, K. & JI, K. S. 2016. Characterization of Cellulose Synthesis in Plant Cells. *Sci World J,* 2016**,** 8641373.

MAO, B., CHENG, Z., LEI, C., XU, F., GAO, S., REN, Y., WANG, J., ZHANG, X., WANG, J., WU, F., GUO, X., LIU, X., WU, C., WANG, H. & WAN, J. 2012. Wax crystal-sparse leaf2, a rice homologue of WAX2/GL1, is involved in synthesis of leaf cuticular wax. *Planta,* 235**,** 39-52.

MATHEW, I. E., DAS, S., MAHTO, A. & AGARWAL, P. 2016. Three Rice NAC Transcription Factors Heteromerize and Are Associated with Seed Size. *Front Plant Sci,* 7**,** 1638.

MATHIEU-RIVET, E., GEVAUDANT, F., CHENICLET, C., HERNOULD, M. & CHEVALIER, C. 2010. The Anaphase Promoting Complex activator CCS52A, a key factor for fruit growth and endoreduplication in Tomato. *Plant Signal Behav,* 5**,** 985-7.

MATILLA, A. J. 2000. Ethylene in seed formation and germination. *Seed Sci Res.* Cambridge Journals.

MAWLONG, I., ALI, K., SRINIVASAN, R., RAI, R. D. & TYAGI, A. 2015. Functional validation of a drought-responsive AP2/ERF family transcription factor-encoding gene from rice in Arabidopsis. *Mol Breeding,* 35.

MEGURO, A. & SATO, Y. 2014. Salicylic acid antagonizes abscisic acid inhibition of shoot growth and cell cycle progression in rice. *Sci Rep,* 4**,** 4555.

MERCHANT, N., LYONS, E., GOFF, S., VAUGHN, M., WARE, D., MICKLOS, D. & ANTIN, P. 2016. The iPlant Collaborative: Cyberinfrastructure for Enabling Data to Discovery for the Life Sciences. *PLoS Biol,* 14**,** e1002342.

MERCHANTE, C. & STEPANOVA, A. N. 2017. The triple response assay and its use to characterize ethylene mutants in Arabidopsis. *In:* BINDER, B. & G SCHALLER, E. (eds.) *Methods in Molecular Biology.* New York, NY: Humana Press.

MINIC, Z. & JOUANIN, L. 2006. Plant glycoside hydrolases involved in cell wall polysaccharide degradation. *Plant Physiol Biochem,* 44**,** 435-49.

MIURA, K., IKEDA, M., MATSUBARA, A., SONG, X. J., ITO, M., ASANO, K., MATSUOKA, M., KITANO, H. & ASHIKARI, M. 2010. OsSPL14 promotes panicle branching and higher grain productivity in rice. *Nat Genet,* 42**,** 545-9.

MIYAZAKI, J. H. & YANG, S. F. 1987. Metabolism of 5-methylthioribose to methionine. *Plant Physiol,* 84**,** 277-81.

MIZOI, J., SHINOZAKI, K. & YAMAGUCHI-SHINOZAKI, K. 2012. AP2/ERF family transcription factors in plant abiotic stress responses. *Biochim Biophys Acta,* 1819**,** 86-96.
MIZUKAMI, Y. & FISCHER, R. L. 2000. Plant organ size control: AINTEGUMENTA regulates growth and cell numbers during organogenesis. *Proc Natl Acad Sci U S A,* 97**,** 942-7.
MIZUTANI, M. 2012. Impacts of diversification of cytochrome P450 on plant metabolism. *Biol Pharm Bull,* 35**,** 824-32.
MOHANTA, T. K., BASHIR, T., HASHEM, A., ABD_ALLAH, E. F., KHAN, A. L. & AL-HARRASI, A. S. 2018. Molecular players of auxin transport systems: advances in genomic and molecular events. *J Plant Interact,* 13**,** 483-495.
MOHANTA, T. K. & MOHANTA, N. 2014. Understanding the role of Oryza sativa OsPILs (PIN like) genes in auxin signalling. *Peerj*.
MOHANTA, T. K., MOHANTA, N. & BAE, H. 2015. Identification and Expression Analysis of PIN-Like (PILS) Gene Family of Rice Treated with Auxin and Cytokinin. *Genes (Basel),* 6**,** 622-40.
MOHANTY, B., HERATH, V., WIJAYA, E., YEO, H. C., DE LOS REYES, B. G. & LEE, D. Y. 2012. Patterns of cis-element enrichment reveal potential regulatory modules involved in the transcriptional regulation of anoxia response of japonica rice. *Gene,* 511**,** 235-42.
MORITA, R., SUGINO, M., HATANAKA, T., MISOO, S. & FUKAYAMA, H. 2015. CO2-responsive CONSTANS, CONSTANS-like, and time of chlorophyll a/b binding protein Expression1 protein is a positive regulator of starch synthesis in vegetative organs of rice. *Plant Physiol,* 167**,** 1321-31.
MUTUKU, J. M., YOSHIDA, S., SHIMIZU, T., ICHIHASHI, Y., WAKATAKE, T., TAKAHASHI, A., SEO, M. & SHIRASU, K. 2015. The WRKY45-Dependent Signaling Pathway Is Required For Resistance against Striga hermonthica Parasitism. *Plant Physiol,* 168**,** 1152-63.
NAFATI, M., CHENICLET, C., HERNOULD, M., DO, P. T., FERNIE, A. R., CHEVALIER, C. & GEVAUDANT, F. 2011. The specific overexpression of a cyclin-dependent kinase inhibitor in tomato fruit mesocarp cells uncouples endoreduplication and cell growth. *Plant J,* 65**,** 543-56.
NAITHANI, S., PREECE, J., D'EUSTACHIO, P., GUPTA, P., AMARASINGHE, V., DHARMAWARDHANA, P. D., WU, G., FABREGAT, A., ELSER, J. L., WEISER, J., KEAYS, M., FUENTES, A. M., PETRYSZAK, R., STEIN, L. D., WARE, D. & JAISWAL, P. 2017. Plant Reactome: a resource for plant pathways and comparative analysis. *Nucleic Acids Res,* 45**,** D1029-D1039.
NAKAMURA, J., YUASA, T., HUONG, T. T., HARANO, K., TANAKA, S., IWATA, T., PHAN, T. & IWAYA, M. 2011. Rice homologs of inducer of CBF expression (OsICE) are involved in cold acclimation. *Plant Biotechnol,* 28**,** 303-309.
NAKANO, T., SUZUKI, K., FUJIMURA, T. & SHINSHI, H. 2006. Genome-wide analysis of the ERF gene family in Arabidopsis and rice. *Plant Physiol,* 140**,** 411-32.
NAKANO, Y., NISHIKUBO, N., SATO-IZAWA, K., MASE, K., KITANO, H., KAJITA, S., DEMURA, T. & KATAYAMA, Y. 2013. Transcription profiling identifies candidate genes for secondary cell wall formation and hydroxycinnamoyl-arabinoxylan biosynthesis in the rice internode. *Plant Biotechnol,* 30**,** 433-446.

NAKASHIMA, K., TAKASAKI, H., MIZOI, J., SHINOZAKI, K. & YAMAGUCHI-SHINOZAKI, K. 2012. NAC transcription factors in plant abiotic stress responses. *Biochim Biophys Acta,* 1819**,** 97-103.

NI, D. H., LI, J., DUAN, Y. B., YANG, Y. C., WEI, P. C., XU, R. F., LI, C. R., LIANG, D. D., LI, H., SONG, F. S., NI, J. L., LI, L. & YANG, J. B. 2014. Identification and utilization of cleistogamy gene cl7(t) in rice (Oryza sativa L.). *J Exp Bot,* 65**,** 2107-17.

OAD, F. C., CRUZ, P. S., MEMON, N., OAD, N. L. & ZIA-UL-HASSAN 2002. Rice Ratooning Management. *J Appl Sci,* 2**,** 29-35.

OHME-TAKAGI, M. & SHINSHI, H. 1995. Ethylene-inducible DNA binding proteins that interact with an ethylene responsive element. *Plant Cell,* 7**,** 173-182.

OHTANI, M., RAMACHANDRAN, V., TOKUMOTO, T., TAKEBAYASHI, A., IHARA, A., MATSUMOTO, T., HIROYAMA, R., NISHIKUBO, N. & DEMURA, T. 2017. Identification of novel factors that increase enzymatic saccharification efficiency in Arabidopsis wood cells. *Plant Biotechnol,* 34**,** 203-206.

OKITA, T. W., HWANG, Y. S., HNILO, J., KIM, W. T., ARYAN, A. P., LARSON, R. & KRISHNAN, H. B. 1989. Structure and expression of the rice glutelin multigene family. *J Biol Chem,* 264**,** 12573-81.

OKUSHIMA, Y., MITINA, I., QUACH, H. L. & THEOLOGIS, A. 2005. AUXIN RESPONSE FACTOR 2 (ARF2): a pleiotropic developmental regulator. *Plant J,* 43**,** 29-46.

OOKAWA, T., INOUE, K., MATSUOKA, M., EBITANI, T., TAKARADA, T., YAMAMOTO, T., UEDA, T., YOKOYAMA, T., SUGIYAMA, C., NAKABA, S., FUNADA, R., KATO, H., KANEKATSU, M., TOYOTA, K., MOTOBAYASHI, T., VAZIRZANJANI, M., TOJO, S. & HIRASAWA, T. 2014. Increased lodging resistance in long-culm, low-lignin gh2 rice for improved feed and bioenergy production. *Sci Rep,* 4**,** 6567.

ORTIZ-LOPEZ, A., CHANG, H. & BUSH, D. R. 2000. Amino acid transporters in plants. *Biochim Biophys Acta,* 1465**,** 275-80.

OUYANG, S., ZHU, W., HAMILTON, J., LIN, H., CAMPBELL, M., CHILDS, K., THIBAUD-NISSEN, F., MALEK, R. L., LEE, Y., ZHENG, L., ORVIS, J., HAAS, B., WORTMAN, J. & BUELL, C. R. 2007. The TIGR Rice Genome Annotation Resource: improvements and new features. *Nucleic Acids Res,* 35**,** D883-7.

PARK, H. L., KIM, T. L., BHOO, S. H., LEE, T. H., LEE, S. W. & CHO, M. H. 2018. Biochemical Characterization of the Rice Cinnamyl Alcohol Dehydrogenase Gene Family. *Molecules,* 23.

PARK, J. 2014. *Identification of rice genes involved in variety-specific tolerance to drought and phosphorus-deficiency.* Doctor of Sciences, ETH Zurich.

PARK, J. A., AHN, J. W., KIM, Y. K., KIM, S. J., KIM, J. K., KIM, W. T. & PAI, H. S. 2005. Retinoblastoma protein regulates cell proliferation, differentiation, and endoreduplication in plants. *Plant J,* 42**,** 153-63.

PAUL, P., AWASTHI, A., RAI, A. K., GUPTA, S. K., PRASAD, R., SHARMA, T. R. & DHALIWAL, H. S. 2012. Reduced tillering in Basmati rice T-DNA insertional mutant OsTEF1 associates with differential expression of stress related genes and transcription factors. *Funct Integr Genomics,* 12**,** 291-304.

PAUWELS, L. & GOOSSENS, A. 2011. The JAZ proteins: a crucial interface in the jasmonate signaling cascade. *Plant Cell,* 23**,** 3089-100.

PAWLAK, S. & DECKERT, J. 2007. Histone modifications under enviornmental stress. *Biol Lett***,** 65-75.

PELEG, Z. & BLUMWALD, E. 2011. Hormone balance and abiotic stress tolerance in crop plants. *Curr Opin Plant Biol,* 14**,** 290-5.
PERILLI, S., DI MAMBRO, R. & SABATINI, S. 2012. Growth and development of the root apical meristem. *Curr Opin Plant Biol,* 15**,** 17-23.
PERSSON, S., WEI, H., MILNE, J., PAGE, G. P. & SOMERVILLE, C. R. 2005. Identification of genes required for cellulose synthesis by regression analysis of public microarray data sets. *Proc Natl Acad Sci U S A,* 102**,** 8633-8.
PETTKO-SZANDTNER, A., CSERHATI, M., BARROCO, R. M., HARIHARAN, S., DUDITS, D. & BEEMSTER, G. T. 2015. Core cell cycle regulatory genes in rice and their expression profiles across the growth zone of the leaf. *J Plant Res,* 128**,** 953-74.
PINYOPICH, A., DITTA, G. S., SAVIDGE, B., LILJEGREN, S. J., BAUMANN, E., WISMAN, E. & YANOFSKY, M. F. 2003. Assessing the redundancy of MADS-box genes during carpel and ovule development. *Nature,* 424**,** 85-8.
PIREYRE, M. & BUROW, M. 2015. Regulation of MYB and bHLH transcription factors: a glance at the protein level. *Mol Plant,* 8**,** 378-88.
PONNU, J., WAHL, V. & SCHMID, M. 2011. Trehalose-6-phosphate: connecting plant metabolism and development. *Front Plant Sci,* 2**,** 70.
POONYARIT, M., MACKILL, D. J. & VERGARA, B. S. 1988. Genetics of photoperiod sensitivity and critical day length in rice. *Crop Sci,* 29**,** 647-652.
QI, J., QIAN, Q., BU, Q., LI, S., CHEN, Q., SUN, J., LIANG, W., ZHOU, Y., CHU, C., LI, X., REN, F., PALME, K., ZHAO, B., CHEN, J., CHEN, M. & LI, C. 2008. Mutation of the rice Narrow leaf1 gene, which encodes a novel protein, affects vein patterning and polar auxin transport. *Plant Physiol,* 147**,** 1947-59.
QI, R. & JOHN, P. C. 2007. Expression of genomic AtCYCD2;1 in Arabidopsis induces cell division at smaller cell sizes: implications for the control of plant growth. *Plant Physiol,* 144**,** 1587-97.
QIAO, H., SHEN, Z., HUANG, S. S., SCHMITZ, R. J., URICH, M. A., BRIGGS, S. P. & ECKER, J. R. 2012. Processing and subcellular trafficking of ER-tethered EIN2 control response to ethylene gas. *Science,* 338**,** 390-3.
QIN, R., ZENG, D., YANG, C., AKHTER, D., ALAMIN, M., JIN, X. & SHI, C. 2018. LTBSG1, a New Allele of BRD2, Regulates Panicle and Grain Development in Rice by Brassinosteroid Biosynthetic Pathway. *Genes (Basel),* 9.
QU, X., YAN, M., ZOU, J., JIANG, M., YANG, K. & LE, J. 2018. A2-type cyclin is required for the asymmetric entry division in rice stomatal development. *J Exp Bot,* 69**,** 3587-3599.
RAINERI, J., WANG, S., PELEG, Z., BLUMWALD, E. & CHAN, R. L. 2015. The rice transcription factor OsWRKY47 is a positive regulator of the response to water deficit stress. *Plant Mol Biol,* 88**,** 401-13.
RAO, K. P., RICHA, T., KUMAR, K., RAGHURAM, B. & SINHA, A. K. 2010. In silico analysis reveals 75 members of mitogen-activated protein kinase kinase kinase gene family in rice. *DNA Res,* 17**,** 139-53.
RAVANEL, S., GAKIERE, B., JOB, D. & DOUCE, R. 1998. The specific features of methionine biosynthesis and metabolism in plants. *Proc Natl Acad Sci U S A,* 95**,** 7805-12.
REDILLAS, M. C., JEONG, J. S., KIM, Y. S., JUNG, H., BANG, S. W., CHOI, Y. D., HA, S. H., REUZEAU, C. & KIM, J. K. 2012. The overexpression of OsNAC9 alters the root

architecture of rice plants enhancing drought resistance and grain yield under field conditions. *Plant Biotechnol J,* 10**,** 792-805.
REGUERA, M., PELEG, Z., ABDEL-TAWAB, Y. M., TUMIMBANG, E. B., DELATORRE, C. A. & BLUMWALD, E. 2013. Stress-Induced Cytokinin Synthesis Increases Drought Tolerance through the Coordinated Regulation of Carbon and Nitrogen Assimilation in Rice. *Plant Physiol,* 163**,** 1609-1622.
REINHARDT, D. & KUHLEMEIER, C. 2002. Plant architecture. *EMBO Rep,* 3**,** 846-51.
RICHARD-MOLARD, C., KRAPP, A., BRUN, F., NEY, B., DANIEL-VEDELE, F. & CHAILLOU, S. 2008. Plant response to nitrate starvation is determined by N storage capacity matched by nitrate uptake capacity in two Arabidopsis genotypes. *J Exp Bot,* 59**,** 779-91.
RICHARDS, D. E., KING, K. E., AIT-ALI, T. & HARBERD, N. P. 2001. How gibberellin regulates plant growth and development: A molecular genetic analysis of gibberellin signaling. *Annu Rev Plant Physiol Plant Mol Biol,* 52**,** 67-88.
RIECHMANN, J. L. & MEYEROWITZ, E. M. 1998. The AP2/EREBP family of plant transcription factors. *Biol Chem,* 379**,** 633-46.
ROBINSON, M. D., MCCARTHY, D. J. & SMYTH, G. K. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics,* 26**,** 139-40.
RODRIGUEZ, F. I., ESCH, J. J., HALL, A. E., BINDER, B. M., SCHALLER, G. E. & BLEECKER, A. B. 1999. A copper cofactor for the ethylene receptor ETR1 from Arabidopsis. *Science,* 283**,** 996-8.
ROMAN, G. & ECKER, J. R. 1995. Genetic analysis of a seedling stress response to ethylene in Arabidopsis. *Philos Trans R Soc Lond B Biol Sci,* 350**,** 75-81.
SABLOWSKI, R. & CARNIER DORNELAS, M. 2014. Interplay between cell growth and cell cycle in plants. *J Exp Bot,* 65**,** 2703-14.
SAGEHASHI, Y., TAKAKU, H. & YATOU, O. 2017. Partial peptides from rice defensin OsAFP1 exhibited antifungal activity against the rice blast pathogen Pyricularia oryzae. *J Pestic Sci,* 42**,** 172-175.
SAMUEL, D., LIU, Y. J., CHENG, C. S. & LYU, P. C. 2002. Solution structure of plant nonspecific lipid transfer protein-2 from rice (Oryza sativa). *J Biol Chem,* 277**,** 35267-73.
SANTNER, A., CALDERON-VILLALOBOS, L. I. & ESTELLE, M. 2009. Plant hormones are versatile chemical regulators of plant growth. *Nat Chem Biol,* 5**,** 301-7.
SASAKI, A., ASHIKARI, M., UEGUCHI-TANAKA, M., ITOH, H., NISHIMURA, A., SWAPAN, D., ISHIYAMA, K., SAITO, T., KOBAYASHI, M., KHUSH, G. S., KITANO, H. & MATSUOKA, M. 2002. Green revolution: a mutant gibberellin-synthesis gene in rice. *Nature,* 416**,** 701-2.
SASAKI, A., ITOH, H., GOMI, K., UEGUCHI-TANAKA, M., ISHIYAMA, K., KOBAYASHI, M., JEONG, D. H., AN, G., KITANO, H., ASHIKARI, M. & MATSUOKA, M. 2003. Accumulation of phosphorylated repressor for gibberellin signaling in an F-box mutant. *Science,* 299**,** 1896-8.
SATO, T. & THEOLOGIES, A. 1989. Cloning the mRNA encoding 1-aminocyclopropane-1-carboxylate synthase, the key enzyme for ethylene biosynthesisin plants. *Proc Natl Acad Sci U S A,* 86**,** 6621-6625.
SATOH, H., SHIBAHARA, K., TOKUNAGA, T., NISHI, A., TASAKI, M., HWANG, S. K., OKITA, T. W., KANEKO, N., FUJITA, N., YOSHIDA, M., HOSAKA, Y., SATO, A.,

UTSUMI, Y., OHDAN, T. & NAKAMURA, Y. 2008. Mutation of the plastidial alpha-glucan phosphorylase gene in rice affects the synthesis and structure of starch in the endosperm. *Plant Cell,* 20**,** 1833-49.
SAZUKA, T., AICHI, I., KAWAI, T., MATSUO, N., KITANO, H. & MATSUOKA, M. 2005. The rice mutant dwarf bamboo shoot 1: a leaky mutant of the NACK-type kinesin-like gene can initiate organ primordia but not organ development. *Plant Cell Physiol,* 46**,** 1934-43.
SCHALLER, G. E. & BLEECKER, A. B. 1995. Ethylene-binding sites generated in yeast expressing the Arabidopsis ETR1 gene. *Science,* 270**,** 1809-11.
SCHALLER, G. E., LADD, A. N., LANAHAN, M. B., SPANBAUER, J. M. & BLEECKER, A. B. 1995. The ethylene response mediatior ETR1 from Arabidopsis forms a disulfide-linked dimer. *J Biol Chem,* 270**,** 12526-12530.
SCHINDELMAN, G., MORIKAMI, A., JUNG, L., BASKIN, T. I., CARPITA, N. C., DERBYSHIRE, P., MCCANN, M. C. & BENFEY, P. N. 2001. COBRA encodes a putative GPI-anchored protein, which is polarly localized and necessary for oriented cell expansion in Arabidopsis. *Genes Dev,* 15**,** 1115-1127.
SEIFI, H. S., VAN BOCKHAVEN, J., ANGENON, G. & HOFTE, M. 2013. Glutamate metabolism in plant disease and defense: friend or foe? *Mol Plant Microbe Interact,* 26**,** 475-85.
SEO, J. S., JOO, J., KIM, M. J., KIM, Y. K., NAHM, B. H., SONG, S. I., CHEONG, J. J., LEE, J. S., KIM, J. K. & CHOI, Y. D. 2011. OsbHLH148, a basic helix-loop-helix protein, interacts with OsJAZ proteins in a jasmonate signaling pathway leading to drought tolerance in rice. *Plant J,* 65**,** 907-21.
SEO, H. S., SONG, J. T., CHEONG, J. J., LEE, Y. H., LEE, Y. W., HWANG, I., LEE, J. S., CHOI, Y. D. 2001. Jasmonic acid carboxyl methyltransferase: A key enzyme for jasmonate-regulated plant responses. *Proc Natl Acad Sci U S A,* 98, 4788-4793.
SEYEDNASROLLAH, F., LAIHO, A. & ELO, L. L. 2015. Comparison of software packages for detecting differential expression in RNA-seq studies. *Brief Bioinform,* 16**,** 59-70.
SHANKAR, A., SINGH, A., KANWAR, P., SRIVASTAVA, A. K., PANDEY, A., SUPRASANNA, P., KAPOOR, S. & PANDEY, G. K. 2013. Gene expression analysis of rice seedling under potassium deprivation reveals major changes in metabolism and signaling components. *PLoS One,* 8**,** e70321.
SHARMA, R., MOHAN SINGH, R. K., MALIK, G., DEVESHWAR, P., TYAGI, A. K., KAPOOR, S. & KAPOOR, M. 2009. Rice cytosine DNA methyltransferases - gene expression profiling during reproductive development and abiotic stress. *FEBS J,* 276**,** 6301-11.
SHARMA, R., TAN, F., JUNG, K.-H., SHARMA, M. K., PENG, Z. & RONALD, P. C. 2011. Transcriptional dynamics during cell wall removal and regeneration reveals key genes involved in cell wall development in rice. *Plant Mol Biol,* 77**,** 391-406.
SHARONI, A. M., NURUZZAMAN, M., SATOH, K., MOUMENI, A., ATTIA, K., VENUPRASAD, R., SERRAJ, R., KUMAR, A., LEUNG, H., ISLAM, A. K. & KIKUCHI, S. 2012. Comparative transcriptome analysis of AP2/EREBP gene family under normal and hormone treatments, and under two drought stresses in NILs setup by Aday Selection and IR64. *Mol Genet Genomics,* 287**,** 1-19.
SHARONI, A. M., NURUZZAMAN, M., SATOH, K., SHIMIZU, T., KONDOH, H., SASAYA, T., CHOI, I. R., OMURA, T. & KIKUCHI, S. 2011. Gene structures, classification and

expression models of the AP2/EREBP transcription factor family in rice. *Plant Cell Physiol,* 52**,** 344-60.
SHI, J., DONG, A. & SHEN, W. H. 2014. Epigenetic regulation of rice flowering and reproduction. *Front Plant Sci,* 5**,** 803.
SHI, J., TAN, H., YU, X. H., LIU, Y., LIANG, W., RANATHUNGE, K., FRANKE, R. B., SCHREIBER, L., WANG, Y., KAI, G., SHANKLIN, J., MA, H. & ZHANG, D. 2011. Defective pollen wall is required for anther and microspore development in rice and encodes a fatty acyl carrier protein reductase. *Plant Cell,* 23**,** 2225-46.
SHI, L., ZHANG, X. B., SHI, Y. F., XU, X., HE, Y., SHAO, G., HUANG, Q. N. & WU, J. L. 2019. OsCDC48/48E complex is required for plant survival in rice (Oryza sativa L.). *Plant Mol Biol,* 100**,** 163-179.
SHULAEV, V., SILVERMAN, P. & RASKIN, I. 1997. Airborne signalling by methyl salicylate in plant pathogen resistance *Nature,* 385**,** 718-721.
SILVEIRA, R. D., ABREU, F. R., MAMIDI, S., MCCLEAN, P. E., VIANELLO, R. P., LANNA, A. C., CARNEIRO, N. P. & BRONDANI, C. 2015. Expression of drought tolerance genes in tropical upland rice cultivars (Oryza sativa). *Genet Mol Res,* 14**,** 8181-200.
SINGH, P. & SINHA, A. K. 2016. A Positive Feedback Loop Governed by SUB1A1 Interaction with MITOGEN-ACTIVATED PROTEIN KINASE3 Imparts Submergence Tolerance in Rice. *Plant Cell,* 28**,** 1127-43.
SINHA, S., KOTASTHANE, A. S. & VERULKAR, S. B. 2017. In silico characterization of QTL regions associated with drought tolerance traits in rice (Oryza sativa L.). *Indian J Genet Pl Br,* 77.
SMITA, S., KATIYAR, A., CHINNUSAMY, V., PANDEY, D. M. & BANSAL, K. C. 2015. Transcriptional Regulatory Network Analysis of MYB Transcription Factor Family Genes in Rice. *Front Plant Sci,* 6**,** 1157.
SOMSSICH, M., JE, B. I., SIMON, R. & JACKSON, D. 2016. CLAVATA-WUSCHEL signaling in the shoot meristem. *Development,* 143**,** 3238-48.
SONG, X., WANG, D., MA, L., CHEN, Z., LI, P., CUI, X., LIU, C., CAO, S., CHU, C., TAO, Y. & CAO, X. 2012. Rice RNA-dependent RNA polymerase 6 acts in small RNA biogenesis and spikelet development. *Plant J,* 71**,** 378-89.
SOUER, E., VAN HOUWLIGEN, A., KLOOS, D., MOL, J. & KOES, R. 1996. The no apical meristem gene of petunia is required for pattern formation in embryos and flowers and is expressed at meristem and primordia boundaries. *Cell,* 85**,** 159-170.
SPANU, P., REINHARDT, D. & BOLLER, T. 1991. Analysis and cloning of the ethylene-forming enzyme from tomato by functional expression of its mRNA in Xenopus laevis oocytes. *EMBO J,* 10**,** 2007-13.
SRIVASTAVA, S., VISHWAKARMA, R. K., ARAFAT, Y. A., GUPTA, S. K. & KHAN, B. M. 2015. Abiotic stress induces change in Cinnamoyl CoA Reductase (CCR) protein abundance and lignin deposition in developing seedlings of Leucaena leucocephala. *Physiol Mol Biol Plants,* 21**,** 197-205.
STURM, A. & TANG, G. 1999. The sucrose-cleaving enzyme of plants are crucial for development, growth and carbon partitioning. *Trends Plant Sci,* 4**,** 401-407.
SUGIMOTO-SHIRASU, K. & ROBERTS, K. 2003. "Big it up": endoreduplication and cell-size control in plants. *Curr Opin Plant Biol,* 6**,** 544-553.

SUI, P., JIN, J., YE, S., MU, C., GAO, J., FENG, H., SHEN, W. H., YU, Y. & DONG, A. 2012. H3K36 methylation is critical for brassinosteroid-regulated plant growth and development in rice. *Plant J,* 70**,** 340-7.

SUI, P., SHI, J., GAO, X., SHEN, W. H. & DONG, A. 2013. H3K36 methylation is involved in promoting rice flowering. *Mol Plant,* 6**,** 975-7.

SUN, C., CHEN, D., FANG, J., WANG, P., DENG, X. & CHU, C. 2014a. Understanding the genetic and epigenetic architecture in complex network of rice flowering pathways. *Protein Cell,* 5**,** 889-98.

SUN, C., FANG, J., ZHAO, T., XU, B., ZHANG, F., LIU, L., TANG, J., ZHANG, G., DENG, X., CHEN, F., QIAN, Q., CAO, X. & CHU, C. 2012. The histone methyltransferase SDG724 mediates H3K36me2/3 deposition at MADS50 and RFT1 and promotes flowering in rice. *Plant Cell,* 24**,** 3235-47.

SUN, L., LI, X., FU, Y., ZHU, Z., TAN, L., LIU, F., SUN, X., SUN, X. & SUN, C. 2013. GS6, a member of the GRAS gene family, negatively regulates grain size in rice. *J Integr Plant Biol,* 55**,** 938-49.

SUN, X., LING, S., LU, Z., OUYANG, Y., LIU, S. & YAO, J. 2014b. OsNF-YB1, a rice endosperm-specific gene, is essential for cell proliferation in endosperm development. *Gene,* 551**,** 214-221.

SUZUKI, K., SUZUKI, N., OHME-TAKAGI, M. & SHINSHI, H. 1998. Immediate early induction of mRNAs for ethylene-responsive transcription factors in tobacco leaf strips after cutting. *Plant J,* 15**,** 657-665.

TAKAHASHI, H., SAIKA, H., MATSUMURA, H., NAGAMURA, Y., TSUTSUMI, N., NISHIZAWA, N. K. & NAKAZONO, M. 2011. Cell division and cell elongation in the coleoptile of rice alcohol dehydrogenase 1-deficient mutant are reduced under complete submergence. *Ann Bot,* 108**,** 253-61.

TAMAKI, S., MATSUO, S., WONG, H. L., YOKOI, S. & SHIMAMOTO, K. 2007. Hd3a protein is a mobile flowering signal in rice. *Science,* 316**,** 1033-6.

TAMAKI, S., TSUJI, H., MATSUMOTO, A., FUJITA, A., SHIMATANI, Z., TERADA, R., SAKAMOTO, T., KURATA, T. & SHIMAMOTO, K. 2015. FT-like proteins induce transposon silencing in the shoot apex during floral induction in rice. *Proc Natl Acad Sci U S A,* 112**,** E901-10.

TANABE, S., ASHIKARI, M., FUJIOKA, S., TAKATSUTO, S., YOSHIDA, S., YANO, M., YOSHIMURA, A., KITANO, H., MATSUOKA, M., FUJISAWA, Y., KATO, H. & IWASAKI, Y. 2005. A novel cytochrome P450 is implicated in brassinosteroid biosynthesis via the characterization of a rice dwarf mutant, dwarf11, with reduced seed length. *Plant Cell,* 17**,** 776-90.

TANAKA, W., TORIBA, T., OHMORI, Y., YOSHIDA, A., KAWAI, A., MAYAMA-TSUCHIDA, T., ICHIKAWA, H., MITSUDA, N., OHME-TAKAGI, M. & HIRANO, H. Y. 2012. The YABBY gene TONGARI-BOUSHI1 is involved in lateral organ development and maintenance of meristem organization in the rice spikelet. *Plant Cell,* 24**,** 80-95.

TANTONG, S., PRINGSULAKA, O., WEERAWANICH, K., MEEPRASERT, A., RUNGROTMONGKOL, T., SARNTHIMA, R., ROYTRAKUL, S. & SIRIKANTARAMAS, S. 2016. Two novel antimicrobial defensins from rice identified by gene coexpression network analyses. *Peptides,* 84**,** 7-16.

TAOKA, K., OHKI, I., TSUJI, H., FURUITA, K., HAYASHI, K., YANASE, T., YAMAGUCHI, M., NAKASHIMA, C., PURWESTRI, Y. A., TAMAKI, S., OGAKI, Y., SHIMADA, C., NAKAGAWA, A., KOJIMA, C. & SHIMAMOTO, K. 2017. 14-3-3 proteins act as intracellular receptors for rice Hd3a florigen. *Nature,* 476**,** 332-335.

TEALE, W. D., PAPONOV, I. A. & PALME, K. 2006. Auxin in action: signalling, transport and the control of plant growth and development. *Nat Rev Mol Cell Biol,* 7**,** 847-59.

THARA, V. K., TANG, X., GU, Y. Q., MARTIN, G. B. & ZHOU, J. M. 1999. Pseudomonas syringae pv tomato induces the expression of tomato EREBP-like genes pti4 and pti5 independent of ethylene, salicylate and jasmonate. *Plant J,* 20**,** 475-83.

THINES, B., KATSIR, L., MELOTTO, M., NIU, Y., MANDAOKAR, A., LIU, G., NOMURA, K., HE, S. Y., HOWE, G. A. & BROWSE, J. 2007. JAZ repressor proteins are targets of the SCF(COI1) complex during jasmonate signalling. *Nature,* 448**,** 661-5.

THOMAS, B. R., ROMERO, G. O., NEVINS, D. J. & RODRIGUEZ, R. L. 2000. New perspectives on the endo-beta-glucanases of glycosyl hydrolase Family 17. *Int J Biol Macromol,* 27**,** 139-44.

TIAN, C., WAN, P., SUN, S., LI, J. & CHEN, M. 2004. Genome-wide analysis of the GRAS gene family in rice and Arabidopsis. *Plant Mol Biol,* 54**,** 519-32.

TIAN, X., LI, X., ZHOU, W., REN, Y., WANG, Z., LIU, Z., TANG, J., TONG, H., FANG, J. & BU, Q. 2017. Transcription Factor OsWRKY53 Positively Regulates Brassinosteroid Signaling and Plant Architecture. *Plant Physiol,* 175**,** 1337-1349.

TONG, H., JIN, Y., LIU, W., LI, F., FANG, J., YIN, Y., QIAN, Q., ZHU, L. & CHU, C. 2009. DWARF AND LOW-TILLERING, a new member of the GRAS family, plays positive roles in brassinosteroid signaling in rice. *Plant J,* 58**,** 803-816.

TORIBA, T. & HIRANO, H.-Y. 2014. TheDROOPING LEAFandOsETTIN2genes promote awn development in rice. *Plant J,* 77**,** 616-626.

TORIBA, T., SUZAKI, T., YAMAGUCHI, T., OHMORI, Y., TSUKAYA, H. & HIRANO, H. Y. 2010. Distinct regulation of adaxial-abaxial polarity in anther patterning in rice. *Plant Cell,* 22**,** 1452-62.

TSUJI, H., NAKAMURA, H., TAOKA, K. & SHIMAMOTO, K. 2013. Functional diversification of FD transcription factors in rice, components of florigen activation complexes. *Plant Cell Physiol,* 54**,** 385-97.

TSUJI, H., TAOKA, K. & SHIMAMOTO, K. 2011. Regulation of flowering in rice: two florigen genes, a complex gene network, and natural variation. *Curr Opin Plant Biol,* 14**,** 45-52.

TUTEJA, N., TRAN, N. Q., DANG, H. Q. & TUTEJA, R. 2011. Plant MCM proteins: role in DNA replication and beyond. *Plant Mol Biol,* 77**,** 537-45.

UJI, Y., TANIGUCHI, S., TAMAOKI, D., SHISHIDO, H., AKIMITSU, K. & GOMI, K. 2016. Overexpression of OsMYC2 results in the up-regulation of early JA-responsive genes and bacterial blight resistance in rice. *Plant Cell Physiol,* 57**,** 1814-27.

UMEDA, M., IWAMOTO, N., UMEDA-HARA, C., YAMAGUCHI, M., HASHIMOTO, J. & UCHIMIYA, H. 1999. Molecular characterization of mitotic cyclins in rice plants. *Mol Gen Genet,* 262**,** 230-8.

VAN ECK, L., DAVIDSON, R. M., WU, S., ZHAO, B. Y., BOTHA, A. M., LEACH, J. E. & LAPITAN, N. L. 2014. The transcriptional network of WRKY53 in cereals links oxidative responses to biotic and abiotic stress inputs. *Funct Integr Genomics,* 14**,** 351-62.

VAN LOON, L. C., GERAATS, B. P. & LINTHORST, H. J. 2006. Ethylene as a modulator of disease resistance in plants. *Trends Plant Sci,* 11**,** 184-91.
VARET, H., BRILLET-GUEGUEN, L., COPPEE, J. Y. & DILLIES, M. A. 2016. SARTools: A DESeq2- and EdgeR-Based R Pipeline for Comprehensive Differential Analysis of RNA-Seq Data. *PLoS One,* 11**,** e0157022.
VISHAL, B., KRISHNAMURTHY, P., RAMAMOORTHY, R. & KUMAR, P. P. 2019. OsTPS8 controls yield-related traits and confers salt stress tolerance in rice by enhancing suberin deposition. *New Phytol,* 221**,** 1369-1386.
WANG, C., TARIQ, R., JI, Z., WEI, Z., ZHENG, K., MISHRA, R. & ZHAO, K. 2019. Transcriptome analysis of a rice cultivar reveals the differentially expressed genes in response to wild and mutant strains of Xanthomonas oryzae pv. oryzae. *Sci Rep,* 9**,** 3757.
WANG, D., PAN, Y., ZHAO, X., ZHU, L., FU, B. & LI, Z. 2011. Genome-wide temporal-spatial gene expression profiling of drought responsiveness in rice. *BMC Genomics,* 12**,** 149.
WANG, G., KONG, H., SUN, Y., ZHANG, X., ZHANG, W., ALTMAN, N., DEPAMPHILIS, C. W. & MA, H. 2004a. Genome-wide analysis of the cyclin family in Arabidopsis and comparative phylogenetic analysis of plant cyclin-like proteins. *Plant Physiol,* 135**,** 1084-99.
WANG, H., HAO, J., CHEN, X., HAO, Z., WANG, X., LOU, Y., PENG, Y. & GUO, Z. 2007. Overexpression of rice WRKY89 enhances ultraviolet B tolerance and disease resistance in rice plants. *Plant Mol Biol,* 65**,** 799-815.
WANG, H., JONES, B., LI, Z., FRASSE, P., DELALANDE, C., REGAD, F., CHAABOUNI, S., LATCHE, A., PECH, J. C. & BOUZAYEN, M. 2005. The tomato Aux/IAA transcription factor IAA9 is involved in fruit development and leaf morphogenesis. *Plant Cell,* 17**,** 2676-92.
WANG, H., ZHOU, Y., GILMER, S., WHITWILL, S. & FOWKE, L. C. 2000. Expression of the plant cyclin-dependent kinase inhibitor ICK1 affects cell division, plant growth and morphology. *Plant J,* 24**,** 613-23.
WANG, K. L. C., LI, H. & ECKER, J. R. 2002. Ethylene Biosynthesis and Signaling Networks. *Plant Cell,* 14**,** S131-S151.
WANG, W., VINOCUR, B., SHOSEYOV, O. & ALTMAN, A. 2004b. Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. *Trends Plant Sci,* 9**,** 244-52.
WANG, Y., GUO, H., LI, H., ZHANG, H. & MIAO, X. 2012. Identification of transcription factors potential related to brown planthopper resistance in rice via microarray expression profiling. *BMC Genomics,* 13**,** 687.
WEI, J., VAN LOON, J. J., GOLS, R., MENZEL, T. R., LI, N., KANG, L. & DICKE, M. 2014. Reciprocal crosstalk between jasmonate and salicylate defence-signalling pathways modulates plant volatile emission and herbivore host-selection behaviour. *J Exp Bot,* 65**,** 3289-98.
WEI, J. Y., LI, A. M., LI, Y., WANG, J., LIU, X. B., LIU, L. S. & XU, Z. F. 2006. Cloning and characterization of an RNase-related protein gene preferentially expressed in rice stems. *Biosci Biotechnol Biochem,* 70**,** 1041-5.
WEI, X., XU, J., GUO, H., JIANG, L., CHEN, S., YU, C., ZHOU, Z., HU, P., ZHAI, H. & WAN, J. 2010. DTH8 suppresses flowering in rice, influencing plant height and yield potential simultaneously. *Plant Physiol,* 153**,** 1747-58.

WEI, Z., HU, W., LIN, Q., CHENG, X., TONG, M., ZHU, L., CHEN, R. & HE, G. 2009. Understanding rice plant resistance to the Brown Planthopper (Nilaparvata lugens): a proteomic approach. *Proteomics,* 9**,** 2798-808.

WEN, X., ZHANG, C., JI, Y., ZHAO, Q., HE, W., AN, F., JIANG, L. & GUO, H. 2012. Activation of ethylene signaling is mediated by nuclear translocation of the cleaved EIN2 carboxyl terminus. *Cell Res,* 22**,** 1613-6.

WERNER, T., MOTYKA, V., STRNAD, M. & SCHMULLING, T. 2001. Regulation of plant growth by cytokinin. *Proc Natl Acad Sci U S A,* 98**,** 10487-92.

WIN, K. T., ZHANG, C. & LEE, S. 2018. Genome-wide identification and description of MLO family genes in pumpkin (Cucurbita maxima Duch.). *Hortic Environ Biote,* 59**,** 397-410.

WU, L., LIU, Z., WANG, J., ZHOU, C. & CHEN, K. 2011. Morphological, anatomical, and physiological characteristics involved in development of the large culm train in rice. *Aust J Crop Sci,* 5**,** 1356-1363.

WU, Z., CHEN, L., YU, Q., ZHOU, W., GOU, X., LI, J. & HOU, S. 2019. Multiple transcriptional factors control stomata development in rice. *New Phytol,* New Phytol.

XIA, H., LUO, Z., XIONG, J., MA, X., LOU, Q., WEI, H., QIU, J., YANG, H., LIU, G., FAN, L., CHEN, L. & LUO, L. 2019. Bi-directional Selection in Upland Rice Leads to Its Adaptive Differentiation from Lowland Rice in Drought Resistance and Productivity. *Mol Plant,* 12**,** 170-184.

XIA, K., OU, X., TANG, H., WANG, R., WU, P., JIA, Y., WEI, X., XU, X., KANG, S. H., KIM, S. K. & ZHANG, M. 2015. Rice microRNA osa-miR1848 targets the obtusifoliol 14alpha-demethylase gene OsCYP51G3 and mediates the biosynthesis of phytosterols and brassinosteroids during development and in response to stress. *New Phytol,* 208**,** 790-802.

XIANG, Y., HUANG, Y. & XIONG, L. 2007. Characterization of stress-responsive CIPK genes in rice for stress tolerance improvement. *Plant Physiol,* 144**,** 1416-28.

XU, G., FAN, X. & MILLER, A. J. 2012. Plant nitrogen assimilation and use efficiency. *Annu Rev Plant Biol,* 63**,** 153-82.

XU, J. J., ZHANG, X. F. & XUE, H. W. 2016. Rice aleurone layer specific OsNF-YB1 regulates grain filling and endosperm development by interacting with an ERF transcription factor. *J Exp Bot,* 67**,** 6399-6411.

XU, J. L., WANG, Y., ZHANG, F., WU, Y., ZHENG, T. Q., WANG, Y. H., ZHAO, X. Q., CUI, Y. R., CHEN, K., ZHANG, Q., LIN, H. X., LI, J. Y. & LI, Z. K. 2015. SS1 (NAL1)- and SS2-Mediated Genetic Networks Underlying Source-Sink and Yield Traits in Rice (Oryza sativa L.). *PLoS One,* 10**,** e0132060.

XU, Y., ZHANG, S., GUO, H., WANG, S., XU, L., LI, C., QIAN, Q., CHEN, F., GEISLER, M., QI, Y. & JIANG DE, A. 2014. OsABCB14 functions in auxin transport and iron homeostasis in rice (Oryza sativa L.). *Plant J,* 79**,** 106-17.

XUE, W., XING, Y., WENG, X., ZHAO, Y., TANG, W., WANG, L., ZHOU, H., YU, S., XU, C., LI, X. & ZHANG, Q. 2008. Natural variation in Ghd7 is an important regulator of heading date and yield potential in rice. *Nat Genet,* 40**,** 761-7.

YAISH, M. W., EL-KEREAMY, A., ZHU, T., BEATTY, P. H., GOOD, A. G., BI, Y. M. & ROTHSTEIN, S. J. 2010. The APETALA-2-like transcription factor OsAP2-39 controls key interactions between abscisic acid and gibberellin in rice. *PLoS Genet,* 6**,** e1001098.

YAMAGUCHI, T., NAGASAWA, N., KAWASAKI, S., MATSUOKA, M., NAGATO, Y. & HIRANO, H. Y. 2004. The YABBY gene DROOPING LEAF regulates carpel specification and midrib development in Oryza sativa. *Plant Cell,* 16**,** 500-9.

YAMAMOTO, S., SUZUKI, K. & SHINSHI, H. 1999. Elicitor-responsive, ethylene-independent activation of GCC box-mediated transcription that is regulated by both protein phosphorylation and dephosphorylation in cultured tobacco cells. *Plant J,* 20.

YAMAMURO, C., IHARA, Y., WU, X., NOGUCHI, T., FUJIOKA, S., TAKATSUTO, S., ASHIKARI, M., KITANO, H. & MATSUOKA, M. 2000. Loss of function of a rice brassinosteroid insensitive1 homolog prevents internode elongation and bending of the lamina joint. *Plant Cell,* 12**,** 1591-606.

YAN, W. H., WANG, P., CHEN, H. X., ZHOU, H. J., LI, Q. P., WANG, C. R., DING, Z. H., ZHANG, Y. S., YU, S. B., XING, Y. Z. & ZHANG, Q. F. 2011. A major QTL, Ghd8, plays pleiotropic roles in regulating grain productivity, plant height, and heading date in rice. *Mol Plant,* 4**,** 319-30.

YAN, Y., STOLZ, S., CHETELAT, A., REYMOND, P., PAGNI, M., DUBUGNON, L. & FARMER, E. E. 2007. A downstream mediator in the growth repression limb of the jasmonate pathway. *Plant Cell,* 19**,** 2470-83.

YANG, C., LI, D., MAO, D., LIU, X., JI, C., LI, X., ZHAO, X., CHENG, Z., CHEN, C. & ZHU, L. 2013. Overexpression of microRNA319 impacts leaf morphogenesis and leads to enhanced cold tolerance in rice (Oryza sativa L.). *Plant Cell Environ,* 36**,** 2207-18.

YANG, W., YOON, J., CHOI, H., FAN, Y., CHEN, R. & AN, G. 2015a. Transcriptome analysis of nitrogen-starvation-responsive genes in rice. *BMC Plant Biol,* 15**,** 31.

YANG, Y. W., CHEN, H. C., JEN, W. F., LIU, L. Y. & CHANG, M. C. 2015b. Comparative Transcriptome Analysis of Shoots and Roots of TNG67 and TCN1 Rice Seedlings under Cold Stress and Following Subsequent Recovery: Insights into Metabolic Pathways, Phytohormones, and Transcription Factors. *PLoS One,* 10**,** e0131391.

YE, H., DU, H., TANG, N., LI, X. & XIONG, L. 2009. Identification and expression profiling analysis of TIFY family genes involved in stress and phytohormone responses in rice. *Plant Mol Biol,* 71**,** 291-305.

YI, S. Y., LEE, H. Y., KIM, H. A., LIM, C. J., KIM, W. B., JANG, H. A., JEON, J.-S. & KWON, S.-Y. 2013. Microarray Analysis of bacterial blight resistance 1 mutant rice infected with Xanthomonas oryzae pv. oryzae. *Plant Breed Biotech,* 1**,** 354-365.

YIN, X., LIU, X., XU, B., LU, P., DONG, T., YANG, D., YE, T., FENG, Y. Q. & WU, Y. 2019. OsMADS18, a membrane-bound MADS-box transcription factor, modulates plant architecture and the ABA response in rice. *J Exp Bot*.

YOKOTANI, N., ICHIKAWA, T., KONDOU, Y., IWABUCHI, M., MATSUI, M., HIROCHIKA, H. & ODA, K. 2013. Role of the rice transcription factor JAmyb in abiotic stress response. *J Plant Res,* 126**,** 131-9.

YOSHIKAWA, T., EIGUCHI, M., HIBARA, K., ITO, J. & NAGATO, Y. 2013. Rice slender leaf 1 gene encodes cellulose synthase-like D4 and is specifically expressed in M-phase cells to regulate cell proliferation. *J Exp Bot,* 64**,** 2049-61.

YU, C., CHEN, H.-M., TIAN, F., BI, Y.-M., STEVEN, R. J., JAN, L. E. & HE, C.-Y. 2015. Identification of differentially-expressed genes of rice in overlapping responses to bacterial infection by Xanthomonas oryzae pv. oryzae and nitrogen deficiency. *J Integr Agr,* 14**,** 888-899.

YU, D., RANATHUNGE, K., HUANG, H., PEI, Z., FRANKE, R., SCHREIBER, L. & HE, C. 2008. Wax Crystal-Sparse Leaf1 encodes a beta-ketoacyl CoA synthase involved in biosynthesis of cuticular waxes on rice leaf. *Planta,* 228**,** 675-85.

YU, G.-H., HUANG, S.-C., HE, R., LI, Y.-Z. & CHENG, X.-G. 2018. Transgenic Rice Overexperessing a Tomato Mitochondrial Phosphate Transporter, SlMPT3;1, Promotes Phosphate Uptake and Increases Grain Yield. *J Plant Biol,* 61**,** 383-400.

YU, H., JIAO, B. & LIANG, C. 2017. Systematic analysis of RNA-seq-based gene co-expression across multiple plants. *bioRxiv*.

YU, Y., STEINMETZ, A., MEYER, D., BROWN, S. & SHEN, W. H. 2003. The tobacco A-type cyclin, Nicta;CYCA3;2, at the nexus of cell division and differentiation. *Plant Cell,* 15**,** 2763-77.

YUENYONG, W., CHINPONGPANICH, A., COMAI, L., CHADCHAWAN, S. & BUABOOCHA, T. 2018. Downstream components of the calmodulin signaling pathway in the rice salt stress response revealed by transcriptome profiling and target identification. *BMC Plant Biol,* 18**,** 335.

ZAREMBINSKI, T. I. & THEOLOGIS, A. 1994. Ethylene biosynthesis and action: a case of conservation. *Plant Mol Biol,* 26**,** 1579-97.

ZHANG, C., LI, C., LIU, J., LV, Y., YU, C., LI, H., ZHAO, T. & LIU, B. 2017a. The OsABF1 transcription factor improves drought tolerance by activating the transcription of COR413-TM1 in rice. *J Exp Bot,* 68**,** 4695-4707.

ZHANG, X., LI, J., LIU, A., ZOU, J., ZHOU, X., XIANG, J., RERKSIRI, W., PENG, Y., XIONG, X. & CHEN, X. 2012. Expression profile in rice panicle: insights into heat response mechanism at reproductive stage. *PLoS One,* 7**,** e49652.

ZHANG, Y., LUO, L., XU, C., ZHANG, Q. & XING, Y. 2006. Quantitative trait loci for panicle size, heading date and plant height co-segregating in trait-performance derived near-isogenic lines of rice (Oryza sativa). *Theor Appl Genet,* 113**,** 361-8.

ZHANG, Y., YU, C., LIN, J., LIU, J., LIU, B., WANG, J., HUANG, A., LI, H. & ZHAO, T. 2017b. OsMPH1 regulates plant height and improves grain yield in rice. *PLoS One,* 12**,** e0180825.

ZHANG, Y., ZHAO, J., LI, Y., YUAN, Z., HE, H., YANG, H., QU, H., MA, C. & QU, S. 2016. Transcriptome Analysis Highlights Defense and Signaling Pathways Mediated by Rice pi21 Gene with Partial Resistance to Magnaporthe oryzae. *Front Plant Sci,* 7**,** 1834.

ZHANG, Y. C., YU, Y., WANG, C. Y., LI, Z. Y., LIU, Q., XU, J., LIAO, J. Y., WANG, X. J., QU, L. H., CHEN, F., XIN, P., YAN, C., CHU, J., LI, H. Q. & CHEN, Y. Q. 2013. Overexpression of microRNA OsmiR397 improves rice yield by increasing grain size and promoting panicle branching. *Nat Biotechnol,* 31**,** 848-52.

ZHANG, Y. J., ZHANG, Y., ZHANG, L. L., HUANG, H. Y., YANG, B. J., LUAN, S., XUE, H. W. & LIN, W. H. 2018. OsGATA7 modulates brassinosteroids-mediated growth regulation and influences architecture and grain shape. *Plant Biotechnol J,* 16**,** 1261-1264.

ZHAO, L., TAN, L., ZHU, Z., XIAO, L., XIE, D. & SUN, C. 2015. PAY1 improves plant architecture and enhances grain yield in rice. *Plant J,* 83**,** 528-36.

ZHAO, L., WANG, P., HOU, H., ZHANG, H., WANG, Y., YAN, S., HUANG, Y., LI, H., TAN, J., HU, A., GAO, F., ZHANG, Q., LI, Y., ZHOU, H., ZHANG, W. & LI, L. 2014. Transcriptional regulation of cell cycle genes in response to abiotic stresses correlates with dynamic changes in histone modifications in maize. *PLoS One,* 9**,** e106070.

ZHOU, Y., CAI, H., XIAO, J., LI, X., ZHANG, Q. & LIAN, X. 2009. Over-expression of aspartate aminotransferase genes in rice resulted in altered nitrogen metabolism and increased amino acid content in seeds. *Theor Appl Genet,* 118**,** 1381-90.

ZHU, X. G., LONG, S. P. & ORT, D. R. 2008. What is the maximum efficiency with which photosynthesis can convert solar energy into biomass? *Curr Opin Biotechnol,* 19**,** 153-9.

ZI, J., ZHANG, J., WANG, Q., ZHOU, B., ZHONG, J., ZHANG, C., QIU, X., WEN, B., ZHANG, S., FU, X., LIN, L. & LIU, S. 2013. Stress responsive proteins are actively regulated during rice (Oryza sativa) embryogenesis as indicated by quantitative proteomics analysis. *PLoS One,* 8**,** e74229.

ZULFUGAROV, I. S., TOVUU, A., KIM, C.-Y., XUAN VO, K. T., KO, S. Y., HALL, M., SEOK, H.-Y., KIM, Y.-K., SKOGSTROM, O., MOON, Y.-H., JANSSON, S., JEON, J.-S. & LEE, C.-H. 2016. Enhanced resistance of PsbS-deficient rice (Oryza sativa L.) to fungal and bacterial pathogens. *J Plant Biol,* 59**,** 616-626.

CHAPTER 5: Conclusion – *mpg1* is a novel rice mutant that experiences ectopic expression of an AP2/EREBP transcription factor producing plants exhibiting a pleiotropic phenotype enhancing plant productivity

Generation and discovery of *mpg1*

The increasing global population has brought noticeable concern to energy and food securities (Chapter 1). Plants potentially offer a means by which energy and food can be sustainably generated. In an effort to try and enhance plant productivity by measurement of biomass per unit area, a strategy for enhancing photosynthesis was investigated (Chapter 2). Specifically, we aimed to alter carbon allocation as a way to enhance net productivity and stimulate greater growth. This was based on the idea that increase phloem loading of sucrose would thereby decrease sugar content in mesophyll cells simultaneously increasing photosynthesis and providing greater levels of carbon to sink tissues. Transgenic rice containing a T-DNA expression cassette consisting of a hyper-active sucrose symporter ($AtSUC1_{H65K}$) being driven by a companion cell specific promoter (*CmGAS1*) were generated.

Interestingly, plants containing the T-DNA expression cassette expressing the transgene resulted in plants with decreased biomass compared to wild-type plants. Since the inception of this study other researchers have attempted to assess the effects of enhanced phloem loading by similar means in *Arabidopsis*. Their results remained similar to our observations, as plants designed to have enhanced phloem loading exhibited reduced biomass compared to wild-type as well. Their studies revealed that these transgenic plants perceived phosphate starvation. The addition of supplemental phosphate to plants engineered to have enhanced phloem loading of sucrose return to growth similar to that of wild-type but do not exceed it (Dasgupta et al., 2014).

They hypothesized this was the result of inappropriate ratios of nutrients, affecting metabolism and growth.

While growing multiple independent transgenic rice plants utilizing the *CmGAS1* promoter directing expression of *AtSUC1*$_{H65K}$, a single plant with noticeably larger biomass was found (Chapter 2). Further investigation of this plant showed that it possessed a bi-laterally truncated version of the T-DNA expression cassette and did not contain the *AtSUC1*$_{H65K}$ transgene. The T-DNA was localized to an intergenic region on chromosome 8 not disrupting any known or functional annotated elements. Sequencing of the insertion revealed that it is 3190bp including the duplicated *35S* promoter driving the gene coding for hygromycin resistance, as well as roughly the first 1kb of the *CmGAS1* promoter sequence. The presence of this T-DNA insertion was monitored and correlated with the increased biomass trait over several segregating generations. Analysis of gene expression proximal to the truncated T-DNA revealed that a gene roughly 7kb downstream of the insertion site exhibited notably higher expression then in wild-type and wild-type null segregant plants, which had little to no expression across vegetative development in leaf and stem tissues. Because of the enhanced biomass observed in this mutant we named it *mpg1* (*Makes-Plants-Gigantic-1*), which also pays tribute to the intended aim for biofuel productions as 'MPG' is a commonly known acronym for 'miles per gallon'. *MPG1* is part of the AP2/EREBP transcription factor superfamily and is a previously undescribed member.

Comprehensive phenotyping of *mpg1*

To better characterize *mpg1* several selfed segregating generations were grown and observed (Chapter 3). The mutant *mpg1* not only generated greater dry biomass, but also produced greater seed yield as well. The biomass accumulation was a result of greater plant

girth, plant height, and leaf size. In particular, the greatest metric was plant girth (larger tillers). The increase in seed yield appears to have come from number of spikelets per plant as *mpg1* plants generated similar number of tillers and panicles to wild-type under optimally grown conditions. The spikelets of *mpg1* had a tendency to also produce long and abundant awns, while wild-type did not. Under greenhouse conditions *mpg1* experiences a delay in flowering by roughly two weeks. Beyond primary growth *mpg1* plants were also able to more successfully ratoon, having higher survival rates and the ability to generate greater biomass and seed yield during the secondary harvest. Besides exhibiting greater growth and seed yield, *mpg1* plants also exhibited a degree of abiotic stress tolerance. *mpg1* plants grown under sub-optimal conditions grew well while growth of wild-type plants were inhibited. Further investigation of *mpg1* under control of specific stressors showed that they were able to accumulate greater biomass compared to wild-type under prolonged drought and salt stress, however experienced a similar time of onset and degree of symptomatic stress response, making it difficult to ascertain whether the greater biomass was a result from enhanced growth or from stress tolerance.

Expression analysis of *MPG1* showed that it was present in leaf and stem tissue throughout development, and in panicle tissue during reproductive development in *mpg1*. Wild-type and wild-type null segregants showed little to no expression in the same tissues. While at first we predicted that the remnant *35S* promoter present in the T-DNA insertion was driving constitutive overexpression of *MPG1* in the mutant, we found little to no expression of *MPG1* in root tissues. This suggests that *MPG1* expression in the mutant is being controlled by a complex interplay of elements not yet understood. Sequences neighboring the site of the T-DNA insertion are likely playing a role in its ectopic expression. A combinatorial effect of the duplicated *35S*

promoter along with surrounding genomic sequences or even the remaining portion of the *CmGAS1* promoter might be driving the expression of *MPG1* resulting in the mutant phenotype.

The delay in developmental transition to flowering in *mpg1* is a challenge while characterizing this mutant. Because of the delay in development, *mpg1* has a longer period of vegetative growth than wild-type plants. It could be that the resulting increase in biomass observed in *mpg1* is due to an increase in vegetative growth phase. Additionally, assessing the effects of prolonged stress response is difficult because *mpg1* and wild-type are under different developmental states during a portion of exposure affecting our understanding of stress response.

Transcriptomic analysis of *mpg1*

Two rounds of RNA-sequencing were performed on homozygous *mpg1* and wild-type null-segregant plants to gauge global transcriptional differences (Chapter 4). Whole tiller tissue was selected for these experiments, as the greatest difference concerning biomass measurements between *mpg1* and wild-type plants is in the form of plant girth. We selected timepoints proximal to the stage were *mpg1* began to outgrow wild-type plants. This difference was noted at around 42 days post-planting, therefore we selected timepoints of 32 and 42 days post-planting. Further dissection of the developmental shift difference in *mpg1* to wild-type revealed that reproductive transition in wild-type occurs around the 42 days post-planting timepoint. Because of this, and differential growth conditions, only minor aspects of the transcriptomic differences were evaluated at this stage - therefore we focused on the 32 days post-planting timepoint to more accurately explore differences of these plants during similar stages of development.

There were a total of 297 differentially regulated genes between *mpg1* and wild-type at 32 days post-planting. Of these genes 106 were up-regulated and 191 were down-regulated. There were several genes that were co-regulated between the 32 and 42 days post planting

timepoints that should be considered potentially critical with generating the phenotype observed in *mpg1* as they might represent continually affected genes throughout growth and development. *MPG1* was amongst this list as the greatest differentially up-regulated gene in *mpg1*. This along with proximity of the gene to the T-DNA insertion site, phenotyping of segregating and backcrossed lines, and similar phenotypic outcomes in individual lines constitutively overexpressing *MPG1* has increased our confidence that the T-DNA insertion dependent ectopic expression of *MPG1* is responsible for he mutant phenotype.

Because the phenotype of *mpg1* is pleiotropic, the differentially regulated genes were parsed by gene ontology enrichment and individually investigated to try to generate hypotheses that could potentially narrow focus of genes and pathways responsible for, or playing a role in, generating the mutant phenotypes. Several interesting gene ontology categories that were enriched from the differentially expressed genes included transcription factor activity, flower development, stress response, metabolism, DNA metabolism, and the cell cycle. Of these, all but DNA metabolism, and cell cycle were up-regulated genes in HM-*mpg1* compared to WT-ns. Categories classically linked to several traits present in *mpg1* were assessed by evaluation of genes associated with defoliation response (as a proxy means to evaluate mechanisms potentially connected to ratooning capability), cell wall formation, and hormone regulation (to investigate biomass and stress response regulators).

Numerous genes were identified that potentially contribute to different characteristics present in the *mpg1* phenotype. The nature of a pleiotropic phenotype could be the result of differential expression of genes responsible for transcription factor activity. *mpg1* had 12 up-regulated and 16 down-regulated genes from this ontological category. These genes could be

causative for driving major phenotypic outcomes. A number of these genes were also found in other categories such as flower development and stress response.

Not surprisingly differentially regulated genes for the ontological categories of flower development and stress response were enriched in *mpg1*. No genes were up-regulated, while 8 genes were down-regulated in *mpg1* within this category. Of the differentially regulated genes involved in flower development several genes stuck out as potentially contributing to the delay in flowering observed in *mpg1*. *mpg1* exhibits down-regulation of two of the flowering activation *FRUITFUL* (FUL)-clade gene members, *OsMADS14* (*Os03g54160*) and *OsMADS18* (*Os07g41370*). These genes play a role in transition of rice from vegetative to reproductive development along with flower formation (Jeon et al., 2000, Kim et al., 2007, Fornara et al., 2004, Yin et al., 2019). The reduced expression of these genes could likely produce the delay in flowering phenotype observed in *mpg1*. Both of these genes also contain GCC-box motifs within -2000 of their TSS, allowing for speculation that MPG1 might directly target and down-regulate these genes.

Thirty up-regulated and zero down-regulated genes were found in the enriched category of stress response. The majority of these genes have been linked to abiotic and biotic stress response, in particular to drought, and bacterial and fungal pathogens. Ten of these genes had greater than a 2.0 $\log_2$-fold change in HM-*mpg1* compared to WT-ns. These genes consist of several transcriptional regulators (JAZ, AP2/EREBP, WRKY), and genes involved in metabolism (3-ketoacyl-CoA synthase, monooxygenase, cinnamoyl CoA reductase, cytochrome P450). Of the genes from this enrichment category, the greatest differentially regulated gene, (*Os11g05380*) encoding a cytochrome P450, experienced a 4.25 $\log_2$-fold change up-regulation in HM-*mpg1* compared to WT-ns. Although this gene doesn't contain a GCC-box in its promoter

region and is unlikely to be a direct target of *MPG1* its high level of differential expression is of potential interest. However, a number of the other genes from this category do contain GCC-box motifs in their promoter region, and could inherently be directly affected by targeting of MPG1.

Interestingly, of the down regulated genes in *mpg1*, the categories of DNA metabolism and cell cycle were also enriched. This was counterintuitive to our expectations because a decrease in activity of these genes would suggest potentially less growth occurring in *mpg1* compared to wild-type. However, this observation could be the result of numerous factors. These genes could be down regulated in *mpg1* due to the developmental shift observed between the mutant and wild-type plants as a cause of faster cell proliferation from arising flowering tissues. Alternatively, these genes could be down regulated due to distorted regulation of the cell cycle in *mpg1* resulting in altered cell size/number potentially causing the enhanced biomass seen in the mutant. Many cyclins, mini chromosome maintenance, and core histone domain containing proteins were affected. Again, a number of these genes possess a GCC-box within their assessed promoter region, potentially acting directly to influence cell cycle regulation.

Biomass is a difficult and complex trait to assess. *mpg1* plants accumulate more overall above ground dry biomass through greater height, tiller width, and leaf size. Based on these traits other classes of genes were investigated to assess their possible influence on these characteristics. In particular, genes involving cell wall biosynthesis and hormone regulation. A number of genes were differentially regulated for lignin biosynthesis, cellulose biosynthesis, and cell wall metabolism. The particular expression of these genes could potentially be contributing to greater biomass in *mpg1*.

Growth metrics are also commonly regulated by hormones. *mpg1* experienced differential expression of genes involved in auxin transport, brassinosteroid signal transduction,

jasmonic acid signal transduction, cytokinin biosynthesis, and ethylene biosynthesis. The differential expression of one or more of these genes might be directing enhanced growth. Of the genes playing a role in hormone regulation, a noticeably high number of ZIM-domain containing proteins (JAZ proteins) were differentially regulated in *mpg1* (Chapter 4, APPENDIX). JAZ proteins function to transcriptionally repress jasmonic acid (JA) responses in plants. With a high number of these genes up-regulated at 32 days post-planting, *mpg1* might have a degree of JA insensitivity during vegetative growth resulting in increased plant growth through reduced influence of the JA defense response pathway plant growth inhibition. In support of this, studies in *Arabidopsis* evaluating the effects of loss of function JAZ mutants exhibited JA sensitivity and reduced plant size, suggesting that JAZ proteins influence growth (Guo et al., 2018). Complementary to this, overexpression of JAZ proteins have led to increased plant growth, and grain size in *Arabidopsis* and rice through desensitizing plants to JA (Hakata et al., 2012, Hakata et al., 2017). Because of the high number of affected ZIM-domain containing proteins it is prospective that *MPG1* is manipulating the JA signaling pathway influencing plant growth in *mpg1*. Additionally, several of these genes contain GCC-box motifs within their promoter regions as well. Investigating the effect of exogenous hormone application could provide evidence to whether these genes are influencing plant growth in *mpg1*.

mpg1 plants also experienced a greater seed yield. During assessment of plants grown under sub-optimal conditions, greater yield accumulation was noted compared to wild-type. This was originally believed to be a result of greater tiller number, as individual tillers are capable of generating panicle tissue. During optimum growth conditions *mpg1* didn't generate a greater number of tillers or panicles yet still exhibited enhanced seed yield (~1.6-fold). Investigation of 1000 grain seed showed that there wasn't a significant difference in grain size or weight between

mpg1 and wild-type. Therefore, only a greater number of spikelets per panicle could explain the increased seed yield observed. Previously, increased grains per panicle have been described through increased accumulation of cytokinins (Ashikari et al., 2005, Li et al., 2013b). *mpg1* exhibited an up-regulation in cytokinin biosynthesis gene, *OsCKX5* (*Os01g56810*). Perhaps the expression of this gene influences in the accumulation of greater grains per panicle in *mpg1*. Cytokinin also may be contributing to overall growth by stimulating the cell cycle (Werner et al., 2001, del Pozo et al., 2005). This gene also contains a GCC-box motif within its promotor region, again allowing for speculation on MPG1's direct influence.

Beyond accumulation of primary growth, *mpg1* exhibited a more robust ratoon response compared to wild-type plants. Ratooning is the agricultural practice of cutting back stems of plants in order to re-generate seed bearing tissues for additional harvests. Not only did *mpg1* plants have a greater survivability-rate compared to wild-type but they were able to accumulate greater biomass, earlier time to flowering, and overall seed yield during second harvests. Although assessments of *mpg1* global gene expression were not performed post-harvest, evaluation of differentially expressed genes during vegetative growth might remain elevated during ratooning and thus lead to discovery of molecular mechanisms controlling base ratooning and elevated growth in general. There have been no previous reports investigating ratooning ability at the transcriptome level, so we compared differentially regulated genes in *mpg1* to genes expressed during response to defoliation as a preliminary base of exploration. Defoliation is similar to ratooning in that the primary leaf and stem tissue are cut away to regenerate tissue. Interestingly, a high number of genes differentially regulated in *mpg1* were similarly co-expressed to genes up-regulated during response to defoliation. Numerous genes including transcription factors, metabolism, cell wall biosynthesis, and hormone regulation were amongst

the co-regulated genes. The ectopic expression of *MPG1* in the mutant might be influencing the expression of these genes. One or many of these genes might play a contributing role to the greater success in ratooning ability or biomass and seed yield accumulation. A number of these genes contain GCC-box motifs in their promoter region and might act as direct targets of MPG1.

Future aims and progression of this study

The evaluation of several individual F_2BC_1 *mpg1* backcrosses showed that the pleiotropic phenotype persisted. Although the presence of the T-DNA and ectopic expression of *MPG1* correlated with the pleiotropic phenotype of *mpg1*, it is imperative that further backcrosses be performed to insure that that no other factors are responsible for generating individual characteristics of *mpg1*.

The generation of rice constitutively overexpressing *MPG1* through use of the *ZmUBI* promoter created plants with larger biomass and delayed flowering compared to other plants that arose from tissue culture not containing the *MPG1* overexpression cassette during the T_0 generation. The generation of aspects of the phenotype in constitutive overexpression transgenics increases our confidence that *MPG1* is responsible for the phenotype observed in *mpg1*. However, plants constitutively overexpressing *MPG1* also resulted in the inability for seeds to germinate in generations subsequent to T_0. That observation shows constitutive overexpression does not replicate the full phenotype. Indeed, the expression of *MPG1* in *mpg1* is not constitutively expressed as little to no expression is measured in root tissue of mutants. *MPG1*'s expression in this mutant is likely spatiotemporally regulated, and its unique pattern or level of expression (present in the mutant) is necessary for successful manifestation of the phenotype. Uncovering the factors controlling *MPG1*'s expression in *mpg1* should allow for appropriate

recapitulation of the phenotype to confirm that its particular expression results in the pleiotropic characteristics present in *mpg1*.

Determining the pattern and level of *MPG1*'s expression in *mpg1* and subsequent recapitulation of the phenotype by independently validating its function through generating new transgenic plants with the desired nature of expression will be time consuming. As a means to support our findings several alternatives using CRISPR-CAS could be performed. The full T-DNA insertion could be deleted in *mpg1* to validate that its presence is responsible for the ectopic expression of *MPG1*. Additionally, the full T-DNA insertion present in *mpg1* could be inserted into wild-type in the same location and fashion found in *mpg1* to verify that its presence results in the same ectopic expression and phenotype as the mutant. Furthermore, if *MPG1* is not necessary for plant growth and survival, *MPG1* could be deleted in *mpg1* to verify if the phenotype can be removed. Together these results could provide greater evidence in support of our hypothesis that the T-DNA insertion present in *mpg1* results in ectopic expression of *MPG1* resulting in the mutant phenotype.

Localizing the expression of *MPG1*, by means of northern blot tissue printing or *in-situ* hybridization could aid in uncovering where and when its expression occurs in the mutant. Further, it might give insight to important sequence elements directing expression in *mpg1*. Sequence elements remaining in the bi-laterally truncated T-DNA and surrounding genomic sequences proximal to the T-DNA insertion likely drive unique expression of *MPG1*. Using CRISPR-CAS, portions of sequence in *mpg1* could be deleted to gauge sequence necessary and sufficient in generating the phenotype. Designing new expression cassettes utilizing different portions of these sequences could also prove useful in trying to recapitulate the phenotype. The effects of site specific integration of the T-DNA could also be important in generating the

phenotype and should also be considered, as perturbed genomic sequences could be disrupting important regulatory elements or chromatin structure influencing the nature of expression. Once the nature of expression is determined, successful replication of the phenotype can be achieved.

Because *MPG1* codes for a transcription factor, identifying its targets will be important. Genes targeted by MPG1 will provide critical information pertaining to what exactly is being affected in *mpg1* to produce the phenotype. Transcriptional regulation is sensitive to environmental queues, stages of development, and tissue specificity. Uncovering MPG1's targets will remain difficult to ascertain until the pattern of *MPG1*'s expression has been determined in the mutant. Although the constitutive overexpression of *MPG1* in rice resulted in loss of germination in T_1 generation transformants, T_0 plants exhibited several traits indicative of *mpg1*. Therefore, alternative means of controlled expression might be useful in evaluating targets of MPG1 under ectopic expression. For example, through the use of an inducible promoter directing expression of *MPG1* (tagged or fused to trackable proteins), ChIP-seq could be performed investigating several tissues and/or timepoints while controlling the induction or repression of *MPG1*. Although this would not be effective in directly evaluating MPG1's targets in the mutant *mpg1*, it could provide primary insight into targeted genes that are influenced by *MPG1*'s ectopic expression. An inducible promoter system could also be used to support that ectopic expression of *MPG1* results in aspects of phenotype originally observed in *mpg1*, while also gaining a better understanding of the tissues or timepoints where *MPG1* expression is necessary, sufficient, or detrimental in formation of the phenotype.

Because *mpg1* saw large differences in the general architecture of the plant (girth, height, leaf size, and presence of awns), it would be useful to evaluate and observe how specific tissues might differ between it and wild-type at a microscopic scale. Analysis of cell size/cell number in

various tissue types should be assessed throughout development to determine if *mpg1*'s accumulation of biomass is from either or both of these factors and where. Due to the vast morphological differences observed in *mpg1* it is likely that the meristem is being influenced in some fashion, and should be an important area of focus.

Evaluation of below-ground tissues could prove useful as well. Till now little formal observations have been made on differences of root tissues in *mpg1* compared to wild-type. General observations have suggested that *mpg1* might accumulate greater amount of root tissue. Dissection and evaluation of these tissues will be enlightening, as the structure, orientation, and general architecture might be affected.

Additional investigations exploring the effect of abiotic and biotic stress should be assessed. Although we observed greater biomass accumulation, seed yield, and a better general appearance of *mpg1* health under sub-optimal conditions, effects of specific stressors revealed a similar degree of symptomatic stress response with the exception of biomass accumulation. The delay in development further impacted our ability to assess prolonged exposure to stress. It is possible that *mpg1* plants process and acquire nutrients differently than wild-type, as the sub-optimal conditions were made up of a different composition of elements for fertilizer solution. Evaluation of the composition of matter in nutrient solutions and subsequent accumulation in plants could provide evidence of stress tolerance or altered nutrient acquisition. Furthermore, heightened levels of *MPG1*'s expression were shown in wild-type plants after exposure to *Xanthomonas oryzae* pv. *oryzae*, suggesting that it might also natively function in biotic stress response as well. Interestingly *mpg1* plants subjected to this pathogen didn't show any sign of enhanced tolerance to infection by measurement of legion length compared to wild-type.

Evaluation of *mpg1*'s effect with additional pathogens could be interesting to evaluate its ability concerning biotic stress.

The original intent of this research was to better understand molecular mechanisms regulating biomass and plant productivity to aid with fuel and food securities. Translational biotechnological assessments could be performed to evaluate the potential for *MPG1*'s use in the commercial sector. *MPG1* is unique to rice sharing no true homolog (having no sequence >56% protein sequence identity) in other plant species. This could mean that *MPG1* might not function in other types of plants, however if *MPG1*'s targets and mechanisms of action remain well conserved in other plant species, its ectopic expression could lead to similar phenotypic outcomes. Other rice specific AP2/EREBP transcription factors have given rise to functional phenotypes in alternate species. Expression of rice *SUB1A* and *SUB1C* alleles in *Arabidopsis* resulted in abiotic stress responses (Pena-Castro et al., 2011). Expression of *OsDREB1D*, a unique AP2 to rice with no true homolog, in *Arabidopsis* results in plants with enhanced abiotic stress tolerance (Zhang et al., 2009). This same phenomenon has also been observed in other plants as well. For example, the novel coconut AP2/EREBP, *SodERF3* which shares extremely low homology with other sequenced genes, when overexpressed in tobacco results in plants with enhanced salt and drought tolerance (Trujillo et al., 2008). This evidence suggests that *MPG1* could serve to improve plant productivity of commercially cultivated crops through its ectopic expression. Although our assessments of constitutive overexpression of *MPG1* in rice resulted in the inability to produce subsequent generations, current investigation by overexpression of *MPG1* through the use of a duplicated 35S promoter in *Arabidopsis* are currently undergoing; T_2 *Arabidopsis* seeds were able to germinate under selective media, provoking further analysis. Additional evaluation of plants with *MPG1* overexpression, *MPG1* ectopic expression, or *mpg1*-

specific expression could be useful in determining its commercial potential. Continued assessment of *mpg1* and the individual components generating its desirable phenotypes could aid in novel discoveries with translational application.

BIBLIOGRAPHY

ASHIKARI, M., SAKAKIBARA, H., LIN, S., YAMAMOTO, T., TAKASHI, T., NISHIMURA, A., ANGELES, E. R., QIAN, Q., KITANO, H. & MATSUOKA, M. 2005. Cytokinin oxidase regulates rice grain production. *Science,* 309**,** 741-5.

DASGUPTA, K., KHADILKAR, A. S., SULPICE, R., PANT, B., SCHEIBLE, W. R., FISAHN, J., STITT, M. & AYRE, B. G. 2014. Expression of Sucrose Transporter cDNAs Specifically in Companion Cells Enhances Phloem Loading and Long-Distance Transport of Sucrose but Leads to an Inhibition of Growth and the Perception of a Phosphate Limitation. *Plant Physiol,* 165**,** 715-731.

DEL POZO, J. C., LOPEZ-MATAS, M. A., RAMIREZ-PARRA, E. & GUTIERREZ, C. 2005. Hormonal control of the plant cell cycle. *Physiol Plantarum,* 123**,** 173-183.

FORNARA, F., PARENICOVA, 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 Physiol,* 135**,** 2207-19.

GUO, Q., YOSHIDA, Y., MAJOR, I. T., WANG, K., SUGIMOTO, K., KAPALI, G., HAVKO, N. E., BENNING, C. & HOWE, G. A. 2018. JAZ repressors of metabolic defense promote growth and reproductive fitness in Arabidopsis. *Proc Natl Acad Sci U S A,* 115**,** E10768-E10777.

HAKATA, M., KURODA, M., OHSUMI, A., HIROSE, T., NAKAMURA, H., MURAMATSU, M., ICHIKAWA, H. & YAMAKAWA, H. 2012. Overexpression of a rice TIFY gene increases grain size through enhanced accumulation of carbohydrates in the stem. *Biosci Biotechnol Biochem,* 76**,** 2129-34.

HAKATA, M., MURAMATSU, M., NAKAMURA, H., HARA, N., KISHIMOTO, M., IIDA-OKADA, K., KAJIKAWA, M., IMAI-TOKI, N., TOKI, S., NAGAMURA, Y., YAMAKAWA, H. & ICHIKAWA, H. 2017. Overexpression of TIFY genes promotes plant growth in rice through jasmonate signaling. *Biosci Biotechnol Biochem,* 81**,** 906-913.

JEON, J. S., LEE, S., JUNG, K. H., JUN, S. H., JEONG, D. H., LEE, J., KIM, C., JANG, S., YANG, K., NAM, J., AN, K., HAN, M. J., SUNG, R. J., CHOI, H. S., YU, J. H., CHOI, J. H., CHO, S. Y., CHA, S. S., KIM, S. I. & AN, G. 2000. T-DNA insertional mutagenesis for functional genomics in rice. *Plant J,* 22**,** 561-70.

KIM, S. L., LEE, S., KIM, H. J., NAM, H. G. & AN, G. 2007. OsMADS51 is a short-day flowering promoter that functions upstream of Ehd1, OsMADS14, and Hd3a. *Plant Physiol,* 145**,** 1484-94.

LI, S., ZHAO, B., YUAN, D., DUAN, M., QIAN, Q., TANG, L., WANG, B., LIU, X., ZHANG, J., WANG, J., SUN, J., LIU, Z., FENG, Y. Q., YUAN, L. & LI, C. 2013. Rice zinc finger protein DST enhances grain production through controlling Gn1a/OsCKX2 expression. *Proc Natl Acad Sci U S A,* 110**,** 3167-72.

PENA-CASTRO, J. M., VAN ZANTEN, M., LEE, S. C., PATEL, M. R., VOESENEK, L. A., FUKAO, T. & BAILEY-SERRES, J. 2011. Expression of rice SUB1A and SUB1C transcription factors in Arabidopsis uncovers flowering inhibition as a submergence tolerance mechanism. *Plant J,* 67**,** 434-46.

TRUJILLO, L. E., SOTOLONGO, M., MENENDEZ, C., OCHOGAVIA, M. E., COLL, Y., HERNANDEZ, I., BORRAS-HIDALGO, O., THOMMA, B. P., VERA, P. & HERNANDEZ, L. 2008. SodERF3, a novel sugarcane ethylene responsive factor (ERF), enhances salt and drought tolerance when overexpressed in tobacco plants. *Plant Cell Physiol,* 49**,** 512-25.

WERNER, T., MOTYKA, V., STRNAD, M. & SCHMULLING, T. 2001. Regulation of plant growth by cytokinin. *Proc Natl Acad Sci U S A,* 98**,** 10487-92.

YIN, X., LIU, X., XU, B., LU, P., DONG, T., YANG, D., YE, T., FENG, Y. Q. & WU, Y. 2019. OsMADS18, a membrane-bound MADS-box transcription factor, modulates plant architecture and the ABA response in rice. *J Exp Bot*.

ZHANG, Y., CHEN, C., JIN, X. F., XIONG, A. S., PENG, R. H., HONG, Y. H., YAO, Q. H. & CHEN, J. M. 2009. Expression of a rice DREB1 gene, OsDREB1D, enhances cold and high-salt tolerance in transgenic Arabidopsis. *BMB Rep,* 42**,** 486-92.

APPENDIX A: Supplementary data – Additional phenotypic assessments of *mpg1* compared to wild-type including collective analysis of several generations and additional preliminary analyses of seed yield characteristics

Introduction

mpg1 is a novel rice mutant that exhibits ectopic expression of an AP2/EREBP transcription factor by result of a T-DNA insertion mutation resulting in increased plant biomass, seed yield, and possible degrees of stress tolerance. The T-DNA insertion responsible for the generation of this mutant integrated in an intergenic region not disrupting any known or annotated functional elements. The T-DNA insertion contains several sequence elements including a duplicated *CaMV35S* promoter driving expression of the gene coding for hygromycin resistance, as well as a portion of a companion cell-specific promoter from the *Cucumis melo CmGAS1* gene. It is currently hypothesized that the insertion is somehow influencing ectopic expression of the AP2/EREBP transcription factor termed *MPG1*, a gene proximal to the T-DNA insertion site. To evaluate the extent of the effects that the ectopic expression culminates in, *mpg1* plants were comprehensively phenotype over the course of several selfed generations. These replicates were performed under both non-optimal and optimal conditions.

Presented here are the assessments of the cumulative measurements taken over several generations of experimentation, as well as preliminary observations of potentially revealing characteristics relevant to seed characteristics not yet fully assessed. To avoid statistical error associated with degrees' of potential variability accompanying these different trials of experimentation, representative populations were selected and reported (Chapter 3). However, to validate confidence and reproducibility of the representative experiments, data from multiple trials of growth experiments are reported.

Materials and Methods

Plant materials

Rice (*Oryza sativa* L. spp. *japonica* cv. Kitaake), including wild-type, and segregating lines of partial T-DNA insertion of expression cassette *CmGAS1pro*::*AtSUC1*$_{H65K}$ - mutant (*mpg1*), were used to assess phenotypic metrics, and expression analyses.

Plant growth conditions

Seeds were placed on germination paper and partially submerged in a 1:1000 dilution of MAXIM XL dual action fungicide (Syngenta) and sealed with parafilm. Seeds were incubated at 30º C under 12 h light cycles, until primary shoot and root development occurred (usually 5-7 days). Seedlings were then transferred to planting medium in the greenhouse. Planting medium (non-optimal) consisted of: 1 part play sand, 4 parts Canadian sphagnum peat moss, 4 parts (Promix) BX, mixed to homogeneity. Plants were either transferred to 3.5" pots until 3-leaf stage where they were genotyped then transplanted into 1.0 gallon pots, or directly into 1.0 gallon experimental pots. Pots were organized in random fashion in a flat or tub with water covered with black plastic and watered until media was fully saturated and pots remained in roughly 3" of standing water. Greenhouse conditions were maintained at 30˚C and 70% RH with a 16h light cycle. Plant chlorosis was monitored and preemptively treated around the 3- to 4-leaf stage using Sprint 330 Iron Chelate at 0.3 g/L water and top-watered. At the same developmental stage plants were fertilized using Scotts Peters Excel 15-5-15 Cal-Mag granular fertilizer at 24.22 g/L water and top-watered. Fertilizer treatment occurred twice weekly until harvest.

The optimal planting media consisted of: 1 part (Profile) Greens Grade porous ceramic particulate, and 1 part (Promix) BX, soil. The contents were mixed to homogeneity, and transferred to 1.0 gallon experimental pots. Pots were organized in random fashion in a flat or

tub with water covered with black plastic and watered until media was fully saturated and pots remained in roughly 3" of standing water. Greenhouse conditions were maintained at 30˚C and 70% RH with a 16h light cycle. Plant chlorosis was monitored and preemptively treated around the 3- to 4-leaf stage using (Sprint) 330 Iron Chelate at 0.3 g/L water and top-watered. At the same developmental stage plants were fertilized using granulized (Technigro) 15-5-15 Plus Cal-Mag at 48.87 g/L and top-watered. Fertilizer treatment occurred twice weekly through maturity until harvest.

Experimental growth replicates

Four independent rounds of experimentation were performed using both non-optimal and optimal conditions. Plant growth experiments were performed based on the availability of greenhouse space, thus plants were grown in differing seasons and were subject to different levels of variability concerning pest outbreaks and/or greenhouse regulatory stability. The experimental trials performed are labeled (1-4) for both non-optimal and optimal conditions. Several stunted variants (not linked to our study, present even in tWT plants) were removed from our analyses as outliers.

The layout for non-optimal trials are as follows. Trial 1 consisted of T_2 plants grown during winter of 2013, and comprised of 72 total plants - 22 tWT, 10 WT-ns, 16 HT-*mpg1*, and 24 HM-*mpg1*. Trial 2 consisted of T_3 and T_4 plants grown during summer of 2013, and comprised of 174 total plants – 0 tWT, 43 WT-ns, 61 HT-*mpg1*, and 70 HM-*mpg1*. Trial 3 consisted of T_4 and T_5 plants grown during spring of 2014, and comprised of 103 total plants – 12 tWT, 11 WT-ns, 24 HT-*mpg1*, and 56 HM-*mpg1*. Trial 4 consisted of T_3 plants grown during fall of 2015, and comprised of 76 total plants – 20 tWT, 20 WT-ns, 17 HT-*mpg1*, and 19 HM-*mpg1*.

The layout for optimal trials are as follows. Trial 1 consisted of T_3 and T_4 plants grown during fall of 2014, and comprised of 86 plants – 12 tWT, 24 WT-ns, 38 HT-*mpg1*, and 22 HM-*mpg1*. Trial 2 consisted of T_3 plants grown during summer of 2016, and comprised of 72 plants – 0 tWT, 11 WT-ns, 35 HT-*mpg1*, and 26 HM-*mpg1*. Trial 3 consisted of T_4 plants grown in summer of 2017, and comprised of 80 plants – 20 tWT, 19 WT-ns, 29 HT-*mpg1*, and 12 HM-*mpg1*. Trial 4 consisted of T_4 plants grown during fall of 2017, and comprised of 75 plants – 20 tWT, 12 WT-ns, 25 HT-*mpg1*, and 18 HM-*mpg1*.

Field trial

Both WT-ns and HM-*mpg1* seeds were germinated and plants were grown under optimal conditions in a local greenhouse (Fort Collins, CO) where they were subjected to a period of 'hardening' by exposure to fans, which were rotated periodically for several days. In June, when plants were around the 4-leaf stage they were transferred to a field plot at Colorado State University's Agricultural Development & Education Center (ARDEC) in Fort Collins, CO. The soil was pre-fertilized and plants were grown to maturity. Plants were irrigated weekly with supplemental hand-watering as necessary. It is important to note that plants were harvested prior to *mpg1*'s ability to fully progress through panicle development and seed filling due to inclement cold weather at the end of the growth experiment.

DNA extraction and genotyping

Young, fresh plant tissue (3-leaf stage) was sampled for DNA extraction and analysis (2-5 cm of leaf-tip). DNA was obtained via mechanical disruption of tissue and Shorty-Buffer extraction. Tissue was flash frozen in liquid nitrogen and disrupted using the Qiagen TissueLyser at 30 rps for a period of 1 minute. Five hundred μL of freshly prepared shorty buffer (0.2 M Tris HCl pH 9.0, 0.4 M LiCl, 25 mM EDTA, and 1.0% SDS) was added to each tissue sample,

vortexed and centrifuged at max speed (13k rpm) for 5 min. Then 350 μL of supernatant was transferred to a new tube containing 400 μL isopropanol, mixed by inverting and centrifuged at max speed for 10 min. The supernatant was discarded and 1mL of 70% ethanol was added to each sample to wash the DNA pellet. Samples were then centrifuged at max speed for 10 minutes, the supernatant was discarded and the tubes were inverted for 30 minutes. The DNA pellet was resuspended in 100 μL of 10 mM Tris-HCl pH 8 and stored short term at 4º C until use.

Haplotyping mutants

To identify *mpg1* plants that were homozygous, heterozygous, or null segregants for the bi-laterally truncated T-DNA insertion, primers were designed in regions directly flanking the site of insertion, as well as primers spanning the integration site into the T-DNA, and used in PCR. The primers flanking the T-DNA insertion were wFLA forward: 5'-GGAAGTTGGAGATGGGAAACA-3', and wFLA reverse: 5'-GGCCTCGTGTGTCAGTAATAA-3'. The primers spanning the genomic region and the T-DNA insertion were wIN forward: 5'-ACACCGGAAGCATAGTCATTT-3', and wIN reverse: 5'-GGTCGCCAACATCTTCTTCT-3'.

RNA-extraction and gene expression analysis

Desired tissue from both stem, leaf, and root across development from various selfed and backcrossed populations (not more than 100 mg) was sampled for RNA extraction and analysis. Tissue was placed in individual 2 mL tubes and flash frozen in liquid nitrogen. Tissue was ground using the TissueLyser (Qiagen) at 30 rps for 1 min, and RNA was extracted using (Qiagen) Plant RNeasy mini-kit. RNA was treated with DNase and purified using the Turbo

DNase kit (Invitrogen). cDNA was synthesized from 1 μg of RNA using SuperScript (Invitrogen).

Specific primers used for expression analysis via RT-PCR for the individual candidate gene of LOC_Os08g41030 (*MPG1*) are 41030 forward: 5'-TCGCCATTGTTCAGCAAGAAGGA-3', and 41030 reverse: 5'-AAGTGCATGACCAAGTACAGA-3'. Housekeeping control primers were designed around actin, more specifically the sequences are, actin forward: 5'-GAGTATGATGAGTCGGGTCCA-3', and actin reverse primer: 5'-ACACCAACAATCCCAAACAGA-3'. PCR was performed in 20 μL reactions using Econo Taq polymerase (Lucigen), 2 μL of cDNA and desired primers under a normal 30-cycle amplification protocol. PCR products were analyzed by electrophoresis on 1% gel-agarose containing ethidium bromide.

Phenotypic analyses

Plants were grown to maturity and several metrics were recorded. Plant height was measured as the length from the planting media to the tip of the tallest leaf. Tiller number was recorded by counting number of stems with true leaves present. Girth was measured as length of diameter or circumference at roughly 5cm above soil line. Time to flowering was measured by the number of days post-planting until panicle emergence (or heading). At harvest height, tiller number, and girth were assessed again. In addition, leaf length, and leaf width were measured. These characteristics were taken from the leaf that provided the highest point of the plant and measured its length from tip to culm, and the width at its widest point. Additionally, panicles inferior to the panicle neck were harvested, dried for 7 days in a 45º C drying oven and weighed to determine seed yield. Plants with panicles removed were cut roughly 5cm above soil line, dried at 45º C for 7-14 days and weighed for a measure of total biomass. Total seed yield with

panicle tissue was recorded, as well as number of panicles. Panicles were also monitored for the presence and length of awn development. Preliminary analysis of panicle architecture were assessed by panicle length, branch number, and total spikelet count per panicle. After harvest residual plant matter remained in growing conditions and ratooning was assessed. Measurements were taken at 41dpp.

Additional preliminary seed yield analyses

Plants were grown to maturity and harvested per standard growth conditions given above. A 1000 grain seed count was performed by removing the hull and weighing 1000 grains from individual treatments. Panicles were assessed by measuring panicle branch number, panicle length, and seeds per panicle through random selection of a single panicle per plant. Additionally, qualitative observations of panicle density were noted on the three eldest panicles per treatment via photograph.

Due to inclement weather seed yield characteristics were not successfully evaluated in our field trial. To establish a better understanding of seed yield potential in the field spikelet number per panicle were counted by totaling the number of spikelets from the three oldest panicles per treatment.

***Xanthomonas* pathogen assay**

Xanthomonas oryzae pv. *oryzae* strain PXO86 was grown on PSA plates at 28 C for 2 days. A new PSA plate was then inoculated from the first to make a lawn. From the lawn plate a loop-full of PXO86 was transferred in to ~6.0 mL of DI H_2O and was mixed via inversion. The resulting suspension concentration was determined and manipulated using a spectrophotometer till the OD at 600nm = 0.2. Scissors were dipped in the suspension and used to clip inoculate 2 leaves on a single plant at 21 days post planting. Eight of each type of plant (tWT, WT-ns, HT-

mpg1, and HM-*mpg1*) were inoculated. Half of the plants were used to gauge gene expression of the candidate gene LOC_Os08g41030 (*MPG1*) by RT-PCR while the other half of the plants were used to assess the PXO86 stress challenge. Samples for assessment of gene expression were taken at time-point 0 (pre- inoculation), 6 hours post-inoculation, and 12 hours post-inoculation. Samples were taken from one of the two inoculated leaves on each plant. At the 6 hour post-inoculation time-point, wild-type plants showed increased expression of candidate gene LOC_Os08g41030 (*MPG1*). The plants that weren't sampled for gene expression were left for a period of 12 days. The degree of the PXO86 infection was determined by the length of the lesion formed from the cut inoculation site.

Statistical analyses

Statistics were calculated for growth metrics by calculating observed means for each treatment across all trials as well as cumulative analysis by calculating least squares means and significance using a type III one-way ANOVA with Kenward-Roger's method at 95% family-wise confidence level in R. Graphical models were made using Microsoft Excel and Powerpoint. Analysis of field trial seed data was performed using a Student's t-test. Graphical models were made using boxplot function in R. Analysis of additional seed characteristics were performed using a one-way ANOVA and Games-Howell post-hoc multiple comparison test at 95% family-wise confidence level. Graphical models were made using Microsoft Excel and Powerpoint.

Results

Under non-optimal conditions *mpg1* generates greater biomass and seed yield than wild-type plants

To more comprehensively assess the phenotypic characteristics of *mpg1*, several selfed segregating generations of *mpg1* were grown in the greenhouse alongside wild-type control

plants. Measurements evaluating terminal growth and seed yield characteristics of *mpg1* under non-optimal conditions can be seen in (Table A.1).

Table A.1: Analysis of biomass and seed yield characteristics from several selfed segregating populations of *mpg1* grown under non-optimal conditions. Measurements of treatments averaged within and across all trials. Observed means were calculated within trials with error represented by standard deviation. Cumulative analysis was taken by least squares means across all treatments with error represented by standard error. Significance between comparisons was conducted using a one-way ANOVA with a Kenward-Rodger's method at a 95% family-wise confidence level.

TRAIT	GENOTYPE	TRIAL 1 (AVG + SD)	TRIAL 2 (AVG + SD)	TRIAL 3 (AVG + SD)	TRIAL 4 (AVG + SD)	LS MEANS + SE	COMPARISONS
Dry Weight (g)							HM-*mpg1*:HT-*mpg1* <0.0001
	tWT	13.1 ± 3.74	-	10.42 ± 4.4	6.59 ± 2.04	14.28 ± 5.28	HM-*mpg1*:tWT <0.0001
	WT-ns	6.54 ± 2.81	14.97 ± 5.35	9.47 ± 4.03	5.23 ± 1.23	6.02 ± 5.13	HM-*mpg1*:WT-ns <0.0001
	HT-*mpg1*	28.28 ± 8.32	49.09 ± 21.0	18.57 ± 13.1	29.37 ± 6.83	32.41 ± 5.05	HT-*mpg1*:tWT <0.0001
	HM-*mpg1*	39.95 ± 16.55	59.5 ± 23.74	36.56 ± 13.26	39.1 ± 12.92	44.3 ± 4.99	HT-*mpg1*:WT-ns <0.0001
							tWT:WT-ns 0.0193
Height (cm)							HM-*mpg1*:HT-*mpg1* 0.0007
	tWT	106.22 ± 5.84	-	94.8 ± 8.27	88.54 ± 5.68	96.5 ± 4.61	HM-*mpg1*:tWT <0.0001
	WT-ns	100.6 ± 5.71	101.2 ± 10.61	94.34 ± 6.85	87.38 ± 8.07	96.7 ± 4.53	HM-*mpg1*:WT-ns <0.0001
	HT-*mpg1*	127.2 ± 9.43	116.48 ± 12.79	98.69 ± 15.35	110.81 ± 7.12	113.1 ± 4.49	HT-*mpg1*:tWT <0.0001
	HM-*mpg1*	131.9 ± 13.16	118.54 ± 6.64	109.53 ± 8.1	115.99 ± 5.32	117.9 ± 4.46	HT-*mpg1*:WT-ns <0.0001
							tWT:WT-ns 0.9998
Tiller number							HM-*mpg1*:HT-*mpg1* 0.0090
	tWT	11.57 ± 7.31	-	24.91 ± 5.93	17.6 ± 4.32	18.6 ± 3.68	HM-*mpg1*:tWT <0.0001
	WT-ns	5.5 ± 3.0	16.02 ± 4.68	25.81 ± 8.78	16.6 ± 3.24	15.3 ± 3.61	HM-*mpg1*:WT-ns <0.0001
	HT-*mpg1*	15.0 ± 5.47	29.83 ± 11.98	26.66 ± 7.77	29.94 ± 5.76	26.2 ± 3.57	HT-*mpg1*:tWT <0.0001
	HM-*mpg1*	17.39 ± 4.87	29.8 ± 10.65	37.58 ± 8.34	33.89 ± 8.66	29.5 ± 3.54	HT-*mpg1*:WT-ns <0.0001
							tWT:WT-ns 0.1791

Girth diameter (cm)						HM-*mpg1*:HT-*mpg1* 0.6362
	tWT	1.94 ± 0.58	-	-	-	2.81 ± 0.91 HM-*mpg1*:tWT 0.0009
	WT-ns	0.97 ± 0.31	2.78 ± 0.93	-	-	1.9 ± 0.88 HM-*mpg1*:WT-ns <0.0001
	HT-*mpg1*	2.88 ± 0.70	4.69 ± 1.52	-	-	3.81 ± 0.88 HT-*mpg1*:tWT 0.0127
	HM-*mpg1*	3.18 ± 0.60	4.88 ± 1.24	-	-	4.02 ± 0.88 HT-*mpg1*:WT-ns <0.0001 tWT:WT-ns 0.0412
Girth circumference (cm)						HM-*mpg1*:HT-*mpg1* <0.0001
	tWT	-	-	10.21 ± 3.18	8.75 ± 1.17	8.51 ± 0.68 HM-*mpg1*:tWT <0.0001
	WT-ns	-	-	9.38 ± 3.74	8.33 ± 0.93	8.99 ± 0.69 HM-*mpg1*:WT-ns <0.0001
	HT-*mpg1*	-	-	11.69 ± 3.57	14.65 ± 1.69	13.09 ± 0.63 HT-*mpg1*:tWT <0.0001
	HM-*mpg1*	-	-	15.25 ± 2.83	16.92 ± 1.82	15.88 ± 0.57 HT-*mpg1*:WT-ns <0.0001 tWT:WT-ns 0.9027
Leaf length (cm)						HM-*mpg1*:HT-*mpg1* 0.2350
	tWT	41.45 ± 8.32	-	34.43 ± 6.15	-	38.1 ± 2.84 HM-*mpg1*:tWT 0.0003
	WT-ns	37.34 ± 6.83	38.08 ± 5.14	38.38 ± 6.25	-	38.4 ± 2.29 HM-*mpg1*:WT-ns <0.0001
	HT-*mpg1*	50.0 ± 5.42	45.5± 13.77	42.65 ± 11.92	-	46.1 ± 2.08 HT-*mpg1*:tWT 0.0230
	HM-*mpg1*	54.47 ± 7.02	47.79 ± 17.61	46.63 ± 11.0	-	49.1 ± 1.92 HT-*mpg1*:WT-ns 0.0006 tWT:WT-ns 0.9996
Leaf width (cm)						HM-*mpg1*:HT-*mpg1* 0.0411
	tWT	1.06 ± 0.3	-	1.25 ± 0.17	-	1.17 ± 0.08 HM-*mpg1*:tWT <0.0001
	WT-ns	1.01 ± 0.31	1.33 ± 0.28	1.16 ± 0.23	-	1.19 ± 0.07 HM-*mpg1*:WT-ns <0.0001
	HT-*mpg1*	1.56 ± 0.15	1.66 ± 0.21	1.5 ± 0.27	-	1.57 ± 0.06 HT-*mpg1*:tWT <0.0001
	HM-*mpg1*	1.65 ± 0.21	1.79 ± 0.41	1.57 ± 0.19	-	1.67 ± 0.06 HT-*mpg1*:WT-ns <0.0001

							tWT:WT-ns 0.9931
Days to heading (dpp)							HM-*mpg1*:HT-*mpg1* <0.0001
	tWT	-	-	56.0 ± 0	48.15 ± 1.55	52.2 ± 4.46	HM-*mpg1*:tWT <0.0001
	WT-ns	-	-	56.0 ± 0	49.6 ± 2.08	53.2 ± 4.65	HM-*mpg1*:WT-ns <0.0001
	HT-*mpg1*	-	-	75.68 ± 6.03	66.0 ± 1.94	70.9 ± 4.64	HT-*mpg1*:tWT <0.0001
	HM-*mpg1*	-	-	82.1 ± 3.97	71.05 ± 3.85	77.0 ± 4.62	HT-*mpg1*:WT-ns <0.0001
							tWT:WT-ns 0.7124
Seed yield (g)							HM-*mpg1*:HT-*mpg1* 0.0094
	tWT	20.21 ± 4.49	-	13.45 ± 11.55	18.55 ± 4.98	22.6 ± 7.8	HM-*mpg1*:tWT <0.0001
	WT-ns	11.46 ± 6.30	26.36 ± 11.0	16.41 ± 7.21	14.87 ± 3.54	13.2 ± 7.65	HM-*mpg1*:WT-ns <0.0001
	HT-*mpg1*	29.32 ± 10.73	58.33 ± 24.14	22.63 ± 13.34	53.22 ± 10.97	40.7 ± 7.57	HT-*mpg1*:tWT <0.0001
	HM-*mpg1*	39.86 ± 16.01	69.82 ± 29.05	27.04 ± 14.71	56.01 ± 19.97	48.1 ± 7.5	HT-*mpg1*:WT-ns <0.0001
							tWT:WT-ns 0.0382
Panicle number							HM-*mpg1*:HT-*mpg1* 0.1201
	tWT	-	-	11.16 ± 6.02	17.36 ± 4.41	13.9 ± 4.91	HM-*mpg1*:tWT <0.0001
	WT-ns	-	-	13.7 ± 3.16	16.42 ± 4.54	14 ± 4.92	HM-*mpg1*:WT-ns <0.0001
	HT-*mpg1*	-	-	13.22 ± 6.09	26.47 ± 6.32	19.8 ±4.89	HT-*mpg1*:tWT 0.0014
	HM-*mpg1*	-	-	17.27 ± 6.27	29.57 ± 7.82	22.7 ± 4.84	HT-*mpg1*:WT-ns 0.0023
							tWT:WT-ns 0.9999

Although degrees of variance are notable between different trials, which can be easily expected with experimentation occurring during different times of the year, phenotypic results still follow the same general trend (*mpg1* showing greater biomass and seed yield characteristics, and delay in flowering compared to wild-type). Under all metrics *mpg1* plants, both heterozygous and homozygous for the T-DNA insertion still generate significantly greater biomass and seed yield compared to wild-type plants (taller plants, greater tiller number, greater girth, larger leaves, greater seed yield, and panicle number). In some measurements, HM-*mpg1* is significantly greater than HT-*mpg1*, suggesting that a dosage effect exists affecting aspects of the phenotype.

Under optimal conditions *mpg1* generates greater biomass, seed yield, and ratooning success compared to wild-type plants

The growth media and fertilizer treatment were found to be non-optimal during initial experimentation (Chapter 3). Optimal growth media and fertilizer treatment were found and used to again evaluate growth and seed yield of *mpg1* compared to wild-type. Besides measurements taken under non-optimal conditions, additional investigation pertaining to ratooning capability was performed. Measurements evaluating the terminal growth and seed yield, as well as ratooning characteristics of *mpg1* under optimal conditions can be seen in (Table A.2).

Table A.2: Analysis of biomass, seed yield, and ratooning characteristics from several selfed segregating populations of *mpg1* grown under optimal conditions. Measurements of treatments averaged within and across all trials. Observed means were calculated within trials with error represented by standard deviation. Cumulative analysis was taken by least squares means across all treatments with error represented by standard error. Significance between comparisons was conducted using a one-way ANOVA with a Kenward-Rodger's method at a 95% family-wise confidence level.

TRAIT	GENOTYPE	TRIAL 1 (AVG + SD)	TRIAL 2 (AVG + SD)	TRIAL 3 (AVG + SD)	TRIAL 4 (AVG + SD)	LS MEANS + SE	COMPARISONS
Dry Weight (g)							HM-*mpg1*:HT-*mpg1* 0.0152
	tWT	13.01 ± 3.02	-	17.93 ± 4.83	12.81 ± 3.48	17.1 ± 4.11	HM-*mpg1*:tWT <0.0001
	WT-ns	9.33 ± 6.85	16.81 ± 4.87	14.63 ± 2.88	9.91 ± 2.97	13.6 ± 4.02	HM-*mpg1*:WT-ns <0.0001
	HT-*mpg1*	27.62 ± 12.0	50.19 ± 15.28	37.57 ± 11.13	34.04 ± 13.80	37.4 ± 3.9	HT-*mpg1*:tWT <0.0001
	HM-*mpg1*	29.87 ± 13.26	50.73 ± 16.38	50.04 ± 11.34	42.67 ± 14.31	42.5 ± 3.99	HT-*mpg1*:WT-ns <0.0001
							tWT:WT-ns 0.3782
Height (cm)							HM-*mpg1*:HT-*mpg1* 0.0335
	tWT	105.07 ± 6.67	-	111.1 ± 6.79	109.23 ± 6.25	111 ± 3.37	HM-*mpg1*:tWT <0.0001
	WT-ns	93.44 ± 11.52	123.22 ± 3.24	106.67 ± 5.45	99.25 ± 5.09	105 ± 3.32	HM-*mpg1*:WT-ns <0.0001
	HT-*mpg1*	119.26 ± 7.24	130.46 ± 4.89	120.18 ± 6.03	124.74 ± 11.50	124 ± 3.24	HT-*mpg1*:tWT <0.0001
	HM-*mpg1*	121.66 ± 9.11	134.04 ± 5.42	122.81 ± 4.28	129.08 ± 11.69	127 ± 3.3	HT-*mpg1*:WT-ns <0.0001
							tWT:WT-ns 0.0003
Tiller number							HM-*mpg1*:HT-*mpg1* 0.9341
	tWT	21.08 ± 9.43	-	24.4 ± 5.72	18.55 ± 4.93	24.9 ± 3.84	HM-*mpg1*:tWT 0.0397
	WT-ns	18 ± 5.67	19.63 ± 3.44	22.63 ± 4.52	19.16 ± 7.19	20.7 ± 3.76	HM-*mpg1*:WT-ns <0.0001
	HT-*mpg1*	30.02 ± 10.4	44.69 ± 18.71	26.13 ± 5.47	23.32 ± 7.44	31.3 ± 3.64	HT-*mpg1*:tWT 0.0038
	HM-*mpg1*	25.9 ± 7.88	42.04 ± 21.02	27.25 ± 5.65	25.83 ± 6.65	30.4 ± 3.72	HT-*mpg1*:WT-ns <0.0001
							tWT:WT-ns 0.1904

Girth circumference (cm)							HM-*mpg1*:HT-*mpg1* 0.0050
	tWT	9.3 ± 1.87	-	9.95 ± 1.51	7.76 ± 1.21	9.8 ± 1.19	HM-*mpg1*:tWT <0.0001
	WT-ns	7.37 ± 1.6	9.74 ± 1.47	9.02 ± 1.16	7.52 ± 1.18	8.51 ± 1.18	HM-*mpg1*:WT-ns <0.0001
	HT-*mpg1*	12.96 ± 2.52	17.88 ± 3.28	14.61 ± 3.2	11.86 ± 2.32	14.38 ± 1.16	HT-*mpg1*:tWT <0.0001
	HM-*mpg1*	12.65 ± 2.53	18.47 ± 3.74	19.22 ± 3.34	13.65 ± 2.17	15.69 ± 1.17	HT-*mpg1*:WT-ns <0.0001
							tWT:WT-ns 0.0551
Days to heading (dpp)							HM-*mpg1*:HT-*mpg1* 0.0014
	tWT	53.5 ± 3.79	-	57.35 ± 3.82	61.95 ± 6.78	56.9 ± 5.76	HM-*mpg1*:tWT <0.0001
	WT-ns	51.83 ± 2.89	-	57.57 ± 2.66	61.08 ± 0.95	58.0 ± 5.76	HM-*mpg1*:WT-ns <0.0001
	HT-*mpg1*	62.55 ± 3.26	-	70.17 ± 3.56	87.64 ± 5.04	73.1 ± 5.73	HT-*mpg1*:tWT <0.0001
	HM-*mpg1*	65.5 ± 4.37	-	70.75 ± 6.69	93.77 ± 2.17	76.9 ± 5.76	HT-*mpg1*:WT-ns <0.0001
							tWT:WT-ns 0.7990
Seed yield (g)							HM-*mpg1*:HT-*mpg1* <0.0001
	tWT	20.62 ± 4.45	-	33.57 ± 9.19	21.08 ± 5.63	24.8 ± 4.06	HM-*mpg1*:tWT <0.0001
	WT-ns	31.21 ± 11.83	31.21 ± 8.88	30.12 ± 7.24	15.77 ± 3.37	22.1 ± 3.95	HM-*mpg1*:WT-ns <0.0001
	HT-*mpg1*	37.15 ± 17.14	34.95 ± 13.95	48.39 ± 11.97	29.61 ± 15.91	37.7 ± 3.77	HT-*mpg1*:tWT <0.0001
	HM-*mpg1*	38.66 ± 16.96	34.33 ± 11.19	56.05 ± 12.36	31.85 ± 16.46	39.5 ± 3.89	HT-*mpg1*:WT-ns <0.0001
							tWT:WT-ns 0.6989
Panicle number							HM-*mpg1*:HT-*mpg1* 0.9994
	tWT	-	-	21.95 ± 5.70	15.1 ± 3.26	18.8 ± 2.3	HM-*mpg1*:tWT 0.9651
	WT-ns	-	19.45 ± 4.29	20.05 ± 4.34	12.75 ± 1.83	17.3 ± 2.28	HM-*mpg1*:WT-ns 0.2133
	HT-*mpg1*	-	20.21 ± 6.58	22.58 ± 4.74	14.64 ± 4.91	19.2 ± 2.21	HT-*mpg1*:tWT 0.9752
	HM-*mpg1*	-	18.37 ± 4.18	24.5 ± 5.37	16.11 ± 4.85	19.3 ± 2.25	HT-*mpg1*:WT-ns 0.1763

							tWT:WT-ns 0.5377
41 days post-ratooning height (cm)							HM-*mpg1*:HT-*mpg1* 0.4248
	tWT	-	-	68.76 ± 24.56	45.38 ± 25.75	56 ± 7.34	HM-*mpg1*:tWT <0.0001
	WT-ns	-	64.38 ± 7.9	64.48 ± 20.46	67.85 ± 21.46	65 ± 7.46	HM-*mpg1*:WT-ns 0.0004
	HT-*mpg1*	-	76.55 ± 12.01	91.43 ± 15.95	65.11 ± 13.17	77.9 ± 6.81	HT-*mpg1*:tWT <0.0001
	HM-*mpg1*	-	78.69 ± 14.36	100.4 ± 6.22	70.25 ± 7.43	82.3 ± 6.96	HT-*mpg1*:WT-ns 0.0075
							tWT:WT-ns 0.2486
41 days post-ratooning tiller number							HM-*mpg1*:HT-*mpg1* 0.0549
	tWT	-	-	10.76 ± 6.57	11.66 ± 7.18	11.8 ± 2.74	HM-*mpg1*:tWT <0.0001
	WT-ns	-	8.0 ± 4.07	9.25 ± 4.17	15.14 ± 7.11	10.7 ± 2.89	HM-*mpg1*:WT-ns <0.0001
	HT-*mpg1*	-	32.03 ± 11.66	28.1 ± 13.48	23.12 ± 10.64	28 ± 1.91	HT-*mpg1*:tWT <0.0001
	HM-*mpg1*	-	31.45 ± 13.81	46.25 ± 13.49	26.83 ± 8.94	33.3 ± 2.16	HT-*mpg1*:WT-ns <0.0001
							tWT:WT-ns 0.9876
41 days post-ratooning girth (cm)							HM-*mpg1*:HT-*mpg1* 0.0018
	tWT	-	-	5.58 ± 2.74	4.52 ± 2.85	5.37 ± 1.22	HM-*mpg1*:tWT <0.0001
	WT-ns	-	4.57 ± 1.65	5.28 ± 1.42	6.22 ± 2.51	5.3 ± 1.18	HM-*mpg1*:WT-ns <0.0001
	HT-*mpg1*	-	13.02 ± 3.03	12.56 ± 4.55	8.77 ± 2.70	11.54 ± 1.03	HT-*mpg1*:tWT <0.0001
	HM-*mpg1*	-	13.39 ± 4.18	18.3 ± 2.52	11.11 ± 2.77	13.81 ± 1.07	HT-*mpg1*:WT-ns <0.0001
							tWT:WT-ns 0.9999
41 days post-ratooning. panicle number							HM-*mpg1*:HT-*mpg1* 0.7778
	tWT	-	-	0.46 ± 0.84	0.66 ± 0.94	1.25 ± 1.57	HM-*mpg1*:tWT <0.0001
	WT-ns	-	0.28 ± 0.69	0.25 ± 0.66	0.42 ± 0.72	0.44 ± 1.59	HM-*mpg1*:WT-ns <0.0001
	HT-*mpg1*	-	9.21 ± 3.3	1.93 ± 2.83	8.44 ± 5.1	6.53 ± 1.46	HT-*mpg1*:tWT <0.0001

HM-*mpg1*	-	6.87 ± 3.0	3.33 ± 1.24	8.0 ± 4.1	5.97 ± 1.49	HT-*mpg1*:WT-ns <0.0001 tWT:WT-ns 0.8456

Growth metrics of *mpg1* were again greater across all measurements compared to wild-type with similar phenotypic patterning. However differences between *mpg1* and wild-type tiller number were not as substantial under optimal conditions. A number of measurements were greater under optimal conditions compared to non-optimal conditions across all treatments, again suggesting that non-optimal conditions were generating a degree of stress. Similarly, some measurements showed a difference between HT-*mpg1* and HM-*mpg1* suggesting that a dosage effect might exist within several of the measured characteristics.

Seed yield characteristics also followed the observed phenotypic pattern as seen under non-optimal conditions, with greater seed yield measured in *mpg1* compared to wild-type. Although, under optimal-conditions, there isn't a significant difference between the panicle number of *mpg1* and wild-type.

Evaluation of harvested plants revealed that re-growth of defoliated plant matter was much more apparent in *mpg1* compared to wild-type (Chapter 3). Further evaluation of specific facets of plant regrowth including height, tiller number, and girth were measured at 41 days post-ratooning. Additionally, evaluation of seed yield characteristics were taken by means of panicle number at this same time.

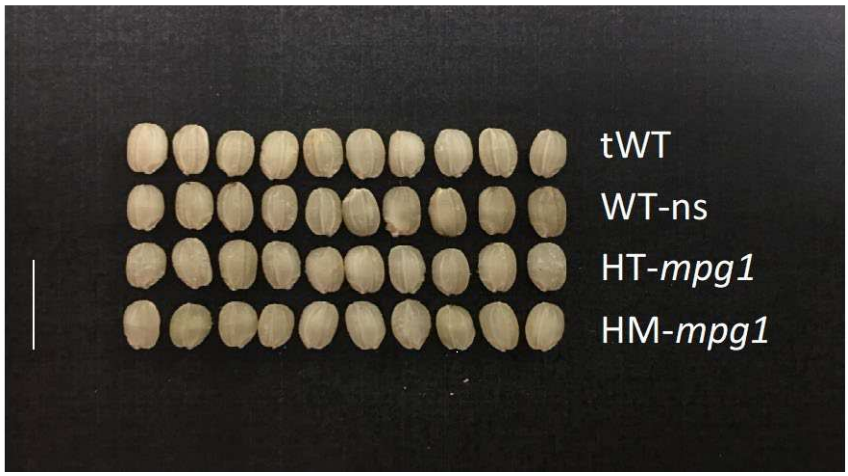
Overall, *mpg1* plants were much more successful at ratooning compared to wild-type (survival rates: trial 2 – WT-ns 63%, HT-*mpg1* 100%, HM-*mpg1* 100%; trial 3- tWT 65%, WT-ns 42%, HT-*mpg1* 100%, HM-*mpg1* 100%; trial 4- tWT 75%, WT-ns 58%, HT-*mpg1* 100%, HM-*mpg1* 100%). Of the surviving plants, all measurements considering plant growth and seed yield post-ratooning were greater in *mpg1* compared to wild-type. The phenotypic patterning also remained similar, with greater biomass and seed yield characteristics, in a post-ratooning state, however, *mpg1* plants were much more successful at regrowth and panicle generation

compared to wild-type, with flowering occurring earlier and producing a greater number of panicles.

Preliminary investigation of seed traits indicates that *mpg1* possesses different panicle morphology than wild-type

Preliminary evaluation of seed yield characteristics were taken to better understand potential mechanisms related to the generation of greater seed yield. Although some growth trials had greater panicle numbers in *mpg1* compared to wild-type, some did not. However, overall seed yield still remained significantly greater in *mpg1* within experiments where panicle numbers were not greater, suggesting that there might be a difference in seed size, or seed number (alteration of panicle architecture by means of panicle length or panicle branch number). To evaluate any differences in grain size a 1000-grain seed count was performed between treatments (Figure A.1).

A



B

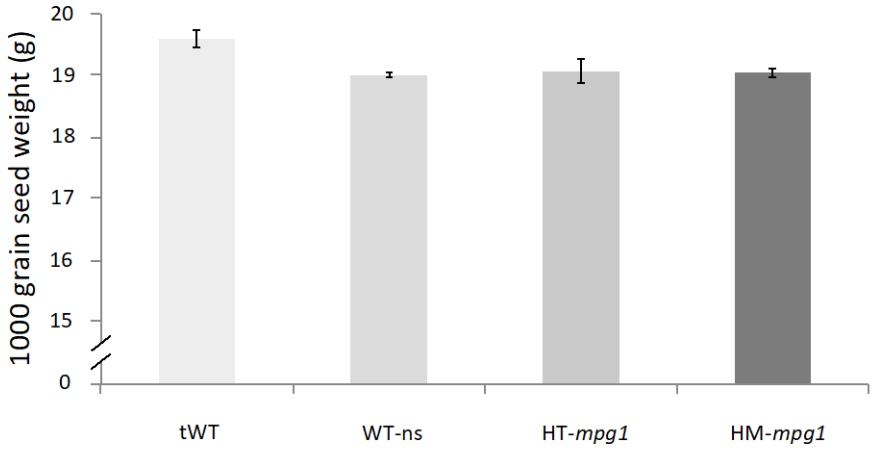


Figure A.1: Preliminary analysis of seed characteristics within a segregating T_2 population of *mpg1* at maturity grown under non-optimal conditions (trial 1).
Measurements of individual seeds were taken to see if there was any size difference in *mpg1*. (A) Photograph of individual de-hulled seeds, and (B) analysis of 1000 grain count. 1000 grain count was taken as a measurement of 1000 total seeds (with roughly one-third from an individual plant). Analysis was conducted using a Student's t-test comparing tWT to other treatments. Error bars represent standard deviation.

Preliminary analysis of the size of *mpg1* seeds compared to wild-type did not statistically differ under a 1000-grain count. This suggests that the generation of greater seed yield observed in our experiments is the result of more seeds per plant. More seeds per plant can result from either more panicles, more seeds per panicle, or both. To better evaluate the determining factor

driving increased seed yield preliminary data on panicle branch number, panicle length, and seeds per panicle were assessed (Figure A.2).

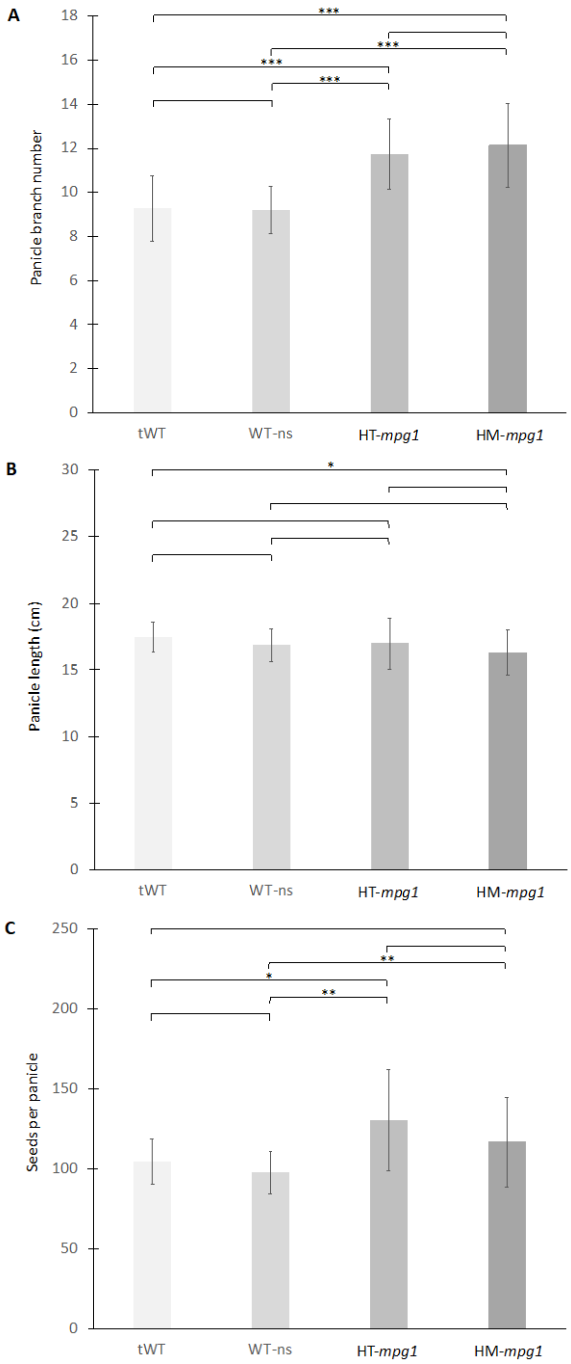


Figure A.2: Preliminary analysis of panicle characteristics within a segregating T_4 and T_5 population of* mpg1 *at maturity grown under non-optimal conditions (trial 3).
(A) Panicle branch number, (B) panicle length, and (C) seeds per panicle. Analysis was conducted using a one-way ANOVA and Games-Howell post-hoc multiple comparison test at

95% family-wise confidence level. '*' indicates $p<0.05$, '**' indicates $p<0.01$, '***' indicates $p<0.001$. Error bars represent standard deviation.

Panicle measurements differed between *mpg1* and wild-type. In terms of general architecture, *mpg1* panicles were not significantly longer than wild-type, however they did have a greater number of branches, and a greater total number of seeds per panicle. Seed size and panicle characteristics were taken under non-optimal conditions and should be repeated under optimal conditions with a greater number of replicates and detail. In (Figure A.3), a qualitative measurement by photograph of mature panicles harvested under optimal conditions.

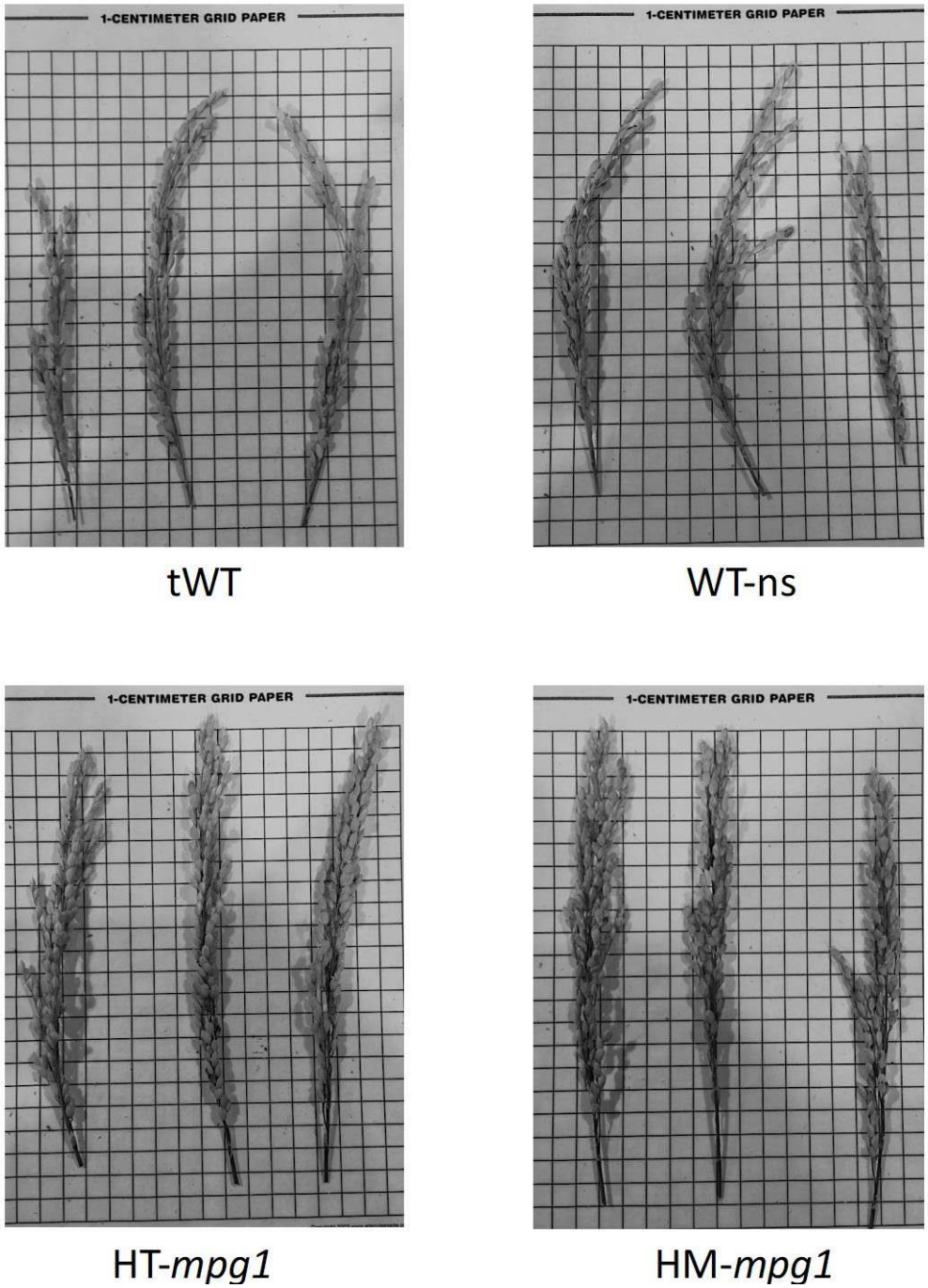


Figure A.3: Qualitative analysis of panicle characteristics within a segregating population of T_4 *mpg1* at maturity grown under optimal conditions (trial 4).
Photographic qualitative assessment spikelet density between *mpg1* plants compared to wild-type.

While qualitative assessment of panicles does not provide a significant understanding of panicle architecture, panicles under optimal growth conditions do appear more full, possibly containing more seeds per panicle.

The *mpg1* field trial, was unable to successfully generate data regarding seed yield due to pre-mature harvesting as a result of inclement weather. However, seed yield potentials were preliminarily assessed by evaluating the number of seeds per panicle on the three oldest panicles on both *mpg1* and wild-type plants (Figure A.4).

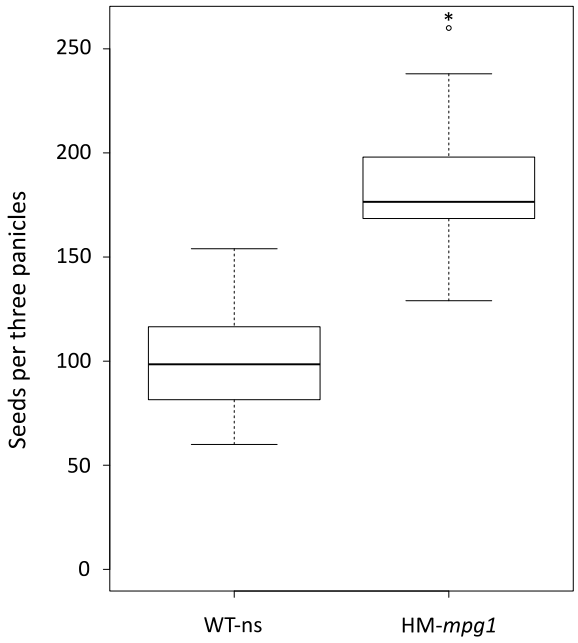


Figure A.4: Analysis of seed-yield metrics within a segregating population of T_4 *mpg1* grown under field conditions.
Seed yield and panicle number were not reported in the field trial due to early harvesting of plants as a result of inclement weather not allowing HM-*mpg1* to complete total panicle growth and seed filling. Therefore, total seed number of the three most mature panicles was recorded in effort to try and gather preliminary data on seed-yield potential in the field. WT-ns (n = 60), HM-*mpg1* (n = 60). Analysis was conducted using a Student's t-test. '*' indicates $p<0.05$.

Although this strategy of evaluation is insufficient in determining terminal seed yield, it did prove insightful as the number of seeds per panicle were significantly greater in *mpg1* compared to wild-type. If the number of panicles generated remained consistent or even higher

between *mpg1* and wild-type it would be reasonable to predict that seed yield of *mpg1* might remain significantly greater under field conditions.

Investigation regarding the native functionality of AP2/ERF transcription factors suggested that they might operate in response to biotic stress, discussed in (Chapter 3, 4). Because of this, we chose to evaluate *mpg1* and wild-type exposed to *Xanthomonas oryzae* pv. *oryzae* as a preliminary assessment of its response to biotic stress (Figure A.5).

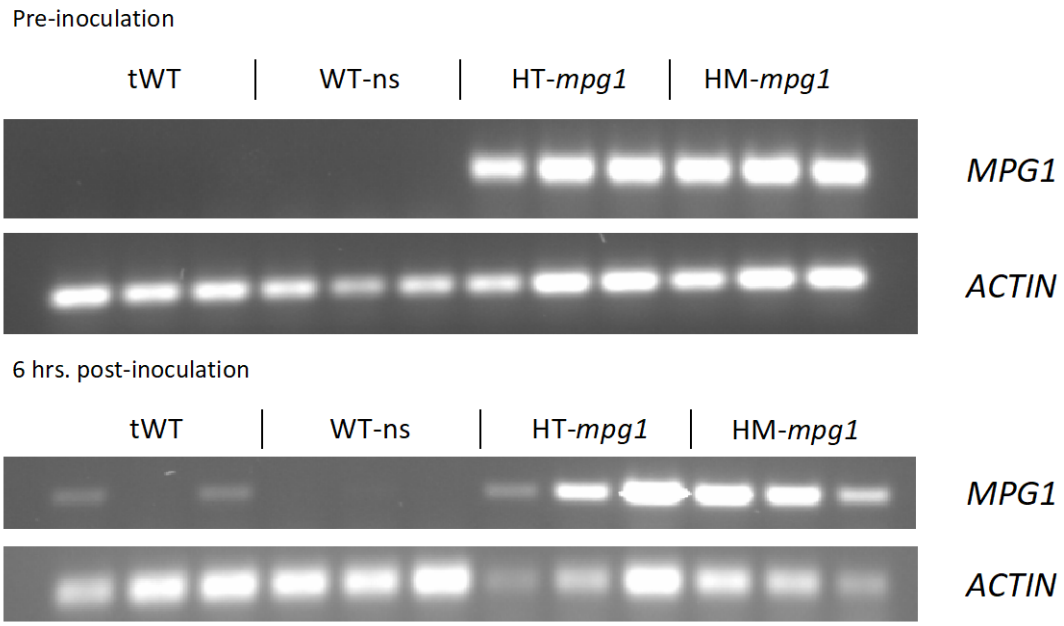


Figure A.5: Expression analysis after *Xanthomonas oryzae* pv. *oryzae* challenge.
RT-PCR gel-electrophoresis image from leaf tissue at 21 days post-planting. There is greater expression of (*MPG1*) in plants that are heterozygous and homozygous for the truncated T-DNA insertion present in *mpg1* compared to wild-type plants which show no expression under pre-inoculated plants. After 6 hours post-inoculation, report of *MPG1* can be seen in tWT and WT-ns plants. Each lane represents a single biological replicate, and the gene coding for actin was used as a control.

Interestingly, there was no difference between legion length between *mpg1* and wild-type plants, suggesting that the ectopic expression of *MPG1* does not influence defense to *Xoo*. However, after inoculation with *Xoo* wild-type expression of the native *MPG1* was found after 6 hours. This indicates that the native role of *MPG1* might function in response to biotic stress.

Discussion

Evaluation of the replicate experiments assessing *mpg1* growth and seed yield accumulation show strong similarity and consistency with phenotypic outcomes between tWT, WT-ns, HT-*mpg1*, and HM-*mpg1* contingent on their respective growth conditions. The variability between experiments can be rationalized by the effects of seasonal influence and greenhouse conditions.

Under non-optimal and optimal conditions, *mpg1* is able to accumulate greater biomass and seed yield. The delay in flowering observed in *mpg1* relative to wild-type also remains consistent. One noteworthy difference between non-optimal conditions and optimal conditions, is resulting tiller number. Perhaps the particular conditions present during non-optimal growth influences branching or tiller formation. The overall differences between growth and seed yield between non-optimal and optimal conditions suggests that a degree of stress was imparted during growth under non-optimal conditions as most of the measurements were greater across all treatments under optimal conditions. Furthermore, the degree of phenotypic difference between *mpg1* and wild-type is greater under non-optimal conditions than under optimal conditions signifying that *mpg1*'s phenotype is linked to stress.

Replicates of stress trials are not provided as the experimental designs differed (pot layout, concentrations of stress, and duration of stress) and could not be accurately assessed in a cumulative manner. The effects of prolonged stress in these experiments did however result in similar phenotypic outcomes as to what was observed in the example provided (Chapter 3). This implies that our strategy of investigating the effects of long-term stress are too great, resulting in an inefficient way to characterize a stress tolerance phenotype in *mpg1*. Additional considerations should be taken concerning the altered shift in development between *mpg1* and

wild-type, as the effects of stress change during different developmental states. Thus, future experimentation should explore the effects of acute stress during similar growth stages between *mpg1* and wild-type to more accurately address *mpg1* under abiotic stress.

Additionally, further characterization of plant defense should be explored. Although *mpg1* plants did not show a significant difference in symptomatic response to *Xoo* challenge compared to wild-type, its infection provoked the expression of the native *MPG1* in wild-type plants. This suggests that *MPG1* might natively function in plant defense response. Investigation into *mpg1*, and the effects of its ectopic expression of *MPG1*, concerning plant-pathogen interaction could prove to be an interesting avenue for future experimentation.

The preliminary analysis of seed characteristics suggests the need for further experimentation. The analysis of initial characteristics implies that differences in panicle architecture may exist between *mpg1* and wild-type. We hypothesize that the increased seed yield observed in *mpg1* is the result of potentially more panicles (variable, with higher incidence under non-optimal conditions), as well as more seeds per panicle (panicle architecture). Investigation into floral meristematic tissue and further comprehensive phenotypic analysis should be performed to better understand the nature of increased seed yield in *mpg1*.

Overall, this data supports that the T-DNA insertion and resulting ectopic expression of *MPG1* directly influences the generation of greater terminal biomass characteristics and seed yield in rice, as well as enhanced ratooning ability. Continued phenotyping could better illuminate our understanding of the extent of these individual characteristics. Speculation and information on some of these characteristics can be found in (Chapters 3-4).

APPENDIX B: Supplementary data – RNA-SEQ analysis pipeline selection and information on differential gene expression results of *mpg1* mutant plants vs. wild-type null segregant *Oryza sativa* in whole tiller tissue at 42 days post-planting

Introduction

The novel rice mutant *mpg1* exhibits ectopic expression of an AP2/EREBP transcription factor by result of a T-DNA insertion mutation resulting in increased plant biomass, seed yield, and possible degrees of stress tolerance. Transcription factors function to modulate transcription of targeted genes. Therefore, altered expression of a transcription factor can result in changed expression of a number of other genes, culminating in large transcriptomic differences. Thus, it is important to investigate global transcriptional differences that result from the ectopic expression of *MPG1* in *mpg1*. Understanding these differences could help identify molecular mechanisms responsible for generating aspects of *mpg1*'s phenotype, providing future directions for this study.

To investigate this, RNA-seq and differential gene expression analyses were conducted between HM-*mpg1* and WT-ns plants at timepoints proximal to when increased biomass characteristics of *mpg1*were observed relative to wild-type. One experiment was conducted using whole tiller tissue at 32 days post-planting (roughly 10 days prior to when *mpg1* shows greater girth than wild-type), and another was conducted using whole tiller tissue at 42 days post-planting (the timepoint where *mpg1* shows greater girth than wild type).

Further investigation regarding the development of these plants revealed that wild-type plants begin their transition from vegetative to reproductive stages of development around 42 days post-planting, while *mpg1* did not. Therefore analysis comparing *mpg1* and wild-type at this timepoint would be difficult to assess, as these samples would not represent a side-by-side

evaluation addressing only the effects of ectopic expression of *MPG1* in *mpg1* against wild-type, but the effects of different stages of development as well. Due to this and differing growth parameters, analysis of transcriptomic differences between *mpg1* and wild-type focused on the 32 days post-planting experiment (Chapter 4). However, elements of the 42 days post-planting experiment were also reported (number of differentially expressed genes, expression surrounding the T-DNA insertion site, and co-expressed differentially regulated genes between 32 and 42 days post-planting).

Presented here are the results of the differential expression analysis, from RNA-seq using whole tiller tissue during the 42 days post-planting time point experiment, used to support data and results in (Chapter 4). Additionally, included in this analysis is the rationale behind the selected RNA-seq analysis pipeline used in our study.

Materials and Methods

Plant materials

Rice (*Oryza sativa* L. spp. *japonica* cv. Kitaake), including wild-type, and segregating lines of partial T-DNA insertion of expression cassette *CmGAS1pro*::$AtSUC1_{H65K}$ - mutant (*mpg1*), were used to assess expression analyses.

Growth conditions

Seeds were placed on germination paper and partially submerged in a 1:1000 dilution of MAXIM XL dual action fungicide (Syngenta) and sealed with parafilm. Seeds were incubated at 30º C under 12 h light cycles, until primary shoot and root development occurred (usually 5-7 days). Seedlings were then transferred to planting medium in the greenhouse. Planting media consisted of: 1 part (Profile) Greens Grade porous ceramic particulate, and 1 part (Promix) BX, soil. The contents were mixed to homogeneity, and transferred to pots. Pots were organized in

random fashion in a flat or tub with water covered with black plastic and watered until media was fully saturated and pots remained in roughly 3" of standing water. Plants were grown in 1.0 gallon size pots in greenhouse conditions maintained at roughly 26˚C and 75% RH with a 16h light cycle. Plants were grown in 0.75 gallon size pots in growthchamber conditions maintained at roughly 26˚C and 80% RH with a 13h light cycle. Plant chlorosis was monitored and preemptively treated around the 3- to 4-leaf stage using Sprint 330 Iron Chelate at 0.3g/L water and top-watered. At the same developmental stage plants were fertilized using granulized (Technigro) 15-5-15 Plus Cal-Mag at 48.87 g/1.0 L and top-watered. Fertilizer treatment occurred twice weekly through maturity until harvest.

DNA extraction

Young, fresh plant tissue (3-leaf stage) was sampled for DNA extraction and analysis (2-5cm of leaf-tip). DNA was obtained via mechanical disruption of tissue and Shorty-Buffer extraction. Tissue was flash frozen in liquid nitrogen and disrupted using the Qiagen TissueLyser at 30 rps for a period of 1 minute. Five hundred uL of freshly prepared shorty buffer (0.2 M Tris HCl pH 9.0, 0.4 M LiCl, 25 mM EDTA, and 1.0% SDS) was added to each tissue sample, vortexed and centrifuged at max speed (13k rpm) for 5 min. Then 350 uL of supernatant was transferred to a new tube containing 400uL isopropanol, mixed by inverting and centrifuged at max speed for 10 min. The supernatant was discarded and 1mL of 70% ethanol was added to each sample to wash the DNA pellet. Samples were then centrifuged at max speed for 10 minutes, the supernatant was discarded and the tubes were inverted for 30 minutes. The DNA pellet was resuspended in 100uL of 10mM Tris-HCl pH 8 and stored short term at 4º C until use.

Haplotyping mutants

To identify *mpg1* plants that were homozygous, heterozygous, or null segregants for the

bi-laterally truncated T-DNA insertion, primers were designed in regions directly flanking the site of insertion, as well as primers spanning the integration site into the T-DNA, and used in PCR. The primers flanking the T-DNA insertion were wFLA forward: 5'-GGAAGTTGGAGATGGGAAACA-3', and wFLA reverse: 5'-GGCCTCGTGTGTCAGTAATAA-3'. The primers spanning the genomic region and the T-DNA insertion were wIN forward: 5'-ACACCGGAAGCATAGTCATTT-3', and wIN reverse: 5'-GGTCGCCAACATCTTCTTCT-3'.

RNA-seq

RNA-extraction and Library Preparation: Two RNA-seq experiments were performed, one on 32 days post-planting material (4 individual plants each from two genotypes grown in greenhouse conditions), and another on 42 days post-planting material (3 individual plants each from two genotypes grown in growth chamber conditions). Whole tiller tissue was selected for analysis, from roughly 5cm above soil line and up (collected mid-day) of two genotypes (HM-*mpg1* and WT-ns. Tissue was placed in individual 50 mL conical tubes and flash frozen in liquid nitrogen. Tissue was ground using the TissueLyser (Qiagen) at 30 rps for 1 min, and RNA was extracted using (Qiagen) Plant RNeasy mini-kit. Desired tissue (not more than 100mg) was sampled for RNA extraction and analysis. RNA was treated with DNase and purified using the Turbo DNase kit (Invitrogen). RNA quality control was verified using a Bioanalyzer (Agilent) and TapeStation (Agilent) for the 32 and 42 days post-planting respectively. Libraries were generated using the TruSeq RNA-seq kit (Illumina) as per manufacture instructions, 125bp paired-end sequencing of the library for 32 days post-planting material was done at the RTSF GENOMICS Core facility at Michigan State University using Illumina Hi seq 2500 system. 75bp paired-end sequencing of the library for 42 days post-planting material was done at the Self-

Service Next Generation Sequencing Core facility at Colorado State University using Illumina NextSeq system.

Mapping of Reads and Identification of the DEG: Resulting reads were assessed for quality control using (FastQC) (Andrews, 2010), where results fell within acceptable parameters. This was followed with FASTQ Toolkit (Illumina-BaseSpace-Labs, 2018) performing quality trimming of anything less than 20 phred score. Due to the inherent nature of RNA-seq pipelines, several methods were originally evaluated (STAR(2.5.3a)-Cufflinks2(2.2.0)-Cuffdiff2(2.2.0) (Dobin et al., 2013, Trapnell et al., 2012) , STAR(2.5.3a)-HTSeq(0.6.1)-DESeq2(1.14.1) (Dobin et al., 2013, Anders et al., 2015, Anders et al. 2010, Love et al., 2014), and STAR(2.5.3a)-HTSeq(0.6.1)-edgeR(3.0) (Dobin et al., 2013, Anders et al., 2015, Robinson et al., 2010)). The analysis pipeline selected utilized STAR (2.5.3a) (Dobin et al., 2013), Htseq-count-merge (0.6.1) (Anders et al., 2015), and edgeR (3.0) (Robinson et al., 2010) through the Cyverse workflow interface (Merchant et al., 2016). Reads were aligned to the Ensembl MSU6.0 version of the rice genome available through (http://rice.plantbiology.msu.edu). Total mapped read count ranged between 40-70 million, and 50-220 reads per sample for 32 and 42 days post-planting RNA-seq experiments respectively.

Hierarchical clustering and heatmaps were created using the R package heatmap.2 with hclust and dendrogram formation functions. Gene co-expression was evaluated using Genevestigatior software (Hruz et al., 2008). Venn diagrams were generated using BioVenn (Hulsen et al., 2008).

Promoter analysis

Assessment of the number of GCC-box (GCCGCC) motif elements present in promoter regions was conducted using R and searching the Ensemble MSU6.0 genome build and

extracting number of motif elements present in the range of -2000 bp from the transcription start site (TSS) for every gene. These numbers were later cross referenced against our generated list of DEG from RNA-seq.

GO enrichment analysis

GO analysis was performed for term enrichment using RiceNetDB (Chen, M. 2013). Single enrichment analysis with GO annotations was performed with a corrected P-value <= 0.01. The genes that are up- or down-regulated for each data set were analyzed separately. To identify various genes that correlate with the pleiotropic phenotype of *mpg1*, several GO terms and their correlating genes were placed in separate tables for partitioning and assessment.

Results

Selection of RNA-seq differential expression analysis pipeline

In both rounds of RNA-seq multiple analysis pipelines were analyzed prior to deeply assessing differential gene outputs. RNA-seq differential gene expression analysis functions simply by aligning processed reads to a reference genome, followed by evaluating the abundance of reads at particular loci, and finalized by statistical assessment of comparing the read abundances between groups or treatments. Each step can significantly alter the outcome of what is deemed significantly different in the output of these pipelines. Assumptions made within these analysis platforms can result in altered numbers of significantly differentially expressed genes, false-positives, or false-negatives. This became apparent after initial assessment of our data using the traditional Tuxedo platform of assessment resulted in a false-negative call for our gene of interest *MPG1*, which has been shown during our investigation to in fact be differentially expressed in *mpg1* compared to wild-type. Numerous pipelines exist, however three were chosen for initial assessment (Figure B1, B2).

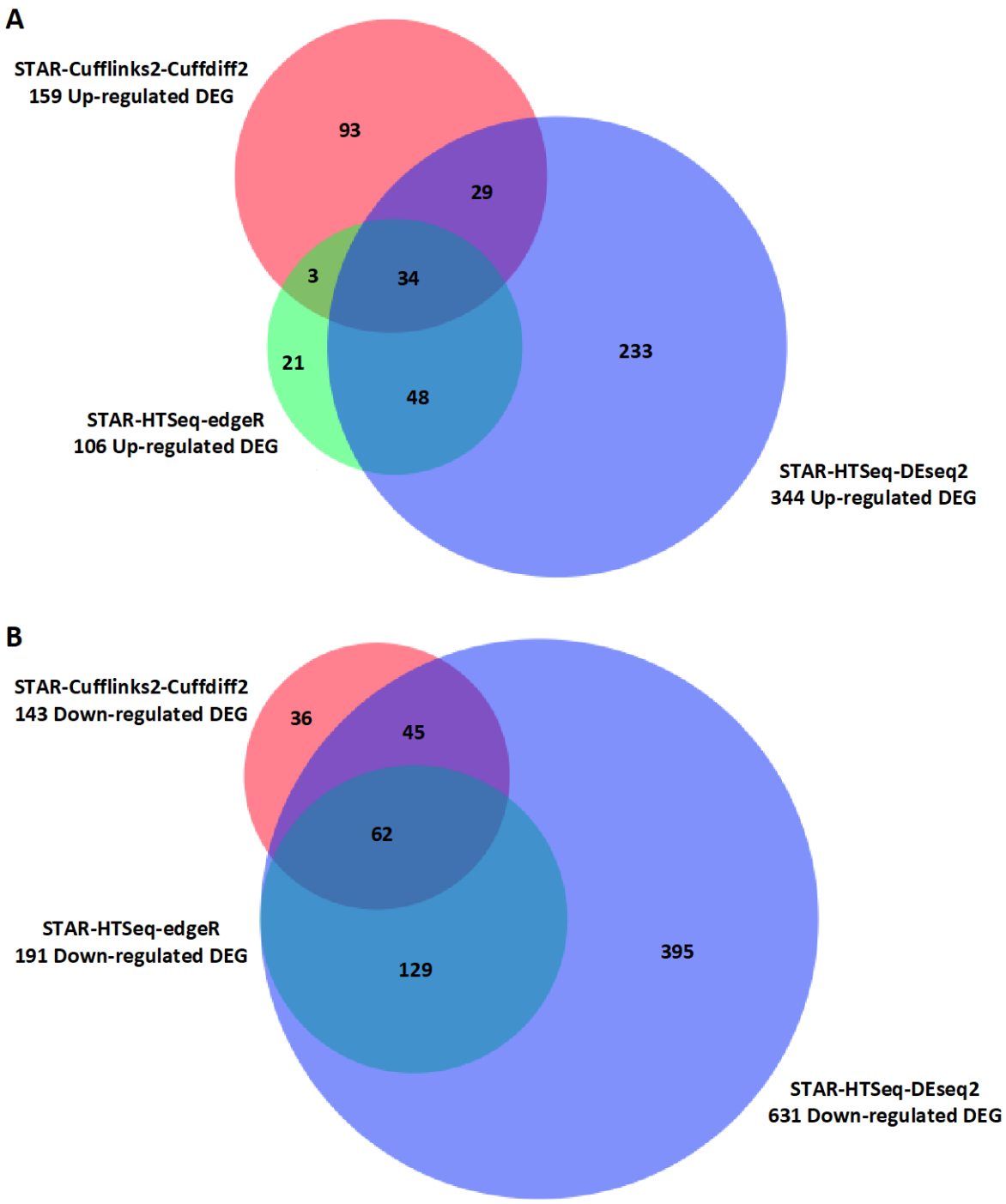


Figure B.1: Analysis of various RNA-seq pipelines utilizing whole tiller tissue from 32 days post-planting.
(A) Venn diagram of resulting calls for differential gene expression from three different RNA-seq pipelines for differentially expressed genes upregulated in HM-mpg1 compared to WT-ns. (B) Venn diagram of resulting calls for differential gene expression from three different RNA-seq pipelines for differentially expressed genes downregulated in HM-mpg1 compared to WT-ns.

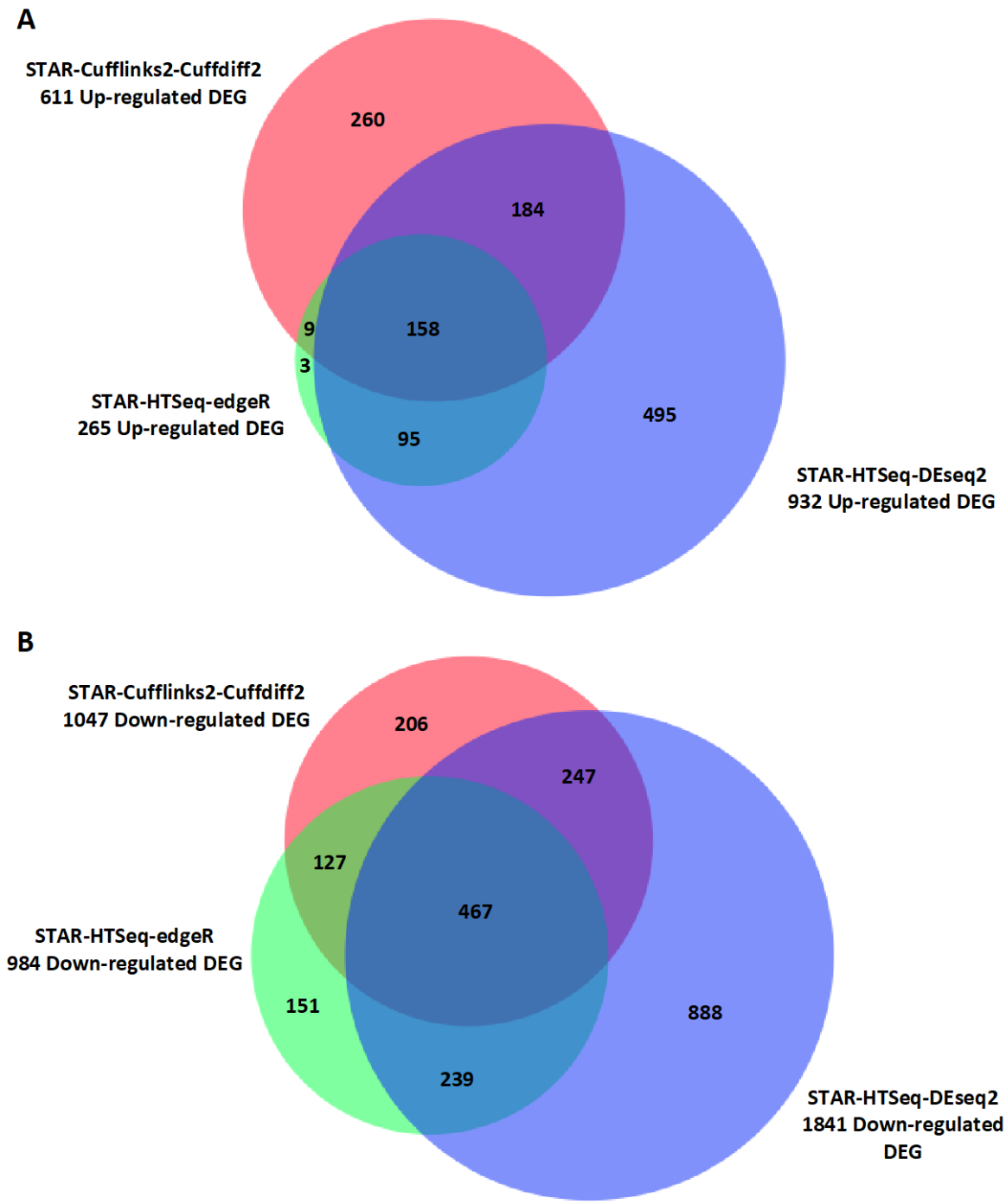


Figure B.2: Analysis of various RNA-seq pipelines utilizing whole tiller tissue from 42 days post-planting.
(A) Venn diagram of resulting calls for differential gene expression from three different RNA-seq pipelines for differentially expressed genes upregulated in HM-mpg1 compared to WT-ns. (B) Venn diagram of resulting calls for differential gene expression from three different RNA-seq pipelines for differentially expressed genes downregulated in HM-mpg1 compared to WT-ns.

Assessment of these various RNA-seq analysis pipelines for differential expression resulted in strikingly different outputs. All of these platforms have been supported and shown successful in determining differentially regulated genes in RNA-seq experiments. Therefore, we opted to select the platform with the best overlap and the least number of differentially regulated genes (with our gene of interest being included) in effort to minimize any artifact that might be occurring in other platforms. This being noted, utilizing the platform we selected could be limiting our view on what is occurring in *mpg1* compared to wild-type that might have been reported using other analysis pipelines that reported a greater number of genes deemed to be significantly differentially expressed.

Information pertaining to differentially expressed genes between HM-*mpg1* and WT-ns at 42 days post-planting

Due to the likelihood that this experiment is evaluating vegetative phase *mpg1* against reproductive phase WT-ns, assessments will not be discussed in detail. The ability to distinguish between the number of genes associated with reproduction and what might be differently expressed due to ectopic expression of *MPG1* remains difficult. The following figures are intended to be used as supplementary information alongside the 32 days post-planting RNA-seq experiment.

Sample variation via hierarchical clustering of normalized reads between samples is shown (Figure B.3). A list of all the differentially regulated genes and relevant information between HM-*mpg1* and WT-ns are provided (Table B.1). Although difficult to extrapolate upon, RiceNETDB was utilized to investigate ontological enrichment categories for this timepoint (Figure B.4-B6). Genes inversely expressed between 32 and 42 days post-planting are also listed (Table B.2).

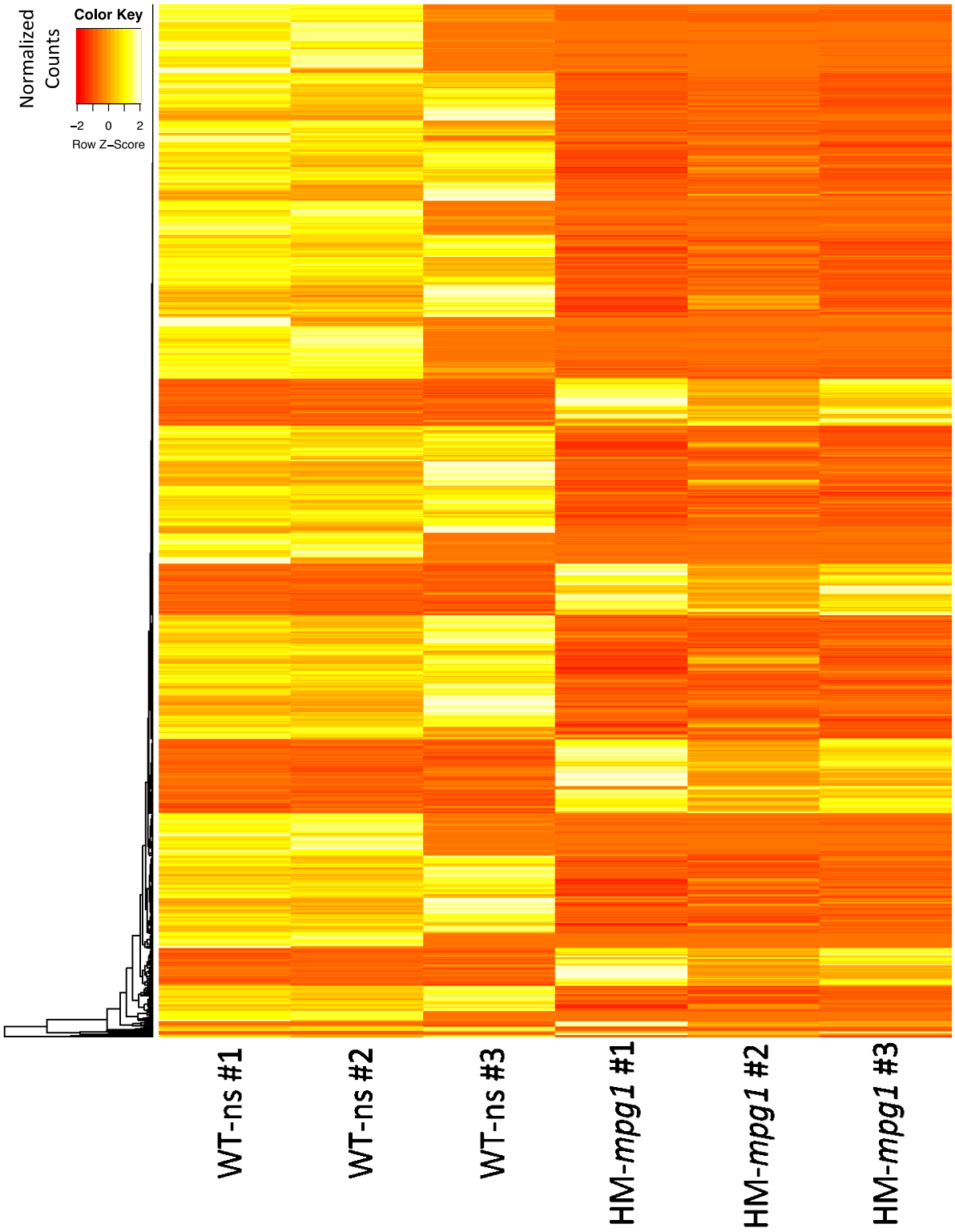


Figure B.3: Heatmap and hierarchical clustering analysis.
Each lane represents an individual biological replicate from 42 days post-planting tissue by normalized counts.

Table B.1: DEG list from RNA-seq analysis of HM-*mpg1* compared to WT-ns within 42 days post-planting tiller tissue.

MSU ID	RAP ID	log2 Fold Change	Corrected p-value	GCC-box	Description
Up-regulated DEG					
LOC_Os08g41030	Os08g0521600	9.478	2.7113E-68	2	AP2 domain containing protein, expressed
LOC_Os01g22570	Os01g0329000	4.506	0.0008668	0	glycerol-3-phosphate acyltransferase, putative, expressed
LOC_Os09g35020	Os09g0522100	4.328	8.372E-23	3	AP2 domain containing protein, expressed
LOC_Os01g22560	Os01g0329000	3.822	0.00345409	2	glycerol-3-phosphate acyltransferase 1, putative, expressed
LOC_Os06g13390	Os06g0242000	3.716	0.00024761	1	SAM dependent carboxyl methyltransferase, putative, expressed
LOC_Os09g28440	Os09g0457900	3.266	8.3877E-12	3	AP2 domain containing protein, expressed
LOC_Os03g22620	Os03g0347900	3.238	1.3213E-12	0	terpene synthase family, metal binding domain containing protein, expressed
LOC_Os03g26530	Os03g0382100	3.168	5.856E-05	0	3-ketoacyl-CoA synthase, putative, expressed
LOC_Os09g35010	Os09g0522000	3.025	3.5267E-13	0	dehydration-responsive element-binding protein, putative, expressed
LOC_Os09g17560	Os09g0344500	2.973	1.1096E-06	0	O-methyltransferase, putative, expressed
LOC_Os01g38980	Os01g0570800	2.946	2.9153E-14	0	calmodulin-binding protein, putative, expressed
LOC_Os04g28620	Os04g0354600	2.634	0.00589952	0	male sterility protein, putative, expressed
LOC_Os03g38470	Os03g0581400	2.586	1.1671E-07	0	GDSL-like lipase/acylhydrolase, putative, expressed
LOC_Os03g37960	Os03g0576600	2.548	0.0001703	1	acyl CoA binding protein, putative, expressed
LOC_Os03g18740	Os03g0299200	2.51	2.7465E-06	0	oxidoreductase, short chain dehydrogenase/reductase family, putative, expressed
LOC_Os03g54130	Os03g0752500	2.449	0.03325479	0	cysteine protease 1 precursor, putative, expressed
LOC_Os06g39370	Os06g0594400	2.434	1.0259E-06	6	OsFBK16 - F-box domain and kelch repeat containing protein, expressed
LOC_Os01g64470	Os01g0864500	2.407	1.7344E-11	1	harpin-induced protein 1 domain containing protein, expressed
LOC_Os05g08860	Os05g0181300	2.397	1.0259E-06	0	expressed protein
LOC_Os09g29710	Os09g0472900	2.381	0.00155453	2	beta-expansin precursor, putative, expressed
LOC_Os02g37300	Os02g0584800	2.369	9.5403E-09	0	heavy metal associated domain containing protein, expressed
LOC_Os02g20540	Os02g0308400	2.364	5.081E-09	1	fasciclin domain containing protein, expressed
LOC_Os11g44380	Os11g0665600	2.34	4.1496E-06	4	expressed protein
LOC_Os06g39110	Os06g0591200	2.33	6.718E-06	2	expressed protein
LOC_Os02g15290	Os02g0251900	2.304	0.02443664	1	VQ domain containing protein, putative, expressed
LOC_Os08g35760	Os08g0460000	2.283	3.1422E-11	0	Cupin domain containing protein, expressed
LOC_Os08g28240	Os08g0369800	2.269	0.00094048	0	carotenoid cleavage dioxygenase, putative, expressed
LOC_Os03g44880	Os03g0651800	2.259	6.6955E-05	0	Cupin domain containing protein, expressed
LOC_Os04g39350	Os04g0469000	2.251	8.4482E-05	6	heavy metal associated domain containing protein, expressed
LOC_Os03g55770	Os03g0766600	2.251	0.03722762	0	expressed protein
LOC_Os01g11730	Os01g0216000	2.213	0.00397005	4	GDSL-like lipase/acylhydrolase, putative, expressed

LOC_Os07g33580	Os07g0519600	2.204	0.00054537	1	cytochrome P450, putative, expressed
LOC_Os10g08026	Os10g0163370	2.201	0.00749639	1	lecithin:cholesterol acyltransferase, putative, expressed
LOC_Os06g20920	Os06g0314600	2.194	0.0001006	0	SAM dependent carboxyl methyltransferase, putative, expressed
LOC_Os04g48290	Os04g0571600	2.176	5.1013E-11	2	MATE efflux family protein, putative, expressed
LOC_Os03g18150	Os03g0292100	2.153	4.9113E-07	1	protein phosphatase 2C, putative, expressed
LOC_Os09g27320	Os09g0445500	2.116	0.00299318	1	ZOS9-13 - C2H2 zinc finger protein, expressed
LOC_Os08g13570	Os08g0232700	2.098	5.0987E-08	0	exo70 exocyst complex subunit family protein, putative, expressed
LOC_Os10g40420	Os10g0551700	2.098	0.00027646	2	LTPL138 - Protease inhibitor/seed storage/LTP family protein precursor, expressed
LOC_Os10g20890	Os10g0349900	2.072	0.0015044	2	LTPL137 - Protease inhibitor/seed storage/LTP family protein precursor, expressed
LOC_Os01g65130	Os01g0871800	2.068	0.00084718	0	peptide transporter, putative, expressed
LOC_Os09g35780	Os09g0526500	2.033	0.03652849	0	BAP2, putative, expressed
LOC_Os07g48280	Os07g0680600	2.008	7.5918E-07	9	expressed protein
LOC_Os12g29160	Os12g0474951	2.002	3.4038E-06	0	LTPL105 - Protease inhibitor/seed storage/LTP family protein precursor, putative, expressed
LOC_Os11g10470	Os11g0210201	1.996	1.0067E-06	0	expressed protein
LOC_Os03g56160	Os03g0772600	1.975	0.000349	1	lectin-like receptor kinase 7, putative, expressed
LOC_Os03g55776	Os03g0766800	1.972	0.00020513	0	expressed protein
LOC_Os09g39410	Os09g0567500	1.963	5.8631E-09	0	male sterility protein, putative, expressed
LOC_Os12g32610	Os12g0510750	1.945	0.00071091	2	expressed protein
LOC_Os02g37290	Os02g0584700	1.927	1.8105E-05	0	heavy metal associated domain containing protein, expressed
LOC_Os08g02230	Os08g0114300	1.917	2.9385E-06	0	FAD-binding and arabino-lactone oxidase domains containing protein, putative, expressed
LOC_Os01g36294	Os01g0543600	1.917	0.00555528	1	cytochrome P450, putative, expressed
LOC_Os09g35030	Os09g0522200	1.907	3.4813E-07	2	dehydration-responsive element-binding protein, putative, expressed
LOC_Os02g44320	Os02g0662100	1.901	4.2913E-07	0	LTPL113 - Protease inhibitor/seed storage/LTP family protein precursor, expressed
LOC_Os05g31620	Os05g0380900	1.895	0.00040817	0	OsCML15 - Calmodulin-related calcium sensor protein, expressed
LOC_Os11g42660	Os11g0646300	1.891	4.4405E-05	1	Leucine Rich Repeat family protein, expressed
LOC_Os08g29660	Os08g0386200	1.888	0.00026802	0	WRKY69, expressed
LOC_Os11g15340	Os11g0260100	1.883	3.2569E-05	1	SAM dependent carboxyl methyltransferase family protein, putative, expressed
LOC_Os03g04070	Os03g0133000	1.878	2.1532E-05	2	no apical meristem protein, putative, expressed
LOC_Os02g31867	Os02g0518000	1.873	3.6443E-05	0	expressed protein
LOC_Os08g36480	Os08g0468100	1.86	0.00115038	0	nitrate reductase, putative, expressed
LOC_Os05g08900	Os05g0181700	1.852	7.383E-05	0	expressed protein
LOC_Os02g02400	Os02g0115700	1.851	9.8729E-08	0	catalase isozyme A, putative, expressed
LOC_Os02g50374	Os02g0736900	1.851	0.00633704	2	expressed protein
LOC_Os02g14490	Os02g0241200	1.83	0.028786	0	MYB family transcription factor, putative, expressed
LOC_Os01g56240	Os01g0768333	1.825	0.00052923	1	OsSAUR2 - Auxin-responsive SAUR gene family member, expressed

LOC_Os04g48350	Os04g0572400	1.824	0.00015451	1	dehydration-responsive element-binding protein, putative, expressed
LOC_Os05g16430	Os05g0253200	1.811	0.00867175	12	SHR5-receptor-like kinase, putative, expressed
LOC_Os01g04280	Os01g0134700	1.804	1.8327E-05	2	calmodulin binding protein, putative, expressed
LOC_Os01g36350	Os01g0544200	1.797	0.00503219	6	cytochrome P450, putative, expressed
LOC_Os05g37600	Os05g0448300	1.794	0.00665632	12	glycerol-3-phosphate acyltransferase, putative, expressed
LOC_Os03g64170	Os03g0859100	1.789	1.2778E-06	0	GDSL-like lipase/acylhydrolase, putative, expressed
LOC_Os03g60570	Os03g0820400	1.787	0.00186946	0	ZOS3-22 - C2H2 zinc finger protein, expressed
LOC_Os01g74040	Os01g0972000	1.785	0.00022375	0	zinc finger, RING-type, putative, expressed
LOC_Os03g41330	Os03g0609500	1.785	0.00137682	0	DUF260 domain containing protein, putative, expressed
LOC_Os06g39120	Os06g0591400	1.784	0.00352997	2	expressed protein
LOC_Os05g39930	Os05g0476700	1.783	3.8019E-06	6	spotted leaf 11, putative, expressed
LOC_Os02g40260	Os02g0616100	1.783	0.00043994	3	uncharacterized protein At4g06744 precursor, putative, expressed
LOC_Os11g10120	Os11g0207400	1.774	0.00491243	0	expressed protein
LOC_Os09g33710	Os09g0511900	1.773	0.00021823	1	Os9bglu33 - beta-glucosidase homologue, similar to G. max hydroxyisourate hydrolase, expressed
LOC_Os01g58640	Os01g0800500	1.773	0.03191788	0	nucleotide pyrophosphatase/phosphodiesterase, putative, expressed
LOC_Os11g45060	Os11g0676050	1.765	0.02545813	1	NB-ARC domain containing protein, expressed
LOC_Os01g36720	Os01g0547600	1.745	0.04879952	0	transporter, major facilitator family, putative, expressed
LOC_Os01g51670	Os01g0714600	1.743	2.8523E-06	6	expressed protein
LOC_Os01g33160	Os01g0516200	1.74	0.02865326	0	stress responsive A/B Barrel domain containing protein, expressed
LOC_Os01g58140	Os01g0793900	1.738	0.00017189	1	expressed protein
LOC_Os03g50210	Os03g0710000	1.724	0.00061423	1	DUF292 domain containing protein, expressed
LOC_Os04g46990	Os04g0556600	1.715	0.00129141	0	cis-zeatin O-glucosyltransferase, putative, expressed
LOC_Os02g44720	Os02g0667300	1.714	0.00030971	2	expressed protein
LOC_Os06g20150	Os06g0306300	1.713	0.0469517	1	peroxidase precursor, putative, expressed
LOC_Os01g70790	Os01g0934100	1.709	2.0674E-05	0	SRC2 protein, putative, expressed
LOC_Os09g23570	Os09g0400500	1.708	0.01103768	1	inactive receptor kinase At2g26730 precursor, putative, expressed
LOC_Os01g72934	None	1.705	0.01627063	0	None
LOC_Os01g62670	Os01g0845100	1.701	0.00021889	0	avr9/Cf-9 rapidly elicited protein, putative, expressed
LOC_Os08g10244	Os08g0202300	1.693	0.01097082	0	retrotransposon protein, putative, unclassified, expressed
LOC_Os01g72530	Os01g0955100	1.692	6.0509E-05	2	OsCML31 - Calmodulin-related calcium sensor protein, expressed
LOC_Os11g41710	Os11g0635500	1.688	6.0509E-05	3	cytochrome P450, putative, expressed
LOC_Os06g21820	Os06g0323100	1.684	0.04473431	10	jasmonate O-methyltransferase, putative, expressed
LOC_Os08g10150	Os08g0201700	1.682	0.00029572	9	SHR5-receptor-like kinase, putative, expressed
LOC_Os02g15280	Os02g0251800	1.669	0.04977843	3	VQ domain containing protein, putative, expressed
LOC_Os09g18159	Os09g0350900	1.667	0.00457387	0	light repressible receptor protein kinase, putative, expressed

LOC_Os02g11130	Os02g0206400	1.662	0.02356323	0	cytokinin-O-glucosyltransferase 3, putative, expressed
LOC_Os07g46480	Os07g0658600	1.654	1.6E-05	1	eukaryotic aspartyl protease domain containing protein, expressed
LOC_Os11g45090	Os11g0675200	1.647	0.00778911	0	NB-ARC domain containing protein, expressed
LOC_Os01g09220	Os01g0186900	1.636	3.5753E-05	0	transposon protein, putative, CACTA, En/Spm sub-class, expressed
LOC_Os11g45190	Os11g0677101	1.631	0.04226209	1	NB-ARC domain containing protein, expressed
LOC_Os02g13510	Os02g0228300	1.625	0.00119421	3	receptor-like protein kinase 5 precursor, putative, expressed
LOC_Os04g27190	Os04g0340300	1.623	4.8074E-05	4	terpene synthase, putative, expressed
LOC_Os10g41330	Os10g0562900	1.62	0.00101144	1	AP2 domain containing protein, expressed
LOC_Os07g40900	Os07g0600000	1.613	0.01434423	0	retrotransposon protein, putative, unclassified, expressed
LOC_Os02g37330	Os02g0585200	1.606	0.00424538	0	heavy metal associated domain containing protein, expressed
LOC_Os06g15170	Os06g0262800	1.599	0.00045115	0	3-ketoacyl-CoA synthase, putative, expressed
LOC_Os04g29680	Os04g0366000	1.589	0.00281142	1	OsWAK38 - OsWAK receptor-like protein kinase, expressed
LOC_Os01g26280	Os01g0364800	1.589	0.00898406	5	OsWAK8 - OsWAK receptor-like protein kinase, expressed
LOC_Os06g14420	Os06g0255400	1.585	0.00152747	0	hydrolase, NUDIX family, domain containing protein, expressed
LOC_Os05g50900	Os05g0586300	1.58	0.00831573	1	helix-loop-helix DNA-binding protein, putative, expressed
LOC_Os06g49220	Os06g0705700	1.574	0.00795992	0	peptide transporter, putative, expressed
LOC_Os02g11110	Os02g0206100	1.569	0.00835678	0	flavonol-3-O-glycoside-7-O-glucosyltransferase 1, putative, expressed
LOC_Os06g13560	Os06g0244000	1.567	0.00013558	0	SAM dependent carboxyl methyltransferase, putative, expressed
LOC_Os06g04240	Os06g0133500	1.567	0.00093073	1	expressed protein
LOC_Os07g07900	Os07g0175300	1.567	0.00299046	1	expressed protein
LOC_Os05g43760	Os05g0513100	1.563	1.999E-05	8	TCP family transcription factor, putative, expressed
LOC_Os01g17050	Os01g0278000	1.56	0.0490105	2	VQ domain containing protein, putative, expressed
LOC_Os11g07200	Os11g0173100	1.556	0.04073249	0	receptor protein kinase CLAVATA1 precursor, putative, expressed
LOC_Os04g43390	Os04g0513400	1.546	0.00340258	0	Os4bglu16 - monolignol beta-glucoside homologue, expressed
LOC_Os05g39720	Os05g0474800	1.532	0.00910967	0	WRKY70, expressed
LOC_Os04g27670	Os04g0344100	1.511	0.00015611	0	terpene synthase family, metal binding domain containing protein, expressed
LOC_Os06g01890	Os06g0108500	1.495	0.00278005	0	MADS-box transcription factor, putative, expressed
LOC_Os10g06000	Os10g0150800	1.487	0.014366	2	POEI15 - Pollen Ole e I allergen and extensin family protein precursor, expressed
LOC_Os03g20090	Os03g0315400	1.479	2.3653E-05	0	MYB family transcription factor, putative, expressed
LOC_Os05g39830	Os05g0476000	1.472	0.02443664	0	expressed protein
LOC_Os05g03972	Os05g0130400	1.468	0.00449765	1	plant protein of unknown function domain containing protein, expressed
LOC_Os12g38270	Os12g0570700	1.466	0.00782317	0	metallothionein, putative, expressed
LOC_Os01g61080	Os01g0826400	1.461	6.0509E-05	1	WRKY24, expressed
LOC_Os11g37700	Os11g0587600	1.459	0.00695215	8	pleiotropic drug resistance protein, putative, expressed

LOC_Os10g31780	Os10g0456100	1.45	0.00109469	1	oxidoreductase, short chain dehydrogenase/reductase family domain containing protein, expressed
LOC_Os07g49140	Os07g0691700	1.444	4.0434E-05	0	expressed protein
LOC_Os01g02550	Os01g0115500	1.444	0.02642388	0	protein kinase domain containing protein, expressed
LOC_Os12g02320	Os12g0115100	1.44	2.0482E-05	4	LTPL12 - Protease inhibitor/seed storage/LTP family protein precursor, expressed
LOC_Os02g33680	Os02g0540700	1.435	0.00025653	0	U-box domain containing protein, expressed
LOC_Os06g03810	Os06g0128800	1.431	0.02457364	0	expressed protein
LOC_Os04g21820	Os04g0286300	1.425	0.01679088	0	OsWAK33 - OsWAK receptor-like protein OsWAK-RLP, expressed
LOC_Os01g29330	Os01g0389700	1.421	0.00374463	0	expressed protein
LOC_Os04g15580	Os04g0226600	1.418	0.03200693	0	serine/threonine-protein kinase receptor precursor, putative, expressed
LOC_Os02g37654	Os02g0589000	1.411	0.01322726	0	lecithin:cholesterol acyltransferase, putative, expressed
LOC_Os04g43440	Os04g0514600	1.399	0.03389018	1	NB-ARC/LRR disease resistance protein, putative, expressed
LOC_Os11g43250	Os11g0653300	1.396	0.00679743	0	Leucine Rich Repeat family protein, expressed
LOC_Os01g39330	Os01g0575200	1.394	0.00129141	0	helix-loop-helix DNA-binding domain containing protein, expressed
LOC_Os06g46950	Os06g0683400	1.385	0.00086837	1	EF hand family protein, putative, expressed
LOC_Os05g47960	Os05g0552800	1.382	0.01318969	3	expressed protein
LOC_Os08g02160	None	1.382	0.01371539	0	no apical meristem protein, putative, expressed
LOC_Os10g36170	Os10g0505700	1.374	4.4602E-05	0	LTPL160 - Protease inhibitor/seed storage/LTP family protein precursor, expressed
LOC_Os10g37760	Os10g0521900	1.368	0.00013351	0	OsRhmbd17 - Putative Rhomboid homologue, expressed
LOC_Os05g09020	Os05g0183100	1.365	0.00126948	1	WRKY67, expressed
LOC_Os03g15230	Os03g0257600	1.355	0.01749839	5	DUF292 domain containing protein, expressed
LOC_Os01g38359	Os01g0564300	1.348	0.0001879	3	peptidyl-prolyl cis-trans isomerase, FKBP-type, putative, expressed
LOC_Os02g50730	Os02g0740700	1.344	0.01749839	1	metalloendoproteinase 1 precursor, putative, expressed
LOC_Os04g58850	Os04g0685300	1.333	0.00089091	1	harpin-induced protein 1 domain containing protein, expressed
LOC_Os11g47809	Os11g0704500	1.33	0.00115655	1	metallothionein, putative, expressed
LOC_Os12g28590	Os12g0471100	1.327	0.01213121	1	ATPase 2, putative, expressed
LOC_Os01g27230	Os01g0369900	1.319	0.04962162	0	12-oxophytodienoate reductase, putative, expressed
LOC_Os05g44420	Os05g0520600	1.315	0.03496467	0	expressed protein
LOC_Os01g66610	Os01g0889900	1.314	0.00216615	0	serine/threonine-protein kinase receptor precursor, putative, expressed
LOC_Os06g22440	Os06g0329900	1.309	0.0177222	1	SAM dependent carboxyl methyltransferase, putative, expressed
LOC_Os04g29790	Os04g0460600	1.309	0.04387977	8	OsWAK40 - OsWAK receptor-like protein OsWAK-RLP, expressed
LOC_Os01g44069	Os01g0631400	1.304	0.00257588	1	glycerol-3-phosphate acyltransferase, putative, expressed
LOC_Os04g38720	Os04g0460600	1.304	0.00409114	0	no apical meristem protein, putative, expressed
LOC_Os10g05980	Os10g0150600	1.304	0.02542294	6	POEI13 - Pollen Ole e I allergen and extensin family protein precursor, expressed
LOC_Os01g11650	Os01g0214800	1.294	0.00357854	0	GDSL-like lipase/acylhydrolase, putative, expressed
LOC_Os04g17650	Os04g0249500	1.289	0.03137516	1	sucrose synthase, putative, expressed

LOC_Os03g19070	Os03g0302800	1.286	0.00580681	0	long cell-linked locus protein, putative, expressed
LOC_Os12g10740	Os12g0210400	1.282	0.00422929	0	leucine-rich repeat family protein, putative, expressed
LOC_Os05g34320	Os05g0415700	1.279	0.00037424	5	beta-hexosaminidase precursor, putative, expressed
LOC_Os04g41310	Os04g0490500	1.278	0.00632498	3	STRUBBELIG-RECEPTOR FAMILY 8 precursor, putative, expressed
LOC_Os03g60650	Os03g0821300	1.272	0.00029604	0	protein phosphatase 2C, putative, expressed
LOC_Os02g48210	Os02g0712700	1.272	0.00114092	0	lectin-like protein kinase, putative, expressed
LOC_Os03g17870	Os03g0288000	1.269	0.02050968	0	metallothionein, putative, expressed
LOC_Os02g40410	Os02g0617100	1.268	0.0062518	2	expressed protein
LOC_Os09g25760	Os09g0425900	1.267	0.00091372	1	tetraspanin family protein, putative, expressed
LOC_Os06g04230	Os06g0133400	1.265	0.00762366	2	expressed protein
LOC_Os01g02560	Os01g0115600	1.252	0.03670787	0	Ser/Thr receptor-like kinase, putative, expressed
LOC_Os05g04020	None	1.244	0.01084727	1	plant protein of unknown function domain containing protein, expressed
LOC_Os06g47800	Os06g0693100	1.241	0.00898855	0	disease resistance protein RGA3, putative, expressed
LOC_Os11g35390	Os11g0558200	1.236	0.04263186	0	MYB family transcription factor, putative, expressed
LOC_Os05g45410	Os05g0530400	1.227	0.04519163	1	HSF-type DNA-binding domain containing protein, expressed
LOC_Os01g06280	Os01g0155500	1.226	0.02355113	0	TKL_IRAK_CrRLK1L-1.4 - The CrRLK1L-1 subfamily has homology to the CrRLK1L homolog, expressed
LOC_Os10g41060	Os10g0560000	1.222	0.01297359	1	expressed protein
LOC_Os03g28300	Os03g0401100	1.22	0.00682096	1	protein kinase domain containing protein, expressed
LOC_Os05g48360	Os05g0557400	1.217	0.00740453	0	membrane attack complex component/perforin/complement C9, putative, expressed
LOC_Os02g50570	Os02g0739100	1.215	0.0018082	0	retrotransposon protein, putative, unclassified, expressed
LOC_Os02g15690	Os02g0256100	1.214	0.0015044	0	polygalacturonase, putative, expressed
LOC_Os08g44270	Os08g0556900	1.209	0.00757152	0	vignain precursor, putative, expressed
LOC_Os02g04130	Os02g0134200	1.208	0.04098777	1	DUF1645 domain containing protein, putative, expressed
LOC_Os02g37700	Os02g0589700	1.207	0.01172683	0	lecithin:cholesterol acyltransferase, putative, expressed
LOC_Os03g12500	Os03g0225900	1.206	0.00332757	1	cytochrome P450, putative, expressed
LOC_Os01g61990	Os01g0837000	1.202	0.00264129	1	ankyrin repeat-containing protein, putative, expressed
LOC_Os06g16040	Os06g0271400	1.198	0.01317667	0	expressed protein
LOC_Os12g36110	Os12g0547600	1.194	0.00269726	0	calmodulin binding protein, putative, expressed
LOC_Os04g33900	Os04g0415600	1.192	0.00741493	0	ctr copper transporter family protein, putative, expressed
LOC_Os10g05750	Os10g0148100	1.192	0.01945666	1	POEI3 - Pollen Ole e I allergen and extensin family protein precursor, expressed
LOC_Os05g05680	Os05g0149400	1.189	0.0078072	0	1-aminocyclopropane-1-carboxylate oxidase, putative, expressed
LOC_Os01g47730	Os01g0667600	1.188	0.00184398	0	ras-related protein, putative, expressed
LOC_Os11g45400	Os11g0679700	1.184	0.03698062	0	glycerol-3-phosphate acyltransferase, putative, expressed
LOC_Os06g44160	Os06g0650900	1.184	0.04083643	0	heat shock protein DnaJ, putative, expressed

LOC_Os09g29200	Os09g0467200	1.183	0.04973018	0	glutathione S-transferase, putative, expressed
LOC_Os04g02754	Os04g0117900	1.182	0.0262397	2	amidase family protein, putative, expressed
LOC_Os01g73170	Os01g0962700	1.18	0.00137155	1	peroxidase precursor, putative, expressed
LOC_Os07g05010	Os07g0143200	1.18	0.01960592	0	helix-loop-helix DNA-binding domain containing protein, expressed
LOC_Os09g37540	Os09g0547500	1.177	0.00410928	0	uncharacterized protein PA4923, putative, expressed
LOC_Os02g44710	Os02g0667100	1.177	0.02111131	0	expressed protein
LOC_Os03g22050	Os03g0339900	1.176	0.00339322	0	CAMK_KIN1/SNF1/Nim1_like.16 - CAMK includes calcium/calmodulin depedent protein kinases, expressed
LOC_Os01g72100	Os01g0949500	1.175	0.00431452	0	OsCML10 - Calmodulin-related calcium sensor protein, expressed
LOC_Os03g10050	Os03g0196600	1.165	0.03986381	1	serine acetyltransferase protein, putative, expressed
LOC_Os03g57310	Os03g0787000	1.162	0.02439702	1	syntaxin, putative, expressed
LOC_Os07g45570	Os07g0650600	1.156	0.01420131	0	expressed protein
LOC_Os06g15910	Os06g0270200	1.154	0.00627963	0	potassium transporter, putative, expressed
LOC_Os03g18910	Os03g0301200	1.15	0.0016315	0	COBRA-like protein 7 precursor, putative, expressed
LOC_Os10g40960	Os10g0558900	1.15	0.03517774	0	oxidoreductase, 2OG-Fe oxygenase family protein, putative, expressed
LOC_Os11g01330	Os11g0104300	1.137	0.03048209	0	expressed protein
LOC_Os10g25010	Os10g0389000	1.135	0.01660545	0	OsCML8 - Calmodulin-related calcium sensor protein, expressed
LOC_Os01g28790	Os01g0384800	1.127	0.01286073	0	PRAS-rich protein, putative, expressed
LOC_Os05g04530	Os05g0135900	1.113	0.04801731	4	IF, putative, expressed
LOC_Os01g65140	Os01g0871900	1.109	0.01651539	1	peptide transporter PTR2, putative, expressed
LOC_Os04g52840	Os04g0619400	1.108	0.00941783	1	tyrosine protein kinase domain containing protein, putative, expressed
LOC_Os01g37910	Os01g0559600	1.102	0.03496467	1	vacuolar-processing enzyme precursor, putative, expressed
LOC_Os08g43400	Os08g0547500	1.079	0.0265012	2	kinesin motor domain containing protein, expressed
LOC_Os12g01360	Os12g0104300	1.071	0.04619293	0	expressed protein
LOC_Os11g18940	Os11g0294400	1.056	0.028786	2	WW domain containing protein, expressed
LOC_Os07g31720	Os07g0500300	1.054	0.00988348	0	GTPase activating protein, putative, expressed
LOC_Os10g02480	Os10g0113900	1.052	0.01994775	1	oxidoreductase, aldo/keto reductase family protein, putative, expressed
LOC_Os12g25200	Os12g0438600	1.05	0.01650437	0	chloride transporter, chloride channel family, putative, expressed
LOC_Os11g37230	Os11g0582100	1.043	0.01295742	1	zinc finger, C3HC4 type domain containing protein, expressed
LOC_Os09g18594	Os09g0355400	1.043	0.02991973	2	protein kinase domain containing protein, expressed
LOC_Os01g60640	Os01g0821600	1.041	0.0328579	0	WRKY21, expressed
LOC_Os01g72290	Os01g0952000	1.037	0.02089696	0	expressed protein
LOC_Os03g19720	Os03g0310800	1.036	0.04513043	1	EF hand family protein, putative, expressed
LOC_Os10g38580	Os10g0529300	1.029	0.04098777	5	glutathione S-transferase, putative, expressed
LOC_Os05g28740	Os05g0355400	1.01	0.02380711	2	universal stress protein domain containing protein, putative, expressed
LOC_Os09g10054	Os09g0272900	1.008	0.04098777	1	disease resistance protein RPS2, putative, expressed

LOC_Os02g54890	Os02g0791500	1.007	0.04008931	2	UDP-glucuronate 4-epimerase, putative, expressed
LOC_Os06g13190	Os06g0239200	1	0.04805946	3	expressed protein
LOC_Os04g35060	Os04g0429800	0.999	0.02572829	0	nicotinate phosphoribosyltransferase family domain containing protein, expressed
LOC_Os03g08970	Os03g0189100	0.988	0.0269411	0	expressed protein
LOC_Os09g13650	Os09g0307300	0.985	0.01746065	0	microtubule-associated protein, putative, expressed
LOC_Os12g22284	Os12g0411700	0.982	0.03968665	2	white-brown complex homolog protein 11, putative, expressed
LOC_Os03g57240	Os03g0786400	0.981	0.01425244	1	ZOS3-19 - C2H2 zinc finger protein, expressed
LOC_Os10g07229	Os10g0159800	0.981	0.02824862	0	dehydrogenase, putative, expressed
LOC_Os12g08760	Os12g0189300	0.979	0.01611781	0	carboxyvinyl-carboxyphosphonate phosphorylmutase, putative, expressed
LOC_Os07g18120	Os07g0281700	0.975	0.02434056	13	aldehyde oxidase, putative, expressed
LOC_Os03g52880	Os03g0738900	0.96	0.02463551	1	BTBN9 - Bric-a-Brac, Tramtrack, Broad Complex BTB domain with non-phototropic hypocotyl 3 NPH3 and coiled-coil domains, expressed
LOC_Os06g24730	Os06g0354700	0.96	0.02700102	0	hydrolase, alpha/beta fold family domain containing protein, expressed
LOC_Os11g02240	Os11g0113700	0.955	0.04338239	2	CAMK_KIN1/SNF1/Nim1_like.4 - CAMK includes calcium/calmodulin depedent protein kinases, expressed
LOC_Os01g60420	Os01g0819700	0.953	0.03986381	0	expressed protein
LOC_Os05g41590	Os05g0495700	0.951	0.0233264	0	glycerol-3-phosphate dehydrogenase, putative, expressed
LOC_Os09g32510	Os09g0501600	0.95	0.04506614	2	BHLH transcription factor, putative, expressed
LOC_Os05g50800	Os05g0585400	0.948	0.02031657	2	expressed protein
LOC_Os03g19020	Os03g0302200	0.942	0.03839376	0	PHD-finger family protein, expressed
LOC_Os03g05620	Os03g0150600	0.935	0.04475803	0	inorganic phosphate transporter, putative, expressed
LOC_Os04g48010	Os04g0568300	0.924	0.02620359	0	WD-40 repeat family protein, putative, expressed
LOC_Os08g41040	Os08g0521800	0.914	0.0490105	4	expressed protein
LOC_Os07g07230	Os07g0166700	0.888	0.03036127	0	protein kinase, putative, expressed
LOC_Os01g54400	Os01g0747800	0.883	0.03523885	0	VQ domain containing protein, putative, expressed
LOC_Os02g08530	Os02g0182600	0.872	0.03739513	0	protein kinase family protein, putative, expressed
LOC_Os06g35160	Os06g0543400	0.862	0.0497077	1	CAMK_KIN1/SNF1/Nim1_like.26 - CAMK includes calcium/calmodulin depedent protein kinases, expressed
LOC_Os03g20720	Os03g0323500	0.843	0.04307111	0	GTPase-activating protein, putative, expressed
Down-regulated DEG					
LOC_Os04g24530	Os04g0310800	-18.624	1.767E-89	1	AMP-binding domain containing protein, expressed
LOC_Os07g37090	Os07g0556800	-16.449	5.7028E-76	1	ribosome inactivating protein, putative, expressed
LOC_Os03g47896	Os03g0683300	-16.302	1.0174E-71	0	ribosome inactivating protein, putative, expressed
LOC_Os03g47910	Os03g0683500	-16.007	8.5012E-70	3	expressed protein
LOC_Os11g37280	Os11g0582500	-15.678	7.8031E-87	1	LTPL68 - Protease inhibitor/seed storage/LTP family protein precursor, expressed
LOC_Os08g43240	Os08g0545800	-15.516	5.3271E-62	6	LTPL97 - Protease inhibitor/seed storage/LTP family protein precursor, expressed

LOC_Os05g05920	Os05g0151100	-15.281	3.451E-53	1	desiccation-related protein PCC13-62 precursor, putative, expressed
LOC_Os11g10910	Os11g0215400	-14.92	7.8012E-42	0	chloroplast nucleoid DNA-binding protein, putative, expressed
LOC_Os08g27210	Os08g0360700	-14.696	2.169E-41	3	LTPL3 - Protease inhibitor/seed storage/LTP family protein precursor, putative, expressed
LOC_Os03g11350	Os03g0212000	-13.997	3.5993E-35	3	UDP-glucoronosyl and UDP-glucosyl transferase, putative, expressed
LOC_Os05g30580	Os05g0368700	-13.914	1.4719E-31	0	OsSub46 - Putative Subtilisin homologue, expressed
LOC_Os06g42860	Os06g0635300	-13.882	2.3373E-20	3	triacylglycerol lipase precursor, putative, expressed
LOC_Os01g12020	Os01g0219500	-13.852	5.307E-72	6	LTPL18 - Protease inhibitor/seed storage/LTP family protein precursor, expressed
LOC_Os03g46110	Os03g0663900	-13.824	3.067E-44	0	LTPL94 - Protease inhibitor/seed storage/LTP family protein precursor, expressed
LOC_Os03g43790	Os03g0639100	-13.808	7.3969E-35	0	expressed protein
LOC_Os12g41640	Os12g0610100	-13.681	5.1492E-33	4	expressed protein
LOC_Os03g48200	Os03g0687400	-13.637	7.4075E-32	20	ribosome inactivating protein, putative, expressed
LOC_Os01g39710	Os01g0579000	-13.613	3.6478E-30	1	expressed protein
LOC_Os06g40550	Os06g0607700	-13.601	4.2158E-69	0	ABC-2 type transporter domain containing protein, expressed
LOC_Os08g41950	Os08g0531700	-13.269	4.5246E-82	0	OsMADS7 - MADS-box family gene with MIKCc type-box, expressed
LOC_Os11g10920	Os11g0215600	-13.242	1.5129E-28	0	carboxyl-terminal proteinase, putative, expressed
LOC_Os11g14900	Os11g0255300	-13.156	8.344E-27	1	thiol protease SEN102 precursor, putative, expressed
LOC_Os03g47890	Os03g0683200	-13.129	4.5521E-27	0	transposon protein, putative, unclassified, expressed
LOC_Os05g33150	Os05g0399700	-12.964	0.00937766	0	CHIT6 - Chitinase family protein precursor, expressed
LOC_Os03g11760	Os03g0216800	-12.924	5.8986E-25	8	polygalacturonase, putative, expressed
LOC_Os06g12350	Os06g0228800	-12.903	1.4604E-16	2	amino acid transporter, putative, expressed
LOC_Os04g48400	Os04g0573100	-12.779	7.183E-86	0	HOTHEAD precursor, putative, expressed
LOC_Os01g70440	Os01g0929600	-12.741	1.0736E-11	5	LEML1 - Anther-specific LEM1 family protein precursor, expressed
LOC_Os03g48220	Os03g0687700	-12.69	1.8118E-68	0	expressed protein
LOC_Os01g41170	Os01g0594900	-12.663	3.6827E-56	1	THION27 - Plant thionin family protein precursor, expressed
LOC_Os09g32370	Os09g0499500	-12.642	3.0565E-21	4	gp176, putative, expressed
LOC_Os03g52160	Os03g0731800	-12.629	1.0018E-20	12	regulatory protein, putative, expressed
LOC_Os01g03670	Os01g0127500	-12.595	2.9783E-82	0	dihydroflavonol-4-reductase, putative, expressed
LOC_Os05g44120	Os05g0517400	-12.582	3.06E-76	5	expressed protein
LOC_Os12g13930	Os12g0242700	-12.576	1.5326E-63	0	3-oxoacyl-reductase, chloroplast precursor, putative, expressed
LOC_Os05g49790	Os05g0573600	-12.459	4.4862E-19	0	DUF538 domain containing protein, putative, expressed
LOC_Os05g49830	Os05g0574000	-12.374	1.2443E-16	1	lipase class 3 family protein, putative, expressed
LOC_Os07g43230	Os07g0625300	-12.353	3.9909E-17	0	skp1 family, tetramerisation domain containing protein, expressed
LOC_Os09g30320	Os09g0480900	-12.342	1.2209E-63	7	BURP domain containing protein, expressed
LOC_Os05g41870	Os05g0498200	-12.233	5.2248E-16	3	glycine-rich cell wall protein, putative, expressed
LOC_Os07g04580	Os07g0138400	-12.115	7.2161E-13	0	zinc finger C-x8-C-x5-C-x3-H type family protein, expressed

LOC_Os05g46190	Os05g0539300	-12.1	7.3837E-15	3	expressed protein
LOC_Os07g09110	Os07g0189400	-12.072	1.6455E-12	0	OsFBX219 - F-box domain containing protein, expressed
LOC_Os02g18750	Os02g0289000	-12.029	8.5282E-14	0	expressed protein
LOC_Os03g09150	Os03g0191700	-11.882	3.0482E-13	0	pumilio-family RNA binding repeat domain containing protein, expressed
LOC_Os07g07420	Os07g0169700	-11.88	4.7395E-12	0	gibberellin 20 oxidase 1-B, putative, expressed
LOC_Os06g44660	Os06g0656800	-11.733	5.9481E-12	9	fasciclin-like arabinogalactan precursor protein, putative, expressed
LOC_Os09g37300	Os09g0545000	-11.55	1.4491E-10	9	transporter, monovalent cation:proton antiporter-2 family, putative, expressed
LOC_Os07g36330	Os07g0547600	-11.537	1.6554E-10	1	OsFBX245 - F-box domain containing protein, expressed
LOC_Os05g51090	Os05g0588500	-11.534	6.5885E-10	2	nodulin MtN3 family protein, putative, expressed
LOC_Os08g06690	Os08g0163900	-11.426	5.8631E-09	1	expressed protein
LOC_Os03g07140	Os03g0167600	-11.415	1.3622E-64	1	male sterility protein, putative, expressed
LOC_Os05g40550	Os05g0484000	-11.368	5.3167E-09	6	expressed protein
LOC_Os09g16010	Os09g0329000	-11.317	3.2918E-31	0	BURP domain containing protein, expressed
LOC_Os03g47460	Os03g0677900	-11.291	3.3132E-37	1	ribosome inactivating protein, putative, expressed
LOC_Os03g24300	Os03g0357500	-11.131	1.1445E-53	0	LTPL1 - Protease inhibitor/seed storage/LTP family protein precursor, expressed
LOC_Os10g34360	Os10g0484800	-11.083	0.00158992	0	stilbene synthase, putative, expressed
LOC_Os01g06400	Os01g0157200	-11.057	6.5179E-08	0	expressed protein
LOC_Os08g31870	Os08g0413000	-11.013	1.4352E-06	1	cell division cycle protein 48, putative, expressed
LOC_Os05g40570	None	-11.011	7.0987E-08	0	expressed protein
LOC_Os05g40500	None	-11.009	5.7914E-08	0	F-box domain containing protein, expressed
LOC_Os07g08530	Os07g0183200	-11.003	1.8794E-07	1	auxin response factor, putative, expressed
LOC_Os07g22850	Os07g0411300	-10.964	3.3214E-59	0	chalcone and stilbene synthases, putative, expressed
LOC_Os05g40620	Os05g0484700	-10.958	2.3575E-07	0	expressed protein
LOC_Os09g32948	Os09g0507200	-10.934	1.1289E-61	1	OsMADS8 - MADS-box family gene with MIKCc type-box, expressed
LOC_Os08g06890	Os08g0166100	-10.877	5.0884E-07	4	expressed protein
LOC_Os03g56490	Os03g0776300	-10.76	1.8324E-06	3	DNA topoisomerase IV subunit A, putative, expressed
LOC_Os05g46050	Os05g0537700	-10.687	1.1059E-05	1	OsFBDUF26 - F-box and DUF domain containing protein, expressed
LOC_Os01g41280	None	-10.682	9.5688E-08	1	OsFBD4 - F-box and FBD domain containing protein, expressed
LOC_Os03g24130	Os03g0356540	-10.671	1.2416E-31	10	CXXC1 - Cysteine-rich protein with CXXC and CXXXC motifs precursor, putative, expressed
LOC_Os11g10890	Os11g0215300	-10.616	3.819E-06	4	expressed protein
LOC_Os06g12330	Os06g0228600	-10.605	4.2642E-23	0	amino acid transporter, putative, expressed
LOC_Os10g27390	Os10g0414000	-10.603	3.5298E-06	1	no apical meristem protein, putative, expressed
LOC_Os02g36070	Os02g0569400	-10.379	4.2807E-21	0	cytochrome P450, putative, expressed
LOC_Os01g52470	Os01g0723000	-10.335	4.6075E-08	12	elongation factor, putative, expressed
LOC_Os02g28970	Os02g0491300	-10.288	4.1921E-34	6	expressed protein

LOC_Os04g42260	Os04g0500900	-10.175	3.413E-31	7	protein phosphatase 2C, putative, expressed
LOC_Os03g48235	Os03g0688000	-10.109	4.7588E-35	0	expressed protein
LOC_Os08g38810	Os08g0496800	-10.043	0.00161377	1	BURP domain containing protein, expressed
LOC_Os02g01190	Os02g0101900	-10.001	4.4602E-05	1	POEI25 - Pollen Ole e I allergen and extensin family protein precursor, expressed
LOC_Os05g41270	Os05g0491900	-9.91	4.2396E-16	7	CAMK_CAMK_like.4 - CAMK includes calcium/calmodulin depedent protein kinases, expressed
LOC_Os10g38050	Os10g0524500	-9.789	6.611E-14	0	HOTHEAD precursor, putative, expressed
LOC_Os03g56890	Os03g0781600	-9.777	3.067E-44	0	expressed protein
LOC_Os07g46210	Os07g0655800	-9.707	1.8212E-05	0	LTPL2 - Protease inhibitor/seed storage/LTP family protein precursor, expressed
LOC_Os10g39980	Os10g0547500	-9.678	7.5903E-05	0	expressed protein
LOC_Os04g33150	Os04g0404400	-9.624	4.7736E-66	1	desiccation-related protein PCC13-62 precursor, putative, expressed
LOC_Os09g32992	Os09g0508250	-9.456	1.2068E-11	0	expressed protein
LOC_Os09g35700	Os09g0525500	-9.428	1.7068E-07	0	LTPL45 - Protease inhibitor/seed storage/LTP family protein precursor, expressed
LOC_Os06g40880	Os06g0611400	-9.136	8.3342E-09	2	polygalacturonase, putative, expressed
LOC_Os01g18870	Os01g0293100	-9.112	3.0986E-33	0	helix-loop-helix DNA-binding domain containing protein, expressed
LOC_Os03g59380	Os03g0808500	-9.106	2.0659E-37	1	LTPL28 - Protease inhibitor/seed storage/LTP family protein precursor, expressed
LOC_Os05g06480	Os05g0156900	-9.033	4.7365E-07	0	inorganic H+ pyrophosphatase, putative, expressed
LOC_Os10g42060	Os10g0570700	-9.028	0.01526344	1	ribosome inactivating protein, putative, expressed
LOC_Os12g42330	Os12g0618000	-8.994	3.3453E-09	1	sublingual apomucin, putative, expressed
LOC_Os02g45530	Os02g0678300	-8.935	1.0618E-05	0	HOTHEAD precursor, putative, expressed
LOC_Os04g50850	Os04g0596200	-8.809	1.18E-11	12	aspartyl protease family, putative, expressed
LOC_Os05g11414	Os05g0203800	-8.758	1.3648E-26	0	OsMADS58 - MADS-box family gene with MIKCc type-box, expressed
LOC_Os08g06700	Os08g0164000	-8.733	1.1745E-13	0	expressed protein
LOC_Os03g07250	Os03g0168600	-8.733	0.00011828	0	cytochrome P450, putative, expressed
LOC_Os12g23670	Os12g0578200	-8.715	4.4332E-07	0	expressed protein
LOC_Os12g38900	Os12g0578200	-8.603	8.5E-35	5	chorismate mutase, chloroplast precursor, putative, expressed
LOC_Os11g02165	None	-8.599	2.8594E-27	0	LTPL57 - Protease inhibitor/seed storage/LTP family protein precursor, expressed
LOC_Os06g40020	Os06g0602400	-8.534	1.662E-23	0	DEAD-box ATP-dependent RNA helicase 52A, putative, expressed
LOC_Os07g43270	None	-8.503	4.7026E-06	2	SKP1-like protein 1B, putative, expressed
LOC_Os04g49150	Os04g0580700	-8.441	1.347E-36	2	OsMADS17 - MADS-box family gene with MIKCc type-box, expressed
LOC_Os09g14550	Os09g0314500	-8.427	1.3943E-08	1	RNA recognition motif containing protein, putative, expressed
LOC_Os07g39640	Os07g0585200	-8.393	5.6152E-20	2	LTPL64 - Protease inhibitor/seed storage/LTP family protein precursor, expressed
LOC_Os07g08650	Os07g0184200	-8.382	3.6427E-10	0	expressed protein
LOC_Os07g09020	Os07g0188000	-8.351	2.2764E-17	1	argonaute, putative, expressed
LOC_Os05g35400	Os05g0428600	-8.35	1.3218E-05	3	DnaK family protein, putative, expressed
LOC_Os03g02700	Os03g0118700	-8.308	1.1191E-21	3	HOTHEAD precursor, putative, expressed

LOC_Os02g33840	Os02g0543200	-8.284	1.1842E-05	2	OsFBX52 - F-box domain containing protein, expressed
LOC_Os05g30390	Os05g0366800	-8.276	1.3354E-05	0	Os5bglu23 - beta-glucosidase homologue, similar to G. max isohydroxyurate hydrolase, likely pseudogene in japonica, expressed
LOC_Os05g34940	Os05g0423400	-8.163	1.6984E-29	4	OsMADS4 - MADS-box family gene with MIKCc type-box, expressed
LOC_Os10g24050	Os10g0382100	-8.161	1.2482E-07	0	ribosome inactivating protein, putative, expressed
LOC_Os12g02105	Os12g0112401	-8.068	2.3041E-24	0	LTPL58 - Protease inhibitor/seed storage/LTP family protein precursor, expressed
LOC_Os01g03170	Os01g0121900	-8.059	5.7485E-13	0	seven in absentia protein family protein, expressed
LOC_Os04g37570	Os04g0448500	-8.057	5.3269E-22	0	aspartic proteinase nepenthesin precursor, putative, expressed
LOC_Os04g58120	Os04g0677600	-7.989	0.00122699	0	CRP2 - Cysteine-rich family protein precursor, expressed
LOC_Os11g47350	Os11g0696400	-7.988	0.00947488	0	beta-D-xylosidase, putative, expressed
LOC_Os03g01700	Os03g0107300	-7.931	3.274E-12	0	expressed protein
LOC_Os03g51690	Os03g0727000	-7.885	0.00108815	5	Homeobox domain containing protein, expressed
LOC_Os02g39000	Os02g0602000	-7.851	7.7516E-19	1	remorin C-terminal domain containing protein, putative, expressed
LOC_Os03g47650	Os03g0680200	-7.795	1.8268E-08	0	ankyrin, putative, expressed
LOC_Os06g13040	Os06g0237400	-7.773	4.6416E-18	0	glycosyl hydrolases family 16, putative, expressed
LOC_Os01g49650	Os01g0691300	-7.681	0.00044385	0	LTPL150 - Protease inhibitor/seed storage/LTP family protein precursor, expressed
LOC_Os06g40770	Os06g0610100	-7.672	4.6011E-06	12	expressed protein
LOC_Os01g38520	Os01g0566000	-7.627	2.13E-10	0	expressed protein
LOC_Os03g11614	Os03g0215400	-7.533	7.1472E-49	3	OsMADS1 - MADS-box family gene with MIKCc type-box, expressed
LOC_Os01g14340	Os01g0245901	-7.464	0.0001342	4	expressed protein
LOC_Os12g10540	Os12g0207000	-7.446	2.0709E-09	0	OsMADS13 - MADS-box family gene with MIKCc type-box, expressed
LOC_Os06g06750	Os06g0162800	-7.433	8.7793E-25	0	OsMADS5 - MADS-box family gene with MIKCc type-box, expressed
LOC_Os11g47670	Os11g0703000	-7.415	0.01341637	0	thaumatin family domain containing protein, expressed
LOC_Os01g34920	Os01g0533400	-7.407	6.9727E-08	0	beta-galactosidase precursor, putative, expressed
LOC_Os03g15340	Os03g0259100	-7.365	2.6373E-26	0	plastocyanin-like domain containing protein, putative, expressed
LOC_Os02g57520	Os02g0820800	-7.287	9.8981E-12	0	DNA binding protein, putative, expressed
LOC_Os01g11130	Os01g0209500	-7.27	1.6609E-14	0	RNA recognition motif containing protein, putative, expressed
LOC_Os03g62830	Os03g0845300	-7.21	8.3392E-15	1	nuclear antigen, putative, expressed
LOC_Os03g45120	Os03g0653900	-7.192	1.1806E-08	2	ribosome inactivating protein, putative, expressed
LOC_Os07g43250	Os07g0625500	-7.137	0.01172615	2	SKP1-like protein 1B, putative, expressed
LOC_Os07g04330	Os07g0136300	-7.126	2.733E-12	0	expressed protein
LOC_Os08g40440	Os08g0515900	-7.101	1.6347E-05	0	dihydroflavonol-4-reductase, putative, expressed
LOC_Os01g38680	Os01g0567600	-7.021	4.5773E-05	0	transporter family protein, putative, expressed
LOC_Os01g55200	Os01g0756700	-6.988	0.0009219	0	potassium channel KAT1, putative, expressed
LOC_Os05g30400	Os05g0366900	-6.963	6.4682E-10	0	expressed protein
LOC_Os02g45770	Os02g0682200	-6.941	1.6708E-29	0	OsMADS6 - MADS-box family gene with MIKCc type-box, expressed

LOC_Os08g12410	Os08g0220400	-6.939	6.869E-10	0	pectinesterase, putative, expressed
LOC_Os07g22680	Os07g0409500	-6.918	6.5814E-07	1	SKP1-like protein 1B, putative, expressed
LOC_Os09g38680	Os09g0559700	-6.889	0.00145104	0	expressed protein
LOC_Os02g45380	Os02g0676400	-6.878	4.3926E-12	0	MATE domain containing protein, expressed
LOC_Os01g40170	Os01g0583900	-6.861	1.9877E-11	0	translation initiation factor, putative, expressed
LOC_Os03g18480	Os03g0296000	-6.821	0.00416833	1	MYB family transcription factor, putative, expressed
LOC_Os01g14030	Os01g0242400	-6.81	2.6349E-07	0	DUF260 domain containing protein, putative, expressed
LOC_Os03g25530	Os03g0371600	-6.765	3.7355E-06	2	expressed protein
LOC_Os01g19694	Os01g0302500	-6.755	0.0004279	3	Homeobox domain containing protein, expressed
LOC_Os06g48740	None	-6.748	0.00083828	0	expressed protein
LOC_Os11g31420	Os11g0513000	-6.72	4.8626E-08	2	expressed protein
LOC_Os08g04140	Os08g0135500	-6.598	7.383E-05	0	X8 domain containing protein, expressed
LOC_Os06g49830	Os06g0712600	-6.473	0.00053458	3	SHI, putative, expressed
LOC_Os01g03840	Os01g0129200	-6.46	7.043E-10	0	ZOS1-02 - C2H2 zinc finger protein, expressed
LOC_Os08g36340	Os08g0466200	-6.392	0.00022367	0	potassium transporter, putative, expressed
LOC_Os08g43290	Os08g0546300	-6.378	0.00010959	2	LTPL44 - Protease inhibitor/seed storage/LTP family protein precursor, expressed
LOC_Os01g52880	Os01g0729400	-6.375	0.00200756	0	leucine-rich repeat family protein, putative, expressed
LOC_Os01g22590	Os01g0329300	-6.367	0.00031584	0	retrotransposon protein, putative, unclassified, expressed
LOC_Os01g56530	Os01g0772100	-6.354	2.107E-05	1	DUF260 domain containing protein, putative, expressed
LOC_Os05g13620	Os05g0223200	-6.348	2.5827E-12	11	RNA recognition motif containing protein, putative, expressed
LOC_Os09g13930	Os09g0309600	-6.322	0.00739745	0	LTPL4 - Protease inhibitor/seed storage/LTP family protein precursor, expressed
LOC_Os05g37060	Os05g0442400	-6.258	1.5893E-07	2	MYB family transcription factor, putative, expressed
LOC_Os04g48210	Os04g0570600	-6.241	0.00017247	0	cytochrome P450, putative, expressed
LOC_Os10g27360	Os10g0413700	-6.218	0.01517517	4	no apical meristem protein, putative, expressed
LOC_Os02g38260	Os02g0596200	-6.191	0.00053458	1	glycosyl hydrolase family 5 protein, putative, expressed
LOC_Os11g31690	Os11g0517200	-6.177	6.0942E-08	0	expressed protein
LOC_Os01g46984	Os01g0659400	-6.109	6.7052E-06	0	expressed protein
LOC_Os04g58590	Os04g0682400	-6.108	7.615E-13	1	RNA recognition motif containing protein, putative, expressed
LOC_Os05g46160	None	-6.088	0.00025951	3	OsFBDUF29 - F-box and DUF domain containing protein, expressed
LOC_Os10g33250	Os10g0471100	-6.072	3.0685E-08	0	WAX2, putative, expressed
LOC_Os06g37680	Os06g0574900	-6.033	4.6179E-05	0	expressed protein
LOC_Os01g41140	Os01g0594500	-5.995	0.01172683	3	THION18 - Plant thionin family protein precursor, expressed
LOC_Os01g03190	Os01g0122200	-5.985	3.1744E-07	3	expressed protein
LOC_Os04g13150	Os04g0208400	-5.938	1.4287E-09	5	OsFBX125 - F-box domain containing protein, expressed
LOC_Os10g32270	Os10g0460500	-5.883	2.3884E-06	0	expressed protein

LOC_Os11g31470	Os11g0513900	-5.874	8.0587E-30	1	expressed protein
LOC_Os12g32450	None	-5.863	0.00241651	6	expressed protein
LOC_Os10g26340	Os10g0403000	-5.841	0.03614384	0	cytochrome P450, putative, expressed
LOC_Os07g03770	Os07g0129700	-5.838	4.4078E-08	1	Homeobox domain containing protein, expressed
LOC_Os08g03682	Os08g0131100	-5.833	0.00014879	0	cytochrome P450, putative, expressed
LOC_Os11g06770	Os11g0168500	-5.818	0.00841307	2	ethylene-responsive transcription factor ERF110, putative, expressed
LOC_Os06g38980	Os06g0589700	-5.813	1.3218E-05	2	polyadenylate-binding protein, putative, expressed
LOC_Os06g12320	Os06g0228500	-5.788	1.3424E-11	0	transmembrane amino acid transporter protein, putative, expressed
LOC_Os02g02820	Os02g0120500	-5.783	0.00011834	0	helix-loop-helix DNA-binding domain containing protein, expressed
LOC_Os07g04350	Os07g0136500	-5.73	4.4765E-14	3	expressed protein
LOC_Os03g11370	Os03g0212300	-5.71	3.9132E-06	0	B3 DNA binding domain containing protein, expressed
LOC_Os10g40530	Os10g0552800	-5.686	2.9902E-33	0	LTPL146 - Protease inhibitor/seed storage/LTP family protein precursor, expressed
LOC_Os09g01670	Os09g0103900	-5.676	4.2406E-07	4	peptidyl-prolyl cis-trans isomerase, FKBP-type, putative, expressed
LOC_Os03g53350	Os03g0745100	-5.658	3.1814E-08	0	anthocyanin 3-O-beta-glucosyltransferase, putative, expressed
LOC_Os07g25740	Os07g0439100	-5.635	9.9987E-06	3	expressed protein
LOC_Os08g45150	Os08g0565900	-5.631	0.00361761	0	GDSL-like lipase/acylhydrolase, putative, expressed
LOC_Os04g57530	Os04g0670900	-5.607	0.01297359	0	transcription factor, putative, expressed
LOC_Os02g42950	Os02g0643200	-5.565	2.1007E-05	0	YABBY domain containing protein, putative, expressed
LOC_Os03g51710	Os03g0727200	-5.549	5.4207E-07	0	homeobox protein knotted-1, putative, expressed
LOC_Os07g35880	Os07g0543100	-5.539	0.0091958	0	beta-amylase, putative, expressed
LOC_Os03g54170	Os03g0753100	-5.475	5.2025E-12	2	OsMADS34 - MADS-box family gene with MIKCc type-box, expressed
LOC_Os04g32960	Os04g0402200	-5.468	2.0198E-09	0	TUDOR protein with multiple SNc domains, putative, expressed
LOC_Os04g44600	Os04g0528200	-5.416	1.3611E-05	1	CRP1 - Cysteine-rich family protein precursor, expressed
LOC_Os11g31500	Os11g0514100	-5.4	2.4654E-10	3	ATP binding protein, putative, expressed
LOC_Os10g05790	Os10g0148700	-5.385	6.8445E-11	0	POEI4 - Pollen Ole e I allergen and extensin family protein precursor, expressed
LOC_Os07g42490	Os07g0616800	-5.378	7.7327E-08	0	sucrose synthase, putative, expressed
LOC_Os03g63240	Os03g0849500	-5.353	4.8037E-08	0	disease resistance protein, putative, expressed
LOC_Os01g55120	Os01g0756000	-5.337	2.815E-06	0	expressed protein
LOC_Os03g15710	Os03g0263600	-5.309	0.00181288	0	strictosidine synthase, putative, expressed
LOC_Os04g59260	Os04g0689000	-5.291	0.01135213	4	peroxidase precursor, putative, expressed
LOC_Os05g15300	Os05g0242600	-5.289	0.00876963	1	expressed protein
LOC_Os06g44240	Os06g0652100	-5.281	0.00129141	2	gp176, putative, expressed
LOC_Os10g42850	Os10g0579400	-5.263	0.00011003	0	WRKY2, expressed
LOC_Os01g02190	Os01g0112400	-5.249	0.00356889	0	aquaporin protein, putative, expressed
LOC_Os11g31450	Os11g0513700	-5.196	5.5724E-07	2	expressed protein

LOC_Os12g27102	Os12g0456700	-5.16	0.00846921	7	glycerophosphoryl diester phosphodiesterase family protein, putative, expressed
LOC_Os01g14430	Os01g0246601	-5.126	0.0165313	0	DNA binding protein, putative, expressed
LOC_Os01g03680	Os01g0127600	-5.118	1.5235E-27	0	BBTI8 - Bowman-Birk type bran trypsin inhibitor precursor, expressed
LOC_Os04g45960	Os04g0543700	-5.102	0.00014023	1	OsSub42 - Putative Subtilisin homologue, expressed
LOC_Os12g36210	Os12g0548401	-5.058	5.5092E-12	0	inhibitor I family protein, putative, expressed
LOC_Os02g01326	Os02g0103600	-5.036	7.8893E-09	1	auxin-induced protein 5NG4, putative, expressed
LOC_Os01g11430	Os01g0212500	-5.033	0.00024536	0	expressed protein
LOC_Os03g38210	Os03g0578900	-5.011	3.6945E-07	1	MYB family transcription factor, putative, expressed
LOC_Os03g52320	Os03g0733600	-5.007	0.03052576	1	GRF-interacting factor 1, putative, expressed
LOC_Os12g18960	Os12g0288000	-4.976	1.4635E-06	0	integral membrane protein DUF6 containing protein, expressed
LOC_Os07g43260	Os07g0625600	-4.94	0.00089979	5	SKP1-like protein 1B, putative, expressed
LOC_Os05g07120	Os05g0163900	-4.926	0.0074275	1	basic helix-loop-helix, putative, expressed
LOC_Os04g45330	Os04g0536300	-4.908	0.03179605	0	YABBY domain containing protein, putative, expressed
LOC_Os03g61290	Os03g0828600	-4.884	0.02240627	3	ATCHX, putative, expressed
LOC_Os04g37680	None	-4.864	1.1842E-05	8	alpha/beta hydrolase fold, putative, expressed
LOC_Os05g45430	Os05g0530701	-4.862	2.5037E-05	0	TOO MANY MOUTHS precursor, putative, expressed
LOC_Os03g31430	Os03g0428200	-4.844	5.0191E-05	1	terpene synthase, N-terminal domain containing protein, expressed
LOC_Os02g36700	Os02g0576600	-4.826	0.02882923	0	sucrose transporter BoSUT1, putative, expressed
LOC_Os12g07840	Os12g0178300	-4.813	3.9527E-06	0	dehydration response related protein, putative, expressed
LOC_Os01g66030	Os01g0883100	-4.73	2.2088E-08	0	OsMADS2 - MADS-box family gene with MIKCc type-box, expressed
LOC_Os04g44740	Os04g0529700	-4.725	5.413E-11	0	glycosyltransferase sugar-binding region containing DXD motif, putative, expressed
LOC_Os01g13120	Os01g0232000	-4.713	1.0806E-05	0	aquaporin protein, putative, expressed
LOC_Os08g05820	Os08g0154250	-4.711	7.6783E-05	1	monocopper oxidase, putative, expressed
LOC_Os04g46444	Os04g0550300	-4.696	0.00910377	1	expressed protein
LOC_Os05g06130	Os05g0153200	-4.693	7.4138E-09	0	transcription factor X1, putative, expressed
LOC_Os08g39250	Os08g0502000	-4.655	5.999E-14	0	EDM2, putative, expressed
LOC_Os06g51370	Os06g0730000	-4.612	0.01374276	0	OsSCP38 - Putative Serine Carboxypeptidase homologue, expressed
LOC_Os01g65590	Os01g0877400	-4.597	0.00040884	1	galactosyltransferase, putative, expressed
LOC_Os12g39310	Os12g0582700	-4.59	4.6127E-33	0	cytochrome P450, putative, expressed
LOC_Os07g04430	Os07g0137000	-4.585	0.00218102	1	expressed protein
LOC_Os02g51730	Os02g0753500	-4.551	2.9775E-21	0	dnaJ homolog subfamily C member 7, putative, expressed
LOC_Os03g37140	Os03g0569000	-4.544	1.5496E-06	0	expressed protein
LOC_Os03g51180	Os03g0721700	-4.527	0.03202558	0	expressed protein
LOC_Os12g32290	Os12g0507700	-4.518	0.00015386	0	expressed protein
LOC_Os02g07650	Os02g0172800	-4.512	0.02957989	3	zinc-binding protein, putative, expressed

LOC_Os07g35610	Os07g0540366	-4.492	3.3657E-13	0	kelch repeat-containing protein, putative, expressed
LOC_Os04g51070	Os04g0599300	-4.488	0.04966906	0	helix-loop-helix DNA-binding domain containing protein, expressed
LOC_Os03g36560	Os03g0563600	-4.458	0.01749839	0	peroxidase precursor, putative, expressed
LOC_Os09g27620	Os09g0449000	-4.435	0.0002053	0	PHD-finger domain containing protein, putative, expressed
LOC_Os04g37520	Os04g0448000	-4.43	0.00332378	6	extracellular ligand-gated ion channel, putative, expressed
LOC_Os01g46270	Os01g0651500	-4.374	0.00016774	10	wax synthase isoform 3, putative, expressed
LOC_Os05g42150	Os05g0500900	-4.361	2.4964E-05	1	OsGH3.4 - Probable indole-3-acetic acid-amido synthetase, expressed
LOC_Os07g37920	Os07g0566500	-4.33	2.9272E-05	2	no apical meristem protein, putative, expressed
LOC_Os06g44300	Os06g0653000	-4.319	3.2119E-10	0	WAX2, putative, expressed
LOC_Os09g36160	Os09g0531600	-4.313	0.00143397	0	LRP1, putative, expressed
LOC_Os08g20730	Os08g0302000	-4.303	1.7365E-05	6	peroxidase precursor, putative, expressed
LOC_Os03g58850	Os03g0803100	-4.253	0.00632498	1	uncharacterized PE-PGRS family protein PE_PGRS3 precursor, putative, expressed
LOC_Os11g37000	Os11g0578500	-4.251	9.0183E-09	0	heat shock protein DnaJ, putative, expressed
LOC_Os09g15700	Os09g0326100	-4.243	8.2171E-09	1	receptor-like protein kinase 5 precursor, putative, expressed
LOC_Os11g06870	Os11g0169700	-4.242	0.01534375	3	glyoxal oxidase-related, putative, expressed
LOC_Os05g03884	Os05g0129700	-4.23	6.2015E-07	0	homeobox protein knotted-1-like 6, putative, expressed
LOC_Os03g50490	Os03g0712800	-4.229	3.6785E-11	0	glutamine synthetase, catalytic domain containing protein, expressed
LOC_Os12g06290	Os12g0160200	-4.204	3.2144E-05	0	expressed protein
LOC_Os09g28520	Os09g0459200	-4.202	0.00324026	1	expressed protein
LOC_Os04g05070	Os04g0137450	-4.201	0.01770198	0	expressed protein
LOC_Os01g34770	Os01g0532100	-4.198	2.1235E-05	3	tetratricopeptide repeat containing protein, putative, expressed
LOC_Os01g49240	Os01g0686300	-4.193	0.00588166	1	limonoid UDP-glucosyltransferase, putative, expressed
LOC_Os01g06560	Os01g0159000	-4.16	0.01089394	2	transcription factor HBP-1b, putative, expressed
LOC_Os05g05354	Os05g0146100	-4.153	8.2171E-09	0	expressed protein
LOC_Os08g15450	Os08g0254300	-4.149	3.7535E-07	1	nodulin, putative, expressed
LOC_Os06g14350	Os06g0254600	-4.112	0.03758326	0	caleosin related protein, putative, expressed
LOC_Os02g01980	Os02g0110000	-4.096	0.00043901	0	GDSL-like lipase/acylhydrolase, putative, expressed
LOC_Os10g36070	Os10g0504650	-4.072	0.00708718	0	LTPL155 - Protease inhibitor/seed storage/LTP family protein precursor, putative, expressed
LOC_Os01g08360	Os01g0178850	-4.048	9.2765E-09	0	hypothetical protein
LOC_Os07g01820	Os07g0108900	-4.047	6.8668E-26	0	OsMADS15 - MADS-box family gene with MIKCc type-box, expressed
LOC_Os01g02000	Os01g0110100	-4.033	4.6821E-05	0	phosphate transporter 1, putative, expressed
LOC_Os05g11610	Os05g0206000	-4.023	2.2307E-05	2	expressed protein
LOC_Os04g59120	Os04g0687800	-3.99	1.7552E-11	0	auxin-induced protein 5NG4, putative, expressed
LOC_Os03g21480	Os03g0332533	-3.974	0.03464877	2	HAD superfamily phosphatase, putative, expressed
LOC_Os08g20200	Os08g0298700	-3.95	0.00575724	0	male sterility protein, putative, expressed

LOC_Os01g64790	Os01g0868000	-3.942	3.5086E-30	6	AP2 domain containing protein, expressed
LOC_Os07g34070	Os07g0524900	-3.922	4.0607E-13	0	auxin-induced protein 5NG4, putative, expressed
LOC_Os05g03390	Os05g0124600	-3.888	3.7586E-13	2	expressed protein
LOC_Os05g13630	Os05g0223300	-3.867	7.0344E-11	0	RNA recognition motif containing protein, putative, expressed
LOC_Os05g07010	Os05g0162800	-3.851	1.8739E-10	2	myb-like DNA-binding domain containing protein, expressed
LOC_Os07g36370	Os07g0548100	-3.849	0.01642916	1	OsFBX247 - F-box domain containing protein, expressed
LOC_Os03g64260	Os03g0860100	-3.833	0.00152747	0	AP2 domain containing protein, expressed
LOC_Os01g03340	Os01g0124200	-3.82	1.7147E-13	0	BBTI4 - Bowman-Birk type bran trypsin inhibitor precursor, expressed
LOC_Os03g55960	Os03g0769500	-3.795	0.0160699	1	EF hand family protein, putative, expressed
LOC_Os08g02220	Os08g0114200	-3.793	0.00332757	0	endoglucanase, putative, expressed
LOC_Os01g66700	Os01g0891000	-3.776	0.01929903	2	beta-hexosaminidase precursor, putative, expressed
LOC_Os07g32406	Os07g0507500	-3.735	1.8857E-06	0	expressed protein
LOC_Os03g56200	Os03g0772900	-3.725	0.00088512	0	expressed protein
LOC_Os03g27850	Os03g0396000	-3.725	0.02009043	0	expressed protein
LOC_Os04g43580	Os04g0516200	-3.716	0.04008931	0	DUF640 domain containing protein, putative, expressed
LOC_Os02g47920	Os02g0709000	-3.715	1.4938E-05	0	ZOS2-16 - C2H2 zinc finger protein, expressed
LOC_Os01g52680	Os01g0726400	-3.696	0.00126177	6	OsMADS32 - MADS-box family gene with MIKCc type-box, expressed
LOC_Os10g40614	Os10g0554800	-3.686	0.00960331	0	LTPL147 - Protease inhibitor/seed storage/LTP family protein precursor, expressed
LOC_Os12g43840	Os12g0634900	-3.655	0.00257751	2	ankyrin repeat domain-containing protein, putative, expressed
LOC_Os03g51470	Os03g0724600	-3.648	2.7723E-06	0	expressed protein
LOC_Os09g30210	Os09g0479400	-3.636	0.01994775	1	expressed protein
LOC_Os06g42120	Os06g0626600	-3.615	0.02802446	0	sulfotransferase domain containing protein, expressed
LOC_Os07g46830	Os07g0663500	-3.608	4.5189E-10	0	sex determination protein tasselseed-2, putative, expressed
LOC_Os04g19750	None	-3.598	0.00695215	1	OsFBL11 - F-box domain and LRR containing protein, expressed
LOC_Os08g06210	Os08g0158600	-3.582	2.5192E-10	0	expressed protein
LOC_Os11g31680	Os11g0516800	-3.578	5.9985E-06	0	retrotransposon protein, putative, unclassified, expressed
LOC_Os07g41650	Os07g0607400	-3.558	0.04002267	0	pectinesterase, putative, expressed
LOC_Os03g59550	Os03g0809900	-3.551	1.9877E-11	0	RNA recognition motif containing protein, putative, expressed
LOC_Os08g41620	Os08g0527900	-3.523	0.00100852	0	ubiquitin carboxyl-terminal hydrolase family protein, expressed
LOC_Os07g09190	Os07g0190000	-3.522	7.4134E-29	0	transketolase, putative, expressed
LOC_Os01g51780	Os01g0715600	-3.517	0.00100852	1	auxin efflux carrier component, putative, expressed
LOC_Os01g52690	Os01g0726700	-3.487	7.9525E-12	0	retrotransposon protein, putative, unclassified, expressed
LOC_Os09g33940	Os09g0514500	-3.486	8.6289E-09	0	expressed protein
LOC_Os04g08824	Os04g0171600	-3.462	0.00146682	1	cytochrome P450, putative, expressed
LOC_Os02g41954	Os02g0630300	-3.445	2.4472E-09	2	gibberellin 2-beta-dioxygenase 7, putative, expressed

LOC_Os08g36910	Os08g0473900	-3.445	0.00024786	7	alpha-amylase precursor, putative, expressed
LOC_Os02g30790	Os02g0511800	-3.438	0.00671954	4	SAR DNA-binding protein-like, putative, expressed
LOC_Os07g02460	Os07g0115500	-3.407	1.0216E-13	1	expressed protein
LOC_Os02g06960	Os02g0165500	-3.39	0.0002053	4	expressed protein
LOC_Os04g01690	Os04g0107600	-3.384	0.00180693	0	pyridoxal-dependent decarboxylase protein, putative, expressed
LOC_Os11g31700	Os11g0517400	-3.38	5.0884E-07	0	expressed protein
LOC_Os04g39470	Os04g0470600	-3.379	0.02271809	2	MYB family transcription factor, putative, expressed
LOC_Os06g07220	Os06g0168700	-3.379	0.04107583	4	LTPL128 - Protease inhibitor/seed storage/LTP family protein precursor, expressed
LOC_Os03g26920	Os03g0386700	-3.373	9.6139E-05	0	OsSCP12 - Putative Serine Carboxypeptidase homologue, expressed
LOC_Os05g50920	Os05g0586500	-3.372	0.00632498	0	transmembrane amino acid transporter protein, putative, expressed
LOC_Os02g21750	Os02g0323000	-3.35	0.001347	0	MDR-like ABC transporter, putative, expressed
LOC_Os03g58600	Os03g0800200	-3.344	1.2926E-13	1	PAZ domain containing protein, putative, expressed
LOC_Os05g12040	Os05g0211100	-3.336	1.1451E-13	1	cytochrome P450 51, putative, expressed
LOC_Os09g30250	Os09g0479900	-3.336	3.0942E-07	0	OsSub58 - Putative Subtilisin homologue, expressed
LOC_Os01g50910	Os01g0705200	-3.321	0.01881434	1	late embryogenesis abundant protein, group 3, putative, expressed
LOC_Os09g36490	Os09g0535400	-3.319	0.00846921	0	expressed protein
LOC_Os11g07950	Os11g0182100	-3.31	2.9087E-09	1	expressed protein
LOC_Os05g08480	Os05g0177500	-3.308	0.00040884	0	cytokinin-O-glucosyltransferase 1, putative, expressed
LOC_Os09g32360	Os09g0499400	-3.253	9.719E-06	4	gp176, putative, expressed
LOC_Os01g56450	Os01g0771000	-3.246	5.7788E-08	0	DUF567 domain containing protein, putative, expressed
LOC_Os01g73770	Os01g0968800	-3.241	0.00015087	0	dehydration-responsive element-binding protein, putative, expressed
LOC_Os01g37460	Os01g0555100	-3.238	0.00083138	0	zinc finger family protein, putative, expressed
LOC_Os01g27160	Os01g0369200	-3.225	0.00026551	0	cullin, putative, expressed
LOC_Os02g38710	Os02g0599200	-3.225	0.00080472	0	protein phosphatase 2C containing protein, expressed
LOC_Os06g40180	Os06g0604300	-3.223	1.4177E-09	2	phospholipase D, putative, expressed
LOC_Os04g47450	Os04g0562500	-3.213	2.0172E-07	0	expressed protein
LOC_Os01g55600	Os01g0761400	-3.203	0.00067569	0	peptide transporter PTR2, putative, expressed
LOC_Os03g03810	Os03g0130300	-3.198	0.00013853	0	DEF8 - Defensin and Defensin-like DEFL family, expressed
LOC_Os05g44340	Os05g0519700	-3.194	0.00016137	1	heat shock protein 101, putative, expressed
LOC_Os01g54570	None	-3.193	9.3401E-06	6	DUF623 domain containing protein, expressed
LOC_Os06g46250	Os06g0675200	-3.192	0.00324026	2	expressed protein
LOC_Os12g13890	Os12g0242100	-3.191	0.02655908	0	retrotransposon protein, putative, unclassified, expressed
LOC_Os11g31705	Os11g0517800	-3.183	0.00036234	1	expressed protein
LOC_Os01g14630	Os01g0248701	-3.181	3.3032E-21	1	polyprenyl synthetase, putative, expressed
LOC_Os01g03320	Os01g0124000	-3.172	1.3903E-05	0	BBTI2 - Bowman-Birk type bran trypsin inhibitor precursor, expressed

LOC_Os03g18030	Os03g0289800	-3.161	1.7949E-09	0	leucoanthocyanidin dioxygenase, putative, expressed
LOC_Os12g32499	Os12g0509400	-3.161	0.01094919	0	expressed protein
LOC_Os06g49050	Os06g0704000	-3.157	1.1571E-05	0	hAT dimerisation domain containing protein, expressed
LOC_Os04g56690	Os04g0662400	-3.156	0.00015611	2	OsSAUR23 - Auxin-responsive SAUR gene family member, expressed
LOC_Os01g51770	Os01g0715500	-3.149	8.0321E-07	10	outer mitochondrial membrane porin, putative, expressed
LOC_Os01g07480	Os01g0169400	-3.144	0.00181288	0	DUF260 domain containing protein, putative, expressed
LOC_Os03g46400	Os03g0666600	-3.143	0.00209423	0	UDP-glucoronosyl and UDP-glucosyl transferase domain containing protein, expressed
LOC_Os01g04620	Os01g0138800	-3.14	1.6947E-07	0	transposon protein, putative, unclassified, expressed
LOC_Os11g37200	Os11g0581900	-3.128	6.8445E-11	0	transmembrane BAX inhibitor motif-containing protein, putative, expressed
LOC_Os07g40220	Os07g0591700	-3.127	7.7038E-08	3	expressed protein
LOC_Os01g16810	Os01g0274800	-3.126	1.1371E-06	0	MYB family transcription factor, putative, expressed
LOC_Os01g63930	Os01g0858200	-3.124	1.0018E-20	0	cytochrome P450, putative, expressed
LOC_Os01g20160	Os01g0307500	-3.122	0.0015044	1	OsHKT1;5 - Na+ transporter, expressed
LOC_Os08g44750	Os08g0561500	-3.118	0.00186439	0	auxin-induced protein 5NG4, putative, expressed
LOC_Os05g03320	Os05g0124000	-3.096	8.9796E-11	3	expressed protein
LOC_Os02g36830	None	-3.09	0.01060636	1	cytokinin-O-glucosyltransferase 2, putative, expressed
LOC_Os07g46920	None	-3.077	1.7992E-05	0	sex determination protein tasselseed-2, putative, expressed
LOC_Os09g34180	Os09g0517600	-3.073	0.01664145	2	formin, putative, expressed
LOC_Os11g02750	Os11g0120300	-3.063	1.0623E-10	1	DUF567 domain containing protein, putative, expressed
LOC_Os03g06610	Os03g0161600	-3.038	5.4519E-09	2	expressed protein
LOC_Os02g02930	Os02g0121700	-3.037	7.8688E-07	0	terpene synthase, putative, expressed
LOC_Os03g58290	Os03g0797300	-3.033	1.3916E-14	0	indole-3-glycerol phosphate lyase, chloroplast precursor, putative, expressed
LOC_Os01g11940	Os01g0218500	-3.03	4.3211E-11	1	osFTL1 FT-Like1 homologous to Flowering Locus T gene; contains Pfam profile PF01161: Phosphatidylethanolamine-binding protein, expressed
LOC_Os03g18010	Os03g0289300	-2.998	0.00214828	0	phospholipase C, putative, expressed
LOC_Os06g40170	Os06g0604200	-2.995	5.9454E-12	0	phospholipase D, putative, expressed
LOC_Os07g36610	Os07g0551600	-2.963	1.0839E-06	0	CSLF9 - cellulose synthase-like family F; beta1,3;1,4 glucan synthase, expressed
LOC_Os09g02660	Os09g0493500	-2.961	0.00029408	0	None
LOC_Os09g32020	Os09g0493500	-2.961	0.0008203	1	ubiquitin fusion degradation protein, putative, expressed
LOC_Os11g40210	Os11g0616300	-2.961	0.01328316	0	remorin C-terminal domain containing protein, putative, expressed
LOC_Os01g03360	Os01g0124401	-2.96	4.608E-05	0	BBTI5 - Bowman-Birk type bran trypsin inhibitor precursor, expressed
LOC_Os03g26650	Os03g0383900	-2.946	7.7705E-05	0	heavy metal-associated domain containing protein, expressed
LOC_Os01g72490	Os01g0954500	-2.936	4.2868E-08	1	LRP1, putative, expressed
LOC_Os06g13140	Os06g0238700	-2.928	3.3565E-06	0	WD domain, G-beta repeat domain containing protein, expressed

LOC_Os09g26260	Os09g0432300	-2.916	2.1447E-09	0	AAA-type ATPase family protein, putative, expressed
LOC_Os01g19370	Os01g0299100	-2.91	0.01409152	0	retrotransposon protein, putative, unclassified, expressed
LOC_Os09g39020	Os09g0563700	-2.899	1.2157E-08	0	N-rich protein, putative, expressed
LOC_Os07g34260	Os07g0526400	-2.886	1.1451E-13	0	chalcone and stilbene synthases, putative, expressed
LOC_Os05g32420	Os05g0390300	-2.886	0.00031029	1	expressed protein
LOC_Os12g02720	Os12g0120100	-2.885	9.3749E-10	1	DUF567 domain containing protein, putative, expressed
LOC_Os03g50960	Os03g0718800	-2.865	2.1438E-07	3	LTPL118 - Protease inhibitor/seed storage/LTP family protein precursor, expressed
LOC_Os07g38890	Os07g0576500	-2.851	0.00578834	5	OsGH3.9 - Probable indole-3-acetic acid-amido synthetase, expressed
LOC_Os06g50950	Os06g0725200	-2.848	9.8668E-13	0	GDSL-like lipase/acylhydrolase, putative, expressed
LOC_Os02g26810	Os02g0467600	-2.84	0.02478847	6	cytochrome P450, putative, expressed
LOC_Os02g46460	Os02g0689900	-2.839	2.5921E-10	2	peptide transporter PTR2, putative, expressed
LOC_Os08g34910	Os08g0450200	-2.832	0.02841961	0	pectinesterase, putative, expressed
LOC_Os11g40030	Os11g0614800	-2.802	0.00184406	15	cyclin-dependent kinase inhibitor, putative, expressed
LOC_Os06g36000	Os06g0553700	-2.8	0.03292939	0	AP2 domain containing protein, expressed
LOC_Os01g10504	Os01g0201700	-2.787	0.00277123	0	OsMADS3 - MADS-box family gene with MIKCc type-box, expressed
LOC_Os06g08380	Os06g0182300	-2.783	0.03464877	1	1,3-beta-glucan synthase component domain containing protein, expressed
LOC_Os07g26630	Os07g0448100	-2.78	0.00200094	2	aquaporin protein, putative, expressed
LOC_Os03g18779	Os03g0299700	-2.778	2.6473E-05	0	expressed protein
LOC_Os02g30910	Os02g0513100	-2.777	0.00343321	2	nodulin MtN3 family protein, putative, expressed
LOC_Os02g06779	Os02g0163533	-2.769	0.02813042	0	expressed protein
LOC_Os01g08380	Os01g0179000	-2.765	6.8328E-11	0	transferase family protein, putative, expressed
LOC_Os07g38130	Os07g0568700	-2.757	3.6273E-05	3	polygalacturonase inhibitor 1 precursor, putative, expressed
LOC_Os09g31390	Os09g0489500	-2.742	3.8276E-05	0	transcription factor, putative, expressed
LOC_Os07g34280	Os07g0526600	-2.717	4.6672E-06	0	CXE carboxylesterase, putative, expressed
LOC_Os01g02160	Os01g0112000	-2.714	0.01291471	0	hydroxyproline-rich glycoprotein family protein, putative, expressed
LOC_Os01g43280	Os01g0620800	-2.706	0.00715913	0	hydroquinone glucosyltransferase, putative, expressed
LOC_Os05g30570	Os05g0368600	-2.691	0.01824633	0	heavy metal-associated domain containing protein, expressed
LOC_Os03g55220	Os03g0759700	-2.69	0.00063298	0	bHelix-loop-helix transcription factor, putative, expressed
LOC_Os01g70870	Os01g0935000	-2.675	0.01041063	0	ZOS1-23 - C2H2 zinc finger protein, expressed
LOC_Os11g29290	Os11g0483000	-2.672	7.7729E-07	0	cytochrome P450, putative, expressed
LOC_Os07g40250	Os07g0592100	-2.67	0.00182136	0	sex determination protein tasselseed-2, putative, expressed
LOC_Os12g05760	None	-2.666	0.01453443	0	DnaK family protein, putative, expressed
LOC_Os11g45740	Os11g0684000	-2.661	2.0214E-13	1	MYB family transcription factor, putative, expressed
LOC_Os09g23620	Os09g0401000	-2.659	2.2932E-05	0	MYB family transcription factor, putative, expressed
LOC_Os06g12450	Os06g0229800	-2.656	0.00979938	0	soluble starch synthase 2-3, chloroplast precursor, putative, expressed

LOC_Os10g40430	Os10g0551800	-2.654	0.00014197	0	LTPL139 - Protease inhibitor/seed storage/LTP family protein precursor, expressed
LOC_Os06g13720	Os06g0246500	-2.643	8.7347E-05	1	dehydrogenase E1 component domain containing protein, expressed
LOC_Os11g37090	Os11g0579900	-2.623	0.0026881	0	pumilio-family RNA binding repeat domain containing protein, expressed
LOC_Os05g04680	Os05g0138000	-2.618	2.3296E-05	1	expressed protein
LOC_Os02g46030	Os02g0685200	-2.608	0.00055049	5	MYB family transcription factor, putative, expressed
LOC_Os09g21290	Os09g0380600	-2.607	0.00042716	0	expressed protein
LOC_Os02g06760	Os02g0163400	-2.59	9.0809E-07	0	expressed protein
LOC_Os01g74480	Os01g0976200	-2.59	0.03475438	1	cupin domain containing protein, expressed
LOC_Os04g44440	Os04g0526000	-2.586	0.01367371	0	TCP family transcription factor, putative, expressed
LOC_Os03g57640	Os03g0790500	-2.583	6.0942E-08	2	gibberellin receptor GID1L2, putative, expressed
LOC_Os03g15360	Os03g0259400	-2.583	0.0160699	2	nmrA-like family domain containing protein, expressed
LOC_Os11g13710	Os11g0241000	-2.58	2.284E-05	0	expressed protein
LOC_Os12g06480	Os12g0161900	-2.575	0.00267956	4	PHD-finger family protein, expressed
LOC_Os09g34230	Os09g0518000	-2.563	0.00013558	0	UDP-glucoronosyl/UDP-glucosyl transferase, putative, expressed
LOC_Os02g56490	Os02g0809300	-2.559	3.5841E-05	1	expressed protein
LOC_Os06g49860	Os06g0712800	-2.554	1.2212E-05	0	methyltransferase, putative, expressed
LOC_Os02g48770	Os02g0719600	-2.553	5.3214E-13	4	SAM dependent carboxyl methyltransferase, putative, expressed
LOC_Os08g42750	Os08g0540400	-2.544	4.8074E-05	1	CAMK_CAMK_like.37 - CAMK includes calcium/calmodulin depedent protein kinases, expressed
LOC_Os08g09380	Os08g0193200	-2.536	0.00900932	2	OsFBX263 - F-box domain containing protein, expressed
LOC_Os12g42280	Os12g0617400	-2.534	6.0613E-07	2	9-cis-epoxycarotenoid dioxygenase 1, chloroplast precursor, putative, expressed
LOC_Os04g48490	Os04g0574200	-2.533	0.00030471	0	fasciclin-like arabinogalactan protein, putative, expressed
LOC_Os08g09950	Os08g0199400	-2.526	2.1544E-06	1	acyl-desaturase, chloroplast precursor, putative, expressed
LOC_Os08g34900	Os08g0450100	-2.525	0.00509478	0	pectinesterase, putative, expressed
LOC_Os12g38460	Os12g0572800	-2.518	0.00107072	1	RNA recognition motif family protein, expressed
LOC_Os09g32440	Os09g0500300	-2.511	0.00683492	8	endonuclease/exonuclease/phosphatase family domain containing protein, expressed
LOC_Os07g05024	Os07g0143500	-2.499	3.6028E-05	0	expressed protein
LOC_Os08g09730	Os08g0197200	-2.493	0.02926344	0	OsFBX273 - F-box domain containing protein, expressed
LOC_Os04g38400	Os04g0456900	-2.491	0.00293292	3	ethylene-insensitive 3, putative, expressed
LOC_Os06g45540	Os06g0666100	-2.486	0.00708718	0	zinc-binding protein, putative, expressed
LOC_Os12g13160	Os12g0233900	-2.485	0.00084718	0	expressed protein
LOC_Os02g10520	Os02g0198700	-2.482	0.0006362	0	OsSub12 - Putative Subtilisin homologue, expressed
LOC_Os03g55030	Os03g0757200	-2.48	0.0003046	2	UDP-glucoronosyl and UDP-glucosyl transferase domain containing protein, expressed
LOC_Os02g48330	Os02g0713900	-2.478	0.03368839	1	3-hydroxy-3-methylglutaryl-coenzyme A reductase, putative, expressed
LOC_Os06g36520	Os06g0560700	-2.478	0.04524219	3	GDSL-like lipase/acylhydrolase, putative, expressed

LOC_Os06g47620	Os06g0691400	-2.476	1.5606E-10	0	peptidase, putative, expressed
LOC_Os07g44290	Os07g0637000	-2.473	8.1389E-10	1	CAMK_KIN1/SNF1/Nim1_like.29 - CAMK includes calcium/calmodulin depedent protein kinases, expressed
LOC_Os02g06754	Os02g0163300	-2.471	0.00018797	0	expressed protein
LOC_Os07g03120	Os07g0123100	-2.453	0.0078072	2	expressed protein
LOC_Os09g33640	Os09g0511200	-2.448	0.00549405	0	expressed protein
LOC_Os09g29130	Os09g0466400	-2.435	0.00324026	0	ZF-HD protein dimerisation region containing protein, expressed
LOC_Os04g23550	Os04g0301500	-2.432	4.0244E-13	0	basic helix-loop-helix family protein, putative, expressed
LOC_Os04g13160	Os04g0208500	-2.421	2.5422E-05	0	OsFBD9 - F-box and FBD domain containing protein, expressed
LOC_Os03g17670	Os03g0285200	-2.419	0.01019917	3	expressed protein
LOC_Os02g53160	Os02g0771400	-2.413	0.02824862	0	tyrosine phosphatase family protein, putative, expressed
LOC_Os05g35266	Os05g0427200	-2.408	5.8707E-07	3	galactosyltransferase, putative, expressed
LOC_Os05g38530	Os05g0427300	-2.397	0.00023886	1	DnaK family protein, putative, expressed
LOC_Os05g03640	Os05g0460000	-2.396	0.00471851	1	flavonol synthase/flavanone 3-hydroxylase, putative, expressed
LOC_Os01g10640	Os01g0203400	-2.393	1.9573E-08	2	expressed protein
LOC_Os08g06370	Os08g0160300	-2.39	0.00038131	0	MYB family transcription factor, putative, expressed
LOC_Os11g05480	Os11g0152700	-2.377	0.00016435	0	transcription factor, putative, expressed
LOC_Os03g59100	Os03g0805700	-2.372	0.04098777	0	pheophorbide a oxygenase, chloroplast precursor, putative, expressed
LOC_Os06g23360	Os06g0341500	-2.361	0.00795383	0	LTPL70 - Protease inhibitor/seed storage/LTP family protein precursor, expressed
LOC_Os02g49560	Os02g0728001	-2.355	0.01349659	0	bZIP transcription factor domain containing protein, expressed
LOC_Os01g50700	Os01g0702500	-2.347	0.00546541	0	dehydrin family protein, expressed
LOC_Os02g29150	Os02g0493100	-2.346	6.1713E-06	3	OsFBO11 - F-box and other domain containing protein, expressed
LOC_Os04g50740	Os04g0593800	-2.341	0.00694574	0	ara54-like RING finger protein, putative, expressed
LOC_Os03g38540	Os03g0582000	-2.334	4.4341E-11	0	folic acid binding protein, putative, expressed
LOC_Os01g42690	Os01g0612500	-2.323	0.00244104	2	OsPOP2 - Putative Prolyl Oligopeptidase homologue, expressed
LOC_Os06g45590	Os06g0666600	-2.321	0.00029216	0	glyceraldehyde-3-phosphate dehydrogenase, putative, expressed
LOC_Os02g51110	Os02g0745100	-2.313	0.00020691	1	aquaporin protein, putative, expressed
LOC_Os01g66590	Os01g0889400	-2.312	0.00018246	2	DUF260 domain containing protein, putative, expressed
LOC_Os06g10990	Os06g0212900	-2.311	0.0026881	0	DnaK family protein, putative, expressed
LOC_Os11g41500	None	-2.301	0.00531914	0	NC domain-containing protein, putative, expressed
LOC_Os04g50176	Os04g0591300	-2.299	0.02404492	3	expressed protein
LOC_Os03g06920	Os03g0165200	-2.286	0.01660545	0	DRD1, putative, expressed
LOC_Os03g09200	Os03g0192300	-2.284	0.00710492	4	domain of unknown function DUF966 domain containing protein, expressed
LOC_Os01g11150	Os01g0209700	-2.281	0.00941771	0	gibberellin 2-beta-dioxygenase, putative, expressed
LOC_Os04g50090	Os04g0590800	-2.28	0.02454545	2	helix-loop-helix DNA-binding protein, putative, expressed
LOC_Os01g66100	Os01g0883800	-2.272	7.7911E-07	0	gibberellin 20 oxidase 2, putative, expressed

LOC_Os05g37250	Os05g0445100	-2.266	1.1441E-11	0	cytochrome P450, putative, expressed
LOC_Os04g19140	Os04g0261400	-2.266	0.03037198	1	expressed protein
LOC_Os08g33150	Os08g0428200	-2.261	0.00515669	0	MYB family transcription factor, putative, expressed
LOC_Os01g60020	Os01g0816100	-2.251	8.3979E-09	0	NAC domain transcription factor, putative, expressed
LOC_Os01g53750	Os01g0739700	-2.248	0.0272195	1	glucan endo-1,3-beta-glucosidase precursor, putative, expressed
LOC_Os03g28110	None	-2.243	0.01371813	0	retrotransposon protein, putative, unclassified, expressed
LOC_Os01g61850	None	-2.239	6.7702E-10	0	None
LOC_Os08g04630	Os08g0141400	-2.232	3.8171E-11	1	external NADH-ubiquinone oxidoreductase 1, mitochondrial precursor, putative, expressed
LOC_Os02g11020	Os02g0204700	-2.232	0.02967275	1	cytochrome P450 72A1, putative, expressed
LOC_Os09g26780	Os09g0439200	-2.231	7.7878E-10	1	zinc-finger protein, putative, expressed
LOC_Os05g03960	Os05g0130300	-2.231	3.7748E-05	5	expressed protein
LOC_Os08g28710	Os08g0374600	-2.226	4.0847E-06	2	receptor protein kinase CRINKLY4 precursor, putative, expressed
LOC_Os12g38170	Os12g0569500	-2.222	2.3031E-05	0	osmotin, putative, expressed
LOC_Os07g37730	Os07g0564500	-2.22	4.5193E-06	1	NADH-ubiquinone oxidoreductase, mitochondrial precursor, putative, expressed
LOC_Os06g48720	Os06g0700700	-2.215	1.0385E-05	0	cadmium/zinc-transporting ATPase 4, putative, expressed
LOC_Os01g02890	Os01g0118300	-2.203	2.428E-05	0	phosphatidylserine synthase, putative, expressed
LOC_Os06g36680	Os06g0562300	-2.2	0.00296962	2	homeobox domain containing protein, expressed
LOC_Os06g49840	Os06g0712700	-2.197	3.9267E-07	1	OsMADS16 - MADS-box family gene with MIKCc type-box, expressed
LOC_Os09g01650	Os09g0103800	-2.171	0.01510902	1	peptidyl-prolyl cis-trans isomerase, FKBP-type, putative, expressed
LOC_Os04g15920	Os04g0229100	-2.168	5.9481E-11	0	dehydrogenase, putative, expressed
LOC_Os01g18120	Os01g0283700	-2.168	1.8962E-07	1	cinnamoyl CoA reductase, putative, expressed
LOC_Os07g48040	Os07g0677400	-2.168	0.00892471	1	peroxidase precursor, putative, expressed
LOC_Os12g19180	Os12g0289600	-2.162	0.04731604	0	ATP binding protein, putative, expressed
LOC_Os01g50940	Os01g0705700	-2.157	1.1441E-11	1	helix-loop-helix DNA-binding domain containing protein, expressed
LOC_Os04g42889	None	-2.156	0.00323664	0	expressed protein
LOC_Os01g07660	Os01g0171200	-2.147	0.0229354	5	expressed protein
LOC_Os04g10350	Os04g0182200	-2.122	1.8616E-10	0	1-aminocyclopropane-1-carboxylate oxidase homolog 2, putative, expressed
LOC_Os09g17630	Os09g0345300	-2.12	1.8324E-06	0	receptor-like protein kinase 2, putative, expressed
LOC_Os12g25660	Os12g0443000	-2.11	0.00025523	0	cytochrome P450, putative, expressed
LOC_Os02g26210	Os02g0460200	-2.11	0.00040884	10	flowering promoting factor-like 1, putative, expressed
LOC_Os01g03310	Os01g0123900	-2.106	4.9113E-07	0	BBTI1 - Bowman-Birk type bran trypsin inhibitor precursor, expressed
LOC_Os02g57350	Os02g0818900	-2.104	0.001347	0	expressed protein
LOC_Os01g12000	Os01g0219300	-2.099	0.00168477	0	expressed protein
LOC_Os05g08620	Os05g0179100	-2.098	0.00613969	0	expressed protein
LOC_Os04g06810	None	-2.094	0.02952702	0	None

LOC_Os02g11040	Os02g0205200	-2.093	0.01478263	1	expressed protein
LOC_Os02g29210	Os02g0494400	-2.092	3.5122E-06	2	ankyrin, putative, expressed
LOC_Os07g40290	Os07g0592600	-2.09	0.000243	6	OsGH3.8 - Probable indole-3-acetic acid-amido synthetase, expressed
LOC_Os09g28370	Os09g0456900	-2.089	0.00164097	0	retrotransposon protein, putative, unclassified, expressed
LOC_Os09g08130	Os09g0255400	-2.083	0.01160403	2	indole-3-glycerol phosphate synthase, chloroplast precursor, putative, expressed
LOC_Os02g44900	Os02g0669900	-2.08	0.00925636	10	expressed protein
LOC_Os05g35380	Os05g0428400	-2.079	0.00395677	0	universal stress protein domain containing protein, putative, expressed
LOC_Os09g30458	Os09g0482660	-2.074	0.02657732	0	subtilisin-like protease, putative, expressed
LOC_Os02g55130	Os02g0794500	-2.073	0.00040884	0	OsSCP10 - Putative Serine Carboxypeptidase homologue, expressed
LOC_Os05g32070	Os05g0386201	-2.067	0.03739513	11	LRP1, putative, expressed
LOC_Os04g46160	Os04g0546000	-2.06	0.00061105	5	expressed protein
LOC_Os11g34710	Os11g0549615	-2.06	0.01840547	0	Ser/Thr protein phosphatase family protein, putative, expressed
LOC_Os03g12510	Os03g0226200	-2.06	0.04491525	3	non-symbiotic hemoglobin 2, putative, expressed
LOC_Os05g37660	Os05g0449200	-2.058	5.7451E-07	4	transferase family protein, putative, expressed
LOC_Os08g41960	Os08g0531900	-2.046	0.01269455	0	OsMADS37 - MADS-box family gene with MIKC* type-box, expressed
LOC_Os06g43600	Os06g0643500	-2.04	0.00097763	1	LTPL129 - Protease inhibitor/seed storage/LTP family protein precursor, expressed
LOC_Os11g38610	Os11g0598800	-2.021	0.00018243	1	expressed protein
LOC_Os12g06080	Os12g0157000	-2.021	0.01679088	0	expressed protein
LOC_Os01g08190	Os01g0177100	-2.017	0.00045115	0	transcriptional corepressor LEUNIG, putative, expressed
LOC_Os05g33140	Os05g0399400	-2.014	0.00110885	0	CHIT5 - Chitinase family protein precursor, expressed
LOC_Os02g06770	Os02g0163500	-2.012	0.0076367	0	expressed protein
LOC_Os06g47830	Os06g0693300	-2.007	0.00953645	1	RPA2C - Putative single-stranded DNA binding complex subunit 2, expressed
LOC_Os04g55840	Os04g0652600	-2.003	0.00139514	0	expressed protein
LOC_Os01g26000	Os01g0361700	-2.003	0.03717824	1	expressed protein
LOC_Os02g04540	Os02g0138000	-2.001	0.03253585	0	retrotransposon protein, putative, unclassified, expressed
LOC_Os03g17230	Os03g0280800	-1.996	0.01283229	2	NAD dependent epimerase/dehydratase family domain containing protein, expressed
LOC_Os10g36420	Os10g0508300	-1.994	0.00080906	0	YABBY domain containing protein, putative, expressed
LOC_Os12g12580	Os12g0226900	-1.991	3.4154E-08	1	NADP-dependent oxidoreductase, putative, expressed
LOC_Os05g51130	Os05g0588900	-1.987	5.2119E-06	0	mitochondrial chaperone BCS1, putative, expressed
LOC_Os08g03825	None	-1.987	0.0289311	6	expressed protein
LOC_Os02g40700	Os02g0620400	-1.984	9.0686E-05	1	enzyme of the cupin superfamily protein, putative, expressed
LOC_Os12g10750	Os12g0210500	-1.98	3.7425E-06	0	ARGOS, putative, expressed
LOC_Os04g41280	Os04g0490300	-1.98	0.00224931	4	ankyrin repeat domain containing protein, expressed
LOC_Os02g51000	Os02g0743800	-1.978	0.00270944	0	CS domain containing protein, putative, expressed
LOC_Os02g47770	Os02g0706600	-1.978	0.01425244	0	ZF-HD protein dimerisation region containing protein, expressed

LOC_Os08g14990	Os08g0248100	-1.977	1.1705E-05	0	receptor-like protein kinase 2 precursor, putative, expressed
LOC_Os04g53190	Os04g0623300	-1.971	1.0857E-08	0	CPuORF12 - conserved peptide uORF-containing transcript, expressed
LOC_Os01g64360	Os01g0863300	-1.969	6.7843E-06	6	MYB family transcription factor, putative, expressed
LOC_Os02g07030	Os02g0166800	-1.963	0.00029604	0	DUF640 domain containing protein, putative, expressed
LOC_Os05g39140	Os05g0467600	-1.962	0.00027135	0	expressed protein
LOC_Os02g14190	Os02g0237100	-1.958	0.01324622	0	spermidine synthase, putative, expressed
LOC_Os03g54160	Os03g0752800	-1.956	6.8484E-09	1	OsMADS14 - MADS-box family gene with MIKCc type-box, expressed
LOC_Os08g43410	Os08g0547600	-1.954	0.00027712	0	LRP1, putative, expressed
LOC_Os08g38210	Os08g0490000	-1.951	1.8317E-09	1	transcription factor BIM2, putative, expressed
LOC_Os03g20350	Os03g0319200	-1.951	0.03996013	1	expressed protein
LOC_Os01g74300	Os01g0974200	-1.948	0.00452655	0	metallothionein, putative, expressed
LOC_Os02g41610	Os02g0625900	-1.946	0.00076121	1	expressed protein
LOC_Os05g01730	Os05g0107900	-1.946	0.00121263	0	drought induced 19 protein, putative, expressed
LOC_Os07g41310	Os07g0604300	-1.935	0.00079154	0	COBRA, putative, expressed
LOC_Os01g08710	Os01g0182700	-1.934	0.00244981	0	WRKY102, expressed
LOC_Os05g43930	Os05g0515500	-1.934	0.00451463	0	O-methyltransferase, putative, expressed
LOC_Os03g01320	Os03g0103300	-1.932	6.7052E-06	0	LTPL116 - Protease inhibitor/seed storage/LTP family protein precursor, expressed
LOC_Os02g09620	Os02g0189300	-1.93	0.04739489	0	GDSL-like lipase/acylhydrolase, putative, expressed
LOC_Os01g45470	Os01g0642200	-1.926	7.8623E-08	0	expressed protein
LOC_Os02g50110	Os02g0733900	-1.923	0.018248	1	expressed protein
LOC_Os06g12500	Os06g0231050	-1.922	0.03088232	0	membrane associated DUF588 domain containing protein, putative, expressed
LOC_Os08g41940	Os08g0531600	-1.92	0.03844937	0	OsSPL16 - SBP-box gene family member, expressed
LOC_Os08g39330	Os08g0503000	-1.916	0.01111964	0	skin secretory protein xP2 precursor, putative, expressed
LOC_Os09g20000	Os09g0364800	-1.911	0.00031667	0	heavy metal-associated domain containing protein, expressed
LOC_Os03g16900	Os03g0277000	-1.911	0.02241163	1	rab GDP dissociation inhibitor alpha, putative, expressed
LOC_Os12g40260	Os12g0594000	-1.909	0.00270944	2	WD-40 repeat family protein, putative, expressed
LOC_Os05g11750	Os05g0207700	-1.909	0.03904576	0	protein kinase, putative, expressed
LOC_Os11g32750	Os11g0531700	-1.907	0.00562381	0	hydrolase, NUDIX family, domain containing protein, expressed
LOC_Os02g52930	Os02g0768300	-1.904	0.00917032	0	integral membrane protein DUF6 containing protein, expressed
LOC_Os05g36260	Os05g0438500	-1.903	0.00264129	0	soluble inorganic pyrophosphatase, putative, expressed
LOC_Os09g27940	Os09g0452800	-1.9	0.00186413	1	aspartic proteinase nepenthesin-1 precursor, putative, expressed
LOC_Os05g35290	Os05g0427400	-1.892	7.4862E-06	1	phenylalanine ammonia-lyase, putative, expressed
LOC_Os05g39150	Os05g0467700	-1.892	0.00979938	0	expressed protein
LOC_Os02g40010	Os02g0613900	-1.888	5.3346E-05	0	phosphoribosyl transferase, putative, expressed
LOC_Os01g21070	Os01g0312800	-1.888	0.04118347	0	endoglucanase, putative, expressed

LOC_Os04g59150	Os04g0688100	-1.887	0.00017994	0	peroxidase precursor, putative, expressed
LOC_Os01g01870	Os01g0108600	-1.886	3.1147E-09	0	helix-loop-helix DNA-binding domain containing protein, expressed
LOC_Os08g31980	Os08g0414700	-1.885	0.00150484	0	trehalose-6-phosphate synthase, putative, expressed
LOC_Os04g57350	Os04g0669300	-1.885	0.00562381	2	EH domain-containing protein 1, putative, expressed
LOC_Os06g08550	Os06g0184500	-1.885	0.01367966	1	BTBN14 - Bric-a-Brac, Tramtrack, Broad Complex BTB domain with non-phototropic hypocotyl 3 NPH3 domain, expressed
LOC_Os06g10210	Os06g0203600	-1.884	1.9573E-08	0	expressed protein
LOC_Os05g29790	Os05g0361500	-1.881	2.3704E-05	0	pectinesterase, putative, expressed
LOC_Os03g18070	Os03g0290300	-1.876	2.8156E-08	2	omega-3 fatty acid desaturase, chloroplast precursor, putative, expressed
LOC_Os01g42810	Os01g0614300	-1.873	0.00064571	0	leaf senescence related protein, putative, expressed
LOC_Os03g03034	Os03g0122300	-1.872	3.1764E-05	2	flavonol synthase/flavanone 3-hydroxylase, putative, expressed
LOC_Os12g05440	Os12g0150200	-1.87	4.9526E-05	0	cytochrome P450, putative, expressed
LOC_Os08g36920	Os08g0474000	-1.869	2.762E-08	1	AP2 domain containing protein, expressed
LOC_Os02g27810	Os02g0479300	-1.869	0.01508946	3	MazG nucleotide pyrophosphohydrolase domain containing protein, expressed
LOC_Os12g05120	Os12g0145900	-1.867	0.00975971	0	receptor kinase, putative, expressed
LOC_Os01g19330	Os01g0298400	-1.867	0.03614384	3	MYB family transcription factor, putative, expressed
LOC_Os11g43520	Os11g0655900	-1.864	0.0221544	0	OsGrx_C17 - glutaredoxin subgroup III, expressed
LOC_Os03g30740	Os03g0421000	-1.86	0.00049251	0	expressed protein
LOC_Os03g17680	Os03g0285300	-1.86	0.01303288	1	expressed protein
LOC_Os07g36570	Os07g0550900	-1.855	0.01217297	0	KI domain interacting kinase 1, putative, expressed
LOC_Os04g46220	Os04g0546800	-1.852	1.8181E-07	0	ethylene-responsive transcription factor, putative, expressed
LOC_Os03g39910	Os03g0596400	-1.848	0.00050052	5	XH domain containing protein, expressed
LOC_Os06g28630	Os06g0480500	-1.847	0.01056792	0	expressed protein
LOC_Os09g03750	Os09g0124800	-1.846	0.00277123	0	ankyrin, putative, expressed
LOC_Os02g03830	Os02g0131000	-1.842	0.03791182	0	expressed protein
LOC_Os06g01840	Os06g0107600	-1.839	0.00798236	1	nodulin, putative, expressed
LOC_Os03g08840	Os03g0187400	-1.838	6.1332E-06	0	zinc finger protein, putative, expressed
LOC_Os06g37080	Os06g0567200	-1.837	0.01092394	0	L-ascorbate oxidase precursor, putative, expressed
LOC_Os02g41840	Os02g0629000	-1.836	0.00467906	2	DUF584 domain containing protein, putative, expressed
LOC_Os05g50610	Os05g0583000	-1.833	0.00240253	0	WRKY8, expressed
LOC_Os05g07740	Os05g0169600	-1.832	0.00057383	1	receptor-like protein kinase 2 precursor, putative, expressed
LOC_Os08g05290	Os08g0148300	-1.831	0.00506134	0	receptor-like protein kinase 5 precursor, putative, expressed
LOC_Os07g09420	Os07g0192000	-1.83	0.00356497	0	ATPase, putative, expressed
LOC_Os07g12340	Os07g0225300	-1.828	6.5035E-06	1	NAC domain-containing protein 67, putative, expressed
LOC_Os11g30760	Os11g0502700	-1.828	0.01054144	0	expressed protein
LOC_Os01g55350	Os01g0758300	-1.827	0.00033801	0	phosphoenolpyruvate carboxylase, putative, expressed

LOC_Os12g39100	Os12g0580600	-1.827	0.02952102	0	expressed protein
LOC_Os08g34510	Os08g0444500	-1.822	0.00180693	0	expressed protein
LOC_Os07g02120	Os07g0111900	-1.822	0.02914244	5	flavin-containing monooxygenase family protein, putative, expressed
LOC_Os03g10390	Os03g0201000	-1.821	0.02865761	1	expressed protein
LOC_Os07g05940	Os07g0154100	-1.82	0.00089163	0	9-cis-epoxycarotenoid dioxygenase 1, chloroplast precursor, putative, expressed
LOC_Os07g35340	Os07g0537900	-1.82	0.01029996	0	TKL_IRAK_DUF26-lc.14 - DUF26 kinases have homology to DUF26 containing loci, expressed
LOC_Os03g01880	Os03g0109300	-1.816	4.2638E-06	1	possible lysine decarboxylase domain containing protein, expressed
LOC_Os04g32480	Os04g0395800	-1.816	0.00039959	2	zinc-finger protein, putative, expressed
LOC_Os08g37456	Os08g0480200	-1.812	0.01269455	4	flavonol synthase/flavanone 3-hydroxylase, putative, expressed
LOC_Os09g38110	Os09g0554200	-1.81	0.00033615	0	RING-H2 finger protein, putative, expressed
LOC_Os03g10210	Os03g0198600	-1.807	0.01137202	1	homeobox domain containing protein, expressed
LOC_Os08g30520	Os08g0395800	-1.802	0.00094048	0	plant protein of unknown function domain containing protein, expressed
LOC_Os09g31080	Os09g0484200	-1.801	1.3496E-05	0	induced stolen tip protein TUB8, putative, expressed
LOC_Os01g13130	Os01g0232100	-1.797	0.0002048	0	aquaporin protein, putative, expressed
LOC_Os03g03370	Os03g0125100	-1.796	0.00053977	1	fatty acid hydroxylase, putative, expressed
LOC_Os01g50400	Os01g0699400	-1.795	0.00260589	0	STE_MEKK_ste11_MAP3K.5 - STE kinases include homologs to sterile 7, sterile 11 and sterile 20 from yeast, expressed
LOC_Os01g18860	Os01g0293000	-1.792	0.00071911	1	S-adenosylmethionine synthetase, putative, expressed
LOC_Os03g58300	Os03g0797400	-1.785	0.01325644	0	indole-3-glycerol phosphate lyase, chloroplast precursor, putative, expressed
LOC_Os05g39880	Os05g0476400	-1.779	0.0027996	0	expressed protein
LOC_Os09g34890	Os09g0520500	-1.773	0.00197284	0	expressed protein
LOC_Os04g54610	Os04g0639000	-1.773	0.00832017	0	expressed protein
LOC_Os03g28940	Os03g0402800	-1.768	3.7567E-07	1	ZIM domain containing protein, putative, expressed
LOC_Os04g49550	Os04g0585000	-1.766	2.3653E-05	0	RING-H2 finger protein ATL2A, putative, expressed
LOC_Os08g30900	Os08g0400000	-1.766	0.00119773	12	YDG/SRA domain containing protein, expressed
LOC_Os10g42410	Os10g0574700	-1.76	1.0839E-06	0	zinc-binding protein, putative, expressed
LOC_Os09g29830	Os09g0474100	-1.755	0.00795837	0	helix-loop-helix DNA-binding domain containing protein, expressed
LOC_Os11g15570	Os11g0261900	-1.752	0.0023924	0	Ser/Thr protein phosphatase family protein, putative, expressed
LOC_Os02g38410	Os02g0597700	-1.748	0.0007813	0	ubiquitin family protein, putative, expressed
LOC_Os06g51380	Os06g0730200	-1.742	5.9329E-06	0	ROOT HAIRLESS 1, putative, expressed
LOC_Os02g42850	Os02g0641300	-1.736	0.00030421	1	MYB family transcription factor, putative, expressed
LOC_Os03g12660	Os03g0227700	-1.73	0.00526759	1	cytochrome P450, putative, expressed
LOC_Os02g32860	Os02g0530600	-1.726	0.00522811	3	poly synthetase 3, putative, expressed
LOC_Os01g10140	Os01g0198000	-1.726	0.00655897	2	RNA-dependent RNA polymerase, putative, expressed
LOC_Os11g05190	Os11g0149400	-1.725	0.00867175	0	phytosulfokines precursor, putative, expressed

LOC_Os01g68870	Os01g0917500	-1.718	3.9132E-06	0	leucine-rich repeat receptor protein kinase EXS precursor, putative, expressed
LOC_Os03g11600	Os03g0215200	-1.711	0.00620404	0	YABBY domain containing protein, putative, expressed
LOC_Os07g46820	Os07g0663300	-1.706	0.00226857	0	CUG-BP- and ETR-3-like factor 5, putative, expressed
LOC_Os04g49000	Os04g0579200	-1.702	4.3582E-06	2	zinc finger, C3HC4 type domain containing protein, expressed
LOC_Os02g56380	Os02g0807900	-1.697	0.00160194	5	OsWAK21 - OsWAK receptor-like cytoplasmic kinase OsWAK-RLCK, expressed
LOC_Os09g32290	Os09g0498400	-1.694	0.01749839	1	FAD dependent oxidoreductase domain containing protein, expressed
LOC_Os07g28850	Os07g0471300	-1.691	0.01088293	0	retrotransposon protein, putative, unclassified, expressed
LOC_Os01g40630	Os01g0588900	-1.689	5.895E-05	0	uncharacterized protein PA4923, putative, expressed
LOC_Os02g48080	Os02g0710500	-1.685	0.00596557	0	cysteine-rich receptor-like protein kinase 7 precursor, putative, expressed
LOC_Os08g40850	Os08g0520000	-1.683	0.00578124	0	mitochondrial carrier protein, putative, expressed
LOC_Os07g44060	Os07g0634400	-1.67	9.6548E-06	0	haloacid dehalogenase-like hydrolase family protein, putative, expressed
LOC_Os03g10260	Os03g0199500	-1.67	0.02604335	1	expressed protein
LOC_Os03g60880	Os03g0823800	-1.67	0.04491525	0	nucleobase-ascorbate transporter, putative, expressed
LOC_Os03g44670	Os03g0649000	-1.669	3.7748E-05	1	expressed protein
LOC_Os06g02410	Os06g0114400	-1.668	0.02109127	0	expressed protein
LOC_Os03g54930	Os03g0756500	-1.668	0.04506614	2	VIP4, putative, expressed
LOC_Os05g07764	Os05g0169800	-1.662	0.00557124	0	carboxyl-terminal proteinase, putative, expressed
LOC_Os12g39130	Os12g0581000	-1.66	0.03358122	3	fibroin heavy chain precursor, putative, expressed
LOC_Os08g44360	Os08g0557800	-1.653	0.0237695	0	male sterility protein 2, putative, expressed
LOC_Os05g02070	Os05g0111300	-1.645	0.00967385	1	expressed protein
LOC_Os12g30520	Os12g0488900	-1.644	0.03898957	0	pumilio-family RNA binding repeat containing protein, expressed
LOC_Os08g28170	Os08g0369000	-1.638	0.04221219	0	nucleobase-ascorbate transporter, putative, expressed
LOC_Os07g06620	Os07g0160100	-1.637	3.944E-05	0	YABBY domain containing protein, putative, expressed
LOC_Os03g19220	Os03g0304800	-1.634	0.00321949	0	expressed protein
LOC_Os04g45730	Os04g0540900	-1.628	5.4926E-07	0	protein kinase domain containing protein, expressed
LOC_Os01g56690	Os01g0773800	-1.624	0.04190467	0	helix-loop-helix DNA-binding domain containing protein, expressed
LOC_Os05g49420	Os05g0569300	-1.623	6.9094E-07	3	transcription factor, putative, expressed
LOC_Os12g38440	Os12g0572500	-1.623	1.4157E-05	1	XH domain containing protein, expressed
LOC_Os05g38290	Os05g0457200	-1.619	0.00152747	1	protein phosphatase 2C, putative, expressed
LOC_Os05g33400	Os05g0402900	-1.618	0.00286996	5	basic 7S globulin precursor, putative, expressed
LOC_Os02g56860	Os02g0813600	-1.615	0.00084206	1	3-ketoacyl-CoA synthase, putative, expressed
LOC_Os06g14670	Os06g0258000	-1.615	0.00579053	0	ODORANT1, putative, expressed
LOC_Os09g38320	Os09g0555500	-1.613	0.00013558	2	phytoene synthase, chloroplast precursor, putative, expressed
LOC_Os06g41030	Os06g0613000	-1.61	0.0194117	0	DUF1680 domain containing protein, putative, expressed
LOC_Os05g44060	Os05g0516700	-1.609	0.00010529	1	expressed protein

LOC_Os06g36920	Os06g0565100	-1.609	0.0165313	8	cytochrome P450, putative, expressed
LOC_Os10g35950	Os10g0503300	-1.608	0.00168459	0	transferase family protein, putative, expressed
LOC_Os05g47770	Os05g0550800	-1.608	0.01254197	0	serine/threonine-protein kinase At1g18390 precursor, putative, expressed
LOC_Os08g04800	Os08g0143600	-1.608	0.02348736	1	triacylglycerol lipase like protein, putative, expressed
LOC_Os12g42910	Os12g0624200	-1.607	0.00057327	0	sodium/calcium exchanger protein, putative, expressed
LOC_Os09g32930	Os09g0506800	-1.605	0.00108815	0	retrotransposon protein, putative, SINE subclass, expressed
LOC_Os09g07154	Os09g0246300	-1.605	0.01289437	3	expressed protein
LOC_Os06g47600	Os06g0691200	-1.604	0.027904	2	thaumatin family domain containing protein, expressed
LOC_Os02g40840	Os02g0621800	-1.598	0.00022375	1	alcohol oxidase, putative, expressed
LOC_Os07g42280	Os07g0614300	-1.593	0.0005595	3	von Willebrand factor type A domain containing protein, expressed
LOC_Os11g31430	Os11g0513500	-1.593	0.03368839	2	expressed protein
LOC_Os02g21700	Os02g0322400	-1.59	0.00110157	0	STE_MEKK_ste11_MAP3K.8 - STE kinases include homologs to sterile 7, sterile 11 and sterile 20 from yeast, expressed
LOC_Os09g25860	Os09g0427125	-1.586	0.00844813	0	expressed protein
LOC_Os07g36590	Os07g0551300	-1.583	1.7906E-05	0	serine/threonine-protein kinase receptor precursor, putative, expressed
LOC_Os07g42994	Os07g0622700	-1.573	0.02182441	3	hydrolase, putative, expressed
LOC_Os03g58580	Os03g0800000	-1.568	0.00756238	2	nodulin, putative, expressed
LOC_Os11g35060	Os11g0552200	-1.566	0.03522271	7	agenet domain containing protein, expressed
LOC_Os11g03230	Os11g0125900	-1.56	7.6185E-05	1	nucleoside-triphosphatase, putative, expressed
LOC_Os05g07940	Os05g0171900	-1.559	0.00018891	6	glyoxalase family protein, putative, expressed
LOC_Os06g46350	Os06g0677000	-1.559	0.00058346	0	PLA IIIA/PLP7, putative, expressed
LOC_Os03g12414	Os03g0225200	-1.558	0.00659814	1	cyclin, putative, expressed
LOC_Os03g47280	Os03g0676400	-1.557	0.00041836	3	VQ domain containing protein, putative, expressed
LOC_Os04g46600	Os04g0551600	-1.553	0.02549256	0	zinc finger protein, putative, expressed
LOC_Os08g35160	Os08g0452900	-1.551	0.00528953	1	heat shock protein DnaJ, putative, expressed
LOC_Os01g68570	Os01g0914000	-1.544	0.02621932	1	expressed protein
LOC_Os03g55800	Os03g0767000	-1.541	7.2435E-06	3	cytochrome P450, putative, expressed
LOC_Os05g30500	Os05g0368000	-1.537	0.00423747	1	expressed protein
LOC_Os05g34830	Os05g0421600	-1.537	0.00658701	3	No apical meristem protein, putative, expressed
LOC_Os12g38850	Os12g0577600	-1.536	0.00014294	0	DUF1336 domain containing protein, expressed
LOC_Os06g50900	Os06g0724800	-1.536	0.00084718	0	expressed protein
LOC_Os10g39660	Os10g0542700	-1.527	0.00039171	1	expressed protein
LOC_Os01g12440	Os01g0224100	-1.52	0.01014473	2	AP2 domain containing protein, expressed
LOC_Os02g52670	Os02g0764700	-1.52	0.02572829	3	AP2 domain containing protein, expressed
LOC_Os04g09720	Os04g0176400	-1.519	0.00556478	0	OsSCP22 - Putative Serine Carboxypeptidase homologue, expressed
LOC_Os01g50100	Os01g0695800	-1.517	0.028786	1	ABC transporter, ATP-binding protein, putative, expressed

LOC_Os01g36580	Os01g0546400	-1.517	0.03045171	0	auxin-induced protein 5NG4, putative, expressed
LOC_Os05g38660	Os05g0461600	-1.512	0.0070836	2	expressed protein
LOC_Os04g54620	Os04g0639100	-1.508	0.0248116	0	expressed protein
LOC_Os10g38870	Os10g0532300	-1.502	0.00179059	3	heavy metal-associated domain containing protein, expressed
LOC_Os04g48950	Os04g0578800	-1.502	0.00549405	0	fringe-related protein, putative, expressed
LOC_Os03g32314	Os03g0438100	-1.5	0.000124	0	allene oxide cyclase 4, chloroplast precursor, putative, expressed
LOC_Os08g07010	Os08g0167000	-1.497	0.02606239	1	ABC-2 type transporter domain containing protein, expressed
LOC_Os01g16980	Os01g0276900	-1.494	0.04506614	0	expressed protein
LOC_Os01g50370	Os01g0699100	-1.493	0.00416833	4	STE_MEKK_ste11_MAP3K.4 - STE kinases include homologs to sterile 7, sterile 11 and sterile 20 from yeast, expressed
LOC_Os08g29570	Os08g0384500	-1.493	0.02598506	0	pleiotropic drug resistance protein 3, putative, expressed
LOC_Os02g12890	Os02g0221900	-1.492	0.00561345	0	cytochrome P450, putative, expressed
LOC_Os11g07460	Os11g0175700	-1.489	0.03776579	0	TCP family transcription factor, putative, expressed
LOC_Os10g38820	Os10g0531900	-1.487	0.04375864	0	bZIP family transcription factor, putative, expressed
LOC_Os09g32540	Os09g0502100	-1.479	0.00260001	5	expressed protein
LOC_Os04g32404	Os04g0395100	-1.479	0.01448497	1	expressed protein
LOC_Os01g43844	Os01g0628700	-1.474	5.5549E-05	0	cytochrome P450 72A1, putative, expressed
LOC_Os05g46350	Os05g0541100	-1.472	0.02880884	4	IQ calmodulin-binding motif domain containing protein, expressed
LOC_Os09g25620	Os09g0424300	-1.47	0.00270944	0	CPuORF8 - conserved peptide uORF-containing transcript, expressed
LOC_Os03g08520	Os03g0183500	-1.468	0.04128117	0	DUF581 domain containing protein, expressed
LOC_Os02g54870	Os02g0791300	-1.467	0.03785779	0	expressed protein
LOC_Os07g48130	Os07g0679000	-1.466	0.0001703	1	potassium transporter, putative, expressed
LOC_Os07g49114	Os07g0691300	-1.465	2.5846E-05	6	wound-induced protein WI12, putative, expressed
LOC_Os03g06330	Os03g0159100	-1.465	4.3616E-05	0	tyrosine protein kinase domain containing protein, putative, expressed
LOC_Os03g10250	Os03g0199300	-1.464	0.02761343	0	expressed protein
LOC_Os02g49720	Os02g0730000	-1.451	0.02970759	0	aldehyde dehydrogenase, putative, expressed
LOC_Os01g70860	Os01g0934900	-1.451	0.03722762	2	esterase, putative, expressed
LOC_Os01g03980	Os01g0130900	-1.45	0.00094048	3	expressed protein
LOC_Os05g06814	Os05g0160600	-1.45	0.00272418	2	expressed protein
LOC_Os02g48094	Os02g0710700	-1.449	0.04919373	0	expressed protein
LOC_Os08g38890	Os08g0497900	-1.447	0.00010881	2	expressed protein
LOC_Os03g41060	Os03g0607200	-1.446	0.00503219	2	GASR2 - Gibberellin-regulated GASA/GAST/Snakin family protein precursor, putative, expressed
LOC_Os02g03960	Os02g0132500	-1.445	1.8783E-05	0	CPuORF1 - conserved peptide uORF-containing transcript, expressed
LOC_Os10g36924	Os10g0513200	-1.442	0.00024705	0	aquaporin protein, putative, expressed
LOC_Os07g05700	Os07g0151800	-1.442	0.00624192	5	expressed protein

LOC_Os05g38270	Os05g0456925	-1.44	0.02083316	1	regulator of chromosome condensation, putative, expressed
LOC_Os01g43460	Os01g0623200	-1.439	0.00447364	3	C4-dicarboxylate transporter/malic acid transport protein, expressed
LOC_Os12g05260	Os12g0147800	-1.438	0.0033067	0	phytosulfokines precursor, putative, expressed
LOC_Os01g19850	Os01g0304100	-1.436	0.02604335	0	amino acid permease family protein, putative, expressed
LOC_Os02g43540	Os02g0651900	-1.435	0.00440587	0	retrotransposon protein, putative, unclassified, expressed
LOC_Os01g51890	Os01g0716800	-1.433	0.00098165	1	endonuclease/exonuclease/phosphatase family domain containing protein, expressed
LOC_Os08g37104	Os08g0476000	-1.429	0.03680922	6	expressed protein
LOC_Os01g09450	Os01g0190300	-1.428	0.00011617	0	OsIAA2 - Auxin-responsive Aux/IAA gene family member, expressed
LOC_Os01g64110	Os01g0860500	-1.426	0.00299046	0	glycosyl hydrolase, putative, expressed
LOC_Os01g51320	Os01g0710200	-1.422	0.00968401	3	peroxisomal N-acetyl-spermine/spermidine oxidase precursor, putative, expressed
LOC_Os01g55610	Os01g0761500	-1.421	0.02485402	1	peptide transporter PTR2, putative, expressed
LOC_Os03g49350	Os03g0700400	-1.419	0.0138223	1	lipozygenase protein, putative, expressed
LOC_Os01g13930	Os01g0241000	-1.414	0.00882829	0	expressed protein
LOC_Os11g34020	Os11g0542100	-1.411	0.01468822	0	protein binding protein, putative, expressed
LOC_Os10g38040	Os10g0524300	-1.409	0.006112	1	lysM domain containing protein, putative, expressed
LOC_Os01g19170	Os01g0296200	-1.408	0.00371513	1	polygalacturonase, putative, expressed
LOC_Os02g27000	Os02g0469300	-1.408	0.00448697	0	ATP-binding region, ATPase-like domain containing protein, expressed
LOC_Os08g33740	Os08g0434500	-1.406	0.01112561	0	CSLA11 - cellulose synthase-like family A, expressed
LOC_Os06g11610	Os06g0219500	-1.397	0.00528953	3	heat shock 22 kDa protein, mitochondrial precursor, putative, expressed
LOC_Os09g23300	Os09g0396900	-1.395	0.04745705	4	integral membrane protein, putative, expressed
LOC_Os01g04050	Os01g0132000	-1.384	0.0015044	2	BBTI12 - Bowman-Birk type bran trypsin inhibitor precursor, expressed
LOC_Os03g44380	Os03g0645900	-1.382	0.00156606	3	9-cis-epoxycarotenoid dioxygenase 1, chloroplast precursor, putative, expressed
LOC_Os03g61780	Os03g0833600	-1.38	0.01132545	0	glucan endo-1,3-beta-glucosidase-related, putative, expressed
LOC_Os07g09710	Os07g0195300	-1.379	0.01642916	0	OsFBX220 - F-box domain containing protein, expressed
LOC_Os03g55760	Os03g0766500	-1.379	0.03895648	0	MYB family transcription factor, putative, expressed
LOC_Os01g49830	Os01g0693400	-1.378	0.00132332	0	B3 DNA binding domain containing protein, expressed
LOC_Os03g11490	Os03g0213600	-1.378	0.01929903	0	expressed protein
LOC_Os11g05690	Os11g0155500	-1.377	0.01103768	1	amino acid permease family protein, putative, expressed
LOC_Os03g12730	Os03g0228800	-1.375	0.0042299	0	receptor protein kinase CLAVATA1 precursor, putative, expressed
LOC_Os07g28890	Os07g0471900	-1.375	0.02534073	0	ethylene-responsive protein related, putative, expressed
LOC_Os04g40100	Os04g0477000	-1.367	0.00408294	0	BTBN11 - Bric-a-Brac, Tramtrack, Broad Complex BTB domain with non-phototropic hypocotyl 3 NPH3 domain, expressed
LOC_Os04g39650	Os04g0472500	-1.366	0.04174739	0	receptor protein kinase CLAVATA1 precursor, putative, expressed
LOC_Os03g15000	Os03g0254900	-1.364	8.5724E-05	0	Zinc finger, C3HC4 type domain containing protein, expressed
LOC_Os03g51610	Os03g0726200	-1.363	0.00273536	0	Inositol 1, 3, 4-trisphosphate 5/6-kinase, putative, expressed

LOC_Os03g10640	Os03g0203700	-1.362	0.00049251	2	calcium-transporting ATPase, plasma membrane-type, putative, expressed
LOC_Os10g40360	Os10g0550900	-1.362	0.01760127	0	proline oxidase, mitochondrial precursor, putative, expressed
LOC_Os10g42430	Os10g0575000	-1.361	0.00025165	1	transcription factor MYC7E, putative, expressed
LOC_Os06g34730	Os06g0538400	-1.36	0.00613997	9	expressed protein
LOC_Os06g41710	Os06g0622000	-1.354	0.00187604	0	CW-type Zinc Finger, putative, expressed
LOC_Os03g49940	Os03g0707200	-1.347	0.00391367	2	integral membrane protein, putative, expressed
LOC_Os03g10800	Os03g0205700	-1.347	0.02626781	0	BTBN4 - Bric-a-Brac, Tramtrack, Broad Complex BTB domain with non-phototropic hypocotyl 3 NPH3 domain, expressed
LOC_Os01g56810	Os01g0775400	-1.346	0.00480689	3	cytokinin dehydrogenase precursor, putative, expressed
LOC_Os01g68650	Os01g0915000	-1.341	0.00979938	1	plant-specific domain TIGR01615 family protein, expressed
LOC_Os02g47810	Os02g0707200	-1.341	0.03436306	1	dof zinc finger domain containing protein, putative, expressed
LOC_Os09g21370	Os09g0381400	-1.339	0.04219844	1	cysteine proteinase EP-B 1 precursor, putative, expressed
LOC_Os12g43350	Os12g0628300	-1.338	0.00795992	0	expressed protein
LOC_Os09g35890	Os09g0528000	-1.337	0.04756435	2	kinesin motor domain containing protein, expressed
LOC_Os03g02280	Os03g0113900	-1.334	0.04112378	0	DUF584 domain containing protein, putative, expressed
LOC_Os03g57880	Os03g0792800	-1.325	0.00062605	1	glucan endo-1,3-beta-glucosidase precursor, putative, expressed
LOC_Os04g45720	Os04g0540600	-1.325	0.02272577	3	aldehyde dehydrogenase, putative, expressed
LOC_Os04g33390	Os04g0406600	-1.324	0.00031667	1	prephenate dehydratase domain containing protein, expressed
LOC_Os04g40560	Os04g0481600	-1.324	0.04052261	0	WD domain, G-beta repeat domain containing protein, expressed
LOC_Os03g03200	Os03g0123800	-1.321	0.00844813	4	hydrolase, alpha/beta fold family protein, putative, expressed
LOC_Os10g35460	Os10g0497700	-1.32	0.00158992	3	COBRA, putative, expressed
LOC_Os03g10090	Os03g0197100	-1.317	0.00029337	1	transporter family protein, putative, expressed
LOC_Os11g12740	Os11g0235200	-1.316	0.00615769	0	peptide transporter PTR2, putative, expressed
LOC_Os02g45200	Os02g0673700	-1.315	0.00376404	1	dof zinc finger domain containing protein, putative, expressed
LOC_Os03g57080	Os03g0784100	-1.315	0.01463078	2	PLA IIIA/PLP7, putative, expressed
LOC_Os03g12820	Os03g0230300	-1.313	0.02458969	1	ATP8, putative, expressed
LOC_Os01g57050	Os01g0778500	-1.313	0.02615537	3	expressed protein
LOC_Os07g17330	Os07g0274700	-1.31	0.04112378	1	B12D protein, putative, expressed
LOC_Os11g07600	Os11g0177400	-1.295	0.01838517	4	ABC-2 type transporter domain containing protein, expressed
LOC_Os07g42370	Os07g0615200	-1.291	0.00080472	0	zinc-finger protein, putative, expressed
LOC_Os05g12320	Os05g0214300	-1.29	0.02009043	0	nodulin MtN3 family protein, putative, expressed
LOC_Os04g45270	Os04g0535200	-1.287	0.00257751	0	aspartyl protease family protein, putative, expressed
LOC_Os03g02460	Os03g0115700	-1.286	0.0177222	0	retinol dehydrogenase, putative, expressed
LOC_Os02g46100	Os02g0686100	-1.284	0.01012594	1	RING-H2 finger protein, putative, expressed
LOC_Os04g01980	Os04g0110400	-1.274	0.02700102	0	receptor protein kinase, putative, expressed
LOC_Os02g03750	Os02g0130200	-1.272	0.04347704	0	polygalacturonase, putative, expressed

LOC_Os04g44354	Os04g0525100	-1.269	0.00624989	0	UDP-glucoronosyl and UDP-glucosyl transferase domain containing protein, expressed
LOC_Os05g37190	Os05g0444200	-1.268	0.01297359	0	ZOS5-08 - C2H2 zinc finger protein, expressed
LOC_Os03g16500	Os03g0271900	-1.268	0.04932644	0	aspartic proteinase nepenthesin precursor, putative, expressed
LOC_Os07g46840	Os07g0663600	-1.267	0.04002267	0	oxidoreductase, short chain dehydrogenase/reductase family domain containing protein, expressed
LOC_Os04g52320	Os04g0613200	-1.252	0.03116323	0	QRT3, putative, expressed
LOC_Os05g46750	Os05g0545300	-1.25	0.03645813	0	STE_MEKK_ste11_MAP3K.18 - STE kinases include homologs to sterile 7, sterile 11 and sterile 20 from yeast, expressed
LOC_Os06g10650	Os06g0208700	-1.249	0.00969432	0	tyrosine phosphatase family protein, putative, expressed
LOC_Os01g67540	Os01g0901600	-1.249	0.0340884	1	AMP-binding domain containing protein, expressed
LOC_Os12g40900	Os12g0601400	-1.246	0.00343826	0	OsIAA31 - Auxin-responsive Aux/IAA gene family member, expressed
LOC_Os01g70820	Os01g0934400	-1.239	0.00594742	0	lumenal PsbP, putative, expressed
LOC_Os06g03830	Os06g0129100	-1.235	0.03139264	0	retinol dehydrogenase, putative, expressed
LOC_Os10g34730	Os10g0489000	-1.232	0.04186899	0	GEM, putative, expressed
LOC_Os05g32440	Os05g0390600	-1.23	0.02841961	0	expressed protein
LOC_Os04g40630	Os04g0482300	-1.228	0.00588166	1	BTBZ4 - Bric-a-Brac, Tramtrack, Broad Complex BTB domain with TAZ zinc finger and Calmodulin-binding domains, expressed
LOC_Os03g45170	Os03g0654400	-1.226	0.00316174	1	amino acid permease, putative, expressed
LOC_Os06g12150	Os06g0225800	-1.226	0.02341263	1	shikimate kinase, putative, expressed
LOC_Os03g51650	Os03g0726500	-1.22	0.00216615	0	membrane protein, putative, expressed
LOC_Os10g37340	Os10g0517500	-1.219	0.00968401	0	cystathionine gamma-synthase, putative, expressed
LOC_Os06g44010	Os06g0649000	-1.218	0.00299014	1	WRKY28, expressed
LOC_Os11g12620	Os11g0233800	-1.215	0.02383758	0	receptor protein kinase CLAVATA1 precursor, putative, expressed
LOC_Os10g25290	Os10g0392400	-1.214	0.00203987	1	ZIM domain containing protein, putative, expressed
LOC_Os06g34830	Os06g0539400	-1.213	0.04933725	20	amino acid permease family protein, putative, expressed
LOC_Os03g59670	Os03g0811400	-1.212	0.03943187	0	basic helix-loop-helix, putative, expressed
LOC_Os03g31690	Os03g0431200	-1.209	0.00220618	5	GCN5-related N-acetyltransferase, putative, expressed
LOC_Os04g41970	Os04g0497200	-1.208	0.03522271	1	endoglucanase, putative, expressed
LOC_Os01g42380	Os01g0609300	-1.206	0.02270464	1	pleiotropic drug resistance protein, putative, expressed
LOC_Os04g59190	Os04g0688300	-1.205	0.04140776	2	peroxidase precursor, putative, expressed
LOC_Os01g07140	Os01g0165200	-1.202	0.02736475	6	kelch repeat-containing protein, putative, expressed
LOC_Os04g36040	Os04g0441800	-1.201	0.01855985	1	peptide transporter PTR2, putative, expressed
LOC_Os01g57420	Os01g0783200	-1.2	0.00108815	4	diacylglycerol kinase, putative, expressed
LOC_Os11g40340	Os11g0618000	-1.199	0.04496639	0	expressed protein
LOC_Os04g54564	Os04g0638300	-1.197	0.01279473	0	expressed protein
LOC_Os07g47960	Os07g0676600	-1.196	0.03228906	0	basic helix-loop-helix domain containing protein, expressed

LOC_Os05g02450	Os05g0115100	-1.195	0.02652911	0	expressed protein
LOC_Os01g50080	Os01g0695700	-1.195	0.03462165	0	MDR-like ABC transporter, putative, expressed
LOC_Os04g50216	Os04g0592600	-1.192	0.00109255	0	SNARE associated Golgi protein, putative, expressed
LOC_Os03g61960	Os03g0835900	-1.188	0.00146705	2	2Fe-2S iron-sulfur cluster binding domain containing protein, expressed
LOC_Os07g44780	Os07g0642200	-1.185	0.00968401	0	GDSL-like lipase/acylhydrolase, putative, expressed
LOC_Os08g33160	Os08g0428400	-1.182	0.00464621	0	ZIM motif family protein, expressed
LOC_Os01g20780	Os01g0309100	-1.182	0.03589501	5	anther-specific protein SF18 precursor, putative, expressed
LOC_Os05g49140	Os05g0566400	-1.179	0.00237247	2	CGMC_MAPKCMGC_2.8 - CGMC includes CDA, MAPK, GSK3, and CLKC kinases, expressed
LOC_Os04g32980	Os04g0402500	-1.175	0.00620404	1	protein binding protein, putative, expressed
LOC_Os07g02970	Os07g0121000	-1.175	0.02270464	5	expressed protein
LOC_Os04g53990	Os04g0631600	-1.173	0.00997952	3	ethylene-responsive protein related, putative, expressed
LOC_Os05g30454	Os05g0367400	-1.173	0.04002267	1	thiamin pyrophosphokinase 1, putative, expressed
LOC_Os09g28210	Os09g0455300	-1.169	0.03762798	0	bHelix-loop-helix transcription factor, putative, expressed
LOC_Os03g02685	Os03g0118450	-1.164	0.04199132	1	lysM domain containing protein, putative, expressed
LOC_Os01g15610	Os01g0260800	-1.162	0.02230605	1	expressed protein
LOC_Os01g54670	Os01g0750500	-1.161	0.01478838	1	coiled-coil domain-containing protein 25, putative, expressed
LOC_Os05g24650	Os05g0311500	-1.16	0.03733384	0	DUF567 domain containing protein, putative, expressed
LOC_Os08g41780	Os08g0529800	-1.151	0.02650701	0	triacylglycerol lipase precursor, putative, expressed
LOC_Os09g32960	Os09g0507400	-1.149	0.00941783	0	expressed protein
LOC_Os10g41550	Os10g0565200	-1.139	0.0160699	0	beta-amylase, putative, expressed
LOC_Os07g14150	Os07g0245100	-1.136	0.01019917	0	cytidine deaminase, putative, expressed
LOC_Os01g63480	Os01g0854000	-1.136	0.018248	0	transferase family protein, putative, expressed
LOC_Os02g40664	Os02g0619600	-1.136	0.0274084	0	zinc finger family protein, putative, expressed
LOC_Os10g31950	Os10g0457600	-1.135	0.0064336	1	3-ketoacyl-CoA thiolase, peroxisomal precursor, putative, expressed
LOC_Os07g34390	Os07g0527800	-1.133	0.02934132	0	ankyrin repeat family protein, putative, expressed
LOC_Os01g62500	Os01g0842600	-1.132	0.02270464	1	OsFtsH3 FtsH protease, homologue of AtFtsH3/10, expressed
LOC_Os01g66010	Os01g0882800	-1.132	0.04634677	0	amino acid transporter, putative, expressed
LOC_Os01g06590	Os01g0159300	-1.121	0.00155453	0	zinc finger, C3HC4 type domain containing protein, expressed
LOC_Os04g32030	Os04g0390100	-1.12	0.00967068	1	heavy metal-associated domain containing protein, expressed
LOC_Os08g04470	Os08g0139400	-1.117	0.02797417	0	U box protein 8, putative, expressed
LOC_Os01g51920	Os01g0717000	-1.115	0.00541949	1	phosphotransferase, putative, expressed
LOC_Os07g38290	Os07g0570600	-1.113	0.04269052	1	plastocyanin-like domain containing protein, putative, expressed
LOC_Os09g25060	Os09g0417600	-1.111	0.01374276	5	WRKY76, expressed
LOC_Os03g10080	Os03g0197000	-1.111	0.01749839	1	expressed protein
LOC_Os02g34410	Os02g0548700	-1.108	0.01311003	0	U-box domain-containing protein, putative, expressed

LOC_Os02g55560	Os02g0799000	-1.107	0.00811629	0	protein phosphatase 2C, putative, expressed
LOC_Os10g22730	Os10g0372800	-1.1	0.04151903	0	expressed protein
LOC_Os07g46350	Os07g0656900	-1.099	0.00535833	0	OsSCP40 - Putative Serine Carboxypeptidase homologue, expressed
LOC_Os03g50970	Os03g0719000	-1.098	0.01398253	0	retrotransposon protein, putative, Ty1-copia subclass, expressed
LOC_Os05g37470	Os05g0447200	-1.092	0.00867746	0	transmembrane amino acid transporter protein, putative, expressed
LOC_Os06g49890	Os06g0713100	-1.091	0.00614683	0	coiled-coil domain-containing protein 90A, mitochondrial precursor, putative, expressed
LOC_Os04g41160	Os04g0488700	-1.08	0.00709091	1	protein kinase, putative, expressed
LOC_Os02g45870	Os02g0683900	-1.079	0.0102525	0	tobamovirus multiplication protein, putative, expressed
LOC_Os09g29490	Os09g0471100	-1.078	0.03904576	0	peroxidase precursor, putative, expressed
LOC_Os05g07220	Os05g0164900	-1.074	0.01554951	0	kelch repeat-containing protein, putative, expressed
LOC_Os10g37190	Os10g0516200	-1.069	0.0036906	0	protein kinase domain containing protein, expressed
LOC_Os01g71320	Os01g0940100	-1.068	0.00968401	1	hexokinase, putative, expressed
LOC_Os11g36200	Os11g0570000	-1.066	0.03564617	4	receptor-like protein kinase 2 precursor, putative, expressed
LOC_Os01g12840	Os01g0228800	-1.055	0.00991059	0	expressed protein
LOC_Os09g12660	Os09g0298200	-1.052	0.01403121	0	glucose-1-phosphate adenylyltransferase large subunit, chloroplast precursor, putative, expressed
LOC_Os01g53710	Os01g0739200	-1.05	0.00495914	0	dual specificity protein phosphatase, putative, expressed
LOC_Os05g29030	Os05g0358500	-1.05	0.01341637	0	protein phosphatase 2C, putative, expressed
LOC_Os01g43480	Os01g0623500	-1.048	0.01520455	2	AAA-type ATPase family protein, putative, expressed
LOC_Os03g46440	Os03g0667100	-1.036	0.00655897	1	BTBA4 - Bric-a-Brac,Tramtrack, Broad Complex BTB domain with Ankyrin repeat region, expressed
LOC_Os02g44880	Os02g0669500	-1.032	0.01290318	2	expressed protein
LOC_Os01g51540	Os01g0712700	-1.031	0.01687998	0	cytidine/deoxycytidylate deaminase, putative, expressed
LOC_Os02g52780	Os02g0766700	-1.026	0.00579053	3	bZIP transcription factor, putative, expressed
LOC_Os02g43700	Os02g0653900	-1.026	0.01685067	0	triacylglycerol lipase like protein, putative, expressed
LOC_Os01g11340	Os01g0211600	-1.025	0.03354676	4	cytochrome P450, putative, expressed
LOC_Os03g45280	Os03g0655400	-1.02	0.02747393	0	dehydrin, putative, expressed
LOC_Os07g43316	Os07g0626200	-1.014	0.03354676	0	SAP domain-containing protein, putative, expressed
LOC_Os03g38970	Os03g0586700	-1.006	0.01174955	0	metal ion binding protein, putative, expressed
LOC_Os01g21250	Os01g0314800	-0.997	0.03739513	0	late embryogenesis abundant protein, putative, expressed
LOC_Os08g33720	Os08g0434300	-0.995	0.01114853	0	lactate/malate dehydrogenase, putative, expressed
LOC_Os04g17100	Os04g0244800	-0.992	0.03157175	5	heavy metal-associated domain containing protein, expressed
LOC_Os01g34480	Os01g0528800	-0.989	0.01651539	0	NAD dependent epimerase/dehydratase family protein, putative, expressed
LOC_Os01g15300	Os01g0256800	-0.985	0.02478847	0	zinc finger helicase family protein, putative, expressed
LOC_Os03g13870	Os03g0242300	-0.984	0.02454545	1	expressed protein
LOC_Os01g60600	Os01g0821300	-0.979	0.02175697	1	WRKY108, expressed

LOC_Os05g50890	Os05g0586200	-0.978	0.02181384	0	OsGH3.5 - Probable indole-3-acetic acid-amido synthetase, expressed
LOC_Os05g48040	Os05g0554000	-0.978	0.03163224	0	MATE efflux family protein, putative, expressed
LOC_Os06g51360	Os06g0729900	-0.977	0.01627063	0	lysM domain containing protein, putative, expressed
LOC_Os04g34170	Os04g0418800	-0.974	0.03815383	0	retrotransposon protein, putative, unclassified, expressed
LOC_Os01g68300	Os01g0910900	-0.972	0.02623076	2	expressed protein
LOC_Os07g07040	Os07g0164800	-0.971	0.02596806	9	erythrocyte binding protein 3, putative, expressed
LOC_Os09g36860	Os09g0539800	-0.966	0.03139264	1	acyl carrier protein, putative, expressed
LOC_Os12g08130	Os12g0181600	-0.964	0.03992053	0	amino acid transporter, putative, expressed
LOC_Os03g15460	Os03g0261100	-0.956	0.03501595	0	expressed protein
LOC_Os09g38040	Os09g0553300	-0.954	0.04906376	6	hydrolase, NUDIX family protein, expressed
LOC_Os07g37320	Os07g0559700	-0.953	0.02080544	0	transporter family protein, putative, expressed
LOC_Os08g38910	Os08g0498400	-0.952	0.03571285	1	caffeoyl-CoA O-methyltransferase, putative, expressed
LOC_Os01g16170	Os01g0266800	-0.95	0.03376178	0	PQ loop repeat domain containing protein, expressed
LOC_Os02g48560	Os02g0716500	-0.949	0.02241163	2	fatty acid desaturase, putative, expressed
LOC_Os02g08420	Os02g0180700	-0.949	0.03346952	0	cinnamoyl CoA reductase, putative, expressed
LOC_Os03g11900	Os03g0218400	-0.947	0.03571285	2	transporter family protein, putative, expressed
LOC_Os11g10760	Os11g0213700	-0.947	0.04425736	0	NBS-LRR disease resistance protein, putative, expressed
LOC_Os05g46460	Os05g0542200	-0.94	0.01419859	1	hydrolase, alpha/beta fold family domain containing protein, expressed
LOC_Os03g12390	Os03g0225100	-0.937	0.01846616	15	STE_MEK_ste7_MAP2K.6 - STE kinases include homologs to sterile 7, sterile 11 and sterile 20 from yeast, expressed
LOC_Os06g14370	Os06g0254700	-0.937	0.03904576	0	caleosin related protein, putative, expressed
LOC_Os05g16824	Os05g0257100	-0.931	0.04473431	0	SHR5-receptor-like kinase, putative, expressed
LOC_Os09g19640	Os09g0360400	-0.93	0.03704143	1	holocarboxylase synthetase, putative, expressed
LOC_Os08g36630	Os08g0470200	-0.929	0.02212521	0	bifunctional monodehydroascorbate reductase and carbonic anhydrasenectarin-3 precursor, putative, expressed
LOC_Os01g65460	Os01g0875500	-0.924	0.04906376	0	beta-galactosidase precursor, putative, expressed
LOC_Os01g15270	Os01g0256500	-0.921	0.04387977	0	expressed protein
LOC_Os08g05910	Os08g0155400	-0.912	0.03839376	0	peptide transporter PTR2, putative, expressed
LOC_Os03g14050	Os03g0244200	-0.911	0.03101115	1	thaumatin-like protein 1 precursor, putative, expressed
LOC_Os04g53620	Os04g0628100	-0.904	0.04118347	6	ubiquitin family protein, putative, expressed
LOC_Os03g11650	Os03g0215700	-0.901	0.03981265	0	ATMAP70 protein, putative, expressed
LOC_Os04g45900	Os04g0542800	-0.898	0.04779946	0	transposon protein, putative, unclassified, expressed
LOC_Os01g13520	Os01g0236300	-0.893	0.02933099	1	auxin response factor 1, putative, expressed
LOC_Os03g08310	Os03g0180800	-0.893	0.04834301	0	ZIM domain containing protein, putative, expressed
LOC_Os01g50310	Os01g0698000	-0.879	0.04322108	0	VIP1 protein, putative, expressed
LOC_Os04g46280	Os04g0547900	-0.877	0.02971114	0	hydrolase, NUDIX family, domain containing protein, expressed

LOC_Os10g39140	Os10g0536400	-0.876	0.02967275	0	flavonol synthase/flavanone 3-hydroxylase, putative, expressed
LOC_Os01g02940	Os01g0119100	-0.87	0.02778032	1	glycosyltransferase protein, putative, expressed
LOC_Os02g54600	Os02g0787300	-0.869	0.03826466	0	STE_MEK_ste7_MAP2K.5 - STE kinases include homologs to sterile 7, sterile 11 and sterile 20 from yeast, expressed
LOC_Os03g08320	Os03g0180900	-0.869	0.03876251	1	ZIM domain containing protein, putative, expressed
LOC_Os01g50760	Os01g0703400	-0.863	0.03839376	0	polyprenyl synthetase, putative, expressed
LOC_Os07g39310	Os07g0581400	-0.862	0.04489348	2	ZOS7-06 - C2H2 zinc finger protein, expressed
LOC_Os01g65920	Os01g0881900	-0.835	0.04602205	0	F-box/LRR-repeat protein 2, putative, expressed
LOC_Os02g04010	Os02g0133000	-0.824	0.04906376	2	cyclin-T1-1, putative, expressed

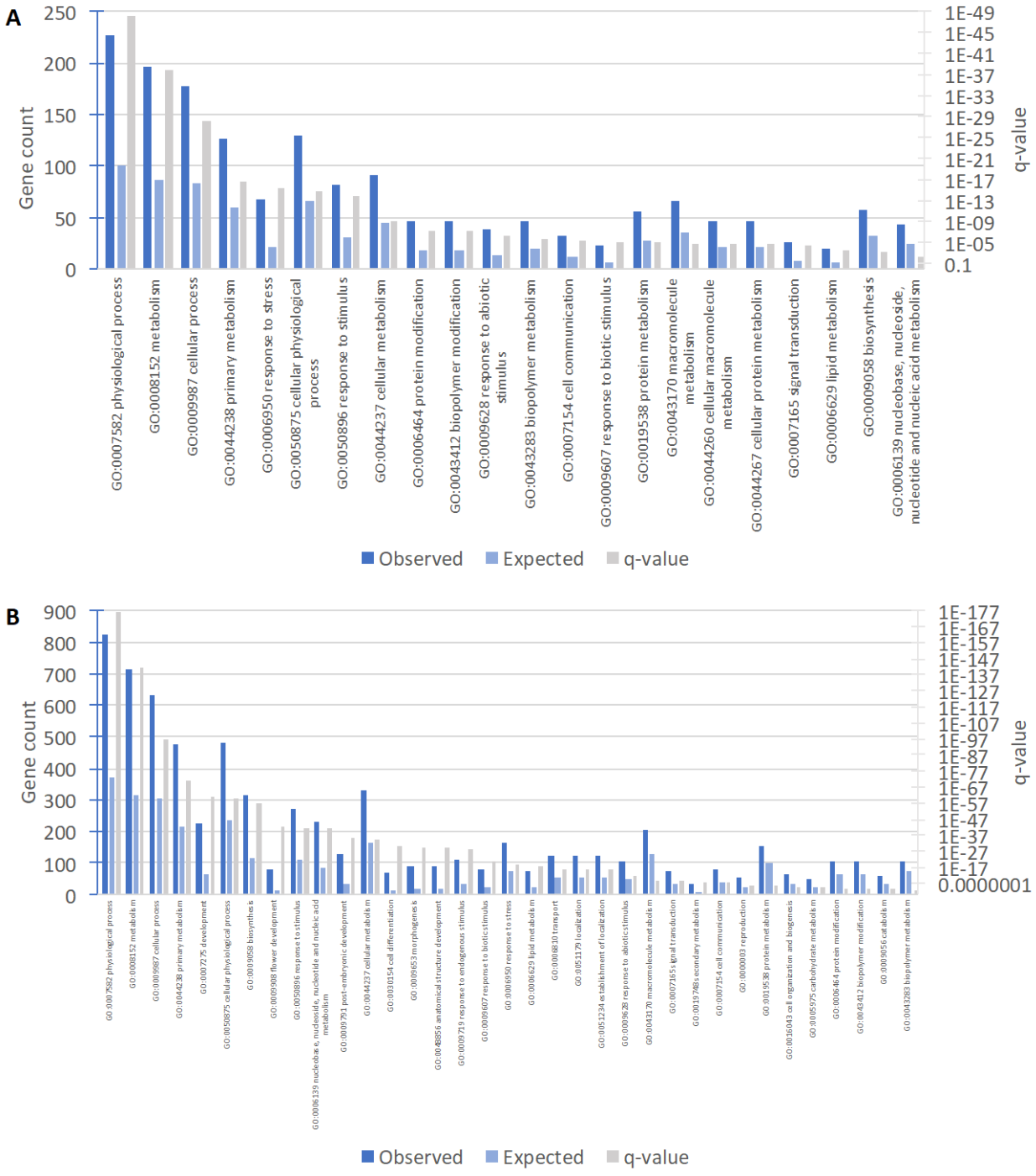


Figure B.4: RiceNETDB biological process gene ontology enrichment analysis from DEG from 42 days post-planting tiller tissue.
(A) up-regulated DEG, (B) down-regulated DEG. RiceNETDB accounts for genes including their differential splice isoforms.

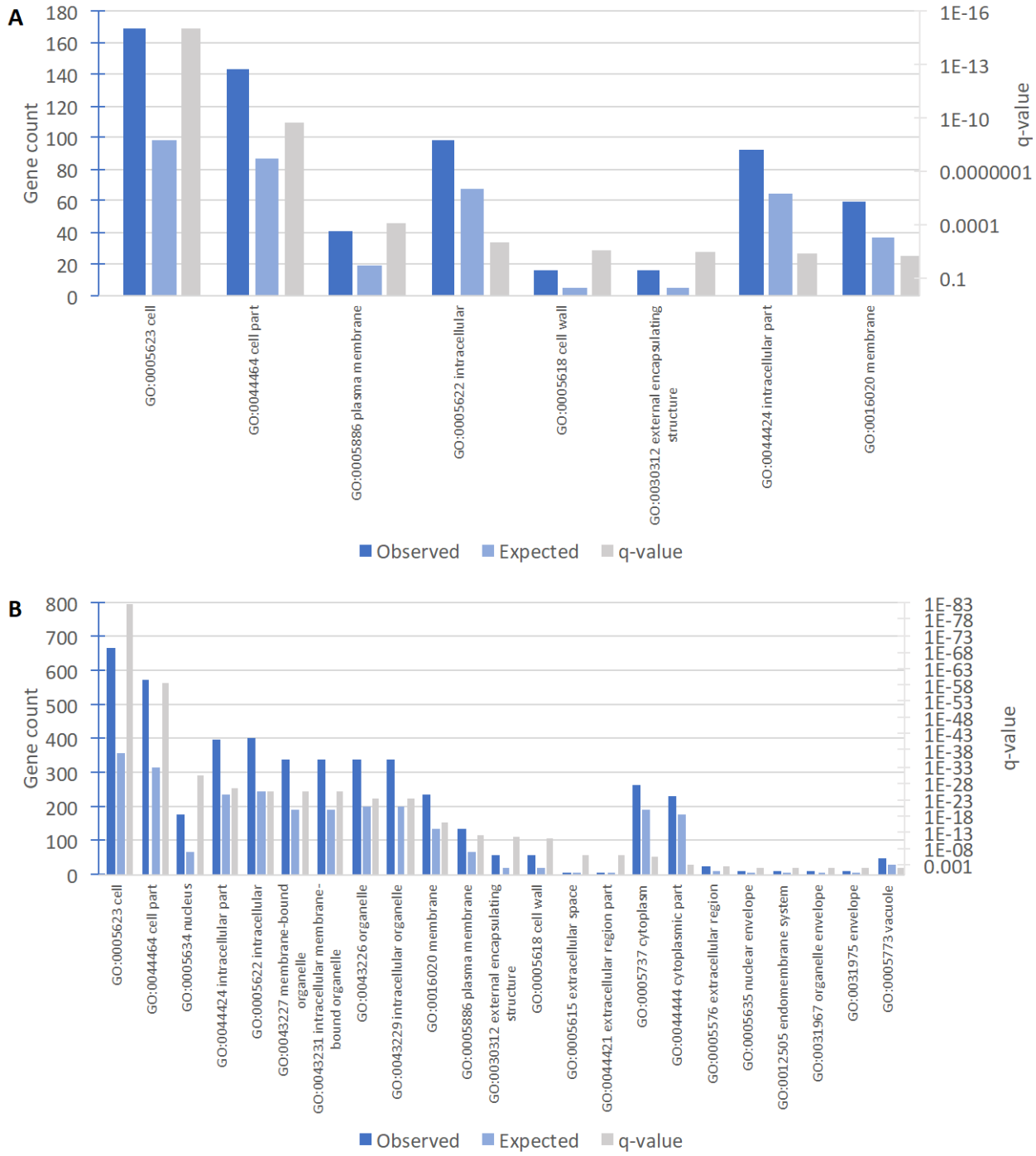


Figure B.5: RiceNETDB cellular localization gene ontology enrichment analysis from DEG from 42 days post-planting tiller tissue.
(A) up-regulated DEG, (B) down-regulated DEG. RiceNETDB accounts for genes including their differential splice isoforms.

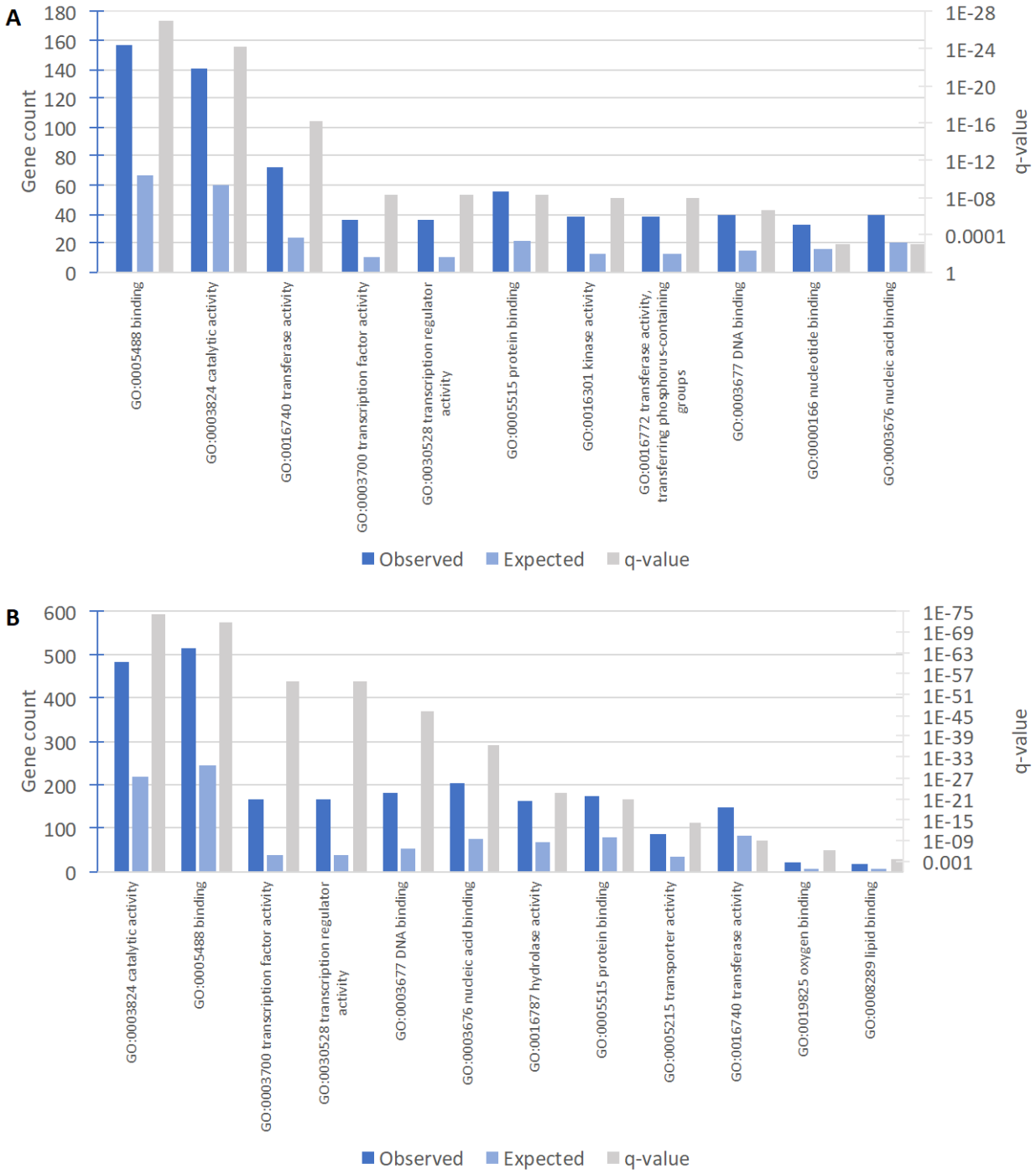


Figure B.6: RiceNETDB molecular function gene ontology enrichment analysis from DEG from 42 days post-planting tiller tissue.
(A) up-regulated DEG, (B) down-regulated DEG. RiceNETDB accounts for genes including their differential splice isoforms.

Table B.2: DEG that were inversely significantly expressed between 32 and 42 days post-planting.

MSU ID	RAP ID	log2 Fold Change (32dpp, 42dpp)	Description
Up-regulated to down-regulated			
LOC_Os03g08310	Os03g0180800	3.476, -0.893	ZIM domain containing protein, putative, expressed
LOC_Os01g60600	Os01g0821300	2.828, -0.979	WRKY108, expressed
LOC_Os04g23550	Os04g0301500	2.828, -2.432	basic helix-loop-helix family protein, putative, expressed
LOC_Os08g28710	Os08g0374600	2.629, -2.226	receptor protein kinase CRINKLY4 precursor, putative, expressed
LOC_Os03g08520	Os03g0183500	2.436, -1.468	DUF581 domain containing protein, expressed
LOC_Os01g64360	Os01g0863300	2.409, -1.969	MYB family transcription factor, putative, expressed
LOC_Os10g25290	Os10g0392400	2.357, -1.214	ZIM domain containing protein, putative, expressed
LOC_Os01g18120	Os01g0283700	2.35, -2.168	cinnamoyl CoA reductase, putative, expressed
LOC_Os05g44060	Os05g0516700	2.151, -1.609	expressed protein
LOC_Os05g12040	Os05g0211100	2.092, -3.336	cytochrome P450 51, putative, expressed
LOC_Os03g08320	Os03g0180900	1.818, -0.869	ZIM domain containing protein, putative, expressed
LOC_Os07g46920	NONE	1.727, -3.077	sex determination protein tasselseed-2, putative, expressed
LOC_Os07g34260	Os07g0526400	1.648, -2.886	chalcone and stilbene synthases, putative, expressed
LOC_Os07g09190	Os07g0190000	1.554, -3.522	transketolase, putative, expressed
LOC_Os11g45740	Os11g0684000	1.495, -2.661	MYB family transcription factor, putative, expressed
LOC_Os02g02930	Os02g0121700	1.482, -3.037	terpene synthase, putative, expressed
LOC_Os02g48770	Os02g0719600	1.417, -2.553	SAM dependent carboxyl methyltransferase, putative, expressed
LOC_Os01g56810	Os01g0775400	1.379, -1.346	cytokinin dehydrogenase precursor, putative, expressed
LOC_Os06g10210	Os06g0203600	1.366, -1.884	expressed protein
LOC_Os04g33390	Os04g0406600	1.148, -1.324	prephenate dehydratase domain containing protein, expressed
LOC_Os03g28940	Os03g0402800	1.134, -1.768	ZIM domain containing protein, putative, expressed
LOC_Os04g15920	Os04g0229100	1.104, -2.168	dehydrogenase, putative, expressed
LOC_Os03g58290	Os03g0797300	1.082, -3.033	indole-3-glycerol phosphate lyase, chloroplast precursor, putative, expressed
LOC_Os10g35460	Os10g0497700	0.972, -1.32	COBRA, putative, expressed
LOC_Os01g61850	NONE	0.905, -2.239	NONE
LOC_Os04g10350	Os04g0182200	0.862, -2.122	1-aminocyclopropane-1-carboxylate oxidase homolog 2, putative, expressed
LOC_Os04g50216	Os04g0592600	0.764, -1.192	SNARE associated Golgi protein, putative, expressed
Down-regulated to up-regulated			
LOC_Os05g04530	Os05g0135900	-0.916, 1.113	IF, putative, expressed

Discussion

The intent of this supplementary data was to provide insight, disclosure, and transparency to the RNA-seq experiment used in association with the results of the 32 days post-planting RNA-seq analysis (Chapter 4). The results here indicate that although treatment groups are likely under different phases of development, samples appear relatively similar within treatment groups showing strong consistency and low levels of variability. A list of inversely regulated genes between 32 and 42 days post-planting is provided as these genes show significantly altered regulation between these two experiments, and might illicit further investigation pertaining to formation of the phenotype or developmental transition. The genes that are differentially expressed affect a variety of ontological functions. Speculation on the outcome of ontological enrichment categories remains challenging, however might be skillfully applied or further dissected during future investigation of *mpg1*.

BIBLIOGRAPHY

ANDERS, S. & HUBER, W. 2010. Differential expression analysis for sequence count data. *Genome Biol,* 11**,** R106.

ANDERS, S., PYL, P. T. & HUBER, W. 2015. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics,* 31**,** 166-9.

ANDREWS, S. 2010. FastQC: a quality control tool for high throughput sequence data. *Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc*.

DOBIN, A., DAVIS, C. A., SCHLESINGER, F., DRENKOW, J., ZALESKI, C., JHA, S., BATUT, P., CHAISSON, M. & GINGERAS, T. R. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics,* 29**,** 15-21.

HRUZ, T., LAULE, O., SZABO, G., WESSENDORP, F., BLEULER, S., OERTLE, L., WIDMAYER, P., GRUISSEM, W. & ZIMMERMANN, P. 2008. Genevestigator v3: a reference expression database for the meta-analysis of transcriptomes. *Adv Bioinformatics,* 2008**,** 420747.

HULSEN, T., DE VLIEG, J. & ALKEMA, W. 2008. BioVenn - a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. *BMC Genomics,* 9**,** 488.

ILLUMINA-BASESPACE-LABS 2018. FASTQ Toolkit. *Available online at: https://www.illumina.com/products/by-type/informatics-products/basespace-sequence-hub/apps/fastq-toolkit.html*.

LOVE, M. I., HUBER, W. & ANDERS, S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol,* 15**,** 550.

MERCHANT, N., LYONS, E., GOFF, S., VAUGHN, M., WARE, D., MICKLOS, D. & ANTIN, P. 2016. The iPlant Collaborative: Cyberinfrastructure for Enabling Data to Discovery for the Life Sciences. *PLoS Biol,* 14**,** e1002342.

ROBINSON, M. D., MCCARTHY, D. J. & SMYTH, G. K. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics,* 26**,** 139-40.

TRAPNELL, C., ROBERTS, A., GOFF, L., PERTEA, G., KIM, D., KELLEY, D. R., PIMENTEL, H., SALZBERG, S. L., RINN, J. L. & PACHTER, L. 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc,* 7**,** 562-78.