GENETIC INSIGHTS INTO STEM NON-STRUCTURAL CARBOHYDRATE DYNAMICS IN CULTIVATED ASIAN RICE, ORYZA SATIVA

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Diane Ran Wang

January 2017

GENETIC INSIGHTS INTO STEM NON-STRUCTURAL CARBOHYDRATE DYNAMICS IN CULTIVATED ASIAN RICE, ORYZA SATIVA

Diane Ran Wang, Ph. D.

Cornell University 2017

Rice plants (*Oryza sativa*) accumulate photo-assimilates in the form of non-structural carbohydrates (NSCs) in their stems prior to heading. These can later be mobilized to supplement photosynthate production during grain-filling. There has been longstanding enthusiasm by rice physiologists in optimizing stem NSC as a strategy for rice improvement. Despite this interest, documented since the 1970s and 80s, very little about the genetic controls regulating NSC accumulation, remobilization, and re-accumulation is known. In this dissertation, we first lay the groundwork for large-scale diversity studies on rice stem NSC. We assess the relationship of stem NSC components with 21 agronomic traits in large-scale, tropical yield trials using 33 breeder-nominated lines, establish an appropriate experimental design for future genetic studies using a Bayesian framework to sample sub-datasets from highly-replicated greenhouse data using 36 genetically diverse genotypes, and use 434 phenotypically divergent rice stem samples to develop two partial least squares (PLS) models using near infrared (NIR) spectra for accurate, rapid prediction of rice stem starch, sucrose, and total non-structural carbohydrates. Secondly, we reveal the genetic architecture that underlies stem NSC dynamics in *tropical japonica* rice using a GWAS approach on two panels complemented by a Near-Isogenic Line library

evaluation. Finally, we present preliminary results for a study on the effects of CO_2, temperature, and CO_2 x temperature interaction effects on structural and non-structural carbohydrate constituents in a New Plant Type accession of rice.

BIOGRAPHICAL SKETCH

Diane Wang was born on February 18, 1988 in Beijing, China. As a child she moved frequently, from several places in Beijing to Queens (NYC), Manhattan (NYC), Columbus (Ohio), Springfield (New Jersey) and Millington (New Jersey). She came to Ithaca, NY to attend Cornell University for college and received her B.S. in Biological Sciences and Entomology in 2010. She started graduate school in August 2011 after working as a technician for ten months and traveling around Asia and Europe for three months.

This work is dedicated to my friend and colleague, Namrata ‘Moni’ Singh. Throughout the final edits of this dissertation, I thought often about her excitement for science, her curiosity for learning, and her amazing strength as she fights acute myeloid leukemia.

ACKNOWLEDGMENTS

I thank my mom, Ling Chen, for teaching me the importance of hard work and my dad, George Wang, for giving me his stubbornness and creativity (whether that be due to G, E, and/or GxE effects!). I am grateful to Susan for her mentorship throughout my Ph.D. program, for the seemingly limitless intellectual freedom she entrusted me with, and for all the diverse learning opportunities she offered that were uniquely challenging and always interesting. Many thanks go to past and present members of the McCouch Lab family for their unfailing camaraderie over the years. Collaborators at Cornell, USDA, NREL, and IRRI supported numerous experiments throughout the course of my Ph.D. program and have been instrumental to the completion of this work. My committee members (Tom Owens, Tim Setter, and Margaret Smith) have lent valuable research insight and career advice. The work described here was funded by NSF Graduate Research Fellowship #DGE- 1144153 and USDA NIFA #2014-67003-21858.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

PREFACE

Chapter 2 was accepted as an original research article in Journal of Experimental Botany (Wang 2016, doi: 10.1093/jxb/erw375). Genotyping efforts described in Chapter 3 will provide genetic variation information in U.S. rice to use for modeling genomic signatures of climate resilience on USDA-NIFA funded project #2014-67003-21858 (Characterization of Genomic Signatures for Rice Crop Resilience in Response to Climate Change in the U.S.). Chapter 4 was prepared as a manuscript for submission to PLoS Genetics. Chapter 5 was published as an original research article in Global Change Biology (Wang 2016; doi: 10.1111/gcb.13279). Information found within Appendices A and B will be the basis of a manuscript with co-authors at USDA (Lewis Ziska) and NREL (Ed Wolfrum).

CHAPTER 1:

INTRODUCTION

Yield potential and the elusive role of stem non-structural carbohydrates

Yield potential is a theoretical ideal developed by crop physiologists based on laboratory and field experiments. It has been defined as the yield of a cultivar under adapted conditions with non-limiting nutrients and controlled stresses (refer to Evans and Fischer, 1999). Difficult to conceptualize, and even harder to measure, it has yet become an endlessly sought-after goal that rice (*Oryza sativa* L.) breeders seek to realize. Attaining, breaking, or raising yield potential features prominently as a mission of the Global Rice Science Partnership (GRiSP), and it was at a GRiSP Yield Potential meeting in 2011 where I was challenged with understanding the genetic basis of rice stem non-structural carbohydrate (NSC) variation for my Ph.D. research. Stem NSCs are carbohydrate reserves that are transiently stored in the rice stem for later utilization; their contribution to grain-filling in the panicle has made them a target of long-standing interest of breeders (Cock and Yoshida, 1972; Slewinski, 2012). However, a synthesis of current literature reveals that NSCs are far from a panacea when it comes to raising yield potential in practice. In fact, a nearly identical rendition of the 2011 GRiSP meeting had played out 18 years prior in 1993 at the International Rice Research Institute (IRRI), when stem NSCs were called out, along with a slew of other traits, some classical and others less obvious, as targets for dedicated investigation aimed at "breaking the yield barrier" (Cassman, 1994). It also became apparent that two schools of thought dominated the arena of rice yield potential research, and proponents of "ideotype" versus "empirical" breeding were fervently

divided. Technological advances during those 18 years, especially in the field of genetics and genomics, afforded an interesting opportunity to shed new light on this old topic.

The journey of stem carbon in a rice plant

Carbohydrates are generated in photosynthetic organs of the rice plant, primarily within the three topmost leaves of a rice cultivar (Yoshida, 1981). Prior to heading, some photosynthates are transported to and accumulate within stems, i.e. the collective culm and leaf sheath, for temporary storage. These reserves are non-structural carbohydrates (NSCs) and are found in the form of starch (polymers of glucose) and sucrose (glucose-fructose disaccharides) (Slewinski, 2012). They are remobilized towards panicles after heading, conditional on environment and genotype, providing building blocks for starch synthesis in the filling grains. While the net change of stem NSCs over the course of a rice plant's lifespan is often neatly summarized by a pre-heading accumulation phase followed by a post-heading remobilization phase, and another brief accumulation phase near physiological maturity consistent with *O. sativa's* semi-perennial nature (**Figure 1.1**, Cock and Yoshida, 1972; Van Dat and Peterson, 1983; Takahashi et al., 2005), these overviews do not capture the full complexities of stem NSC dynamics.

Stem NSC levels in rice are mediated by a number of enzymes. Major players include catalysts of starch synthesis/breakdown and sucrose synthesis/breakdown or proteins that function in sucrose translocation. Expression studies have reported correlations of rice stem NSC with changes in timing and levels of transcription for

genes encoding the following proteins: ADP glucose pyrophosphorylase (AGPase) large subunit, granule bound starch synthase (GBSS), soluble starch synthase (SSS), starch branching enzyme (BE), cytosolic fructose-1,6-bisphophatase (cytosolic FBPase) α-amylase, β-amylase, and sucrose transporter (SUT) (Hirose et al., 1999; Hirose et al., 2006; Chen and Wang, 2008).

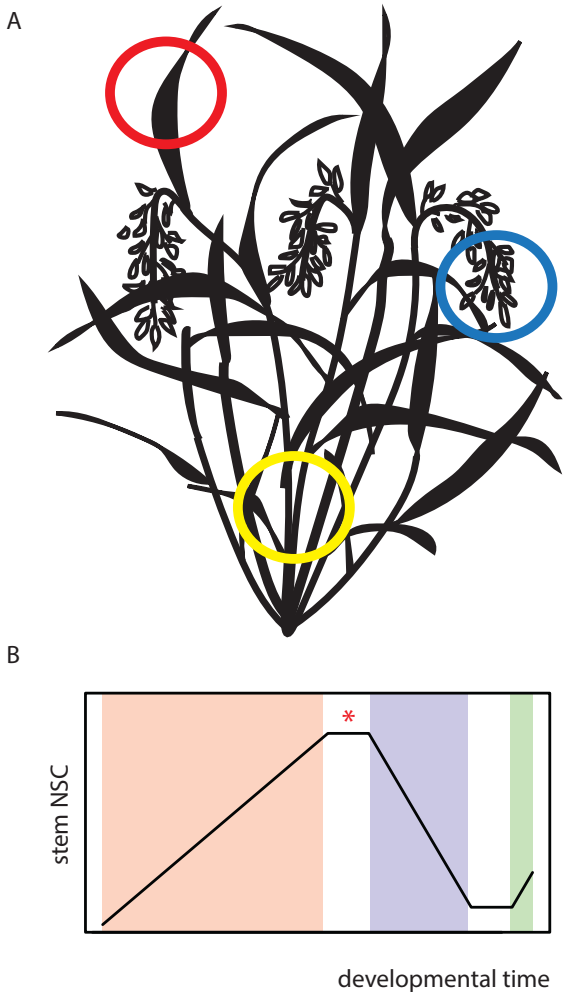


Figure 1.1. Overview of rice stem NSC dynamics. A) Rice plant with three organs of interest circled: red- photosynthetic leaf; yellow- stem; blue- panicle. B) Net changes in stem NSC over time. Red shading indicates period of NSC accumulation in stems, purple shading indicates period of rapid NSC net loss (remobilization to other organs), and green shading indicates end-of-season re-accumulation of stem NSCs. Red asterisk marks heading.

Corresponding studies investigating the change in activity levels according to stem NSC accumulation and remobilization found associations with these enzymes: α-amylase, AGPase, BE, SSS, GBSS, plastidial FBPase, and sucrose synthase (SuSy, SuSase) (Perez et al., 1971; Watanabe et al., 1997; Ishimaru et al., 2004; He et al., 2005; Hirano et al., 2005; Fu et al., 2011). Additional work on a mutant lacking the large subunit of AGPase found drastically reduced starch accumulation in rice culms (Cook et al., 2012). **Figure 1.2** summarizes the genes and enzymes reported to associate with stem NSC levels.

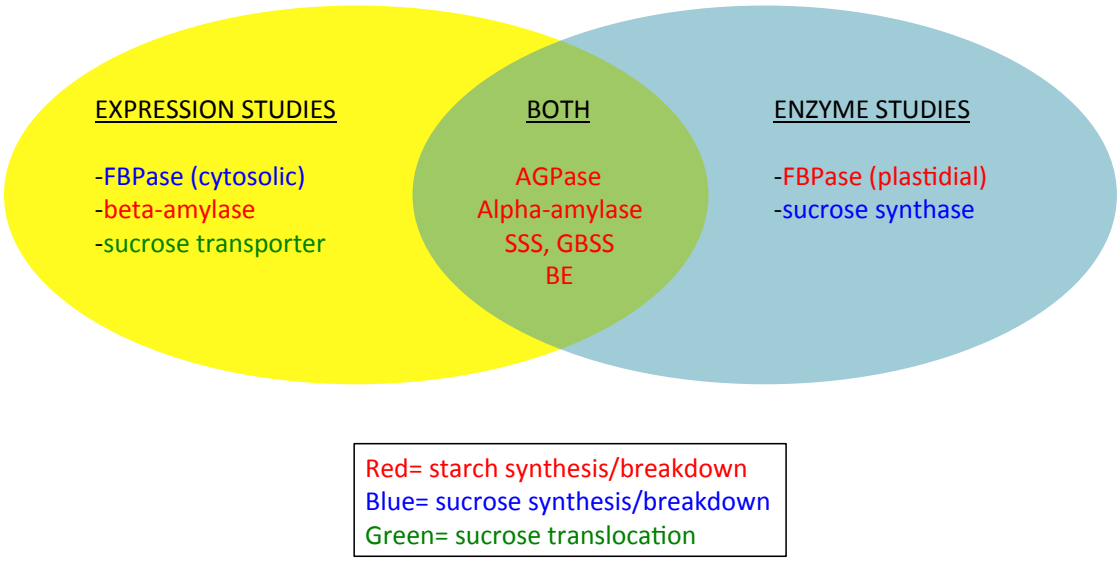


Figure 1.2. Summary of genes and enzymes with documented experimental correlations with rice stem NSC levels.

Stem NSC linkages

Previous work shows that NSC dynamics are well coordinated with other physiological activities (e.g., nitrogen usage, response to water availability and temperature) although underlying pathways are still unclear. These studies are outlined below.

Nitrogen- Heavy nitrogen application suppresses translocation of stem reserves to the grain and decreases its contribution to grain-filling across varieties (Hirano et al., 2005; Fu et al., 2011; Pan et al., 2011). High N levels lead to increased biomass and tiller number and delay senescence, a process associated with carbon remobilization. It has been reported that the decrease in stem reserve translocation due to high nitrogen application occurs along with decreased sink strength (number of endosperm cells), SuSase, AGPase, and weight of inferior spikelets (Fu et al., 2011). Hirano et al. in 2005 found that decreased starch accumulation due to high N was correlated with decreased starch BE activity when documenting differences in NSC dynamics between three cultivars. Negative association found between delayed senescence due to high nitrogen application and stem NSC contribution to grain may have implications for the utility of the 'stay-green' phenotype targeted by irrigated lowland rice breeders.

Water deficit/heat- Opposite to the effect of nitrogen on NSC remobilization, water deficit and high temperatures have been found to increase the contribution of stem reserves to grain filling, coordinated with more rapid grain-filling and leaf senescence

(Yang et al., 2001; Yang et al., 2002; Kim et al., 2011; Morita and Nakano, 2011). Researchers have suggested that stem NSC content at heading contributes to heat tolerance in rice with respect to grain ripening in a comparison of Japanese cultivars with similar heading/maturity dates and sink sizes but different tolerances to high temperature (Morita and Nakano, 2011). The action of ABA and cytokinin may mediate the remobilization of carbon reserves in water stressed rice (Yang et al., 2002); enzymes found to associate with NSC levels under water or heat stress are alpha amylase and sucrose phosphate synthase (Yang et al., 2001; Yang et al., 2002). This positive effect of water deficit on stem carbohydrate remobilization to the grain supports the agronomic practice of withholding water during grain-filling to promote drying down of the rice crop.

When examined together, the effects of nitrogen application and water stress on NSC remobilization are combinatorial. In a 2x2 factorial design experiment (2 watering schemes and 2 nitrogen levels), the water-stressed/normal nitrogen treatment remobilized the most stem reserve, the well-watered/high nitrogen treatment remobilized the least, and the other treatments (water-stressed/high nitrogen and well-watered/normal nitrogen) led to responses in between(Yang et al., 2001).

Climate variability and the potential role for stem NSCs

The relationships between stem NSCs and external variables such as nitrogen, water, and temperature give rise to negative positive linkages that could be potentially exploited in varietal crop improvement in the face of climate variability. However, given the complexities of stem NSC dynamics and how fundamental the processes of

carbohydrate allocation and utilization are, it is more probable that stem NSCs will be revealed as an important functional trait underlying, or even serving as a physiological marker for, adaptive (favorable) variation rather than as a direct target for breeding. For example, it was discovered that submergence tolerance due to the *SUB1A* gene was attributed to optimal regulation of NSC levels during submergence, leaving more shoot NSCs post-submergence thereby leading to faster recovery (Singh et al., 2014). Furthermore, since carbohydrate partitioning is a characteristic of the degree of perenniality in rice (Zhao et al., 2008), stem NSC levels may have implications for some avant garde breeding efforts, such as the perennial grains movement (https://landinstitute.org) or breeding with perennial wild relatives. Leveraging modern genomic tools to understand the basis of phenotypic variation of rice stem NSC builds upon decades of physiological study of this fundamental process and gets us one step closer to practical application.

This dissertation seeks to assess the utility of transient stem non-structural carbohydrates in varietal rice improvement by exploring the genetic basis of its phenotypic variability. Chapter 2 reports the relationship of stem NSCs with yield and agronomic traits under tropical Replicated Yield Trial (RYT) conditions, explores optimal experimental design for evaluating stem NSC genetics under greenhouse conditions, and develops a more rapid means of estimating rice stem NSC levels using near-infrared spectroscopy (NIR) models. Chapter 3 describes the genotyping efforts and population structure analysis of a historical U.S. rice panel that is used for association analysis in Chapter 4. Chapter 4 attempts to clarify the genetic architecture of stem NSC levels using a genome-wide association analysis approach on two related

but distinct germplasm panels (a U.S. rice panel and a *tropical japonica* diversity panel) in addition to a QTL study using an interspecific Chromosome Segment Substitution Line library in a *tropical japonica* genomic background. Chapter 5 explores the effects of CO_2, temperature, and their interaction on yield-related traits on a panel of geographically and genetically diverse rice accessions under growth chamber conditions. Finally, the Appendix describes preliminary work that links Chapter 5 with carbohydrate reserves by exploring the effects of CO_2, temperature and their interaction on leaf and stem NSCs in a 'New Plant Type' (NPT) rice ideotype.

REFERENCES

1. Evans, L. T. & Fischer, R. A. Yield potential: its definition, measurement, and significance. *Crop Sci.* **39,** 1544–1551 (1999).
2. Cock, J. H. & Yoshida, S. Accumulation of 14C-labelled Carbohydrate Before Flowering and its Subsequent Redistribution and Respiration in the Rice Plant. *Proc. Crop Sci. Soc. Japan* **41,** (1972).
3. Slewinski, T. L. Non-structural carbohydrate partitioning in grass stems: a target to increase yield stability, stress tolerance, and biofuel production. *J. Exp. Bot.* **63,** 4647–4670 (2012).
4. Cassman, K. G. *Breaking the yield barrier: Proceedings of a workshop on rice yield potential in favorable environments, IRRI, 29 November-4 December 1993*. (Int. Rice Res. Inst., 1994).
5. Yoshida, S. Physiological analysis of rice yield. *Fundam. rice Crop Sci.* 231–251 (1981).
6. Takahashi, S. *et al.* Microarray Analysis of Sink-Source Transition in Rice Leaf Sheaths. *Breed. Sci.* **55,** 153–162 (2005).
7. Van Dat, T. & Peterson, M. L. Performance of Near-Isogenic Genotypes of Rice Differing in Growth Duration. II. Carbohydrate Partitioning During Grain Filling. *Crop Sci.* **23,** (1983).
8. Chen, H.-J. & Wang, S.-J. Molecular regulation of sinksource transition in rice leaf sheaths during the heading period. *Acta Physiol. Plant.* **30,** 639–649 (2008).
9. Hirose, T., Endler, A. & Ohsugi, R. Gene Expression of Enzymes for Starch

and Sucrose Metabolism and Transport in Leaf Sheaths of Rice (Oryza sativa L.) during the Heading Period in Relation to the Sink to Source Transition. *Plant Prod. Sci.* **2,** (1999).

10. Hirose, T., Ohdan, T., Nakamura, Y. & Terao, T. Expression profiling of genes related to starch synthesis in rice leaf sheaths during the heading period. *Physiol. Plant.* **128,** 425–435 (2006).
11. He, H. Y., Koike, M., Ishimaru, T., Ohsugi, R. & Yamagishi, T. *Temporal and spatial variations of carbohydrate content in rice leaf sheath and their varietal differences*. **8,** (Crop Science Society of Japan, 2005).
12. Ishimaru, K., Kosone, M., Sasaki, H. & Kashiwagi, T. Leaf contents differ depending on the position in a rice leaf sheath during sink-source transition. *Plant Physiol. Biochem.* **42,** 855–860 (2004).
13. Perez, C. M., Palmiano, E. P., Baun, L. C. & Juliano, B. O. Starch Metabolism in the Leaf Sheaths and Culm of Rice. *Plant Physiol.* **47,** (1971).
14. Watanabe, Y., Nakamura, Y. & Ishii, R. Relationship between Starch Accumulation and Activities of the Related Enzymes in the Leaf Sheath as a Temporary Sink Organ in Rice Oryza sativa. *Funct. Plant Biol.* **24,** 563–569 (1997).
15. Fu, J., Huang, Z., Wang, Z., Yang, J. & Zhang, J. Pre-anthesis non-structural carbohydrate reserve in the stem enhances the sink strength of inferior spikelets during grain filling of rice. *F. Crop. Res.* **123,** 170–182 (2011).
16. Hirano, T., Saito, Y., Ushimaru, H. & Michiyama, H. The Effect of the Amount of Nitrogen Fertilizer on Starch Metabolism in Leaf Sheath of Japonica and

Indica Rice Varieties during the Heading Period. *Plant Prod. Sci.* **8,** 122–130 (2005).

17. Cook, F. R., Fahy, B. & Trafford, K. A rice mutant lacking a large subunit of ADP-glucose pyrophosphorylase has drastically reduced starch content in the culm but normal plant morphology and yield. *Funct. Plant Biol.* **39,** 1068–1078 (2012).

18. Pan, J. *et al.* Relationships of non-structural carbohydrates accumulation and translocation with yield formation in rice recombinant inbred lines under two nitrogen levels. *Physiol. Plant.* **141,** 321–331 (2011).

19. Kim, J. *et al.* Relationship between grain filling duration and leaf senescence of temperate rice under high temperature. *F. Crop. Res.* **122,** 207–213 (2011).

20. Yang, J., Zhang, J., Wang, Z., Zhu, Q. & Wang, W. Remobilization of carbon reserves in response to water deficit during grain filling of rice. *F. Crop. Res.* **71,** 47–55 (2001).

21. Yang, J., Zhang, J., Wang, Z., Zhu, Q. & Liu, L. Abscisic acid and cytokinins in the root exudates and leaves and their relationship to senescence and remobilization of carbon reserves in rice subjected to water stress during grain filling. *Planta* **215,** 645–652 (2002).

22. Morita, S. & Nakano, H. Nonstructural Carbohydrate Content in the Stem at Full Heading Contributes to High Performance of Ripening in Heat-Tolerant Rice Cultivar Nikomaru. *Crop Sci.* **51,** 818–828 (2011).

23. Singh, S., Mackill, D. J. & Ismail, A. M. Physiological basis of tolerance to complete submergence in rice involves genetic factors in addition to the SUB1

gene. *AoB Plants* **6,** plu060 (2014).

24. Zhao, M., Acuna, T. L. B., Lafitte, H. R., Dimayuga, G. & Sacks, E. Perennial hybrids of Oryza sativa/Oryza rufipogon: Part II. Carbon exchange and assimilate partitioning. *F. Crop. Res.* **106,** 214–223 (2008).

CHAPTER 2:

ROBUST PHENOTYPING STRATEGIES FOR EVALUATION OF STEM NON-STRUCTURAL CARBOHYDRATES (NSC) IN RICE

HIGHLIGHT

Experimental design and phenotyping methods were explored to enable efficient evaluation of rice stem non-structural carbohydrates for genetic studies. Evidence was found for phenology-dependent yield and stem NSC relationships.

ABSTRACT

Rice plants (*Oryza sativa*) accumulate excess photo-assimilates in the form of non-structural carbohydrates (NSCs) in their stems prior to heading that can later be mobilized to supplement photosynthate production during grain-filling. Despite longstanding interest in stem NSC for rice improvement, the dynamics of NSC accumulation, remobilization, and re-accumulation that have genetic potential for optimization have not been systematically investigated. Here we conducted 3 pilot experiments to lay the groundwork for large-scale diversity studies on rice stem NSC. We assessed the relationship of stem NSC components with 21 agronomic traits in large-scale, tropical yield trials using 33 breeder-nominated lines, established an appropriate experimental design for future genetic studies using a Bayesian framework to sample sub-datasets from highly-replicated greenhouse data using 36 genetically diverse genotypes, and used 434 phenotypically divergent rice stem samples to

develop two partial least squares (PLS) models using near infrared (NIR) spectra for accurate, rapid prediction of rice stem starch, sucrose, and total non-structural carbohydrates. We find evidence that stem reserves are most critical for short-duration varieties and suggest that pre-heading stem NSC is worthy of further experimentation for breeding early maturing rice.

INTRODUCTION

Stem non-structural carbohydrates (NSCs) have long elicited interest from physiologists and breeders across a diversity of economically important grass species. By serving as a temporary sink to store excess photo-assimilates during vegetative growth, the grass stem can also fulfill the role of a source organ during grain filling and maturation, providing carbohydrates for these *in vivo* processes and/or in the form of harvestable end product (Slewinski, 2012). In sucrose-accumulating sugarcane, stems are collected principally for their soluble sugars to support bioenergy production and human consumption (Moore, 1995; Waclawovsky et al., 2010). For perennial forage species such as ryegrass, post-grazing stem NSC reserves enable critical biomass regeneration for the next grazing period (Fulkerson and Donaghy, 2001). NSC reserves stored in the vegetative parts of perennial grasses during the fall are also important to winter survival of temperate species (Slewinski, 2012). For cereals cultivated for grain production, pre-anthesis stem carbohydrate stores can be remobilized to the grain to buffer yields against suboptimum environmental conditions that limit leaf photosynthesis. In subtropical climates where rice is grown as a ratoon crop, NSC reserves that remain in the stem after the first harvest determine the speed

of subsequent vegetative re-growth, flowering, and grain filling of the second crop (Slewinski, 2012) The evident diversity of roles that grass stem NSCs play across species suggests the potential to also optimize their capacities within species through breeding.

For *Oryza sativa* (cultivated Asian rice), knowledge about stem NSC's direct link to yield performance and the genetic controls that underlie its dynamics are still superficial. Rice stems preferentially store starch and sucrose prior to heading, but lose these reserves rapidly following panicle exertion when grain-filling is prioritized energetically; these data support the idea that the stem experiences a sink to source transition at heading (Cock and Yoshida, 1972; Chen and Wang, 2008). There is also evidence of re-accumulation of stem NSCs as the panicle loses sink strength nearing grain-filling completion (Van Dat and Peterson, 1983; Kashiwagi et al., 2006). From their evaluation of two near-isogenic genotypes (*cv.* Calrose 76 and its spontaneous mutant, ED7) to determine yield and carbohydrate-partitioning dynamics throughout grain-filling, Van Dat and Peterson postulated that pre-anthesis stem reserves may have been critical for yield realization in the short-duration genotype, ED7, but not useful for Calrose 76. Despite general acceptance of the temporal and spatial patterns of rice stem NSC, estimates for the overall contribution of these stored carbohydrates to final grain yield vary widely across studies. While some document significant contribution of stem reserves to final yield (up to 40%) (Van Dat and Peterson, 1983; Samonte et al., 2001; Kanbe et al., 2009), others report no such association, pointing to sink limitation (unpublished data cited in Setter et al., 1994). Several studies show evidence of enhanced contribution to grain-filling during suboptimum conditions (e.g.

water deficit or heat stress) coordinated with more rapid grain-filling and leaf senescence (Yang et al., 2000; Yang et al., 2002; Kim et al., 2011; Morita and Nakano, 2011). Management conditions that delay senescence (e.g. heavy nitrogen application) seem to have the opposite effect and suppress translocation of stem reserves, in effect decreasing their grain-filling contribution (Hirano et al., 2005; Fu et al., 2011; Pan et al., 2011). Contrasting with wheat research in which the agronomic role of stem fructans was established early and now informs wheat physiological breeding (Bidinger et al., 1977; Pask et al., 2012; Reynolds et al., 2012), studies on rice have not been able to clearly define aspects of stem reserves that are genetically tractable or capable of optimization and thereby valuable for varietal improvement.

The burgeoning availability of open-access genetic resources for rice (IRGSP, 2005; Jacquemin et al., 2013; Li et al., 2014; Schatz et al., 2014; Duitama et al., 2015; McCouch et al., 2016) supports systematic large-scale studies to dissect the genetic architecture underlying stem NSC dynamics . The MSUv7 genome assembly (http://rice.plantbiology.msu.edu/) has been annotated with gene models across the 12 rice chromosomes, and over 100 predicted enzyme-coding genes have putative catalytic involvement in starch and sucrose biosynthesis, degradation, and transport (Dharmawardhana et al., 2013). Despite these candidates and other genes involved in related developmental pathways (e.g. vascular bundle formation) that likely confer nuances to stem NSC, only very large quantitative trait loci (QTL) that span many megabases of the rice genome have been discovered so far using bi-parental mapping populations.(Nagata et al., 2002; Kashiwagi and Ishimaru, 2004; Kashiwagi et al., 2006; Kashiwagi et al., 2008; Kanbe et al., 2009). No QTLs have been reported in the

literature using genome-wide association studies (GWAS) for rice stem NSC.

We consider two practical constraints to successful QTL identification for the suite of transient, dynamic traits associated with stem NSC. The first issue is that the experimental design parameters needed to effectively evaluate NSC for genetic studies have not been systematically explored. Design features such as the number of replicates required to adequately estimate the mean or dispersion have not been studied for rice stem NSC; the solution likely differs from highly heritable traits such as plant height or heading date but this requires investigation. The second challenge is that the effort involved in traditional analytical chemistry techniques may be prohibitively laborious for larger-scale studies. Carbohydrates are typically assayed using a series of enzyme-mediated reactions to break down polymers into monomers that can then be measured using a UV-VIS spectrophotometer (Trinder, 1969). A previous attempt was made to apply near-infrared (NIR) spectra models to predict stem NSC in rice (Batten et al., 1993), but most advances in the application of indirect methods for NSC determination have taken place in wheat (Ruuska et al., 2006; Wang et al., 2011; Dreccer et al., 2014).

For this study, we began by examining the role of rice stem NSC under optimal growing conditions in irrigated lowland yield trials using a breeder-nominated panel of elite germplasm. We next phenotyped a set of diverse germplasm under greenhouse conditions to generate a highly replicated dataset from which to extract sub-datasets to establish optimal experimental design. Finally, we developed two partial least squares (PLS) models to accurately predict rice stem constituents using near-infrared (NIR) spectral data generated from a semi-automated Thermo-Antaris

Autosampler as a high-throughput alternative to wet chemistry analysis. Our overall goal was to streamline a methodology to support evaluation of rice stem NSC for future genetic studies under controlled (greenhouse) conditions.

MATERIALS AND METHODS

Plant materials: Two sets of rice germplasm were evaluated in this study. The first was a breeder's panel (BP), consisting of thirty-three elite accessions nominated by the International Rice Research Institute (IRRI) and PhilRice, evaluated under field conditions in Los Baños, Philippines. The panel consisted primarily of *indica* genotypes (n=25) along with two *tropical japonica* accessions, four *indica/tropical japonica* derivatives, and two entries that harbored introgressions from wild relatives. Entries were grouped into maturity groups for analyses: Early (<120 days, 10 entries), Medium (120-129 days, 10 entries), and Late (≥130 days, 13 entries).

The second set was a small diversity panel evaluated under greenhouse conditions in Ithaca, NY. This diversity set was comprised of 30 diverse *O. sativa* accessions (20 *indica* and 10 *tropical japonica*) along with 6 contrasting lines (described below) from the breeder's panel. Germplasm information is summarized in **Supp. Table 2.1** and **Supp. Table 2.2**.

Selection of contrasting breeding lines for diversity study: To select multivariate contrasting genotypes out of the 33 genotypes evaluated in the field, we used a Principal Components Analysis based on the line means for four NSC traits (starch levels at flowering and maturity, and sucrose levels at flowering and maturity) and

hierarchical clustering with two linkage methods (complete and average). We selected six genotypes based on membership across clusters in both the complete linkage and average linkage hierarchical clustering trees and wide distribution across the PC1-PC2 plane resulting from the PCA (Fig. 3). The chosen subset was: PS5 (HHZ12-DT10-SAL1-DT1; Green Super Rice), PS6 (IR60; *indica* variety with low chalkiness), PS13 (IR78049-25-2-2-2; *tropical japonica* New Plant Type), PS14 (IR78222-20-7-148-2-B; pyramided QTL for blast-resistance), PS22 (IRRI127; *indica* with tungro resistance), and PS37 (Teqing 1; *indica* from China).

Genetic information: The six BP lines included in the diversity evaluation were genotyped along with 54 control varieties with known subpopulation identities using the Genotyping-By-Sequencing (GBS) platform at 96-plex with *ApeK1* enzyme digestion as the basis for assigning subpopulation designations. Using PCA, the 6 BP genotypes were analyzed along with 54 control varieties using 6,125 genome-wide SNP markers that had 100% call rate on the sixty total accessions (**Supp. Fig. 2.5**). For studying the effect of experimental replication on heritability estimates, we used publicly available genotype information from 700,000 SNPs generated using the High-Density Rice Array (McCouch et al., 2016). Publicly available genotypes were available for 31 of the 36 individuals in the Diversity Panel.

Evaluation and sampling under field conditions: The breeder's panel was planted in Replicated Yield Trials using three replications during the 2012 Dry Season (2012DS) at the International Rice Research Institute in Los Baños, Philippines. The

beginning of the growing season (January-March) displayed consistent precipitation and temperatures, while April at the IRRI Farm experienced significantly higher solar radiation and warmer maximum temperatures (IRRI, 2012). Each replication plot followed Replicated Yield Trial standards 6m x 2m (10 rows x 30 plants). Seeds were sown on 21 December 2011 and seedlings were transplanted on 10 January 2012. Irrigated lowland growing conditions were maintained and fertilizer was applied on the following dates: 6 January (basal application, 30-30-30 NPK), 31 January (topdress, Urea 45kg H/ha), 16 February (topdress), and 29 February (topdress). Days to heading (DTH) was scored on the day that 50% of the individuals in a replication plot exerted panicles and occurred from 14 February to 28 March. Seeds were harvested as plots matured, from 3 April to 30 April.

Four categories of traits were evaluated: phenological traits (days to heading, growth duration, grain-filling duration), resource allocation traits (biomass component dry weights, plant height, tiller and panicle number), physiological traits (flag leaf chlorophyll content at maturity, stem carbohydrates at heading and maturity, and senescence score at maturity) and yield components (yield, harvest index, percent grain-filling, 1000 grain weight, and harvest index). See **Supp. Table 2.1** for complete list and details on trait measurement methodologies. To sample plants for trait measurements at heading and maturity stages, five hills were chosen randomly from each plot per sampling point and pulled entirely out of the ground, while yield was measured on $4m^2$ harvested area per plot (equivalent to 25 hills) and harvest index measured on $1m^2$ harvested area per plot (equivalent to 25 hills). All grain weight values are calibrated to 14% moisture content. For chlorophyll determination, five flag

leaves were sampled per rep at maturity. Leaves were freeze-dried and chlorophyll extracted overnight with cold 80% acetone aqueous solution. Absorbance was read at 663, 652, and 645 nm using a spectrophotometer and chlorophyll concentration in ppm (mg/L) was calculated (per Bruuinsma, 1963).

For carbohydrate sampling, all five randomly selected hills per plot sampled for trait evaluation on the day of flowering were combined. Panicles and leaf blades were immediately removed; stems were chopped manually using scissors into ~1cm long pieces and subsequently subsampled for 20% of the original mass and flash frozen in liquid nitrogen. These samples were then freeze-dried and finely ground using a sample mill. Subsamples of 200mg of the freeze-dried, ground tissue were taken and subjected to carbohydrate analysis per the Yoshida method (Cock and Yoshida, 1972). The proportion of TNC remobilized from heading to maturity (TNC_{RMB}) relative to the amount accumulated at heading was calculated following

$$TNC_{RMB} = 100\% * \frac{(TNC_{hd} \times STM_{hd} - TNC_{mt} \times STM_{hd})}{(TNC_{hd} \times STM_{hd})}$$

where *TNC* is the percentage of total non-structural carbohydrates by dry weight at heading (*hd*) or maturity (*mt*), *STM* is the dry weight of stems from five hills collected at heading (*hd*) or maturity (*mt*). This is an indirect estimate of the TNC remobilized, as it reflects the net difference between heading and maturity. It therefore does not account for any changes that might occur between the two sampling points (e.g. re-accumulation of TNC at the end of grain-filling).

Evaluation and sampling under greenhouse conditions: The small diversity panel

was grown in Guterman greenhouse in Ithaca, NY during spring 2013 in a RCBD with 20 replicates per accession (36 accessions total, see Plant Materials). Seeds were sown on March 5 and seedlings transplanted into eight-inch pots set into flooded tanks two weeks after. Greenhouse conditions were managed using 11 hours light (29^{o}C)/13 hours dark (24^{o}C) at 55% constant humidity.

Days to heading was scored when the first panicle per plant emerged at least 50% out of its sheath and the following additional traits were evaluated: tiller number (counted at maturity), senescence score (according to SES, IRRI), stem dry weight measured on a tiller basis after removal of the panicle and leaf blades. Sampling for stem carbohydrates was done for a single sampled tiller per replicate per accession at two time points (heading and maturity), resulting in 20 biological replicates per accession.

For carbohydrate analysis, at each stem sampling point, tillers were excised from individual plants, and the leaf blades and panicle were removed. The remaining stems (culm and leaf sheath collectively) were microwaved immediately post-sampling to destroy respiration enzymes and then dried slowly in a 65^{o}C oven over the course of several days until constant weight. Stem samples were initially coarse-ground using a conventional coffee grinder and then milled finely to a 0.2mesh with an Udy Cyclone Sample Mill (Udy Corporation, Fort Collins, CO, USA). ~25mg subsamples (exact weights recorded) were measured into 1.2mL tubes in a 96-tube format assayed for carbohydrates: sugars were extracted using 80% EtOH (and sucroses broken down via invertase) and the remaining starches gelatinized and digested into glucose using amyloglucosidase/α-amylase. All carbohydrates were

assayed utilizing peroxidase and glucose oxidase via the Trinder reaction and compared against known standards of glucose and sucrose at 490nm using a 96-well plate-reading UV-VIS spectrophotometer (Trinder, 1969).

Near infrared spectral measurements: A total of 434 samples were used in the NIRS calibration study. These samples were selected in two batches: the first (n=277) were chosen to span the range of measured wet chemistry values in the diversity panel experiment with representation across the 36 different genotypes and two sampling points, and the second batch (n=157, also with wet chemistry data) were selected randomly across genotypes and sampling points in two experiments: 357 samples from the current diversity study (germplasm described above) and 77 samples from a separate unpublished evaluation on a set of Chromosome Segment Substitution Lines (germplasm described in Arbelaez et al., 2015)..Choosing from the diversity study allowed us to sample across genetically divergent individuals that may display different background spectral signatures, while selecting across experiments conducted in two years allowed us to sample across environmental differences that may impact background spectral variation. Samples were packed in 2-dram borosilicate scintillation vials and scanned using a Thermo Antaris II FT-NIR Spectrometer with a 40-position autosampler carousel (Thermo Fisher Scientific, Waltham, MA). Each sample was scanned 128 times (wavenumber range: 3,300-12,000 cm^{-1}).

Statistical analysis: All statistical analyses were coded and carried out in R

(https://www.R-project.org/). Bayesian approaches were used for experimental design optimization and these methods are fully described in **Supp. Fig. 2.6-2.7**. These analyses involved nine traits (DTH- days to heading, STCH_HD- starch, heading, SUC_HD- sucrose, heading, SEN- senescence score, TIL- tiller number, WT- stem weight, GLC_HV- glucose, maturity, SUC_HV- sucrose, maturity, STCH_HV- starch, maturity). Model fitting for NIR data were undertaken using functions from the following R packages: signal (http://r-forge.r-project.org/projects/signal/), prospectr (Stevens and Lopez, 2013), and pls (Mevik et al, 2015). For calibration, resultant spectral data were first limited to the 4000-9000 wave number range and subjected to mathematical pre-treatment using the standard-normal-variate (SNV) scatter correction and a first-derivative Savitzky-Golay smoothing (n=25 points). Using these pre-treated spectra, the 434 samples were divided into a calibration set (n=300) and an external validation set (n=134) using the Kennard-Stone algorithm. Two partial least square (PLS) models were developed: a PLS1 model and a PLS2 model. The PLS1 model predicted a single response variable, total non-structural carbohydrate (TNC = starch + sucrose). The more generalized PLS2 model predicted two variables, starch and sucrose, as a single multivariate model. Small subsets of the 300 calibration samples were excluded as outliers (18 for the PLS1 model and 17 for the PLS2 model) in the final models. These outliers were identified based on the difference between initial prediction of the fully cross-validated model and actual wet chemistry and excluded if they were more than two times the RMSECV value for the initial model. Root mean squared errors for calibration (RMSEC), full cross-validation (RMSECV), and prediction of an independent test set (RMSEP) were calculated

$$RMSE = \sqrt{\frac{1}{N}\Sigma(\widehat{Y}_i - Y_{i,ref})^2}$$

where N is the number of samples in each population, $\widehat{Y}_i$ the predicted value and $Y_{i,ref}$ the measured value. Reference method (i.e. wet chemistry) uncertainties were estimated as twice the average standard deviation (per Wolfrum et al., 2013) on an external set of material (n=150) using triplicated technical replicates. **Table 2.3** summarizes the results of the modeling effort.

RESULTS

Correlation blocks of agronomic characteristics in breeder-nominated lines

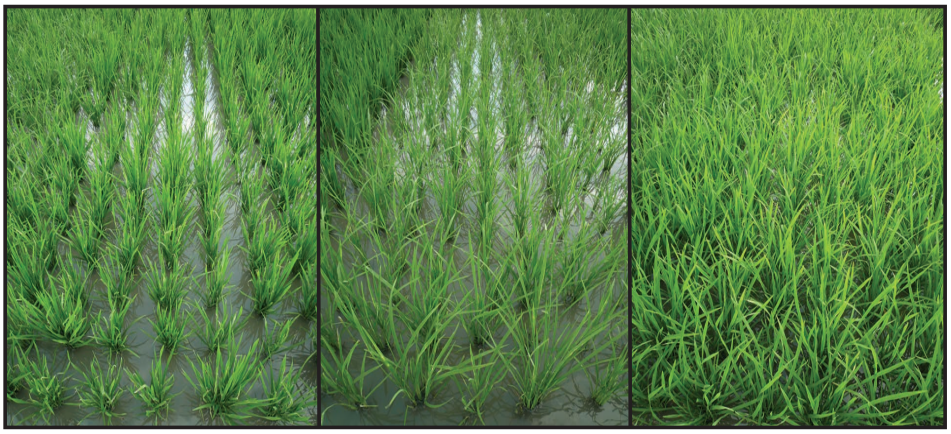
During 2012 dry season the breeder's panel exhibited wide variation in growth duration, which spanned from 104 to 141 days after sowing, and other growth-related phenotypes (e.g. canopy growth rate **Fig. 2.1A**). Evidence of strong phenological linkage was observed for biomass traits. Days to heading, growth duration, and grain-filling duration were positively correlated with dry weight measurements (panicle weight, stem weight, leaf weight at heading, stem weight at maturity and 25 hill straw weight at maturity), while these same phenological measures were negatively correlated with "count" data (panicle number, tiller number at heading and maturity) (see **Fig. 2.1B** for correlation blocks). Flag leaf area, plant height, and straw dry weight at maturity were also negatively associated with these count traits. There was no relationship between phenology and yield or yield components except for a weak negative correlation of days to heading and harvest index, an indirect outcome of the positive relationship between phenology and straw biomass, which contributes to the

denominator of harvest index. Of the stem NSC traits, we observed a highly significant association between phenological measurements and stem sucrose at maturity, indicating that in this panel, observations with a longer growing season either retained or re-accumulated more stem sucrose.

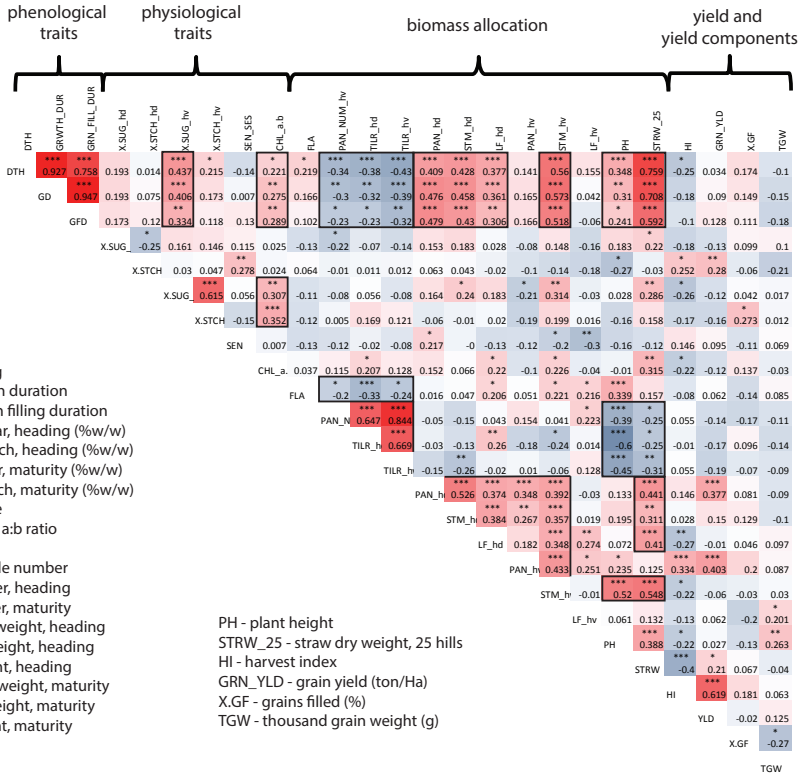
Out of the panel, BP30 (*cv* IR60) had the lowest yield with an average of 3.72 ton/Ha while BP 26 (*cv* Teqing) had the highest yield with 6.9 ton/Ha average. Not including seed weight-related traits that are simply alternative expressions of yield, e.g. harvest index and panicle weight at harvest, we found two traits significantly correlated with grain yield: stem starch at heading (potential source for grain-filling) and panicle weight at heading (potential sink size for grain-filling). Interestingly, we also observed a significant negative relationship between thousand grain weight and percentage of grains filled, meaning entries with heavier, completely filled grains tended to have a smaller proportion of filled grains overall.

Figure 2.1. Phenotyping of the breeder's panel in a replicated yield trial. (A) BP entries exhibited a wide range of phenotypic variation for many traits including developmental rate as depicted by these photos of the field trial taken at the same time point in the same replicate field for three entries. From left to right: BP 30, BP 6, and BP 31. (B) Correlation matrix of all directly measured traits ($p<0.05$ *; <0.01 **, <0.001 ***). Red and blue squares indicate positive and negative correlations, respectively.

A



B



DTH - days to heading
GRWTH_DUR - growth duration
GRN_FILL_DUR - grain filling duration
X.SUG_hd - stem sugar, heading (%w/w)
X.STCH_hd- stem starch, heading (%w/w)
X.SUG_hv- stem sugar, maturity (%w/w)
X.STCH_hv- stem starch, maturity (%w/w)
SEN_SES - senescence
CHL_a.b - chlorophyll a:b ratio
FLA - flag leaf area
PAN_NUM_hv - panicle number
TILR_hd - tiller number, heading
TILR_hv - tiller number, maturity
PAN_hd- panicle dry weight, heading
STM_hd - stem dry weight, heading
LF_hd - leaf dry weight, heading
PAN_hv - panicle dry weight, maturity
STM_hv - stem dry weight, maturity
LF_hv - leaf dry weight, maturity

PH - plant height
STRW_25 - straw dry weight, 25 hills
HI - harvest index
GRN_YLD - grain yield (ton/Ha)
X.GF - grains filled (%)
TGW - thousand grain weight (g)

Linkage between stem NSC and yield performance is strongest for short duration rice

At heading, the BP accumulated an average of 16.7% total non-structural carbohydrates (TNC) and retained 7.7% at physiological maturity (**Supp. Table 2.1**). Observed differences in stem NSC at heading and at maturity provided evidence that patterns of starch accumulation and remobilization were synchronized with development and differed significantly across the two sampling points (paired t-test, $p<2.2e-16$). Stem soluble sugar levels, on the other hand, were not different between heading and maturity.

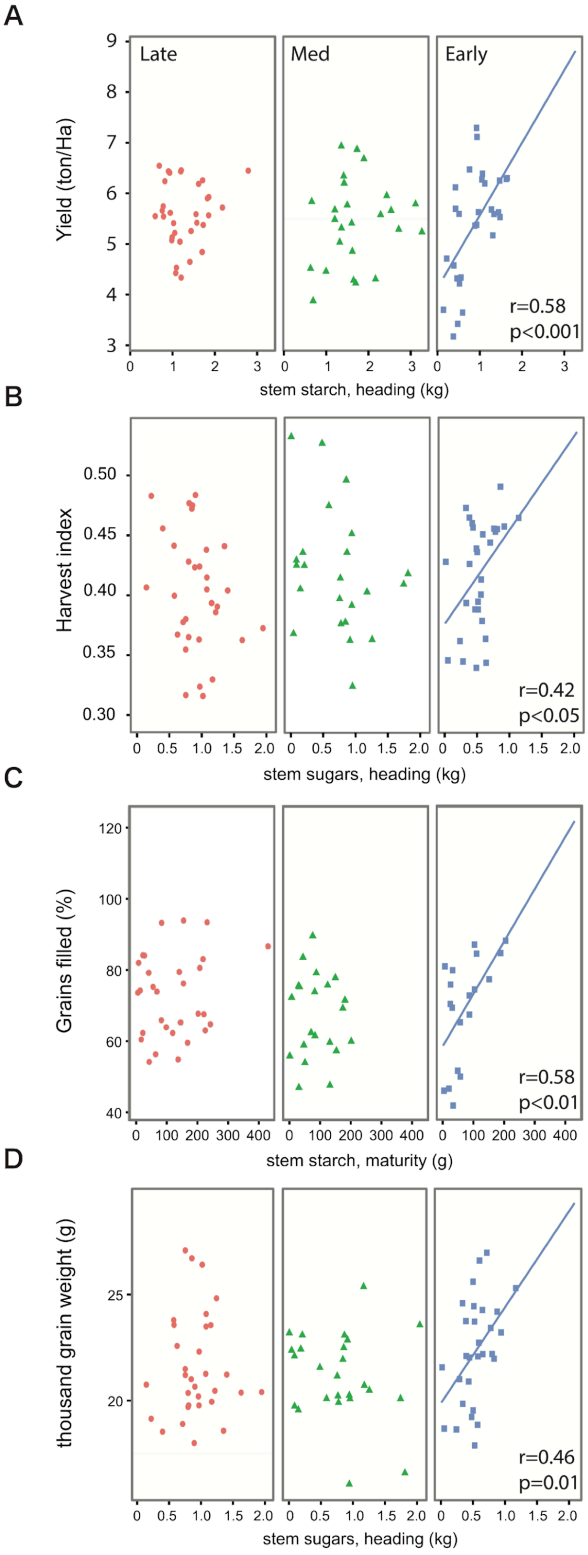
Despite the fact that all entries were adapted to tropical conditions and selected for this study because they play key roles in irrigated lowland rice breeding, there was a surprisingly wide range of phenotypic variation with respect to stem NSC across the BP. At heading, BP 11 (*cv* BR29, a Bangladeshi *indica* mega-variety) accumulated the least amount of total non-structural carbohydrate TNC (8.62%) while BP 25 (*cv* Teqing Acc. IRGC78727, a Chinese *indica*) accumulated the most at 24.62%. While most entries had lost most of their stem reserves by the end of the season, BP 18 (*cv* NSIC212, an *indica* high-yielding variety released in the Philippines) harbored 13.11% TNC at maturity, due to high levels of soluble sugars.

Although per se effects of growth duration or flowering time were not observed for NSC traits or yield traits, we wondered if the relationships between NSC and yield might vary due to phenology, as previously suggested (Van Dat and Peterson, 1983). We addressed this question by grouping entries into three maturity groups, Early, Medium, and Late (see **Methods** for definitions). To further investigate the positive

relationship observed between grain yield and relative starch at heading, we re-expressed relative stem NSC traits (% w/w) in absolute terms (dry weight per 5 hills sampled) as these are more appropriate representations of the total carbon incorporated into non-structural components of biomass. Relationships between NSC and the four yield components studied (yield, harvest index, percent grain filling, and thousand-grain weight) differed significantly across maturity groups. Analyzing Early maturity group on its own revealed significant associations of at least one NSC component for every yield trait, while the isolating the Medium and Late groups uncovered no further relationships between stem NSC and yield components (e.g. **Fig. 2.2**). One surprising outcome, given that yield, harvest index, and thousand grain weight were all associated with NSC traits at heading, was that grain-filling percentage was correlated with NSC at maturity.

To explore the hypothesis that there might be a physiological trade-off between carbohydrate remobilization and the Stay-Green phenotype at maturity, we examined the distribution of starch remobilization across senescence classes (**Supp. Table 2.2** for scoring method). The distribution of TNC remobilized shifted upwards with increasing senescence (**Supp. Fig. 2.2**), and a significant difference in the means of Group 3 (the most "Stay-Green") and Group 9 (the most senescent) was observed (one-tailed Welch's t-test, $p= 0.02$). This result suggests that there is a negative physiological linkage between carbohydrate remobilization and the Stay-Green condition in these lines, a relevant finding given that both traits are targets of interest for IRRI's irrigated lowland breeding program.

Figure 2.2. Phenology-dependent trait relationships. Early (square), Medium (triangle), and Late (circle) entries displayed contrasting relationships of NSC with yield traits for four yield components: A) yield, B) harvest index, C) grain-filling percentage, and D) thousand grain weight.


A
Late
Med
Early
Yield (ton/Ha)
r=0.58
p<0.001
stem starch, heading (kg)
B
Harvest index
r=0.42
p<0.05
stem sugars, heading (kg)
C
Grains filled (%)
r=0.58
p<0.01
stem starch, maturity (g)
D
thousand grain weight (g)
r=0.46
p=0.01
stem sugars, heading (kg)

Multivariate classification of breeder's panel by stem NSC components

Principal Component Analysis (PCA) using stem NSC components as variables revealed that PC 1 and 2 collectively explained ~80% of the total phenotypic variance observed in the thirty-three entries of the Breeder's Panel. Mapping the trait vectors onto the PC1-PC2 plane showed that PC1 was most aligned with stem NSC levels at maturity while PC2 was nearly collinear with the percentage of stem NSC at heading. Most entries clustered together (black individuals, **Fig. 2.3**, but twelve fell beyond the main group in the PC1-PC2 plane. Of these twelve, seven formed a second group (blue individuals, **Fig. 2.3**) that was characterized by low starch levels at heading and low starch and sugar at maturity. At the other end of the PC2-axis were two individuals, BP 26 and BP 29 (yellow individuals, **Fig. 2.3**), which also retained little NSC at maturity but accumulated a large amount of starch at heading. These two entries represent the best "remobilizers" of stem carbohydrates in the panel due to the large net loss of NSC between heading and maturity. In contrast, BP 5 and BP 18 (pink, **Fig. 2.3**) retained nearly the same amount of TNC at maturity as they had accumulated at heading, indicating either lack of remobilization or a re-accumulation at maturity. Interestingly, BP 25 (green, **Fig. 2.3**) and BP 26, different accessions of the same variety Teqing, did not cluster with each other as we had expected; both accumulated the same amount of starch at heading but BP 25 had much more sucrose at heading and maturity. Upon closer inspection we discovered that in our yield trial, BP 25 took an average of 4 days longer to flower than BP 26 but matured in the same time, indicating a faster grain-filling period. It was also less senescent at maturity than BP 26. Hierarchical clustering using the average linkage method supported the PCA

findings overall (**Supp. Fig. 2.1**). There was no significant association between multivariate clustering by stem NSC components and other metadata on the rice accessions, such as subpopulation identity (*indica* vs. *tropical japonica*) or line classification (breeding line vs. released variety).

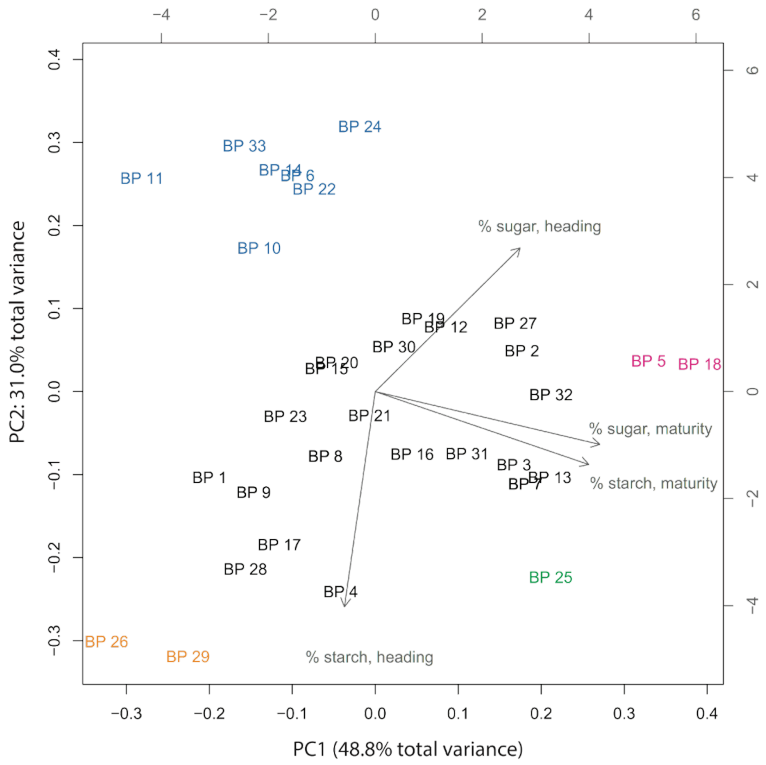


Figure 2.3. Principal Component Analysis (PCA) of Breeder's Panel. PCA was performed using mean NSC trait data and trait loadings are indicated by gray arrows. Individuals (breeding lines) are colored according to groupings, defined by visual inspection of the PCA biplot: black text indicate the main cluster of individuals while blue, yellow, pink, and green text define minor clusters.

NSC variation across diverse rice germplasm under greenhouse conditions

In contrast with the results of the field evaluation of the IRRI BP, no coordinated net starch loss was observed in the greenhouse trial of diverse germplasm (**Supp. Table 2.3** for germplasm details); distributions of starch levels at heading did not differ significantly from overall starch distribution at maturity for any of the six BP lines selected for greenhouse evaluation (see **Methods** for selection process) (**Supp. Fig. 2.3A, D**). This may be due to sustained stem starch deposition throughout grain-filling due to continuous fertilization in our greenhouse management conditions in conjunction with lower competition for radiation associated with spaced planting. Sugar level, however, was significantly lower at maturity than at heading in the greenhouse (**Supp. Fig. 2.3B, E**). Relative to 20 diverse rice accessions that were concurrently evaluated in the greenhouse, the IRRI BP lines accumulated more sucrose at heading but retained more starch at maturity, and the IRRI lines tended to cluster together at in one tail of the phenotypic spectrum (**Supp. Fig. 2.4B-C**).

Because rice has a deeply stratified subpopulation structure, we next wanted to compare trait distributions of the 6 BP entries directly against diverse germplasm that shared their genetic grouping. To determine the subpopulation identity of the 6 BP entries, we genotyped them using Genotyping-By-Sequencing (GBS) and analyzed the accessions along with fifty-four controls with known subpopulation identity (**Supp. Fig. 2.5**). Out of the six BP lines, four (BP 14, BP 25, BP 29, BP 30) belonged to the *indica* subpopulation, reflecting this subpopulation's importance to irrigated lowland rice breeding. Of the remaining two, BP 6 was classified as an admixed *indica* + *tropical japonica* and BP 5 was an admixed *temperate japonica* + *tropical japonica*.

These results are consistent with pedigree information of BP 6 (*indica* x *tropical japonica* derived for blast resistance) and BP 5 (*temperate* x *tropical japonica*-derived New Plant Type). We found a significant difference between the mean performance of BP *indicas* and the more diverse *indicas* for levels of starch at heading, sucrose at heading, and starch at maturity ($p<0.01$ for all, Welch's t-test). A significant difference was also observed in the mean of the BP *tropical japonica* and the more diverse *tropical japonicas* for levels of both starch and sucrose at heading and maturity ($p<0.01$ for all). This supports previous observations that under greenhouse conditions, these breeding lines sample only a portion of the phenotypic variation found in an expanded panel of diverse rice germplasm, despite the wide range of NSC variation observed in the BP.

Effect of experimental replication on genetic parameters of rice stem NSC

We next carried out a greenhouse evaluation of stem NSC in diverse rice accessions using 20 replicates per accession. The highly replicated nature of this study allowed us to explore questions related to experimental design to understand the effect of replicate size (RS) on estimates of distribution and genetic parameters for stem NSC traits. Using these data in conjunction with genetic information on 700,000 genome-wide SNPs from the High-Density Rice Array (McCouch et al., 2016), we took a Bayesian approach to estimate heritability (broad and narrow), genetic and environmental correlations, line mean, genomic estimated breeding value (GEBV), line dispersion parameters, and variance components (additive, non-additive, and error). Narrow sense heritability estimates for starch and sucrose at heading were 0.56

and 0.61, respectively, while at maturity they were slightly lower at 0.49 (starch) and 0.58 (sucrose) (**Table 2.1**). Genetic and environmental correlations of trait pairs are summarized in **Table 2.2**.

Table 2.1. Heritability of rice stem NSC traits. Point estimates for narrow-sense heritability of stem NSC traits are shown here. Values are calculated based on genotypic data on 700K SNPs on 31 *indica* and *tropical japonica* individuals and phenotype data from 20 biological replicates per accession under greenhouse conditions. The labels 'hd' and 'mt' indicate heading and maturity sampling points, respectively.

trait	**h^2**
Starch, hd	0.56
Starch, mt	0.49
Sucrose, hd	0.61
Sucrose, mt	0.58

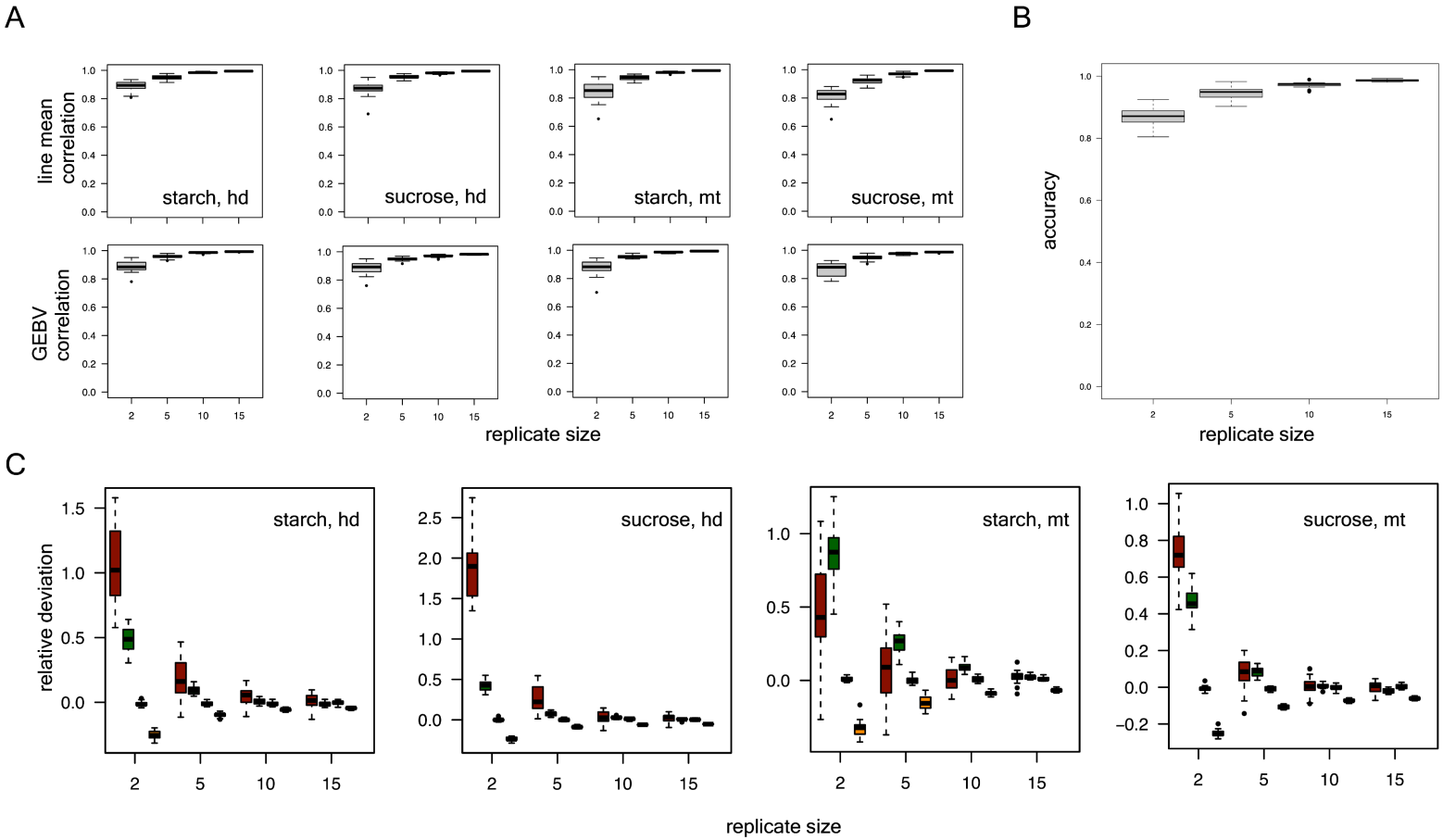
Table 2.2. Genetic and environmental correlations for full replicate dataset. In each trait-trait pairwise sector of the matrix, the middle number indicates the point estimate while the top and bottom values indicate the upper and lower credible interval bounds resulting from Bayesian analyses with Gaussian error distributions. Estimation using Bayesian methods with Student-t error distribution yielded similar results.

	DTH	**Starch, hd**	**Sucrose, hd**	**Tiller no.**	**Stem wt**	**Sucrose, mt**	**Starch, mt**
DTH	1	0.792	0.745	0.706	0.631	0.718	0.849
	1	0.468	0.333	0.415	0.337	0.291	0.617
	1	-0.197	-0.33	-0.0241	-0.0314	-0.381	0.0719
Starch, hd	0.134	1	0.709	0.689	0.799	0.68	0.805
	0.0878	1	0.191	0.222	0.469	0.0777	0.464
	0.0418	1	-0.533	-0.42	-0.196	-0.586	-0.283
Sucrose, hd	0.15	0.234	1	0.752	0.695	0.732	0.684
	0.0604	0.199	1	0.341	0.218	0.247	0.123
	-0.0262	0.168	1	-0.37	-0.442	-0.554	-0.553
Tiller no.	0.141	0.186	0.13	1	0.516	0.665	0.567
	0.0541	0.105	0.0446	1	0.0606	0.179	0.0417
	-0.033	0.0173	-0.0406	1	-0.369	-0.484	-0.498
Stem wt	0.112	0.131	0.192	-0.0569	1	0.741	0.807
	0.0282	0.0466	0.108	-0.111	1	0.326	0.509
	-0.0567	-0.039	0.0232	-0.162	1	-0.36	-0.0605
Sucrose, mt	0.0476	0.072	0.193	0.104	0.239	1	0.667
	-0.0348	-0.0131	0.115	0.0167	0.168	1	0.111
	-0.124	-0.0978	0.0246	-0.0646	0.0969	1	-0.551
Starch, mt	0.203	0.116	0.217	0.00827	0.617	0.31	1
	0.119	0.0317	0.136	-0.0774	0.562	0.278	1
	0.0328	-0.0544	0.0527	-0.161	0.503	0.251	1

To evaluate the effect of experimental replication size on trait parameters, we simulated at four different replicate sizes (RS = 2, 5, 10, 15) by subsampling from the full dataset (see **Methods**). With respect to line mean, improvements in accuracy (correlation between subsampled data and full dataset) were observed between RS of

two and five for starch and sucrose phenotypes, while there was not much gain in accuracy beyond five replicates (**Fig. 2.4A**). A similar trend was observed for GEBV estimates, except for starch at maturity, which benefitted from an RS increase up to 10. In addition to point estimates, we were interested in the effect of RS on uncertainty estimates. We chose coefficient of variation (CV) to describe trait dispersion and found that an RS of at least 10 was necessary to come within 1.5 times the full set CV estimate (**Supp. Fig. 2.6**). For genetic correlation estimates, an increase from two to five replicates improved the median accuracy but an RS of 10 yielded a much tighter distribution of accuracies compared with an RS of five (**Fig. 2.4B**). Finally, we used our highly replicated data to inquire about the effect of replicate size on variance component estimates and narrow sense heritability (**Fig. 2.4C**). Analysis of relative deviation of sub-sampled data versus full data value revealed that error variances were overestimated at small replicate sizes while narrow sense heritabilities were underestimated when compared to corresponding values from the full dataset for starch and sucrose traits.

Figure 2.4. Effect of experimental replication on trait parameters. Effect of replicate size on (A) line mean and GEBV, (B) genetic correlation, and (C) variance component and narrow sense heritability estimates (boxes from left to right per replicate size: error variance, non-additive genetic variance, additive genetic variance, heritability).


A
line mean correlation
GEBV correlation
starch, hd
sucrose, hd
starch, mt
sucrose, mt
replicate size
B
accuracy
replicate size
C
relative deviation
starch, hd
sucrose, hd
starch, mt
sucrose, mt
replicate size

Utility of NIR spectral data to predict NSC in rice stem samples

As NIR models have proven useful for assaying wheat stem fructans (Wang et al., 2011), we assessed the potential of using NIR predictions to determine rice stem NSC (**Fig. 2.5**). From 434 rice stem samples selected to represent across the range of variation detected in primary analytical data, we divided samples into either the calibration set or the external validation set using the Kennard-Stone algorithm. This resulted in 300 and 134 samples in the calibration and validation sets, respectively, which spanned the multi-dimensional spectral space (**Supp. Fig. 2.8**). Two models were fit for three primary NSC variables of interest: a PLS-1 model for TNC, and a PLS-2 model for both starch and sucrose. Overall, the models predicted TNC and starch with high accuracy, resulting in validation R^2 values of 0.92 and 0.96, respectively. The PLS-2 model yielded a validation R^2 of 0.76 for sucrose, with a satisfactory RMSEP of 0.02 that is comparable to the reference method uncertainty of 0.018 (see **Methods** for definitions). Final model results are summarized in **Table 2.3** and **Fig. 2.6**. The average throughput for scanning on the Thermo-Antaris FT-NIR spectrometer using an autosampler carousel was approximately 40 samples/hour, with each sample being scanned 128 times. This estimate includes the time necessary for packing and cleaning vials and loading/unloading the scanning carousel. Our results support the use of NIR prediction models to determine NSC levels in rice stems, conditional on a reliable primary analytical method for calibration.

Table 2.3. Summary statistics for NIR calibration models

	PLS-1 model	PLS-2 model	
	TNC	sucrose	starch
calibration samples	300	300	300
independent validation samples	134	134	134
outliers removed*	18	17	17
Principal components in model	8	10	10
R^2 calibration	0.96	0.86	0.96
R^2 cross validation	0.95	0.84	0.95
R^2 independent validation	0.92	0.74	0.96
RMSEC	0.02	0.015	0.015
RMSECV	0.02	0.016	0.016
RMSEP validation	0.025	0.02	0.013
Reference method uncertainty	0.024	0.018	0.022

*from calibration set

Figure 2.5. NIR protocol for assaying rice stem NSC. Steps taken for using near-infrared spectroscopy for rapid determination of rice stem NSC components.

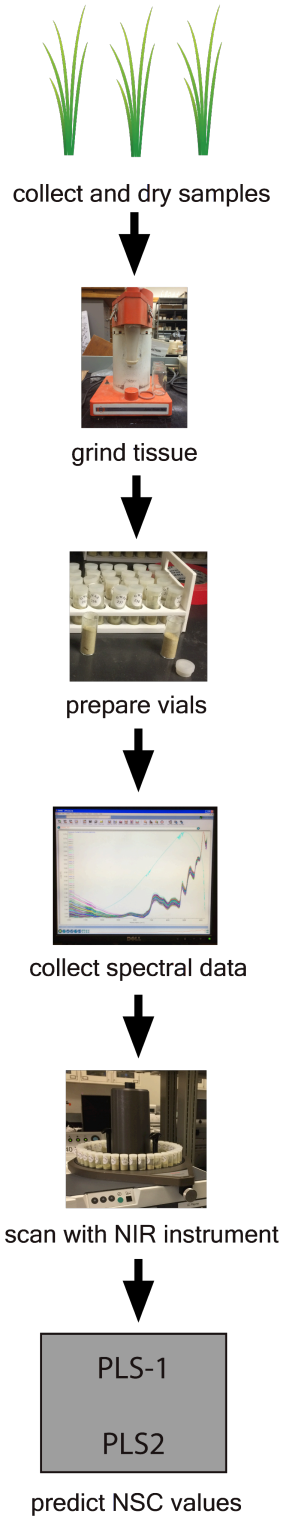
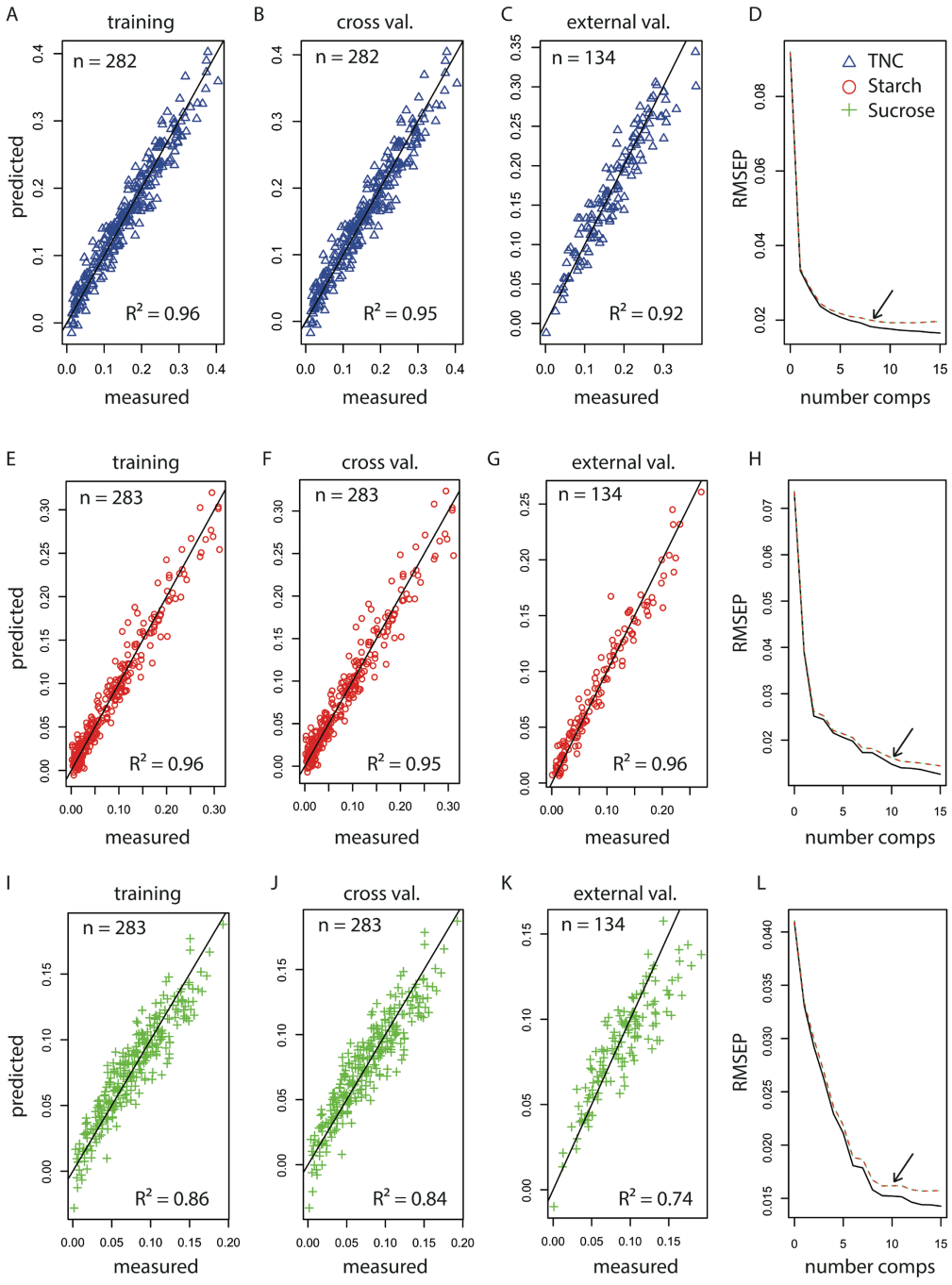


Figure 2.6. Results from NIR PLS calibration models. Predicted vs. measured values for training, cross validation, and external validation sets of the PLS1 TNC model (A-C), PLS2 model for starch (E-G) and PLS2 model for sucrose (I-K). RMSEP values as a function of model component number for PLS1 TNC (d), PLS2 starch (H) and PLS2 sucrose (L). Black arrows in RMSEP panels indicate the number of components included in final calibration models.


A
training
n = 282
R² = 0.96
predicted
measured
B
cross val.
n = 282
R² = 0.95
measured
C
external val.
n = 134
R² = 0.92
measured
D
TNC
Starch
Sucrose
RMSEP
number comps
E
training
n = 283
R² = 0.96
predicted
measured
F
cross val.
n = 283
R² = 0.95
measured
G
external val.
n = 134
R² = 0.96
measured
H
RMSEP
number comps
I
training
n = 283
R² = 0.86
predicted
measured
J
cross val.
n = 283
R² = 0.84
measured
K
external val.
n = 134
R² = 0.74
measured
L
RMSEP
number comps

DISCUSSION

Heading stem NSC as a potential breeding target for rice

Results of our field evaluation on the breeder's panel suggest that the amount of stem carbohydrates accumulated by heading would be the most appropriate target for rice varietal improvement with respect to NSC traits. Post-heading stem reserve mobilization appears to be primarily a function of sink strength and environment/management (Yang et al., 2000; Chen and Wang, 2008; Kim et al., 2011; Morita and Nakano, 2011), while levels measured at maturity can be confounded by additional photo-assimilate re-accumulation at the end of grain-filling. The extents to which rice accumulates, remobilizes, and re-accumulates stem NSC are each affected by genotype, environment, and genotype x environment interaction, however, earlier traits (e.g. accumulation measured at heading stage) are probably the most genetically tractable, evidenced here by higher narrow-sense heritabilities with heading NSC traits.

In our study, grain yield was significantly associated with relative starch at heading (%w/w) across the panel. When evaluating Early, Medium, and Late maturity groups separately, we find that the linkages between stem NSC and yield traits (yield, harvest index, grain-filling percentage, and thousand-grain weight) were strongest for the Early maturity group while generally insignificant for the Medium and Late groups. These associations all involved NSC heading traits, except for grain-filling percentage that was positively correlated with stem starch at maturity. Since the level of stem NSC at maturity is a combination of NSC retained from heading and any end-of-season re-accumulation, this last outcome may due to one or both of the following:

1) an indirect consequence of sink limitation (e.g. low number of total spikelets) that prematurely resulted in a high percentage of grains filled and a decrease in panicle sink strength that led to re-accumulation of stem NSCs, and 2) adequate photoassimilate production from flag leaves during grain-filling that reduced the need for pre-anthesis stem NSC. In this evaluation, there was no overall association of grain-filling percentage with grain yield, a possible demonstration of the diverse strategies (and therefore constraints) for yield formation taken by this collection of elite germplasm; some entries may be sink limited in the panicle, while others may have experienced flag leaf source limitations.

In addition to finding maturity group-level differences, we documented a curious contrast in NSC characteristics between BP 25 and BP 26, two accessions of *cv.* Teqing entered into IRRI gene bank. Documentation suggests these accessions were contributed into the gene bank 17 years apart (http://www.iris.irri.org/) and represent similar but not genetically identical versions of the named variety, *cv* Teqing, from China. In our yield trial BP 25 and BP 26 displayed similar yield (6.3 and 6.29 ton/Ha, respectively), harvest index (0.46 for both), and growth duration (130 days for both) but differed in phenology (flowering time and grain-filling duration) and stem NSC patterns. These two nearly isogenic accessions are a valuable resource to help disentangle the physiological and genetic linkages between NSC and phenology.

Observed dependency of NSC-yield relationships on phenology in our work here on tropical rice is consistent with prior work on temperate-adapted near-isogenic lines and the hypothesis that carbohydrates accumulated prior to heading may be more critical for yield formation in plants with a shorter lifespan (Van Dat and

Peterson,1983). We rationalize that short-duration varieties do not have a long vegetative growth period to produce lavish amounts of leaf biomass that may later serve as source organs to generate photo-assimilates concurrently with grain-filling. This may underlie the tight relationships between pre-heading stem reserves and yield attributes in BP's Early maturity group and raises the possibility of tailoring phenologically-dependent source-sink relationships in rice breeding.

Previous studies have demonstrated physiological linkages within the triad of non-structural carbohydrate remobilization, monocarpic senescence, and the Stay-Green condition, with the former two acting oppositely from the latter. Environmental factors that hasten senescence and enhance NSC remobilization (e.g. water limitation and temperature increase) are also common climate change variables of concern, which highlights future potential of breeding for high accumulation of NSC at heading to buffer against increased within-season environmental variability. However, long-term expectations should be tempered until effects of other climatic parameters (e.g. rising atmospheric CO_2) on pre-heading stem NSC are better understood, as they may have unexpected interaction effects (Moya et al., 1998; Baker, 2004; Ziska et al., 2014; Wang et al., 2016).

Evaluating the genetic control of rice stem NSC

From the results of our greenhouse evaluation, over half of the phenotypic variation found in rice stem NSC due to genetic factors was contributed by additive genetic variance. This suggests that direct selection for rice stem NSC is possible. Other strategies to optimize stem carbohydrate dynamics may include indirect selection via correlated traits or indices. These approaches require that target traits are not merely

phenotypically correlated with stem NSC, but strongly genetically correlated. Genetic correlation arises due to pleiotropy or gametic phase disequilibrium, and this metric is often overlooked in crop physiology studies that focus on phenotypic correlation. Here we found that stem starch and sucrose have positive genetic correlation suggesting that it would be possible to select simultaneously for increased starch and sucrose. This may reflect the fact that starch synthesis in the rice stem depends on glucoses that result from sucrose breakdown, so a greater sucrose influx represents greater potential for starch accumulation. Using genetic correlations, we also find evidence that genotypes with few but heavy tillers tend to accumulate greater proportions of stem starch by the time they flower, suggesting that tiller number and weight have potential as indirect indicators of stem starch levels at heading.

Our simulation study indicates that five replicates are adequate for estimating line mean, GEBV, and variance components for rice stem NSC traits under greenhouse conditions. The mobile nature of stem NSC reserves, expressed as seasonal fluctuations in response to internal or external cues, gives rise to the idea that uncertainty measures of these traits are biologically relevant and may have a tractable genetic basis. Here we show that twice as many experimental replicates are required to estimate uncertainty measures (e.g. line CV) as are sufficient to estimate line mean and GEBV. Despite this greater replication requirement (10 replicates), studying the genetic basis of line CV is still possible under conditions similar to those in which this experiment was carried out. Here we assessed 36 genotypes replicated 20 times. Assuming the same planting constraints (pot size, spacing, etc), evaluating ten replicates necessary for line CV estimation would allow us to consider 72 genotypes,

equivalent to a medium-sized bi-parental mapping population in rice.

Applicability of NIRS for large-scale prediction of stem NSC

Given the time and labor-saving attributes of NIRS prediction for grass biomass constituents relative to traditional analytical chemistry, the reason underlying the apparent lag of NIRS application for predicting rice stem NSC is unclear. One explanation may be that multidisciplinary expertise is needed for optimal model development. Lack of domain knowledge in either NIR spectroscopy or in the application area may pose limitations on the long-term success of using NIRS models for prediction. Here, we show that NIRS calibration models can accurately predict rice stem NSC constituents across genetically and experimentally diverse samples. To our knowledge, only one study has been published that has assessed the utility of NIRS for rice stem NSC prediction, using 61 rice stem samples for starch only (Batten et al., 1993). An R^2 of 0.98 and standard error of performance of 1.5 (% starch) were reported, however, it appears that these values are likely results from cross-validation. Prediction of an unknown set of samples is critically dependent upon adequate representation of the calibration set across the spectral space of the prediction samples. To that end, for our calibration we made sure to select samples as diverse as possible, choosing across wet chemistry values, sampling points (heading versus maturity), genetic identity, and experimental conditions (two experiments across two years). With that, we expect we can predict rice stem NSC composition in future experiments under similar greenhouse conditions for large-scale genetic studies. Research avenues to further increase efficiency of rice stem NSC phenotyping may include improving

NIR prediction by tailoring calibration sets to specific prediction samples using local models (as in Godin et al., 2015) or implementing non-destructive, in-field methods of NSC determination such as hyperspectral reflectance as demonstrated in wheat (Dreccer et al., 2014).

The possibility of optimizing stem NSC dynamics for rice varietal improvement, an idea whose roots go back several decades, grows more relevant with rising concerns about the impacts of climate variability. As demonstrated previously (Yang et al., 2002; Kim et al., 2011; Morita and Nakano, 2011), post-heading temperature and water availability stresses shift rice dependencies heavily towards pre-heading carbohydrate reserves. Understanding the genetic architecture underlying stem NSC accumulation, remobilization, and re-accumulation may be key to its utility not only for rice but also for other economically important grass species.

Supplemental Table 2.1. Phenotype and germplasm information for breeder's panel

Entry	Designation	Description	Heading			Maturity		
			sugars (%)	starch (%)	TNC (%)	sugars (%)	starch (%)	TNC (%)
BP 1	AS996	IR 64 x O. rupipogon GID 707288; Acid sulpahate; BPH, Cold tolerant	4.14	13.06	17.20	3.57	0.96	4.53
BP 2	IR71700-247-1-1-2	NSIC Rc 214 Indica variety relased in the Philippines	8.92	10.97	19.90	7.52	1.89	9.41
BP 3	IR72	Dwarf, high tillering indica variety	7.14	12.46	19.60	8.09	2.08	10.16
BP 4	IR77186-148-3-4-3	Indica/japonica derivative with submergence tolerance	2.98	13.44	16.43	6.88	1.54	8.42
BP 5	IR78049-25-2-2-2	Tropical japonica with strong stems, large panicle, 100% grain filling, delayed senescence, but hard to thresh	7.03	6.31	13.34	9.55	3.01	12.56
BP 6	IR78222-20-7-148-2-B	(pyramided QTL-blast resistance) - (Chi, OXO, HSP90, POX, OXOLP)	6.75	4.84	11.59	3.74	1.05	4.78
BP 7	IR79643-39-2-2-3	Elite indica line	6.29	11.80	18.10	9.26	1.93	11.19
BP 8	IR83142-B-19-B	GSR drought tolerant	4.44	10.95	15.40	5.47	1.48	6.95
BP 9	IR83704-78-2-3-2	Indica- High yielding elite line during 2011 DS	3.60	12.26	15.86	5.43	0.85	6.28
BP 10	IR83704-86-3-3-1	Indica- High yielding elite line during 2011 DS	7.19	8.91	16.10	4.07	0.42	4.49
BP 11	BR29	Boro rice mega variety in Bangladesh (Indica)	4.49	4.12	8.62	2.77	0.06	2.83
BP 12	IR87520-44-3-1-2	Indica/japonica IRBB60 derivative	6.34	7.60	13.93	8.13	1.25	9.39
BP 13	IRBB66	IR24 background with 5 Xa genes	8.38	14.24	22.62	8.30	2.16	10.46
BP 14	IRRI127	Indica with tungro resistance	5.98	4.16	10.13	5.86	0.10	5.97
BP 15	MTU-1115	Indica high yielding from Maruteru	7.62	12.91	20.54	5.67	0.60	6.27
BP 16	NSICRc158	Indica/japonica derivative released variety in the Philippines	6.58	12.96	19.55	7.17	1.34	8.51
BP 17	NSICRc160	Indica HYV relased in the Philippines	3.09	13.00	16.09	6.28	0.97	7.25
BP 18	NSIC212	Indica HYV relased in the Philippines	7.95	6.97	14.92	9.77	3.34	13.11
BP 19	NSICRc222	Indica HYV relased in the Philippines	8.17	10.50	18.67	7.34	0.90	8.23
BP 20	NSICRC238	Indica HYV relased in the Philippines	5.62	9.63	15.25	7.17	0.53	7.70
BP 21	PSBRc18	Indica HYV relased in the Philippines	7.28	13.39	20.68	6.62	0.85	7.48
BP 22	PSBRc68	Indica HYV relased in the Philippines with submergence tolerance	9.45	9.46	18.91	4.69	0.26	4.95
BP 23	PSBRc82	Indica HYV relased in the Philippines	4.08	9.92	14.00	6.37	0.64	7.01
BP 24	TDK1	Mega var in Laos	9.52	6.71	16.23	5.80	0.26	6.06
BP 25	Tequing 1.AcclRGC78727	Indica from China	7.63	16.83	24.46	9.54	1.92	11.47
BP 26	Tequing 3.AcclRGC117912	Indica from China	1.68	17.04	18.72	3.91	0.29	4.20
BP 27	Federroz50	Indica variety in Columbia (Marker analysis indica Japonica genotype)	5.68	5.21	10.89	8.69	2.01	10.70
BP 28	YTL126	IR64 NIL with large panicle	3.22	14.56	17.78	5.53	0.87	6.40
BP 29	HHZ12-DT10-SAL1-DT1	GSR elite line	2.96	18.04	21.00	3.90	1.01	4.91
BP 30	IR60	Indica variety with low chalkiness	5.74	7.83	13.57	5.48	1.90	7.38
BP 31	IR64	Popular indica variety	7.57	13.42	20.99	6.94	1.89	8.82
BP 32	IR68	Indica variety with big grains	6.70	8.82	15.52	9.37	2.01	11.38
BP 33	IR71033-121-15	*O. sativa x O. minuta* (101141) BPH resistant introgression line	7.35	5.66	13.01	3.29	0.42	3.71
Mean			6.11	10.55	16.65	6.43	1.24	7.67

Supplemental Table 2.2. Trait measurement methodology in breeder's panel

			Sampling Point	
Measurement		**Sampling**	**50% heading**	**Harvest**
Phenological traits	days to heading	--	--	--
	growth duration	--	--	--
	grain-filling duration (growth dur - days to heading)	--	--	--
Resource allocation traits	biomass allocation (dry weights of stem, panicle, leaf)	5 hills/rep	x	x
	plant height	5 hills/rep		x
	flag leaf area	5 flag leaves/rep		x
	tiller number	5 hills/rep	x	x
	panicle number	5 hills/rep		x
Physiological traits	flag leaf chlorophyll content	5 flag leaves/rep		x
	stem NSC	5 hills/rep	x	x
	senescence score, visual scoring according to SES*	whole plot		x
Yield components	yield	100 hills/rep = $4m^2$		x
	harvest index	25 hills/rep = $1m^2$		x
	percent grain-filling	5 hills/rep		x
	1000 grain weight	1000 filled grns		x

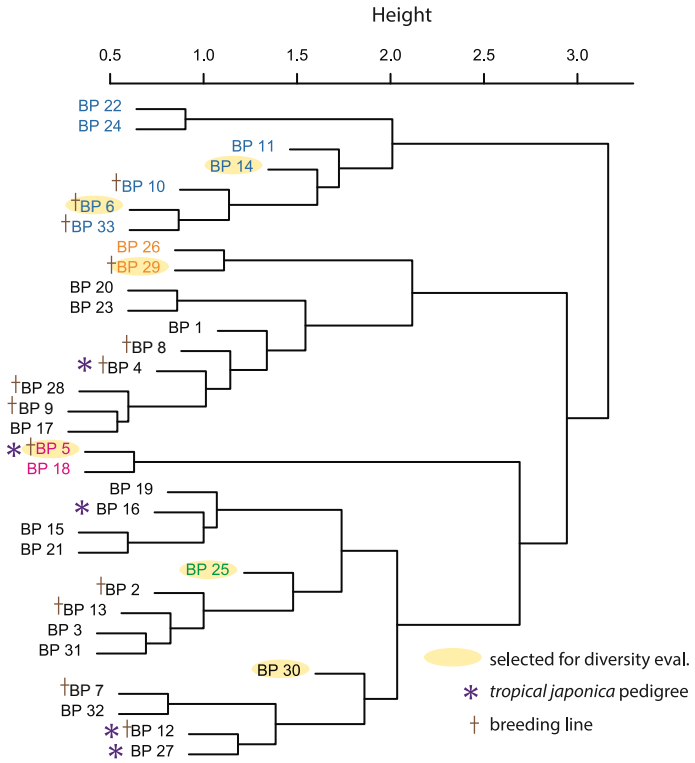
*SES is the Standard Evaluation System for rice (IRRI)

Supplemental Table 2.3. Germplasm information for diversity panel

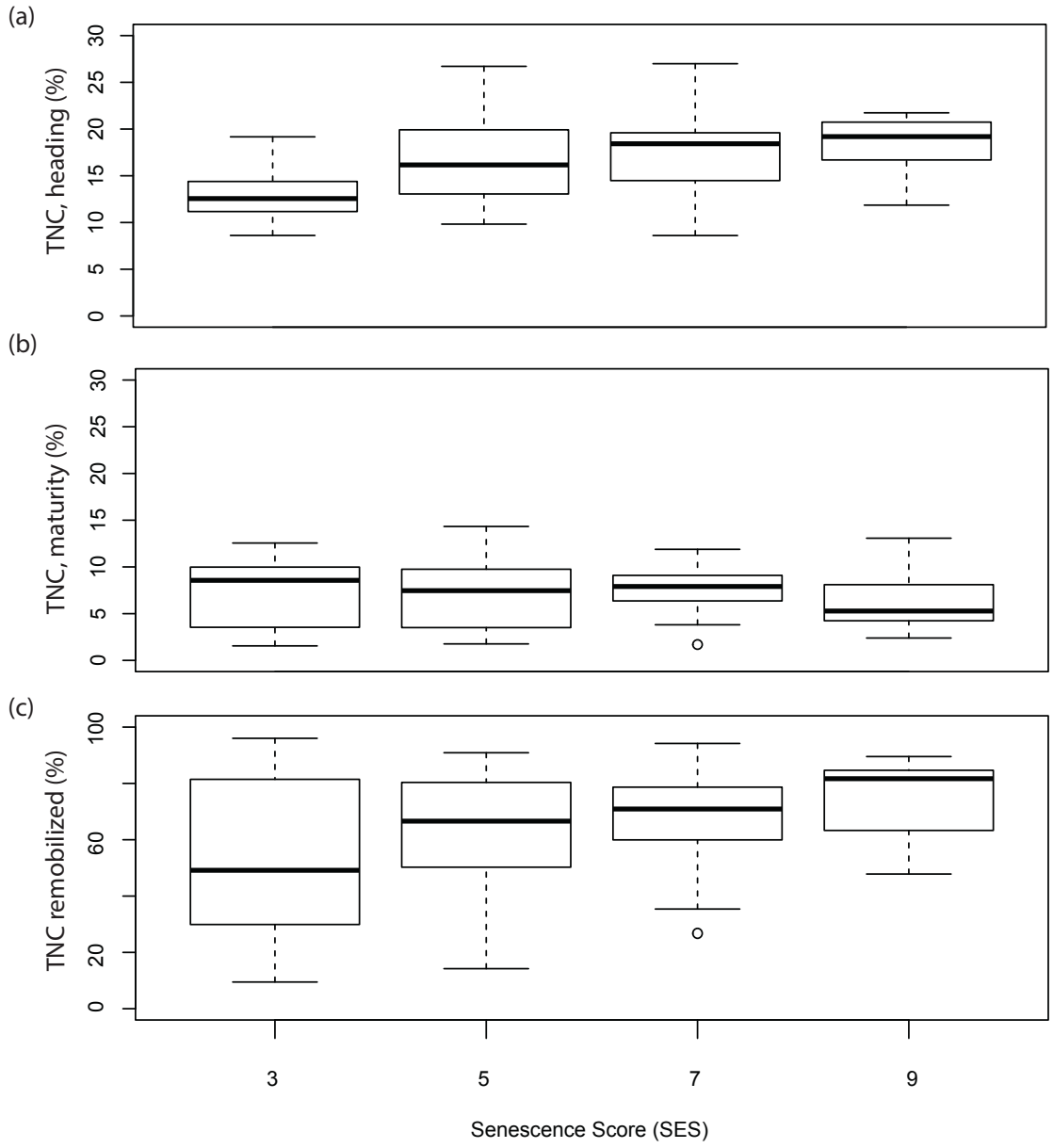
Designation*	Variety Name
NSFTV17	Binulawan
NSFTV59	Gogo Lempuk
NSFTV75	Jambu
NSFTV90	Kiang-Chou-Chiu
NSFTV107	NSF-TV 107
NSFTV130	Peh-Kuh-Tsao-Tu
NSFTV137	RTS14
NSFTV150	Sultani
NSFTV189	Criollo La Fria
NSFTV203	Radin Ebos 33
NSFTV209	Tchibanga
NSFTV222	Paraiba Chines Nova
NSFTV223	Priano Guaira
NSFTV226	IRAT 44
NSFTV235	Sze Guen Zim
NSFTV239	WAB 502-13-4-1
NSFTV252	Djimoron
NSFTV254	Hon Chim
NSFTV284	IR-44595
NSFTV285	Tox 782-20-1
NSFTV313	BR24
NSFTV325	EMATA A 16-34
NSFTV337	Sabharaj
NSFTV339	Yodanya
NSFTV356	JC 117
NSFTV385	Nira
NSFTV397	Cybonnet
NSFTV616	RT0034
NSFTV628	Jefferson
NSFTV644	IR64
BP 29	HHZ12-DT10-SAL1-DT1
BP 30	IR60
BP 25	Tequing 1.AccIRGC78727
BP 14	IRRI127
BP 5	IR78049-25-2-2-2
BP 6	IR78222-20-7-148-2-B

*NSFTV prefixes refer to accessions from the Rice Diversity Panel 1 (Zhao 2009)
BP prefixes refer to accessions from the Breeder's Panel, this paper

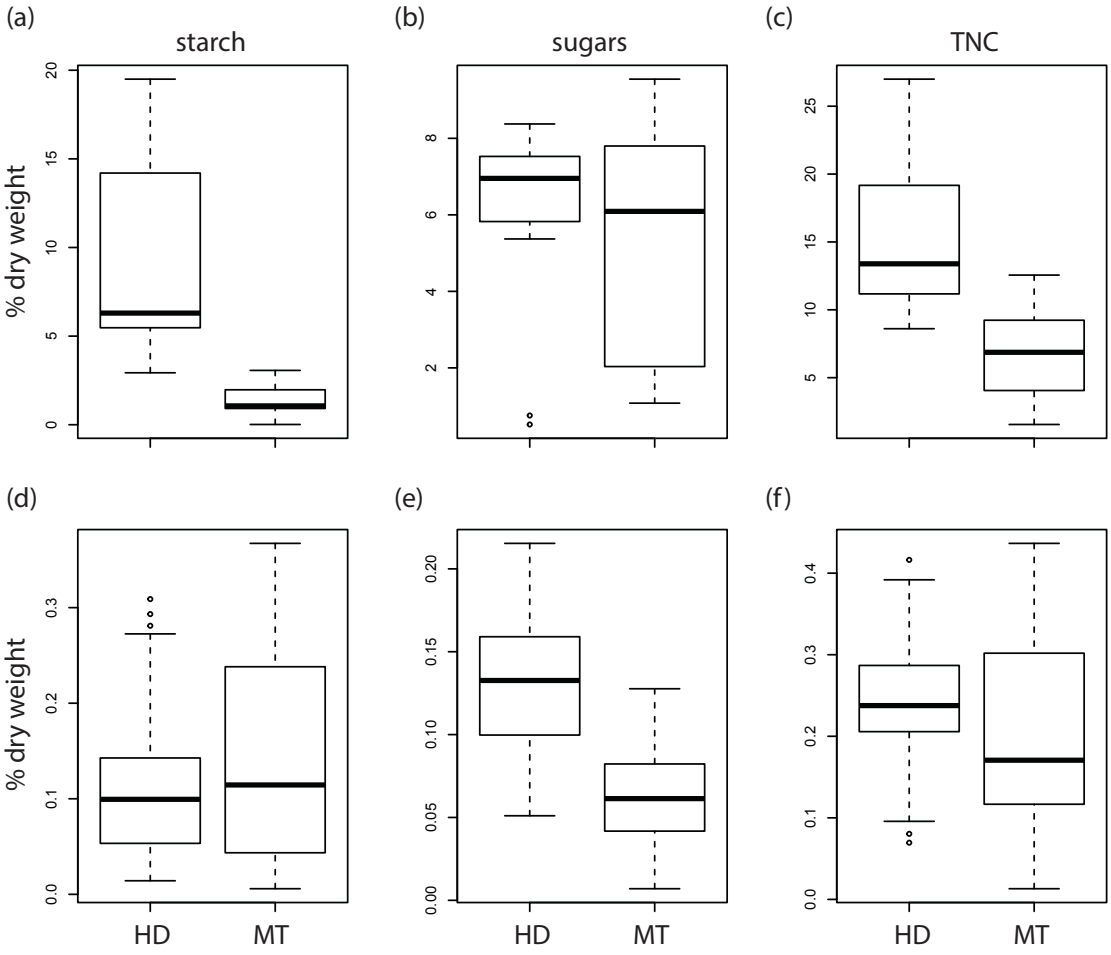
Supplemental Figure 2.1. Hierarchical clustering of breeder's panel using NSC traits. Hierarchical clustering analysis of breeder's panel using NSC traits. Results of hierarchical clustering using average linkage on scaled line means of NSC trait data on 33 breeding lines. Names are colored as in **Fig. 2.3**. Additional annotations indicate meta-information: yellow ovals are individuals selected for further evaluation with diverse germplasm, purple asterisks indicate entries with *tropical japonica* pedigree, and brown crosses refer to breeding lines (versus released variety).



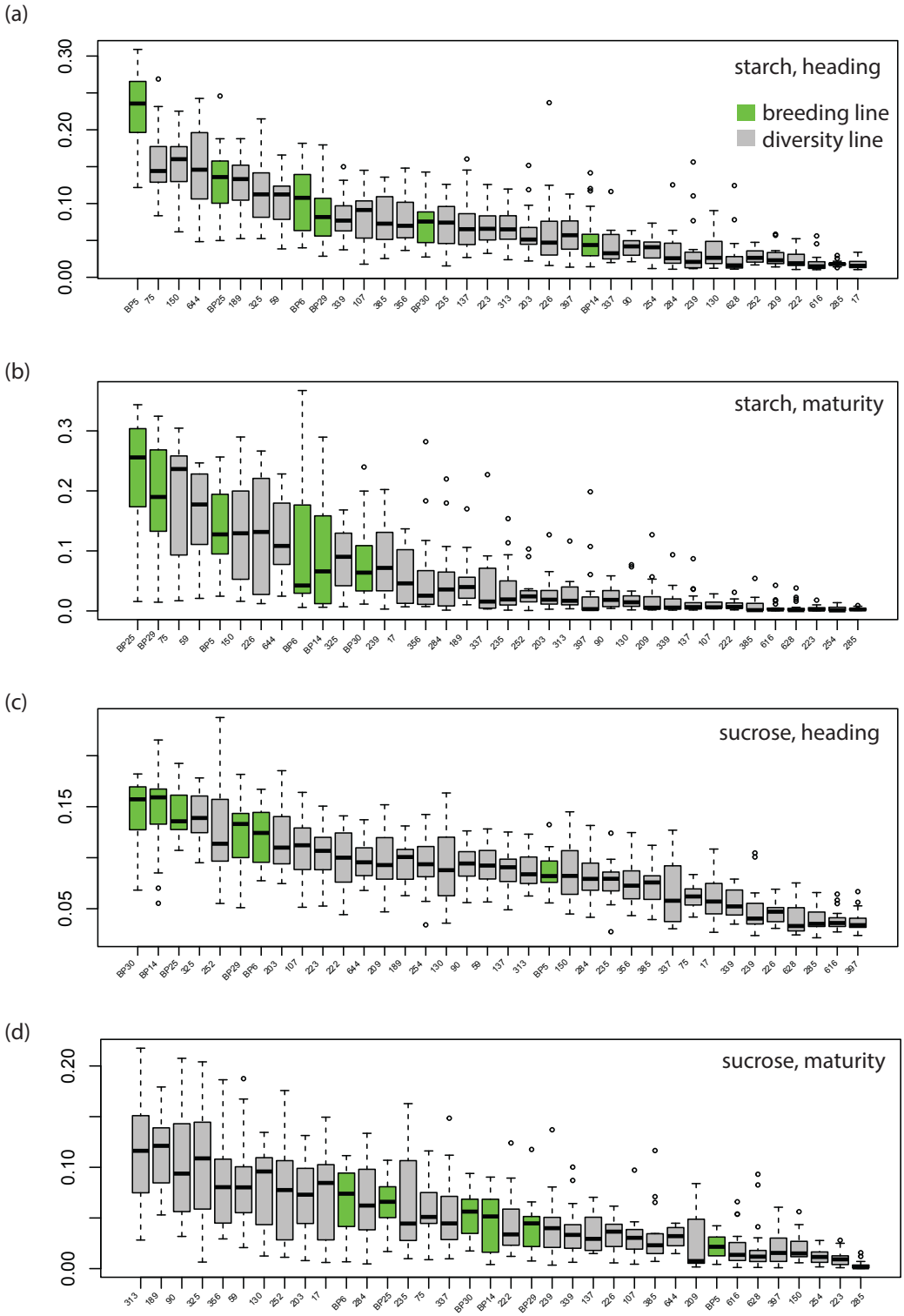
Supplemental Figure 2.2. NSC distribution by senescence class.



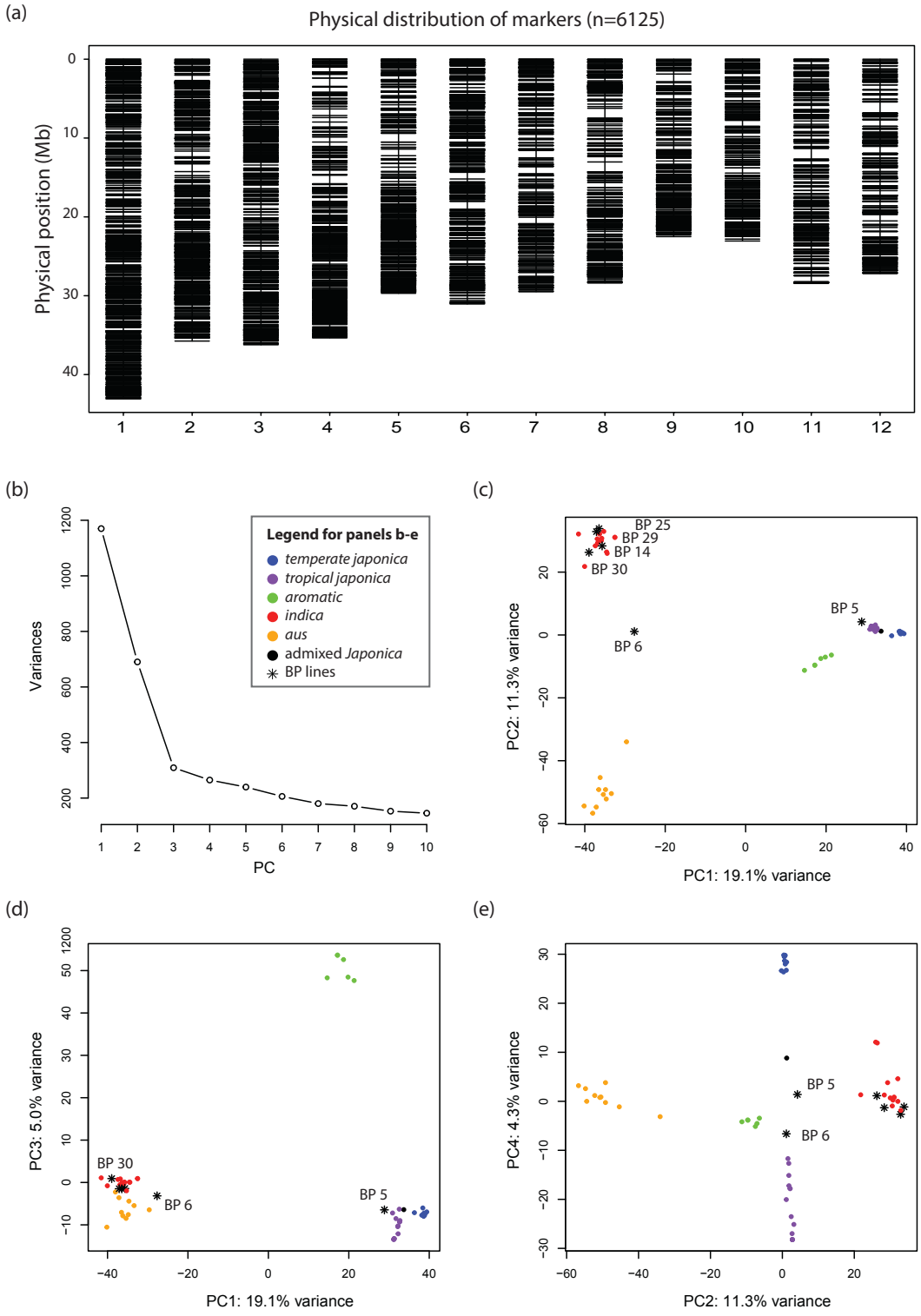
Supplemental Figure 2.3. Distribution of NSC traits by sampling point for 6 selected BP entries in the IRRI field trial (a-c) and greenhouse diversity screen (d-f)



Supplemental Figure 2.4. Accession-specific distributions of NSC traits ordered by decreasing line mean



Supplemental Figure 2.5. Genotyping results of six selected breeding lines. (a) Genome distribution of 6,125 SNPs. (b-e) Results from Principal Components Analysis.



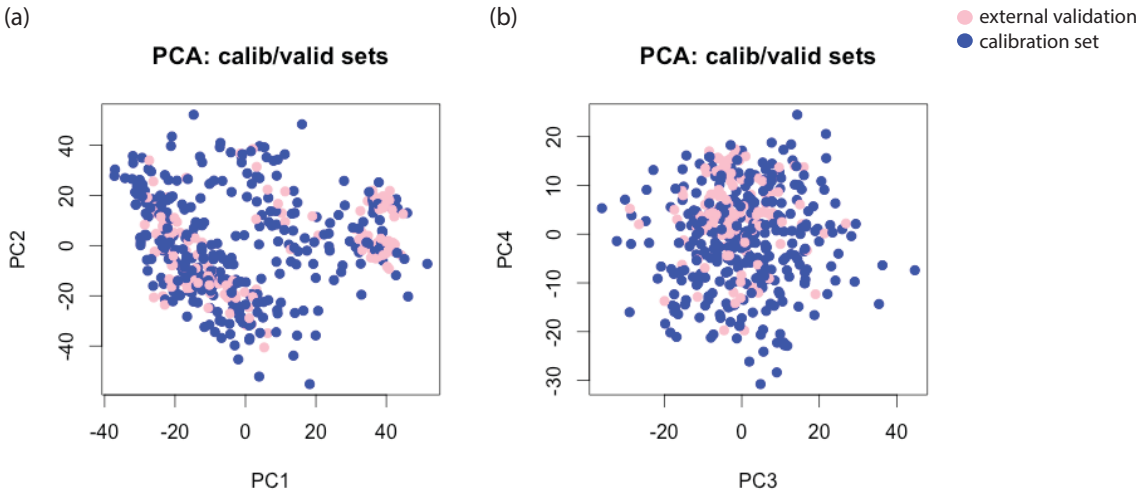
Supplemental Figure 2.6. Full description of Bayesian analyses for experimental design study

Please refer to the published article for this information.

Supplemental Figure 2.7. Zip file containing scripts and related files for Bayesian analyses

Please refer to the published article for this information.

Supplemental Figure 2.8. Comparison of calibration and validation sets for NIR calibration



REFERENCES

Arbelaez JD, Moreno LT, Singh N, Tung C-W, Maron LG, Ospina Y, Martinez CP, Grenier C, Lorieux M, McCouch S (2015) Development and GBS-genotyping of introgression lines (ILs) using two wild species of rice, *O. meridionalis* and *O. rufipogon*, in a common recurrent parent, *O. sativa* cv. Curinga. Mol Breed 35: 1–18

Baker JT (2004) Yield responses of southern US rice cultivars to CO_2 and temperature. Agric For Meteorol 122: 129–137

Batten GD, Blakeney AB, McGrath VB, Ciaverella S (1993) Non-structural carbohydrate: Analysis by near infrared reflectance spectroscopy and its importance as an indicator of plant growth. Plant Soil 155/156: 269–272

Bidinger F, Musgrave RB, Fischer RA (1977) Contribution of stored pre-anthesis assimilate to grain yield in wheat and barley. Nature 270: 431–433

Bruuinsma J (1963) The quantitative analysis of chlorophylls a and b in plant extracts. Photochem Photobiol 2: 241–249

Chen H-J, Wang S-J (2008) Molecular regulation of sinksource transition in rice leaf sheaths during the heading period. Acta Physiol Plant 30: 639–649

Cock JH, Yoshida S (1972) Accumulation of 14C-labelled Carbohydrate Before Flowering and its Subsequent Redistribution and Respiration in the Rice Plant. Proc. Crop Sci. Soc. Japan 41:

Van Dat T, Peterson ML (1983) Performance of Near-Isogenic Genotypes of Rice Differing in Growth Duration. II. Carbohydrate Partitioning During Grain Filling. Crop Sci. 23:

Dharmawardhana P, Ren L, Amarasinghe V, Monaco M, Thomason J, Ravenscroft D, McCouch S, Ware D, Jaiswal P (2013) A genome scale metabolic network for rice and accompanying analysis of tryptophan, auxin and serotonin biosynthesis regulation under biotic stress. Rice 6: 1–15

Dreccer MF, Barnes LR, Meder R (2014) Quantitative dynamics of stem water soluble carbohydrates in wheat can be monitored in the field using hyperspectral reflectance. F Crop Res 159: 70–80

Duitama J, Silva A, Sanabria Y, Cruz DF, Quintero C, Ballen C, Lorieux M, Scheffler B, Farmer A, Torres E (2015) Whole genome sequencing of elite rice cultivars as a comprehensive information resource for marker assisted selection. PLoS One 10: e0124617

Fu J, Huang Z, Wang Z, Yang J, Zhang J (2011) Pre-anthesis non-structural carbohydrate reserve in the stem enhances the sink strength of inferior spikelets during grain filling of rice. F Crop Res 123: 170–182

Fulkerson WJ, Donaghy DJ (2001) Plant-soluble carbohydrate reserves and senescence - key criteria for developing an effective grazing management system for ryegrass-based pastures: a review. Aust J Exp Agric 41: 261–275

Godin B, Agneessens R, Delcarte J, Dardenne P (2015) Prediction of chemical characteristics of fibrous plant biomasses from their near infrared spectrum: comparing local versus partial least square models and cross-validation versus independent validations. J Near Infrared Spectrosc 23: 1–14

Hirano T, Saito Y, Ushimaru H, Michiyama H (2005) The Effect of the Amount of Nitrogen Fertilizer on Starch Metabolism in Leaf Sheath of *Japonica* and *Indica* Rice Varieties during the Heading Period. Plant Prod Sci 8: 122–130

IRRI (2012) Weather Summary. IRRI Annu. Rep.

Jacquemin J, Bhatia D, Singh K, Wing RA (2013) The International Oryza Map Alignment Project: development of a genus-wide comparative genomics platform to help solve the 9 billion-people question. Curr Opin Plant Biol 16: 147–156

Kanbe T, Sasaki H, Aoki N, Yamagishi T, Ohsugi R (2009) The QTL Analysis of RuBisCO in Flag Leaves and Non-Structural Carbohydrates in Leaf Sheaths of Rice Using Chromosome Segment Substitution Lines and Backcross Progeny F2 Populations. Plant Prod Sci 12: 224–232

Kashiwagi T, Ishimaru K (2004) Identification and functional analysis of a locus for improvement of lodging resistance in rice. Plant Physiol 134: 676–683

Kashiwagi T, Madoka Y, Hirotsu N, Ishimaru K (2006) Locus prl5 improves lodging resistance of rice by delaying senescence and increasing carbohydrate reaccumulation. Plant Physiol Biochem 44: 152–157

Kashiwagi T, Togawa E, Hirotsu N, Ishimaru K (2008) Improvement of lodging resistance with QTLs for stem diameter in rice (Oryza sativa, L.). TAG Theor Appl Genet 117: 749–757

Kim J, Shon J, Lee C-K, Yang W, Yoon Y, Yang W-H, Kim Y-G, Lee B-W (2011) Relationship between grain filling duration and leaf senescence of temperate rice under high temperature. F Crop Res 122: 207–213

Laza MRC, Peng S, Akita S, Saka H (2003) Contribution of Biomass Partitioning and Translocation to Grain Yield under Sub-Optimum Growing Conditions in Irrigated Rice. Plant Prod Sci 6: 28–35

Li J-Y, Wang J, Zeigler RS (2014) The 3,000 rice genomes project: new opportunities and challenges for future rice research. Gigascience 3: 1–3

McCouch SR, Wright MH, Tung C-W, Maron LG, McNally KL, Fitzgerald M, Singh N, DeClerck G, Agosto-Perez F, Korniliev P, et al (2016) Open access resources for genome-wide association mapping in rice. Nat Commun 7: doi:10.1038/ncomms10532

Moore PH (1995) Temporal and Spatial Regulation of Sucrose Accumulation in the Sugarcane Stem. Funct Plant Biol 22: 661–679

Morita S, Nakano H (2011) Nonstructural Carbohydrate Content in the Stem at Full Heading Contributes to High Performance of Ripening in Heat-Tolerant Rice Cultivar Nikomaru. Crop Sci 51: 818–828

MOYA TB, Ziska LH, NAMUCO OS, OLSZYK D (1998) Growth dynamics and genotypic variation in tropical, field‐grown paddy rice (*Oryza sativa* L.) in response to increasing carbon dioxide and temperature. Glob Chang Biol 4: 645–656

Nagata K, Shimizu H, Terao T (2002) Quantitative Trait Loci for Nonstructural Carbohydrate Accumulation in Leaf Sheaths and Culms of Rice (Oryza sativa L.) and their Effects on Grain Filling. Breed Sci 52: 275–283

Pan J, Cui K, Wei D, Huang J, Xiang J, Nie L (2011) Relationships of non-structural carbohydrates accumulation and translocation with yield formation in rice recombinant inbred lines under two nitrogen levels. Physiol Plant 141: 321–331

Pask A, Pietragalla J, Mullan D, Reynolds M (2012) Physiological Breeding II: A Field Guide to Wheat Phenotyping. Mexico, D.F.: CIMMYT

Project IRGS (2005) The map-based sequence of the rice genome. Nature 436: 793–800

Reynolds M, Pask A, Mullan D (2012) Physiological Breeding I: Interdisciplinary Approaches to Improve Crop Adaptation. Mexico, D.F.: CIMMYT

Ruuska SA, Rebetzke GJ, van Herwaarden AF, Richards RA, Fettell NA, Tabe L, Jenkins CLD (2006) Genotypic variation in water-soluble carbohydrate accumulation in wheat. Funct Plant Biol 33 : 799–809

Samonte SOPB, Wilson LT, MCCLUNG AM, Tarpley L (2001) Seasonal Dynamics of Nonstructural Carbohydrate Partitioning in 15 Diverse Rice Genotypes. Crop Sci 41: 902–909

Schatz MC, Maron LG, Stein JC, Wences AH, Gurtowski J, Biggers E, Lee H, Kramer M, Antoniou E, Ghiban E (2014) Whole genome de novo assemblies of three divergent strains of rice, Oryza sativa, document novel gene space of aus and indica. Genome Biol 15: 1–16

Setter TL, Peng S, Kirk GJD, Virmani SS, Kropff MJ, Cassman KG (1994) Physiological considerations and hybrid rice. Break yield barrier Int Rice Res Inst, Los Baños, Philipp 39–62

Slewinski TL (2012) Non-structural carbohydrate partitioning in grass stems: a target to increase yield stability, stress tolerance, and biofuel production. J Exp Bot 63: 4647–4670

Trinder P (1969) Determination of blood glucose using an oxidase-peroxidase system with a non-carcinogenic chromogen. J Clin Pathol 22: 158–161

Waclawovsky AJ, Sato PM, Lembke CG, Moore PH, Souza GM (2010) Sugarcane for bioenergy production: an assessment of yield and regulation of sucrose content. Plant Biotechnol J 8: 263–276

Wang DR, Bunce JA, Tomecek MB, Gealy D, McClung A, McCouch SR, Ziska LH (2016) Evidence for divergence of response in Indica, Japonica, and wild rice to high CO2 x temperature interaction. Glob Chang Biol 22: 2620–2632

Wang Z, Liu X, Li R, Chang X, Jing R (2011) Development of Near-Infrared Reflectance Spectroscopy Models for Quantitative Determination of Water-Soluble Carbohydrate Content in Wheat Stem and Glume. Anal Lett 44: 2478–2490

Wolfrum E, Payne C, Stefaniak T, Rooney W, Dighe N, Bean B, Dahlberg J (2013) Multivariate calibration models for sorghum composition using near-infrared spectroscopy. Contract 303: 275–3000

Yang J, Zhang J, Huang Z, Zhu Q, Wang L (2000) Remobilization of Carbon Reserves is Improved by Controled Soil-drying during Grain-filling of Wheat. Crop Sci 40: 1645

Yang J, Zhang J, Liu L, Wang Z, Zhu Q (1999) Carbon Remobilization and Grain Filling in Japonica / Indica Hybrid Rice Subjected to Postanthesis Water Deficits. 102–109

Yang J, Zhang J, Wang Z, Zhu Q, Liu L (2002) Abscisic acid and cytokinins in the root exudates and leaves and their relationship to senescence and remobilization of carbon reserves in rice subjected to water stress during grain filling. Planta 215: 645–652

Yang J, Zhang J, Wang Z, Zhu Q, Wang W (2001) Remobilization of carbon reserves in response to water deficit during grain filling of rice. F Crop Res 71: 47–55

Ziska LH, Tomecek MB, Gealy DR (2014) Assessment of cultivated and wild, weedy rice lines to concurrent changes in CO2 concentration and air temperature: determining traits for enhanced seed yield with increasing atmospheric CO2. Funct Plant Biol 41: 236–243

CHAPTER 3:

GENOTYPING AND POPULATION STRUCTURE ANALYSIS OF U.S. RICE

INTRODUCTION

This chapter describes the population structure of a rice panel (n=153) that reflects the historical diversity of germplasm grown in the southern U.S.. The genotyping effort was undertaken as part of a USDA-AFRI funded project (USDA-NIFA #2014-67003-21858: Characterization of Genomic Signatures for Rice Crop Resilience in Response to Climate Change in the U.S.) to model historical market yield (county-level yields) on genetic markers and weather variables from 1970-2014, with the objectives of identifying key climatic and genetic factors, and their interactions, that determine rice yields in the Southern U.S. In addition to 106 accessions that map directly to varieties grown from 1970-2014 (the time period of modeling interest), we included 47 accessions that were either new varieties, parents used in modern crosses, parents used for older crosses, or U.S. rice progenitors. Taken together, these 153 accessions effectively represent the entire Southern U.S. rice gene pool. For the purpose of this chapter, I only present the genotypic and population structure analysis of this panel, based on Single Nucleotide Polymorphism (SNP) markers, as these data will be used in the following chapter on stem non-structural carbohydrate genetic architecture. Results indicate that genetic variation is largely aligned with market class (long, medium or short grain), supporting previous findings from a U.S. rice population structure study that used Simple Sequence Repeat markers (Lu et al., 2005).

MATERIALS AND METHODS

Plant material: Germplasm information is found in **Supplemental Table 3.1** and an overview is provided in **Figure 3.1**. FULL set refers to the entire panel of 153 accessions while ACREAGE refers to the subset of 106 accessions that map to the variety acreage datasets that will be used for future modeling. Seeds were imported

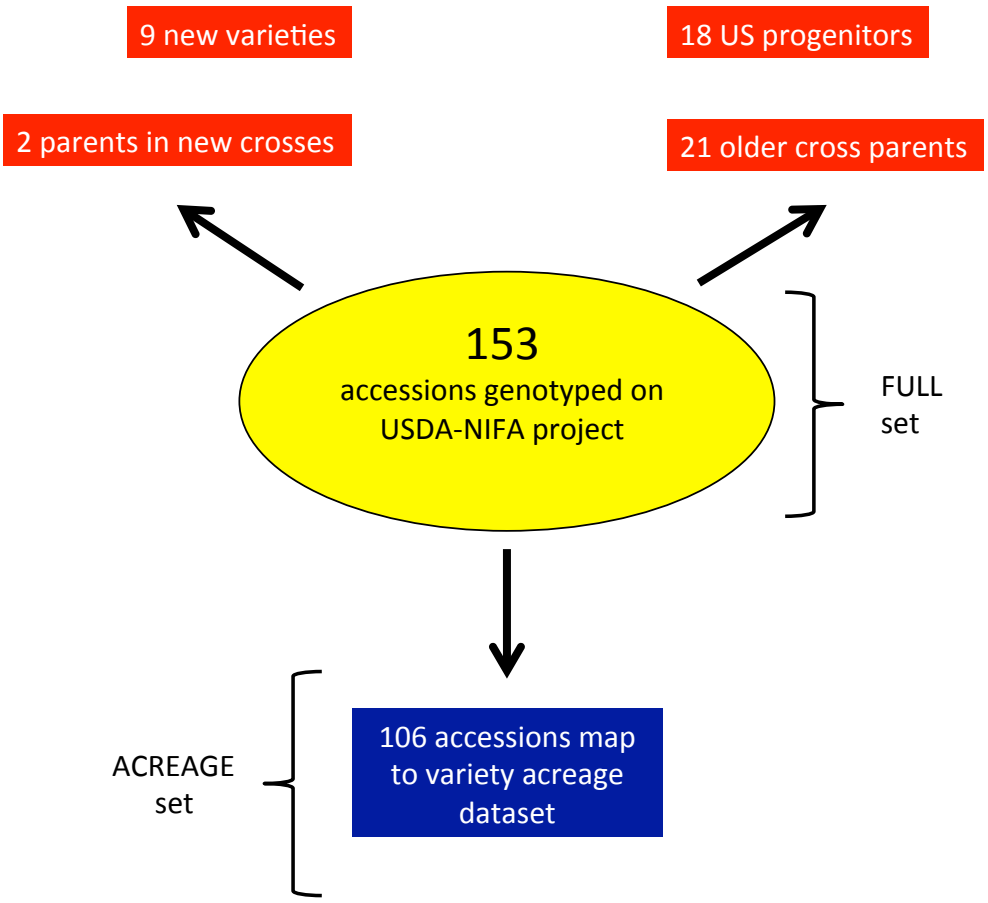


Figure 3.1. Overview of the U.S. historical rice panel. FULL set refers to the entire panel of 153 accessions while ACREAGE refers to the subset of 106 accessions that map to the variety acreage datasets that will be used for future modeling. Note: these categories are not mutually exclusive e.g. an accession may be relevant to variety acreage names and is an older cross parent.

from Dale Bumpers National Rice Research Center (DBNRRC) and used for genotypic analysis, except for 48 accessions that had been previously genotyped on the high-density rice array (McCouch et al., 2016), in which case the seed source from the McCouch seed bank at Cornell University was used instead of the seed from DBNRRC. Plants were grown in Guterman Greenhouse in Ithaca, NY during the summer of 2014 and tissue collected from a single plant per accession. Purified seeds (selfed seeds to eliminate residual heterozygosity) were then collected from the individual that was genotyped, and sent back to DBNRRC for amplification, photographing, and distribution.

Genotyping: DNA was extracted using Qiagen DNeasy kits (Qiagen, Hilden, Germany). Samples were genotyped using 96-plex Genotyping-By-Sequencing (GBS) (Elshire et al., 2011). Sample preparation and bioinformatics analysis to obtain SNP calls followed the procedure described by Spindel et al. (2013). Genotypes for the FULL set were called along with 54 *O. sativa* subpopulation controls at minor allele frequency (MAF) of 0.01. Missing data were imputed using a TASSEL plugin, FastImputationBitFixedWindowPlugin. The resulting dataset yielded 111,500 SNP sites at 0.01 MAF across the 207 accessions.

Analysis: fastSTRUCTURE software (Raj et al., 2014) was employed to infer the ancestry of the FULL panel along with the 54 *O. sativa* subpopulation controls with the unfiltered dataset of 111,500 SNP variants. Principal Component Analysis using R statistical software (https://www.**r**-project.org/) and distance tree generation using Geneious 7.0.4 (Kearse et al., 2012) were done using filtered sets of SNPs (see **Table 3.1**). SNPs were filtered with TASSEL5. LD-pruning was performed using Plink1.9

with 25-SNP sliding windows, pruned at r^2>0.5 with step size = 5 SNPs.

Table 3.1. Number of SNPs used for different population structure analyses.

Analysis	Call rate	Minor allele freq.	Number of sites
fastSTRUCTURE	NA	NA	111,500
Distance tree	90%	NA	9004
PCA	90%	0.05	4704

RESULTS

Population structure of U.S. rice

Three methods were employed to examine high-level population structure in U.S. rice: fastSTRUCTURE, distance analysis, and PCA. These analyses utilized different subsets of SNPs (see **Methods** and **Table 3.1**) but results were congruent. fastSTRUCTURE determined the ideal number of model components (or populations) to be six, using the full set of 111,500 markers. At K=6, the *aus, indica, aromatic*, and *temperate japonica* subpopulation control individuals, whose subpopulation memberships were determined using HDRA information (McCouch et al., 2016), were distinguished (**Figure 3.2**). Interestingly, *tropical japonica* controls were divided into two populations, referred to here as *tropical japonica-1* and *tropical japonica-2*. The split is also observed in the U.S. accessions. This extra separation within the *tropical japonicas* is likely due to the overwhelming representation of this subpopulation in the panel as a whole. From these results, we see that U.S. rice has mainly *tropical japonica* ancestry. There are also 16 US rice varieties classified as *indica* and 6 as

temperate japonica, however, these are primarily varieties used as parents in crosses rather than as varieties for production. In addition, there were 13 admixed individuals (i.e. <0.8 inferred ancestry for a single group) (**Table 3.2**). Relationships between U.S. rice accessions and subpopulation controls using a subset of 9,004 sites (**Figure 3.3**) and 4,074 sites (**Figure 3.4**) revealed largely the same picture.

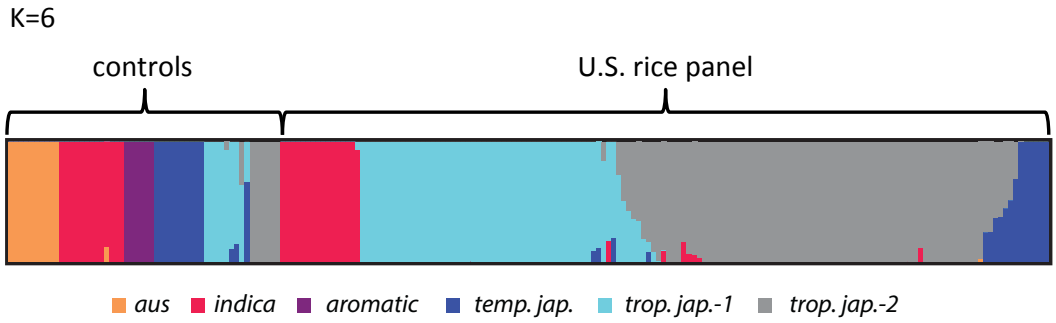


Figure 3.2. fastSTRUCTURE analysis of 153 U.S. rice accessions and 54 diverse *O. sativa* subpopulation controls at K=6. Number of model components (K=6) was selected using fastSTRUCTURE's built-in algorithm. Population controls are displayed on the left side of the plot and U.S. accessions on the right.

Figure 3.3 Distance tree of 153 U.S. rice accessions and *indica*, *tropical japonica*, and *temperate japonica* subpopulation controls. Colored branches indicate subpopulation control individuals (red = *indica*, dark blue= *temperate japonica*, light blue = *tropical japonica*) and black are U.S. individuals.

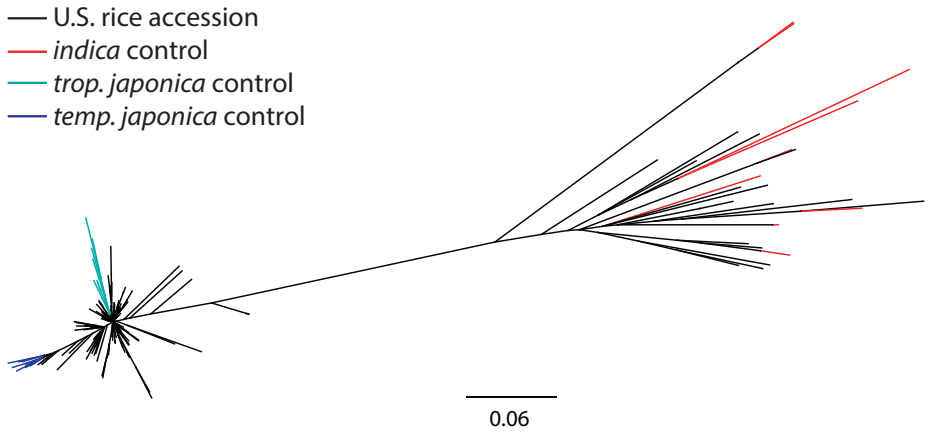


Figure 3.4. Principal components analysis of 153 rice accessions and 54 diverse *O. sativa* subpopulation controls. The first two components explained 32% of the total variation. Subpopulation controls are colored (red=*indica*, orange=*aus*, purple=*aromatic*, dark blue= *temperate japonica*, light blue= *tropical japonica*) and U.S. accessions are gray.

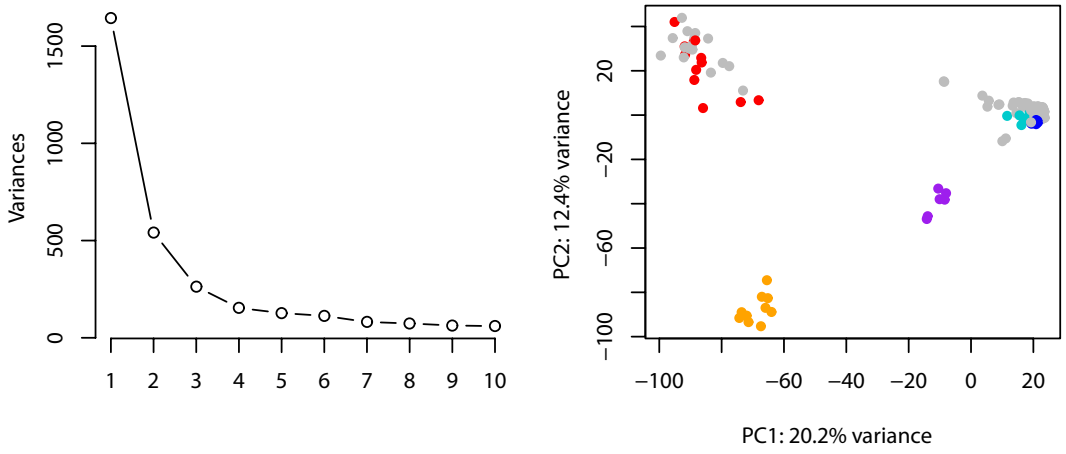


Table 3.2. Inferred ancestry at K=6 of U.S. rice panel (n=153) genotyped using Genotyping-By-Sequencing. Column 'subpop' indicates assigned subpopulation based on this analysis. The 54 subpopulation control individuals are not included in this table but were used to associate fastStructure populations to the five known subpopulations of cultivated Asian rice. Individuals that are strictly Southern U.S. rice progenitors or parents used in crosses between 1980-present (i.e., not found on acreage reports for this time period) are denoted with an asterisk.

name	aus	ind	aro	tej	trj-1	trj-2	subpop
IR36	0.000001	0.999993	0.000001	0.000001	0.000001	0.000001	IND *
IRGA409	0.000001	0.999993	0.000001	0.000001	0.000001	0.000001	IND *
JASMINE_85	0.000001	0.999993	0.000001	0.000001	0.000001	0.000001	IND
JES	0.000001	0.999993	0.000001	0.000001	0.000001	0.000001	IND
RONDO	0.000001	0.999993	0.000001	0.000001	0.000001	0.000001	IND
SHUFENG_121-1655	0.000001	0.999993	0.000001	0.000001	0.000001	0.000001	IND
LA110	0.000002	0.999992	0.000002	0.000002	0.000002	0.000002	IND *
NIRA	0.000002	0.999992	0.000002	0.000002	0.000002	0.000002	IND *
ZHE733	0.000002	0.999992	0.000002	0.000002	0.000002	0.000002	IND *
TADUCAN	0.000002	0.999991	0.000002	0.000002	0.000002	0.000002	IND *
TAICHUNG NATIVE	0.000002	0.999991	0.000002	0.000002	0.000002	0.000002	IND *
TESANAI 2	0.000002	0.999991	0.000002	0.000002	0.000002	0.000002	IND
GUICHAO-2	0.000002	0.99999	0.000002	0.000002	0.000002	0.000002	IND *
MILAGRO_FILIPINO	0.000002	0.999989	0.000002	0.000002	0.000002	0.000002	IND
TEQING	0.000002	0.999989	0.000002	0.000002	0.000002	0.000002	IND *
SLO17	0.000002	0.9385	0.000002	0.000002	0.061491	0.000002	IND *
EDITH	0.000001	0.000001	0.000001	0.000001	0.999994	0.000001	TRJ-1 *
BLUE_ROSE_SUPREME	0.000001	0.000001	0.000001	0.000001	0.999993	0.000002	TRJ-1
BLUE_ROSE	0.000001	0.000001	0.000001	0.000001	0.999993	0.000001	TRJ-1
CAROLINA_GOLD	0.000001	0.000001	0.000001	0.000001	0.999993	0.000001	TRJ-1 *
CAROLINA_GOLD	0.000001	0.000001	0.000001	0.000001	0.999993	0.000001	TRJ-1 *
CAROLINA_GOLD_SELECT	0.000001	0.000001	0.000001	0.000001	0.999993	0.000001	TRJ-1 *
LADY_WRIGHT_SELECTION	0.000001	0.000001	0.000001	0.000001	0.999993	0.000001	TRJ-1 *
NOVA	0.000001	0.000001	0.000001	0.000001	0.999993	0.000001	TRJ-1
TEMPLETON	0.000001	0.000001	0.000001	0.000001	0.999993	0.000001	TRJ-1
BLUEBELLE	0.000002	0.000002	0.000002	0.000002	0.999992	0.000002	TRJ-1
BRAZOS	0.000002	0.000002	0.000002	0.000002	0.999992	0.000002	TRJ-1
CAFFEY	0.000002	0.000002	0.000002	0.000002	0.999992	0.000002	TRJ-1

EARL	0.000002	0.000002	0.000002	0.000002	0.999992	0.000002	TRJ-1
HONDURAS	0.000002	0.000002	0.000002	0.000002	0.999992	0.000002	TRJ-1 *
MELROSE	0.000002	0.000002	0.000002	0.000002	0.999992	0.000002	TRJ-1
NORTHROSE	0.000002	0.000002	0.000002	0.000002	0.999992	0.000002	TRJ-1 *
TP49	0.000002	0.000002	0.000002	0.000002	0.999992	0.000002	TRJ-1
ANTONIO	0.000002	0.000002	0.000002	0.000002	0.999991	0.000002	TRJ-1
BOWMAN	0.000002	0.000002	0.000002	0.000002	0.999991	0.000002	TRJ-1
CL111	0.000002	0.000002	0.000002	0.000002	0.999991	0.000002	TRJ-1
GULFROSE	0.000002	0.000002	0.000002	0.000002	0.999991	0.000002	TRJ-1
JUPITER	0.000002	0.000002	0.000002	0.000002	0.999991	0.000002	TRJ-1
LACROSSE	0.000002	0.000002	0.000002	0.000002	0.999991	0.000002	TRJ-1 *
LAGRUE	0.000002	0.000002	0.000002	0.000002	0.999991	0.000002	TRJ-1
LEMONT	0.000002	0.000002	0.000002	0.000002	0.999991	0.000002	TRJ-1
NEPTUNE	0.000002	0.000002	0.000002	0.000002	0.999991	0.000002	TRJ-1
PRESIDIO	0.000002	0.000002	0.000002	0.000002	0.999991	0.000002	TRJ-1
SATURN	0.000002	0.000002	0.000002	0.000002	0.999991	0.000002	TRJ-1
STARBONNET	0.000002	0.000002	0.000002	0.000002	0.999991	0.000002	TRJ-1
TORO	0.000002	0.000002	0.000002	0.000002	0.999991	0.000002	TRJ-1
VEGOLD	0.000002	0.000002	0.000002	0.000002	0.999991	0.000002	TRJ-1
VISTA	0.000002	0.000002	0.000002	0.000002	0.999991	0.000002	TRJ-1
ARKROSE	0.000002	0.000002	0.000002	0.000002	0.99999	0.000002	TRJ-1
BELLE_PATNA	0.000002	0.000002	0.000002	0.000002	0.99999	0.000002	TRJ-1
BLUEBONNET	0.000002	0.000002	0.000002	0.000002	0.99999	0.000002	TRJ-1
BOND	0.000002	0.000002	0.000002	0.000002	0.99999	0.000002	TRJ-1
COLUSA	0.000002	0.000002	0.000002	0.000002	0.99999	0.000002	TRJ-1
DELLA_2	0.000002	0.000002	0.000002	0.000002	0.99999	0.000002	TRJ-1
DIXIEBELLE	0.000002	0.000002	0.000002	0.000002	0.99999	0.000002	TRJ-1
IMPROVED_BLUE_ROSE	0.000002	0.000002	0.000002	0.000002	0.99999	0.000002	TRJ-1 *
LAFITTE	0.000002	0.000002	0.000002	0.000002	0.99999	0.000002	TRJ-1
NATO	0.000002	0.000002	0.000002	0.000002	0.99999	0.000002	TRJ-1
NOVA_76	0.000002	0.000002	0.000002	0.000002	0.99999	0.000002	TRJ-1 *
ORION	0.000002	0.000002	0.000002	0.000002	0.99999	0.000002	TRJ-1
BELLEMONT	0.000002	0.000002	0.000002	0.000002	0.999989	0.000002	TRJ-1
MERCURY	0.000002	0.000002	0.000002	0.000002	0.999989	0.000002	TRJ-1
RICO1	0.000002	0.000002	0.000002	0.090693	0.9093	0.000002	TRJ-1
PECOS	0.000002	0.000002	0.000002	0.112461	0.887533	0.000002	TRJ-1 *
PANDA	0.000001	0.000001	0.000001	0.002378	0.844152	0.153466	TRJ-1 *
V7817	0.000002	0.173648	0.000002	0.000002	0.826344	0.000002	TRJ-1

PIROGUE	0.000002	0.000002	0.000002	0.202158	0.797836	0.000002	admix
REXARK_	0.000002	0.000002	0.000002	0.000002	0.727002	0.272991	admix *
CAROLINA_GOLD_SELECT	0.000002	0.000002	0.000002	0.000002	0.508398	0.491596	admix *
FORTUNA_	0.000001	0.000001	0.000001	0.000001	0.424693	0.575302	admix *
SINAMPAGA_SELECTION	0.000001	0.000001	0.000001	0.000001	0.361341	0.638653	admix *
HILL_LONG_GRAIN	0.000001	0.000001	0.000001	0.000001	0.3389	0.661094	admix *
DELLA	0.000002	0.000002	0.000002	0.000002	0.192767	0.807227	TRJ-2
ZENITH	0.000002	0.000002	0.000002	0.081814	0.085447	0.832734	TRJ-2 *
SABINE	0.000002	0.000002	0.000002	0.000002	0.082407	0.917586	TRJ-2
CL121	0.000001	0.000001	0.000001	0.000001	0.010141	0.989853	TRJ-2
A-301	0.000001	0.092066	0.000001	0.000001	0.003758	0.904171	TRJ-2 *
NECHES	0.000001	0.000002	0.000002	0.000002	0.000523	0.999471	TRJ-2 *
ROSEMONT	0.000001	0.000001	0.000001	0.000001	0.000388	0.999607	TRJ-2
HIDALGO	0.000004	0.000004	0.000004	0.000004	0.000004	0.999981	TRJ-2
CB801	0.000002	0.168603	0.000002	0.000002	0.000002	0.831389	TRJ-2
JAZZMAN	0.000002	0.067485	0.000002	0.000002	0.000002	0.932505	TRJ-2
TRENASSE	0.000002	0.056116	0.000002	0.000002	0.000002	0.943878	TRJ-2
NEWREX	0.000002	0.032706	0.000002	0.000002	0.000002	0.967287	TRJ-2
AHRENT	0.000002	0.000002	0.000002	0.000002	0.000002	0.999992	TRJ-2
ALAN	0.000002	0.000002	0.000002	0.000002	0.000002	0.999992	TRJ-2
BELLEVUE	0.000002	0.000002	0.000002	0.000002	0.000002	0.999992	TRJ-2
CENTURY_PATNA_231	0.000002	0.000002	0.000002	0.000002	0.000002	0.999992	TRJ-2
CL131	0.000002	0.000002	0.000002	0.000002	0.000002	0.999992	TRJ-2
CL142_AR	0.000002	0.000002	0.000002	0.000002	0.000002	0.999992	TRJ-2
CL161	0.000002	0.000002	0.000002	0.000002	0.000002	0.999992	TRJ-2
CL162	0.000002	0.000002	0.000002	0.000002	0.000002	0.999992	TRJ-2
COLORADO	0.000002	0.000002	0.000002	0.000002	0.000002	0.999992	TRJ-2
CYBONNET	0.000002	0.000002	0.000002	0.000002	0.000002	0.999992	TRJ-2
DAWN	0.000002	0.000002	0.000002	0.000002	0.000002	0.999992	TRJ-2
DELLROSE	0.000002	0.000002	0.000002	0.000002	0.000002	0.999992	TRJ-2
FRANCIS	0.000002	0.000002	0.000002	0.000002	0.000002	0.999992	TRJ-2
JACKSON	0.000002	0.000002	0.000002	0.000002	0.000002	0.999992	TRJ-2
JAZZMAN_2	0.000002	0.000002	0.000002	0.000002	0.000002	0.999992	TRJ-2
JODON	0.000002	0.000002	0.000002	0.000002	0.000002	0.999992	TRJ-2
L201	0.000002	0.000002	0.000002	0.000002	0.000002	0.999992	TRJ-2 *
LABELLE	0.000002	0.000002	0.000002	0.000002	0.000002	0.999992	TRJ-2
LEBONNET	0.000002	0.000002	0.000002	0.000002	0.000002	0.999992	TRJ-2

MERMENTAU	0.000002	0.000002	0.000002	0.000002	0.000002	0.999992	TRJ-2
MILLIE	0.000002	0.000002	0.000002	0.000002	0.000002	0.999992	TRJ-2
NEWBONNET	0.000001	0.000002	0.000002	0.000002	0.000002	0.999992	TRJ-2
REXMONT	0.000002	0.000002	0.000002	0.000002	0.000002	0.999992	TRJ-2
ROY_J	0.000002	0.000002	0.000002	0.000002	0.000002	0.999992	TRJ-2
RT7015	0.000002	0.000002	0.000002	0.000002	0.000002	0.999992	TRJ-2
TEXMONT	0.000002	0.000002	0.000002	0.000002	0.000002	0.999992	TRJ-2
CHENIERE	0.000002	0.000002	0.000002	0.000002	0.000002	0.999991	TRJ-2
CL151	0.000002	0.000002	0.000002	0.000002	0.000002	0.999991	TRJ-2
CL171_AR	0.000002	0.000002	0.000002	0.000002	0.000002	0.999991	TRJ-2
GULFMONT	0.000002	0.000002	0.000002	0.000002	0.000002	0.999991	TRJ-2
L202	0.000002	0.000002	0.000002	0.000002	0.000002	0.999991	TRJ-2
LACASSINE	0.000002	0.000002	0.000002	0.000002	0.000002	0.999991	TRJ-2
MADISON	0.000002	0.000002	0.000002	0.000002	0.000002	0.999991	TRJ-2
MAYBELLE	0.000002	0.000002	0.000002	0.000002	0.000002	0.999991	TRJ-2
PRISCILLA	0.000002	0.000002	0.000002	0.000002	0.000002	0.999991	TRJ-2
REX	0.000002	0.000002	0.000002	0.000002	0.000002	0.999991	TRJ-2
STG06L	0.000002	0.000002	0.000002	0.000002	0.000002	0.999991	TRJ-2
TAGGART	0.000002	0.000002	0.000002	0.000002	0.000002	0.999991	TRJ-2
CL152	0.000002	0.000002	0.000002	0.000002	0.000002	0.99999	TRJ-2
COCODRIE	0.000002	0.000002	0.000002	0.000002	0.000002	0.99999	TRJ-2
CATAHOULA	0.000002	0.000002	0.000002	0.000002	0.000002	0.999988	TRJ-2
BONNET_73	0.000002	0.000002	0.000002	0.000372	0.000002	0.999621	TRJ-2
TORO	0.000002	0.000002	0.000002	0.00217	0.000002	0.997822	TRJ-2
MARS	0.000002	0.000002	0.000002	0.246432	0.000002	0.753561	admix
NOVA_66	0.000002	0.000002	0.000002	0.246958	0.000002	0.753036	admix *
BENGAL	0.000002	0.000002	0.000002	0.365784	0.000002	0.634209	admix
CL261	0.000002	0.000002	0.000002	0.376287	0.000002	0.623706	admix
MEDARK	0.000002	0.000002	0.000002	0.446265	0.000002	0.553728	admix
ARBORIO	0.000002	0.000002	0.000002	0.513941	0.000002	0.486053	admix
SABER	0.000001	0.12005	0.000001	0.000001	0.000001	0.879944	TRJ-2
Cypress2	0.000001	0.000001	0.000001	0.000001	0.000001	0.999994	TRJ-2
ADAIR	0.000001	0.000001	0.000001	0.000001	0.000001	0.999993	TRJ-2 *
BANKS	0.000001	0.000001	0.000001	0.000001	0.000001	0.999993	TRJ-2
DREW	0.000001	0.000001	0.000001	0.000001	0.000001	0.999993	TRJ-2
JEFFERSON	0.000001	0.000001	0.000001	0.000001	0.000001	0.999993	TRJ-2
KATY	0.000001	0.000001	0.000001	0.000001	0.000001	0.999993	TRJ-2
KAYBONNET	0.000001	0.000001	0.000001	0.000001	0.000001	0.999993	TRJ-2
SIERRA	0.000001	0.000001	0.000001	0.000001	0.000001	0.999993	TRJ-2

SKYBONNET	0.000001	0.000001	0.000001	0.000001	0.000001	0.999993	TRJ-2
TEBONNET	0.000001	0.000001	0.000001	0.000001	0.000001	0.999993	TRJ-2
WELLS	0.000001	0.000001	0.000001	0.000001	0.000001	0.999993	TRJ-2
LEAH	0.021654	0.000001	0.000001	0.000001	0.000001	0.978341	TRJ-2
CALROSE	0.000001	0.000001	0.000001	0.695841	0.000002	0.304153	admix
TAINAN_IKU_487	0.000001	0.000001	0.000001	0.998717	0.000001	0.001278	TEJ *
CALORO	0.000002	0.000002	0.000002	0.999991	0.000002	0.000002	TEJ
CHINESE	0.000002	0.000002	0.000002	0.999992	0.000002	0.000002	TEJ *
EARLY_WATARIBUNE	0.000002	0.000002	0.000002	0.999992	0.000002	0.000002	TEJ *
KOSHIHIKARI	0.000002	0.000002	0.000002	0.999992	0.000002	0.000002	TEJ *
NORTAI	0.000001	0.000001	0.000001	0.999993	0.000001	0.000001	TEJ

Genetic identity of ACREAGE set reflects market class

We next separated out the ACREAGE set (n=103 after removal of three *indica* individuals) and examined their relationships without subpopulation controls. Individuals were highly related, as reflected by the nearly radial distance tree (**Figure 3.5**), but there was still some structure. When annotating with available metadata on this panel (breeding program and market class), we found that this structure was primarily associated with market class (i.e. grain length) rather than breeding program.

Generation of datasets to use for yield modeling

Finally, genotype sub-datasets on the ACREAGE varieties were generated to enable yield modeling with historical market data. Call rate (0.8 and 0.9), minor allele frequency (0.01 and 0.05) and LD-pruning (non-pruned and pruned) were adjusted to produce eight sub-datasets. The number of SNP sites ranged from 426 to 8,755 (**Tables 3.3-3.4**).

Figure 3.5. Sub-subpopulation structure of U.S. rice is aligned with market class not breeding program. Note: 3 *indica* accessions dropped for this analysis (cv. Rondo, Milagro Filipino, and Jasmine85).

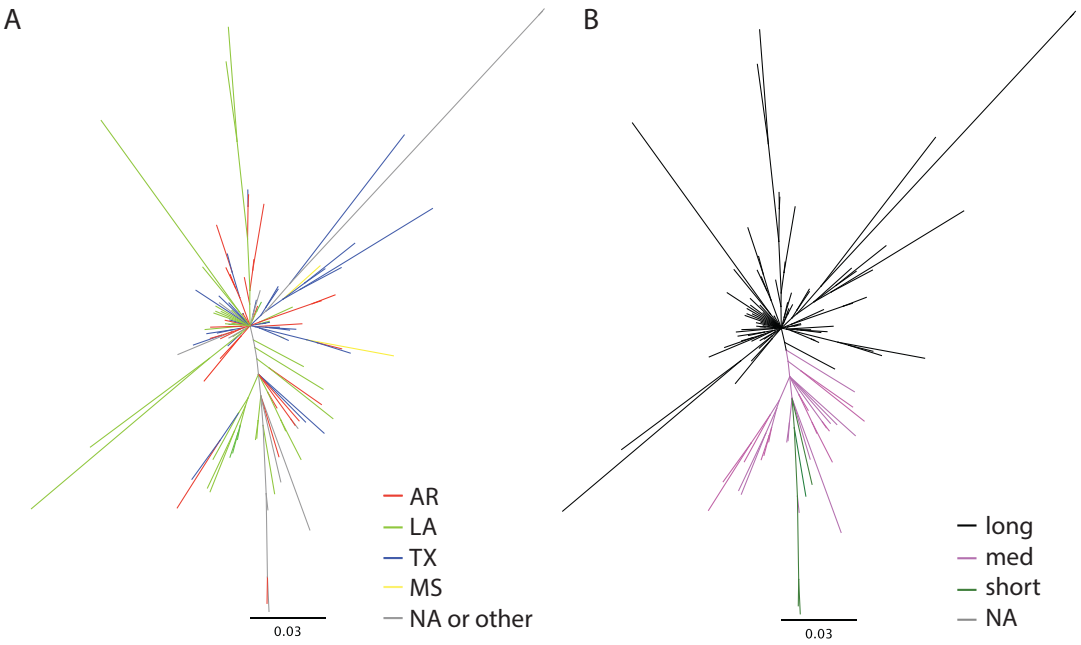


Table 3.2. Non-LD pruned datasets on ACREAGE set. Filtering on call rate and MAF based on the full non-LD pruned dataset of 111,500 sites.

Call rate	Minor allele freq.	Number of sites
0.8	0.01	8,755
0.8	0.05	2,807
0.9	0.01	6,051
0.9	0.05	2,016

Table 3.3 LD pruned datasets on ACREAGE set. Filtering on call rate and MAF based on the full LD pruned dataset of 14719. See Methods for LD pruning parameters.

Call rate	Minor allele freq.	Number of sites
0.8	0.01	1,340
0.8	0.05	426
0.9	0.01	837
0.9	0.05	277

Genotype information on this panel, both FULL and ACREAGE sets, is a valuable resource that enables at once a retrospective analysis of what markers or haplotypes have persisted through the relatively short history of Southern U.S. rice breeding and a forward look at what genomic signatures are potentially useful for present and future climate conditions. Work to understand the relationships among individuals at a local genomic level is ongoing (e.g. admixture analysis, local imputation).

APPENDIX

Supplemental Table 3.1. Germplasm information on 153 U.S. rice accessions.

Sample Name	Purpose
A-301	older cross parent
ADAIR	older cross parent
AHRENT	on acreage list
ALAN	on acreage list
Antonio	on acreage list
Arborio	on acreage list
Arkrose	on acreage list
BANKS	on acreage list
BELLE PATNA	on acreage list
Lemont	on acreage list
BELLEMONT	on acreage list
BELLEVUE	on acreage list
Bengal	on acreage list
Blue Rose	on acreage list
Blue Rose Supreme	US pedigree progenitor, on acreage list
BLUEBELLE	on acreage list
BLUEBONNET	on acreage list
BOND	on acreage list
BONNET 73	on acreage list
BOWMAN	on acreage list
BRAZOS	on acreage list
Caffey	on acreage list
Caloro	US pedigree progenitor, on acreage list
CALROSE	on acreage list
CAROLINA GOLD SELECT	US pedigree progenitor
Carolina Gold	US pedigree progenitor
Carolina Gold	US pedigree progenitor
Carolina Gold Select	US pedigree progenitor
CATAHOULA	on acreage list
CB801	on acreage list
Century Patna 231	on acreage list
CHENIERE	on acreage list
Chinese	US pedigree progenitor
CL111	on acreage list

CL121	on acreage list
CL131	on acreage list
CL142 AR	on acreage list
CL151	on acreage list
CL152	on acreage list
CL161	on acreage list
CL162	new variety
CL171 AR	on acreage list
CL261	on acreage list
Cocodrie	on acreage list
Colorado	on acreage list
Colusa	on acreage list
Cybonnet	parent in new crosses, on acreage list
Cypress	on acreage list
DAWN	on acreage list
Della	older cross parent, on acreage list
DELLA 2	new variety
DELLROSE	on acreage list
DIXIEBELLE	on acreage list
DREW	on acreage list
EARL	on acreage list
Early Wataribune	US pedigree progenitor
Edith	US pedigree progenitor
collected Shufeng 121-1655	new variety/germplasm
Fortuna	US pedigree progenitor
Francis	on acreage list
Guichao-2	older cross parent
GULFMONT	on acreage list
Gulfrose	on acreage list
HILL LONG GRAIN	older cross parent
Hidalgo	on acreage list
Honduras	US pedigree progenitor
IMPROVED BLUE ROSE	US pedigree progenitor
Inan Iku 487	US pedigree progenitor
IR36	older cross parent
IRGA409	older cross parent
JACKSON	on acreage list
Jasmin 85	on acreage list
JAZZMAN	on acreage list
JAZZMAN 2	on acreage list

Jefferson	on acreage list
JES	new variety
JODON	on acreage list
JUPITER	on acreage list
Katy	on acreage list
Kaybonnet	on acreage list
Koshihikari	older cross parent
L-201	older cross parent
L202	on acreage list
LA110	older cross parent
LABELLE	on acreage list
LACASSINE	on acreage list
Lacrosse	older cross parent
LAFITTE	on acreage list
LaGrue	on acreage list
Leah	on acreage list
LEBONNET	on acreage list
MADISON	on acreage list
MARS	on acreage list
MAYBELLE	on acreage list
MEDARK	on acreage list
MERCURY	on acreage list
MERMENTAU	on acreage list
MILAGRO FILIPINO	on acreage list
MILLIE	on acreage list
Nira	US pedigree progenitor
NATO	on acreage list
Neches	older cross parent
MELROSE	on acreage list
NEPTUNE	on acreage list
NEWBONNET	on acreage list
NEWREX	on acreage list
Nortai	on acreage list
NORTHROSE	older cross parent
Nova	on acreage list
NOVA 66	older cross parent
NOVA 76	older cross parent
Sinampaga Selection	US pedigree progenitor
ORION	on acreage list
Panda	US pedigree progenitor

Pecos	older cross parent
PIROGUE	on acreage list
Presidio	on acreage list
PRISCILLA	on acreage list
REX	on acreage list
REXARK	older cross parent
REXMONT	on acreage list
RICO1	on acreage list
Rondo	on acreage list
Rosemont	on acreage list
Roy J	new variety
RT7015	on acreage list
Saber	on acreage list
SABINE	on acreage list
Saturn	on acreage list
Sierra	on acreage list
SKYBONNET	on acreage list
SLO17	US pedigree progenitor
STARBONNET	on acreage list
STG06L-35-061	new variety/germplasm
Taducan	older cross parent
Taggart	new variety
Taichung Native	older cross parent
TEBONNET	on acreage list
TEMPLETON	new variety
Teqing	older cross parent
TESANAI 2	new variety
TEXMONT	on acreage list
TORO	on acreage list
TORO-2	on acreage list
TP49	on acreage list
TRENASSE	on acreage list
V7817	on acreage list
VEGOLD	on acreage list
VISTA	on acreage list
Wells	on acreage list
Lady Wright Selection	US pedigree progenitor
ZENITH	older cross parent
Zhe733	parent in new crosses

REFERENCES

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS One **6**: e19379

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics **28**: 1647–1649

Lu H, Redus MA, Coburn JR, Rutger JN, McCOUCH SR, Tai TH (2005) Population structure and breeding patterns of 145 US rice cultivars based on SSR marker analysis. Crop Sci **45**: 66–76

McCouch SR, Wright MH, Tung C-W, Maron LG, McNally KL, Fitzgerald M, Singh N, DeClerck G, Agosto-Perez F, Korniliev P, et al (2016) Open access resources for genome-wide association mapping in rice. Nat Commun 7: 10532 | DOI: 10.1038/ncomms10532

Raj A, Stephens M, Pritchard JK (2014) fastSTRUCTURE: variational inference of population structure in large SNP data sets. Genetics **197**: 573–589

Spindel J, Wright M, Chen C, Cobb J, Gage J, Harrington S, Lorieux M, Ahmadi N, McCouch S (2013) Bridging the genotyping gap: using genotyping by sequencing (GBS) to add high-density SNP markers and new value to traditional bi-parental mapping and breeding populations. Theor Appl Genet **126**: 2699–2716

CHAPTER 4:

GENOME-WIDE ASSOCIATION ANALYSES REVEAL THE GENETIC ARCHITECTURE OF STEM NON-STRUCTURAL CARBOHYDRATES IN *TROPICAL JAPONICA* RICE

INTRODUCTION

Genome-wide association studies (GWAS) in plants offer a complementary approach to traditional mapping methods that use bi-parental populations for quantitative trait loci (QTL) discovery. They leverage historical recombination events, leading to higher mapping resolution. This in turn has positive implications for both applied and basic avenues, e.g. marker-assisted selection and candidate gene identification, respectively. The wide genomic net cast by a single round of GWA analysis can yield many candidate regions significantly associated with the trait(s) of interest, however the value of these candidate loci cannot be realized without follow-up evaluation. Here we are interested in using GWA to elucidate the genetic architecture underlying stem non-structural carbohydrates (NSCs) in cultivated Asian rice (*Oryza sativa*), a globally important inbreeding crop species. We undertake the strategy of repeated evaluations on genetically similar (within a single subpopulation) panels in different greenhouse seasons, with the rationale that resulting significant loci that co-localize across these assessments represent the most promising regions for future investigation.

Non-structural carbohydrates refer to the carbohydrate components of plants that do not make up the cell wall. NSCs can be readily broken down, built up, or mobilized; they have the capacity to act as reserves to buffer against suboptimal

environmental conditions that limit photosynthetic activity (Slewinski, 2012). In rice, the net movement of stem NSCs follows: 1) carbohydrates from photosynthesis un-utilized for primary sinks (e.g., actively growing organs) are stored in the stem as starch and sucrose prior to heading, 2) they are remobilized to panicles during grain-filling, and 3) they can be re-accumulated in the stem for a brief period near crop maturity as the panicle loses sink strength. Previous studies investigating the relationship of stem NSC to agronomic traits have linked these carbohydrate reserves to yield potential under optimal conditions (Van Dat and Peterson, 1983) and grain-filling enhancement (Nagata et al., 2002). Additional reports documented the ability of stored stem NSCs to buffer performance under heat and water stress (Yang et al., 2001; Yang et al., 2002; Kim et al., 2011; Morita and Nakano, 2011). These observations highlight the potential importance of stem NSC for rice yield stability, especially under variable climatic conditions during grain-filling.

While there is extensive literature on protein- and transcript-level variation that reflect temporal patterns of rice stem NSC (Perez et al., 1971; Watanabe et al., 1997; Hirose et al., 1999; Ishimaru et al., 2004; He et al., 2005; Hirano et al., 2005; Hirose et al., 2006; Chen and Wang, 2008; Fu et al., 2011; Cook et al., 2012), there exists comparatively less work that describes phenotypic variation observed across accessions (Van Dat and Peterson, 1983; Conocono et al., 1995; Samonte et al., 2001; Hirano et al., 2005; Arai-Sanoh et al., 2011; Pan et al., 2011) and even fewer that try to disentangle the genetic differences that underlie this variation (Nagata et al., 2002; Kashiwagi et al., 2006; Kanbe et al., 2009). Of the genetic studies that have been conducted, they all evaluated bi-parental mapping populations derived from *Japonica*

(*temperate japonica* subpopulation specifically) x *Indica* (*indica* and *aus* subpopulations) crosses, with additional backcrossing to the *temperate japonica* parent. Although these studies did identify significant QTLs associated with stem NSC, candidate regions span many megabases and their applicability is likely limited to the gene pool of *O. sativa* adapted to temperate rice-growing regions such as Japan, northern China, Italy, Chile, and California (United States). Since the majority of rice production occurs worldwide in the tropics and sub-tropics, additional investigation is needed to address whether tropical rice shares these QTL.

In the current study we focus on *tropical japonica*, a subpopulation of the *Japonica* Varietal Group that is critical to rice production in Indonesia, Southern United States, West Africa and Latin America. We first reveal a coarse understanding of stem NSC genetic architecture by performing GWAS on a panel of Southern U.S. rice (n=133). These varieties represent a relatively narrow gene pool within *tropical japonica* with high linkage disequilibrium (LD) due to shared ancestry from common progenitors. Rice varieties grown in the U.S. are derived from a relatively small number of introductions brought into the country over the course of its <150 year rice production history (Dilday, 1990). To evaluate the consistency of resulting stem NSC QTL in other tropical/sub-tropical germplasm, we undertake two additional evaluations. We perform GWAS on a collection of genetically diverse accessions representative of the greater *tropical japonica* gene pool (n=175) (**Supp. Table 4.1**) using near-infrared spectroscopy (NIRS)-predicted phenotypes. This simultaneously provides an opportunity to validate U.S. rice results through a higher-resolution lens due to reduced LD and tests the performance of NIR prediction models for rice stem

NSC with a previously developed calibration (refer to Chapter 2 of this dissertation). Finally we conduct QTL mapping using an interspecific bi-parental Chromosome Segment Substitution Line (CSSL) library (n=48) in a *tropical japonica* genomic background (cv. Curinga) (Arbelaez et al., 2015) as a second validation method. This population is comprised of Near-Isogenic Lines, so results can lead easily into fine-mapping studies. By implementing a repetitive approach of evaluating related but not genetically identical material, we identify QTL on chromosomes one, five, and eleven that consistently associate with stem NSC traits in *tropical japonica* rice and highlight promising candidate genes underlying these regions.

RESULTS AND DISCUSSION

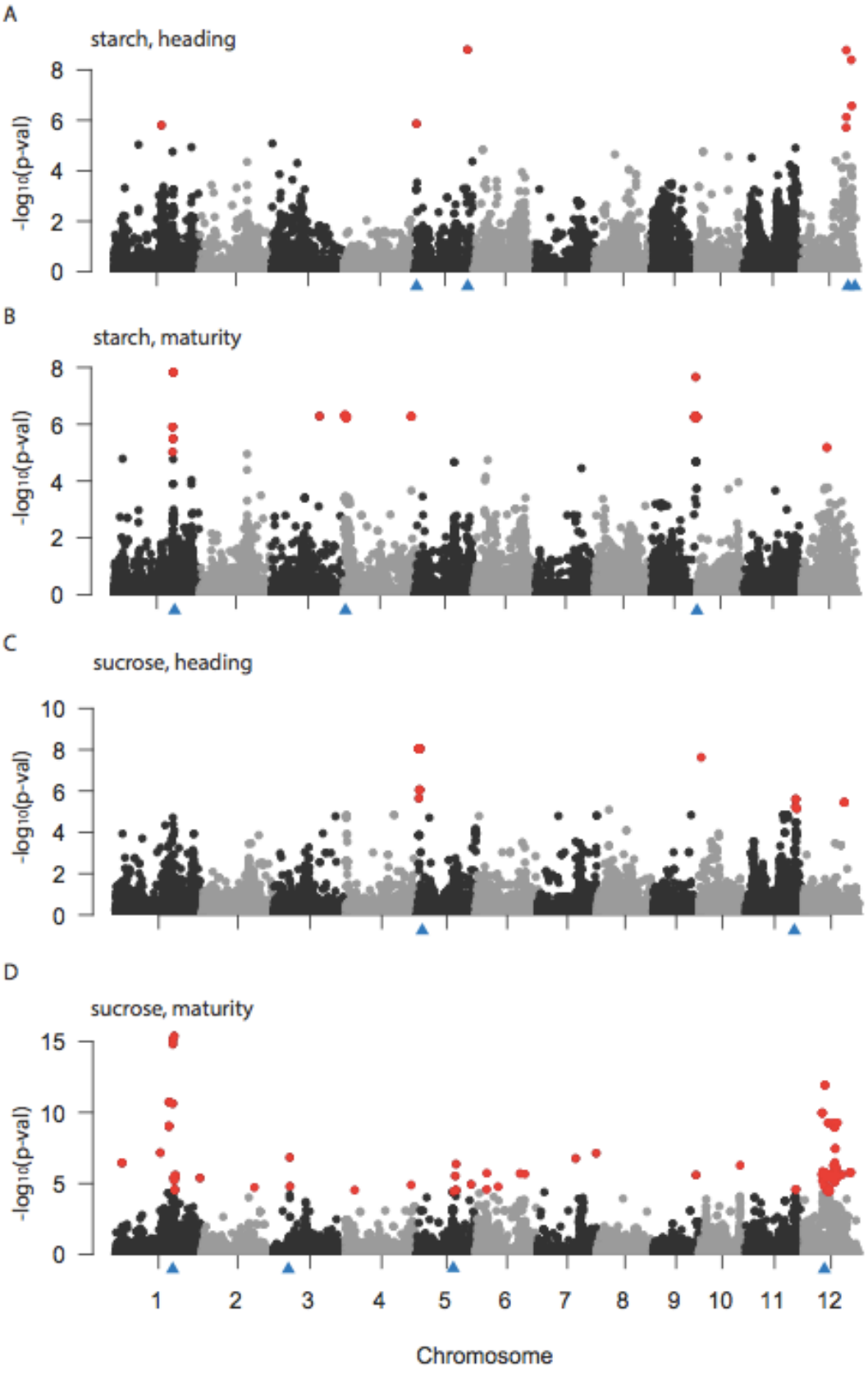
Stem NSC association analysis in U.S. rice

To gain a rough understanding of rice stem NSC genetic architecture, a collection of 126 *tropical japonica* accessions (hereafter referred to as US-TRJ) and seven *temperate japonica* varieties that represent the historical genetic diversity of U.S. rice production (**Supp. Table 4.1**) were phenotyped for four NSC component traits (starch and sucrose at heading and maturity). In U.S. rice breeding programs, crossing occurs primarily within market classes that are defined by grain type, i.e., long, medium or short grain. This practice is reflected in a partitioning of genetic variation by market class (Lu et al., 2005). We first checked to see if stem NSC phenotype distributions within US-TRJ differed across long and medium grain types (note: short grain varieties in this study belong to the *temperate japonica* subpopulation, and are grown

primarily in California), which would lead to a certainty of false positives due to genetic population structure. There was no observed alignment of NSC phenotypes with long and medium grain market classes (**Supp. Figure 4.1**). We proceeded with association mapping for the four NSC traits but still included population stratification controls (see Methods) and a stringent False Discovery Rate threshold (0.01) as precautions against false positives. The analyses were performed using 38,618 genome-wide Single Nucleotide Polymorphism (SNP) markers in two panels: *US-TRJ* (126 *tropical japonica* varieties) and *US-JAPONICA* (all 133 U.S. accessions, including 126 *tropical* and seven *temperate japonica* varieties, that were phenotyped for this study). Results in the two panels agreed (**Fig. 4.1** and **Supp. Figure 4.2**); this was expected due to nearly identical germplasm composition.

From *US-TRJ* results, we cataloged 13 regions of association that were each supported by multiple significant SNPs (**Fig. 4.1, Supp. Table 4.2**). There were four loci associated with starch at heading on chromosomes 5 and 12, three associated with starch at maturity on chromosomes 1, 4, and 9, two associated with sucrose at heading on chromosomes 5 and 11, and four associated with sucrose at maturity on chromosomes 1, 3, 5, and 12. To test for the effect of phenology on stem NSC genetic architecture, we repeated *US-TRJ* GWA analyses with the addition of a days-to-heading (DTH) covariate. For starch at maturity, sucrose at heading, and sucrose at maturity, resultant GWAS p-values corresponded well with previous analyses that did not account for DTH (r= 0.90-0.99). For starch at heading, most of the genetic signal

Figure 4.1. Genetic architecture of stem NSC in U.S. *tropical japonica* rice. Genome-wide association results for stem NSC traits: A) starch at heading, B) starch at maturity, C) sucrose at heading, and D) sucrose at maturity. Results shown here are for the *US-TRJ* panel, comprised of 126 *tropical japonica* accessions. Red circles indicate SNPs significant at 0.01 FDR and blue triangles mark peaks supported by multiple significant SNPs; note that some SNPs may appear visually to be a single marker but are actually multiple overlapping markers.


A
starch, heading
B
starch, maturity
C
sucrose, heading
D
sucrose, maturity
-log10(p-val)
Chromosome
1
2
3
4
5
6
7
8
9
10
11
12

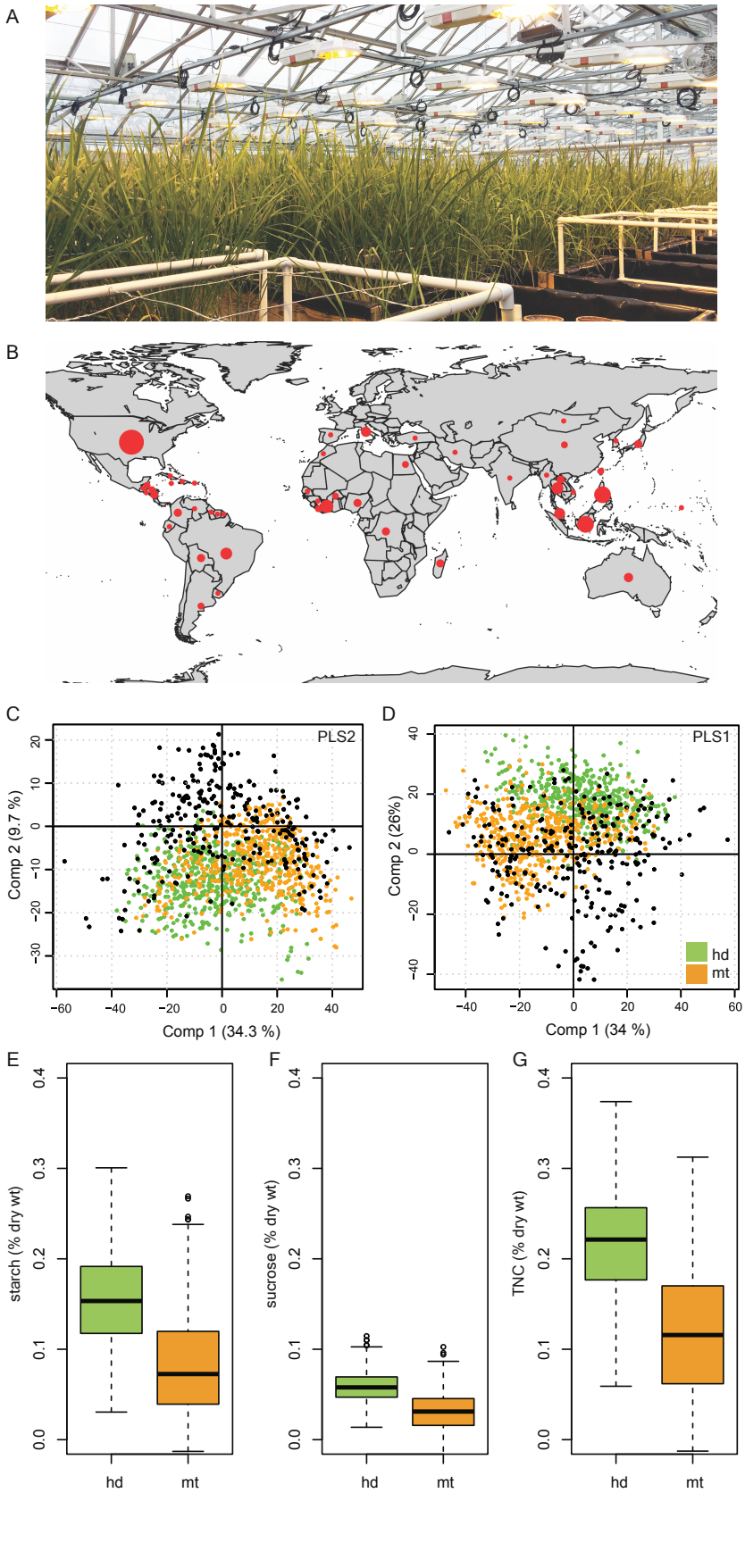
was muted by the addition of the DTH covariate; however, one QTL on chromosome 12 remained significant at 0.05 FDR (**Supp. Figure 4.4**).

Of all *US-TRJ* QTL for stem NSC, only a single region appeared to co-localize across multiple traits; around the midpoint of chromosome 1 was a significant locus associated with both sucrose and starch at maturity. The most significant SNP (msSNP) for sucrose at maturity (S1_29495819) was found at position 29,495,819 bp and the msSNP for starch at maturity (S1_29976759) was 500 kb downstream at position 29,976,759 bp. Despite the large physical distance between these two msSNPs, it is possible that they may be tracking the same causative factor due to high LD in U.S. rice (**Supp. Figure 4.7**). The remaining 11 QTL appeared to be trait specific. While performing GWAS using high LD populations such as *US-TRJ* requires lower marker density for QTL discovery, its trade-off is low mapping resolution, so we could not acceptably make further inferences without additional evaluations.

NSC prediction using NIR information in diverse *tropical japonica*

To gain a finer understanding of the genetic architecture underlying stem NSC in *tropical japonica* rice and to evaluate whether the 13 significant genomic QTL discovered in U.S. rice were common beyond these regional breeding programs, we phenotyped a geographically and genetically diverse *tropical japonica* panel of 175 accessions (**Fig. 4.2A-B, Supp. Table 4.1**). As a collection of diverse landraces, this panel has much lower LD than *US-TRJ* (**Supp. Figure 4.7**), thereby supporting greater mapping resolution. In lieu of wet chemistry assays, we scanned the set of *tropical*

Figure 4.2. Evaluation of diverse *tropical japonica* rice accessions for stem non-structural carbohydrates using predictions by NIR spectral information. Greenhouse grow-out (A), geographical origin of 175 rice accessions (B). Distribution of 976 prediction samples (green and orange circles) and 300 calibration samples (black circles) in PC1 and PC2 of PLS model space using spectral data for PLS2 (C) and PLS1 (D) models. Distribution of predicted values, expressed in percentage dry weight, at heading (hd) and maturity (mt) for starch (E), sucrose (F), and TNC (G).


A
B
C
PLS2
Comp 2 (9.7 %)
Comp 1 (34.3 %)
D
PLS1
Comp 2 (26%)
Comp 1 (34 %)
hd
mt
E
starch (% dry wt)
F
sucrose (% dry wt)
G
TNC (% dry wt)
hd
mt

japonica stem samples for NIR reflectance and made predictions of NSC components using the resultant spectral data as a higher-throughput approach to determine stem constituents (see Methods). Starch and sucrose levels (%w/w) were jointly predicted using a single PLS-2 model while total non-structural carbohydrates, TNC (%w/w), was predicted separately using a PLS-1 model. We found that the spectral data measured on our prediction set were well-aligned to that of the calibration set (**Fig. 2C**), and so we were able to predict the three NSC constituents across two developmental points with very low uncertainty. Predicted values at heading and maturity were consistent with expected distributions (i.e. higher means at heading than maturity) (**Fig. 2D**), and were adjusted for stem dry weight to yield absolute estimates of NSC per tiller for subsequent GWA analyses.

Association analysis of *tropical japonica* rice stem NSC

Out of 700,000 genome-wide SNP markers found on the high-density rice array (HDRA) (McCouch et al., 2016) 222,689 markers had minor allele frequencies greater than 0.05 in the *tropical japonica* panel and were used in conjunction with six NIR-predicted traits (starch, sucrose, and TNC at heading and maturity) for GWAS. Heritability estimates for these traits were lower than previously reported (Wang 2016), which may be due to greenhouse anomalies observed during this evaluation (**Supp. Table 4.4**). These genome scans yielded a total of nine loci of interest (**Supp. Figure 4.6, Supp. Table 4.5**): three were associated with starch-at-heading, two with starch-at-maturity, two with TNC-at-heading, one with TNC-at-maturity and one QTL on chromosome 1 that was associated with both starch-at-maturity and sucrose-at-

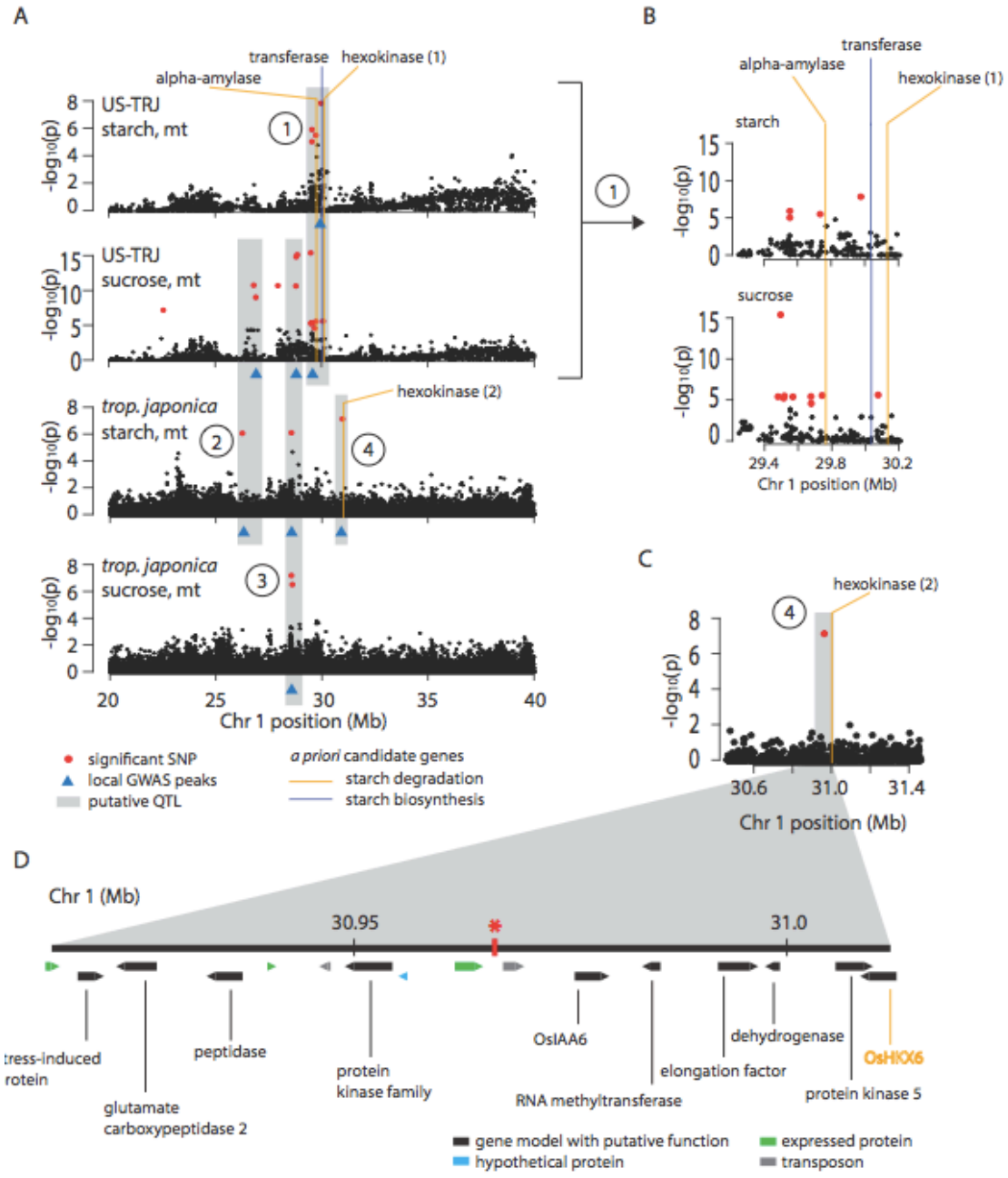
maturity. Local LD decay was calculated for each of these regions (see **Methods**, **Supp. Table 4.5**, **Supp. Figure 4.7**), resulting in values that ranged from 33.85 to 252.78 kb on either side of the msSNP (corresponding to between six and 52 non-TE gene models). Of these nine loci, two appeared to co-localize with peaks resulting from U.S. GWA analyses and warranted further investigation; they were found on chromosomes 1 and 11 (**Table 4.1**).

Table 4.1. GWAS peaks common across U.S. and *tropical japonica* diversity panels

Trait (*trj*)	**Marker (*trj*)**	**Chr**	**Marker position**	**Local LD (kb)**	**Region size (kb)**	**# gene models (non-TE)**	**Candidate gene(s)**
starch, mt	SNP-1.30962744.	1	30963789	48.99	97.98	13	Hexokinase (OsHXK6)
TNC, hd	SNP-11.24633058.	11	25140140	252.87	505.74	53	3 sugar transporters

Upon closer study, the region on chromosome 1 at ~29-30Mb, which was associated with starch-at-maturity and sucrose-at-maturity in both the *tropical japonica* and *US-TRJ* panels, appeared to be a cluster of several distinct sub-QTL. We tagged four potential sub-regions of significance (QTLs 1-4, **Fig. 4.3A**) and inspected them individually. QTL-1 was found only in US-TRJ and spanned a ~500kb physical interval (**Fig. 4.3B**), consistent with the high LD in U.S. rice (**Supp. Figure 4.7**). The other three sub-QTL were found either in the *tropical japonica* or both the US-TRJ

Figure 4.3. Dissection of a region on chromosome 1 associated with stem NSC at maturity. Association analysis uncovered a significant peak on chromosome 1 at position 29-30 Mb in U.S. rice and in diverse *tropical japonica* for starch and sucrose at maturity (Fig. 4.1 and Supp. Fig. 4.6). A) Zoom-in examination of the target region (chr1: 20-40 Mb) reveals up to 4 distinct QTL (gray boxes labeled 1-4, circled). QTL-1 (msSNPs S1_29976759 and S1_29495819) is significantly associated with starch and sucrose in *US-TRJ* but not in *tropical japonica*. QTL-2 may be shared between sucrose in *US-TRJ* and starch in *tropical japonica*. QTL-3 is significantly associated with sucrose in *US-TRJ* and both starch and sucrose in *tropical japonica*. QTL-4 is significant only in starch-at-maturity in *tropical japonica*. Vertical lines mark the locations of *a priori* candidate genes and line colors indicate the candidate genes' biochemical pathway (orange= starch degradation, blue = starch biosynthesis genes). B) A close-up view of QTL-1. Three *a priori* genes (alpha-amylase = LOC_Os01g51754; transferase = LOC_Os01g52250; hexokinase(1) = LOC_Os01g52450) are found in the vicinity of QTL-1. C) A 1 Mb region surrounding the msSNP (SNP-1.30962744) in QTL-4. The gray box delimits the local region that is in LD with the msSNP (red circle). D) A close-up of the LD region. 13 non-TE gene models were found here (gray boxes depict TEs while black and orange boxes depict non-TE gene models). This region encompasses one *a priori* candidate gene model that encodes a second hexokinase (LOC_Os01g52450, *OsHXK6*). The red asterisk marks the location of the msSNP (SNP-1.30962744).


A
transferase
hexokinase (1)
alpha-amylase
US-TRJ
starch, mt
US-TRJ
sucrose, mt
hexokinase (2)
trop. japonica
starch, mt
trop. japonica
sucrose, mt
-log10(p)
Chr 1 position (Mb)
significant SNP
local GWAS peaks
putative QTL
a priori candidate genes
starch degradation
starch biosynthesis
B
transferase
alpha-amylase
hexokinase (1)
starch
sucrose
C
hexokinase (2)
D
Chr 1 (Mb)
30.95
31.0
tress-induced
rotein
glutamate
carboxypeptidase 2
peptidase
protein
kinase family
OsIAA6
RNA methyltransferase
elongation factor
dehydrogenase
protein kinase 5
OsHKX6
gene model with putative function
hypothetical protein
expressed protein
transposon

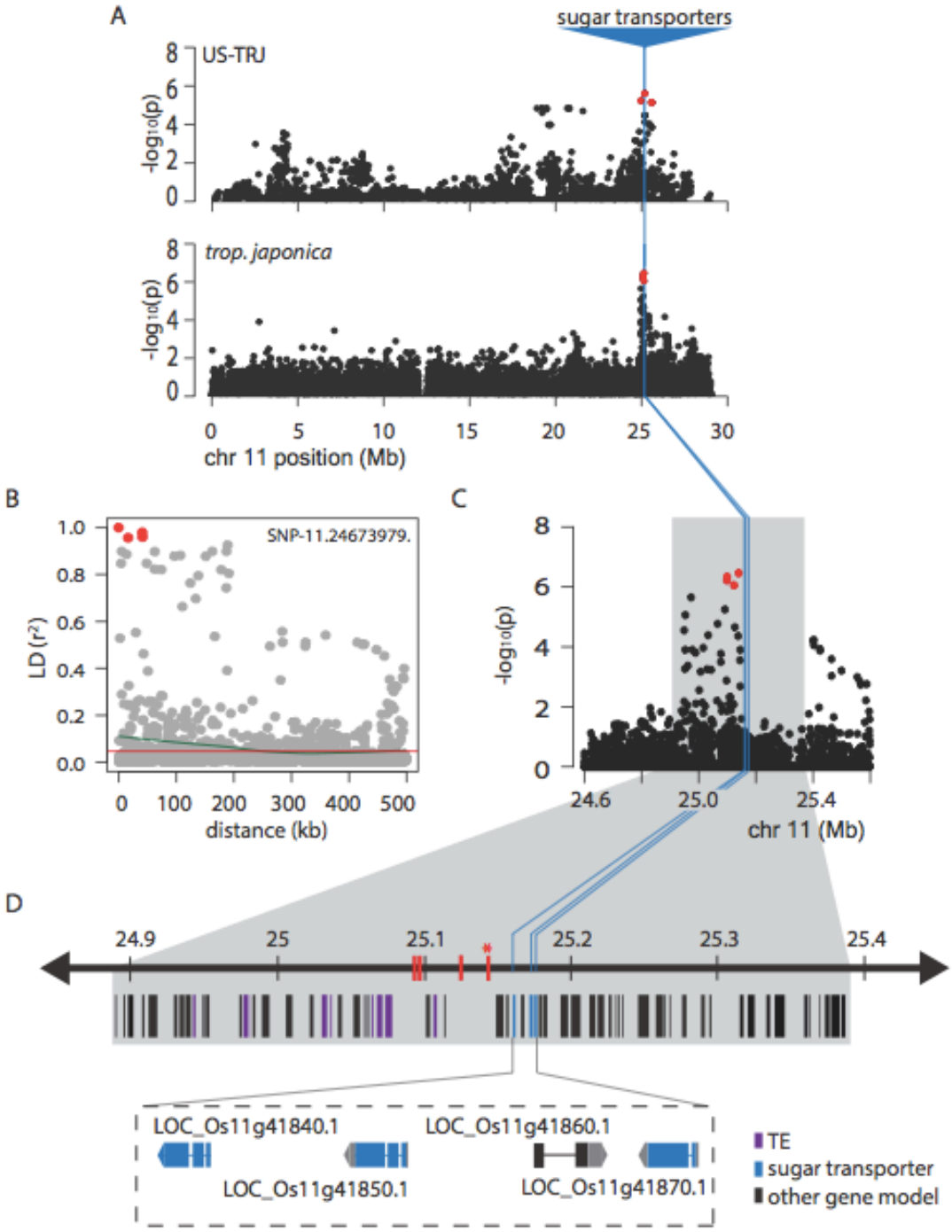
and *tropical japonica* panels, which allowed us to investigate at a finer scale due to reduced LD in the *tropical japonica* panel. These three loci did not co-localize in the same LD blocks (**Supp. Figure 4.9**), supporting the idea that this mega-locus on chromosome 1 is made up of several distinct QTL. For all three msSNPs (SNP-1.26243779, SNP-1.28557570, and SNP-1.30962744) the minor allele in the population was associated with increased stem NSC (**Supp. Figure 4.10**).

Using the HDRA SNPs that were polymorphic at a minor allele frequency above 5% in the *tropical japonica* panel we estimated local LD decay for the msSNPs of QTLs 3 and 4 (see **Methods** and **Supp. Table 4.5**); the regions encompassed 24 and 13 non-TE gene models, respectively. No genes with obvious carbohydrate functions were found within the region of LD for QTL 3 but QTL 4 encompassed one *a priori* candidate gene, a hexokinase gene (*OsHXK6*) (**Fig. 4.3C**). Plant hexokinases primarily function to catalyze the phosphorylation of alpha-D-glucose that results from starch breakdown, but perhaps more interestingly, they also have been shown to mediate sugar signaling by binding glucose along with protein factors to repress gene transcription (Jang et al., 1997; Granot, 2008; Cho et al., 2009). In rice there are ten hexokinase genes and *OsHXK6* is one of the two documented to play a sugar-sensing role (Cho et al., 2009). Specifically, OsHXK6 was shown to repress transcription of rice alpha-amylase genes, contributing to a negative feedback under high glucose levels. In addition to *OsHXK6*, the LD region of QTL-4 also included 12 other non-TE gene models (**Fig. 4.3D**, **Supp. Table 4.6**) but none of them had apparent roles related to carbohydrate metabolism, leaving *OsHXK6* as the strongest candidate.

We next investigated the QTL on chromosome 11 at position 25 Mb (**Fig.**

4.4A, C). This region was significant for sucrose in the *US-TRJ* panel and TNC in the *tropical japonica* panel. The msSNP here (SNP-11.24673979) was in a region of local LD that spanned 252.87 kb to each side in *tropical japonica* (**Fig. 4.4B**), and the minor allele was associated with increased stem NSC (**Supp. Figure 4.10B**). The region of LD (chr 11: 24,907,499-25.290,939 bp) harbored 12 TEs and 52 non-TE gene models (**Supp. Table 4.7**). Though none were genes collected in our *a priori* candidate gene search (**Supp. Table 4.6**), there were three tandemly arrayed sugar transporters (LOC_Os11g41840, LOC_Os11g41850, and LOC_Os11g41870) (**Fig. 4.4D**). These genes are part of the polyol/monosaccharide transporter (PLT/PMT) family of monosaccharide sugar transporters (MSTs) in rice (Johnson and Thomas, 2007). MST families are ancient gene families that have expanded and diverged both across- and within-species due to duplication events; this has given rise to a wide range of functional variation including phloem transport and plant defense (Johnson and Thomas, 2007; Slewinski, 2011). In rice there are six PLT genes, four of which are tandemly arrayed on chromosome 11. Three of those four underlie the significant region associated with NSC at heading in *US-TRJ* and *tropical japonica* rice in this study. The molecular functions of these three transporters have yet to be characterized but from results of the current GWAS, it would be of interest to elucidate whether their roles (structure and regulation) have diverged in the context of rice stem NSC accumulation at heading in a future investigation. Of the 54 gene models in this region, these three PLT sugar transporters are the most promising candidates likely to impact levels of rice stem NSC.

Figure 4.4. Three sugar transporters underlie a significant peak for NSC at heading. A) Association analysis revealed a significant peak on chromosome 11 at position 25 Mb in both U.S. *tropical japonica* rice (*US-TRJ*) for sucrose at heading and in diverse *tropical japonica* for total non-structural carbohydrates at heading. Significant SNPs are colored red. B) LD decays ~192kb away from the most significant SNP in the *tropical japonica* panel. The red line in (B) marks the critical r^2 threshold below. C) A close-up of the candidate region. The gray box marks the region in LD with the msSNP (SNP-11.24673979), corresponding to chr 11: 24,907,499-25,290,939 bp, and blue lines indicate the locations of three tandemly arrayed sugar transporters (LOC_Os11g41840.1, LOC_Os11g41850.1, and LOC_Os11g41870.1). D) In total 38 non-TE gene models were found in this LD region. The red lines denote the four significant SNPs detected at this locus and the asterisk marks the msSNP in relation to the genes in the region.


A
sugar transporters
US-TRJ
trop. japonica
-log10(p)
chr 11 position (Mb)
B
SNP-11.24673979.
LD (r²)
distance (kb)
C
-log10(p)
chr 11 (Mb)
D
LOC_Os11g41840.1
LOC_Os11g41850.1
LOC_Os11g41860.1
LOC_Os11g41870.1
TE
sugar transporter
other gene model

We were next interested to examine the allele frequencies of SNP-1.30962774 and SNP-11.24673979 across a wider spectrum of *O. sativa* diversity, beyond *tropical japonica*. Analysis of allele frequencies across 1,493 diverse *O. sativa* accessions from the Rice Diversity Panels 1 and 2 (McCouch et al., 2016) revealed that the alleles of both msSNPs that were associated with increased NSC occurred at low frequencies in *tropical japonica* (7% of n=279 and 10% of n=340 for SNP-1.30962774 and SNP-11.24673979, respectively). The trait-increasing allele for SNP-1.30962774 was found at higher frequencies in the *indica* and *aromatic* subpopulations (29% and 36%), and the trait-increasing allele of SNP-11.24673979 was similiarly found at highest frequency in *indica* and *aus* (19.5% and 26%) (**Supp. Table 4.8**). It remains to be determined whether or not these alleles, which occur at low frequency in *tropical japonica,* are the result of introgression from other subpopulations.

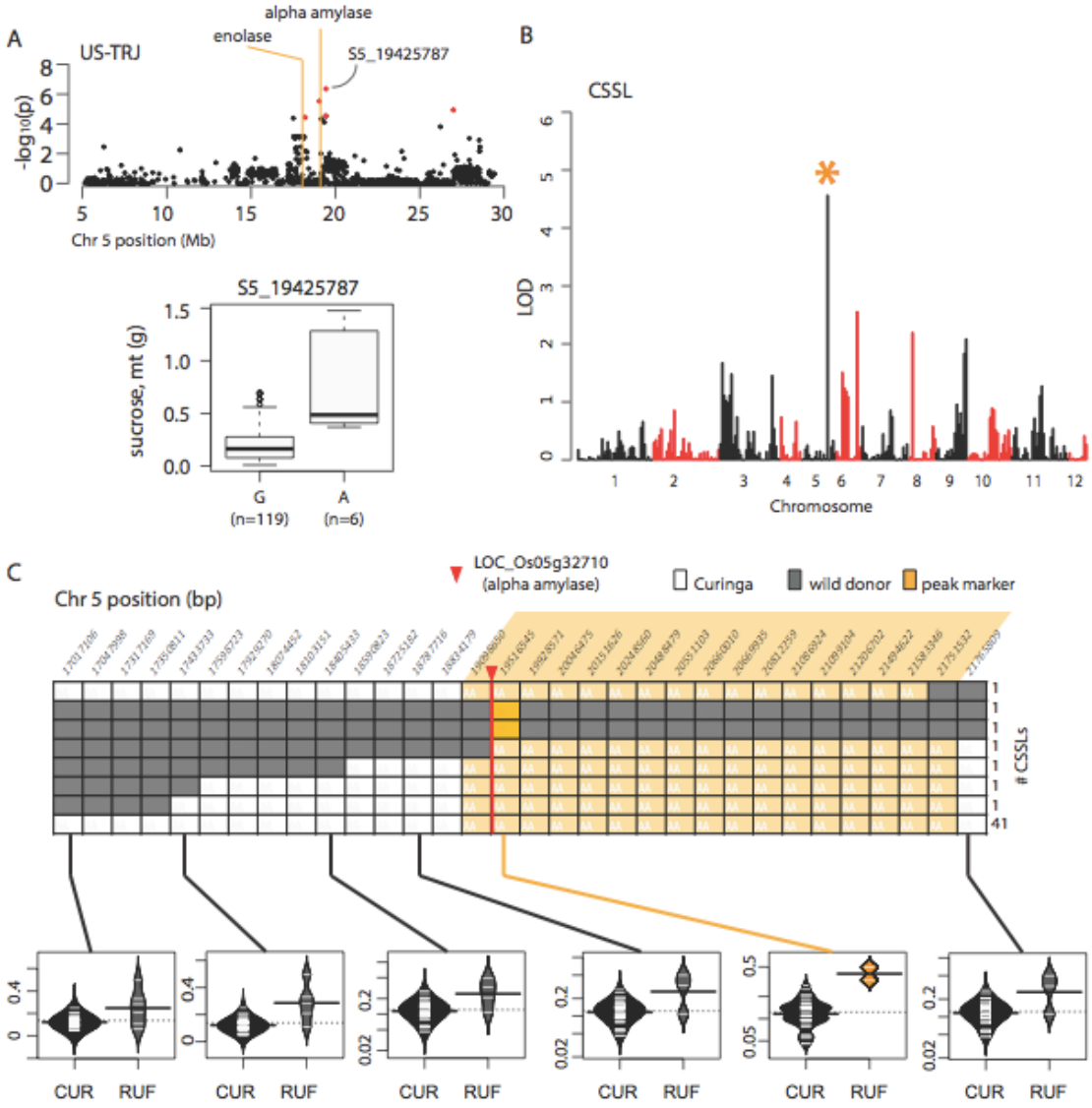
QTL mapping in Curinga x wild CSSL population

We next performed QTL mapping in a CSSL population derived from an interspecific cross between cv. Curinga, an upland *tropical japonica* variety from Brazil, and a wild *aus*-like *O. rufipogon* accession of rice (Arbelaez et al., 2015; Kim et al., 2016). Here we were interested in whether QTL associated with stem NSC traits in the CSSLs co-localized with any regions identified from GWAS in the *US-TRJ* and *tropical japonica* panels. Of the two QTLs identified (**Supp. Table 4.9**), one marker on chromosome 5 at position 19.5 Mb that associated with sucrose-at-maturity co-localized with a GWAS peak in *US-TRJ*, which was also associated with sucrose-at-maturity. Despite low numbers of individuals carrying the donor genotype relative to the recurrent

parent genotype at any particular locus in the CSSL population, we documented a significant effect on stem sucrose-at-maturity for a large introgression at position 19.1-21.75 Mb where the *O. rufipogon* allele significantly increases the phenotype. In other words, CSSL lines that harbored a wild *aus*-like introgression across the target region had values beyond (greater than) the distribution of the rest of the CSSL lines that carried the *tropical japonica* genotype (**Fig. 4.5C-D**). This introgression contained a large number of genes (61 TEs and 339 non-TEs) so further fine mapping is required to narrow down the target region and determine the causative gene(s). To lay the foundation for this work, line CURUF25, one of the two CSSLs that harbored the donor allele at the region of interest, was backcrossed again to Curinga. We confirmed the genotype of a single F_1 plant using a 6K SNP array and generated F_2 seed.

The *US-TRJ* GWAS peak centered at 19.5 Mb on chromosome 5 mapped within the introgressed region of interest in the CSSLs, and contained an *a priori* gene, LOC_Os05g32710 (**Fig. 4.5A-B**). The other 339 gene models in the introgressed region did not have obvious carbohydrate metabolism roles so we looked further into LOC_Os05g32710. This gene was initially predicted to encode an alpha amylase, based on our initial *a priori* candidate screen using RiceCyc and the MSUv7 genome assembly (Dharmawardhana et al., 2013). However, after studying previous work involving this gene, we found that LOC_Os05g32710 is a rice homolog of maize *sugary1* (*su1*) and Arabidopsis *isa2;* it encodes an isoamylase-2 debranching enzyme, ISA2 (Kharabian-Masouleh et al., 2011; Streb and Zeeman, 2014). Isoamylases are

Figure 4.5. Chromosome five is associated with stem sucrose levels at maturity. (A) Association analysis results for chromosome 5 in U.S. *tropical japonica* rice (*US-TRJ*). Several significant SNPs (red circles) localized near two *a priori* candidate genes, enolase (LOC_Os05g31110) and isoamylase-2 (LOC_Os05g32710) (orange lines). The boxplot shows sucrose-at-maturity distributions for the two alleles of the primary peak's msSNP, S5_19425787. (B) Mapping results from the CSSL population (cv. Curinga x *O. rufipogon* wild donor) for sucrose-at-maturity revealed a significant marker (c5_19516545; annotated by orange asterisk) on chromosome five at 19.5 Mb. (C) Graphical genotypes showing wild introgressions in CSSL lines in the region around SNP S5_19425787 using data from a 6K liquid array platform. Sucrose level distributions are shown below for individuals with recurrent parent genotype "CUR", and donor parent genotype "RUFI", at six SNP markers across the region. The peak SNP (S5_19425787) is indicated in bright orange and the greater candidate region is colored in pale orange. The red arrow marks the location of the *a priori* gene, *OsISA2*, in relation to marker data on the CSSLs.


A
US-TRJ
alpha amylase
enolase
S5_19425787
-log10(p)
Chr 5 position (Mb)
S5_19425787
sucrose, mt (g)
G
(n=119)
A
(n=6)
B
CSSL
LOD
Chromosome
C
Chr 5 position (bp)
LOC_Os05g32710
(alpha amylase)
Curinga
wild donor
peak marker
CSSLs
CUR
RUF

debranching enzyme complexes that function in the starch biosynthesis pathway to mediate production of amylopectin, branched starch molecules (Tetlow et al., 2004). In rice, there are two isoamylase protein complexes (i.e., oligomers): an ISA1 homo-oligomer and an ISA1-ISA2 hetero-oligomer. While the ISA1 homo-oligomer is essential for endosperm amylopectin production, the ISA1-ISA2 hetero-oligomer is likely not (Utsumi et al., 2011). However, both *ISA1* and *ISA2* were able to partially complement leaf starch phenotypes of *isa1* and *isa2* single and double mutants in dicot *Arabidopsis* (Streb and Zeeman, 2014), indicating broad functionality of the rice *ISA2* gene. No studies have investigated the role of rice *ISA2* in the plant stem.

We compared sequences across a ~3kb region that included the *ISA2* gene and a 1.5 kb putative promoter region in the CSSL parents, Curinga and *O. rufipogon*. We found one SNP in the coding sequence and 3 single base-pair indels in the putative promoter (**Supp. Fig. 4.14**). Interestingly, we also found a region in the putative promoter just upstream of the *ISA2* start site that was rich in consecutive missing data points in Curinga. To clarify whether this result was specific to Curinga, or was more generally a feature of *tropical japonica* genomes, we examined re-sequencing information from 18 other *tropical japonica* accessions and discovered that the same region was characterized by missing data and variable alignments of the short NGS reads against the reference genome. In contrast, the single exon of the gene itself and the remaining part of the promoter region upstream of the variable region aligned well and were largely devoid of missing data. This pattern of disruptive alignment with a reference genome typically reflects a structurally- or hyper-variable region, and is likely associated with the presence of repetitive elements (e.g. Simple Sequence

Repeats, SSR, or transposable elements, TEs) in the reference genome and/or the query sequences. Viewing this region in the MSUv7 genome browser (http://rice.plantbiology.msu.edu/cgi-bin/gbrowse/rice/), we identified a previously annotated gg(9) SSR here (**Supp. Fig. 4.14**), adding to our confidence that there is a repetitive element in the *ISA2* promoter. These results raise the possibility that there may be transcriptional variation for *ISA2* in rice due to the presence of repeat motif variation in the promoter of the *ISA2* gene.

In rice stems, starch is present as both amylose and amylopectin (Blakeney and Matheson, 1984), whose sets of enzymes for synthesis and breakdown are not identical; however, differentiating between these two forms has not been a focus of rice stem NSC studies, including the present study, due to the extensive wet chemistry assays that are required. Rather, this work has been primarily reserved for studies on rice endosperm, an economically important trait with regards to grain quality. It may be possible that in certain rice varieties, one form of starch is more easily broken down for translocation out of the stem than the other. Further investigation is required to determine the causative factor(s) underlying the observed variation in sucrose-at-maturity in the CSSL lines and *US-TRJ* and if it turns out to be *ISA2*, clarifying the relationship between amylopectin production and stem sucrose levels may be important to understanding the genetics underlying stem NSC levels in rice.

The work presented here is the most comprehensive genetic study on rice stem NSC to date regarding the germplasm that was evaluated, genetic resources that were leveraged, and analytical methods that were employed. Its primary objectives were to gain a clearer picture of the genetic architecture of rice stem NSC at the genome-level

and to generate hypotheses for testing at the gene-level. One curious observation made during our study is that increased stem NSC, regardless of the specific trait (starch, sucrose, TNC) or sampling point (heading, maturity) was always associated with a SNP allele that was rare in *tropical japonica* (**Supp. Fig. 4.10, Fig. 4.5A**). These potentially favorable alleles were found more frequently in other subpopulations (*indica, aus,* and *aromatic*) (**Supp. Tables 4.8 and 4.10**). For the chromosome 5 QTL discovered in *US-TRJ* GWAS, out of the six accessions found to carry these alleles, five (cv. Carolina Gold, Blue Rose, Blue Rose Supreme, and Lady Wright Selection) have unknown pedigree and were selected from the earliest introductions in the United States prior to 1930; they were used as parents in some of the earliest crosses in the history of U.S. rice breeding efforts (**Supp. Fig. 4.12**) (Lu et al., 2005), however, this particular allele did not persist into the gene pool of modern rice varieties. Many intriguing questions remain. Can these rare alleles in the early plant selections be traced back to introgressions from other subpopulations? If so, do these observations hint at a fundamental divergence of stem NSC physiology between rice subpopulations? While genome-wide association analyses neither prove causality nor directly address evolutionary relationships, they give rise to tantalizing hypotheses to drive future work. By examining NSC variation across diverse panels of germplasm at both the population level and the gene level, we have identified several promising QTL for stem NSC in the *tropical japonica* gene pool. These QTL and the candidate genes associated with them open the door to more targeted gene discovery efforts and provide plant breeders with immediate targets for marker-assisted crop improvement. Further work is needed to understand the molecular mechanisms and dynamic

regulatory signals governing the synthesis, storage, breakdown and remobilization of NSC within a plant, and to begin to harness the potential of stem NSC for enhancing rice yield stability in the face of changing and highly variable climatic conditions.

MATERIALS AND METHODS

Plant materials: Three sets of germplasm were evaluated in this study: 1) *US-TRJ*, a U.S. rice panel made up of 133 varieties representative of the *tropical* and *temperate japonica* gene pools of U.S. breeding programs (refer to Chapter 3 of this dissertation), 2) a diversity panel composed of 136 *tropical japonica* and 39 admixed *Japonica* accessions from the Rice Diversity Panels 1 and 2 (McCouch et al., 2016), and 3) an interspecific Chromosome Segment Substitution Line (CSSL) library, which comprised 48 near-isogenic lines of cv. Curinga (subpopulation *tropical japonica*) that each harbored different genomic introgressions from *O. rufipogon* accession IRGC105491 (Arbelaez et al., 2015). Additional information on these panels may be found in **Supp. Table 4.1**.

Genotyping information: Genotype data on these panels were generated in previous or ongoing studies. The *tropical japonica* diversity panel had been genotyped for 700,000 Single Nucleotide Polymorphisms (SNP) using the High Density Rice Array (HDRA) (McCouch et al., 2016) and the Curinga/*O. rufipogon* CSSL population had been previously characterized using a 6K SNP array (Arbelaez et al., 2015). For the U.S. panel, raw genotyping-by-sequencing (GBS) data were generated as part of a separate, larger study and these reads were re-analyzed here to provide SNP data for GWAS. The GBS pipeline followed the procedure outlined previously (Spindel et al.,

2013). Due to the high proportion of missing information inherent in GBS data, we filtered the GBS dataset for 80% call rate and 0.05 minor allele frequency to generate a SNP dataset of 38,618 sites for performing GWAS on the U.S. varieties. Data at these 38,618 markers on the U.S. lines may be found in NCBI's dbSNP after publication.

Evaluation conditions, sample processing, and phenotypes: Growth management, destructive sampling, and sample processing followed conditions outlined previously (Chapter 2 of this dissertation). Briefly, plant stems were collected at two time-points: heading and physiological maturity. The U.S. panel was un-replicated while the *tropical japonica* panel and CSSL evaluation measures were averaged across three replicates per accession and two replicates per accession, respectively. All NSC phenotypes were expressed in absolute terms (g) on a per tiller basis. In total, the U.S. panel evaluated four phenotypes (stem starch and stem sucrose at heading and maturity). In the *tropical japonica* panel, six phenotypes were evaluated (starch, sucrose, and total non-structural carbohydrates at heading and maturity). In the CSSLs, four phenotypes were evaluated (starch and sucrose at heading and maturity).

Near-infrared spectroscopy measurements and carbohydrate assays: NSC components (relative starch and sucrose, % w/w) and total non-structural carbohydrates (TNC, %w/w) were predicted for the 976 samples (485 from heading and 491 from maturity) that were part of the *tropical japonica* diversity panel using near-infrared spectral data. Calibration models were developed previously on a set of 434 samples of diverse rice stems; the PLS-1 model predicted univariate TNC while the PLS-2 model predicted starch and sucrose components at once (refer to Chapter 2

of this dissertation). The 976 prediction samples were packed in 2-dram borosilicate scintillation vials and scanned using a Thermo Antaris II FT-NIR Spectrometer with a 40-position autosampler carousel (Thermo Fisher Scientific, Waltham, MA). Each sample was scanned 128 times (wavenumber range: 3,300-12,000 cm-1). For predictions, we used functions found in the following R packages: pls, signal, and prospectr.

Selection of *a priori* candidate genes: 120 *a priori* candidate genes were identified based on membership in the following pathways according to RiceCyc (Dharmawardhana et al., 2013): sucrose degradation, sucrose synthesis, starch degradation, starch synthesis. All pathways were classified under carbohydrate metabolism>starch and sucrose metabolism. An additional five members of the rice sucrose transporter gene family (OsSUT1-5,) were added to this list, resulting in a total of 125 *a priori* candidate genes (**Supplemental Table 4.3**).

Statistical analysis: Data handling, analysis, and plotting were performed using R statistical software (https://www.R-project.org/). Kinship matrix calculations and genome-wide association analyses were carried out using functions within the rrBLUP R package. For U.S. rice GWAS, two principal components (PCs) were included in *ALL* and *JAPONICA* mixed models to control for population structure and one PC was included in the *TRJ* model. Visualization of results (Manhattan and qq plots) utilized the qqman R package. To deal with multiple testing inherent in GWAS, we set a Benjamini-Hochberg FDR threshold of 1% for the U.S. rice analysis and 10% for the *tropical japonica* analysis (Benjamini and Hochberg, 1995). A more stringent FDR was used for the U.S. rice GWAS due to potentially uncontrolled family structure

diagnosed in QQ plots. Local LD (r^2) was calculated for the msSNP within each association peak resulting from the *tropical japonica* GWAS runs using Plink1.9 (https://www.cog-genomics.org/plink2, Chang et al., 2015). A critical r^2 was first determined as the 95th percentile of the distribution of all pairwise inter-chromosomal r^2 (i.e. unlinked) with the significant SNP marker. Next, pairwise r^2 values for all markers within 1Mb of the msSNP were plotted against physical distance from the msSNP. These data were then fit to a second-degree loess curve as described previously (Breseghello and Sorrells, 2006; Laido et al., 2014). The intercept of the resultant curve and the critical r^2 value was determined as the point of local LD decay (**Supplemental Fig. 4.11**). In the *tropical japonica* GWAS, the point of LD decay could not be determined for two of these loci; the msSNPs appeared unlinked to even their closest adjacent SNPs and the loess curve never intercepted the critical r^2 value. Candidate regions were further explored using the GWAS Viewer Tool (McCouch et al., 2016). Stepwise regression within IciMapping v4.0 software (http://www.isbreeding.net/software/?type=detail&id=14, RSTEP-LRT option) was implemented to map QTL for stem NSC traits in the interspecific CSSL population on line means. Markers with LOD greater than 3.0 were declared significant. For sequence analysis of *OsIsa2* gene and putative promoter region, sequences at chromosome 5 from 19154970 bp – 19159100 bp were extracted for cv. Curinga from a previously published re-sequencing dataset (Duitama et al., 2015) and for O. rufipogon IRGC105491 from an unpublished re-sequencing dataset from the McCouch Lab. All sites that were successfully called with respect to the reference genome were aligned and polymorphisms between the two accessions were extracted.

The alignments will be deposited in NCBI's SRA (short read archives) upon publication.

APPENDIX

Supplemental Table 4.1. Germplasm information

Name	subpopulation	ID used in reference	Panel	Reference
A 301	TRJ-2	A 301	US rice	Chapter 3
ADAIR	TRJ-2	ADAIR	US rice	Chapter 3
AHRENT	TRJ-2	AHRENT	US rice	Chapter 3
ALAN	TRJ-2	ALAN	US rice	Chapter 3
ANTONIO	TRJ-1	ANTONIO	US rice	Chapter 3
ARBORIO	admix	ARBORIO	US rice	Chapter 3
ARKROSE	TRJ-1	ARKROSE	US rice	Chapter 3
BANKS	TRJ-2	BANKS	US rice	Chapter 3
BELLE_PATNA	TRJ-1	BELLE_PATNA	US rice	Chapter 3
BELLEMONT	TRJ-1	BELLEMONT	US rice	Chapter 3
BELLEVUE	TRJ-2	BELLEVUE	US rice	Chapter 3
BENGAL	admix	BENGAL	US rice	Chapter 3
BLUEBONNET	TRJ-1	BLUEBONNET	US rice	Chapter 3
BLUE_ROSE	TRJ-1	BLUE_ROSE	US rice	Chapter 3
BLUE_ROSE_SUPREME	TRJ-1	BLUE_ROSE_SUPRE ME	US rice	Chapter 3
BLUEBELLE	TRJ-1	BLUEBELLE	US rice	Chapter 3
BOND	TRJ-1	BOND	US rice	Chapter 3
BONNET_73	TRJ-2	BONNET_73	US rice	Chapter 3
BOWMAN	TRJ-1	BOWMAN	US rice	Chapter 3

BRAZOS	TRJ-1	BRAZOS	US rice	Chapter 3
CAFFEY	TRJ-1	CAFFEY	US rice	Chapter 3
CALORO	TEJ	CALORO	US rice	Chapter 3
CALROSE	admix	CALROSE	US rice	Chapter 3
CAROLINA_GOLD	TRJ-1	CAROLINA_GOLD	US rice	Chapter 3
CAROLINA_GOLD	TRJ-1	CAROLINA_GOLD	US rice	Chapter 3
CAROLINA_GOLD_SELECT	TRJ-1	CAROLINA_GOLD_SELECT	US rice	Chapter 3
CAROLINA_GOLD_SELECT	TRJ-1	CAROLINA_GOLD_SELECT	US rice	Chapter 3
CB801	TRJ-2	CB801	US rice	Chapter 3
CENTURY_PATNA_231	TRJ-2	CENTURY_PATNA_231	US rice	Chapter 3
CHENIERE	TRJ-2	CHENIERE	US rice	Chapter 3
CHINESE	TEJ	CHINESE	US rice	Chapter 3
CL111	TRJ-1	CL111	US rice	Chapter 3
CL121	TRJ-2	CL121	US rice	Chapter 3
CL131	TRJ-2	CL131	US rice	Chapter 3
CL142_AR	TRJ-2	CL142_AR	US rice	Chapter 3
CL151	TRJ-2	CL151	US rice	Chapter 3
CL152	TRJ-2	CL152	US rice	Chapter 3
CL161	TRJ-2	CL161	US rice	Chapter 3
CL162	TRJ-2	CL162	US rice	Chapter 3
CL171_AR	TRJ-2	CL171_AR	US rice	Chapter 3
CL261	admix	CL261	US rice	Chapter 3

COCODRIE	TRJ-2	COCODRIE	US rice	Chapter 3
COLORADO	TRJ-2	COLORADO	US rice	Chapter 3
COLUSA	TRJ-1	COLUSA	US rice	Chapter 3
CYBONNET	TRJ-2	CYBONNET	US rice	Chapter 3
DAWN	TRJ-2	DAWN	US rice	Chapter 3
DELLA	TRJ-2	DELLA	US rice	Chapter 3
DELLA_2	TRJ-1	DELLA_2	US rice	Chapter 3
DELLROSE	TRJ-2	DELLROSE	US rice	Chapter 3
DIXIEBELLE	TRJ-1	DIXIEBELLE	US rice	Chapter 3
DREW	TRJ-2	DREW	US rice	Chapter 3
EARL	TRJ-1	EARL	US rice	Chapter 3
EARLY_WATARIBUNE	TEJ	EARLY_WATARIBUNE	US rice	Chapter 3
EDITH	TRJ-1	EDITH	US rice	Chapter 3
FORTUNA_	admix	FORTUNA_	US rice	Chapter 3
FRANCIS	TRJ-2	FRANCIS	US rice	Chapter 3
GULFMONT	TRJ-2	GULFMONT	US rice	Chapter 3
GULFROSE	TRJ-1	GULFROSE	US rice	Chapter 3
HILL_LONG_GRAIN	admix	HILL_LONG_GRAIN	US rice	Chapter 3
HONDURAS	TRJ-1	HONDURAS	US rice	Chapter 3
IMPROVED_BLUE_ROSE	TRJ-1	IMPROVED_BLUE_ROSE	US rice	Chapter 3
TAINAN_IKU_487	TEJ	TAINAN_IKU_487	US rice	Chapter 3
JACKSON	TRJ-2	JACKSON	US rice	Chapter 3
JAZZMAN_2	TRJ-2	JAZZMAN_2	US rice	Chapter 3

JEFFERSON	TRJ-2	JEFFERSON	US rice	Chapter 3
JODON	TRJ-2	JODON	US rice	Chapter 3
JUPITER	TRJ-1	JUPITER	US rice	Chapter 3
KATY	TRJ-2	KATY	US rice	Chapter 3
KAYBONNET	TRJ-2	KAYBONNET	US rice	Chapter 3
KOSHIHIKARI	TEJ	KOSHIHIKARI	US rice	Chapter 3
L-201	TRJ-2	L-201	US rice	Chapter 3
L202	TRJ-2	L202	US rice	Chapter 3
LABELLE	TRJ-2	LABELLE	US rice	Chapter 3
LACASSINE	TRJ-2	LACASSINE	US rice	Chapter 3
LACROSSE	TRJ-1	LACROSSE	US rice	Chapter 3
LADY_WRIGHT_SELECTION	TRJ-1	LADY_WRIGHT_SELECTION	US rice	Chapter 3
LAFITTE	TRJ-1	LAFITTE	US rice	Chapter 3
LAGRUE	TRJ-1	LAGRUE	US rice	Chapter 3
LEAH	TRJ-2	LEAH	US rice	Chapter 3
LEBONNET	TRJ-2	LEBONNET	US rice	Chapter 3
LEMONT	TRJ-1	LEMONT	US rice	Chapter 3
MADISON	TRJ-2	MADISON	US rice	Chapter 3
MARS	admix	MARS	US rice	Chapter 3
MAYBELLE	TRJ-2	MAYBELLE	US rice	Chapter 3
MEDARK	admix	MEDARK	US rice	Chapter 3
MELROSE	TRJ-1	MELROSE	US rice	Chapter 3
MERCURY	TRJ-1	MERCURY	US rice	Chapter 3
MERMENTAU	TRJ-2	MERMENTAU	US rice	Chapter 3

MILLIE	TRJ-2	MILLIE	US rice	Chapter 3
NATO	TRJ-1	NATO	US rice	Chapter 3
NECHES	TRJ-2	NECHES	US rice	Chapter 3
NEPTUNE	TRJ-1	NEPTUNE	US rice	Chapter 3
NEWBONNET	TRJ-2	NEWBONNET	US rice	Chapter 3
NEWREX	TRJ-2	NEWREX	US rice	Chapter 3
NORTAI	TEJ	NORTAI	US rice	Chapter 3
NORTHROSE	TRJ-1	NORTHROSE	US rice	Chapter 3
NOVA	TRJ-1	NOVA	US rice	Chapter 3
NOVA_66	admix	NOVA_66	US rice	Chapter 3
NOVA_76	TRJ-1	NOVA_76	US rice	Chapter 3
ORION	TRJ-1	ORION	US rice	Chapter 3
PANDA	TRJ-1	PANDA	US rice	Chapter 3
PECOS	TRJ-1	PECOS	US rice	Chapter 3
PIROGUE	admix	PIROGUE	US rice	Chapter 3
PRESIDIO	TRJ-1	PRESIDIO	US rice	Chapter 3
PRISCILLA	TRJ-2	PRISCILLA	US rice	Chapter 3
REX	TRJ-2	REX	US rice	Chapter 3
REXARK_	admix	REXARK_	US rice	Chapter 3
REXMONT	TRJ-2	REXMONT	US rice	Chapter 3
RICO1	TRJ-1	RICO1	US rice	Chapter 3
ROSEMONT	TRJ-2	ROSEMONT	US rice	Chapter 3
ROY_J	TRJ-2	ROY_J	US rice	Chapter 3
RT7015	TRJ-2	RT7015	US rice	Chapter 3
SABER	TRJ-2	SABER	US rice	Chapter 3

SABINE	TRJ-2	SABINE	US rice	Chapter 3
SATURN	TRJ-1	SATURN	US rice	Chapter 3
SIERRA	TRJ-2	SIERRA	US rice	Chapter 3
SINAMPAGA_SELECTION	admix	SINAMPAGA_SELECTION	US rice	Chapter 3
SKYBONNET	TRJ-2	SKYBONNET	US rice	Chapter 3
STARBONNET	TRJ-1	STARBONNET	US rice	Chapter 3
STG06L-35-61	TRJ-2	STG06L-35-61	US rice	Chapter 3
TAGGART	TRJ-2	TAGGART	US rice	Chapter 3
TEBONNET	TRJ-2	TEBONNET	US rice	Chapter 3
TEMPLETON	TRJ-1	TEMPLETON	US rice	Chapter 3
TEXMONT	TRJ-2	TEXMONT	US rice	Chapter 3
TORO	TRJ-1	TORO	US rice	Chapter 3
TORO-2	TRJ-2	TORO-2	US rice	Chapter 3
TP49	TRJ-1	TP49	US rice	Chapter 3
TRENASSE	TRJ-2	TRENASSE	US rice	Chapter 3
V7817	TRJ-1	V7817	US rice	Chapter 3
VEGOLD	TRJ-1	VEGOLD	US rice	Chapter 3
VISTA	TRJ-1	VISTA	US rice	Chapter 3
WELLS	TRJ-2	WELLS	US rice	Chapter 3
ZENITH	TRJ-2	ZENITH	US rice	Chapter 3
CURUF-1	tropical japonica	RUF1	CSSL	Arbelaez et al 2015
CURUF-2	tropical japonica	RUF2	CSSL	Arbelaez et al 2015
CURUF-3	tropical japonica	RUF3	CSSL	Arbelaez et al 2015
CURUF-4	tropical japonica	RUF4	CSSL	Arbelaez et al 2015

CURUF-5	tropical japonica	RUF5	CSSL	Arbelaez et al 2015
CURUF-6	tropical japonica	RUF6	CSSL	Arbelaez et al 2015
CURUF-7	tropical japonica	RUF7	CSSL	Arbelaez et al 2015
CURUF-8	tropical japonica	RUF8	CSSL	Arbelaez et al 2015
CURUF-9	tropical japonica	RUF9	CSSL	Arbelaez et al 2015
CURUF-10	tropical japonica	RUF10	CSSL	Arbelaez et al 2015
CURUF-11	tropical japonica	RUF11	CSSL	Arbelaez et al 2015
CURUF-12	tropical japonica	RUF12	CSSL	Arbelaez et al 2015
CURUF-13	tropical japonica	RUF13	CSSL	Arbelaez et al 2015
CURUF-14	tropical japonica	RUF14	CSSL	Arbelaez et al 2015
CURUF-15	tropical japonica	RUF15	CSSL	Arbelaez et al 2015
CURUF-16	tropical japonica	RUF16	CSSL	Arbelaez et al 2015
CURUF-17	tropical japonica	RUF17	CSSL	Arbelaez et al 2015
CURUF-18	tropical japonica	RUF18	CSSL	Arbelaez et al 2015
CURUF-19	tropical japonica	RUF19	CSSL	Arbelaez et al 2015
CURUF-20	tropical japonica	RUF20	CSSL	Arbelaez et al 2015
CURUF-21	tropical japonica	RUF21	CSSL	Arbelaez et al 2015
CURUF-22	tropical japonica	RUF22	CSSL	Arbelaez et al 2015
CURUF-23	tropical japonica	RUF23	CSSL	Arbelaez et al 2015
CURUF-24	tropical japonica	RUF24	CSSL	Arbelaez et al 2015
CURUF-25	tropical japonica	RUF25	CSSL	Arbelaez et al 2015
CURUF-26	tropical japonica	RUF26	CSSL	Arbelaez et al 2015
CURUF-27	tropical japonica	RUF27	CSSL	Arbelaez et al 2015
CURUF-28	tropical japonica	RUF28	CSSL	Arbelaez et al 2015
CURUF-29	tropical japonica	RUF29	CSSL	Arbelaez et al 2015

CURUF-30	tropical japonica	RUF30	CSSL	Arbelaez et al 2015
CURUF-31	tropical japonica	RUF31	CSSL	Arbelaez et al 2015
CURUF-32	tropical japonica	RUF32	CSSL	Arbelaez et al 2015
CURUF-33	tropical japonica	RUF33	CSSL	Arbelaez et al 2015
CURUF-34	tropical japonica	RUF34	CSSL	Arbelaez et al 2015
CURUF-35	tropical japonica	RUF35	CSSL	Arbelaez et al 2015
CURUF-36	tropical japonica	RUF36	CSSL	Arbelaez et al 2015
CURUF-37	tropical japonica	RUF37	CSSL	Arbelaez et al 2015
CURUF-38	tropical japonica	RUF38	CSSL	Arbelaez et al 2015
CURUF-39	tropical japonica	RUF39	CSSL	Arbelaez et al 2015
CURUF-40	tropical japonica	RUF40	CSSL	Arbelaez et al 2015
CURUF-41	tropical japonica	RUF41	CSSL	Arbelaez et al 2015
CURUF-42	tropical japonica	RUF42	CSSL	Arbelaez et al 2015
CURUF-43	tropical japonica	RUF43	CSSL	Arbelaez et al 2015
CURUF-44	tropical japonica	RUF44	CSSL	Arbelaez et al 2015
CURUF-45	tropical japonica	RUF45	CSSL	Arbelaez et al 2015
CURUF-46	tropical japonica	RUF46	CSSL	Arbelaez et al 2015
CURUF-47	tropical japonica	RUF47	CSSL	Arbelaez et al 2015
CURUF-48	tropical japonica	RUF48	CSSL	Arbelaez et al 2015
NSFTV175	tropical-japonica	NSFTV175	diversity	McCouch et al 2016
318	tropical-japonica	NSFTV384	diversity	McCouch et al 2016
325	tropical-japonica	NSFTV381	diversity	McCouch et al 2016
583	tropical-japonica	NSFTV176	diversity	McCouch et al 2016
Arabi	admixed-japonica	NSFTV244	diversity	McCouch et al 2016
ARC 12701::IRGC22267-1	admixed-japonica	IRGC121595	diversity	McCouch et al 2016

Asse Y Pung	tropical-japonica	NSFTV8	diversity	McCouch et al 2016
Azucena	tropical-japonica	NSFTV635	diversity	McCouch et al 2016
B6616A4-22-Bk-5-4	tropical-japonica	NSFTV167	diversity	McCouch et al 2016
BAKUNG (H)::IRGC60220-1	tropical-japonica	IRGC121193	diversity	McCouch et al 2016
Baldo	admixed-japonica	NSFTV264	diversity	McCouch et al 2016
Bengal	admixed-japonica	NSFTV622	diversity	McCouch et al 2016
Blue Rose	admixed-japonica	NSFTV20	diversity	McCouch et al 2016
Blue Rose Supreme	admixed-japonica	NSFTV182	diversity	McCouch et al 2016
BLUEBONNET 50::IRGC1811-1	tropical-japonica	IRGC121874	diversity	McCouch et al 2016
Boa Vista	tropical-japonica	NSFTV183	diversity	McCouch et al 2016
BORUBI::IRGC25181-1	tropical-japonica	IRGC121875	diversity	McCouch et al 2016
British Honduras Creole	tropical-japonica	NSFTV185	diversity	McCouch et al 2016
C57-5043	tropical-japonica	NSFTV187	diversity	McCouch et al 2016
Caawa/Fortuna 6-103-15	tropical-japonica	NSFTV22	diversity	McCouch et al 2016
Carolina Gold 12034	tropical-japonica	NSFTV25	diversity	McCouch et al 2016
Carolina Gold Sel	tropical-japonica	NSFTV26	diversity	McCouch et al 2016
Cenit	tropical-japonica	NSFTV342	diversity	McCouch et al 2016
CHA LOY OE::C1	admixed-japonica	IRGC122004	diversity	McCouch et al 2016
CI 11011	tropical-japonica	NSFTV389	diversity	McCouch et al 2016
CINA::IRGC27116-1	tropical-japonica	IRGC121905	diversity	McCouch et al 2016
Cocodrie	tropical-japonica	NSFTV396	diversity	McCouch et al 2016
Creole	tropical-japonica	NSFTV347	diversity	McCouch et al 2016
Cuba 65	tropical-japonica	NSFTV37	diversity	McCouch et al 2016
Cybonnet	tropical-japonica	NSFTV397	diversity	McCouch et al 2016

Cypress	tropical-japonica	NSFTV647	diversity	McCouch et al 2016
D 4-136::IRGC31051-1	tropical-japonica	IRGC121876	diversity	McCouch et al 2016
Dam	admixed-japonica	NSFTV40	diversity	McCouch et al 2016
DAVAO::IRGC8244-C1	tropical-japonica	IRGC122031	diversity	McCouch et al 2016
Della	tropical-japonica	NSFTV391	diversity	McCouch et al 2016
DINOLORES::IRGC67431-1	tropical-japonica	IRGC121195	diversity	McCouch et al 2016
Doble Carolina Rinaldo Barsani	admixed-japonica	NSFTV305	diversity	McCouch et al 2016
Dourado Agulha	tropical-japonica	NSFTV46	diversity	McCouch et al 2016
Dragon Eyeball 100	admixed-japonica	NSFTV631	diversity	McCouch et al 2016
Edith	tropical-japonica	NSFTV392	diversity	McCouch et al 2016
Estrela	admixed-japonica	NSFTV237	diversity	McCouch et al 2016
Fortuna	tropical-japonica	NSFTV54	diversity	McCouch et al 2016
Fossa Av	tropical-japonica	NSFTV193	diversity	McCouch et al 2016
Francis	tropical-japonica	NSFTV632	diversity	McCouch et al 2016
GANIGI::IRGC48698-C1	tropical-japonica	IRGC122052	diversity	McCouch et al 2016
Gerdeh	admixed-japonica	NSFTV55	diversity	McCouch et al 2016
GOGO::IRGC43390-C1	tropical-japonica	IRGC122063	diversity	McCouch et al 2016
Gogo Lempuk	tropical-japonica	NSFTV59	diversity	McCouch et al 2016
Gotak Gatik	admixed-japonica	NSFTV60	diversity	McCouch et al 2016
Guatemala 1021	tropical-japonica	NSFTV352	diversity	McCouch et al 2016
H256-76-1-1-1	tropical-japonica	NSFTV251	diversity	McCouch et al 2016
Hiderisirazu	admixed-japonica	NSFTV266	diversity	McCouch et al 2016
Honduras	tropical-japonica	NSFTV65	diversity	McCouch et al 2016
IAC 1111::IRGC39050-1	tropical-japonica	IRGC124476	diversity	McCouch et al 2016
IAC 25	tropical-japonica	NSFTV69	diversity	McCouch et al 2016

Iguape Cateto	tropical-japonica	NSFTV70	diversity	McCouch et al 2016
IITA 135	tropical-japonica	NSFTV286	diversity	McCouch et al 2016
ILANG-ILANG::GERVEX 509-C1	tropical-japonica	IRGC121747	diversity	McCouch et al 2016
IMMARANG::IRGC44455-1	tropical-japonica	IRGC124461	diversity	McCouch et al 2016
WC 4419	tropical-japonica	NSFTV214	diversity	McCouch et al 2016
IR 47684-05-1-B::C1	tropical-japonica	IRGC121749	diversity	McCouch et al 2016
IR 47686-09-01-B-1::C1	tropical-japonica	IRGC122095	diversity	McCouch et al 2016
IR 63372-8::C1	tropical-japonica	IRGC121755	diversity	McCouch et al 2016
IR 68704-145-1-1-B::C1	tropical-japonica	IRGC122108	diversity	McCouch et al 2016
IRAT 109::GERVEX 4988-C1	tropical-japonica	IRGC121761	diversity	McCouch et al 2016
IRAT 177	tropical-japonica	NSFTV73	diversity	McCouch et al 2016
IRAT 2::GERVEX 606-C1	tropical-japonica	IRGC121766	diversity	McCouch et al 2016
IRAT 212::GERVEX 7698-C1	tropical-japonica	IRGC121763	diversity	McCouch et al 2016
IRAT 216::GERVEX 7702-C1	tropical-japonica	IRGC121764	diversity	McCouch et al 2016
IRAT 257::C1	tropical-japonica	IRGC121765	diversity	McCouch et al 2016
IRAT 335::C1	tropical-japonica	IRGC122120	diversity	McCouch et al 2016
IRAT 362::GERVEX 8712-C1	tropical-japonica	IRGC121767	diversity	McCouch et al 2016
IRAT 364::GERVEX 8714-C1	tropical-japonica	IRGC121768	diversity	McCouch et al 2016
IRAT 366::GERVEX 8716-C1	tropical-japonica	IRGC121769	diversity	McCouch et al 2016
IRAT 380::C1	tropical-japonica	IRGC121770	diversity	McCouch et al 2016
IRAT 13	tropical-japonica	NSFTV195	diversity	McCouch et al 2016
IRAT 44	tropical-japonica	NSFTV226	diversity	McCouch et al 2016
NSFTV362	tropical-japonica	NSFTV362	diversity	McCouch et al 2016
Jambu	tropical-japonica	NSFTV75	diversity	McCouch et al 2016

Jefferson	tropical-japonica	NSFTV628	diversity	McCouch et al 2016
Kaniranga	tropical-japonica	NSFTV84	diversity	McCouch et al 2016
Katy	tropical-japonica	NSFTV625	diversity	McCouch et al 2016
Kaybonnet	tropical-japonica	NSFTV624	diversity	McCouch et al 2016
KENDINGA 5 H::IRGC60310-C1	tropical-japonica	IRGC122134	diversity	McCouch et al 2016
Keriting Tingii	admixed-japonica	NSFTV87	diversity	McCouch et al 2016
KHAO DAM	admixed-japonica	IRGC117513	diversity	McCouch et al 2016
Khao do ngoi::IRGC29772-1	admixed-japonica	IRGC121208	diversity	McCouch et al 2016
KIKILONG::IRGC71539-1	tropical-japonica	IRGC121391	diversity	McCouch et al 2016
Kinastano	tropical-japonica	NSFTV92	diversity	McCouch et al 2016
KING::GERVEX 1655-C1	tropical-japonica	IRGC121777	diversity	McCouch et al 2016
KOMOJAMANITRA::GERVEX 8453-C1	tropical-japonica	IRGC121779	diversity	McCouch et al 2016
KU115	tropical-japonica	NSFTV96	diversity	McCouch et al 2016
KYEEMA::GERVEX 1656-C1	tropical-japonica	IRGC121782	diversity	McCouch et al 2016
L 204::GERVEX 1658-C1	admixed-japonica	IRGC121783	diversity	McCouch et al 2016
L-202	tropical-japonica	NSFTV98	diversity	McCouch et al 2016
LAC 23	tropical-japonica	NSFTV99	diversity	McCouch et al 2016
Lacrosse	admixed-japonica	NSFTV100	diversity	McCouch et al 2016
Lady Wright Seln	tropical-japonica	NSFTV394	diversity	McCouch et al 2016
LaGrue	tropical-japonica	NSFTV621	diversity	McCouch et al 2016
Leah	tropical-japonica	NSFTV198	diversity	McCouch et al 2016
Lemont	tropical-japonica	NSFTV101	diversity	McCouch et al 2016
Ligerito	tropical-japonica	NSFTV350	diversity	McCouch et al 2016

Llanero 501	tropical-japonica	NSFTV308	diversity	McCouch et al 2016
LUDAN::IRGC64189-C1	tropical-japonica	IRGC121787	diversity	McCouch et al 2016
Manzano	tropical-japonica	NSFTV309	diversity	McCouch et al 2016
Mojito Colorado	tropical-japonica	NSFTV770	diversity	McCouch et al 2016
Moroberekan	tropical-japonica	NSFTV108	diversity	McCouch et al 2016
MUT IAC 25-44-807::IRGC68799-1	tropical-japonica	IRGC121672	diversity	McCouch et al 2016
N 3::GERVEX 547-C1	admixed-japonica	IRGC121797	diversity	McCouch et al 2016
NENG NAH::IRGC78275-1	admixed-japonica	IRGC121450	diversity	McCouch et al 2016
Nortai	admixed-japonica	NSFTV388	diversity	McCouch et al 2016
Nova	admixed-japonica	NSFTV114	diversity	McCouch et al 2016
NSFTV107	tropical-japonica	NSFTV107	diversity	McCouch et al 2016
NSFTV116	tropical-japonica	NSFTV116	diversity	McCouch et al 2016
NSFTV199	tropical-japonica	NSFTV199	diversity	McCouch et al 2016
NSFTV27	tropical-japonica	NSFTV27	diversity	McCouch et al 2016
NSFTV89	tropical-japonica	NSFTV89	diversity	McCouch et al 2016
Okshitmayin	admixed-japonica	NSFTV335	diversity	McCouch et al 2016
ORYZICA SABANA 6::C1	tropical-japonica	IRGC121801	diversity	McCouch et al 2016
OS 6 (WC 10296)	tropical-japonica	NSFTV395	diversity	McCouch et al 2016
OS6	tropical-japonica	NSFTV120	diversity	McCouch et al 2016
Ostiglia	admixed-japonica	NSFTV121	diversity	McCouch et al 2016
P. TINGAGEW DAYKET QAY DAYON::IRGC8046-1	tropical-japonica	IRGC121930	diversity	McCouch et al 2016
Padi Pagalong	tropical-japonica	NSFTV274	diversity	McCouch et al 2016
PALAWAN::C1	tropical-japonica	IRGC121803	diversity	McCouch et al 2016

Panda	admixed-japonica	NSFTV629	diversity	McCouch et al 2016
Pate Blanc Mn 1	tropical-japonica	NSFTV201	diversity	McCouch et al 2016
Patna	admixed-japonica	NSFTV281	diversity	McCouch et al 2016
Pato De Gallinazo	admixed-japonica	NSFTV128	diversity	McCouch et al 2016
Pecos	admixed-japonica	NSFTV618	diversity	McCouch et al 2016
Peek::IRGC11821-1	admixed-japonica	IRGC121469	diversity	McCouch et al 2016
PI 298967-1	admixed-japonica	NSFTV218	diversity	McCouch et al 2016
PIN KAEO	tropical-japonica	IRGC117560	diversity	McCouch et al 2016
PR 304	tropical-japonica	NSFTV377	diversity	McCouch et al 2016
Pratao	tropical-japonica	NSFTV202	diversity	McCouch et al 2016
PRIMAVERA::C1	tropical-japonica	IRGC121809	diversity	McCouch et al 2016
R 101	tropical-japonica	NSFTV310	diversity	McCouch et al 2016
RANAU KADAI::IRGC71604-1	tropical-japonica	IRGC121635	diversity	McCouch et al 2016
Rikuto Norin 21	admixed-japonica	NSFTV364	diversity	McCouch et al 2016
Rosemont	tropical-japonica	NSFTV619	diversity	McCouch et al 2016
S4542A3-49B-2B12	tropical-japonica	NSFTV139	diversity	McCouch et al 2016
Saber	tropical-japonica	NSFTV630	diversity	McCouch et al 2016
Saku	tropical-japonica	NSFTV280	diversity	McCouch et al 2016
Saturn	tropical-japonica	NSFTV140	diversity	McCouch et al 2016
Sinaguing	tropical-japonica	NSFTV149	diversity	McCouch et al 2016
Sinampaga Selection	tropical-japonica	NSFTV147	diversity	McCouch et al 2016
Spring	tropical-japonica	NSFTV399	diversity	McCouch et al 2016
Sultani	tropical-japonica	NSFTV150	diversity	McCouch et al 2016
Tia Bura	tropical-japonica	NSFTV258	diversity	McCouch et al 2016

TOANG::IRGC19144-1	tropical-japonica	IRGC121523	diversity	McCouch et al 2016
Tokyo Shino Mochi	admixed-japonica	NSFTV211	diversity	McCouch et al 2016
Tondok	tropical-japonica	NSFTV164	diversity	McCouch et al 2016
Tox 782-20-1	tropical-japonica	NSFTV285	diversity	McCouch et al 2016
Trembese	tropical-japonica	NSFTV165	diversity	McCouch et al 2016
Upland	tropical-japonica	NSFTV375	diversity	McCouch et al 2016
Varyla	tropical-japonica	NSFTV273	diversity	McCouch et al 2016
Vialone	admixed-japonica	NSFTV265	diversity	McCouch et al 2016
VIETNAM 1::C1	admixed-japonica	IRGC122281	diversity	McCouch et al 2016
WAB 501-11-5-1	tropical-japonica	NSFTV240	diversity	McCouch et al 2016
WAB 502-13-4-1	tropical-japonica	NSFTV239	diversity	McCouch et al 2016
WAB 56-104	tropical-japonica	NSFTV238	diversity	McCouch et al 2016
Wanica	tropical-japonica	NSFTV379	diversity	McCouch et al 2016
NSFTV212	tropical-japonica	NSFTV212	diversity	McCouch et al 2016
WC 3397	tropical-japonica	NSFTV213	diversity	McCouch et al 2016
WC 4443	tropical-japonica	NSFTV215	diversity	McCouch et al 2016
WC 521	admixed-japonica	NSFTV236	diversity	McCouch et al 2016
Wells	tropical-japonica	NSFTV170	diversity	McCouch et al 2016
YANCAOUSSA::IRGC16071-C1	tropical-japonica	IRGC121859	diversity	McCouch et al 2016
YRL-1	admixed-japonica	NSFTV217	diversity	McCouch et al 2016
Canella De Ferro	tropical-japonica	NSFTV23	diversity	McCouch et al 2016
KAMPAI 15::IRGC78246-1	admixed-japonica	IRGC121378	diversity	McCouch et al 2016
LAI::IRGC15023-1	admixed-japonica	IRGC121405	diversity	McCouch et al 2016
PADI TERAS::IRGC25554-1	tropical-japonica	IRGC121926	diversity	McCouch et al 2016

Padi Kasalle	tropical-japonica	NSFTV122	diversity	McCouch et al 2016
POPONG::IRGC13297-1	tropical-japonica	IRGC121475	diversity	McCouch et al 2016

Supplemental Table 4.2. GWAS results from *US-TRJ* panel

trait	chr	peak msSNP	pos	-log10(p)
starch, hd	5	S5_26025809	26025809	8.79556
	5	S5_578975	578975	5.86207
	12	S12_22283218	22283218	8.776116
	12	S12_24716046	24716046	8.393052
starch, mt	1	S1_29976759	29976759	7.831426
	9	S9_22045460	22045460	7.663395
	4	S4_94653	94653	6.325673
sucrose, hd	5	S5_704530	704530	8.047205
	11	S11_25167461	25167461	5.601342
sucrose, mt	1	S1_29495819	29495819	15.38403
	5	S5_19425787	19425787	6.364683
	12	S12_10707252	10707252	11.93229
	3	S3_8125367	8125367	6.835099

Supplemental Table 4.3. ***A priori*** **candidate gene list.**

	Gene name	pathway	chr	start pos	end pos	Reaction id	Enzymatic activity
1	LOC_Os01g09460	starch_degrad	1	4820104	4823129	RXN-1685	hexokinase, putative, expressed
2	LOC_Os01g13550	starch_degrad	1	7578091	7582923	RXN-1827	beta-amylase, putative, expressed
3	LOC_Os01g15910	suc_degrad	1	8958411	8967207	GLUC1PURIDYLTRANS-RXN	UTP--glucose-1-phosphate uridylyltransferase, putative
4	LOC_Os01g22900	suc_degrad	1	12870228	12874264	3.2.1.26-RXN	beta-fructofuranosidase
5	LOC_Os01g24950	starch_degrad	1	14059572	14064905	RXN-1685	exocyst complex component 6, putative, expressed
6	LOC_Os01	suc_biosyn	1	15568776	15573377	SUCROSE-	sucrose-

	g27880					PHOSPHATASE-RXN	phosphatase, putative, expressed
7	LOC_Os01g51754	starch_degrad	1	29760719	29770037	RXN-1823	alpha-amylase
8	LOC_Os01g52450	starch_degrad	1	30131348	30135287	RXN-1685	hexokinase
9	LOC_Os01g53930	starch_degrad	1	31009006	31014001	RXN-1685	hexokinase
10	LOC_Os01g62840	suc_degrad	1	36395486	36398185	GLUC1PURIDYLTRANS-RXN	mannose-1-phosphate guanyltransferase, putative, expressed
11	LOC_Os01g62850	starch_degrad	1	36398958	36401777	RXN-1827	beta-amylase
12	LOC_Os01g63270	starch_degrad	1	36670221	36676478	RXN-1826	transferase, transferring glycosyl groups
13	LOC_Os01g66940	suc_degrad	1	38884741	38887723	FRUCTOKINASE-RXN	fructokinase
14	LOC_Os01g69030	suc_biosyn	1	40101404	40107186	SUCROSE-PHOSPHATE-SYNTHASE-	sucrose-phosphate synthase

						RXN	
15	LOC_Os01g71320	starch_degrad	1	41305317	41314527	RXN-1685	phosphopyruvate hydratase
16	LOC_Os01g73580	suc_degrad	1	42626321	42630440	3.2.1.26-RXN	hydrolase, hydrolyzing O-glycosyl compounds
17	LOC_Os02g01590	suc_degrad	2	343057	347196	3.2.1.26-RXN	beta-fructofuranosidase
18	LOC_Os02g03320	suc_degrad	2	1341178	1344290	3.2.1.26-RXN	plant neutral invertase domain containing protein, expressed
19	LOC_Os02g03690	starch_degrad	2	1539488	1548149	RXN-1827	2-oxoglutarate decarboxylase
20	LOC_Os02g05030	suc_biosyn	2	2388399	2394346	SUCROSE-PHOSPHATASE-RXN	sucrose-phosphatase, putative, expressed
21	LOC_Os02	suc_biosyn	2	4708493	4716880	SUCROSE-	sucrose-

	g09170					PHOSPHATE-SYNTHASE-RXN	phosphate synthase
22	LOC_Os02g17500	suc_degrad	2	10074581	10081512	PGLUCISOM-RXN	transporter family protein, putative, expressed
23	LOC_Os02g33110	suc_degrad	2	19682543	19687250	3.2.1.26-RXN	beta-fructofuranosidase
24	LOC_Os02g34560	suc_degrad	2	20717887	20721920	3.2.1.26-RXN	beta-fructofuranosidase
25	LOC_Os02g52480	starch_degrad	2	32111311	32112949	RXN-1685	cyclin-dependent kinase inhibitor, putative, expressed
26	LOC_Os02g52700	starch_degrad	2	32243146	32245056	RXN-1823	alpha-amylase
27	LOC_Os02g52710	starch_degrad	2	32248279	32250180	RXN-1823	alpha-amylase
28	LOC_Os02g58480	suc_degrad	2	35754946	35761781	SUCROSE-SYNTHASE-	UDP-glycosyltran

						RXN	sferase
29	LOC_Os03g04770	starch_degrad	3	2272965	2276401	RXN-1827	beta-amylase
30	LOC_Os03g11050	suc_degrad	3	5680647	5685409	GLUC1PURIDYLTRANS-RXN	UTP-glucose-1-phosphate uridylyltransferase
31	LOC_Os03g16150	suc_degrad	3	8909153	8912958	GLUC1PURIDYLTRANS-RXN	mannose-1-phosphate guanyltransferase, putative, expressed
32	LOC_Os03g20020	suc_degrad	3	11283340	11287896	3.2.1.26-RXN	beta-fructofuranosidase
33	LOC_Os03g22120	suc_degrad	3	12674462	12680657	SUCROSE-SYNTHASE-RXN	sucrose synthase, putative, expressed
34	LOC_Os03g22790	starch_degrad	3	13161913	13164014	RXN-1827	beta-amylase
35	LOC_Os03g28330	suc_degrad	3	16301277	16307632	SUCROSE-SYNTHASE-RXN	UDP-glycosyltransferase

36	LOC_Os03g50480	starch_degrad	3	28813570	28820161	PHOSPHOGLUCMUT-RXN	phosphoglucomutase
37	LOC_Os03g55090	starch_degrad	3	31332396	31339163	RXN-1826	phosphorylase
38	LOC_Os03g56460	suc_degrad	3	32175133	32182101	PGLUCISOM-RXN	glucose-6-phosphate isomerase
39	LOC_Os04g17650	suc_degrad	4	9661934	9667057	SUCROSE-SYNTHASE-RXN	UDP-glycosyltransferase
40	LOC_Os04g24430	suc_degrad	4	14008245	14013440	SUCROSE-SYNTHASE-RXN	UDP-glycosyltransferase
41	LOC_Os04g33040	starch_degrad	4	20006128	20009276	RXN-1823	alpha-amylase
42	LOC_Os04g45290	suc_degrad	4	26765199	26771893	3.2.1.26-RXN	beta-fructofuranosidase
43	LOC_Os04g56920	suc_degrad	4	33940929	33943427	3.2.1.26-RXN	beta-fructofuranosidase
44	LOC_Os05g05270	suc_biosyn	5	2595449	2599141	SUCROSE-PHOSPHATASE-RXN	sucrose-phosphatase, putative, expressed
45	LOC_Os05	starch_degrad	5	5337195	5341210	RXN-1685	hexokinase,

	g09500						putative, expressed
46	LOC_Os05g31110	starch_degrad	5	18075301	18081280	RXN-1685	phosphopyruvate hydratase
47	LOC_Os05g32710	starch_degrad	5	19154970	19157600	RXN-1823	alpha-amylase
48	LOC_Os05g38250	starch_degrad	5	22425361	22427795	RXN-1827	beta-amylase
49	LOC_Os05g44760	starch_degrad	5	26017295	26022937	RXN-1685	hexokinase
50	LOC_Os05g45590	starch_degrad	5	26418621	26422556	RXN-1685	hexokinase
51	LOC_Os06g09450	suc_degrad	6	4796286	4804246	SUCROSE-SYNTHASE-RXN	UDP-glycosyltransferase
52	LOC_Os06g14510	suc_degrad	6	8152235	8159385	PGLUCISOM-RXN	glucose-6-phosphate isomerase
53	LOC_Os06g26234	starch_degrad	6	15335687	15365531	RXN-1823	alpha-amylase
54	LOC_Os06g28194	starch_degrad	6	16011323	16032869	PHOSPHOGLUCMUT-RXN	expressed protein
55	LOC_Os06g43630	suc_biosyn	6	26242005	26250269	SUCROSE-PHOSPHATE-SYNTHASE-	sucrose-phosphate synthase

						RXN	
56	LOC_Os06g45980	starch_degrad	6	27866741	27874406	RXN-1685	hexokinase
57	LOC_Os06g46284	starch_degrad	6	28045147	28058082	RXN-2141	hydrolase, hydrolyzing O-glycosyl compounds
58	LOC_Os06g49970	starch_degrad	6	30262778	30266915	RXN-1823	alpha-amylase
59	LOC_Os07g09890	starch_degrad	7	5256320	5259795	RXN-1685	hexokinase
60	LOC_Os07g17180	starch_degrad	7	10109609	10111749	RXN-1685	no apical meristem protein, expressed
61	LOC_Os07g26540	starch_degrad	7	15293156	15294805	RXN-1685	hexokinase, putative, expressed
62	LOC_Os07g26610	starch_degrad	7	15338913	15351164	PHOSPHOGLUCMUT-RXN	phosphoglucomutase
63	LOC_Os07g35880	starch_degrad	7	21471055	21473327	RXN-1827	beta-amylase
64	LOC_Os07g35940	starch_degrad	7	21503815	21506356	RXN-1827	beta-amylase
65	LOC_Os07g42490	suc_degrad	7	25429639	25435182	SUCROSE-SYNTHASE-	UDP-glycosyltran

						RXN	sferase
66	LOC_Os07g43390	starch_degrad	7	25980425	25984955	RXN-1828	4-alpha-glucanotransferase
67	LOC_Os07g46790	starch_degrad	7	27960440	27968017	RXN-1828	4-alpha-glucanotransferase
68	LOC_Os07g47120	starch_degrad	7	28178236	28181579	RXN-1827	beta-amylase
69	LOC_Os08g02120	suc_degrad	8	701246	704015	FRUCTOKINASE-RXN	kinase
70	LOC_Os08g13930	suc_degrad	8	8342181	8346667	GLUC1PURIDYLTRANS-RXN	mannose-1-phosphate guanyltransferase, putative, expressed
71	LOC_Os08g20660	suc_biosyn	8	12411932	12424769	SUCROSE-PHOSPHATE-SYNTHASE-RXN	sucrose-phosphate synthase
72	LOC_Os08g36900	starch_degrad	8	23335165	23337151	RXN-1823	alpha-amylase
73	LOC_Os08g36910	starch_degrad	8	23340676	23343533	RXN-1823	alpha-amylase
74	LOC_Os08	suc_degrad	8	23623080	23628557	PGLUCISOM-	glucose-6-

	g37380					RXN	phosphate isomerase
75	LOC_Os09g07480	starch_degrad	9	3745683	3749470	RXN-1827	beta-amylase
76	LOC_Os09g23560	starch_degrad	9	14001192	14003182	RXN-1685	aryl-alcohol dehydrogenase
77	LOC_Os09g25460	starch_degrad	9	15251886	15256328	RXN-1685	transferase family protein, putative, expressed
78	LOC_Os09g28400	starch_degrad	9	17288993	17291295	RXN-1823	alpha-amylase
79	LOC_Os09g28420	starch_degrad	9	17296166	17305076	RXN-1823	alpha-amylase
80	LOC_Os09g29070	suc_degrad	9	17648471	17653621	PGLUCISOM-RXN	glucose-6-phosphate isomerase
81	LOC_Os09g29404	starch_degrad	9	17851430	17868669	RXN-1823	alpha-amylase
82	LOC_Os09g38030	suc_degrad	9	21917319	21922425	GLUC1PURIDYLTRANS-RXN	UTP-glucose-1-phosphate uridylyltransferase

83	LOC_Os09g39570	starch_degrad	9	22713280	22717110	RXN-1827	beta-amylase
84	LOC_Os10g11140	starch_degrad	10	6167111	6174631	PHOSPHOGLUCMUT-RXN	phosphoglucomutase
85	LOC_Os10g32810	starch_degrad	10	17180757	17183465	RXN-1827	beta-amylase
86	LOC_Os10g35020	starch_degrad	10	18682977	18685571	RXN-1685	glycosyltransferase, putative, expressed
87	LOC_Os10g41550	starch_degrad	10	22324198	22327665	RXN-1827	beta-amylase
88	LOC_Os11g07440	suc_degrad	11	3739630	3743621	3.2.1.26-RXN	beta-fructofuranosidase
89	LOC_Os11g12810	suc_biosyn	11	7255766	7263164	SUCROSE-PHOSPHATE-SYNTHASE-RXN	transferase, transferring glycosyl groups
90	LOC_Os12g04000	starch_degrad	12	1658206	1660737	RXN-1685	auxin efflux carrier component, putative, expressed
91	LOC_Os02g3670	suc_transport	2	22146420	22150147		OsSUT 5

92	LOC_Os02g58080	suc_transport	2	35548813	35553460		OsSUT 4
93	LOC_Os03g07480	suc_transport	3	3797691	3804132		OsSUT 1
94	LOC_Os10g26470	suc_transport	10	13778798	13785215		OsSUT 3
95	LOC_Os12g44380	suc_transport	12	27513939	27518251		OsSUT 2
96	LOC_Os01g44220	starch_biosyn	1	25353769	25362199	GLUC1PADENYLTRANS-RXN	glucose-1-phosphate adenylyltransferase
97	LOC_Os01g52250	starch_biosyn	1	30030997	30041476	GLYCOGENSYN-RXN	transferase, transferring glycosyl groups
98	LOC_Os02g17500	starch_biosyn	2	10074581	10081512	PGLUCISOM-RXN	transporter family protein, putative, expressed
99	LOC_Os02g32660	starch_biosyn	2	19355748	19367127	RXN-7710	1,4-alpha-glucan branching enzyme
100	LOC_Os02	starch_biosyn	2	31232888	31238210	GLYCOGENSY	transferase,

	g51070					N-RXN	transferring glycosyl groups
101	LOC_Os03g50480	starch_biosyn	3	28813570	28820161	PHOSPHOGLUCMUT-RXN	phosphoglucomutase
102	LOC_Os03g52460	starch_biosyn	3	30099369	30104596	GLUC1PADENYLTRANS-RXN	UTP-glucose-1-phosphate uridylyltransferase
103	LOC_Os03g56460	starch_biosyn	3	32175133	32182101	PGLUCISOM-RXN	glucose-6-phosphate isomerase
104	LOC_Os04g53310	starch_biosyn	4	31750955	31759581	GLYCOGENSYN-RXN	transferase, transferring glycosyl groups
105	LOC_Os05g45720	starch_biosyn	5	26485807	26494112	GLYCOGENSYN-RXN	transferase, transferring glycosyl groups
106	LOC_Os05g50380	starch_biosyn	5	28871794	28877277	GLUC1PADENYLTRANS-RXN	glucose-1-phosphate adenylyltransferase
107	LOC_Os06	starch_biosyn	6	1765622	1770656	GLYCOGENSY	starch

	g04200					N-RXN	synthase, putative, expressed
108	LOC_Os06g06560	starch_biosyn	6	3079059	3086808	GLYCOGENSYN-RXN	transferase, transferring glycosyl groups
109	LOC_Os06g12450	starch_biosyn	6	6748358	6753338	GLYCOGENSYN-RXN	transferase, transferring glycosyl groups
110	LOC_Os06g14510	starch_biosyn	6	8152235	8159385	PGLUCISOM-RXN	glucose-6-phosphate isomerase
111	LOC_Os06g26234	starch_biosyn	6	15335687	15365531	RXN-7710	1,4-alpha-glucan branching enzyme
112	LOC_Os06g28194	starch_biosyn	6	16011323	16032869	PHOSPHOGLUCMUT-RXN	expressed protein
113	LOC_Os06g51084	starch_biosyn	6	30897376	30905815	RXN-7710	1,4-alpha-glucan-branching enzyme, chloroplast precursor,

							putative, expressed
114	LOC_Os07g13980	starch_biosyn	7	7982501	7989283	GLUC1PADENYLTRANS-RXN	glucose-1-phosphate adenylyltransferase large subunit, putative, expressed
115	LOC_Os07g22930	starch_biosyn	7	12916277	12924325	GLYCOGENSYN-RXN	transferase, transferring glycosyl groups
116	LOC_Os07g26610	starch_biosyn	7	15338913	15351164	PHOSPHOGLUCMUT-RXN	phosphoglucomutase
117	LOC_Os08g09230	starch_biosyn	8	5352105	5363367	GLYCOGENSYN-RXN	transferase, transferring glycosyl groups
118	LOC_Os08g25734	starch_biosyn	8	15666201	15672594	GLUC1PADENYLTRANS-RXN	glucose-1-phosphate adenylyltransferase
119	LOC_Os08g37380	starch_biosyn	8	23623080	23628557	PGLUCISOM-RXN	glucose-6-phosphate isomerase

120	LOC_Os08g40930	starch_biosyn	8	25892391	25900576	RXN-7710	1,4-alpha-glucan branching enzyme
121	LOC_Os09g12660	starch_biosyn	9	7245359	7249816	GLUC1PADENYLTRANS-RXN	glucose-1-phosphate adenylyltransferase
122	LOC_Os09g29070	starch_biosyn	9	17648471	17653621	PGLUCISOM-RXN	glucose-6-phosphate isomerase
123	LOC_Os09g29404	starch_biosyn	9	17851430	17868669	RXN-7710	1,4-alpha-glucan branching enzyme
124	LOC_Os10g11140	starch_biosyn	10	6167111	6174631	PHOSPHOGLUCMUT-RXN	phosphoglucomutase
125	LOC_Os10g30156	starch_biosyn	10	15673128	15681124	GLYCOGENSYN-RXN	transferase, transferring glycosyl groups

Supplemental Table 4.4. Narrow-sense heritability of NIR-predicted NSC traits.

trait	**h^2**
Starch, hd	0.25
Sucrose, hd	0.21
TNC, hd	0.26
Starch, mt	0.14
Sucrose, mt	0.21
TNC, mt	0.16

Supplemental Table 4.5. GWAS results from *tropical japonica* panel.

chr	trait	msSNP	pos	baseline r2	LD (kb)	size flanking SNP (kb)	non-TE genes
3	starch, hd	SNP-3.31861332.	31868456	0.05459995	49.84	99.68	15
3		SNP-3.36253278.	36260409	0.03475884	76.59	153.18	20
5		SNP-5.16472032.	16529551	0.0391062	33.85	67.7	6
1	suc, mt and starch, mt	SNP-1.28557570.	28558616	0.04371349	61.21	122.42	24
1	starch, mt	SNP-1.30962744.	30963789	0.02624896	48.99	97.98	13
1		SNP-1.26243779.	26244824	0.0533055	NA	NA	NA
2	TNC, hd	SNP-2.25384431.	25390301	0.04900148	152.43	304.86	39
11		SNP-11.24673979.	25140140	0.04693588	252.87	505.74	52
5	TNC, mt	SNP-5.3775943.	3775967	0.06031838	52.71	105.42	11

Supplemental Table 4.6. Gene models within LD of chr 1 msSNP SNP-1.30962744.

Gene Model	Description
LOC_Os01g53790	stress-induced protein, putative, expressed
LOC_Os01g53800	glutamate carboxypeptidase 2, putative, expressed
LOC_Os01g53810	peptidase, putative, expressed
LOC_Os01g53820	expressed protein
LOC_Os01g53830	transposon protein, putative, CACTA, En/Spm sub-class, expressed
LOC_Os01g53840	protein kinase family protein, putative, expressed
LOC_Os01g53850	hypothetical protein
LOC_Os01g53860	expressed protein
LOC_Os01g53870	transposon protein, putative, unclassified, expressed
LOC_Os01g53880	OsIAA6 - Auxin-responsive Aux/IAA gene family member, expressed
LOC_Os01g53890	RNA methyltransferase, TrmH family protein, putative, expressed
LOC_Os01g53900	elongation factor, putative, expressed
LOC_Os01g53910	dehydrogenase, putative, expressed
LOC_Os01g53920	receptor-like protein kinase 5 precursor, putative, expressed
LOC_Os01g53930	hexokinase, putative, expressed

Supplemental Table 4.7. Gene models within LD of chr 11 msSNP SNP-11.24633058.

Gene Model	Description
LOC_Os11g41500	Note NC domain-containing protein, putative, expressed
LOC_Os11g41510	Note expressed protein
LOC_Os11g41520	Note hypothetical protein
LOC_Os11g41530	Note expressed protein
LOC_Os11g41540	disease resistance RPP8-like protein 3, putative, expressed
LOC_Os11g41550	expressed protein
LOC_Os11g41560	OsFBX427 - F-box domain containing protein, expressed
LOC_Os11g41570	OsFBX428 - F-box domain containing protein, expressed
LOC_Os11g41580	transposon protein, putative, unclassified, expressed
LOC_Os11g41590	expressed protein
LOC_Os11g41600	expressed protein
LOC_Os11g41610	40S ribosomal protein S29, putative, expressed
LOC_Os11g41620	transposon protein, putative, unclassified, expressed
LOC_Os11g41630	transposon protein, putative, CACTA, En/Spm sub-class, expressed
LOC_Os11g41640	helix-loop-helix DNA-binding domain containing protein, expressed
LOC_Os11g41650	adenylyl-sulfate kinase, putative, expressed
LOC_Os11g41670	expressed protein
LOC_Os11g41680	cytochrome P450, putative, expressed
LOC_Os11g41690	retrotransposon protein, putative, unclassified, expressed
LOC_Os11g41700	transposon protein, putative, unclassified, expressed

LOC_Os11g41710	cytochrome P450, putative, expressed
LOC_Os11g41720	TNP1, putative, expressed
LOC_Os11g41730	transposon protein, putative, CACTA, En/Spm sub-class, expressed
LOC_Os11g41740	retrotransposon protein, putative, unclassified
LOC_Os11g41750	retrotransposon protein, putative, Ty3-gypsy subclass, expressed
LOC_Os11g41760	retrotransposon protein, putative, Ty3-gypsy subclass, expressed
LOC_Os11g41770	expressed protein
LOC_Os11g41780	transposon protein, putative, CACTA, En/Spm sub-class, expressed
LOC_Os11g41800	hypothetical protein
LOC_Os11g41820	RNA recognition motif containing protein, expressed
LOC_Os11g41830	proteophosphoglycan ppg4, putative, expressed
LOC_Os11g41840	transporter family protein, putative, expressed
LOC_Os11g41850	transporter family protein, putative, expressed
LOC_Os11g41860	OsFBX429 - F-box domain containing protein, expressed
LOC_Os11g41870	transporter family protein, putative, expressed
LOC_Os11g41875	membrane protein, putative, expressed
LOC_Os11g41880	potyvirus VPg interacting protein, putative, expressed
LOC_Os11g41890	RNA recognition motif containing protein, putative, expressed
LOC_Os11g41900	acyltransferase, putative, expressed
LOC_Os11g41910	GTP-binding protein, putative, expressed

LOC_Os11g41920	expressed protein
LOC_Os11g41940	expressed protein
LOC_Os11g41950	expressed protein
LOC_Os11g41960	hypothetical protein
LOC_Os11g41970	hypothetical protein
LOC_Os11g41990	armadillo/beta-catenin-like repeat family protein, expressed
LOC_Os11g42000	FRA10AC1, putative, expressed
LOC_Os11g42010	NBS-LRR type disease resistance protein, putative, expressed
LOC_Os11g42020	expressed protein
LOC_Os11g42030	expressed protein
LOC_Os11g42040	Note non-TIR-NBS-LRR type resistance protein, putative, expressed
LOC_Os11g42050	Note expressed protein
LOC_Os11g42060	Note Leucine Rich Repeat family protein, expressed
LOC_Os11g42070	Note Leucine Rich Repeat family protein, expressed
LOC_Os11g42080	Note expressed protein
LOC_Os11g42090	Note Leucine Rich Repeat family protein, expressed
LOC_Os11g42100	Note Leucine Rich Repeat family protein, expressed
LOC_Os11g42110	Note expressed protein
LOC_Os11g42120	Note expressed protein
LOC_Os11g42130	Note transposon protein, putative, unclassified
LOC_Os11g42140	Note transposon protein, putative, unclassified, expressed
LOC_Os11g42150	Note hypothetical protein

Supplemental Table 4.8. SNP frequencies across RDP-1 and RDP-2 diverse *O. sativa* (McCouch et al., 2016).

SNP-1.30962744.

subpop	admixed	*temperate-japonica*	*aromatic*	*indica*	*tropical-japonica*	admixed-japonica	*aus*	admixed-indica
AA	0.88	0.92	0.64	0.71	0.93	0.96	0.87	0.81
BB	0.12	0.08	0.36	0.29	0.07	0.04	0.13	0.19

SNP-11.24633058.

subpop	admixed	*temperate-japonica*	*aromatic*	*indica*	*tropical-japonica*	admixed-japonica	*aus*	admixed-indica
AA	0.88	0.99	0.89	0.79	0.90	0.94	0.71	0.97
BA	0.06	0.01	0.07	0.03	0.00	0.00	0.06	0.00
BB	0.06	0.00	0.04	0.18	0.10	0.06	0.23	0.03

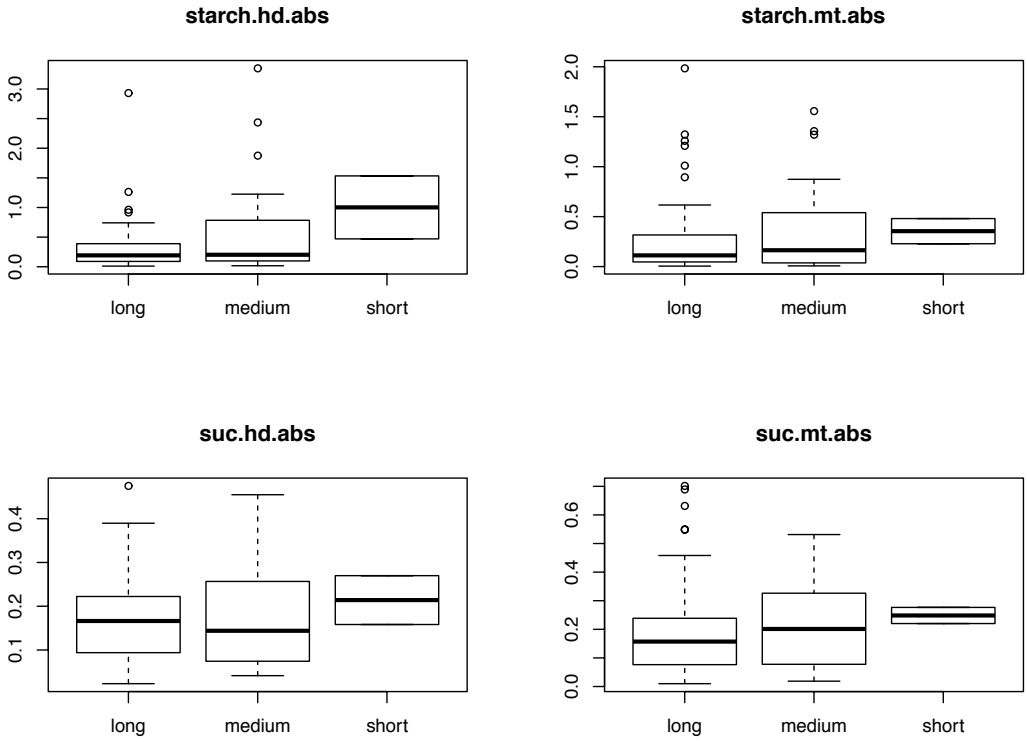
Supplemental Table 4.9. Output of IciMapping of Curinga x O. rufipogon CSSL population for NSC traits

Trait Name	Marker Name	LOD	PVE(%)	Add
Sucrose, heading	c5_21751532	5.431	36.2838	0.377
Sucrose, maturity	c5_19516545	4.5664	30.9295	0.405

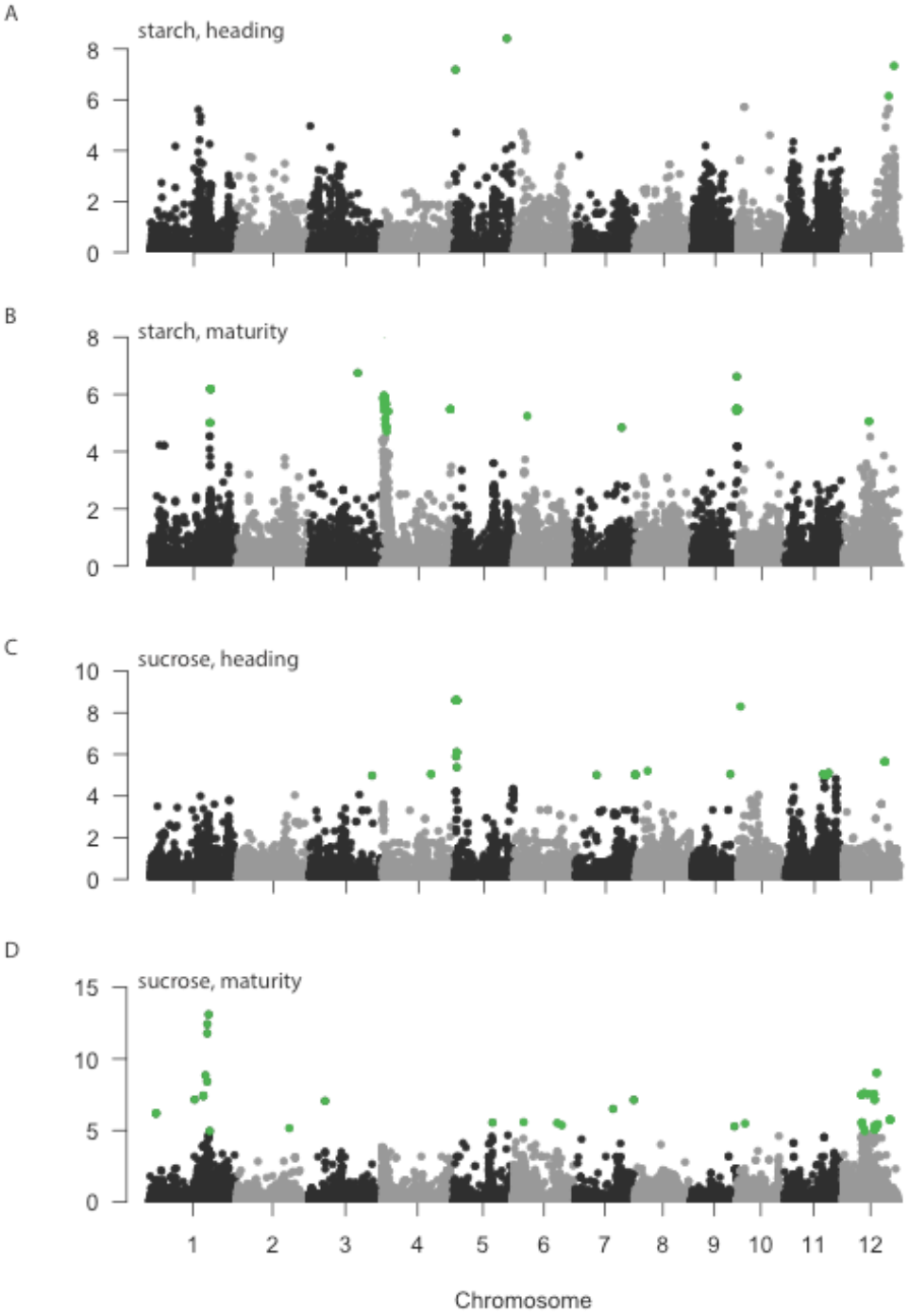
Supplemental Table 4.10. Allele counts of SNP S5_19425787 across diverse *O. sativa* subpopulation control samples.

Varietal Group	Allele G	Allele A
Japonica	18	3
Indica	1	7

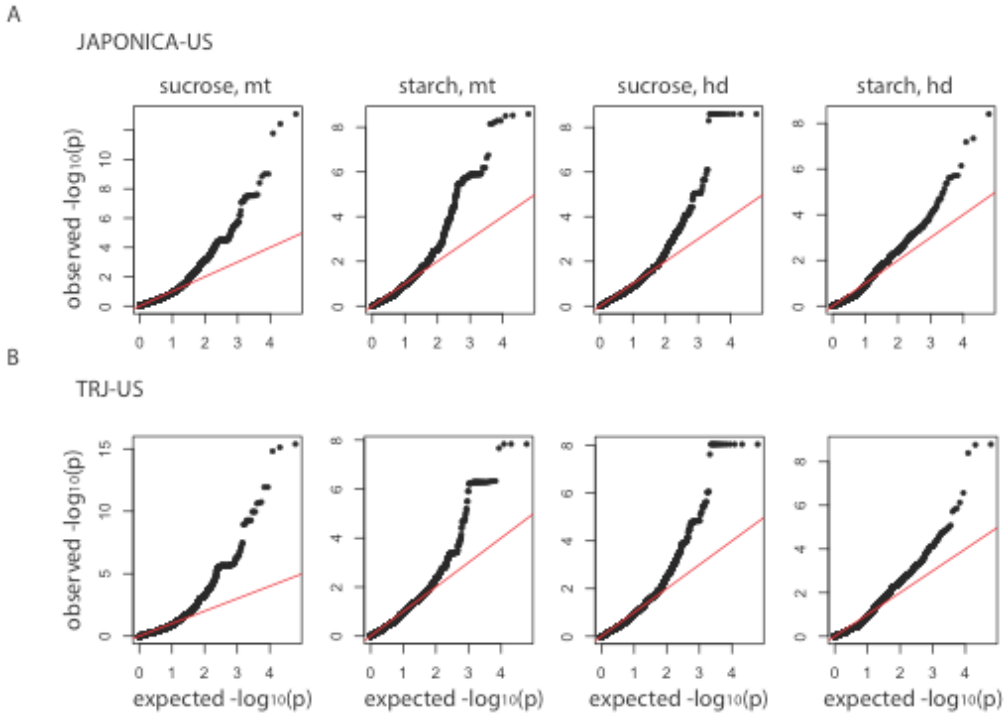
Supplemental Figure 4.1. Stem NSC distributions across market classes of U.S. rice. Long (n=71), medium (n=23), and short (n=2) grain length classes.



Supplemental Figure 4.2. GWAS results for NSC traits in the *US-JAPONICA* rice panel. Green points represent markers that passed a 1% FDR threshold.

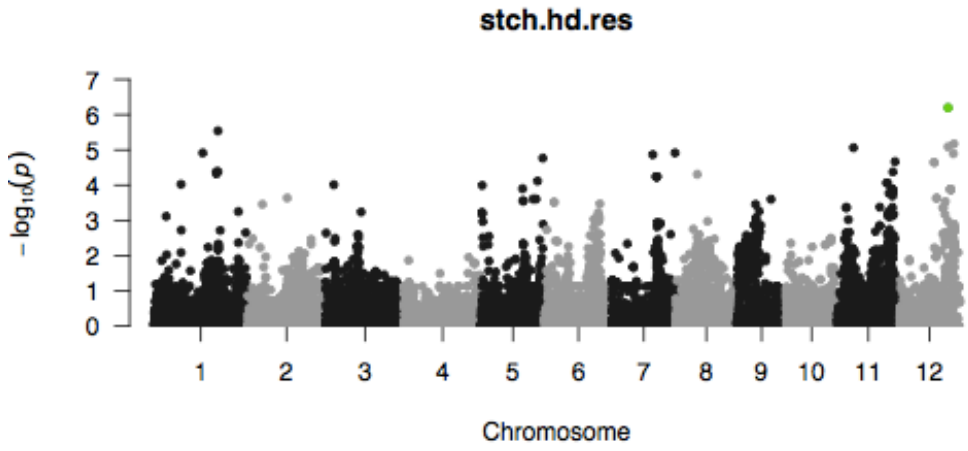


Supplemental Figure 4.3. QQ plots from stem NSC GWAS on *US-JAPONICA* and *US-TRJ* panels.

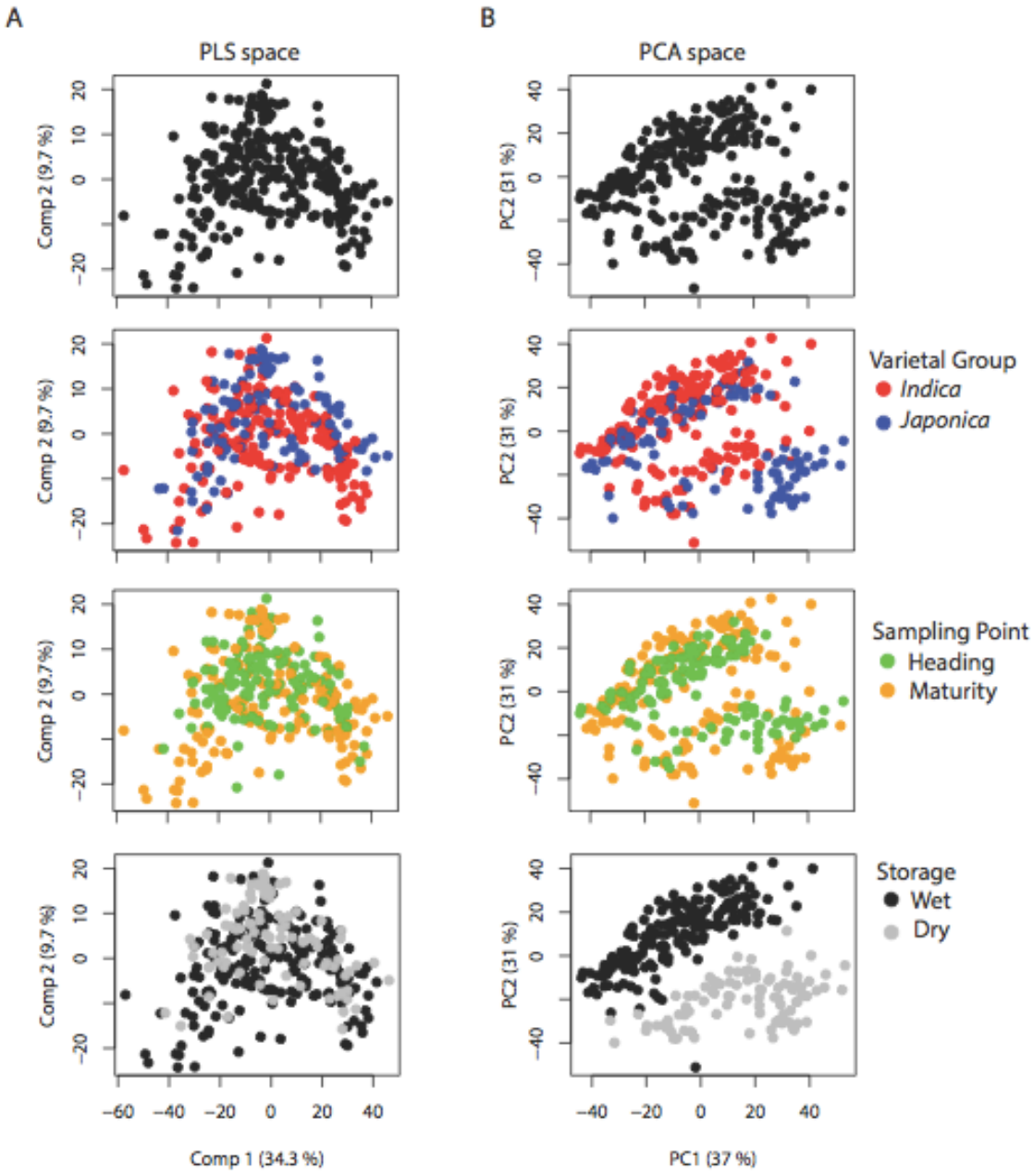


Supplemental Figure 4.4. Investigation into the effect of days to heading on stem NSC GWAS results. Days to heading (DTH) was incorporated as a covariate in GWAS on the TRJ-US panel for stem starch and sucrose at heading and maturity. Upper table: Correlations of resultant marker p-values from genome scans with and without DTH covariate are displayed below. Results with and without DTH covariate agreed well for NSC traits except starch at heading, which was the most affected by the DTH covariate. Lower Figure: manhattan plot for GWAS results of starch-at-heading with DTH covariate (0.05 FDR).

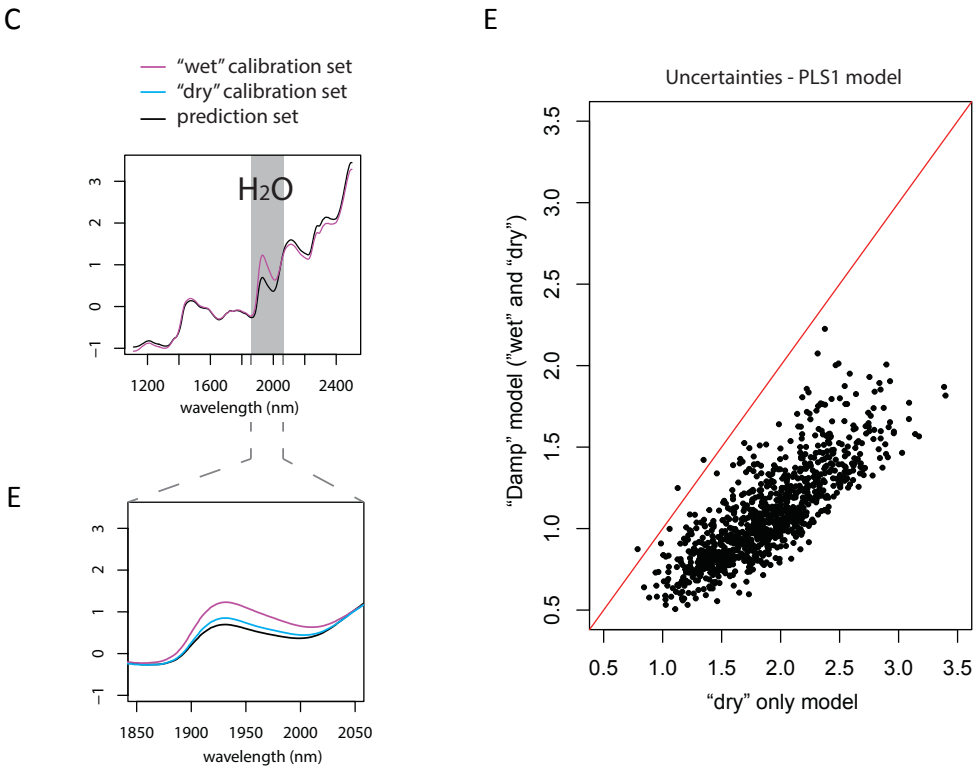
Trait	Pearson's correlation coefficient
Starch, hd	0.77
Starch, mt	0.90
Sucrose, hd	0.93
Sucrose, mt	0.99



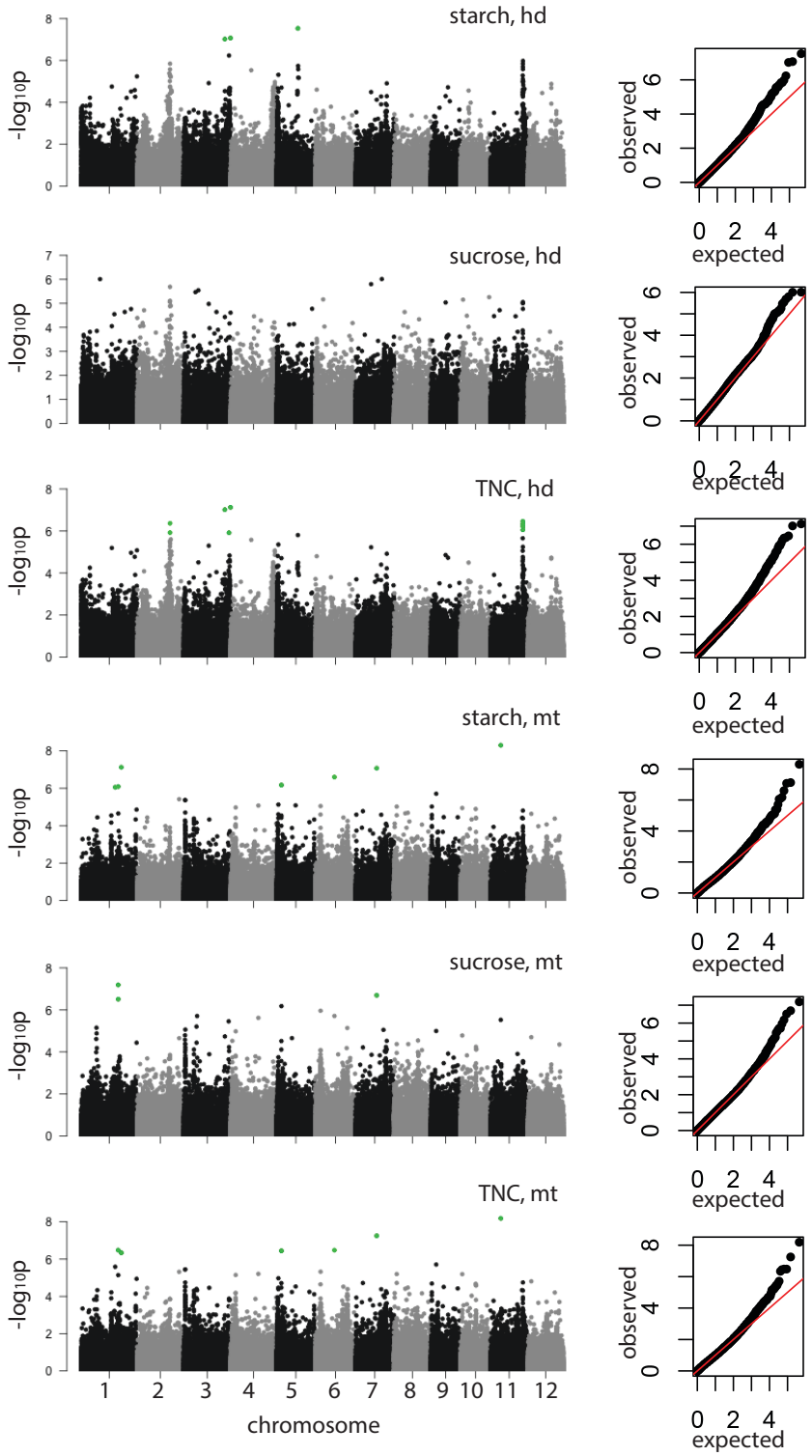
Supplemental Figure 4.5. 283 samples used for the NIR calibration model colored by various metadata. A-B) Varietal group, sampling point, and storage condition are annotated for spectral projections onto PCA space and PLS model space. Storage condition "wet" refers to samples that underwent one round of drying after sampling and before grinding and subsequently stored under ambient conditions in a laboratory in the NE U.S. (i.e. humid summers) while "dry" samples underwent a second round of drying post-grinding and were stored in laboratory conditions in the arid Western U.S. All samples were scanned together on the same spectrophotometer. Major structure in PCA space of the calibration set was attributed to storage condition, however this structure was not seen when projected into the first two components of PLS space. C) Differences in storage condition was due primarily with differences in the spectrum associated with water. D) The prediction set fell between the "dry" and "wet" calibration sets in the range of 1850-2050nm associated with water and was more closely aligned with the "dry" set. E) Despite more closer relationship to the "dry" set, the prediction set was better predicted using a calibration set of combined wet and dry than dry samples alone.


A
PLS space
B
PCA space
Comp 2 (9.7 %)
PC2 (31 %)
Comp 1 (34.3 %)
PC1 (37 %)
Varietal Group
Indica
Japonica
Sampling Point
Heading
Maturity
Storage
Wet
Dry

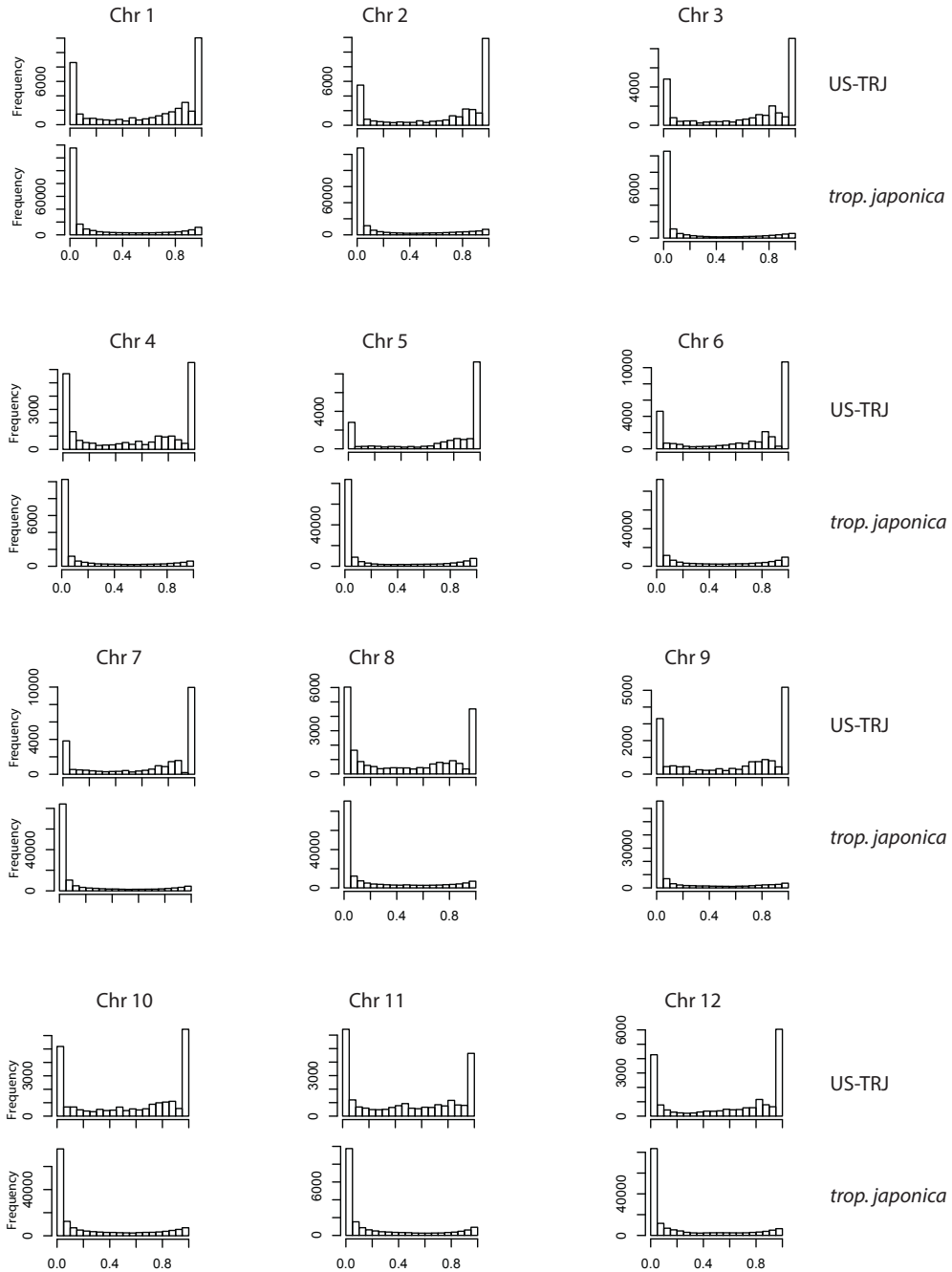
Supplemental Figure 4.5 continued



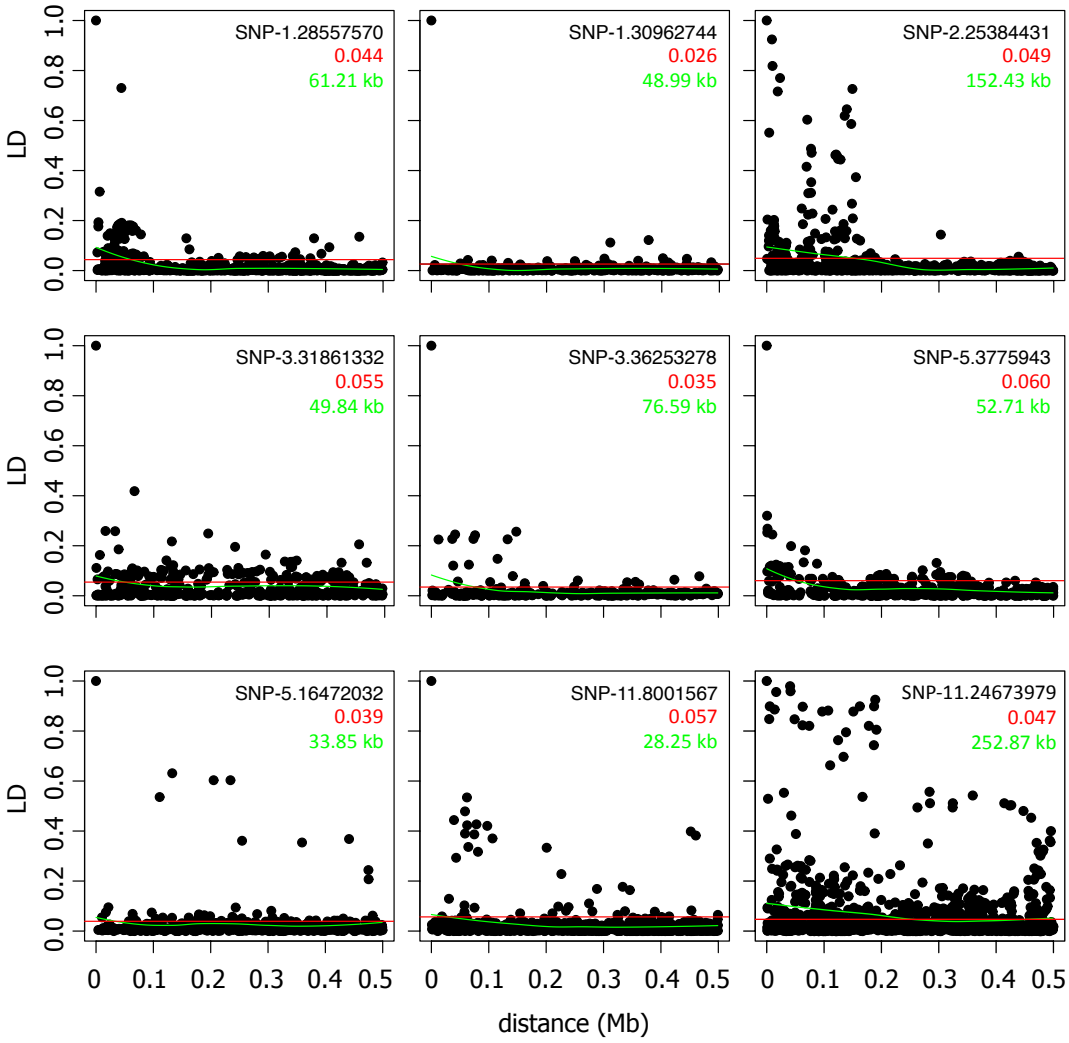
Supplemental Figure 4.6. GWAS results for *tropical japonica* diversity panel.



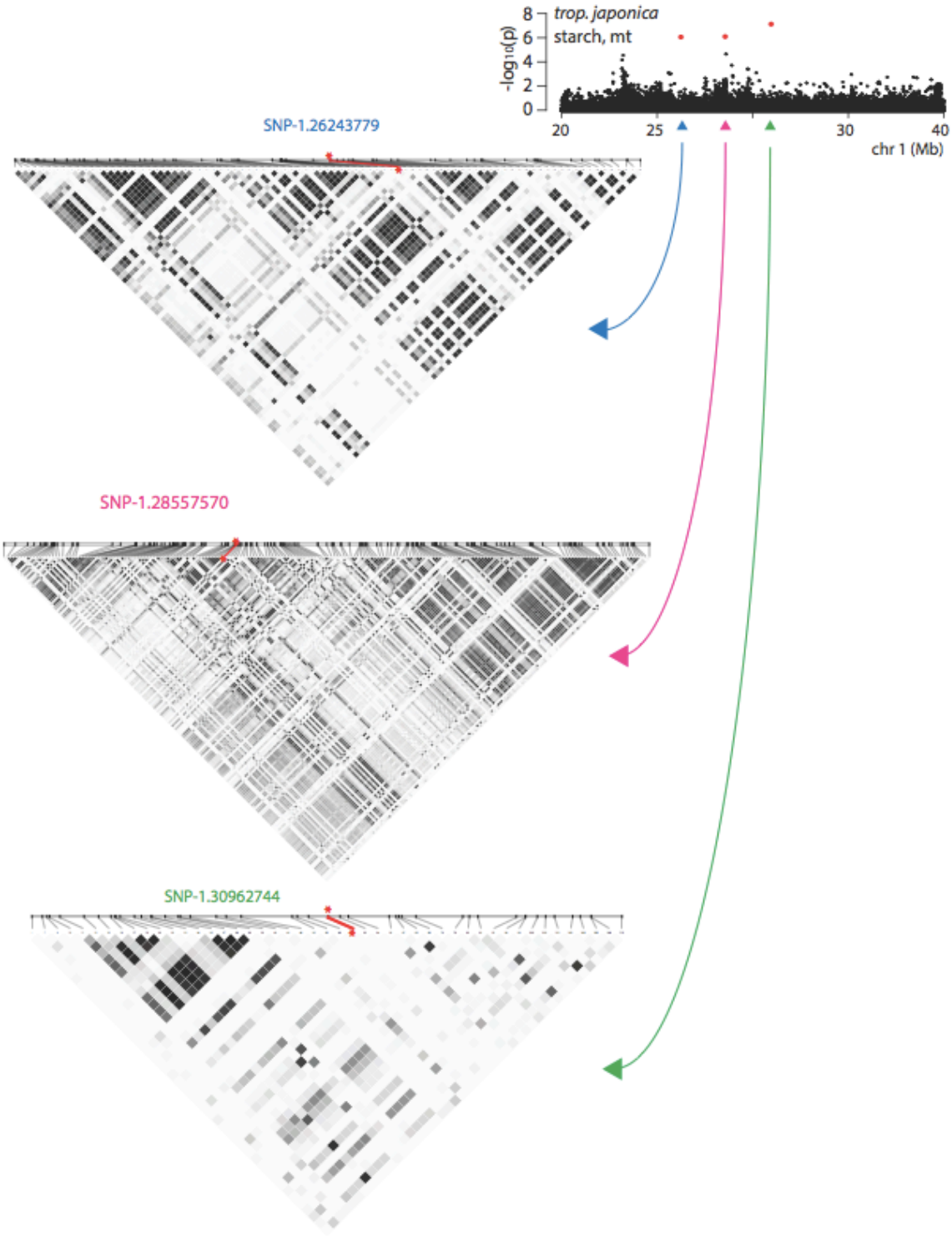
Supplemental Figure 4.7. LD of US-TRJ versus diverse *tropical japonica.* Pairwise LD across the rice 12 chromosomes was calculated for all SNP pairs with distances 500bp-1Mb in U.S. rice and *tropical japonica* panels. U.S. rice shows much higher frequency of high LD SNP pairs.



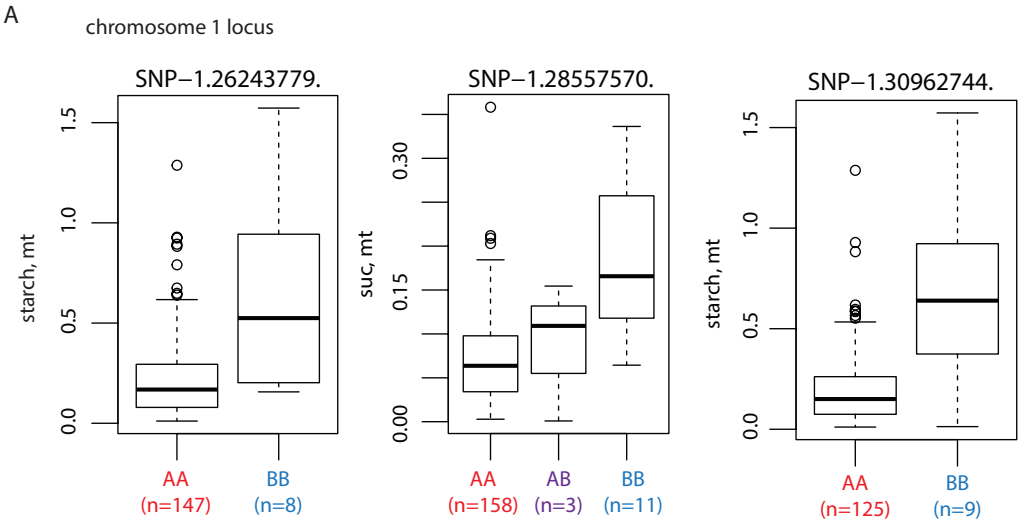
Supplemental Figure 4.8. Local LD decay around msSNPs of association peaks. SNP-specific local LD. See Methods.



Supplemental Figure 4.9. Local LD heatmaps of significant regions in the chromosome 1 locus. LD (r^2) is displayed (1 = black, 0 = white, 0-1 = shades of gray, per Haploview visualization).



Supplemental Figure 4.10. Phenotypic distributions grouped by genotype class for associated traits of two significant peaks found from both U.S. and *tropical japonica* GWAS. AA indicates the major allele and BB indicates the minor allele at each marker. Names of accession observed to carry the minor BB allele or that are heterozygous (AB) are listed below each plot in A and adjacent to the plot in B.

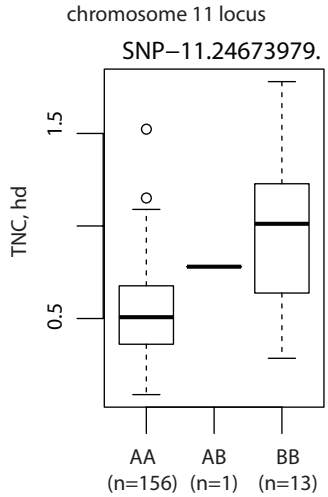


BB:
Kendinga 5 H
Vialone
Padi Kasalle
Sinaguing
Carolina Gold Sel.
Cybonnet
Ostiglia
NSFTV362

BB:
Yancaoussa
D 4-136
P. Tingagew Dayket Qay Dayon
Davao
Immarang
Moroberekan
Sinaguing
Tondok
Iguape Cateto
Pate Blanc Mn 1
Fossa Av

AB:
Ranau Kadai
IRAT 366
325

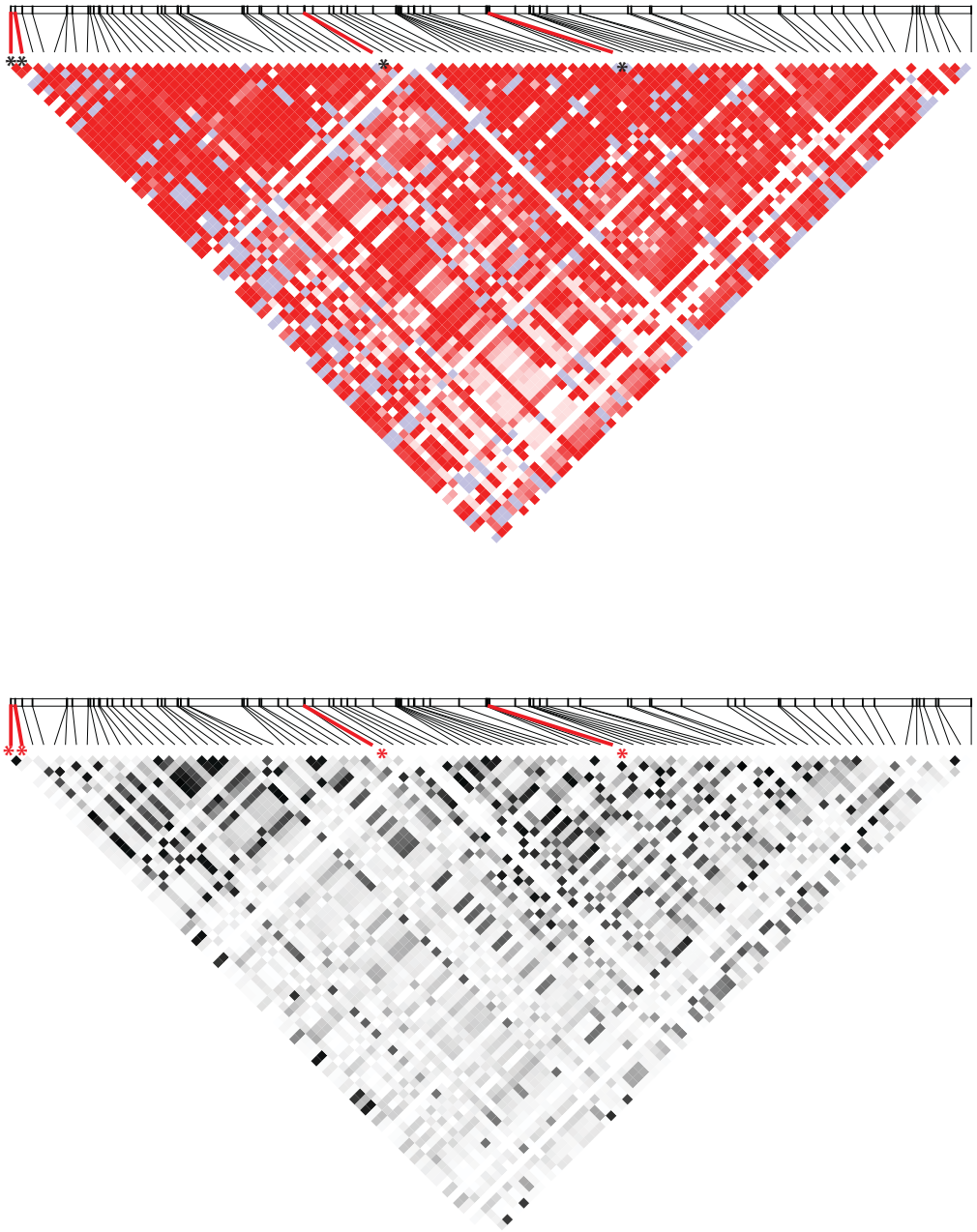
BB:
IRAT 366
IRAT 364
D 4-136
Blue Rose Supreme
Kinastano
Padi Kassalle
Sinaguing
Canella De Ferro
583



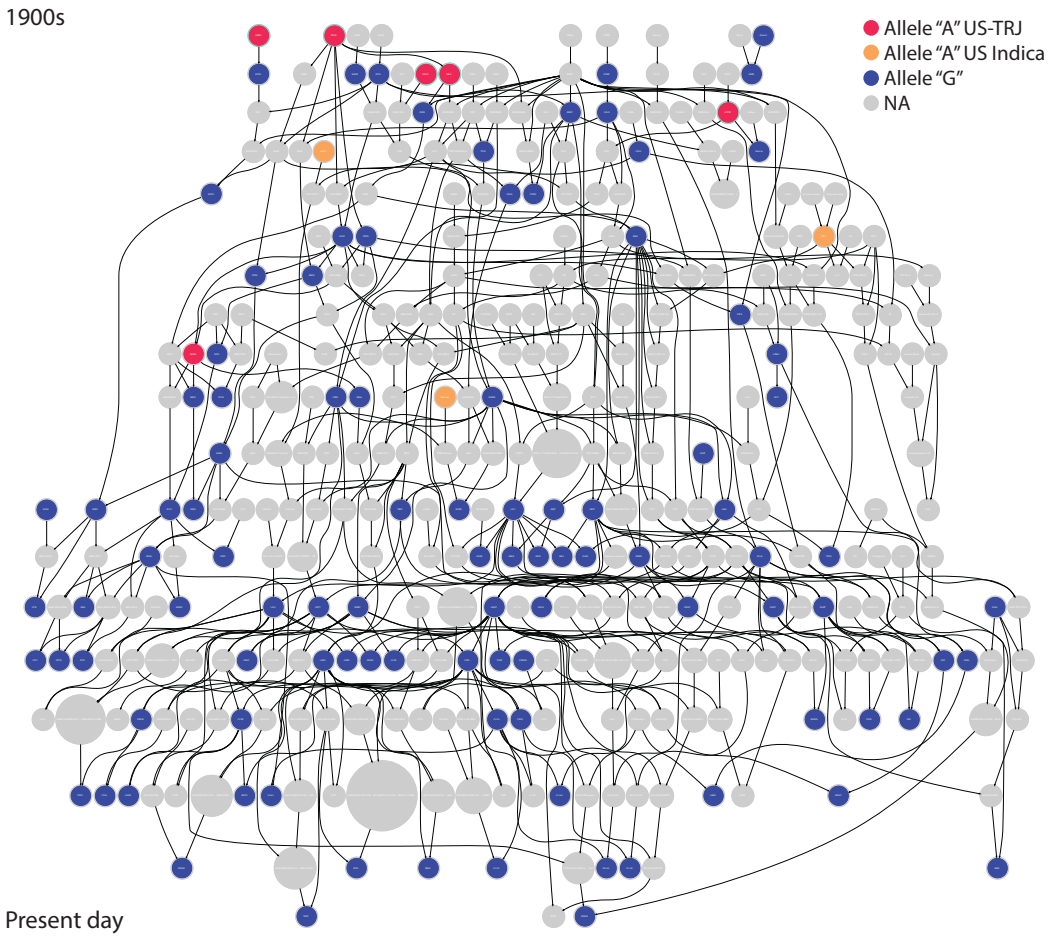
BB:
Peek
Padi Teras
Ilang-Ilang
Primavera
Palawan
Kendinga
Bluebonnet
CI 11011
Sinaguing
Tondok
C57-5043
Fortuna
583

AB:
D 4-136

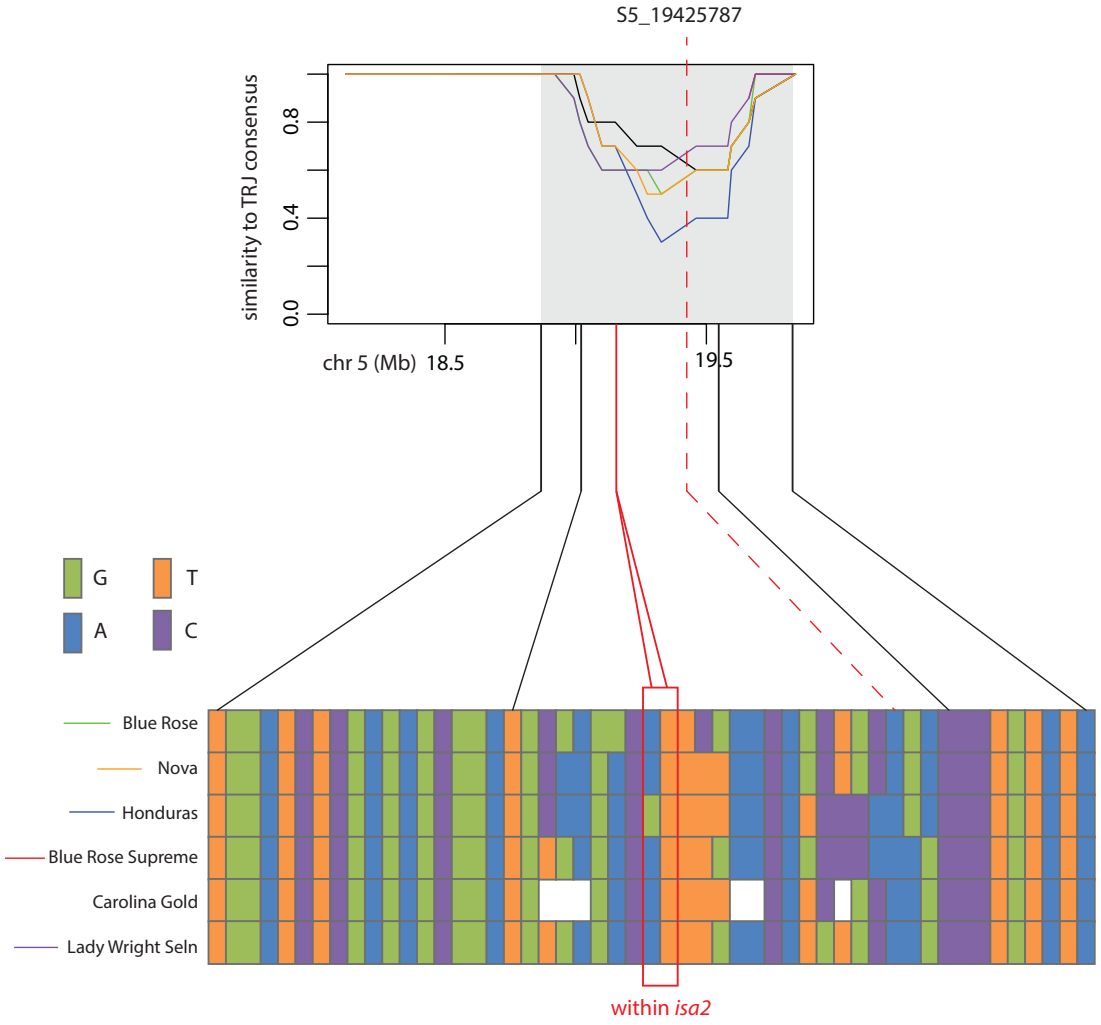
Supplemental Figure 4.11. LD plot of chr 11 of a region associated with starch-at-heading. Plot shows the region from25098867 bp to 25176821 bp. Red lines indicate positions of four significant SNPs (see Fig. 4). Top panel displays D' statistic and bottom panel displays r^2 values.



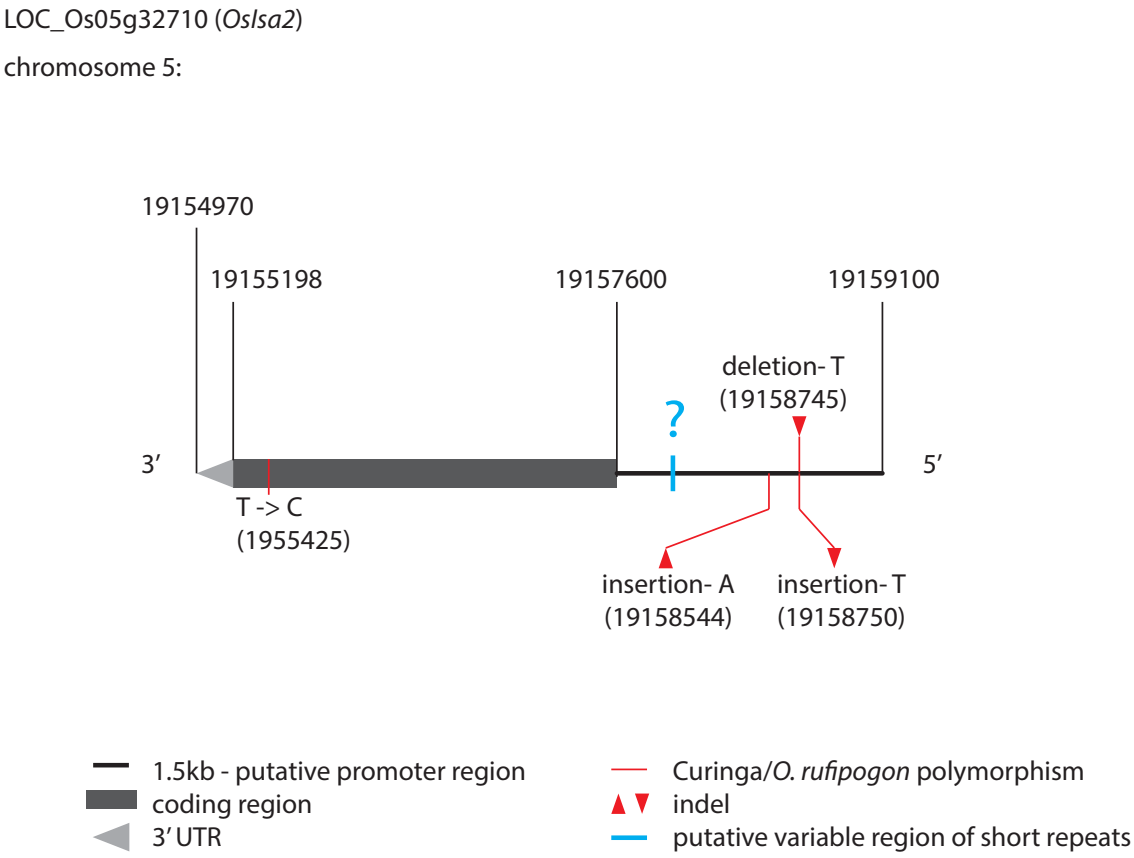
Supplemental Figure 4.12. U.S. rice pedigree colored by S5_19425787 genotype.



Supplemental Figure 4.13. Haplotype analysis of a region on chromosome five associated with sucrose-at-maturity in U.S. rice. Six accessions harbored the minor allele at msSNP, S5_19425787, that was associated with increased sucrose-at-maturity levels. All six accessions were varieties of the *tropical japonica-1* subpopulation according to FastStructure analysis. Top: Markers from 18-20Mb were extracted for these varieties along with 27 *indica* accessions and 30 *trj*-1 accessions. A consensus sequence was determined for *indica* and *trj*-1 and filtered to keep only markers that distinguished the two sequences. Next, similarity (defined as number of shared alleles/total markers) of each of these accessions with the trj-1 consensus sequence was calculated using windows of 10 SNPs, with step size 2 variants. Carolina Gold had missing data and was omitted for this analysis. Bottom: haplotypes of these six accessions.



Supplemental Figure 4.14. Sequence polymorphisms between cv. Curinga and O. rufipogon IRGC 105491 from re-sequencing data for LOC_Os05g32710, isoamylase. Target region includes gene (chr 5: 19157600 - 19154970 bp) and 1.5kb upstream putative promoter (chr 5: 19159100 - 19157600 bp). The blue annotation marks a region of variable alignment against the reference genome, Nipponbare, when analyzing re-sequencing information from Curinga and 18 other *tropical japonica* accessions. This type of result is suggestive of a variable region due to repetitive elements, such as Simple Sequence Repeats (SSRs).



REFERENCES

Arai-Sanoh Y, Ida M, Zhao R, Yoshinaga S, Takai T, Ishimaru T, Maeda H, Nishitani K, Terashima Y, Gau M, et al (2011) Genotypic variations in non-structural carbohydrate and cell-wall components of the stem in rice, sorghum, and sugar vane. Biosci Biotechnol Biochem **75**: 1104–1112

Arbelaez JD, Moreno LT, Singh N, Tung C-W, Maron LG, Ospina Y, Martinez CP, Grenier C, Lorieux M, McCouch S (2015) Development and GBS-genotyping of introgression lines (ILs) using two wild species of rice, O. meridionalis and O. rufipogon, in a common recurrent parent, O. sativa cv. Curinga. Mol Breed **35**: 1–18

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B 289–300

Blakeney AB, Matheson NK (1984) Some Properties of the Stem and Pollen Starches of Rice. Starch - Stärke **36**: 265–269

Breseghello F, Sorrells ME (2006) Association mapping of kernel size and milling quality in wheat (Triticum aestivum L.) cultivars. Genetics **172**: 1165–1177

Chang CC, Chow CC, Tellier LCAM, Vattikuti S, Purcell SM, Lee JJ (2015) Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience **4**: 1

Chen H-J, Wang S-J (2008) Molecular regulation of sinksource transition in rice leaf sheaths during the heading period. Acta Physiol Plant **30**: 639–649

Cho J-I, Ryoo N, Eom J-S, Lee D-W, Kim H-B, Jeong S-W, Lee Y-H, Kwon Y-K,

Cho M-H, Bhoo SH, et al (2009) Role of the Rice Hexokinases OsHXK5 and OsHXK6 as Glucose Sensors. Plant Physiol **149**: 745–759

Conocono EA, Egdane JA, Setter TL (1995) Estimation of Canopy Photosynthesis in Rice by Means of Daily Increases in Leaf Carbohydrate Concentrations. Crop Sci **38**: 987–995

Cook FR, Fahy B, Trafford K (2012) A rice mutant lacking a large subunit of ADP-glucose pyrophosphorylase has drastically reduced starch content in the culm but normal plant morphology and yield. Funct Plant Biol **39**: 1068–1078

Van Dat T, Peterson ML (1983) Performance of Near-Isogenic Genotypes of Rice Differing in Growth Duration. II. Carbohydrate Partitioning During Grain Filling. Crop Sci. 23:

Dharmawardhana P, Ren L, Amarasinghe V, Monaco M, Thomason J, Ravenscroft D, McCouch S, Ware D, Jaiswal P (2013) A genome scale metabolic network for rice and accompanying analysis of tryptophan, auxin and serotonin biosynthesis regulation under biotic stress. Rice **6**: 1–15

Dilday RH (1990) Contribution of ancestral lines in the development of new cultivars of rice. Crop Sci **30**: 905–911

Duitama J, Silva A, Sanabria Y, Cruz DF, Quintero C, Ballen C, Lorieux M, Scheffler B, Farmer A, Torres E (2015) Whole genome sequencing of elite rice cultivars as a comprehensive information resource for marker assisted selection. PLoS One **10**: e0124617

Fu J, Huang Z, Wang Z, Yang J, Zhang J (2011) Pre-anthesis non-structural carbohydrate reserve in the stem enhances the sink strength of inferior spikelets

during grain filling of rice. F Crop Res **123**: 170–182

Granot D (2008) Putting plant hexokinases in their proper place. Phytochemistry **69**: 2649–2654

He HY, Koike M, Ishimaru T, Ohsugi R, Yamagishi T (2005) Temporal and spatial variations of carbohydrate content in rice leaf sheath and their varietal differences. Crop Science Society of Japan, Tokyo, JAPON

Hirano T, Saito Y, Ushimaru H, Michiyama H (2005) The Effect of the Amount of Nitrogen Fertilizer on Starch Metabolism in Leaf Sheath of Japonica and Indica Rice Varieties during the Heading Period. Plant Prod Sci **8**: 122–130

Hirose T, Endler A, Ohsugi R (1999) Gene Expression of Enzymes for Starch and Sucrose Metabolism and Transport in Leaf Sheaths of Rice (Oryza sativa L.) during the Heading Period in Relation to the Sink to Source Transition. Plant Prod. Sci. 2:

Hirose T, Ohdan T, Nakamura Y, Terao T (2006) Expression profiling of genes related to starch synthesis in rice leaf sheaths during the heading period. Physiol Plant **128**: 425–435

Ishimaru K, Kosone M, Sasaki H, Kashiwagi T (2004) Leaf contents differ depending on the position in a rice leaf sheath during sink-source transition. Plant Physiol Biochem **42**: 855–860

Jang J-C, León P, Zhou L, Sheen J (1997) Hexokinase as a sugar sensor in higher plants. Plant Cell **9**: 5–19

Johnson DA, Thomas MA (2007) The Monosaccharide Transporter Gene Family in Arabidopsis and Rice: A History of Duplications, Adaptive Evolution, and

Functional Divergence. Mol Biol Evol **24** : 2412–2423

Kanbe T, Sasaki H, Aoki N, Yamagishi T, Ohsugi R (2009) The QTL Analysis of RuBisCO in Flag Leaves and Non-Structural Carbohydrates in Leaf Sheaths of Rice Using Chromosome Segment Substitution Lines and Backcross Progeny F2 Populations. Plant Prod Sci **12**: 224–232

Kashiwagi T, Madoka Y, Hirotsu N, Ishimaru K (2006) Locus prl5 improves lodging resistance of rice by delaying senescence and increasing carbohydrate reaccumulation. Plant Physiol Biochem **44**: 152–157

Kharabian-Masouleh A, Waters DLE, Reinke RF, Henry RJ (2011) Discovery of polymorphisms in starch-related genes in rice germplasm by amplification of pooled DNA and deeply parallel sequencing†. Plant Biotechnol J **9**: 1074–1085

Kim J, Shon J, Lee C-K, Yang W, Yoon Y, Yang W-H, Kim Y-G, Lee B-W (2011) Relationship between grain filling duration and leaf senescence of temperate rice under high temperature. F Crop Res **122**: 207–213

Laido G, Marone D, Russo MA, Colecchia SA, Mastrangelo AM, De Vita P, Papa R (2014) Linkage disequilibrium and genome-wide association mapping in tetraploid wheat (Triticum turgidum L.). PLoS One **9**: e95211

Lu H, Redus MA, Coburn JR, Rutger JN, McCOUCH SR, Tai TH (2005) Population structure and breeding patterns of 145 US rice cultivars based on SSR marker analysis. Crop Sci **45**: 66–76

McCouch SR, Wright MH, Tung C-W, Maron LG, McNally KL, Fitzgerald M, Singh N, DeClerck G, Agosto-Perez F, Korniliev P, et al (2016) Open access resources for genome-wide association mapping in rice. Nat Commun 7:

Morita S, Nakano H (2011) Nonstructural Carbohydrate Content in the Stem at Full Heading Contributes to High Performance of Ripening in Heat-Tolerant Rice Cultivar Nikomaru. Crop Sci **51**: 818–828

Nagata K, Shimizu H, Terao T (2002) Quantitative Trait Loci for Nonstructural Carbohydrate Accumulation in Leaf Sheaths and Culms of Rice (Oryza sativa L.) and their Effects on Grain Filling. Breed Sci **52**: 275–283

Pan J, Cui K, Wei D, Huang J, Xiang J, Nie L (2011) Relationships of non-structural carbohydrates accumulation and translocation with yield formation in rice recombinant inbred lines under two nitrogen levels. Physiol Plant **141**: 321–331

Perez CM, Palmiano EP, Baun LC, Juliano BO (1971) Starch Metabolism in the Leaf Sheaths and Culm of Rice. Plant Physiol. 47:

Samonte SOPB, Wilson LT, MCCLUNG AM, Tarpley L (2001) Seasonal Dynamics of Nonstructural Carbohydrate Partitioning in 15 Diverse Rice Genotypes. Crop Sci **41**: 902–909

Slewinski TL (2012) Non-structural carbohydrate partitioning in grass stems: a target to increase yield stability, stress tolerance, and biofuel production. J Exp Bot **63**: 4647–4670

Slewinski TL (2011) Diverse functional roles of monosaccharide transporters and their homologs in vascular plants: a physiological perspective. Mol Plant **4**: 641–662

Spindel J, Wright M, Chen C, Cobb J, Gage J, Harrington S, Lorieux M, Ahmadi N, McCouch S (2013) Bridging the genotyping gap: using genotyping by

sequencing (GBS) to add high-density SNP markers and new value to traditional bi-parental mapping and breeding populations. Theor Appl Genet **126**: 2699–2716

Streb S, Zeeman SC (2014) Replacement of the endogenous starch debranching enzymes ISA1 and ISA2 of Arabidopsis with the rice orthologs reveals a degree of functional conservation during starch synthesis. PLoS One **9**: e92174

Tetlow IJ, Morell MK, Emes MJ (2004) Recent developments in understanding the regulation of starch metabolism in higher plants. J Exp Bot **55**: 2131–2145

Utsumi Y, Utsumi C, Sawada T, Fujita N, Nakamura Y (2011) Functional diversity of isoamylase oligomers: the ISA1 homo-oligomer is essential for amylopectin biosynthesis in rice endosperm. Plant Physiol **156**: 61–77

Wang DR, Wolfrum EJ, Virk P, Ismail A, Greenberg AJ, McCouch SR (2016) Robust phenotyping strategies for evaluation of stem non-structural carbohydrates (NSC) in rice. J Ex Bot (in press)

Watanabe Y, Nakamura Y, Ishii R (1997) Relationship between Starch Accumulation and Activities of the Related Enzymes in the Leaf Sheath as a Temporary Sink Organ in Rice Oryza sativa. Funct Plant Biol **24**: 563–569

Yang J, Zhang J, Wang Z, Zhu Q, Liu L (2002) Abscisic acid and cytokinins in the root exudates and leaves and their relationship to senescence and remobilization of carbon reserves in rice subjected to water stress during grain filling. Planta **215**: 645–652

Yang J, Zhang J, Wang Z, Zhu Q, Wang W (2001) Remobilization of carbon reserves in response to water deficit during grain filling of rice. F Crop Res **71**:

47–55

Zhao M, Acuna TLB, Lafitte HR, Dimayuga G, Sacks E (2008) Perennial hybrids of Oryza sativa/Oryza rufipogon: Part II. Carbon exchange and assimilate partitioning. F Crop Res **106**: 214–223

CHAPTER 5:
EVIDENCE FOR DIVERGENCE OF RESPONSE IN *INDICA*, *JAPONICA*, AND WILD RICE TO HIGH CO_2 X TEMPERATURE INTERACTION

ABSTRACT

High CO_2 and high temperature have an antagonistic interaction effect on rice yield potential and present a unique challenge to adapting rice to projected future climates. Understanding how the differences in response to these two abiotic variables are partitioned across rice germplasm accessions may be key to identifying potentially useful sources of resilient alleles for adapting rice to climate change. In this study, we evaluated eleven globally diverse rice accessions under controlled conditions at two carbon dioxide concentrations (400 and 600 ppm) and four temperature environments (29^oC day/21^oC night; 29^oC day/21^oC night with additional heat stress at anthesis; 34^oC day/26^oC night; and 34^oC day/26^oC night with additional heat stress at anthesis) for a suite of traits including five yield components, five growth characteristics, one phenological trait, and four photosynthesis-related measurements. Multivariate analyses of mean trait data from these eight treatments divide our rice panel into two primary groups consistent with the genetic classification of INDICA/INDICA-like and JAPONICA populations. Overall, we find that the productivity of plants grown under elevated [CO_2] was more sensitive (negative response) to high temperature stress compared with that of plants grown under ambient [CO_2] across this diversity panel. We report differential response to CO_2 x temperature interaction for INDICA/INDICA-like and JAPONICA rice accessions and find preliminary evidence

for the beneficial introduction of exotic alleles into cultivated rice genomic background. Overall, these results support the idea of using wild or currently unadapted gene pools in rice to enhance breeding efforts to secure future climate change adaptation.

INTRODUCTION

The recent rise in atmospheric CO_2 concentration (~23% since 1970) has been proposed as an environmental resource that can be exploited to improve crop performance by converting additional CO_2 into harvestable yield for C_3 crops such as cultivated Asian rice, *Oryza sativa* (Shimono et al., 2009; Ziska et al., 2012b; Hasegawa et al., 2013). For agricultural crops, increasing carbon dioxide levels have been shown to stimulate both vegetative growth and seed yield (Amthor, 2001; Kimball et al., 2002; Poorter and Navas, 2003). In rice, significant [CO_2] x cultivar interactions for growth and yield traits have been reported in a number of field and chamber studies (Moya et al., 1998; Hasegawa et al., 2013; Shimono et al., 2014; Wang et al., 2015), underscoring the potential of genetic improvement to target photoassimilate allocation towards seed yield in response to increased [CO_2].

Despite the observed positive effects elevated CO_2 per se may have on rice yield, research has also shown that other climatic variables, such as the projected rise in global temperature, can negate these benefits. Past studies examining the effects of [CO_2] x temperature interaction on rice reproductive potential provide evidence that increasing temperature can negate [CO_2] stimulatory effects on seed yield and that elevated CO_2 levels can exacerbate yield losses relative to ambient conditions as

temperatures rise (Moya et al., 1998; Baker, 2004; Madan et al., 2012; Ziska et al., 2014). Temperature increases, expressed either as mean seasonal temperature or temperature extremes, may be particularly consequential for reproductive physiology in rice (Matsui et al., 1997). The optimal temperature range for floral development in rice is 23-26^{o}C, and temperatures above 35-36^{o}C during anthesis can result in significant (>50%) floral sterility (Baker et al., 1995). The response of reducing stomatal conductance under elevated carbon dioxide conditions, while increasing water use efficiency, reflects decreased capacity for transpiration and evaporative cooling and can lead to higher canopy temperatures relative to air temperature in rice (Matsui et al., 1997). While overall rice response to concurrent rises in [CO_2] and temperature is unfavorable, what is yet to be determined is the corresponding genetic or phenological basis for the response variability within *O. sativa* and if this differs in its wild relatives (e.g. *O. rufipogon*).

Rice's global distribution, inbreeding nature, and contrasting domestication histories across ecological and cultural landscapes have resulted in what is observed today as deep population substructure within the species. Traditional morphological classification of *O. sativa* into two distinct Varietal Groups, *Indica* and *Japonica*, is additionally divided using genetic markers into 5 subpopulations (*indica* and *aus* within *Indica*; *temperate japonica*, *tropical japonica*, and *aromatic* subpopulations within *Japonica*), each harboring a unique suite of adaptations (Garris et al., 2005; Zhao et al., 2011; Huang et al., 2012). Though rice-producing regions have historically relied on single subpopulations (e.g. *indica* in South Asia and China, *tropical japonica* in Latin America and the Southern U.S.A., and *temperate japonica*

in Northeast Asia), there is increasing need to diversify the sources of alleles utilized for breeding of new rice varieties for changing climates. Wild ancestors and well-differentiated subpopulations of *O. sativa* present genetic pools that are each enriched for different alleles that may confer resilience under changing climatic conditions.

Previous studies of rice response to [CO_2] x temperature interaction have focused either on single cultivars (Baker et al., 1992; Cai et al., 2015) or very limited germplasm panels typically encompassing a single subpopulation or breeding program (Moya et al., 1998; Baker, 2004; Ziska et al., 2014). For this study, we assembled and evaluated a panel of eleven genetically and geographically distinct rice accessions under controlled conditions in a full factorial experiment at current and mid-century projections of atmospheric carbon dioxide concentration (400 and 600 ppm), two different growth temperatures (29/21 and 34/26^{o}C), and with or without anthesis heat stress at 37/29^{o}C. We address the following questions related to the variability of rice yield responses to [CO_2] x temperature interaction: 1) What are the overall effects of CO_2, temperature, and their interaction on biogeographically distinct accessions, and do they support past findings on single cultivars? 2) Do we observe broad response differentiation within the panel, and how does that relate to known genetic differentiation? 3) Are there individual cultivars that can actually respond positively to [CO_2] and temperature increases?

MATERIALS AND METHODS

Plant material. A panel of eleven diverse photoperiod insensitive accessions of rice was assembled for this study based on geographical and genetic diversity (**Table 5.1**).

The panel represents four subpopulations (*aus*, *indica*, *temperate japonica* and *tropical japonica*) within the two Varietal Groups of cultivated Asian rice, along with one wild accession (an *aus*-like *O. rufipogon* from Malaysia) and one weedy accession (an *indica*-like U.S. red rice (Gealy et al., 2012; Ziska et al. 2012b)). The IL 43-1-2 accession was developed from introgressing the *O. rufipogon* accession (IRGC 105491) into a *tropical japonica* recurrent parent, Jefferson (Imai et al. 2013). All seeds were obtained from the Genetic Stocks *Oryza* collection at the Dale Bumpers National Rice Research Center in Stuttgart, Arkansas (http://www.ars.usda.gov/Main/docs.htm?docid=23562).

Table 5.1. Species, country of origin, and subpopulation association of 11 diverse rice accessions evaluated in this study.

	Accession	**Species**	**Origin**	**Comments**
INDICA	AR-1995-StgS (PI 653423)	*O. sativa*	USA	*indica*-like Stuttgart strawhull red rice
	IR64	*O. sativa*	Philippines	*indica* mega-variety widely grown in South Asia
	Teqing	*O. sativa*	China	High-yielding *indica* variety
	IRGC 105491	*O. rufipogon*	Malaysia	*aus*-like wild donor of IL 43-1-2
	DJ123	*O. sativa*	Bangladesh	*aus* subpopulation
	Shirkati	*O. sativa*	Afghanistan	*aus* subpopulation
JAPONICA	Geumobyeo	*O. sativa*	South Korea	*temperate japonica* subpopulation
	Nipponbare	*O. sativa*	Japan	*temperate japonica* reference genome, reportedly heat tolerant (Matsui et al., 2000)
	IL 43-1-2	*O. sativa*	USA	Jefferson/IRGC 105491 BC_3 introgression line (Imai et al 2013)
	Jefferson	*O. sativa*	USA	Long-grain *tropical japonica* cultivar developed for the U.S.A.
	WAB 56-104	*O. sativa*	Ivory Coast	*tropical japonica* variety used in African breeding programs

Experimental design and growth conditions. Environmental growth chambers (EGC Corp., Chagrin Falls, OH, USA) were used and a split plot design with two [CO_2] treatments as main plots and four temperature treatments (growth and with or without the extreme high temperature at anthesis [AT]) as split plots was applied. Experimental units were represented by a single plant and the entire experiment was repeated, with each repetition considered a replication after chambers were reassigned to a different [CO_2] x temperature treatment. The experiment was replicated three times over 30 months. Plants within the chamber were rotated every two weeks until grain filling to avoid chamber effects. Prior to, and at the end of each replication, [CO_2], humidity, PAR, and temperature were evaluated to assess within and between chamber variability.

Light, measured as photosynthetically active radiation (PAR), was programmed to vary diurnally with the highest PAR (~900 μmol m^2 s^{-1}) occurring during the afternoon at maximum temperature. The daily light period was 12 h with PAR supplied by a mixture of high-pressure sodium and metal halide lamps, averaging 22.3 mol m^{-2} day^{-1} for all chambers. Relative humidity was programmed at 65% daytime, 75% night. A TC-2 controller was used to monitor subsamples of air and the [CO_2] was determined from an absolute infrared gas analyzer (WMA-4, PP Systems, Amesbury, MA, USA). Based on the value, injection of either CO_2 or CO_2-free air was then added to the chamber to maintain the programmed [CO_2] set point. Daily averages of [CO_2] were 410±13 and 613±22 μmol mol^{-1} (ppm) for the ambient and elevated [CO_2] treatments respectively for the study.

Seeds of each rice accession were sown into 11.9 L pots filled with vermiculite

and thinned to one seedling 5-10 days after emergence for a given experimental unit. The large pot size was used to prevent root binding and allowed drainage. Pots were spaced at 20 per m^{-2} and watered to the drip point daily with a complete nutrient solution containing 14.5 mmol m^{-3} nitrogen until seed maturity (Robinson 1984).

Treatments. Mean growth temperature was varied in one of two diurnal modes, 1) from an overnight low of 21°C to an afternoon high of 29°C, or 2) an overnight low of 26°C with an afternoon high of 34°C. The 29 °C day/21°C night condition has been shown in a number of studies to be optimal for rice cultivars and represents the control group for this study (Ziska et al., 2014). For each temperature regime, CO_2 values were set at either 400 or 600 ppm. These concentrations approximate both the current (ambient) [CO_2] and that anticipated by 2060 using the A1FI model (http://www.ipcc-data.org/observ/ddc_co2.html). Because increases in mean temperature may also result in extreme high temperatures (EHTs), an additional high temperature treatment (37 °C day/29°C night) was tested for each of the four [CO_2] x growth temperature treatment combination during anthesis. At anthesis, half of all plants for a given [CO_2] x growth temperature treatment were transferred into an adjacent chamber at the elevated 37 °C day/29°C night for ~72 h, tagged, and transferred back to the original [CO_2] and temperature condition.

Growth and yield evaluations. At weekly intervals following emergence, tiller number was determined for all experimental units. At maturity, identified when >90% of seed on the plant had turned light brown, individual plants were cut at the base and the days to maturity from sowing (DTM), plant tiller number and panicle number were recorded. All leaves were removed from each plant and the leaf area of a subsample of

15-20 leaves was determined photometrically (Li-Cor 3000, Li-Cor Lincoln, NE, USA) as an estimate of leaf area on a whole plant basis (Ziska and McClung, 2008). A subset of 3 panicles per plant was hand-threshed, and the seed collected. Immature seed was identified as light green in color and soft to the touch, whereas mature seed was light brown and firm. Subsamples of 50 immature and 50 mature seed were counted and then weighed to determine individual seed weight and proportion of immature:mature seed was calculated. Biomass of all above-ground plant parts, leaves, panicles, stems and tillers was then determined by dry weight following drying at 65^{o}C until a constant dry mass was obtained.

Photosynthesis measurements. Leaf gas exchange was measured during the late vegetative stage on a single leaf of all accessions under the four [CO_2] x temperature growth conditions, using a CIRAS-2 portable photosynthesis system (PP-Systems, Amesbury, MA). Intact, fully expanded upper canopy leaves that were fully exposed to light *in situ* were enclosed in a narrow-leaf cuvette with programmed control of temperature, humidity, light and [CO_2]. Steady-state values of photosynthesis, stomatal conductance, sub-stomatal [CO_2], leaf temperature, leaf to air water vapor pressure difference and photosynthetic photon flux density (PPFD) were recorded. All measurements were conducted within a few hours of midday. For estimates of midday photosynthesis and stomatal conductance, the PPFD was 1000 mol m^{-2} s^{-1}, [CO_2] was either 380 or 580 ppm and the leaf temperature was either 28 or 33^{o}C, depending on the temperature regime. For the measurements at limiting and at high [CO_2], the PPFD was 1500 mol m^{-2} s^{-1}, the leaf temperature was 30^{o}C, and measurements were made at external an [CO_2] of 90, 180, and 800 ppm. For all measurements, the leaf to air water

vapor pressure difference (VPD) was less than 1.5 kPa. For assessment of acclimation, we analyzed two key parameters that describe leaf photosynthesis: V_{cmax}, the maximum rate of Rubisco carboxylase activity, and J_{max}, the maximum rate of photosynthetic electron transport. In this study we measured the initial slope of the response of photosynthesis to sub-stomatal carbon dioxide concentration (proportional to V_{cmax}) and the rate of photosynthesis at a high [CO_2] (proportional to J_{max}) under standardized temperature and light conditions.

Statistical analysis. All statistical analyses were performed in R. Raw data were first visualized to examine presence of outliers and distribution. Right-skewed variables were log transformed and observations beyond 3 times the median distance from the median were winsorized to the median +/- 3 times the median distance of observations to the median. Pearson's correlation coefficients were calculated for all pairwise combinations of traits and asymptotic p-values extracted using the H_{misc} package. Principal Components Analysis (PCA) was performed using accession mean data collected across each of the eight treatment combinations (two [CO_2] levels x four temperature regimes). The correlation matrix rather than covariance matrix was used in the PCA to account for different units across the traits. Agglomerative hierarchical clustering was performed on the Euclidean distance matrix from scaled line mean data with two linkage methods: complete and average. Linear mixed models were fit with lme4 and functions from the lmerTest package were used to extract p-values associated with the fixed effects using Satterthwaite approximated degrees of freedom. Each linear mixed model followed the form

$$\mathbf{y = X\beta + Zu + \varepsilon}$$

where **β** represented the vector of fixed effects parameters (see S4 for fixed effect variable selection), and **u,** the vector of random effect coefficients. Fitting this varying-intercepts model allowed us to mitigate any differences in general agronomic fitness across the accessions that may result in varying accession-level means. Three types of temperature variables were tested per yield trait. The first predictor set uses temperature regime (Tm) as a fixed effect with four separate levels: 1) Normal growth temperature without anthesis stress (Norm GT/Norm AT), 2) normal growth temperature with anthesis stress (Norm GT/High AT), 3) high growth temperature without additional anthesis stress (High GT/Norm AT), and 4) high growth temperature with additional anthesis stress (High GT/High AT). The second predictor set focuses on growth temperature (GT) and includes it as a fixed effect with two levels (Norm GT and High GT). The third set tests specifically the effect of additional anthesis heat stress (variable name AT) and includes it as a fixed effect with two levels (Norm AT and High AT). All three models also include an interaction term between $[CO_2]$ and the temperature regime. One model was selected for each trait based on lowest Akaike Information Criterion and other variables (DTM, GROUP) were added as necessary to account for phenology and INDICA/INDICA-like versus JAPONICA distributional differences (see **Supp. Fig. 5.5** for variable selection overview). For analysis of photosynthesis measurements, we first fit a varying intercepts model with $[CO_2]$, GT, and their interaction as fixed effects and accession (ACC) as a random effect. Because ACC did not account for any variance (data not shown), we dropped the ACC term and fit a linear regression model with only fixed effects.

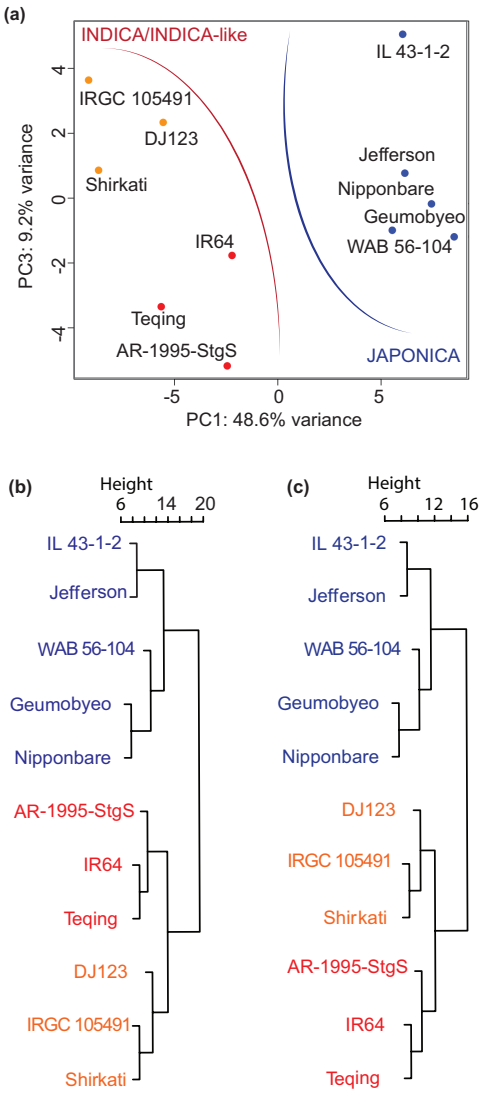
RESULTS

Part I: Multivariate classification of rice accessions

To appropriately model the effects of [CO_2] x temperature in our study, it was first necessary to characterize our data to find correlations and related individuals. Results from exploratory analyses are described below and histograms showing trait distributions are available as **Supp. Fig. 5.1**.

Relationship of yield and growth traits. Five yield-related traits (panicle number/plant, total panicle weight/plant, seed weight/panicle, 50 seed weight, and proportion of immature grains), five growth-related traits (plant height, tiller number, specific leaf area, total plant weight (i.e. above ground biomass), and panicle to tiller ratio,), and one phenological trait (days to maturity) were evaluated on the panel of 11 diverse rice accessions (**Table 5.1**) across eight treatment combinations (two [CO_2] levels x four temperature regimes). Among the five yield traits, panicle number per plant and total panicle weight per plant, proportion of immature grains and

Figure 5.1. Multivariate classification of rice panel from trait information across varying CO_2 and temperature levels. Results using accession means of eleven yield and biomass traits collected from two CO_2 x four temperature combinations. **(a)** Principal Components Analysis plot of PC1 (x-axis) and PC3 (y-axis). Blue font indicate JAPONICA accessions, while orange are *aus*-specific and red are *indica*-specific of the INDICA/INDICA-like accessions. Hierarchical clustering results utilizing **(b)** complete and **(c)** average linkage methods.



seed weight per panicle, and 50 seed weight and seed weight per panicle display the three strongest significant pairwise relationships (r=0.71, -0.62, and 0.44, respectively, $p<0.05$) (**S2**). Among the five growth traits, plant weight is positively correlated with tiller number and plant height (r=0.83 and 0.62, respectively, $p<0.05$), and negatively correlated with the proportion of immature grains (measured as the ratio of immature to mature spikelets) (r=-0.31, $p<0.05$).

Trait-based classification of rice panel is consistent with genetic grouping. Due to significant correlations among traits described in the previous section, we chose to classify our rice accessions using a multivariate approach. To assess whether the trait data gathered from these growth chamber experiments clustered genotypes into previously characterized genetic subpopulations, we performed a Principal Components Analysis using the correlation matrix resulting from accession means for all 11 traits. The first three PCs collectively accounted for 73.8% of total variance; PC1 explains 48.6%, PC2 explains 16.0%, and PC3 explains 9.2%. Together, PC1 and PC3 clustered genotypes into groups that are consistent with previously identified genetic populations (**Fig. 5.1a, Supp. Fig. 5.3**). The first PC dissected the *Indica* Varietal Group (*aus*: Shirkati and DJ123, *indica*: IR64 and Teqing) and *Indica*-like accessions (*aus*-like wild IRGC 105491 and *indica*-like weedy red rice AR-1995-StgS), collectively referred to as the INDICA/INDICA-like group for the remainder of this paper, from the *Japonica* Varietal Group (Jefferson, IL 43-1-2, Nipponbare, WAB 56-104, and Geumobyeo), which is referred to as the JAPONICA group. The third PC distinguished the *aus* subpopulation from the *indica* subpopulation within the INDICA/INDICA-like group. One interesting observation is that PC3 additionally

isolated IL 43-1-2 from the rest of JAPONICA. IL 43-1-2 is an introgression line nearly genetically identical to Jefferson except for small genomic regions from IRGC 105491, its *aus*-like *O. rufipogon* donor. In contrast with PC1 and PC3, PC2 does not separate genotypes in a way that was clearly aligned with genetic information, but rather separates individuals within a Varietal Group based on unknown components of the multivariate trait complex (**Supp. Fig. 5.3**).

For an additional perspective on multivariate classification of the rice panel using these experimental trait data, we next performed agglomerative hierarchical clustering using the same accession means across the eight treatments (two [CO_2] x four temperature regimes) for the 11 traits. Results from clustering using complete and average linkage clearly support the major PC1 division between INDICA/INDICA-like and JAPONICA accessions (**Fig. 5.1b**). Despite clear multivariate differentiation of INDICA/INDICA-like and JAPONICA rice genotypes, overall univariate distributions of the five yield traits did not appear to be different between the two groups, except for panicle number and panicle weight (**Supp. Fig. 5.4**).

Part II: Linear mixed model results for yield-related traits

Next, we were interested in testing the effects of [CO_2] and temperature on the five yield-related traits (seed weight/panicle, panicle weight/plant, panicle number, 50 seed weight, and proportion of immature grain) on our panel. To account for correlated observations arising from experimental replication (i.e. observations of the same accession should be more related than observations of different accessions), linear mixed models (n=290 observations) were fit with accession (ACC) as a random effect,

which allowed for a varying intercept per accession. For each yield trait, different sets of fixed effects variables were selected to explain observed variation (see Methods for selection process). Overall, we observe that productivity under ambient [CO_2] conditions relative to elevated [CO_2] was less affected by temperature across this diversity panel, supporting previous findings on single cultivars or limited panels (Baker, 2004; Baker et al., 1992; Ziska et al., 1996). This was true with respect to four yield-related traits (seed weight/panicle, panicle weight/plant, panicle number, and proportion of immature grain) described below.

Increasing temperature at anthesis decreases panicle seed weight under high [CO_2]. Seed weight per panicle is a yield component of rice and under most conditions is directly proportional to total seed yield per plant. Main effects of [CO_2] and Tm on panicle seed weight were both significant while the F test on their interaction term was insignificant (**Fig. 5.2a, Supp. Table 5.1**). The random term ACC explained a large portion (25.7%) of the variance not explained by fixed effects, indicating that panicle seed weight of accessions responded differently to the treatments (**Supp. Table 5.1**). Under ambient [CO_2], pairwise comparisons of Least-Squares (LS) means of each temperature regime were not significantly different. In contrast, under elevated [CO_2] conditions, the Norm_GT/Norm_AT panicle seed weight LS mean (equivalent to 5.13g) was significantly greater than LS means for panicle seed weight at all other three temperature regimes (High_GT/Norm_AT = 2.09g; Norm_GT/High_AT = 1.29g; High_GT/High_AT = 1.14g) (**Fig. 5.2a, Supp. Table 5.1**). These results suggest that elevated [CO_2] in combination with heat stress results in a reduction in seed weight.

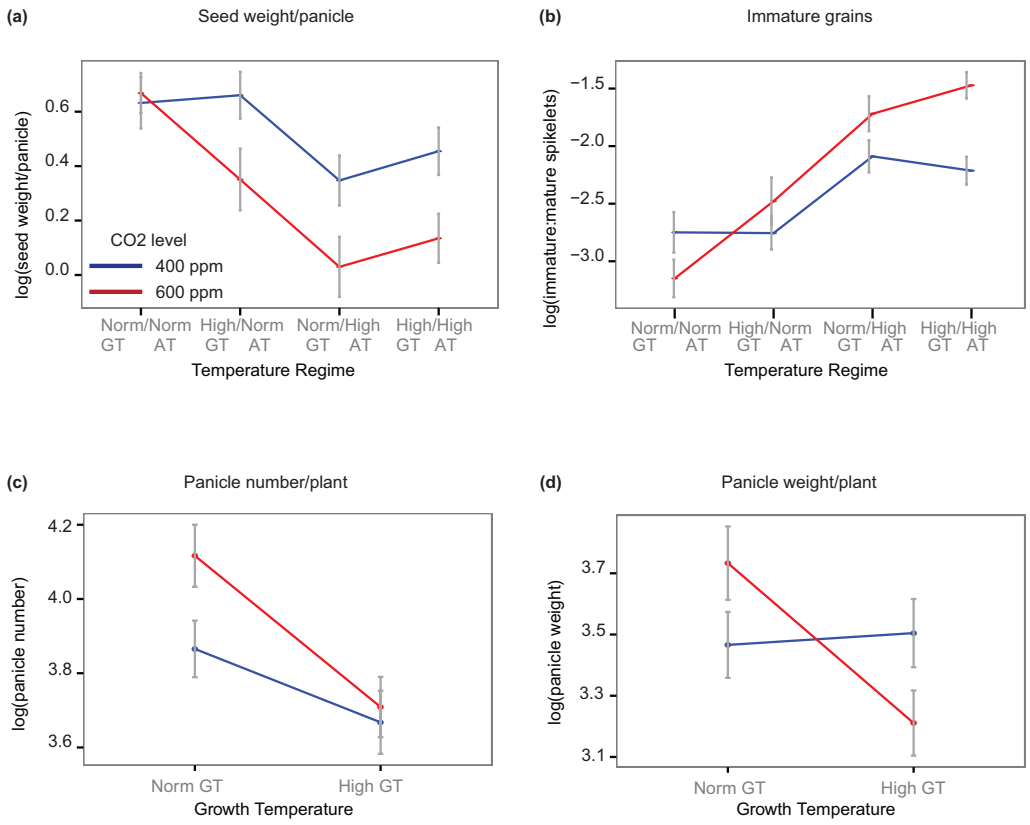


Figure 5.2. Interaction plots of means for four yield component traits. Traces indicate [CO_2] (blue=400 ppm, red= 600 ppm) and temperature regimes are displayed along the x-axis.

Temperature and [CO_2] interact to affect the proportion of filled grains. Proportion of completely filled grain is critical to yield as well as grain quality; partially filled grain reduces grain milling yields and lowers market value. The mixed model selected for proportion of immature grains evaluated Tm, [CO_2], and Tm x [CO_2] as fixed effects and ACC as a random effect. The interaction term was significant (**Fig. 5.2b, Supp. Table 5.2**) and a comparison of group LS means revealed that under ambient [CO_2] conditions, only anthesis temperature stress affects the proportion of immature grain. In contrast, under elevated [CO_2] conditions, the control group, Norm_GT/Norm_AT,

contrasted significantly with all other temperature regimes. When back calculated to untransformed data, this was equivalent to a proportion of immature grain of 0.0006 (Norm_GT/Norm_AT), 0.0035 (High_GT/Norm_AT), 0.018 (Norm_GT/High_AT), or 0.037 (High_GT/High_AT). From these data, temperature rise (either during growth duration or at anthesis) increases the proportion of immature grains. Immature grain as measured in this study are partially-filled, green kernels; hence, observed increased proportion of immature grain could be due to an increase in sink production (i.e. more spikelets/plant) under high [CO_2]/high GT, leading to source deficiency.

Increased growth temperature negates stimulatory effect of elevated [CO_2] on total panicle number per plant. The number of panicles on a rice plant affects yield by contributing to the number of seed produced. Under elevated [CO_2] conditions, panicle number may be expected to increase due to the stimulatory effect of [CO_2] on tiller production(Shimono et al., 2009; Shimono, 2011). To evaluate how temperature affects panicle number under higher [CO_2], we fit a linear mixed model using GT, [CO_2], and their interaction. We also included GROUP to account for observed INDICA/INDICA-like versus JAPONICA differences. The main effects of both [CO_2] and GT were significant for panicle number (**Fig. 5.2c, Supp. Table 5.3**); the elevated [CO_2] condition results in a decrease in 0.56 panicle/plant on average relative to ambient [CO_2] condition while elevated growth temperature condition results in a decrease of 0.46 panicle/plant on average relative to normal growth temperature condition. LS means comparisons between the four treatments (two [CO_2] x two GT) revealed that under normal growth temperatures (29^o day/21 ^{o} night), elevated [CO_2] stimulated panicle number but under increased growth temperatures (34^o day/26 ^{o}

night), the stimulatory effect disappeared.

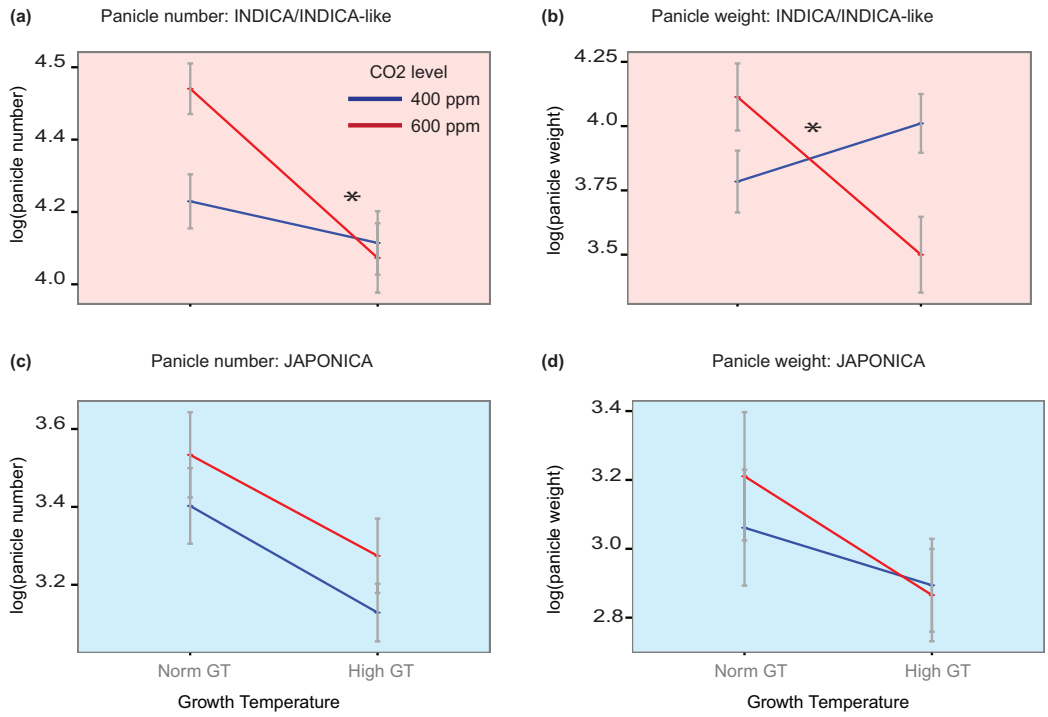
High growth temperature decreases total panicle weight/plant under high [CO_2] conditions. For total panicle weight per plant, we fit a linear mixed model using GT and GROUP and found a significant [CO_2] x GT interaction (**Fig. 5.2d, Supp. Table 5.4**). Under ambient [CO_2] conditions, growth temperature did not affect total panicle weight per plant, but under elevated [CO_2], the increased growth temperature decreased total panicle weight. The decreased total panicle weight under high [CO_2] and high growth temperature may be related to a decrease in panicle number/plant (**Supp. Table 5.3**), a decrease in spikelet number/panicle, and/or an increased proportion of immature grains (**Supp. Table 5.2**).

Part III: Group- and accession-level contrasts

INDICA/INDICA-like and JAPONICA rice respond differentially to [CO_2] and growth temperature with respect to panicle number and panicle weight per plant. From results above, we observed that GROUP had a significant main effect on two traits: panicle number/plant and panicle weight/plant (**Supp. Tables 5.3, 5.4**). To determine if INDICA/INDICA-like and JAPONICA also responded differently under varying [CO_2] and GT conditions (Tm without AT treatment), we fit the models for panicle number and panicle weight separately for each group and were surprised to see that the JAPONICA group (130 observations, five accessions) does not respond significantly to [CO_2] nor [CO_2] x GT for either trait. In contrast, the [CO_2] x GT term is significant in INDICA/INDICA-like (160 observations, six accessions)) for both traits. Although there were no differences in response to temperature under ambient

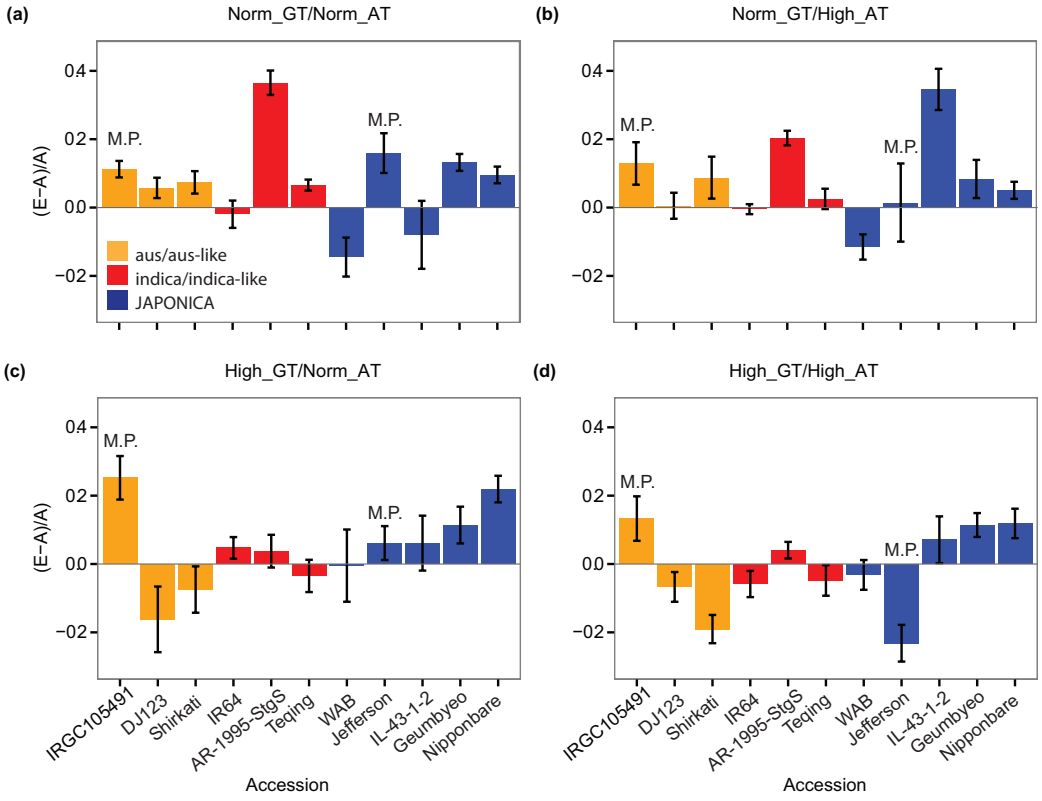
[CO_2] conditions, under elevated [CO_2], high GT resulted in decreased panicle number and panicle weight (**Fig. 5.3a, Supp. Tables 5.5, 5.6**). Notably, the percent variance explained by ACC after fitting the fixed effects is much higher in the JAPONICA models than in the INDICA/INDICA-like models: 21.04% (panicle number) and 12.3% (panicle weight) for JAPONICA versus 5.9% (panicle number) and 0.81% (panicle weight) for INDICA/INDICA-like. These data provide evidence that INDICA/INDICA-like and JAPONICA rice groups respond differently to [CO_2] and temperature. Given that the INDICA/INDICA-like accessions were negatively impacted by high temperatures under elevated [CO_2], these results suggest that they will not respond favorably to predicted future climate change scenarios in the tropics. In contrast, JAPONICA accessions are less responsive or sensitive to climate changes, and despite their lower overall yield performance in this study, they may carry genetic factors that could enhance the stability of yield in the INDICA/INDICA-like group under high [CO_2] and temperature (**Fig. 5.3**).

Figure 5.3. Differential response of INDICA/INDICA-like versus JAPONICA rice to [CO_2] and temperature for panicle number and panicle weight. Interaction plots of means for panicle number/plant and panicle weight/plant. Traces show [CO_2] while growth temperature (Norm_GT=29^o day/21^o night, Hi_GT= 34^o day/26^o night) is displayed along the x-axes. INDICA/INDICA-like varieties are significantly affected by [CO_2] and growth temperature interaction for both panicle traits while JAPONICA varieties are not significantly affected by this interaction for either yield trait (see Results for model outputs). Asterisk (*) indicates significant interaction effect.



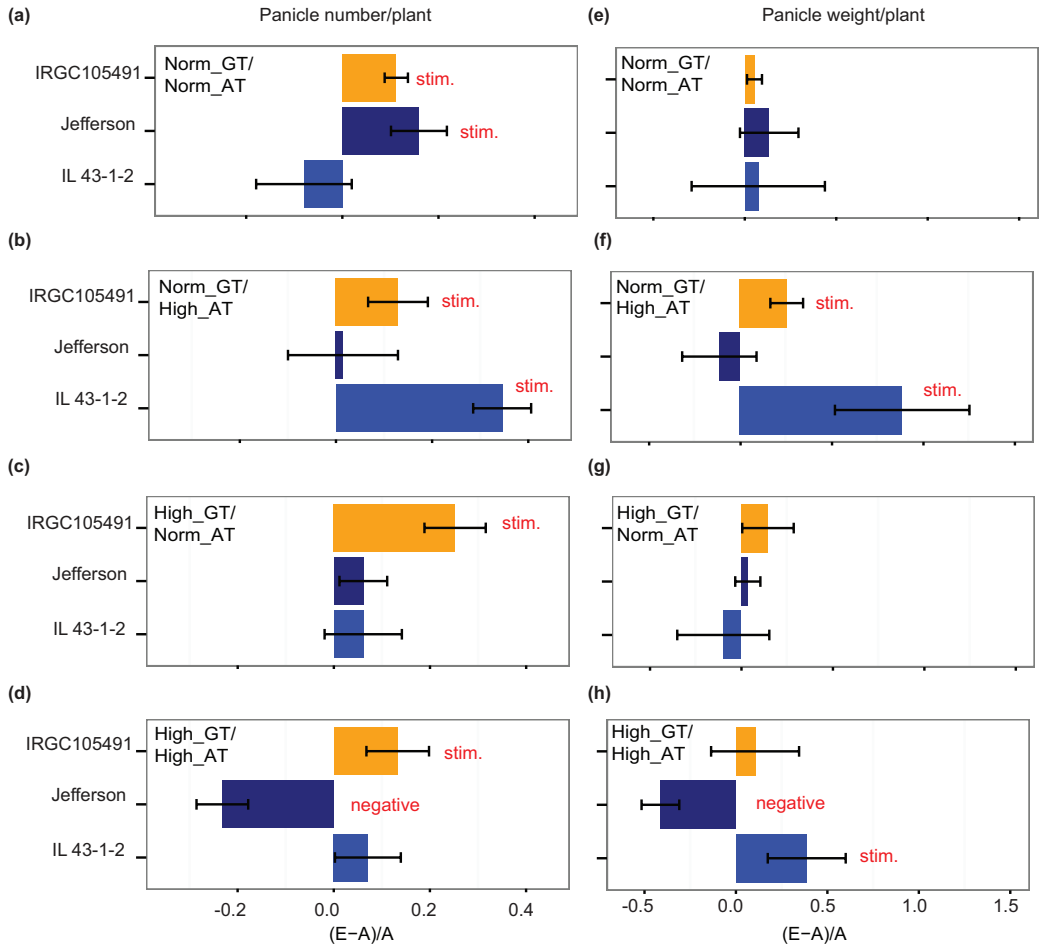
Inter-accession variability in [CO_2] stimulation response across temperature regimes. We next assessed the pattern of stimulation variability across individual accessions for panicle number/plant, which displayed a stimulatory response to elevated [CO_2] at the control temperature condition (**Fig. 5.2c**). Relative stimulation was calculated using the following equation: (E-A)/A, where E and A are the value of the measured parameter at 600 and 400 ppm CO_2, respectively (see Ziska et al., 2012a). Under the control temperature condition, eight accessions were positively stimulated, two accessions showed no response and one accession (cv. Jefferson) had a negative response (**Fig. 5.4a**). On the whole, increases in either growth temperature or anthesis temperature decreased the number of accessions that were positively stimulated by elevated carbon dioxide, however, JAPONICA accessions (blue bars, **Fig. 5.4**) tended to remain positively stimulated across the different temperature regimes. Interesting to note is the resiliency of IRGC 105491, an *aus*-like wild accession, which consistently displayed positive [CO_2] stimulation for panicle number/plant across all four temperature regimes in contrast to other *aus* accessions. AR-1995-StgS, an *indica*-like weedy rice accession, also outperformed the two cultivated *indicas* at three out of the four temperature regimes. Thus, the two accessions that are not developed cultivars, IRGC 105491 and AR-1995-StgS, appeared to have the most resilient response in panicle number/plant when challenged with different climate scenarios.

Figure 5.4. Relative stimulation by [CO_2] enrichment for panicle number/plant for individual rice accessions across four temperature treatments. Relative stimulation is calculated by (E-A)/A where E and A are panicle number per plant at 600 and 400 ppm, respectively. Bars indicate standard error of the mean (E-A)/A for a given accession and treatment. 'M.P' ('Mapping Parent') marks accessions that are the parents used to develop IL 43-1-2.



Comparison of an introgression line, IL 43-1-2, and its two parents. Included in our diversity panel were two genetically divergent parents (*O. sativa* cv. Jefferson and *O. rufipogon* acc. IRGC 105491) of a mapping population (Thomson et al., 2003) and an introgression line (IL 43-1-2) derived from this population that outyields Jefferson in field trials (Imai et al., 2013). While the cultivated parent displayed stimulation in only one scenario (Normal_GT/Normal_AT for panicle number, **Fig. 5.5a**), the wild donor parent IRGC105491 was positively stimulated under most conditions (**Fig. 5.5a-d, f**). Under the control temperature regime, we detected no improvement of the introgression line, IL 43-1-2, relative to Jefferson for stimulation of either panicle number/plant or panicle weight/plant (**Fig. 5.5a, e**). However, under increased temperatures, IL 43-1-2 outperformed Jefferson (**Fig. 5.5b, f, e, h**). The most compelling comparison is for panicle weight/plant under High_GT/High_AT conditions in which Jefferson displayed a negative response while IL 43-1-2 was positively stimulated. Contrasts between Jefferson and IL 43-1-2 must primarily be due to one or more of the donor introgressions from IRGC 105491 in the genetic background of the cultivated parent.

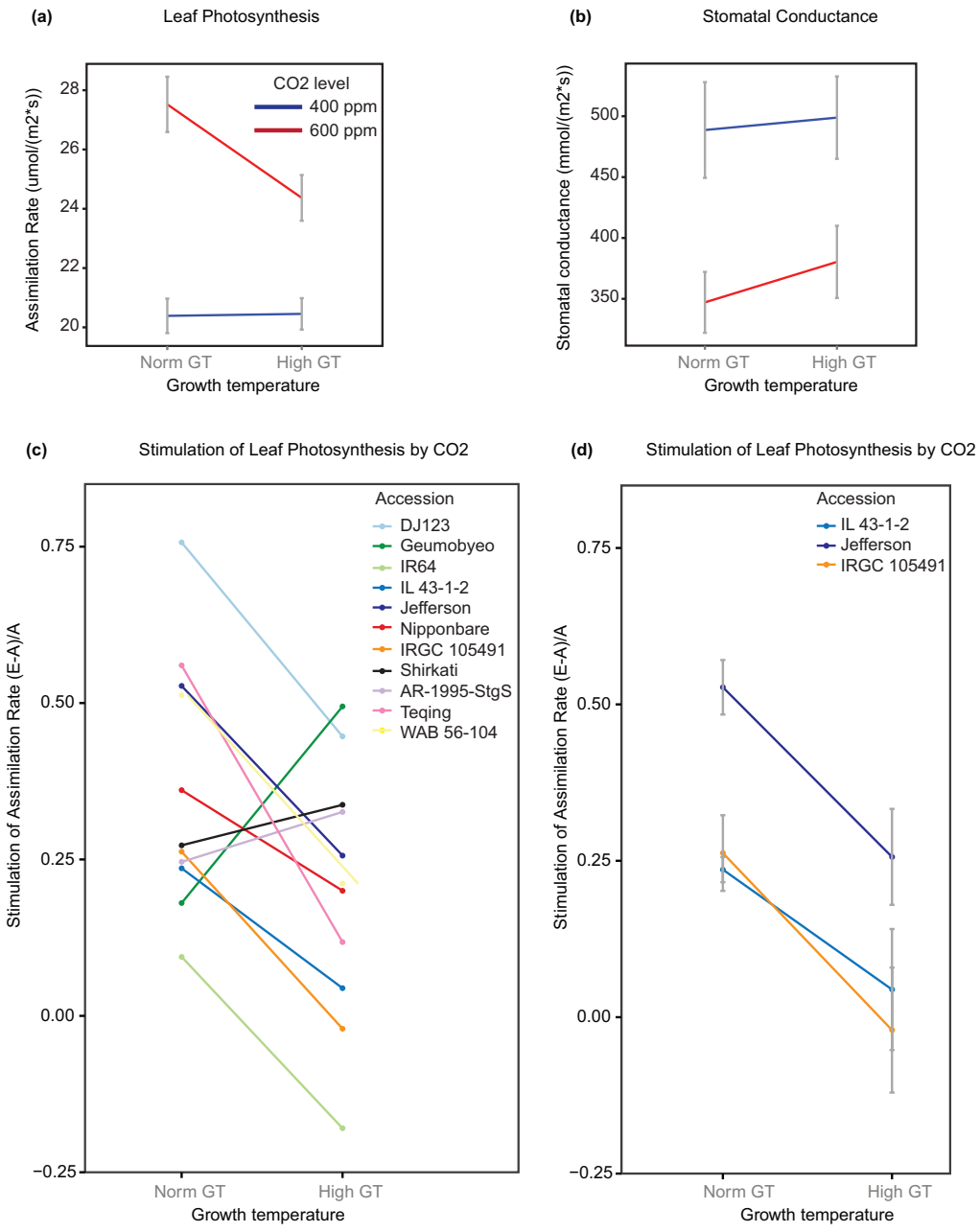
Figure 5.5. Comparison of [CO_2] stimulation of an introgression line (IL 43-1-2) with its two parents. IL43-1-2 is an introgression line genetically identical to Jefferson except for 5 introgressions from *aus*-like wild donor, IRGC105491 (Imai et al. 2013). Panicle number/plant **(a-d)** and panicle weight/plant **(e-h)** are compared across the three accessions. Favorable transgressive variation for IL 43-1-2 is evident under Norm_GT/High_AT (**b, f**) and High_GT/High_AT (**d**). Labels 'negative' and 'stim' are added if the se interval did not include 0; 'negative' indicates a negative (E-A)/A value while 'stim' indicates a positive value.



Part IV: Investigation into photosynthetic activity as a basis for rice response differences

Single-leaf photosynthesis response to [CO_2] and growth temperature. In addition to yield and growth evaluation, we conducted measurements of leaf gas exchange during the late vegetative stage on all accessions across all four combinations of [CO_2] and growth temperatures. Leaves were measured for two purposes: 1) to determine the effect of [CO_2] and growth temperature on steady-state values of photosynthesis, and 2) to assess the degree of acclimation, if any, of leaves conditioned to high [CO_2]. Addressing the first objective, leaves under elevated [CO_2] are expected to have increased photosynthetic rates (A, μmol m^{-2} s^{-1}) and decreased stomatal conductance (g_s, mmol m^{-2} s^{-1}). In this study, we found a significant [CO_2] x GT interaction for leaf photosynthetic rate: under the normal growth condition (29^o day/21^o) the effect of [CO_2] is greater than at elevated growth temperature (**Fig. 5.6a, Supp. Table 5.8**). Stomatal conductance decreased significantly, as expected, with increased [CO_2] but was not affected by [CO_2] x GT (**Fig. 5.6b**). Previous studies have suggested photosynthetic acclimation of leaves grown at high [CO_2] for extended periods, but from our data we detected no significant effect of growth condition [CO_2], GT, or [CO_2] x GT on either parameter, indicating lack of photosynthetic acclimation of the plants (see Materials and Methods).

Figure 5.6. Responses of photosynthesis-related traits to [CO_2] and temperature. Plants grown under elevated [CO_2] (red traces = 600 ppm) display higher leaf photosynthesis (**a**) and lower stomatal conductance (**b**) than plants grown under ambient [CO_2] (blue traces = 400 ppm). The overall positive response of leaf photosynthesis under elevated [CO_2] (i.e. stimulation) decreases with increased growth temperature but this response varies across accessions (**c**). (**d**) The same analysis as in (**c**) but a direct comparison of Jefferson, IL 43-1-2, and IRGC 105491.



To see if photosynthetic activity may underlie patterns of [CO_2] stimulation of yield-related traits across our panel, we assessed [CO_2] stimulation on photosynthetic rate under normal and high growth temperature conditions (**Fig. 5.6c-d**). Stimulation of single leaf photosynthetic rate was generally lower at elevated growth temperature, however, three accessions stood out as outliers: Geumobyeo (JAPONICA), Shirkati (INDICA/INDICA-like: *aus*), and AR-1995-StgS (INDICA/INDICA-like: *indica*). Interestingly, these three individuals show equal or higher stimulation at the higher growth temperature condition (**Fig. 5.6c, Supp. Fig. 5.6**). Overall, there was no clear-cut pattern in the panel between single-leaf photosynthetic rate response stimulation by [CO_2] and yield/growth trait response (data not shown). However, comparing IL 43-1-2 with its two parents, Jefferson and IRGC 105491, we observe that photosynthetic rate stimulation by [CO_2] across the two growth temperatures in the introgression line is nearly identical to that of its wild donor, contrasting with cv. Jefferson. This gives credence to the possibility that there may be foundational physiological differences between these two nearly genetically identical individuals (cv. Jefferson and IL 43-1-2), and suggests that maintaining greater capability of photosynthetic stimulation by [CO_2] at high temperatures does not necessarily contribute to stimulation of yield-related panicle traits (**Fig. 5.6d, Fig. 5.5**). Our data underscore the complex relationship between basic physiological processes and morphological/yield expression.

DISCUSSION

[CO_2] and temperature must be considered together when evaluating rice under

projected climate change scenarios. The antagonistic interaction effect of high [CO_2] and high temperature, reported by this and other studies (Baker et al., 1989; Baker et al., 1992; Ziska et al., 1996; Baker, 2004), undermines efforts to identify sources of elevated [CO_2] responsiveness *per se* for rice yield enhancement (Hasegawa et al., 2013; Zhu et al., 2015). The combination of high [CO_2] and high temperature is an abiotic stress in its own right and presents a unique challenge for breeding; projected scenarios for these factors are not yet expressed, precluding identification of sources of adaptation from existing production environments. In addition, growth chambers or Free-Air CO_2 Enrichment (FACE) setups ideal for physiological studies are not scalable to accommodate the massive volume of lines evaluated by breeding programs, making direct population improvement for high [CO_2]/high temperature resiliency unfeasible. Instead, unraveling the genetic architecture of [CO_2] x temperature response under experimental conditions to model and identify sources of adaptation may be the best way to ensure a future repository of useful alleles with which breeders can hedge against uncertain climatic conditions.

The current study uses genetic diversity as a framework to explore rice responses to long and short-term temperature increases (growth and anthesis, respectively) concurrent with elevated [CO_2]. Consistent with previous findings on limited germplasm, we observed that plants grown under elevated [CO_2] are indeed more sensitive to increases in temperature than plants grown under current [CO_2] levels for seed weight/panicle, panicle number/plant, panicle weight/plant, and proportion of immature grain across a panel of eleven genetically diverse accessions. It is possible that reduced stomatal conductances at elevated [CO_2] may result in

higher tissue temperatures which is further exacerbated by higher growth temperatures.

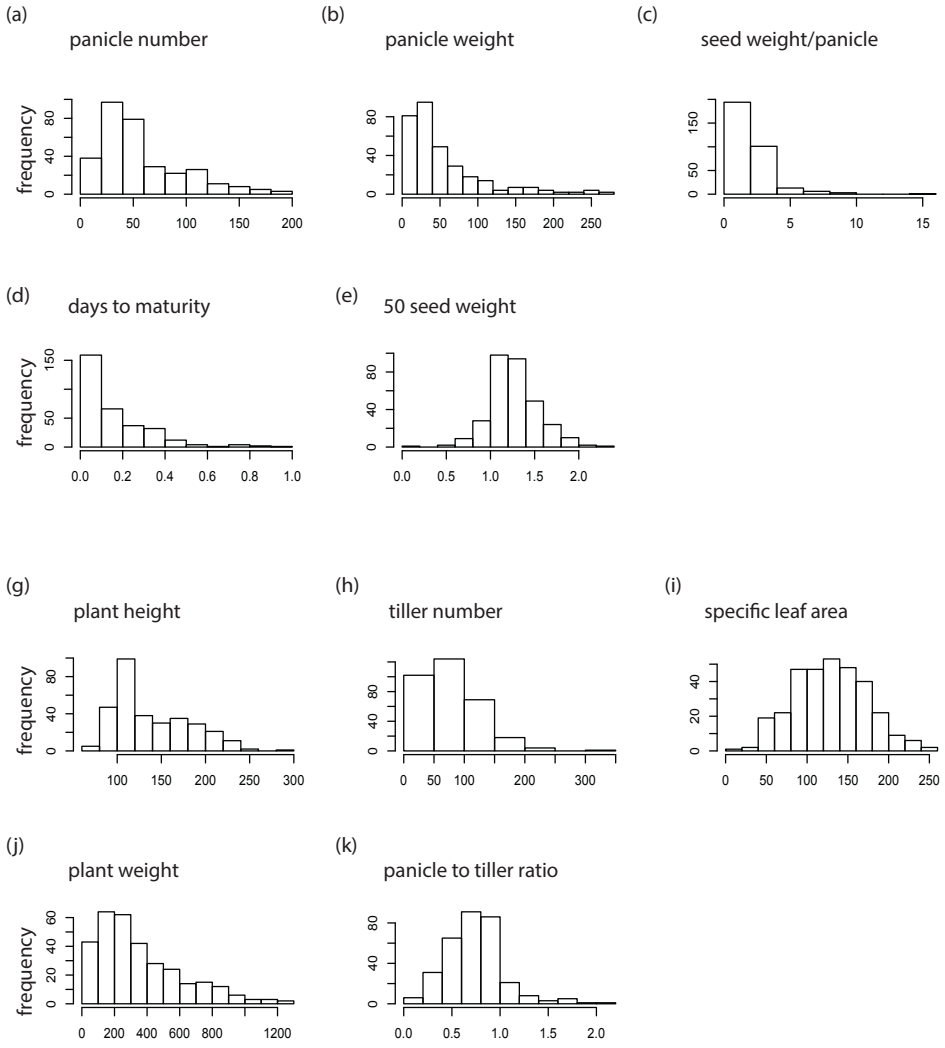
Our results provide preliminary evidence that introduction of exotic alleles from wild ancestors may benefit *O. sativa* cultivar performance under increasing [CO_2] and rising temperature. IRGC 105491, the genetically *aus*-like wild relative to *O. sativa,* displayed positive panicle number stimulation to elevated [CO_2] under all temperature conditions evaluated. When IRGC 105491 was used as a donor in backcrosses to cv. Jefferson (a JAPONICA), the derived IL 43-1-2, which carries five genomic regions inherited from the wild parent, showed enhanced stimulation of panicle number/plant and panicle weight/plant over Jefferson under temperature stress conditions. This finding may underlie the PCA result of IL 43-1-2's separation from the rest of JAPONICA along PC3 (**Fig. 5.1a**) and further motivates interest in this interspecific introgression line, which has previously been shown to have 27.7% yield improvement over Jefferson in multi-state yield evaluations across two years (Imai et al., 2013). Before the present study, physiological parameters had not been compared between Jefferson and its IL. The observation that IL 43-1-2 displays photosynthetic rate stimulation similar to its wild donor parent, at least under experimental conditions, motivates deeper studies into physiological differences between the IL and cv. Jefferson as a basis for IL 43-1-2's superior performance under field conditions.

Contrary to previous suggestions that plasticity in response to [CO_2] would be desirable (Zhu et al., 2015), we find that among the eleven global rice accessions that were evaluated, the more plastic INDICA/INDICA-like group responded in a detrimental manner at high [CO_2] /high growth temperature while JAPONICA

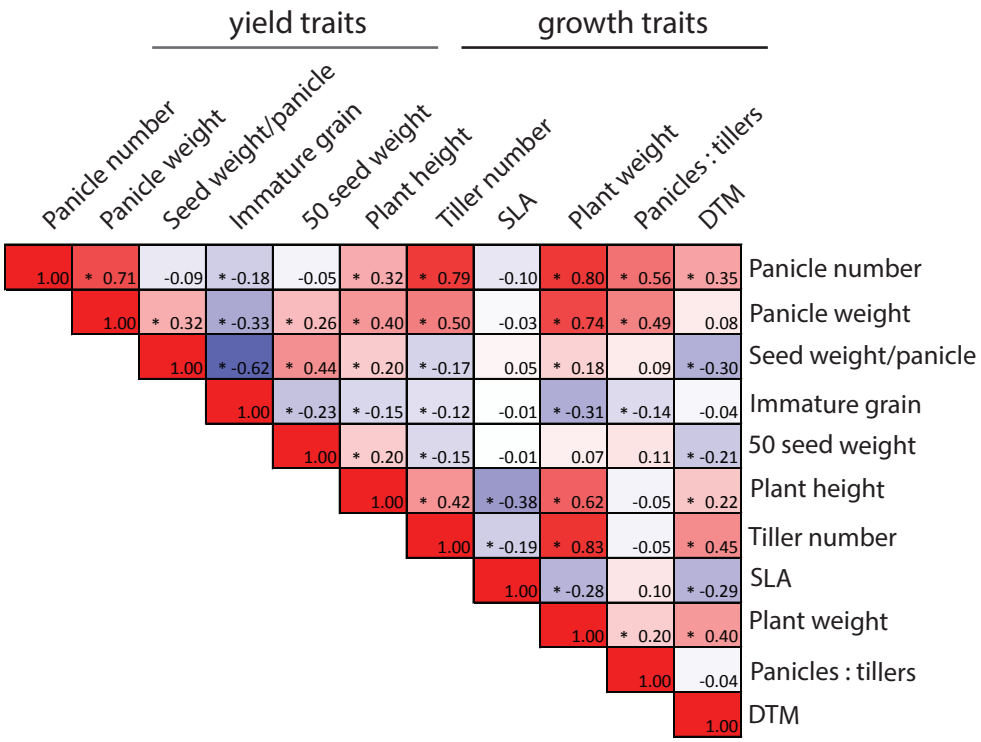
displayed stability overall. Yet the overall yield potential of INDICA/INDICA-like accessions was generally higher than the JAPONICA group, regardless of the treatment. This indicates that germplasm from both Varietal Groups may have limitations in responding to climate change, while recombining diversity across divergent gene pools through breeding may offer a means to overcome such potential genetic and physiological bottlenecks. Favorable responses, such as increased resilience or beneficial plasticity, in the face of climate change may come from crosses between subpopulations or between species, as demonstrated in this study. Although there is wide diversity within *Indica* and Japonica populations, they are also distinctly different genetically, in their geographic adaptation, and in other intrinsic characters (eg. grain quality) that are preferred by local markets. Breeders will be challenged to reach beyond their current gene pools to develop new cultivars that break physiological barriers to response to climate change while maintaining traits that are important to local rice production systems and consumer markets.

APPENDIX

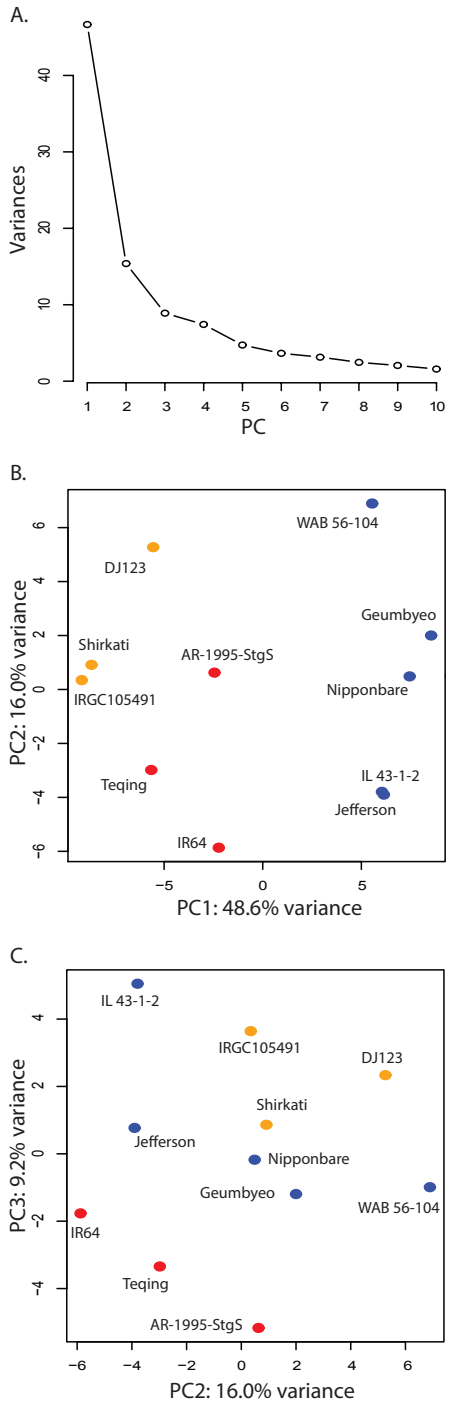
Supplemental Figure 5.1: Trait histograms for yield and growth related traits with non-transformed data.



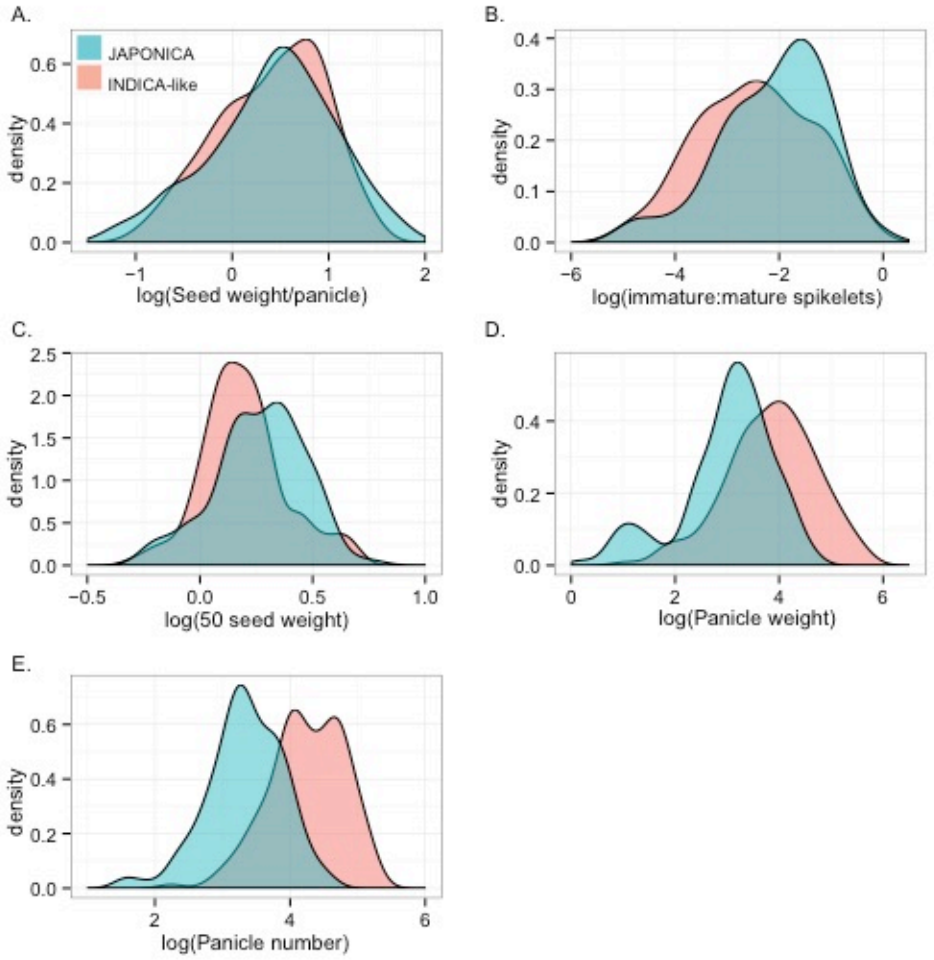
Supplemental Figure 5.2: Correlation matrix of eleven traits evaluated on eleven rice accessions across two CO_2 levels and four temperature regimes. Asterisk indicates asymptotic p-values < 0.05.



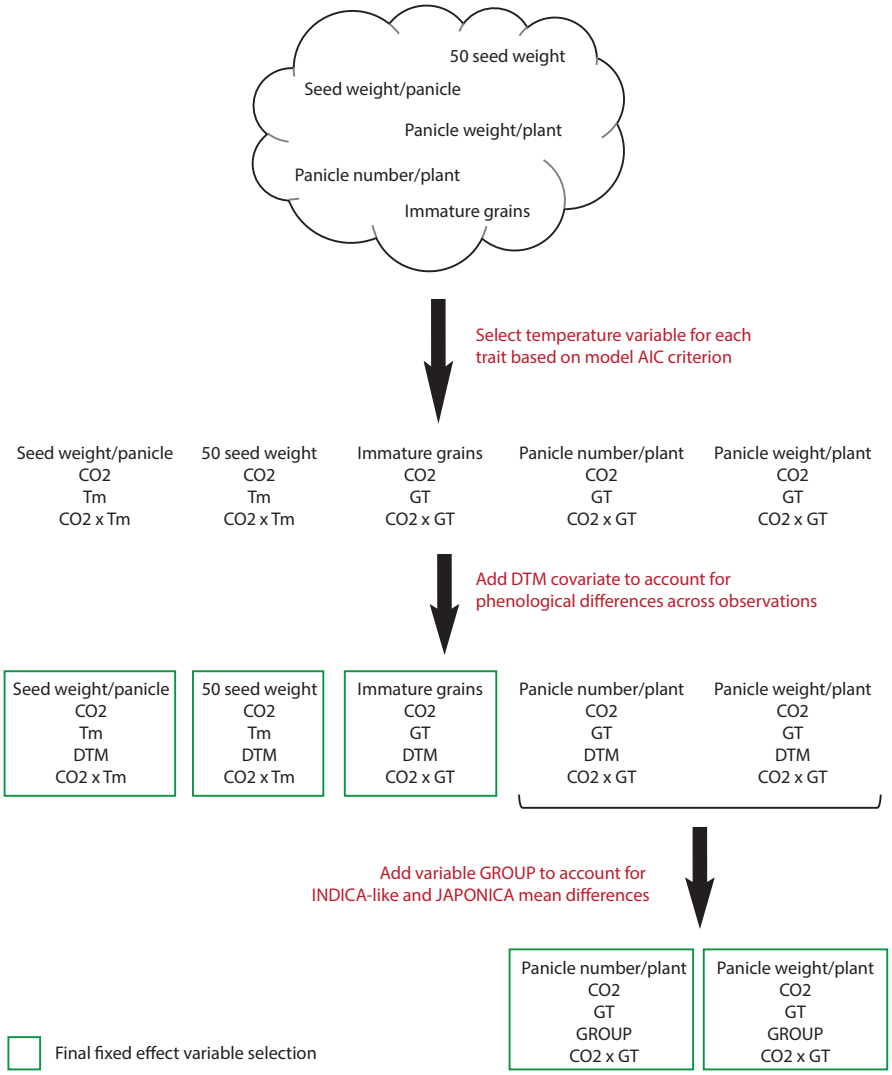
Supplemental Figure 5.3: Additional PCA plots using line mean trait data across 8 treatment combinations (two [CO_2] x four Tm). A) Scree plot displaying variances explained by each Principal Component. B) PC1 versus PC2 and C) PC2 versus PC3.



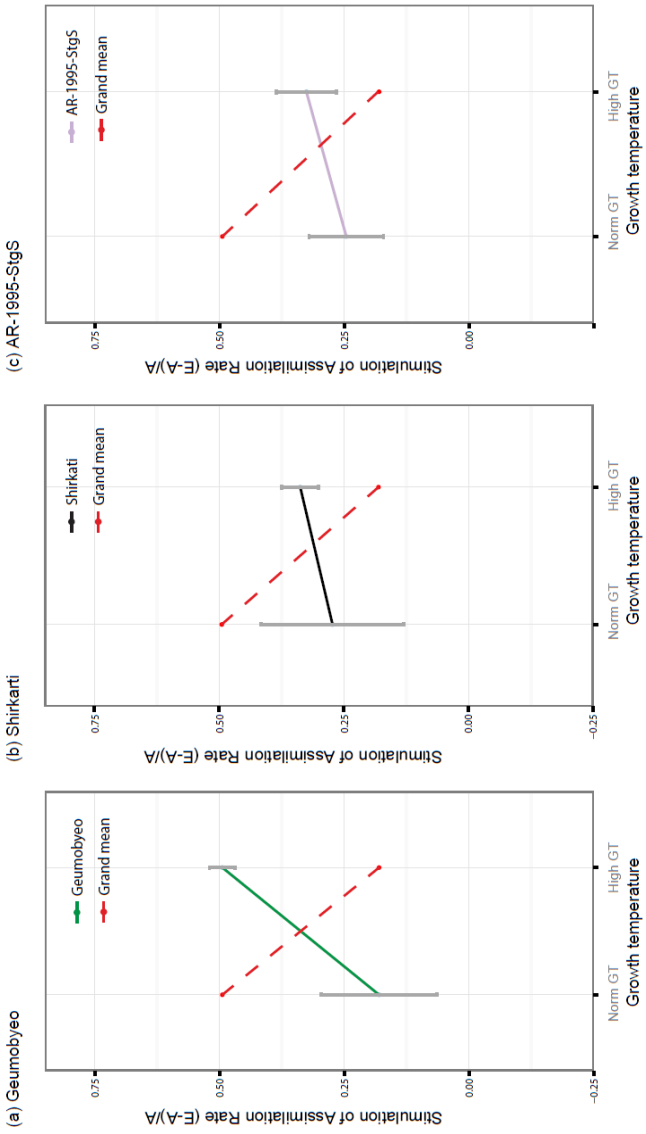
Supplemental Figure 5.4: Density plots for yield traits. Distributions are colored according to INDICA/INDICA-like and JAPONICA groupings. No differences in distributions except for panicle weight/plant (D) and panicle number/plant (E).



Supplemental Figure 5.5: Overview of the variable selection process for fixed effect variables.



Supplemental Figure 5.6: Stimulation of single-leaf photosynthetic rate, (E-A)/A, for three outlier accessions: (a) Geumobyeo (*Japonica*), (b) Shirkati (*aus*), (c) AR-1995-StgS (*indica*-like). Dotted red lines indicate grand means across all 11 accessions.



Supplemental Table 5.1: Mixed model results for seed weight/panicle.

Fixed Effects

	Estimate	Std. Error	df	t value	Pr(>|t|)	
(Intercept)	5.27	0.75	245.93	7.05	1.77E-11	***
Amb_CO2	-0.01	0.11	275.27	-0.11	9.16E-01	
Hi_GT_Hi_AT	-0.66	0.11	276.77	-5.95	7.99E-09	***
Hi_GT_Norm_AT	-0.40	0.11	275.86	-3.63	3.41E-04	***
Norm_GT_Hi_AT	-0.60	0.11	275.26	-5.46	1.05E-07	***
DTM	-0.93	0.15	256.19	-6.21	2.14E-09	***
Amb_CO2/Hi_GT_Hi_AT	0.32	0.16	275.13	2.06	4.04E-02	*
Amb_CO2/Hi_GT_Norm_AT	0.29	0.16	275.28	1.79	7.39E-02	.
Amb_CO2/Norm_GT_Hi_AT	0.36	0.16	275.19	2.29	2.26E-02	*

Random Effects

	Variance	% Variance
ACC (Intercept)	0.08	25.66
Residual	0.23	74.34
Total	0.31	

Fixed Effects Type III ANOVA

	Sum Sq	Mean Sq	NumDF	DenDF	F.value	Pr(>F)	
CO2	3.84	3.84	1.00	275.39	16.88	5.25E-05	***
Tm	10.48	3.49	3.00	277.37	15.37	2.79E-09	***
DTM	8.76	8.76	1.00	256.19	38.55	2.14E-09	***
CO2 x Tm	1.50	0.50	3.00	275.25	2.20	8.87E-02	.

LSMeans

CO2	Tm	lsmean	SE	df	lower.CL	upper.CL	
Hi_CO2	Hi_GT/Hi_AT	0.06	0.11	26.33	-0.18	0.29	1
Hi_CO2	Norm_GT/Hi_AT	0.11	0.12	27.06	-0.12	0.35	1
Hi_CO2	Hi_GT/Norm_AT	0.32	0.11	26.23	0.08	0.55	12
Amb_CO2	Hi_GT/Hi_AT	0.37	0.12	30.85	0.13	0.61	123
Amb_CO2	Norm_GT/Hi_AT	0.46	0.12	27.47	0.22	0.70	23
Amb_CO2	Hi_GT/Norm_AT	0.59	0.12	32.54	0.34	0.84	23
Amb_CO2	Norm_GT/Norm_AT	0.70	0.12	27.57	0.46	0.94	3
Hi_CO2	Norm_GT/Norm_AT	0.71	0.11	26.70	0.48	0.95	3

Supplemental Table 5.2: Mixed model results for immature grains

Fixed Effects

	Estimate	Std. Error	df	t value	Pr(>|t|)	
(Intercept)	-5.54	1.34	170.88	-4.13	5.63E-05	***
Amb_CO2	0.37	0.21	274.53	1.81	7.21E-02	.
Hi_GT_Hi_AT	1.74	0.21	277.15	8.40	2.22E-15	***
Hi_GT_Norm_AT	0.71	0.21	275.57	3.47	5.95E-04	***
Norm_GT_Hi_AT	1.42	0.21	274.51	6.90	3.64E-11	***
DTM	0.49	0.27	174.94	1.80	7.42E-02	.
Amb_CO2/Hi_GT_Hi_AT	-1.11	0.30	274.27	-3.76	2.07E-04	***
Amb_CO2/Hi_GT_Norm_AT	-0.65	0.30	274.55	-2.16	3.18E-02	*
Amb_CO2/Norm_GT_Hi_AT	-0.77	0.29	274.41	-2.65	8.56E-03	**

Random Effects

	Variance	% Variance
ACC (Intercept)	0.13	13.68
Residual	0.80	86.32
Total	0.93	

Fixed Effects Type III ANOVA

	Sum Sq	Mean Sq	NumDF	DenDF	F.value	Pr(>F)	
CO2	4.86	4.86	1.00	274.78	6.05	1.45E-02	*
Tm	67.40	22.47	3.00	278.06	27.95	7.77E-16	***
DTM	2.59	2.59	1.00	174.94	3.23	7.42E-02	.
CO2 x Tm	12.03	4.01	3.00	274.50	4.99	2.20E-03	*

LSMeans

CO2	Tm	lsmean	SE	df	lower.CL	upper.CL	
Hi_CO2	Norm_GT/Norm_AT	-3.17	0.18	48.02	-3.53	-2.80	1
Amb_CO2	Norm_GT/Norm_AT	-2.79	0.18	49.93	-3.16	-2.43	12
Amb_CO2	Hi_GT/Norm_AT	-2.73	0.19	61.43	-3.12	-2.34	123
Hi_CO2	Hi_GT/Norm_AT	-2.45	0.18	46.69	-2.82	-2.09	23
Amb_CO2	Hi_GT/Hi_AT	-2.17	0.19	57.41	-2.55	-1.79	234
Amb_CO2	Norm_GT/Hi_AT	-2.15	0.18	49.20	-2.52	-1.78	34
Hi_CO2	Norm_GT/Hi_AT	-1.75	0.18	48.72	-2.11	-1.38	45
Hi_CO2	Hi_GT/Hi_AT	-1.43	0.18	46.46	-1.80	-1.07	5

Supplemental Table 5.3: Mixed model results for panicle number/plant

Fixed Effects

	Estimate	Std. Error	df	t value	Pr(>|t|)	
(Intercept)	3.04	0.80	223.60	3.81	1.78E-04	***
Amb_CO2	-0.25	0.08	279.46	-3.04	2.60E-03	**
Hi_GT	-0.34	0.08	282.55	-4.02	7.41E-05	***
JAPONICA	-0.84	0.14	9.73	-5.97	1.53E-04	***
DTM	0.29	0.16	234.94	1.83	6.78E-02	.
Amb_CO2/Hi_GT	0.21	0.12	279.17	1.76	7.87E-02	.

Random Effects

	Variance	% Variance
ACC (Intercept)	0.04	13.69
Residual	0.26	86.31
Total	0.30	

Fixed Effects Type III ANOVA

	Sum Sq	Mean Sq	NumDF	DenDF	F.value	Pr(>F)	
CO2	1.56	1.56	1.00	279.33	6.08	1.42E-02	*
GT	3.43	3.43	1.00	285.71	13.38	3.03E-04	***
GROUP	9.15	9.15	1.00	9.73	35.64	1.53E-04	***
DTM	0.86	0.86	1.00	234.94	3.37	6.78E-02	.
CO2 x GT	0.80	0.80	1.00	279.17	3.12	7.87E-02	.

LSMeans

CO2	GT	lsmean	SE	df	lower.CL	upper.CL	
Amb_CO2	Hi_GT	3.65	0.09	25.51	3.47	3.83	1
Hi_CO2	Hi_GT	3.69	0.08	20.95	3.51	3.87	1
Amb_CO2	Norm_GT	3.78	0.09	21.66	3.60	3.96	1
Hi_CO2	Norm_GT	4.03	0.08	21.03	3.85	4.21	2

Supplemental Table 5.4: Mixed model results for panicle weight/plant

Fixed Effects

	Estimate	Std. Error	df	t value	Pr(>|t|)	
(Intercept)	6.92	1.25	156.06	5.54	1.24E-07	***
Amb_CO2	-0.25	0.14	279.93	-1.82	7.01E-02	.
Hi_GT	-0.57	0.14	284.17	-4.12	4.95E-05	***
JAPONICA	-0.97	0.17	10.32	-5.54	2.20E-04	***
DTM	-0.56	0.25	160.06	-2.28	2.41E-02	*
Amb_CO2/Hi_GT	0.52	0.19	279.47	2.70	7.37E-03	**

Random Effects

	Variance	% Variance
ACC (Intercept)	0.05	6.62
Residual	0.69	93.38
Total	0.74	

Fixed Effects Type III ANOVA

	Sum Sq	Mean Sq	NumDF	DenDF	F.value	Pr(>F)	
CO2	0.02	0.02	1.00	279.74	0.03	8.67E-01	
GT	5.86	5.86	1.00	287.52	8.54	3.76E-03	**
GROUP	21.10	21.10	1.00	10.32	30.72	2.20E-04	***
DTM	3.56	3.56	1.00	160.06	5.18	2.41E-02	*
CO2 x GT	5.01	5.01	1.00	279.47	7.29	7.37E-03	**

LSMeans

CO2	GT	lsmean	SE	df	lower.CL	upper.CL	
Hi_CO2	Hi_GT	3.13	0.12	32.80	2.89	3.37	1
Amb_CO2	Hi_GT	3.41	0.13	42.67	3.16	3.66	12
Amb_CO2	Norm_GT	3.45	0.12	34.65	3.21	3.69	12
Hi_CO2	Norm_GT	3.70	0.12	33.61	3.46	3.93	2

Supplemental Table 5.5: Results of mixed models for INDICA/INDICA-like group

A. Panicle Number

Fixed Effects

	Estimate	Std. Error	df	t value	Pr(>|t|)	
(Intercept)	2.31	1.08	63.45	2.14	3.59E-02	*
Amb_CO2	-0.33	0.11	155.05	-2.99	3.26E-03	*
Hi_GT	-0.44	0.11	155.84	-3.96	1.15E-04	***
DTM	0.44	0.21	63.70	2.07	4.23E-02	*
Amb_CO2/Hi_GT	0.38	0.16	154.82	2.40	1.77E-02	*

Random Effects

	Variance	% Variance
ACC (Intercept)	0.02	5.88
Residual	0.26	94.12
Total	0.28	

Type III ANOVA

	Sum Sq	Mean Sq	NumDF	DenDF	F.value	Pr(>F)	
CO2	0.77	0.77	1.00	154.79	2.96	8.75E-02	
GT	2.27	2.27	1.00	158.08	8.71	3.64E-03	*
DTM	1.12	1.12	1.00	63.70	4.29	4.23E-02	*
CO2 x GT	1.50	1.50	1.00	154.82	5.75	1.77E-02	*

B. Panicle Weight

Fixed Effects

	Estimate	Std. Error	df	t value	Pr(>|t|)	
(Intercept)	5.53	1.56	31.79	3.54	1.26E-03	*
Amb_CO2	-0.32	0.18	155.83	-1.77	7.90E-02	
Hi_GT	-0.64	0.18	156.44	-3.52	5.69E-04	***
DTM	-0.28	0.31	31.38	-0.91	3.71E-01	
Amb_CO2/Hi_GT	0.82	0.26	155.34	3.14	2.03E-03	*

Random Effects

	Variance	% Variance
ACC (Intercept)	0.01	0.81
Residual	0.70	99.19
Total	0.70	

Type III ANOVA

	Sum Sq	Mean Sq	NumDF	DenDF	F.value	Pr(>F)	
CO2	0.35	0.35	1.00	155.54	0.50	4.81E-01	
GT	1.91	1.91	1.00	158.67	2.74	9.98E-02	
DTM	0.57	0.57	1.00	31.38	0.82	3.71E-01	
CO2 x GT	6.86	6.86	1.00	155.34	9.85	2.03E-03	*

Supplemental Table 5.6: Results of mixed models for JAPONICA group

A. Panicle Number

Fixed Effects

	Estimate	Std. Error	df	t value	Pr(>|t|)	
(Intercept)	2.51	1.09	119.50	2.29	2.35E-02	*
Amb_CO2	-0.15	0.12	120.95	-1.22	2.24E-01	
Hi_GT	-0.21	0.13	122.51	-1.61	1.10E-01	
DTM	0.21	0.22	122.49	0.94	3.50E-01	
Amb_CO2/Hi_GT	0.00	0.18	120.90	-0.01	9.90E-01	

Random Effects

	Variance	% Variance
ACC (Intercept)	0.07	21.04
Residual	0.26	78.96
Total	0.33	

Type III ANOVA

	Sum Sq	Mean Sq	NumDF	DenDF	F.value	Pr(>F)	
CO2	0.76	0.76	1.00	121.00	3.01	8.53E-02	
GT	1.13	1.13	1.00	123.55	4.47	3.66E-02	*
DTM	0.22	0.22	1.00	122.49	0.88	3.50E-01	
CO2 x GT	0.00	0.00	1.00	120.90	0.00	9.90E-01	

B. Panicle Weight

Fixed Effects

	Estimate	Std. Error	df	t value	Pr(>|t|)	
(Intercept)	6.39	1.75	105.71	3.65	4.03E-04	***
Amb_CO2	-0.16	0.20	120.70	-0.78	4.36E-01	
Hi_GT	-0.47	0.21	123.23	-2.21	2.89E-02	*
DTM	-0.66	0.36	108.80	-1.83	7.03E-02	
Amb_CO2/Hi_GT	0.17	0.29	120.63	0.60	5.50E-01	

Random Effects

	Variance	% Variance
ACC (Intercept)	0.09	12.29
Residual	0.68	87.71
Total	0.77	

Type III ANOVA

	Sum Sq	Mean Sq	NumDF	DenDF	F.value	Pr(>F)	
CO2	0.17	0.17	1.00	120.79	0.25	6.16E-01	
GT	3.76	3.76	1.00	124.53	5.57	1.98E-02	*
DTM	2.26	2.26	1.00	108.80	3.34	7.03E-02	
CO2 x GT	0.24	0.24	1.00	120.63	0.36	5.50E-01	

Supplemental Table 5.7: Data for gas exchange measurements

chamber	variety	IS	A at hi CO2	A at growth	run	gs growth	Vcmax 30C	Jmax 30
cool elev	Geumobyeo	0.268	37.900	19.500	1	265.000	358.395	196.330
cool elev	TeQing	0.488	47.700	27.600	1	322.000	645.604	247.055
cool elev	Nipponbare	0.162	35.100	23.100	1	337.000	218.806	181.838
cool elev	WAB	0.176	32.700	27.000	1	355.000	237.622	169.415
cool elev	DJ125	0.415	44.900	23.900	1	305.000	549.878	232.562
cool elev	Shirkati	0.227	35.700	20.700	1	265.000	304.573	184.943
cool elev	IR 64	0.169	37.200	26.300	1	333.000	227.865	192.707
cool elev	Jefferson	0.329	44.700	27.600	1	300.000	438.171	231.527
cool elev	STGS	0.400	39.600	26.800	1	301.000	530.500	205.130
cool elev	488-B	0.389	38.100	25.700	1	322.000	515.967	197.366
cool elev	Jeff/Nils	0.275	41.700	28.200	1	257.000	367.000	215.999
cool elev	NPT	0.350	48.100	28.700	1	264.000	465.100	249.126
warm elev	Geumobyeo	0.372	44.600	27.400	1	460.000	493.876	231.010
warm elev	TeQing	0.214	26.900	26.300	1	404.000	286.736	139.394
warm elev	Nipponbare	0.220	27.100	28.400	1	352.000	295.060	140.430
warm elev	WAB	0.389	33.700	23.800	1	305.000	515.967	174.591
warm elev	DJ125	0.294	41.600	28.100	1	395.000	392.433	215.482
warm elev	Shirkati	0.196	32.600	27.500	1	366.000	263.213	168.898
warm elev	IR 64	0.255	30.800	23.600	1	390.000	341.184	159.581
warm elev	Jefferson	0.176	34.200	25.600	1	375.000	236.867	177.179
warm elev	STGS	0.221	33.900	24.700	1	316.000	296.929	175.626
warm elev	488-B	0.552	35.500	25.300	1	354.000	729.814	183.908
warm elev	Jeff/Nils	0.094	18.200	25.400	1	388.000	130.051	94.363
warm elev	NPT	0.535	36.700	26.300	1	355.000	706.796	190.119
cool amb	Geumobyeo	0.229	29.400	24.600	1	275.000	306.271	152.334
cool amb	TeQing	0.429	30.700	20.800	1	255.000	568.432	159.063
cool amb	Nipponbare	0.230	39.100	22.600	1	320.000	308.140	202.542
cool amb	WAB	0.229	38.100	21.300	1	296.000	306.271	197.366
cool amb	DJ125	0.187	38.100	13.800	1	309.000	251.460	197.366
cool amb	Shirkati	0.364	45.500	21.300	1	325.000	482.936	235.668
cool amb	IR 64	0.196	38.400	23.000	1	296.000	263.334	198.918
cool amb	Jefferson	0.200	31.900	16.800	1	288.000	268.900	165.274
cool amb	STGS	0.191	23.800	19.200	1	270.000	257.009	123.349
cool amb	488-B	0.410	47.800	20.600	1	355.000	544.031	247.573
cool amb	Jeff/Nils	0.294	44.200	23.600	1	305.000	392.433	228.939
cool amb	NPT	0.483	47.200	24.300	1	305.000	639.064	244.467
warm amb	Nipponbare	0.339	40.700	21.800	1	362.000	451.227	210.823
warm amb	WAB	0.278	33.200	17.900	1	326.000	371.088	172.003
warm amb	DJ125	0.291	33.900	19.900	1	347.000	387.438	175.626
warm amb	Shirkati	0.334	34.200	20.700	1	250.000	444.663	177.179
warm amb	IR 64	0.500	36.000	22.900	1	377.000	661.300	186.496
warm amb	Jefferson	0.490	37.600	21.600	1	300.000	648.220	194.778
warm amb	STGS	0.198	29.300	19.900	1	300.000	265.710	151.817
warm amb	488-B	0.307	33.300	22.000	1	310.000	408.420	172.521

Supplemental Table 5.7 continued

warm amb	Jeff/Nils	0.184	31.600	21.300	1	360.000	248.553	163.722
warm amb	NPT	0.491	39.400	22.300	1	360.000	649.926	204.094
warm amb	Geumobyeo	0.251	30.400	19.400	1	362.000	336.068	157.510
warm amb	TeQing	0.267	36.100	23.700	1	367.000	356.100	187.014
cool elev	Geumobyeo	0.220	34.900	19.300	2	73.000	295.060	180.802
cool elev	TeQing	0.258	30.600	45.100	2	359.000	344.764	158.546
cool elev	Nipponbare	0.190	15.800	26.400	2	121.000	255.820	81.941
cool elev	WAB	0.186	24.700	37.400	2	287.000	250.588	128.007
cool elev	DJ125	0.243	15.700	35.400	2	104.000	325.144	81.423
cool elev	Shirkati	0.378	26.900	39.200	2	258.000	501.724	139.394
cool elev	IR 64	0.145	32.700	22.300	2	63.000	196.960	169.415
cool elev	Jefferson	0.350	45.000	28.600	2	250.000	465.100	233.080
cool elev	STGS	0.380	38.200	26.600	2	250.000	504.340	197.883
cool elev	488-B	0.400	39.000	26.100	2	275.000	530.500	202.024
cool elev	Jeff/Nils	0.282	41.900	28.500	2	235.000	376.156	217.034
cool elev	NPT	0.231	43.300	32.900	2	73.000	309.448	224.281
warm elev	Geumobyeo	0.213	41.400	27.800	2	332.000	285.904	214.446
warm elev	TeQing	0.265	46.700	32.200	2	391.000	353.920	241.879
warm elev	Nipponbare	0.203	32.200	23.300	2	199.000	272.824	166.827
warm elev	WAB	0.213	37.600	23.600	2	256.000	285.904	194.778
warm elev	DJ125	0.300	45.400	18.100	2	120.000	399.700	235.150
warm elev	Shirkati	0.319	41.300	24.100	2	210.000	424.552	213.929
warm elev	IR 64	0.281	25.500	18.300	2	139.000	374.848	132.148
warm elev	Jefferson	0.116	41.500	24.700	2	245.000	159.028	214.964
warm elev	STGS	0.250	17.900	16.500	2	95.000	334.300	92.810
warm elev	488-B	0.103	17.200	12.600	2	89.000	142.024	89.187
warm elev	Jeff/Nils	0.247	44.200	35.100	2	209.000	330.376	228.939
warm elev	NPT	0.221	46.300	31.800	2	474.000	296.368	239.809
cool amb	Geumobyeo	0.193	37.200	19.000	2	319.000	259.744	192.707
cool amb	TeQing	0.207	42.300	24.400	2	954.000	278.056	219.105
cool amb	Nipponbare	0.179	26.700	12.400	2	180.000	241.432	138.359
cool amb	WAB	0.281	37.400	17.100	2	380.000	374.848	193.742
cool amb	DJ125	0.205	21.700	13.200	2	151.000	275.440	112.479
cool amb	Shirkati	0.248	28.000	23.000	2	748.000	331.684	145.088
cool amb	IR 64	0.075	19.400	22.000	2	686.000	105.400	100.574
cool amb	Jefferson	0.216	36.000	20.300	2	609.000	289.828	186.496
cool amb	STGS	0.241	32.700	26.700	2	971.000	322.528	169.415
cool amb	488-B	0.268	29.500	23.100	2	503.000	357.844	152.852
cool amb	Jeff/Nils	0.283	41.800	21.900	2	565.000	377.464	216.517
cool amb	NPT	0.241	38.800	27.000	2	644.000	322.528	200.989
warm amb	Nipponbare	0.166	40.000	27.000	2	998.000	224.428	207.200
warm amb	WAB	0.169	29.800	24.900	2	863.000	228.352	154.405
warm amb	DJ125	0.171	23.400	15.000	2	218.000	230.968	121.278
warm amb	Shirkati	0.171	28.100	19.800	2	651.000	230.968	145.606
warm amb	IR 64	0.209	37.700	20.000	2	589.000	280.672	195.295
warm amb	Jefferson	0.187	40.000	22.700	2	530.000	251.896	207.200

Supplemental Table 5.7 continued

warm amb	STGS	0.197	37.500	22.300	2	833.000	264.976	194.260
warm amb	488-B	0.144	28.600	18.900	2	385.000	195.652	148.194
warm amb	Jeff/Nils	0.247	39.900	24.200	2	902.000	330.376	206.682
warm amb	NPT	0.153	35.900	25.900	2	880.000	207.424	185.978
warm amb	Geumobyeo	0.187	30.400	18.700	2	325.000	251.896	157.510
warm amb	TeQing	0.159	40.900	26.700	2	680.000	215.272	211.858
cool elev	Geumobyeo			20.900	0	646.000		
cool elev	TeQing			25.900	0	383.000		
cool elev	Nipponbare			28.200	0	583.000		
cool elev	WAB			25.900	0	540.000		
cool elev	DJ125			25.200	0	452.000		
cool elev	Shirkati			21.800	0	591.000		
cool elev	IR 64			23.300	0	641.000		
cool elev	Jefferson			27.500	0	430.000		
cool elev	STGS			26.100	0	414.000		
cool elev	488-B			24.700	0	414.000		
cool elev	Jeff/Nils			26.600	0	523.000		
cool elev	NPT			28.200	0	590.000		
warm elev	Geumobyeo			27.600	0	558.000		
warm elev	TeQing			23.000	0	585.000		
warm elev	Nipponbare			28.700	0	532.000		
warm elev	WAB			23.700	0	650.000		
warm elev	DJ125			28.300	0	514.000		
warm elev	Shirkati			26.900	0	449.000		
warm elev	IR 64			21.500	0	349.000		
warm elev	Jefferson			25.200	0	590.000		
warm elev	STGS			22.300	0	444.000		
warm elev	488-B			23.900	0	451.000		
warm elev	Jeff/Nils			25.600	0	939.000		
warm elev	NPT			25.200	0	595.000		
cool amb	Geumobyeo			24.600	0	642.000		
cool amb	TeQing			18.000	0	611.000		
cool amb	Nipponbare			22.100	0	789.000		
cool amb	WAB			21.300	0	542.000		
cool amb	DJ125			21.100	0	570.000		
cool amb	Shirkati			19.900	0	446.000		
cool amb	IR 64			20.700	0	615.000		
cool amb	Jefferson			17.700	0	582.000		
cool amb	STGS			17.900	0	660.000		
cool amb	488-B			16.900	0	373.000		
cool amb	Jeff/Nils			21.900	0	936.000		
cool amb	NPT			23.300	0	802.000		
warm amb	Nipponbare			18.200	0	558.000		
warm amb	WAB			15.900	0	590.000		
warm amb	DJ125			16.600	0	450.000		
warm amb	Shirkati			18.200	0	537.000		

Supplemental Table 5.7 continued

warm amb	IR 64	22.200	0	414.000
warm amb	Jefferson	15.800	0	543.000
warm amb	STGS	17.000	0	467.000
warm amb	488-B	22.200	0	539.000
warm amb	Jeff/Nils	17.800	0	630.000
warm amb	NPT	18.900	0	416.000
warm amb	Geumobyeo	17.300	0	558.000
warm amb	TeQing	22.500	0	541.000

Supplemental Table 5.8: Results of linear regression for photosynthesis-related traits: A) Single-leaf photosynthetic rate and B) stomatal conductance.

A. Photosynthetic rate

Fixed Effects

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	20.388	0.723	28.208	<2E-16	***
Hi_CO2	7.133	1.022	6.979	1.43E-10	***
Hi_GT	0.067	1.022	0.065	9.48E-01	
Amb_CO2/Hi_GT	-3.221	1.446	-2.228	2.76E-02	*

ANOVA

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
CO2	1	1006.517	1006.517	58.388	4.44E-12	***
GT	1	78.664	78.664	4.563	3.46E-02	*
CO2 x GT	1	85.604	85.604	4.966	2.76E-02	*
Residuals	128	2206.525	17.238			

B. Stomatal conductance

Fixed Effects

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	488.667	32.414	15.076	<2E-16	***
Hi_CO2	-141.576	45.841	-3.088	2.47E-03	**
Hi_GT	10.182	45.841	0.222	8.25E-01	
Amb_CO2/Hi_GT	23.061	64.828	0.356	7.23E-01	

ANOVA

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
CO2	1	558090.070	558090.070	16.096	0.0001	***
GT	1	15556.730	15556.730	0.449	0.504	
CO2 x GT	1	4387.280	4387.280	0.127	0.723	
Residuals	128	4438063.640	34672.370			

REFERENCES

Amthor JS (2001) Effects of atmospheric CO_2 concentration on wheat yield: review of results from experiments using various approaches to control CO_2 concentration. F Crop Res 73: 1–34

Baker JT (2004) Yield responses of southern US rice cultivars to CO_2 and temperature. Agric For Meteorol 122: 129–137

Baker, J.T., K.J. Boote, and L.H. Allen, Jr. 1995. Potential climate change effects on rice: Carbon dioxide and temperature. p. 31–47. In C. Rosenzweig et al.(ed.) Climate change and agriculture: Analysis of potential international impacts. ASA Spec. Publ. 59. ASA, CSSA, and SSSA, Madison,WI

Baker JT, Allen LH, Boote KJ (1992) Temperature effects on rice at elevated CO_2 concentration. J Exp Bot 43: 959–964

Baker JT, Allen LH, Boote KJ, Jones P, Jones JW (1989) Response of soybean to air temperature and carbon dioxide concentration. Crop Sci 29: 98–105

Cai C, Yin X, He S, Jiang W, Si C, Struik PC, Luo W, Li G, Xie Y, Xiong Y (2015) Responses of wheat and rice to factorial combinations of ambient and elevated CO_2 and temperature in FACE experiments. Glob. Chang. Biol.

Furuta T, Uehara K, Angeles-Shim RB, Shim J, Ashikari M, Takashi T (2014) Development and evaluation of chromosome segment substitution lines (CSSLs) carrying chromosome segments derived from Oryza rufipogon in the genetic background of Oryza sativa L. Breed Sci 63: 468

Garris AJ, Tai TH, Coburn J, Kresovich S, McCouch S (2005) Genetic structure and diversity in Oryza sativa L. Genetics 169: 1631–1638

Gealy DH, Agrama H, Jia MH (2012) Genetic Analysis of Atypical U.S. Red Rice Phenotypes: Indications of Prior Gene Flow in Rice Fields? Weed Sci 60: 451–461

Hasegawa T, Sakai H, Tokida T, Nakamura H, Zhu C, Usui Y, Yoshimoto M, Fukuoka M, Wakatsuki H, Katayanagi N (2013) Rice cultivar responses to

elevated CO_2 at two free-air CO_2 enrichment (FACE) sites in Japan. Funct Plant Biol 40: 148–159

Huang X, Kurata N, Wei X, Wang Z-X, Wang A, Zhao Q, Zhao Y, Liu K, Lu H, Li W, et al (2012) A map of rice genome variation reveals the origin of cultivated rice. Nature 490: 497–501

Imai I, Kimball JA, Conway B, Yeater KM, McCouch SR, McClung A (2013) Validation of yield-enhancing quantitative trait loci from a low-yielding wild ancestor of rice. Mol Breed 32: 101–120

Kimball BA, Kobayashi K, Bindi M (2002) Responses of agricultural crops to free-air CO_2 enrichment. Adv Agron 77: 293–368

Madan P, Jagadish SVK, Craufurd PQ, Fitzgerald M, Lafarge T, Wheeler TR (2012) Effect of elevated CO2 and high temperature on seed-set and grain quality of rice. J Exp Bot 63: 3843–3852

Matsui T, Namuco OS, Ziska LH, Horie T (1997) Effects of high temperature and CO_2 concentration on spikelet sterility in indica rice. F Crop Res 51: 213–219

Matsui T, Omasa K, Horie T (2000). High temperature at flowering inhibits swelling of pollen grains, a driving force for thecae dehiscence in rice (Oryza sativa L.). Plant Prod Sci 1: 430-434.

Mitchell RAC, Mitchell VJ, Driscoll SP, Franklin J, Lawlor DW (1993) Effects of increased CO_2 concentration and temperature on growth and yield of winter wheat at two levels of nitrogen application. Plant Cell Environ 16: 521–529

Moya TB, Ziska LH, NAMUCO OS, OLSZYK D (1998) Growth dynamics and genotypic variation in tropical, field-grown paddy rice (Oryza sativa L.) in response to increasing carbon dioxide and temperature. Glob Chang Biol 4: 645–656

Poorter H, Navas M (2003) Plant growth and competition at elevated CO_2: on winners, losers and functional groups. New Phytol 157: 175–198

Shimono H (2011) Rice genotypes that respond strongly to elevated CO_2 also respond

strongly to low planting density. Agric Ecosyst Environ 141: 240–243

Shimono H, Okada M, Yamakawa Y, Nakamura H, Kobayashi K, Hasegawa T (2009) Genotypic variation in rice yield enhancement by elevated CO_2 relates to growth before heading, and not to maturity group. J Exp Bot 60: 523–532

Shimono H, Ozaki Y, Jagadish KS V, Sakai H, Usui Y, Hasegawa T, Kumagai E, Nakano H, Yoshinaga S (2014) Planting geometry as a pre-screening technique for identifying CO_2 responsive rice genotypes: a case study of panicle number. Physiol Plant 152: 520–528

Thomson MJ, Tai TH, McClung AM, Lai XH, Hinga ME, Lobos KB, Xu Y, Martinez CP, McCouch SR (2003) Mapping quantitative trait loci for yield, yield components and morphological traits in an advanced backcross population between Oryza rufipogon and the Oryza sativa cultivar Jefferson. Theor Appl Genet 107: 479–493

Wang J, Wang C, Chen N, Xiong Z, Wolfe D, Zou J (2015) Response of rice production to elevated [CO2] and its interaction with rising temperature or nitrogen supply: a meta-analysis. Clim Change 130: 529–543

Zhao K, Tung C-W, Eizenga GC, Wright MH, Ali ML, Price AH, Norton GJ, Islam MR, Reynolds A, Mezey J (2011) Genome-wide association mapping reveals a rich genetic architecture of complex traits in Oryza sativa. Nat Commun 2: 467

Zhu C, Xu X, Wang D, Zhu J, Liu G (2015) An indica rice genotype showed a similar yield enhancement to that of hybrid rice under free air carbon dioxide enrichment. Sci. Rep. 5:

Ziska LH, Bunce JA, Shimono H, Gealy DR, Baker JT, Newton PCD, Reynolds MP, Jagadish KS V, Zhu C, Howden M (2012a) Food security and climate change: on the potential to adapt global crop production by active selection to rising atmospheric carbon dioxide. Proc R Soc London B Biol Sci rspb20121005

Ziska LH, Gealy DR, Tomecek MB, Jackson AK, Black HL (2012b) Recent and projected increases in atmospheric CO2 concentration can enhance gene flow

between wild and genetically altered rice (Oryza sativa). PLoS One 7: e37522

Ziska LH, Manalo PA, Ordonez RA (1996) Intraspecific variation in the response of rice (Oryza sativa L.) to increased CO2 and temperature: growth and yield response of 17 cultivars. J Exp Bot 47: 1353–1359

Ziska LH, McClung A (2008) Differential Response of Cultivated and Weedy (Red) Rice to Recent and Projected Increases in Atmospheric Carbon Dioxide. Agron J 100: 1259

Ziska LH, Tomecek MB, Gealy DR (2014) Assessment of cultivated and wild, weedy rice lines to concurrent changes in CO2 concentration and air temperature: determining traits for enhanced seed yield with increasing atmospheric CO2. Funct Plant Biol 41: 236–243

Appendix A: Accumulation of stem non-structural carbohydrate in a New Plant Type rice accession, IR78049-25-2-2-2: Effects of plant density, CO_2, and temperature

INTRODUCTION

Ideotype breeding to increase rice yield potential is a strategy that supplements empirical selection for yield. In contrast to empirical breeding, the ideotype approach is a prescriptive method whereby plants are bred to capture a model set of morphological and architectural characteristics (e.g., erect leaves and moderate tillering capability) chosen based on plant physiological models to optimize yield (Jennings 1964, Donald 1968). The 1960s saw a change in the plant architecture of tropical rice from tall landraces to short-statured, semi-dwarf varieties responsive to fertilizer inputs, resulting in enhanced harvest index and dramatic yield boosts. In the late 1980s, concepts for further improvement of rice yield potential based on physiological and morphological changes became known as the New Plant Type (NPT), and breeding efforts were placed into generating lines that had low tillering capacity, thick stems with dark-green leaves, and a target of 200-250 spikelets/panicle (Cassman, 1994). *Tropical japonica* rice varieties were used as parents for these crosses and ideotype lines were produced by the early 1990s. Unfortunately, poor grain-filling led to low yields and these First Generation NPT lines could not be released (Peng et al., 2008). Instead, these lines are valued as donors of preferred traits such as strong stems and stay-green physiology (personal communication). Here, we evaluate a breeder-favored NPT line (IR78049-25-2-2-2) from the International Rice Research Institute for its ability to accumulate non-structural carbohydrates (NSC)

under varying planting density, CO_2, and growth temperature.

METHODS

Plant material and genotyping: NPT IR78049-25-2-2-2 (IRIS 18-335212) seeds were imported from the International Rice Research Institute. Pedigree information on the NPT was extracted from the International Rice Information System (IRIS). 54 diverse *O. sativa* from Rice Diversity Panel 1 (RDP1) (Zhao et al., 2011) from Chapter 3 of this dissertation were used as controls in genetic analyses. Genome-wide Single Nucleotide Polymorphism (SNP) information for the NPT and *O. sativa* controls were obtained using Genotyping-By-Sequencing technology (Elshire et al., 2011). Raw data were analyzed and SNPs extracted using the TASSEL4.0 pipeline. A total of 3,446 SNPs were selected based on perfect call rate across all 55 individuals and used for genetic analysis.

Plant density experiment: NPT individuals were grown in Percival growth chambers at Cornell University (12h day-length; 29^{o}C-day/21 ^{o}C -night) under three plant density treatments: high (7.4 cm spacing measured from the center of one pot to the center of the next), medium (10.5 cm spacing), and low (14.8 cm spacing). Under each condition, identical deep pots (6.5 cm diameter x 30 cm deep) were used in order to separate the effect of available underground space from the effect of aboveground plant density. 20 biological replicates of the NPT per density treatment were grown in racks (4 x 5 pot formation) at the appropriate spacing and individual plants were rotated weekly prior to flowering to minimize border effects. Whole plants were sampled at flowering and stems separated as described in the following section. We

measured 3 morphological traits (tiller number at heading stage, plant height, total stem dry weight), scored days to heading, and recorded tillering dynamics during the first 70 days after sowing (DAS).

CO_2 and temperature experiment: The NPT was evaluated under growth chamber conditions along with 11 other diverse rice accessions that were grown for a separate study. Briefly, the rice panel was grown under two CO_2 concentrations (400 ppm and 600 ppm) and four temperature regimes (29°C day/21°C night; 29°C day/21°C night with additional heat stress at anthesis; 34°C day/26°C night; and 34°C day/26°C night with additional heat stress at anthesis). A suite of traits (five yield components, five growth phenotypes, two photosynthesis measurements, and one phenological trait) was evaluated on all accessions. Specifics on chamber conditions and trait measurements are detailed previously (Wang et al., 2016). In addition to traits described above, extra replicates of the NPT were grown under each treatment combination for NSC determination. Whole plants were sampled at the flowering stage and all stems were separated from the leaf blades, panicles, and roots. The entire collective stem biomass for an individual plant was ground, subsampled, and used for NSC determination.

NSC determination: Stem material from each individual was ground in two stages, first with a coffee grinder and then with an Udy mill, to a fine powder. Samples were mixed thoroughly and then subsampled for three separate replicates per individual. Sugars were extracted using 80% EtOH (and sucroses broken down via invertase) and the remaining starches gelatinized and digested into glucose using amyloglucosidase/α-amylase. All carbohydrates were assayed utilizing peroxidase and

glucose oxidase via the Trinder reaction and compared against known standards of glucose and sucrose at 490nm using a 96-well plate-reading UV-VIS spectrophotometer (Trinder 1969). Three technical replicates for the assay were used per sample.

Statistical analysis: For genetic analyses, Principal Components Analysis was performed on the SNP correlation matrix using R. A consensus neighbor-joining (NJ) tree resulting from 100 bootstrap replications was generated using Geneious 7.0.4.

RESULTS

In this appendix, results are summarized as figures and tables on the following pages.

Figure A1. Genetic relationship of NPT IR78049-25-2-2-2 and diverse *O. sativa*: A) Pedigree of NPT IR78049-25-2-2-2 (purple font indicates *tropical japonica* parent and blue indicates *temperate japonica* parent). B) PC2 versus PC4 plot of a Principal Components Analysis using 3,446 SNPs on the NPT and 54 diverse rice accessions across *O. sativa*'s 5 subpopulations. Colors represent previously known subpopulation membership (see legend). PC2 separates *indica* and *aus* of the *Indica* Varietal Group and PC4 pulls apart *temperate japonica* from *tropical japonica*. The NPT is indicated by the black asterisk near the middle of the plot. C) Neighbor-joining (NJ) tree of the *Japonica* Varietal Group excluding *aromatics* (10 *temperate japonica*, 10 *tropical japonica*, 1 admixed *Japonica*, and the NPT) based on 100 bootstrap replications.

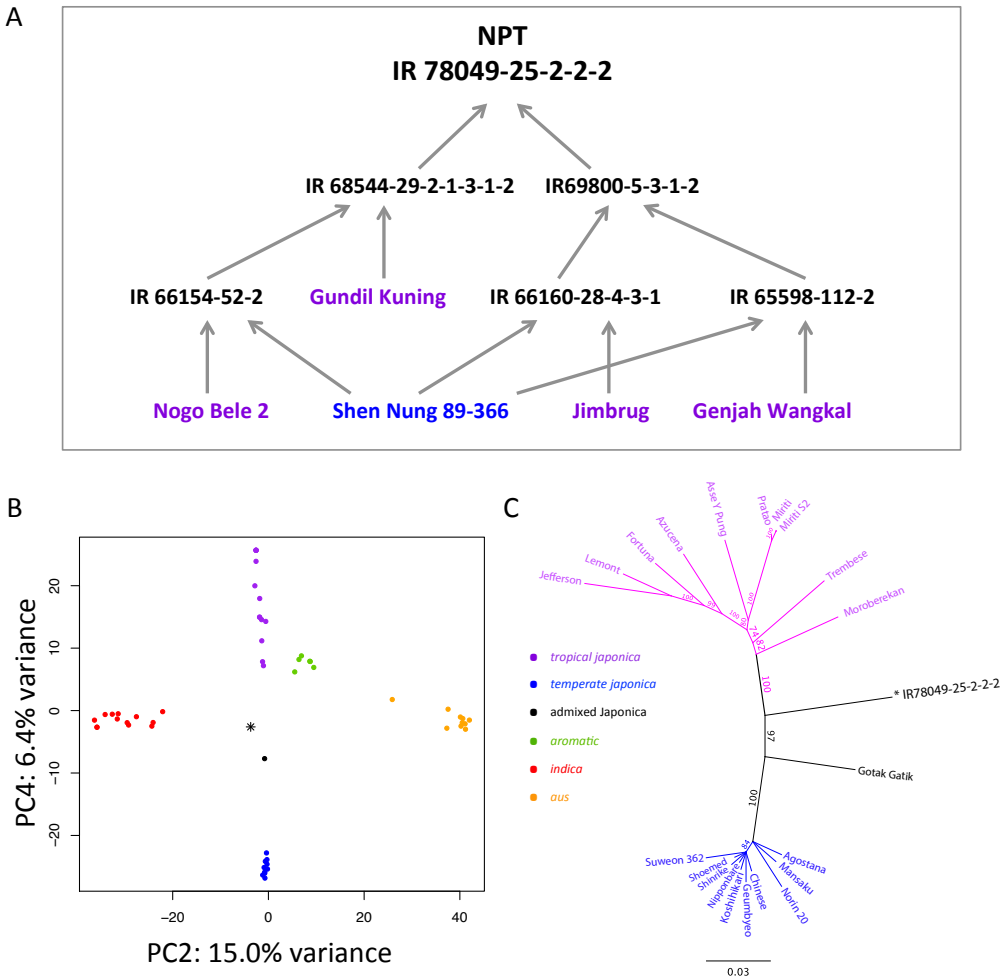


Figure A2. Overview of aboveground plant density experiment: A) Tillering rate of high, medium, and low density NPT plants (n=20 per density), B) Proportion of stem starch at heading stage of high, medium, and low density NPT plants, and C) Summary of results from density experiment. Photos were taken (57 days after sowing) of NPT individuals planted at High, Medium, and Low densities (left to right). Bar represents 25cm. *Plant angle was observed visually but not tested formally.

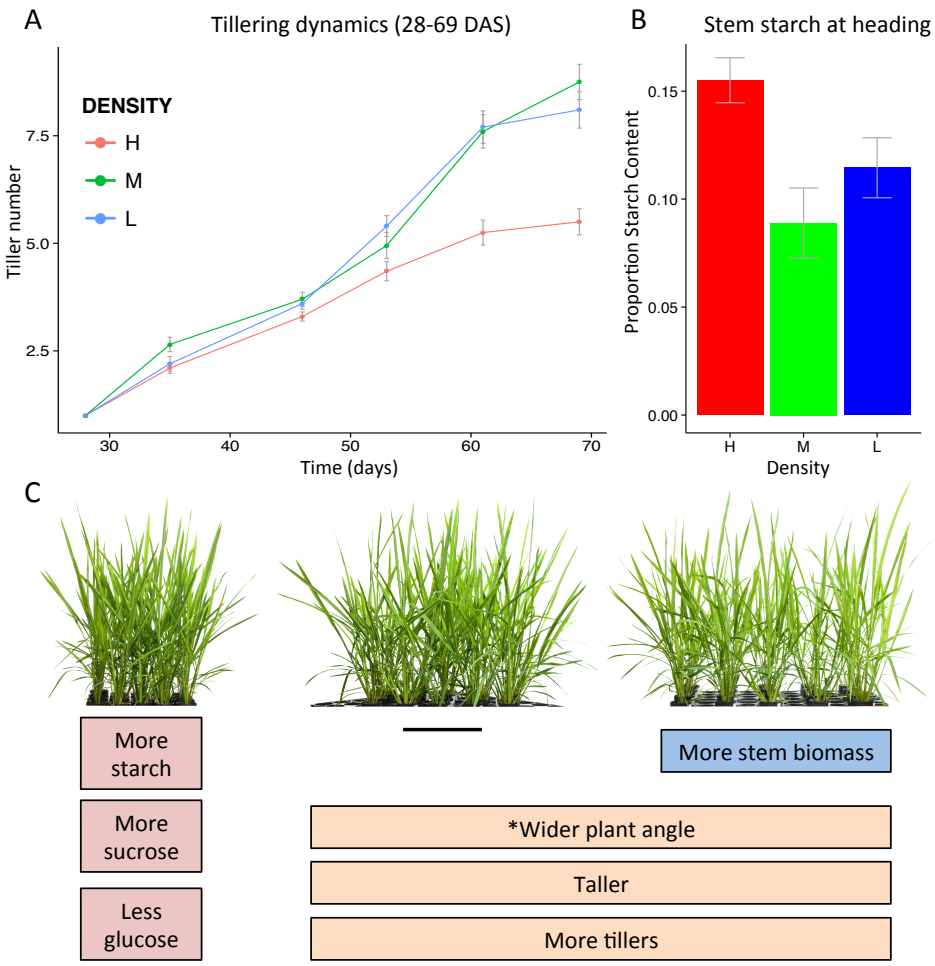


Figure A3. NPT response to varying CO_2 and temperature: A) Interaction plot of stem total non-structural carbohydrates (TNC) at anthesis under varying CO_2 and growth temperature. Traces indicate CO2 condition (blue= 400 ppm, red= 600 ppm) and x-axis displays the growth temperature (Norm GT = 29 ºC day/21ºC night, High GT = 34 ºC day/26 ºC night). Asterisk indicates significant CO2 x GT interaction. B) Mean single-leaf photosynthetic rates at two CO_2 levels and 2 growth temperatures (gray bars: s.e. of the mean).

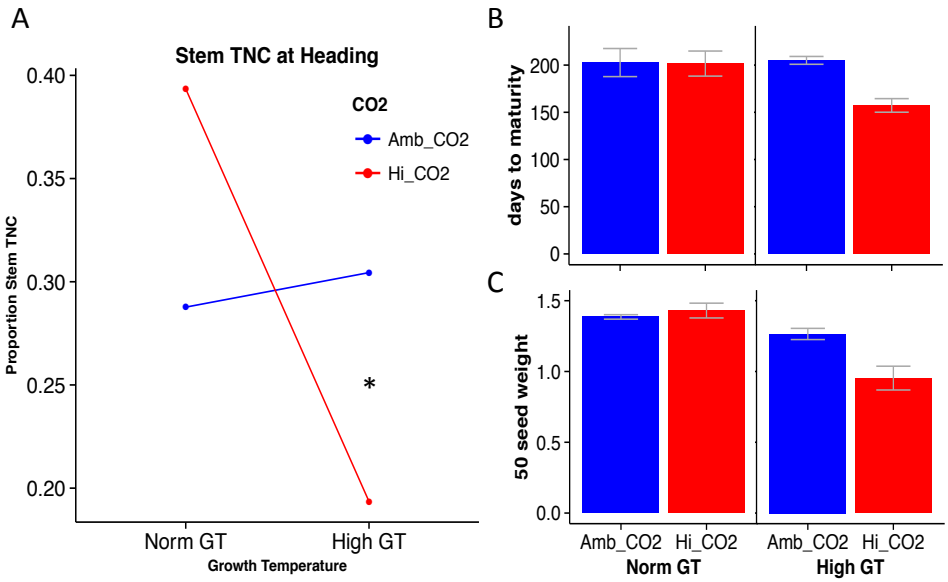


Figure A4. Additional PCA results: A) Distribution of the 3,446 SNP markers generated using Genotyping-By-Sequencing, B) Scree plot from Principal Components Analysis, and PCA results: C) PC1 versus PC3 plot and D) PC1 versus PC2 plot.

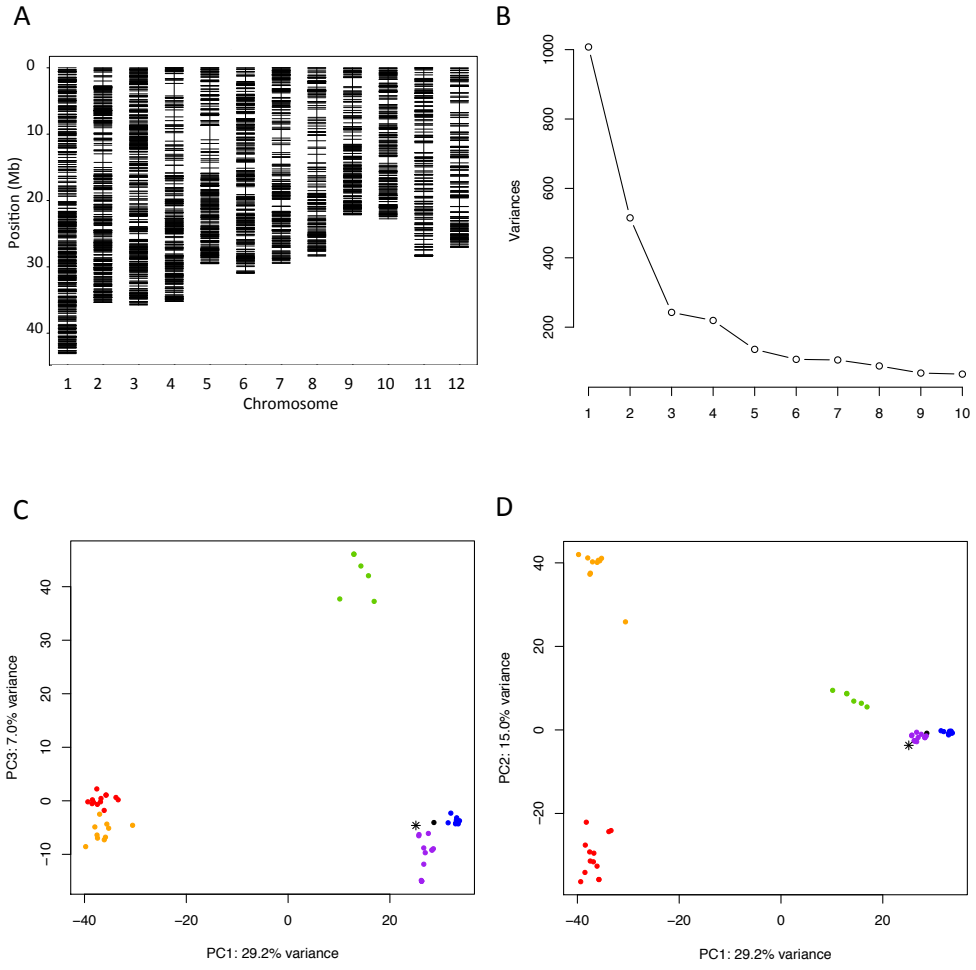


Figure A5. Trait associations from plant density experiment. A) Trait correlation matrix. Asterisk indicates asymptotic p-value < 0.05. B) One representative individual from the three planting densities: High, Medium, and Low (left to right). Bar indicates 25 centimeters. Lower densities were observed to have a wider plant angle.

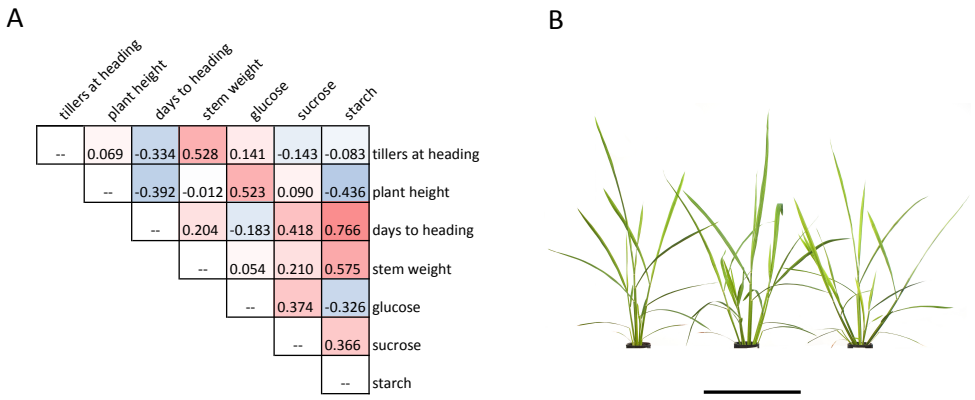


Table A1. MANCOVA and ANOVA results from density experiment: A) Tests for MANCOVA significance with starch, sucrose, and glucose as the response and B) univariate ANOVA tests.

A

	df	Pillai's Trace	Approx F	num df	den df	Pr(>F)	
DENSITY	2	0.56522	7.2223	6	110	1.63E-06	***
DTH	1	0.61316	28.5307	3	54	3.44E-11	***
Residuals	56						

	df	Wilk's Lambda	Approx F	num df	den df	Pr(>F)	
DENSITY	2	0.47059	8.2393	6	108	2.38E-07	***
DTH	1	0.38684	28.5307	3	54	3.44E-11	***
Residuals	56						

	df	Hotelling-Lawley Trace	Approx F	num df	den df	Pr(>F)	
DENSITY	2	1.0489	9.2653	6	106	3.72E-08	***
DTH	1	1.585	28.5307	3	54	3.44E-11	***
Residuals	56						

	df	Roy's Greatest Root	Approx F	num df	den df	Pr(>F)	
DENSITY	2	0.9705	17.792	3	55	3.40E-08	***
DTH	1	1.585	28.531	3	54	3.44E-11	***
Residuals	56						

B

Response: starch

	df	Sum Sq	Mean Sq	F value	Pr(>F)	
DENSITY	2	0.044413	0.022207	14.549	8.16E-06	***
DTH	1	0.128305	0.128305	84.058	9.65E-13	***
Residuals	56	0.085477	0.001526			

Response: sucrose

	df	Sum Sq	Mean Sq	F value	Pr(>F)	
DENSITY	2	0.0002782	0.0001391	0.7309	0.486023	
DTH	1	0.002922	0.00292201	15.353	0.000245	***
Residuals	56	0.010658	0.00019032			

Response: glucose

	df	Sum Sq	Mean Sq	F value	Pr(>F)	
DENSITY	2	1.65E-04	8.25E-05	17.4812	1.26E-06	***
DTH	1	9.49E-07	9.49E-07	0.2009	0.6557	
Residuals	56	2.64E-04	4.72E-06			

Table A2. Results of linear regression of stem and leaf NSC traits on CO2, GT, and CO2 x GT: A) stem TNC, B) stem starch, C) stem sucrose, D) leaf starch, E) leaf sucrose, F) leaf TNC. Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.'

A

Fixed Effect	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	0.288	0.049	5.826	2.02E-05	***
High CO2	0.106	0.064	1.659	1.15E-01	
High GT	0.017	0.064	0.262	7.96E-01	
High CO2 x High GT	-0.217	0.087	-2.478	2.40E-02	*

B

Fixed Effect	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	0.267	0.047	5.720	2.50E-05	***
High CO2	0.073	0.060	1.221	2.39E-01	
High GT	-0.007	0.060	-0.113	9.11E-01	
High CO2 x High GT	-0.171	0.082	-2.070	5.40E-02	.

C

Fixed Effect	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	0.015	0.006	2.252	0.0378	*
High CO2	0.034	0.008	4.091	0.00076	***
High GT	0.026	0.008	3.130	0.0061	**
High CO2 x High GT	-0.047	0.012	-4.095	0.00075	***

D

Fixed Effect	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	0.020	0.013	1.485	0.15581	
High CO2	0.046	0.017	2.702	0.01513	*
High GT	0.016	0.017	0.924	0.36849	
High CO2 x High GT	-0.056	0.024	-2.369	0.02997	*

E

Fixed Effect	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	0.054	0.006	9.010	6.98E-08	***
High CO2	-0.023	0.008	-2.934	9.27E-03	**
High GT	-0.023	0.008	-2.936	9.23E-03	**
High CO2 x High GT	0.016	0.011	1.524	1.46E-01	

F

Fixed Effect	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	0.079	0.016	4.849	0.00015	***
High CO2	0.021	0.021	0.975	0.34303	
High GT	-0.008	0.021	-0.382	0.70712	
High CO2 x High GT	-0.038	0.029	-1.310	0.20773	

Figure A6. Interaction plots of stem and leaf NSC traits: Traces indicate CO2 condition (blue= 400 ppm, red= 600 ppm) and x-axis displays the growth temperature (Norm GT = 29 ºC day/21ºC night, High GT = 34 ºC day/26 ºC night). Y axes represent A) leaf starch, B) stem starch, C) leaf sucrose, and D) stem sucrose. Asterisk represents significance of the CO2 x GT interaction effect: Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.'

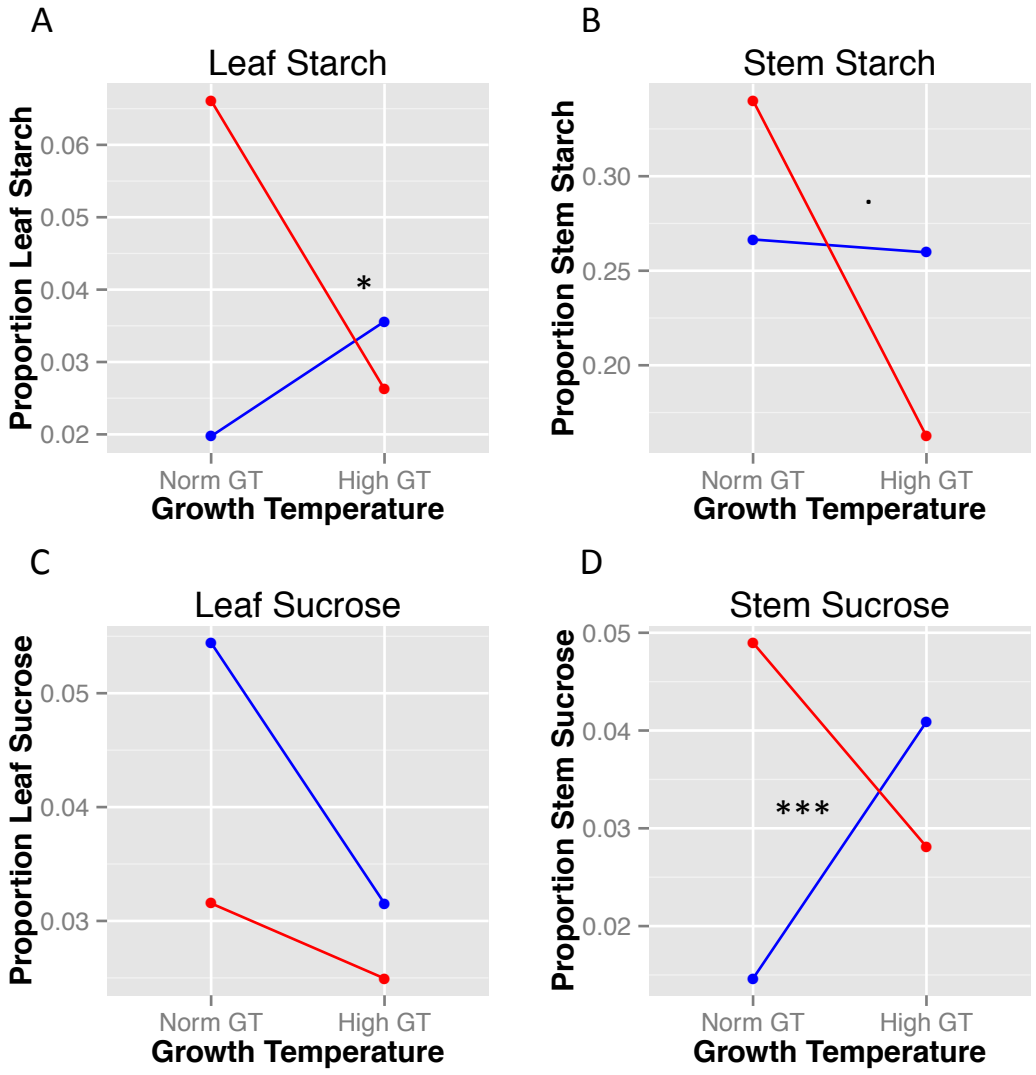


Table A3. Results of linear regression of days to maturity and 50 seed weight on CO2, GT, and CO2 x GT: A) Days to maturity and B) 50 seed weight. Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.'

A

Fixed Effect	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	1.265	0.058	21.755	2.16E-15	***
High CO2	-0.312	0.079	-3.934	8.22E-04	***
Norm GT	0.121	0.082	1.465	1.58E-01	
High CO2 x Norm GT	0.357	0.117	3.045	6.40E-03	**

B

Fixed Effect	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	205.000	10.348	19.810	1.30E-14	***
High CO2	-47.714	14.102	-3.383	2.95E-03	**
Norm GT	-2.333	14.634	-0.159	8.75E-01	
High CO2 x Norm GT	46.648	20.843	2.238	3.68E-02	*

REFERENCES

Cassman KG (1994) Breaking the Yield Barrier: Proceedings of a workshop on rice yield potential in favorable environments. International Rice Research Institute

Donald CM (1968) The breeding of crop ideotypes. Euphytica 17: 385–403

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS One 6: e19379

Peter R Jennings (1964) Plant Type as a Rice Breeding Objective. Crop Science 4:13-15

Peng S, Khush GS, Virk P, Tang Q, Zou Y (2008) Progress in ideotype breeding to increase rice yield potential. F Crop Res 108: 32–38

Wang DR, Bunce JA, Tomecek MB, Gealy D, McClung A, McCouch SR, Ziska LH (2016) Evidence for divergence of response in Indica, Japonica, and wild rice to high CO2 x temperature interaction. Glob Chang Biol 22: 2620–2632

Zhao K, Tung C-W, Eizenga GC, Wright MH, Ali ML, Price AH, Norton GJ, Islam MR, Reynolds A, Mezey J (2011) Genome-wide association mapping reveals a rich genetic architecture of complex traits in Oryza sativa. Nat Commun 2: 467

Appendix B: Preliminary analytical chemistry results of biomass carbohydrates in a New Plant Type rice accession, IR78049-25-2-2-2

Description: New Plant Type samples evaluated for NSCs in Appendix A were sent to National Renewable Energy Lab where they were processed for a suite of both structural and non-structural biomass constituents. Results are presented in the tables of this appendix. Once data are finalized, we aim to address the following questions:

- How do CO_2, temperature, and CO_2 x temperature affect relative contribution of biomass constituents in the NPT leaf and stem?
- How do CO_2, temperature, and CO_2 x temperature affect the relationship between structural and non-structural carbohydrate components of the NPT stem and leaf?

Table B1. Metadata on samples evaluated for biomass constituents. Conditions are fully described in Chapter 5 of this dissertation.

Sample	Organ	CO_2	Temperature
CO2_Tm_1_leaf	leaf	Current	High
CO2_Tm_1_stem	stem	Current	High
CO2_Tm_10_leaf	leaf	High	High
CO2_Tm_10_stem	stem	High	High
CO2_Tm_11_leaf	leaf	Current	High
CO2_Tm_11_stem	stem	Current	High
CO2_Tm_12_leaf	leaf	High	Average
CO2_Tm_12_stem	stem	High	Average
CO2_Tm_13_leaf	leaf	Current	Average
CO2_Tm_13_stem	stem	Current	Average
CO2_Tm_14_leaf	leaf	High	High
CO2_Tm_14_stem	stem	High	High
CO2_Tm_15_leaf	leaf	High	High
CO2_Tm_15_stem	stem	High	High
CO2_Tm_16_leaf	leaf	High	Average
CO2_Tm_16_stem	stem	High	Average
CO2_Tm_17_leaf	leaf	High	Average
CO2_Tm_17_stem	stem	High	Average
CO2_Tm_18_leaf	leaf	Current	Average
CO2_Tm_18_stem	stem	Current	Average
CO2_Tm_19_leaf	leaf	Current	Average
CO2_Tm_19_stem	stem	Current	Average
CO2_Tm_2_leaf	leaf	Current	High
CO2_Tm_2_stem	stem	Current	High
CO2_Tm_20_leaf	leaf	Current	High
CO2_Tm_20_stem	stem	Current	High
CO2_Tm_21_leaf	leaf	Current	High
CO2_Tm_21_stem	stem	Current	High
CO2_Tm_3_leaf	leaf	High	Average
CO2_Tm_3_stem	stem	High	Average
CO2_Tm_4_leaf	leaf	High	Average
CO2_Tm_4_stem	stem	High	Average
CO2_Tm_5_leaf	leaf	High	High
CO2_Tm_5_stem	stem	High	High

CO2_Tm _6_leaf	leaf	High	Average
CO2_Tm _6_stem	stem	High	Average
CO2_Tm _7_leaf	leaf	High	High
CO2_Tm _7_stem	stem	High	High
CO2_Tm _8_leaf	leaf	Current	Average
CO2_Tm _8_stem	stem	Current	Average
CO2_Tm _9_leaf	leaf	Current	High
CO2_Tm _9_stem	stem	Current	High

Table B2. Preliminary results of biomass analysis. Blue cells denote technical triplicates and green cells mark control samples. Samples with red text did not pass NREL's QC protocol and will be re-analyzed.

Master Ref		Sample Description	%Total Ash	%Structural Inorganics	%Non-structural Inorganics	%Total Protein	%Structural Protein	%Non-structural Protein	% Sucrose	% Free Glucose	%Free Sucrose	%Water Extractable Others	%Ethanol Extractives	% Lignin	% Glucan	% Xylan	% Galactan	% Arabinan	% Fructan	Acetyl	Total %
BATCH 1																					
1		CO2_TM-1 STEM	10.99	2.35	8.64	8.37	3.48	4.90	2.39	1.09	2.18	8.16	2.68	7.98	35.28	11.04	1.01	2.32	0.00	1.52	95.00
2		CO2_TM-1 STEM	10.99	2.27	8.72	8.00	3.48	4.53	2.39	1.13	1.70	9.56	2.61	7.83	35.26	11.16	1.07	2.33	0.00	1.42	95.46
3		CO2_TM-1 STEM	10.99	2.44	8.55	8.23	3.70	4.54	2.37	1.13	1.56	9.08	2.87	7.88	35.80	11.01	1.05	2.32	0.00	1.47	95.77
4		CO2_TM-1 LEAF	12.79	3.97	8.81	14.86	11.05	3.81	3.36	1.58	1.75	7.93	4.88	10.91	20.95	10.20	0.93	2.35	0.00	0.99	93.48
5		CO2_TM-2 STEM	11.30	1.77	9.52	7.50	2.89	4.60	3.21	1.62	2.02	8.19	2.48	8.47	33.64	11.71	0.98	2.21	0.00	1.59	94.90
6		CO2_TM-2 LEAF	10.39	3.67	6.72	13.52	10.22	3.30	3.87	1.69	1.69	6.96	4.77	11.69	22.95	10.72	1.01	2.60	0.00	1.03	92.90
7		CO2_TM-3 STEM	11.19	1.79	9.40	6.62	2.67	3.96	3.74	1.19	1.55	9.74	2.24	7.56	36.94	10.26	1.04	2.13	0.00	1.54	95.73
8		CO2_TM-3 LEAF	12.61	3.49	9.12	13.85	9.79	4.05	2.38	1.25	1.32	6.95	4.89	10.31	24.24	11.57	0.94	2.47	0.00	1.24	94.01
9		CO2_TM-3 LEAF	12.61	3.25	9.36	14.35	10.01	4.34	1.78	1.53	1.57	6.24	4.95	10.16	24.73	11.85	0.98	2.54	0.00	1.29	94.58
10		CO2_TM-3 LEAF	12.61	3.38	9.24	14.40	9.45	4.95	1.90	1.49	1.56	5.59	4.91	10.26	24.00	11.54	0.94	2.48	0.00	1.17	92.86
11		CO2_TM-4 STEM	10.54	1.90	8.64	6.53	2.81	3.72	3.25	1.50	1.63	10.33	2.29	7.70	36.67	10.36	0.98	2.07	0.00	1.45	95.28
12		CO2_TM-4 LEAF	13.60	3.72	9.88	14.31	8.97	5.34	3.00	1.50	1.49	5.75	4.15	10.54	23.37	10.72	0.99	2.48	0.00	1.14	93.04
13		NIST 8491	4.15	3.59	0.56		0.56		0.08	0.00	0.00	3.21	1.93	24.69	38.23	21.17	0.75	1.84	0.00	3.29	99.92
14		CO2_TM-5 STEM	15.38	3.69	11.70	8.19	3.20	4.99	1.64	1.28	2.12	8.24	2.44	8.47	30.29	11.52	1.05	2.48	0.00	1.51	94.63
15		CO2_TM-5 STEM	15.38	3.68	11.71	8.19	3.23	4.96	1.62	1.24	1.56	9.62	2.38	8.56	29.76	11.58	1.02	2.43	0.00	1.52	94.86
16		CO2_TM-5 STEM	15.38	3.71	11.67	8.10	3.44	4.66	1.65	1.25	1.99	7.78	2.44	8.63	30.48	11.92	1.04	2.50	0.00	1.57	94.73
17		CO2_TM-5 LEAF	13.90	6.00	7.90	12.83	8.90	3.94	2.60	1.04	1.12	8.42	3.87	11.44	21.86	10.83	1.07	2.81	0.00	1.11	92.92
18		CO2_TM-6 STEM	6.57	1.34	5.23	4.69	2.69	2.01	5.37	0.99	1.37	10.78	3.89	9.44	38.20	10.36	1.29	2.64	0.00	1.42	97.01
19		CO2_TM-6 LEAF	10.11	3.70	6.41	10.99	8.31	2.68	5.47	1.26	1.42	10.10	3.71	11.77	23.77	10.39	1.04	2.65	0.00	1.02	93.70
20		CO2_TM-7 STEM	13.88	2.27	11.62	7.87	3.37	4.50	3.25	0.96	1.32	11.80	2.28	8.74	29.42	11.01	1.04	2.36	0.00	1.49	95.43
21		CO2_TM-7 LEAF	11.59	3.48	8.11	13.62	10.59	3.03	3.14	1.21	1.45	8.60	4.28	11.32	23.00	10.80	0.89	2.44	0.00	1.07	93.41
22		NIST 8491	3.86	3.67	0.19		1.08		0.06	0.00	0.00	3.77	1.92	24.92	37.40	20.84	0.73	1.76	0.00	3.25	99.60
23		CO2_TM-8 STEM	7.90	1.22	6.67	6.76	2.75	4.02	4.08	2.14	3.06	12.95	2.18	7.38	33.66	10.28	1.10	2.41	0.00	1.37	95.27
24		CO2_TM-8 STEM	7.90	1.18	6.71	6.85	2.74	4.11	4.71	1.77	2.40	13.67	2.48	6.79	33.08	10.22	1.08	2.40	0.00	1.34	94.69
25		CO2_TM-8 STEM	7.90	1.25	6.65	6.95	2.79	4.15	4.67	1.89	2.48	12.68	2.18	8.30	33.80	10.32	1.11	2.40	0.00	1.33	96.09
26		CO2_TM-9 LEAF	10.29	3.03	7.27	13.85	10.13	3.71	0.26	1.51	2.31	10.88	4.36	10.42	22.75	11.02	0.93	2.64	0.00	1.04	92.25
27		CO2_TM-9 STEM	8.52	2.93	5.59	6.35	3.32	3.03	4.21	0.68	0.80	14.30	2.64	8.01	34.28	9.75	1.18	2.37	0.00	1.36	94.45
28		CO2_TM-9 LEAF	10.91	4.19	6.72	12.01	10.05	1.95	5.19	1.06	0.93	12.24	3.40	12.53	22.21	9.63	1.06	2.38	0.00	0.95	94.50
29		CO2_TM-10 STEM	11.16	2.63	8.53	7.13	3.49	3.64	2.47	0.66	0.74	11.47	2.16	8.76	32.86	11.46	1.20	2.76	0.00	1.46	94.28
30		CO2_TM-10 LEAF	12.75	4.02	8.73	13.71	9.93	3.78	3.03	0.51	0.63	9.09	2.31	12.92	22.86	10.80	0.99	2.47	0.00	1.09	93.15
31		CO2_TM-10 LEAF	12.75	3.72	9.03	13.52	9.79	3.74	3.04	0.51	0.61	9.61	2.55	12.65	22.77	10.78	1.00	2.51	0.00	1.02	93.31
32		CO2_TM-10 LEAF	12.75	3.86	8.89	13.48	9.69	3.78	2.86	0.58	0.67	9.79	2.96	12.55	22.76	10.82	0.96	2.48	0.00	1.11	93.76
33		CO2_TM-11 STEM	11.29	3.03	8.26	6.67	2.87	3.80	2.45	0.75	1.17	12.70	2.32	7.95	32.49	11.05	1.15	2.54	0.00	1.48	94.03
34		CO2_TM-11 LEAF	12.94	4.36	8.58	13.11	8.72	4.39	2.85	0.85	1.11	7.52	3.73	10.99	22.73	10.77	1.00	2.49	0.00	1.08	91.17
35		NIST 8491	4.49	4.14	0.34		0.56		0.11	0.01	0.04	3.40	1.79	26.85	33.32	16.37	0.67	1.78	0.00	3.14	92.55
BATCH 2																					
36		CO2_TM-12 STEM	7.77	1.60	6.17	5.66	2.82	2.83	3.21	0.74	1.16	13.36	3.54	8.01	38.25	10.45	1.20	2.39	0.00	1.43	97.17
37		CO2_TM-12 STEM	7.77	1.56	6.21	5.43	2.85	2.57	3.25	0.72	1.01	14.79	3.24	10.98	35.43	8.93	1.18	2.25	0.00	1.41	96.40
38		CO2_TM-12 STEM	7.77	1.65	6.12	5.37	2.78	2.58	3.48	0.70	1.00	14.70	3.32	7.80	37.87	10.35	1.22	2.40	0.00	1.43	97.32
39		CO2_TM-13 LEAF	11.47	3.20	8.27	12.93	9.75	3.18	3.47	1.99	2.52	11.31	4.06	12.22	22.85	9.99	0.95	2.32	0.00	1.00	96.86
40		CO2_TM-13 STEM	9.96	1.87	8.10	7.08	3.00	4.08	2.67	1.63	2.53	14.64	2.39	8.21	32.31	10.90	1.13	2.47	0.00	1.47	97.40
41		CO2_TM-13 LEAF	12.19	4.17	8.02	14.95	10.96	3.99	0.55	1.27	2.17	9.80	4.13	11.50	22.72	10.86	1.00	2.58	0.00	1.04	94.76
42		CO2_TM-14 STEM	12.23	2.41	9.82	7.54	3.45	4.09	2.69	0.40	0.55	8.81	2.23	9.76	34.29	12.86	1.08	2.47	0.00	1.71	96.61
43		CO2_TM-14 LEAF	12.93	4.09	8.84	12.05	8.91	3.14	1.50	0.31	0.43	8.09	3.45	12.38	26.12	12.31	1.09	2.67	0.00	1.21	94.57
44		CO2_TM-14 LEAF	12.93	4.37	8.56	13.62	9.46	4.16	1.47	0.31	0.41	7.70	3.44	12.87	26.38	12.41	1.15	2.75	0.00	1.25	96.88
45		CO2_TM-14 LEAF	12.93	4.39	8.54	13.20	8.47	4.73	1.43	0.34	0.45	6.52	3.43	12.11	25.09	11.79	1.09	2.59	0.00	1.22	92.21
46		CO2_TM-15 STEM	10.83	1.65	9.18	7.50	2.58	4.92	4.32	0.93	1.08	8.87	2.20	9.33	33.32	12.00	1.16	2.51	0.00	1.60	96.02
47		CO2_TM-15 LEAF	12.80	4.01	8.79	13.06	8.47	4.59	2.10	0.46	0.55	7.45	3.54	11.88	23.72	12.07	1.22	2.90	0.00	1.25	95.00
48		NIST 8491	3.45	4.64	0.81		0.56		0.07	0.00	0.00	3.42	1.92	25.07	37.27	21.28	0.81	2.01	0.00	3.35	101.21
49		CO2_TM-16 STEM	9.19	1.74	7.45	6.30	2.65	3.65	5.04	0.68	1.55	9.35	2.08	7.80	39.63	10.34	1.10	2.16	0.00	1.50	96.73
50		CO2_TM-16 STEM	9.19	1.71	7.48	6.35	2.99	3.36	4.95	0.68	1.39	9.43	2.03	7.54	39.35	10.50	1.14	2.22	0.00	1.47	96.45
51		CO2_TM-16 STEM	9.19	1.77	7.42	6.03	2.93	3.10	4.87	0.69	1.00	10.88	2.14	7.60	38.76	10.31	1.14	2.69	0.00	1.49	96.79
52		CO2_TM-17 LEAF	13.10	4.38	8.73	12.70	10.29	2.40	4.07	0.73	0.96	9.55	3.84	11.98	22.96	10.49	1.09	2.57	0.00	1.12	95.17
53		CO2_TM-17 STEM	7.21	1.48	5.73	5.34	2.98	2.36	2.65	0.65	1.01	9.16	2.28	8.38	43.77	10.91	1.16	2.60	0.00	1.69	96.82
54		CO2_TM-17 LEAF	13.89	4.84	9.05	12.42	8.63	3.79	2.64	0.49	1.56	8.31	4.40	11.02	24.05	10.56	1.07	2.82	0.00	1.12	94.33
55		CO2_TM-18 STEM	14.62	2.00	12.62	7.36	2.74	4.62	6.69	1.02	1.33	7.42	2.29	8.75	29.22	11.90	1.30	2.66	0.00	1.72	96.29
56		CO2_TM-18 LEAF	15.01	5.65	9.37	14.03	9.47	4.56	1.53	0.69	1.02	6.92	4.90	10.89	23.26	11.03	1.09	2.88	0.00	1.16	94.40
57		CO2_TM-19 STEM	9.98	1.36	8.63	6.07	2.79	3.28	4.40	0.71	1.01	8.41	2.14	9.69	37.59	11.23	1.18	2.41	0.00	1.55	96.39
58		CO2_TM-19 LEAF	10.77	3.61	7.16	14.12	10.29	3.83	1.03	0.53	0.83	7.33	4.73	12.77	25.16	11.92	1.07	2.73	0.00	1.21	94.20
59		CO2_TM-20 STEM	13.17	2.54	10.62	7.96	3.02	4.94	2.85	0.75	0.88	9.13	2.45	8.76	32.32	12.00	1.05	2.57	0.00	1.59	95.50
60		CO2_TM-20 LEAF	13.53	5.86	7.67	12.51	8.90	3.61	1.97	0.48	0.62	8.34	3.75	12.16	24.24	11.52	1.02	2.86	0.00	1.18	94.18
61		NIST 8491	3.91	3.71	0.20		0.52		0.07	0.00	0.00	3.64	1.81	24.56	38.39	21.60	0.79	1.93	0.00	3.37	100.61
62		CO2_TM-21 STEM	12.46	2.50	9.95	7.59	2.99	4.60	2.75	0.61	1.38	10.49	2.42	9.05	31.76	11.69	1.20	2.77	0.00	1.62	95.78
63		CO2_TM-21 STEM	12.46	2.27	10.18	7.64	3.21	4.43	2.78	0.59	0.76	11.06	2.50	8.95	32.17	11.47	1.10	2.66	0.00	1.56	95.71
64		CO2_TM-21 STEM	12.46	2.45	10.01	7.41	3.24	4.16	2.70	0.59	0.75	11.46	2.53	10.78	31.49	11.45	1.11	2.74	0.00	1.61	97.07
65		CO2_TM-21 LEAF	11.97	4.53	7.44	13.11	9.38	3.73	2.15	0.86	0.98	8.53	3.75	12.21	24.46	11.31	1.00	2.90	0.00	1.18	94.40
66		NIST 8491	3.86	3.38	0.48		0.52		0.11	0.04	0.06	3.31	1.93	28.07	37.71	21.32	0.80	2.16	0.00	3.42	103.29