EXPLORING THE GENETIC AND PHYSIOLOGICAL MECHANISMS OF Al TOLERANCE IN RICE (Oryza sativa L.)

by Adam Nicholas Famoso

EXPLORING THE GENETIC AND PHYSIOLOGICAL MECHANISMS OF ALUMINUM TOLERANCE IN RICE (*Oryza sativa* L.)

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

Adam Nicholas Famoso

August 2010

EXPLORING THE GENETIC AND PHYSIOLOGICAL MECHANIMS OF ALUMINUM TOLERANCE IN RICE (*Oryza sativa* L.)

Adam Nicholas Famoso, Ph.D.

Cornell University 2010

Aluminum (Al) is the most abundant metal in the earth's crust and is predominantly found as a key component of soil clays. Under highly acidic soil conditions (pH<5.0) it is solubilized to Al^{3+}, which is highly phytotoxic. Al^{3+} causes a rapid inhibition of root growth that leads to a reduced and stunted root system, thus having a direct effect on the ability of a plant to acquire both water and nutrients. Over 50% of the world's potentially arable lands are acidic and therefore Al toxicity is one of the major limitations to global agriculture. This dissertation examines the genetic and physiological mechanisms of Al tolerance in rice and demonstrates that rice is significantly more tolerant than other cereals. Physiological experiments demonstrate that rice utilizes a novel physiological mechanism that does not involve the well described root tip exclusion of rhizosphere Al mediated by organic acid exudation that has been identified in a number of other crop species. Genetic analysis was conducted using both QTL and whole genome association analysis on a panel of 385 *Oryza sativa* accessions. This analysis identified that the *Japonica* varietal group is significantly more Al tolerant than the *Indica* varietal group and identified loci that confer tolerance both within and across rice subpopulations. Furthermore, admixture analysis identified distinct *Japonica* loci that increase *Indica* Al tolerance to levels at or above that of *Japonica*. The results from this suggest that rice is a unique model to study Al tolerance in cereals and will lead to experiments to further understand Al tolerance and toxicity and develop crop varieties with enhanced Al tolerance.

BIOGRAPHICAL SKETCH

Adam was born July 23, 1980 in Princeton, New Jersey and grew up in Hamilton, NJ where he graduated from Steinert High School in 1998. He then attended Paul Smith's College where he obtained an associate's degree in Urban Forestry in 2000. In 2001 he worked as an arborist in Jacksonville, Florida and then traveled across the country for three months. In 2002 he enrolled in the Horticulture program at the Pennsylvania State University at University Park. While at Penn State Adam became interested in genetics and plant breeding and focused his studies on plant physiology and genetics. In 2004 he graduated with "High Distinction" with a Bachelor's degree in Horticulture and a minor in Biology and enrolled in the PhD program in Plant Breeding and Genetics at Cornell University. While at Cornell, Adam conducted international research at CIAT in Cali, Colombia and at IRRI in Los Banos, Philippines and attended courses and/or conferences in Mozambique, South Africa, and The Philippines. He also proposed, designed and co-taught three genetics courses through the Cornell Prison Education Program at the Auburn and Cayuga Correctional Facilities. Adam will begin his post-graduate career as a Research Scientist for Pioneer-Hi Bred, working on molecular breeding of rice for Pioneer's Asian Rice Research Program and will be based between Des Moines, Iowa and Los Banos, Philippines.

ACKNOWLEDGMENTS

I have deep gratitude to my parents for their endless support and giving me the freedom to design and pursue my own directions in academics, travel, and life. My academic path to a PhD was not a traditional one and would not have been possible without their encouragement and patience.

For Susan and Leon, who allowed me to work between their labs and gain a wide range of knowledge in both genetics and plant physiology. The encouragement, mentorship, enthusiasm, and support I received during my PhD has been invaluable to my development as a scientist and I am truly grateful for this support. To my committee, Tim and Ronnie, I want to thank them for their valuable feedback and contributions on my research and development as a student. To Margaret, I appreciate her willingness to serve as a proxy for my defense and for the teaching skills I have learned from her through various courses in which I was a student, TA, and instructor.

I am thankful to the members of the McCouch and Kochian labs, The Plant Breeding and Genetics Department, and the Cornell community for providing an exciting and positive academic and research environment and for their valuable critical analysis and useful suggestions and comments on my work.

I am grateful to Pioneer Hi-Bred, The Departments of Plant Breeding and Genetics and Plant Biology, and the USDA for funding of my degree. I also acknowledge a research grant funded by the USDA AFRI program (Award Number: 2009-02273) and the Generation Challenge Programme (GCP) for research funding support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

PREFACE

There are two widely recognized varietal groups of *Oryza sativa*, *Indica* and *Japonica*. These varietal groups have been further classified into genetic subpopulations, with *Indica* varietal group containing the *indica* and *aus* subpopulations and the *Japonica* varietal group containing the *tropical japonica, temperate japonica,* and *aromatic* subpopulations. To distinguish between the varietal groups and subpopulations, the first letter will be capitalized when referring to varietal groups and lowercase letters will be used when referring to subpopulations.

Chapter 2 was published as an original research article in *Plant Physiology* (Famoso et al., DOI:10.1104/pp.110.156794). Chapter 3 is being submitted as an original research article to *Plant Cell*. Chapter 4 is an outline of the future experiments that will be conducted as part of a funded USDA AFRI grant.

CHAPTER 1:

INTRODUCTION

The advent of agriculture has allowed for vast human population growth. The hunter-gather civilizations sustained approximately 4 million people, whereas modern agriculture currently sustains over 6.5 billion people globally (Tilman et al., 2002). The global population is increasing at a rate of approximately 160 people each minute and it is expected that the global population will reach 8 billion by 2025 (Hoisington et al., 1999; Dyson, 1999). It is estimated that agriculture production will have to increase 100% over the next 50 years to feed the projected population (Tilman et al., 2002). To meet this increased demand, production between 2000-2050 will have to equal the amount of food produced from the dawn of agriculture through the twentieth century (~10,000 years) (Hoisington et al., 1999). Increases in population will continue to reduce the availability of arable land, thus, it is not possible to meet future food demands by simply expanding the land under cultivation. As arable land diminishes, it will be necessary to increase productivity on less than desirable land. Increased production of cereal species (Poaceae) will be one of the primary means to meet the increased food demands, as cereals account for one-half of the world's cropland and account for ~2/3 of direct and indirect human calories (Dyson, 1999; http://faostat.fao.org/site/339/default.aspx).

Importance and production of rice

Rice (*Oryza sativa L.*) is the world's most important calorie source, serving as the main staple for about half of the world's population and providing 20% of the calories directly consumed by humans (Khush, 1997; Zeigler and Barclay, 2008). Rice is grown in more than 100 countries on all inhabited continents, under diverse

conditions, such as irrigated paddies, rainfed lowland and upland environments (Zeigler and Barclay, 2008). The different environments in which rice is produced are associated with specific biotic and abiotic constraints, including drought, flooding, nutrient deficiencies and toxicities, and disease and insect pressure. Around half of the land area planted to rice is irrigated, yet irrigated rice accounts for 75% of total production (Zeigler and Barclay, 2008). Rainfed rice accounts for the remaining ~50% of land area under rice production and is widespread throughout South and Southeast Asia, Africa and Latin America, often on farms of resource-poor farmers (Zeigler and Barclay, 2008). As freshwater supplies and arable land become increasingly limited, rainfed rice production systems will likely surpass irrigated systems in terms of area under cultivation. Therefore, increasing the productivity of rainfed systems will be critical to meeting the increasing demand for rice. This will ultimately require development of varieties that are resistant to the constraints of rainfed systems.

Aluminum toxicity

One of the major limitations to rainfed rice production is aluminum (Al) toxicity. Aluminum is the most abundant metal in the earth's crust. Comprising approximately 7% of the soil (Wolt, 1994), Al is predominantly found as a key component of soil clays. Under highly acidic soil conditions (pH<5.0) it is solubilized to Al^{3+}, which is highly phytotoxic. Al^{3+} causes a rapid inhibition of root growth that leads to a reduced and stunted root system. Thus, it has a direct effect on the ability of a plant to acquire both water and nutrients. Approximately 30% of the world's total land area and over 50% of the world's potentially arable land are acidic. The majority of acid soils are found in the tropics and subtropics, which are estimated to account for 60% of the world's acid soils (von Uexkull et al, 1995). Al toxicity is the foremost limitation to

crop production for 38% of farmland in Southeast Asia, 31% in Latin America and approximately 20% in East Asia, Sub-Saharan Africa and North America (Wood et al., 2000). Acid soils are also a significant limitation to U.S. agriculture, with approximately 135 million hectares of land in the U.S. being classified as highly acidic. Acid soils will become more of an issue with the changing environment, as weathering of soils, intensive agricultural practices, and acid rain are all contributors to soil acidification (Sumner and Noble, 2003). Future agricultural productivity in the U.S. and the world will be impacted by the limitations of acid soils, which are exceeded only by drought stress with regards to abiotic stress (von Uexkull et al, 1995).

Cultural approaches, such as applying lime ($CaCO_3$), can ameliorate some of the limitations of acidic soils, leading to yield increases of as much as 100% (Sumner and Noble, 2003). However, liming is only effective at increasing the pH in the upper soil profile and is generally ineffective when the subsoil is acidic (Marschner, 1995). It has been estimated that 75% of the world's acidic soils are affected by subsoil acidity, making it infeasible to completely solve the limitations of acid soils through agronomic approaches. Liming is also problematic in many regions of the world, due to financial and/or infrastructural limitations. Thus, developing Al tolerant crops capable of tolerating acid soils is a priority for breeding programs around the world. To achieve the full genetic potential in breeding Al tolerant crops, it is critical that we understand the genetic, molecular, and physiological mechanisms underlying this trait.

Physiology of Al Tolerance:

There are two well-recognized classes of physiological mechanisms that confer the ability of plants to survive toxic levels of Al. Exclusion mechanisms serve to exclude

Al from entering the root, while internal sequestration allows the plant to accumulate Al in the symplasm (Miyasaka et al, 1991; Delhaize and Ryan, 1995; Kochian, 1995; Kochian et al, 2004). Experimental evidence from a number of crop species supports the exclusion mechanism (Ryan et al, 1993; Ma et al, 2001; Ryan et al., 2001) although this tolerance mechanism has not been demonstrated in rice. Two independent studies investigating Al tolerance mechanisms in rice have demonstrated that susceptible varieties have significantly increased levels of Al accumulation in the root apex, which has been clearly identified as the site of Al toxicity. However, there were no differences in organic acid exudation or rhizosphere pH when Al tolerant and Al sensitive varieties were compared (Ma et al., 2002; Yang et al., 2008). These studies have only focused on one tolerant and one susceptible genotype and therefore a more thorough analysis needs to be conducted to confirm that the exclusion mechanism is functioning in rice Al tolerance.

The exclusion mechanism has been demonstrated to be mediated by the Al-activated exudation of organic acids such as citrate, oxalate, or malate from the root apex. These organic acids chelate Al in the rhizosphere, reducing the concentration and toxicity of Al (Ma et al., 2001). Physiological studies have identified malate and citrate anion efflux transporters in wheat (*Triticum aestivum* L.*)* and maize (*Zea mays* L.), respectively, that are involved in Al tolerance in these species and are activated by Al (Zhang et al., 2001; Piñeros et al., 2002). Thus, some of the physiological variation observed within and between species may be attributed to the density of the organic acid transporter proteins in the root plasma membrane, their permeability to specific organic acids, or their activation by Al. The strongest evidence demonstrating the importance of Al-activated transporter proteins comes from the cloning of Al tolerance genes encoding organic acid anion membrane transporters in wheat (*ALMT1*) (Sasaki

et al., 2004) and sorghum (*Sorghum bicolor* L. Moench.) (*SbMATE*) (Magalhaes et al, 2007). Organic acids have also been shown to be involved in an internal Al tolerance mechanism in Buckwheat (*Fagopyrum esculentum* Moench.) (Ma et al., 1998). In rice, a significant increase in root citrate release was observed upon Al treatment in both a tolerant and susceptible variety; however, there was no a significant difference in exudation between the two varieties, thus citrate exudation could not explain the differences in tolerance (Ma et al., 2002).

Although organic acids have been shown to have a major role in Al tolerance, other exclusion mechanisms have been identified in *Arabidopsis thaliana* and there is clear evidence that tolerance in maize cannot be fully explained through root organic acid release (Degenhardt et al., 1998; Ezaki et al, 2001; Piñeros et al., 2005). In an *Arabidopsis* mutant, an Al-induced increase in rhizosphere pH, due to an influx of H^+ ions, contributes to Al tolerance (Degenhardt et al., 1998). In maize, organic acid (citrate) release accounts for some, but not all of its ability to exclude Al and confer tolerance (Piñeros et al., 2005). This strongly suggests that multiple tolerance and exclusion mechanisms exist in maize.

Cell wall composition and structure has been implicated in plant Al tolerance (Yang, 2008; Eticha, 2005; Mimmo, 2009), but has yet to be carefully investigated, especially with regards to studies in multiple genotypes or across species. Cell wall expansion has also been demonstrated to be greatly inhibited by Al, but it is still not known whether this is a component of Al toxicity, or if there is variation between Al tolerant and Al susceptible varieties for cell expansion in response to Al. Despite rice being one of the most Al tolerant cereals, very little is known about the physiological mechanisms of Al tolerance in rice.

Genetics of Al Tolerance:

High levels of genetic variation for Al tolerance have been observed both within and between species (Wu et al., 1997; Piñeros et al., 2005; Garvin and Carver, 2003; Famoso et al., 2010). In wheat, sorghum, and barley (*Hordeum vulgare* L.), Al tolerance is inherited as a simple trait, controlled by one or a few genes (Sasaki et al, 2004; Magalhaes et al, 2004; Minella and Sorrells, 1992). However, in maize, rice, and *Arabidopsis,* tolerance is quantitatively inherited (Ninamango-Cardenas et al., 2003; Hoekenga et al., 2003; Nguyen et al., 2001 and 2002).

As mentioned above, two Al tolerance genes have been cloned to date, in wheat and sorghum. The wheat resistance gene, *ALMT1*, encodes an Al-activated, malate transporter (Sasaki et al., 2004). The sorghum resistance gene, Alt_{SB}, encodes a member of the multidrug and toxic compound extrusion (MATE) family and is an Al-activated, root citrate efflux transporter (Magalhaes et al, 2007). The coding sequences (open reading frames) from susceptible and tolerant parents are identical in both cases, but in each case, tolerance is highly correlated to high levels of gene expression.

Eight QTL studies on Al tolerance have been reported in rice using 7 different inter- and intra-specific mapping populations (Wu et al., 2000; Nguyen et al., 2001, 2002, 2003; Ma et al., 2002; Xue et al., 2006, 2007). These studies identified a total of 33 QTLs on 11 different chromosomes, with three regions, on chromosomes 1, 3, and 9, being detected in multiple studies. An Al tolerance QTL on rice chromosome 1, which was identified in all published studies, shares extensive conservation with sorghum chromosome 3 (Klein et al., 2003), the location of the cloned sorghum Al tolerance gene, Alt_{SB}, (Magalhaes et al, 2007).

When the sorghum *SbMATE* sequence was used as a query against the rice genome (using the BLAST module in the Gramene database, www.gramene.org), a rice ortholog to *SbMATE, OsMATE,* was identified on chromosome 1, in roughly the same region as four previously published (but not all overlapping) QTLs (Nguyen, 2001, 2002; Ma, 2002; Wu, 2000). Several other members of the *SbMATE-like* (MATE) gene family also co-localize with previously published Al tolerance QTL. Yet, despite the coincidence between the positions of these MATE genes and previously identified QTLs, physiological analysis of organic acid exudation and root-tip Al exclusion provide no support for the hypothesis that the MATE gene family is directly responsible for Al tolerance in rice. This begs the question of what genes underlie the QTLs and what mechanism(s) are functionally responsible for the high levels of Al tolerance that distinguish rice from other cereals.

A second major Al tolerance QTL was identified on rice chromosome 3, in a homoeologous region corresponding to wheat chromosome 4DL (Gale and Devos, 1998), which is the location of *ALMT1*, the cloned wheat Al tolerance gene (Sasaki et al, 2004). However, when the wheat *ALMT1* gene sequence is used as a query against the rice genome, it identifies a homologue on rice chromosome 4 that does not overlap with any previously published QTLs. Furthermore, when the wheat *ALMT1* gene was transformed into rice, transgene expression was associated with an Al-activated release of malate from rice roots, but this was not correlated with an increase in tolerance to Al in the transgenic lines (Sasaki et al, 2004). The authors suggested that the innately high levels of Al tolerance in rice may mask the detoxifying effect of the malate released from the roots. When all the evidence is taken together, including physiological analysis of organic acid exudation and root-tip Al exclusion, it suggests that the *ALMT1* gene family is not involved in Al tolerance in rice.

Three genes underlying three different Al-sensitive rice mutants, *STAR1* and *STAR2* (Sensitive to Al rhizotoxicity) and *ART1* (Aluminum rhizotoxicity 1), that lead to Al sensitivity have recently been cloned and demonstrated to be involved in rice Al tolerance (Huang et al, 2009, Yamaji et al, 2009). The gene products of *STAR1* and *STAR2* are components of a bacterial-type ATP binding cassette (ABC) transporter, are expressed mainly in the roots, and are transcriptionally activated by exposure to Al. Loss of function of either gene results in hypersensitivity to Al. Huang et al. (2009) demonstrated that the STAR1/STAR2 transporter complex transports UDP-glucose out of the root cytoplasm into the cell wall, and the addition of exogenous UDP-glucose partially rescued the mutant Al sensitivity phenotype. *ART1* is a novel C2H2-type zinc finger-type transcription factor that specifically regulates expression of Al-inducible rice genes, some of which appear to be related to Al tolerance. *ART1* is localized to the nucleus of all root cells, is constitutively expressed in roots, and is not affected by Al treatment. *ART1* interacts with the promoter region of *STAR1* and microarray analysis identified 30 down-stream transcripts regulated by *ART1*, some of which are involved in internal and external Al detoxification (Yamaji et al., 2009). Although these three genes are required for root growth in Al, they do not map to any of the Al tolerance QTL previously identified in rice. This suggests that these genes are involved in basal Al tolerance and possibly suggests that UDP-glucose and cell wall modification may be important for general rice root growth in the presence of Al (Huang et al, 2009, Yamaji et al, 2009).

Two other Al sensitive, ABC transporter mutants (*als1* and *als3*) have been positionally cloned and characterized in *Arabidopsis* (Larsen et al. 2005, 2007). *ALS3* encodes a phloem-localized ABC transporter-like protein that is localized to the plasma membrane (Larsen et al., 2005). *ALS1* encodes a half type ABC transporter

that is localized to the vacuole membrane and is expressed throughout the plant (Larsen et al., 2007). When knocked out, both genes reduce Al tolerance in *Arabidopsis* and it appears that they each may be part of a multi-component mechanism, as overexpression of each gene individually does not increase Al tolerance. Although the physiological mechanism(s) of Al tolerance associated with *ALS1* and *ALS3* are not clear, the expression of these genes in both the root tip and up through the vasculature system, suggests that they may function by translocating Al away from sensitive sites to the vacuole and/or less sensitive tissues. Although a sequestration mechanism may also be involved in Al tolerance in rice, BLAST sequence analysis of *ALS1* and *ALS3* indicates that the rice homologs for these genes do not overlap with rice Al tolerance QTL.

Rice domestication and genetic structure

There are two cultivated species of rice; *Oryza sativa*, which originated in Asia but is grown throughout the world, and *Oryza glaberrima*, which is found predominantly in Africa. Within *O. glaberrima*, three genetic subpopulations have been identified, as well as admixtures between *O. glaberrima* and *O. sativa*. The genetic subpopulations in *O. glaberrima* are associated with morphological and physiological variation in traits related to ecological adaptations corresponding to the upland, floating, and non-floating types (Semon et al., 2005). Modern *O. sativa* rice has undergone extensive selective pressure since its domestication ~10,000 years ago (Kovach et al., 2007). It has a complex domestication history, with multiple domestications from a pre-differentiated ancestral gene pool (*O. rufipogon*) followed by significant gene flow among and between subpopulations (Zhou et al., 2003; Barbier, 1989; Zhu and Ge, 2005; Ma and Bennetzen, 2004; Vitte et al., 2004, Londo et al., 2006; Kovach and McCouch, 2008). This evolutionary history led to the formation of two major varietal

groups, *Indica* and *Japonica* (Oka, 1988). These two groups have been further divided into five major genetic subpopulation groups, based on isozymes and DNA markers (SSR, SNPs, Indels, etc.). The *indica* and *aus* subpopulations trace their ancestry from the *Indica* varietal group and the *tropical japonica, temperate japonica,* and *aromatic [Group V]*) subpopulations trace their ancestry through the *Japonica* varietal group (Glaszmann, 1987; Garris et al., 2005; Zhao et al., 2010). Subpopulation differences are evident at the gene and sequence level throughout the genome (Liu et al., 2007) and also in terms of morphological, developmental and physiological trait performance (Champoux et al., 1995; Lilley et al, 1996; Garris et al. 2003; Xu et al., 2009; Oka, 1988).

Genetic resources of rice

Due to its abundant genetic and genomic resources, its economic and social importance, and its relatively small genome (390 Mb), rice was the first plant crop genome to be fully sequenced and remains a model monocot species. The rice genome is six times smaller than maize and 40 times smaller than wheat (Jung et al., 2008). Both *Japonica* and *Indica* genome sequences are available in rice (Goff et al., 2002; Yu et al., 2002). The *Japonica* sequence (cv. Nipponbare) was the result of an international effort based on a BAC-by-BAC sequencing approach (IRGSP, 2005), while the *Indica* sequence (cv. 93-11) was a shot-gun re-sequencing effort with reads assembled using the cv. Nipponbare genome as a template (Goff et al., 2002; Yu et al., 2002). Hundreds of diverse rice genotypes are currently being resequenced and aligned to the existing genome sequences. Annotation of the Nipponbare (*Japonica*) genome (MSU6) predicts that rice contains 56,797 genes, 40,577 of which are Non-TE-related genes (http://rice.plantbiology.msu.edu/riceInfo/info.shtml**).** In addition to a fully sequenced and annotated genome, there are also many complementary, publicly

available platforms for identifying gene functions in rice, including numerous mapping populations, microarrays, tilling arrays, gene-indexed mutants, gene silencing constructs, and transient assays (reviewed by Jung et al., 2008). Although "functional genomics" approaches are very effective in determining gene function, they are less effective at identifying genes and alleles that have agricultural application, as they generally rely on loss-of-function knockout alleles. However, the identification of gene function through loss-of-function mutants does provide a reservoir of candidate genes to investigate for natural variation underlying a phenotype of interest.

QTL mapping

To date, the most widely used approach to identify genes underlying complex traits in rice has relied on bi-parental quantitative trait loci (QTL) mapping. The concept of QTL mapping was first demonstrated by Sax et al. (1923) based on the observation that bean color, a quantitative trait, was significantly associated with the quantitative trait of bean size. However, until the development of molecular markers, there was not sufficient genetic map coverage to dissect quantitative traits. With the development of a saturated molecular marker map in tomato, it became possible to associate quantitative phenotypes with molecular markers that segregated as qualitative traits throughout the genome (Patterson et al., 1988). Since the first rice QTL study in 1994 (Wang et al.) through 2009, >8600 rice QTLs have been identified in 617 published reports (Ni et al., 2009; Yamamoto et al., 2009; www.gramene.org). Through the development of near isogenic lines (NILs) it is possible to convert a QTL into a Mendelian factor, which can facilitate the fine-mapping and identification of the underlying gene(s) (Martin, 1993). This approach has been successful in identifying genes underlying many QTLs of agronomic traits in rice, including heading date

(Yano et al., 2000; Kojima et al., 2002; Takahashi et al., 2001; Doi et al., 2004; Xue et al., 2008), submergence tolerance (Xu et al., 2006), seed number (Ashikari et al., 2005), UVB resistance (Ueda et al., 2005), salt tolerance (Ren et al., 2005), seed shattering (Konishi et al., 2006; Li et al., 2006), growth habit (Tan et al., 2008; Jin et al., 2008), seed length (Fan et al., 2006), seed width (Song et al., 2007; Shomura et al., 2008), low temperature germinability (Fujino et al., 2008), and regeneration ability (Nishimura et al., 2005).

Although bi-parental QTL mapping has been successful in identifying genes underlying quantitative trait variation, it is a labor intensive process that often requires development of experimental populations specific for the trait of interest. The basis of QTL mapping is to identify a statistical association between a specific genetic marker(s) and a phenotype. Bi-parental QTL mapping is effective in delimiting a gene(s) contributing to a quantitative trait to a genetic interval of 10-30cM, depending on population size and marker density. Mapping populations are also useful for NIL development, fine-mapping, and gene cloning, as backcross populations can be initiated prior to any QTL knowledge. The major limitation of QTL mapping is that it only examines two alleles of a gene and will not detect genes that have the same allele in the parents. Furthermore, it can only investigate the effect of the alleles in the genetic backgrounds of the mapping parents.

Whole genome association mapping

Association analysis (or association mapping) is a powerful method for dissecting quantitative traits. Although this approach was initially used to map genes in humans (Hirschhorn and Daly, 2005), it has been successfully utilized to associate polymorphisms to both quantitative and qualitative traits in maize and *Arabidopsis*

(Thornsberry, 2001; Wilson, 2004; Harjes et al 2008; Whitt, 2002; Szalma, 2005; Palaisa, 2003; Zhao et al., 2007). Similar to QTL mapping, association mapping uses statistical models to identify markers that are linked to a trait of interest. The major differences between QTL mapping and association mapping are the choice of germplasm and the number of alleles being examined. Whereas QTL mapping is based on experimental populations derived from two parents, such that only the parental alleles are segregating in the population, association mapping utilizes natural populations or collections of germplasm whose relationships to each other are not clear and theoretically examines all the alleles present in that population. Both QTL and association mapping rely upon linkage of a marker to a gene controlling a trait of interest, and the resolution in both cases is limited by the number of recombination events that can be monitored. In the best-case scenario of QTL mapping with intermated recombinant inbred lines (RILs), there are ~6-10 generations in which recombination can occur. Because association mapping uses natural populations, it takes advantage of all the historical recombination events that have occurred over time in the population, since divergence from a common ancestor. Depending on the germplasm used for association mapping, this approach may provide higher resolution than QTL mapping for the same number of individuals (~200). This prediction is contingent on understanding and controlling for population substructure that may be present. In addition, association mapping can identify regions associated with a trait of interest that may not have been polymorphic in the bi-parental QTL mapping population and provides an understanding of whether different alleles and/or genes are segregating in different sub-populations, offering some information on the evolutionary history of the trait. Thus, association mapping is a complementary approach to QTL mapping and fine-mapping efforts in plants.

The degree of linkage disequilibrium (LD) and population structure can vary considerably between species and these differences have direct effects on association mapping. Population structure can also result in spurious associations between markers and phenotypes, thus increasing the risk of type 1 error (Yu et al, 2006). Inflated LD values are observed in most rice subpopulations, compared to maize, human, and *Arabidopsis*, largely attributed to the high levels of self-pollination coupled with the domestication bottleneck in rice. The slow rate of LD decay in rice makes it possible to perform whole genome association studies, but limits the resolution of association mapping. For example, the resolution of association mapping in maize is at the gene level (1-3 genes), whereas in the *indica* and *aus* subpopulations of rice, the resolution is typically ~20-50 genes (Garris et al., 2003; Mather et al., 2007; Rakshit et al, 2007; Zhao et al., 2010), and for the *japonica* subpopulation, resolution is ten times lower (> 1,000 genes). When compared to the typical resolution for a QTL study of 250 rice lines, association mapping tends to provide higher resolution for the same number of individuals, though this is not always the case. QTL analysis typically provides between 10-20 cM resolution (1cM = ~250kb), and in the *indica* and *aus* subpopulations, association mapping will provide between 10-20 times higher resolution for a population of similar size. Although association mapping generally increases mapping resolution compared to QTL mapping, the resolution is limited by extent of LD in the target region. Therefore, to identify the gene(s) underlying a QTL, fine-mapping and experimental populations are still required. However, if association mapping is used for QTL discovery, fine-mapping can then be focused on a well resolved region.

To leverage the genetic diversity of rice, a diversity of panel of 400 *O. sativa* accessions, representing the genetic diversity of domesticated rice, was recently

compiled and genotyped with 44,000 SNPs (~1 SNP/10 kb) (Tung et al., submitted). Statistical mixed-models have been developed to account for deep subpopulation structure and inbreeding structure in rice and facilitate whole-genome association (GWA) analysis (Yu et al., 2006). This panel has been phenotyped for >25 agronomic, physiological and domestication related traits and significant loci have been associated with all traits, many of which co-localize with previously published QTLs and/or identified genes. Furthermore, admixed lines have proven to be powerful in identifying and confirming loci associated with subpopulation specific alleles (Zhao et al., 2010).

Phenotypes for dissecting quantitative traits

Most traits of agricultural significance are quantitative in nature, controlled by multiple genes/loci. Quantitative traits are difficult to dissect genetically for two main reasons; insufficient genotypic data and inadequate precision of phenotyping (Benfey and Mitchell-Olds, 2008). With the technological advances in genome sequencing over the last 10 years, the limitation of dissecting quantitative traits has shifted from incomplete genetic data to insufficient phenotyping platforms. While the cost and efficiency of genotyping has become more automated and considerably cheaper over the years, the cost of phenotyping in field trials is still time-consuming, labor intensive and expensive (Montes et al., 2007). To genetically dissect quantitative traits, it is necessary to have reliable and reproducible quantitatively measured phenotypes. It is also critical that the phenotyping platform is high-throughput, due to the large number of genotypes and replications necessary to obtain an accurate phenotype for each genotype. While breeders have selected for quantitative trait improvement in breeding populations using visual evaluation and trait indexing (often based on categorical assessments) that allow them to discard the bad and keep the best, this is not sufficient

to dissect the underlying genetic basis of a trait. To dissect the biochemical, developmental or physiological mechanism(s) underlying quantitative traits, biologists conduct detailed experiments to evaluate fine-scaled phenotypic variation. However, it is often only possible to characterize a few genotypes at this level of resolution. Thus, more precise approaches to phenotyping are often not applicable for screening thousands of plants. To efficiently utilize the extensive genomic data and mapping platforms available in rice, it is necessary to develop high-throughput, low cost quantitative phenotyping platforms as the basis for dissecting the genetic architecture underlying traits of interest.

REFERENCES

Ashikari M, Sakakibara H, Lin S, Yamamoto T, Takashi T, Nishimura A, Angeles ER, Qian Q, Kitano H, Matsuoka M (2005) Cytokinin oxidase regulates rice grain production. Science 309: 741-745

Barbier P (1989) Genetic variation and ecotypic differentiation in the wild rice species O. rufipogon. II. Influence of the mating system and life history traits on the genetic structure of populations. Japanese Journal of Genetics 64: 273-285

Benfey PN, Mitchell-Olds T (2008) From genotype to phenotype: systems biology meets natural variation. Science 320: 495-497

Brar DS, Khush GS (1997) Alien introgression in rice. Plant Mol Biol 35: 35-47

Champoux MC, Wang G, Sarkarung S, Mackill DJ, O'Toole JC, Huang N, McCouch S (1995) Locating genes associated with root morphology and drought avoidance in rice via linkage to molecular markers. Theoretical and Applied Genetics 90: 969-981

Degenhardt J, Larsen PB, Howell SH, Kochian LV (1998) Aluminum resistance in the *Arabidopsis* mutant alr-104 is caused by an aluminum-induced increase in rhizosphere pH. Plant Physiol 117: 19-27

Delhaize E, Ryan PR (1995) Aluminum Toxicity and Tolerance in Plants. Plant Physiol 107: 315-321

Doi K, Izawa T, Fuse T, Yamanouchi U, Kubo T, Shimatani Z, Yano M, Yoshimura A (2004) Ehd1, a B-type response regulator in rice, confers short-day promotion of flowering and controls FT-like gene expression independently of Hd1. Genes Dev 18: 926-936

Dyson T (1999) World food trends and prospects to 2025. Proc Natl Acad Sci U S A 96: 5929-5936

Ezaki B, Katsuhara M, Kawamura M, Matsumoto H (2001) Different mechanisms of four aluminum (Al)-resistant transgenes for Al toxicity in *Arabidopsis*. Plant Physiol 127: 918-927

Famoso A, Clark R, Shaff J, Craft E, McCouch S, Kochian L (2010) Development of a novel aluminum tolerance phenotyping platform used for comparisons of cereal Al tolerance and investigations into rice Al tolerance mechanisms. Plant Physiology. Published on June 10, 2010; 10.1104/pp.110.156794

Fan C, Xing Y, Mao H, Lu T, Han B, Xu C, Li X, Zhang Q (2006) GS3, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein. Theor Appl Genet 112: 1164-1171

Fujino K, Sekiguchi H, Matsuda Y, Sugimoto K, Ono K, Yano M (2008) Molecular identification of a major quantitative trait locus, qLTG3-1, controlling low-temperature germinability in rice. Proc Natl Acad Sci U S A 105: 12623-12628

Gale MD, Devos KM (1998) Comparative genetics in the grasses. Proc Natl Acad Sci U S A 95: 1971-1974

Garris AJ, McCouch SR, Kresovich S (2003) Population structure and its effect on haplotype diversity and linkage disequilibrium surrounding the xa5 locus of rice (Oryza sativa L.). Genetics 165: 759-769

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

Garvin DF, Carver BF (2003) Role of the genotype in tolerance to acidity and aluminum toxicity. In Z Rengel, ed, Handbook of soil acidity. Marcel Dekker, New York

Glaszmann JC (1987) Isozymes and classification of Asian rice varieties. TAG 74: 21-30

Goff SA, Ricke D, Lan TH, Presting G, Wang R, Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H, Hadley D, Hutchison D, Martin C, Katagiri F, Lange BM, Moughamer T, Xia Y, Budworth P, Zhong J, Miguel T, Paszkowski U, Zhang S, Colbert M, Sun WL, Chen L, Cooper B, Park S, Wood TC, Mao L, Quail P, Wing R, Dean R, Yu Y, Zharkikh A, Shen R, Sahasrabudhe S, Thomas A, Cannings R, Gutin A, Pruss D, Reid J, Tavtigian S, Mitchell J, Eldredge G, Scholl T, Miller RM, Bhatnagar S, Adey N, Rubano T, Tusneem N, Robinson R, Feldhaus J, Macalma T, Oliphant A, Briggs S (2002) A draft sequence of the rice genome (Oryza sativa L. ssp. japonica). Science 296: 92-100

Harjes CE, Rocheford TR, Bai L, Brutnell TP, Kandianis CB, Sowinski SG, Stapleton AE, Vallabhaneni R, Williams M, Wurtzel ET, Yan J, Buckler ES (2008) Natural genetic variation in lycopene epsilon cyclase tapped for maize biofortification. Science 319: 330-333

Hirschhorn JN, Daly MJ (2005) Genome-wide association studies for common diseases and complex traits. Nat Rev Genet 6: 95-108

Hoekenga OA, Vision TJ, Shaff JE, Monforte AJ, Lee GP, Howell SH, Kochian LV (2003) Identification and characterization of aluminum tolerance loci in *Arabidopsis* (Landsberg erecta x Columbia) by quantitative trait locus mapping. A physiologically simple but genetically complex trait. Plant Physiol 132: 936-948

Hoisington D, Khairallah M, Reeves T, Ribaut JM, Skovmand B, Taba S, Warburton M (1999) Plant genetic resources: what can they contribute toward increased crop productivity? Proc Natl Acad Sci U S A 96: 5937-5943

Huang CF, Yamaji N, Mitani N, Yano M, Nagamura Y, Ma JF (2009) A bacterial-type ABC transporter is involved in aluminum tolerance in rice. Plant Cell 21: 655-667

Jin J, Huang W, Gao JP, Yang J, Shi M, Zhu MZ, Luo D, Lin HX (2008) Genetic control of rice plant architecture under domestication. Nat Genet 40: 1365-1369

Jung KH, An G, Ronald PC (2008) Towards a better bowl of rice: assigning function to tens of thousands of rice genes. Nat Rev Genet 9: 91-101

Khush GS (1997) Origin, dispersal, cultivation and variation of rice. Plant Mol Biol 35: 25-34

Klein PE, Klein RR, Vrebalov J, Mullet JE (2003) Sequence-based alignment of sorghum chromosome 3 and rice chromosome 1 reveals extensive conservation of gene order and one major chromosomal rearrangement. Plant J 34: 605-621

Kochian LV (1995) Cellular Mechanisms of Aluminum Toxicity and Resistance in Plants. Annu Rev Plant Physiol Plant Mol Biol 46: 237-260

Kochian LV, Hoekenga OA, Piñeros MA (2004) How do crop plants tolerate acid soils? Mechanisms of aluminum tolerance and phosphorous efficiency. Annu Rev Plant Biol 55: 459-493

Kojima S, Takahashi Y, Kobayashi Y, Monna L, Sasaki T, Araki T, Yano M (2002) Hd3a, a rice ortholog of the *Arabidopsis* FT gene, promotes transition to flowering downstream of Hd1 under short-day conditions. Plant Cell Physiol 43: 1096-1105

Konishi S, Izawa T, Lin SY, Ebana K, Fukuta Y, Sasaki T, Yano M (2006) An SNP caused loss of seed shattering during rice domestication. Science 312: 1392-1396

Kovach MJ, McCouch SR (2008) Leveraging natural diversity: back through the bottleneck. Curr Opin Plant Biol 11: 193-200

Kovach MJ, Sweeney MT, McCouch SR (2007) New insights into the history of rice domestication. Trends Genet 23: 578-587

Larsen PB, Geisler MJ, Jones CA, Williams KM, Cancel JD (2005) als3encodes a phloem-localized abc transporter-like protein that is required for aluminum tolerance in *Arabidopsis*. Plant Journal 41: 353

Larsen PB, Cancel J, Rounds M, Ochoa V (2007) *Arabidopsis* ALS1 encodes a root tip and stele localized half type ABC transporter required for root growth in an aluminum toxic environment. Planta 225: 1447-1458

Li C, Zhou A, Sang T (2006) Rice domestication by reducing shattering. Science 311: 1936-1939

Lilley JM, Ludlow MM, McCouch S, O'Toole JC (1996) Locating QTL for osmotic adjustment and dehydration tolerance in rice. Journal of Experimental Botany 47: 1427-1436

Liu X, Lu T, Yu S, Li Y, Huang Y, Huang T, Zhang L, Zhu J, Zhao Q, Fan D, Mu J, Shangguan Y, Feng Q, Guan J, Ying K, Zhang Y, Lin Z, Sun Z, Qian Q, Lu Y, Han B (2007) A collection of 10,096 indica rice full-length cDNAs reveals highly expressed sequence divergence between Oryza sativa indica and japonica subspecies. Plant Mol Biol 65: 403-415

Londo JP, Chiang YC, Hung KH, Chiang TY, Schaal BA (2006) Phylogeography of Asian wild rice, Oryza rufipogon, reveals multiple independent domestications of cultivated rice, Oryza sativa. Proc Natl Acad Sci U S A 103: 9578-9583

Ma J, Bennetzen J (2004) Rapid recent growth and divergence of rice nuclear genome. Proceedings of the National Academy of Sciences of the United States of America 101: 12404-12410

Ma J, Hiradate S, Matsumoto H (1998) High aluminum resistance in buckwheat. Ii. Oxalic acid detoxifies aluminum internally. Plant Physiol 117: 753-759

Ma JF, Ryan PR, Delhaize E (2001) Aluminium tolerance in plants and the complexing role of organic acids. Trends Plant Sci 6: 273-278

Ma JF, Shen R, Zhao Z, Wissuwa M, Takeuchi Y, Ebitani T, Yano M (2002) Response of rice to Al stress and identification of quantitative trait loci for Al tolerance. Plant Cell Physiol 43: 652-659

Magalhaes JV, Garvin DF, Wang Y, Sorrells ME, Klein PE, Schaffert RE, Li L, Kochian LV (2004) Comparative mapping of a major aluminum tolerance gene in sorghum and other species in the poaceae. Genetics 167: 1905-1914

Magalhaes JV, Liu J, Guimaraes CT, Lana UG, Alves VM, Wang YH, Schaffert RE, Hoekenga OA, Piñeros MA, Shaff JE, Klein PE, Carneiro NP, Coelho CM, Trick HN, Kochian LV (2007) A gene in the multidrug and toxic compound

extrusion (MATE) family confers aluminum tolerance in sorghum. Nat Genet 39: 1156-1161

Marschner H (1995) Mineral Nutrition of Higher Plants. Academic Press, London

Martin GB, Brommonschenkel SH, Chunwongse J, Frary A, Ganal MW, Spivey R, Wu T, Earle ED, Tanksley SD (1993) Map-based cloning of a protein kinase gene conferring disease resistance in tomato. Science 262: 1432-1436

Mather K, Caicedo A, Polato N, Olsen K, McCouch S, Purugganan M (2007) The Extent of Linkage Disequilibrium in Rice (Oryza sativa L.). Genetics 177: 2223-2232

Mimmo T, Marzadori C, Gessa CE (2009) Does the degree of pectin esterification influence aluminum absorption by the root apoplast? Plant Soil 314: 159-168

Minella E, Sorrells ME (1992) Aluminum tolerance in barley: genetic relationships among genotypes of diverse origin. Crop Science 32: 593-598

Miyasaka SC, Buta JG, Howell RK (1991) Mechanism of Aluminum Tolerance in Snapbeans: Root Exudation of Citric Acid. Plant Physiol 96: 737-743

Montes JM, Melchinger AE, Reif JC (2007) Novel throughput phenotyping platforms in plant genetic studies. Trends Plant Sci 12: 433-436

Nguyen VT, Burow MD, Nguyen HT, Le BT, Le TD, Paterson AH (2001) Molecular mapping of genes conferring aluminum tolerance in rice (*Oryza sativa I.*). TAG 102: 1002-1010

Nguyen BD, Brar DS, Bui BC, Nguyen TV, Pham LN, Nguyen HT (2003) Identification and mapping of the QTL for aluminum tolerance introgressed from the new source, Oryza Rufipogon Griff., into indica rice (Oryza sativa L.). TAG 106: 583-593

Nguyen VT, Nguyen BD, Sarkarung S, Martinez C, Paterson AH, Nguyen HT (2002) Mapping of genes controlling aluminum tolerance in rice: comparison of different genetic backgrounds. Mol Genet Genomics 267: 772-780

Ni J, Pujar A, Youens-Clark K, Yap I, Jaiswal P, Tecle I, Tung CW, Ren L, Spooner W, Wei X, Avraham S, Ware D, Stein L, McCouch S (2009) Gramene QTL database: development, content and applications. Database (Oxford) 2009: bap005

Ninamango-Cardenas FE, Guimaraes CT, Martins PR, Parentoni SN, Carneiro NP, Lopes MA, Moro JR, Paiva E (2003) Mapping QTLs for aluminum tolerance in maize. Euphytica 130: 223-232

Nishimura A, Ashikari M, Lin S, Takashi T, Angeles ER, Yamamoto T, Matsuoka M (2005) Isolation of a rice regeneration quantitative trait loci gene and its application to transformation systems. Proc Natl Acad Sci U S A 102: 11940-11944

Oka HI (1988) Origin of cultivated rice. Japan Scientific Societies Press ;Amsterdam; Elsevier, Tokyo; New York

Palaisa KA, Morgante M, Williams M, Rafalski A (2003) Contrasting effects of selection on sequence diversity and linkage disequilibrium at two phytoene synthase loci. Plant Cell 15: 1795-1806

Paterson AH, Lander ES, Hewitt JD, Peterson S, Lincoln SE, Tanksley SD (1988) Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. Nature 335: 721-726

Piñeros MA, Magalhaes JV, Carvalho Alves VM, Kochian LV (2002) The physiology and biophysics of an aluminum tolerance mechanism based on root citrate exudation in maize. Plant Physiol 129: 1194-1206

Piñeros MA, Shaff JE, Manslank HS, Alves VM, Kochian LV (2005) Aluminum resistance in maize cannot be solely explained by root organic acid exudation. A comparative physiological study. Plant Physiol 137: 231-241

Rakshit S, Rakshit A, Matsumura H, Takahashi Y, Hasegawa Y, Ito A, Ishii T, Miyashita NT, Terauchi R (2007) Large-scale DNA polymorphism study of Oryza sativa and O. rufipogon reveals the origin and divergence of Asian rice. TAG 114: 731-743

Ren ZH, Gao JP, Li LG, Cai XL, Huang W, Chao DY, Zhu MZ, Wang ZY, Luan S, Lin HX (2005) A rice quantitative trait locus for salt tolerance encodes a sodium transporter. Nat Genet 37: 1141-1146

Ryan PR, DiTomaso JM, Kochian LV (1993) Aluminum toxicity in roots: an investigation of spatial sensitivity and the role of the root cap. Journal of Experimental Botany 44: 437-446

Ryan P, Delhaize E, Jones D (2001) Function and Mechanism of Organic Anion Exudation from Plant Roots. Annu Rev Plant Physiol Plant Mol Biol 52: 527-560

Sasaki T, Yamamoto Y, Ezaki B, Katsuhara M, Ahn SJ, Ryan PR, Delhaize E, Matsumoto H (2004) A wheat gene encoding an aluminum-activated malate transporter. Plant J 37: 645-653

Sax K (1923) The Association of Size Differences with Seed-Coat Pattern and Pigmentation in PHASEOLUS VULGARIS. Genetics 8: 552-560

Semon M, Nielsen R, Jones MP, McCouch SR (2005) The population structure of African cultivated rice Oryza glaberrima (Steud.): evidence for elevated levels of linkage disequilibrium caused by admixture with O. sativa and ecological adaptation. Genetics 169: 1639-1647

Shomura A, Izawa T, Ebana K, Ebitani T, Kanegae H, Konishi S, Yano M (2008) Deletion in a gene associated with grain size increased yields during rice domestication. Nat Genet 40: 1023-1028

Song XJ, Huang W, Shi M, Zhu MZ, Lin HX (2007) A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. Nat Genet 39: 623-630

Sumner ME, Noble AD (2003) Soil Acidification: The World Story. In Z Rengel, ed, The Handbook of Soil Acidity. Marcel Dekker, Inc, New York, NY, pp 1-28

Szalma SJ, Buckler ES,4th, Snook ME, McMullen MD (2005) Association analysis of candidate genes for maysin and chlorogenic acid accumulation in maize silks. TAG 110: 1324-1333

Takahashi Y, Shomura A, Sasaki T, Yano M (2001) Hd6, a rice quantitative trait locus involved in photoperiod sensitivity, encodes the alpha subunit of protein kinase CK2. Proc Natl Acad Sci U S A 98: 7922-7927

Tan L, Li X, Liu F, Sun X, Li C, Zhu Z, Fu Y, Cai H, Wang X, Xie D, Sun C (2008) Control of a key transition from prostrate to erect growth in rice domestication. Nat Genet 40: 1360-1364

Thornsberry JM, Goodman MM, Doebley J, Kresovich S, Nielsen D, Buckler ES,4th (2001) Dwarf8 polymorphisms associate with variation in flowering time. Nat Genet 28: 286-289

Tilman D, Cassman KG, Matson PA, Naylor R, Polasky S (2002) Agricultural sustainability and intensive production practices. Nature 418: 671-677

Ueda T, Sato T, Hidema J, Hirouchi T, Yamamoto K, Kumagai T, Yano M (2005) qUVR-10, a major quantitative trait locus for ultraviolet-B resistance in rice, encodes cyclobutane pyrimidine dimer photolyase. Genetics 171: 1941-1950

Vitte C, Ishii T, Lamy F, Brar D, Panaud O (2004) Genomic paleontology provides evidence for two distinct origins of Asian rice (Oryza sativa L.). Mol Genet Genomics 272: 504-511

von Uexkull HR, Mutert E (1995) Global extent, development and economic impact of acid soils. In RA Date, NJ Grundon, GE Raymet, ME Probert, eds, Plant-Soil Interactions at Low pH: Principles and Management. Kluwer Academic, Dordrecht, Netherlands, pp 5-19

Wang GL, Mackill DJ, Bonman JM, McCouch SR, Champoux MC, Nelson RJ (1994) RFLP mapping of genes conferring complete and partial resistance to blast in a durably resistant rice cultivar. Genetics 136: 1421-1434

Whitt SR, Wilson LM, Tenaillon MI, Gaut BS, Buckler ES,4th (2002) Genetic diversity and selection in the maize starch pathway. Proc Natl Acad Sci U S A 99: 12959-12962

Wilson LM, Whitt SR, Ibanez AM, Rocheford TR, Goodman MM, Buckler ES,4th (2004) Dissection of maize kernel composition and starch production by candidate gene association. Plant Cell 16: 2719-2733

Wolt J (1994) Soil Solution Chemistry: Applications to Environmental Science and Agriculture. John Wiley & Sons INC, New York, New York

Wood S, Sebastian K, Scherr SJ (2000) Soil Resource Condition. Pilot Analysis of Global Ecosystems. International Food Policy Research Institute and the World Resources Institute, Washington D.C., pp 45-54

Wu P, Liao CY, Hu B, Yi KK, Jin WZ, Ni J, He C (2000) QTLs and epistatis for aluminum tolerance in rice (*Oryza sativa* L.) at different seedling stages. Theoretical and Applied Genetics 100: 1303

Wu P, Zhao B, Yan J, Luo A, Wu Y, Senadihra D (1997) Genetic control of seedling tolerance to aluminum toxicity in rice. Euphytica 289-293

Xu K, Xu X, Fukao T, Canlas P, Maghirang-Rodriguez R, Heuer S, Ismail AM, Bailey-Serres J, Ronald PC, Mackill DJ (2006) Sub1A is an ethylene-response-factor-like gene that confers submergence tolerance to rice. Nature 442: 705-708

Xu Y, This D, Pausch RC, Vonhof WM, Coburn JR, Comstock JP, McCouch SR (2009) Leaf-level water use efficiency determined by carbon isotope discrimination in rice seedlings: genetic variation associated with population structure and QTL mapping. Theoretical and Applied Genetics 118: 1065-1081

Xue Y, Wan J, Jiang L, Wang C, Liu L, Zhang YM, Zhai H (2006) Identification of quantitative trait loci associated with aluminum tolerance in rice. Euphytica 150: 37-45

Xue W, Xing Y, Weng X, Zhao Y, Tang W, Wang L, Zhou H, Yu S, Xu C, Li X, Zhang Q (2008) Natural variation in Ghd7 is an important regulator of heading date and yield potential in rice. Nat Genet 40: 761-767

Xue Y, Jiang L, Su N, Wang JK, Deng P, Ma JF, Zhai HQ, Wan JM (2007) The genetic basic and fine-mapping of a stable quantitative-trait loci for aluminium tolerance in rice. Planta 227: 255-262

Yamaji N, Huang CF, Nagao S, Yano M, Sato Y, Nagamura Y, Ma JF (2009) A Zinc Finger Transcription Factor ART1 Regulates Multiple Genes Implicated in Aluminum Tolerance in Rice. Plant Cell

Yamamoto T, Yonemaru J, Yano M (2009) Towards the understanding of complex traits in rice: substantially or superficially. 16: 1-14

Yang JL, Li YY, Zhang YJ, Zhang SS, Wu YR, Wu P, Zheng SJ (2008) Cell wall polysaccharides are specifically involved in the exclusion of aluminum from the rice root apex. Plant Physiol 146: 602-611

Yano M, Katayose Y, Ashikari M, Yamanouchi U, Monna L, Fuse T, Baba T, Yamamoto K, Umehara Y, Nagamura Y, Sasaki T (2000) Hd1, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the *Arabidopsis* flowering time gene CONSTANS. Plant Cell 12: 2473-2484

Yu J, Hu S, Wang J, Wong GK, Li S, Liu B, Deng Y, Dai L, Zhou Y, Zhang X, Cao M, Liu J, Sun J, Tang J, Chen Y, Huang X, Lin W, Ye C, Tong W, Cong L, Geng J, Han Y, Li L, Li W, Hu G, Huang X, Li W, Li J, Liu Z, Li L, Liu J, Qi Q, Liu J, Li L, Li T, Wang X, Lu H, Wu T, Zhu M, Ni P, Han H, Dong W, Ren X, Feng X, Cui P, Li X, Wang H, Xu X, Zhai W, Xu Z, Zhang J, He S, Zhang J, Xu J, Zhang K, Zheng X, Dong J, Zeng W, Tao L, Ye J, Tan J, Ren X, Chen X, He J, Liu D, Tian W, Tian C, Xia H, Bao Q, Li G, Gao H, Cao T, Wang J, Zhao W, Li P, Chen W, Wang X, Zhang Y, Hu J, Wang J, Liu S, Yang J, Zhang G, Xiong Y, Li Z, Mao L, Zhou C, Zhu Z, Chen R, Hao B, Zheng W, Chen S, Guo W, Li G, Liu S, Tao M, Wang J, Zhu L, Yuan L, Yang H (2002) A draft sequence of the rice genome (Oryza sativa L. ssp. indica). Science 296: 79-92

Yu J, Pressoir G, Briggs WH, Vroh Bi I, Yamasaki M, Doebley JF, McMullen MD, Gaut BS, Nielsen DM, Holland JB, Kresovich S, Buckler ES (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat Genet 38: 203-208

Zeigler RS, Barclay A (2008) The Relevance of Rice. Rice 1: 3-10

Zhang WH, Ryan PR, Tyerman SD (2001) Malate-permeable channels and cation channels activated by aluminum in the apical cells of wheat roots. Plant Physiol 125: 1459-1472

Zhao K, Wright M, Kimball K, Eizenga G, McClung A, Kovach M, Tyagi W, Ali L, Tung C, Reynolds A, Bustamante C, McCouch S (2010) Genomic Diversity and Introgression in *O. sativa* Reveal the Impact of Domestication and Breeding on the Rice Genome. PLOS One 5:

Zhao K, Aranzana MJ, Kim S, Lister C, Shindo C, Tang C, Toomajian C, Zheng H, Dean C, Marjoram P, Nordborg M (2007) An *Arabidopsis* example of association mapping in structured samples. PLoS Genet 3: e4

Zhou H, Xie Z, Ge S (2003) Microsatellite analysis of genetic diversity and population genetic structure of a wild rice (Oryza rufipogon Griff.) in China. Theoretical and Applied Genetics 107: 332-339

Zhu Q, Ge S (2005) Phylogenetic relationships among A-genome species of the genus Oryza revealed by intron sequences of four nuclear genes. New Phytologist 167: 249-265

CHAPTER 2:
DEVELOPMENT OF A NOVEL ALUMINUM TOLERANCE PHENOTYPING PLATFORM USED FOR COMPARISON OF CEREAL Al TOLERANCE AND INVESTIGATIONS INTO RICE Al TOLERANCE MECHANISMS

ABSTRACT

The genetic and physiological mechanisms of aluminum (Al) tolerance have been well studied in certain cereal crops, and Al tolerance genes have been identified in sorghum and wheat. Rice (*Oryza sativa* L.) has been reported to be highly Al tolerant; however, a direct comparison of rice and other cereals has not been reported and the mechanisms of rice Al tolerance are poorly understood. To facilitate Al tolerance phenotyping in rice, a high-throughput imaging system and root quantification computer program was developed, permitting quantification of the entire root system, rather than just the longest root. Additionally, a novel hydroponic solution was developed and optimized for Al tolerance screening in rice and compared to the Yoshida's rice solution commonly used for rice Al tolerance studies. To gain a better understanding of Al tolerance in cereals, comparisons of Al tolerance across cereal species were conducted at four Al concentrations using seven to nine genetically diverse genotypes of wheat, maize, sorghum, and rice. Rice was significantly more tolerant than maize, wheat, and sorghum at all Al concentrations, with the mean Al tolerance level for rice found to be 2-6-fold greater than that in maize, wheat, and sorghum. Physiological experiments were conducted on a genetically diverse panel of >20 rice genotypes spanning the range of rice Al tolerance and compared to two maize genotypes to determine if rice utilizes the well-described Al tolerance mechanism of root tip Al exclusion mediated by organic acid exudation. Results clearly demonstrate

that the extremely high levels of rice Al tolerance are mediated by a novel mechanism, which is independent of root tip Al exclusion.

INTRODUCTION

Aluminum (Al) is the most abundant metal in the earth's crust, comprising approximately 7% of the soil (Wolt, 1994). Al is predominantly found as a key component of soil clays; however, under highly acidic soil conditions (pH<5.0), Al^{3+} is solubilized into the soil solution and is highly phytotoxic. Al^{3+} causes a rapid inhibition of root growth that leads to a reduced and stunted root system, thus having a direct effect on the ability of a plant to acquire both water and nutrients. Approximately 30% of the world's total land area and over 50% of potentially arable lands are acidic, with the majority (60%) found in the tropics and subtropics (von Uexkull and Mutert, 1995). Thus, acidic soils are a major limitation to crop production, particularly in the developing world.

As a whole, cereal crops (*Poaceae*) provide an excellent model for studying Al tolerance because of their abundant genetic resources, large, active research communities, and importance to agriculture. In addition, work in one cereal species can rapidly translate into impact throughout the family. Previous research has focused on understanding the genetic and physiological mechanisms of Al tolerance in maize, sorghum, and wheat. The most recognized physiological mechanism conferring Al tolerance in plants involves exclusion of Al from the root tip (Miyasaka et al, 1991; Delhaize and Ryan, 1995; Kochian, 1995; Kochian et al, 2004 a, b). The exclusion mechanism is primarily mediated by Al-activated exudation of organic acids such as malate, citrate, or oxalate from the root apex, the site of Al toxicity (Ryan et al, 1993;

Ma et al, 2001; Ryan et al., 2001). These organic acids chelate Al in the rhizosphere, reducing the concentration and toxicity of Al at the growing root tip (Ma et al., 2001). Phosphate has also been identified as a class of root exudates involved in cation chelation and can therefore be considered a potential exudate involved in Al exclusion from the root tip (Pellet et al., 1996).

Al-activated malate and citrate anion efflux transporters have been cloned from wheat (*ALMT1*) (Sasaki et al., 2004) and sorghum (*SbMATE*) (Magalhaes et al, 2007), and root citrate efflux transporters have been implicated in Al tolerance in maize (Zhang et al., 2001; Piñeros et al., 2001). Recently, a maize homolog of sorghum *SbMATE* was shown to be the root citrate efflux transporter that plays a role in maize Al tolerance (Maron et al., 2010). Although organic acids have been shown to play a major role in Al tolerance in these species, another exclusion mechanism has been identified in an *Arabidopsis* mutant, where a root mediated increase in rhizosphere pH lowers the Al^{3+} activity and thus participates in Al exclusion from the root apex (Degenhardt et al., 1998). Furthermore, there is clear evidence that tolerance in maize cannot be fully explained by organic acid release (Piñeros et al., 2005). These types of findings strongly suggest that multiple Al tolerance mechanisms exist in plants.

Rice has been reported to be the most Al-tolerant cereal crop under field conditions, capable of withstanding significantly higher concentrations of Al than other major cereals (Foy, 1988). Despite this fact, very little is known about the physiological mechanisms of Al tolerance in rice. Two independent studies have identified increased Al accumulation in the root apex in susceptible compared to Al tolerant rice

varieties, but no differences were observed in organic acid exudation or rhizosphere pH (Ma et al., 2002; Yang et al., 2008). These studies suggest that rice may contain novel physiological and/or genetic mechanisms that confer significantly higher levels of Al tolerance than those found in other cereals. A more thorough analysis is required to clarify the mechanism of Al tolerance in rice.

Cultivated rice (*O. sativa*) is characterized by deep genetic divergence between the two major varietal groups: *Indica* and *Japonica* (Londo et al., 2006; Hu et al., 2006; Garris et al., 2005; Dally and Second, 1990). Extensive selection pressure over the last 10,000 years has resulted in the formation of five genetically distinct subpopulations: *indica* and *aus* within the *Indica* varietal group, and *temperate japonica*, *tropical japonica* and *aromatic/groupV*, within the *Japonica* varietal group (Garris et al., 2005; Caicedo et al., 2007; Zhao and McCouch, personal communication). (Note: when referring to varietal groups, the first letter will be capitalized, while lower case letters will be used to refer to the subpopulation groups.) Subpopulation differences in trait performance are often significant, particularly with respect to biotic and abiotic stress (Champoux et al., 1995; Lilley et al, 1996; Garris et al. 2003; Xu et al., 2009). This can lead to confusion because trait or performance differences may be confounded with subpopulation structure, leading to false positives (Type 1 error) (Devlin and Roeder, 1999; Pritchard and Donnelly, 2001; Yu et al., 2006; Zhao et al., 2007). Therefore, it is important to consider the subpopulation origin of genotypes being compared when studying the genetics and physiology of Al tolerance in rice.

Al tolerance screening is typically conducted by comparing root growth of seedlings grown in hydroponic solutions, with and without Al (Sasaki, 2004; Magalhaes et al., 2004; Piñeros et al, 2001). Sorghum and maize are often screened for Al tolerance in Magnavaca's nutrient solution (Magalhaes et al., 2004; Piñeros et al, 2001, 2005), while rice seedlings are typically grown in Yoshida's solution (Yoshida, 1978). Furthermore, Al concentrations used to screen for Al tolerance in maize (222μM), sorghum (148μM), and wheat (100μM) are significantly lower than those used for screening Al tolerance in rice (1,112-1,482 μM) (Wu et al., 2000; Nguyen VT et al., 2001, 2002; Nguyen BD et al., 2003). These differences in chemical composition of the nutrient solutions make it difficult to directly compare plant response to Al across these cereals. In rice, the high Al concentrations required to observe significant differences in root growth between susceptible and resistant varieties also complicate Al tolerance screening due to the precipitation of Al, along with other elements. The result is that control (-Al) and treatment (+Al) solutions may differ with regard to essential mineral nutrients that react with Al, leading to differences in growth not directly attributable to Al. Additionally, because the active form of Al that is toxic to root growth is Al^{3+}, any Al that precipitates out of solution has no effect on root growth (Kochian, 2004a). In a hydroponic solution, Al may be found in one of four forms, (1) as free Al^{3+} where it actively inhibits root growth, (2) precipitated with other elements and essentially unavailable to inhibit plant growth, (3) as different hydroxyl monomers of Al which are not believed to be toxic to roots (Parker et al, 1988), or (4) complexed with other elements in an equilibrium between its active and inactive states. The degree to which Al inhibits root growth is primarily dependent upon the activity of free Al^{3+} in solution (Kochian, 2004a).

The objectives of this study were to: 1) develop and optimize a suitable nutrient solution and high-throughput Al tolerance screening method for rice; 2) quantify and compare differences in Al tolerance between maize, sorghum, wheat, and rice; and 3) use the developed screening methods to determine if rice utilizes the organic acid-mediated Al exclusion mechanism that is observed in maize, sorghum, and wheat.

RESULTS

Optimization of nutrient solution composition for Al tolerance screening in rice

To establish a hydroponic solution for screening Al tolerance in rice seedlings that would allow us to compare levels of tolerance between rice and other cereals, we modified the Magnavaca's nutrient solution (Magnavaca et al., 1987) that has been previously used for maize and sorghum Al tolerance research (Magalhaes et al., 2004; Piñeros et al, 2001, 2005). Modifications were made to ensure a sufficient supply of essential nutrients and to minimize the chemical interactions between Al and other mineral species in the nutrient solution at the high Al concentrations needed for rice. Using the chemical speciation program Geochem-EZ (Shaff et al., 2010; http://www.plantmineralnutrition.net/Geochem/geochem%20home.htm) and Inductively Coupled Plasma Emission Spectrometry (ICP-ES) analysis, we first analyzed the chemical composition of the Yoshida's solution to understand what was causing the visible precipitate that was always observed when Al was supplied at necessarily high concentrations above 1 mM $AlCl_3$. These concentrations have been shown to cause a measurable inhibition of rice root growth (Wu et al., 2000; Nguyen VT et al., 2001, 2002; Nguyen BD et al., 2003). Geochem-EZ predicted that when 1,297μM $AlCl_3$ (35ppm) was added to Yoshida's solution, it would only result in a free Al^{3+} activity of 116 μM, with the significant reduction in Al^{3+} activity due to Al

interaction with HPO_4^{2-} , SO_4^{2-}, and citrate (citrate was used as the Fe chelate). The observed precipitation in the Yoshida's solution was predicted by Geochem-EZ to be due to Al precipitating with the high concentrations of SO_4^{2-} and HPO_4^{2-} in the solution, and Fe precipitating with HPO_4^{2-}. ICP-ES analysis of Yoshida's control and Al-containing solutions confirmed that, in addition to the differences in Al concentrations, there were significant differences in P and Fe availability. Chemical analysis of the nutrient solution to determine soluble and precipitated minerals identified that available P was reduced 85.8% (from 321 to 45.6 µM) and available Fe was reduced 85.2% (from 35.8 µM to 5.3 µM) (Table 2.1). Furthermore, 40% of the total Al added to the solution was lost as a precipitate.

Table 2.1. Comparison of total (supplied) and soluble P, Fe, and Al in Modified Magnavaca's and Yoshida's solutions: Total P and Fe is the total concentration of phosphate and chelated Fe provided in the control (-Al) nutrient solutions. Available P and Fe is the concentration of soluble P and Fe measured in the Al-containing solutions using ICP-ES after centrifugation to pellet out precipitated P and Fe. %P and %Fe Decrease is the difference in soluble P or Fe concentrations between control and Al-treated nutrient solutions. Total Al is the concentration of Al added to the treatment solution (as $AlCl_3$), and Soluble Al is the amount of soluble Al (not precipitated) in each nutrient solution as determined by centrifugation followed by ICP-ES analysis. The % soluble Al quantifies the percent of the total Al added that is in a soluble state as determined by chemical analysis. The Al^{3+} activity values in the last column were predicted using the GEOCHEM-EZ speciation program based on chemical equilibrium constants for each nutrient solution.

	Total P	Available P	% P Decrease	Total Fe	Available Fe	% Fe Decrease	Total Al	Soluble Al	% Soluble Al	Al^{3+} Activity (GEOCHEM-EZ)
Mod. Mag.	47.8 µM	34.8 µM	27.2 %	77 µM	68.1 µM	11.5 %	540 µM	517µM	95.7 %	160 µM
Yos.	321 µM	45.6 µM	85.8 %	35.8 µM	5.3 µM	85.2 %	1297 µM	775 µM	59.7 %	116 µM

To address these problems, we developed an optimized nutrient solution, hereafter referred to as Modified Magnavaca's solution, which minimizes the concentration of Al necessary to elicit significant levels of root growth inhibition in rice seedlings (Table 2.2). We accomplished this by reducing the ionic strength of the nutrient solution and reducing the interactions between Al and other mineral ions.

Table 2.2: Nutrient composition of the Modified Magnavaca's nutrient solution optimized for rice Al tolerance screening. Key differences between this solution and the standard rice Al tolerance screening Yoshida's solution include reduced P and S concentrations and an Fe-HEDTA chelate, replacing the citrate Fe chelate used in Yoshida's solution

Compound	Concentration
KCl	1mM
NH_4NO_3	1.5mM
CaCl	1mM
KH_2PO_4	45µM
$MgSO_4$	200µM
$Mg(NO_3)_2$	500µM
MgCl2	155µM
$MnCl_24H_2O$	11.8µM
H_3BO_3	33µM
$ZnSO_47H_2O$	3.06µM
$CuSO_45H_2O$	0.8µM
$Na_2MoO_4H_2O$	1.07µM
Fe-HEDTA	77µM

Geochem-EZ predictions and preliminary plant growth studies suggested that the optimal total concentration of Al required for screening of Al tolerance in rice would be 540 μM $AlCl_3$ (~60% less than Yoshida's solution), yielding an Al^{3+} activity of 160 μM in the Modified Maganvaca's solution. Our solution also contained much lower levels of the ions that most strongly interact with Al^{3+} (SO_4^{2-} and $H_2PO_4^-$), and used HEDTA as the Fe chelate, rather than citrate, which preferentially binds Al over Fe. ICP-ES analysis demonstrated that when the Modified Magnavaca's +Al treatment solution is compared to the control (-Al) solution, only 4.3% of the total Al was precipitated (in contrast to 40% in Yoshida's solutions), available P was only reduced 27% (in contrast to 85.8% in Yoshida's solution), and available Fe was reduced by 11.5% (in contrast to 85.2% in Yoshida's solution) (Table 2.1).

Plant growth experiments were conducted using seven diverse rice genotypes to investigate whether the two nutrient solutions had different effects on seedling root growth under control conditions (-Al) and whether there were differential responses to the total Al added to each solution. The average total root growth (TRG) of the seven genotypes after three days of growth in the two control solutions were virtually identical, 60.58 cm in Modified Magnavaca solution and 59.47 cm in Yoshida's solution (Figure 2.1A). However, when the same $AlCl_3$ concentrations were used in the two treatment solutions, the average TRG of the seven genotypes was 40-50% less in the Modified Magnavaca's solution than in the Yoshida's solution. At a total concentration of 540 μM $AlCl_3$, TRG averaged 30.34 cm in the modified Magnavaca's solution (RRG=0.50) and 49.95 cm in the Yoshida's solution (RRG=0.84). At the 1,297 μM $AlCl_3$ concentration, root growth averaged 17.12 cm in modified Magnavaca's (RRG=0.28) and 33.38 cm in the Yoshida's solution (RRG=0.56)

(Figure 2.1A). These results demonstrate that inhibition of TRG is not determined by the total amount of Al added to a solution, but rather by the activity of available Al^{3+} in the solution. Figure 2.1A and 2.1B display the correlation coefficients of TRG as a function of total Al added (R^2=0.40) and available soluble Al (not precipitated) (R^2=-0.86), demonstrating that available soluble Al is a much better predictor of root growth inhibition than total Al added.

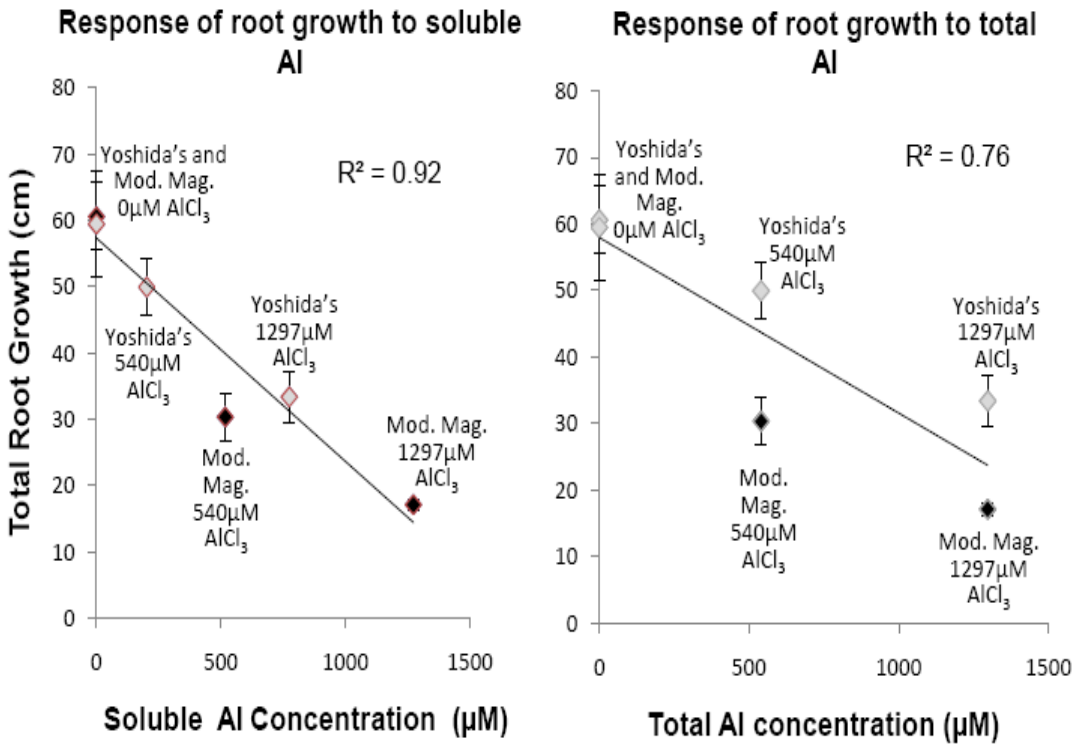


Figure 2.1. Mean total root growth (+/- sd) of seven rice genotypes in Yoshida's (grey diamonds) and Modified Magnavaca's (black diamonds) control and Al solutions. The Al concentrations represent previously reported concentrations for rice Al tolerance screening in Yoshida's (1297µM) and concentrations for Modified Magnavaca's (540µM) determined in this study. In control solutions (0µM) total root growth is identical. **A)** Total root growth in response to the concentration of soluble Al in Yoshida's and Modified Magnavaca nutrient solutions (R^2=0.92) **B)** Total root growth in response to concentration of total Al in Yoshida's and Modified Magnavaca's nutrient solutions. (R^2=0.76).

Al tolerance phenotyping platform

Relative root growth (RRG) of the longest root is the most commonly used parameter for estimating Al tolerance in cereals. We compared estimates of Al tolerance based on RRG of the longest root and RRG of the total root system to determine whether the longest root measurement would serve as a useful proxy for estimating the inhibition of Al on total root growth of rice seedlings. Using 225 genetically diverse *O. sativa* accessions (a subset of an association mapping panel) (Zhao and McCouch, personal communication), the correlation coefficient for the relationship between RRG of the longest root and RRG of the total root system was R^2=0.172 (Supplemental Figure S2.1). Based on this analysis, it was determined that the RRG of the longest root was not a good proxy for RRG of the total root system because a genotype may appear tolerant based on longest root measurements when, in fact, total root growth is inhibited (Figure 2.2). To obtain accurate estimations of total root growth, we used a custom root digital imaging system developed in our labs to quantify root length parameters for the thin, fibrous root systems of rice. The system was based on digital photography and semi-automatic measurement of individual primary, secondary, and tertiary roots using RootReader2D software (see Materials and Methods for details). In this system, the length of the total root system can be reliably measured and we are able to capture high quality digital images of each root system.

Figure 2.2. Rice root system image used for quantification. Example where growth of the longest root in an Al grown (right) and control grown (left) rice seedling is similar, but total root growth is significantly different. Images are of plants representative of the mean growth of the longest root in control (-Al) and treatment (+Al) solutions for genotype NSF4. The mean longest root growth was 1.8 (+/-0.14) cm in control solution and 2.0 (+/-0.18) cm in treatment solution. However, the mean total root growth was 50.29 (+/-7.3) cm in control and 27.10 (+/-2.47) cm in treatment. The mean longest root RRG was 1.11, however the mean total root RRG was only 0.54.

Comparison of Al tolerance between cereal species

When Al tolerance was directly compared between diverse genotypes of maize, sorghum, wheat, and rice at three Al^{3+} activity levels (8.75 µM, 27 µM, and 160 µM), rice was consistently more tolerant than the other cereals, maize was intermediate, and sorghum and wheat were the most sensitive (Figure 2.3). The genotypes used in this analysis were selected to represent the range of known Al tolerance within each species; that is, we selected varieties classified as Al sensitive, intermediate, and tolerant for each species to ensure adequate representation of variation within as well as between the species. At all Al^{3+} activities, the order of Al tolerance among the four cereal species was consistent: *rice* > *maize* > *sorghum* ≥ *wheat*.

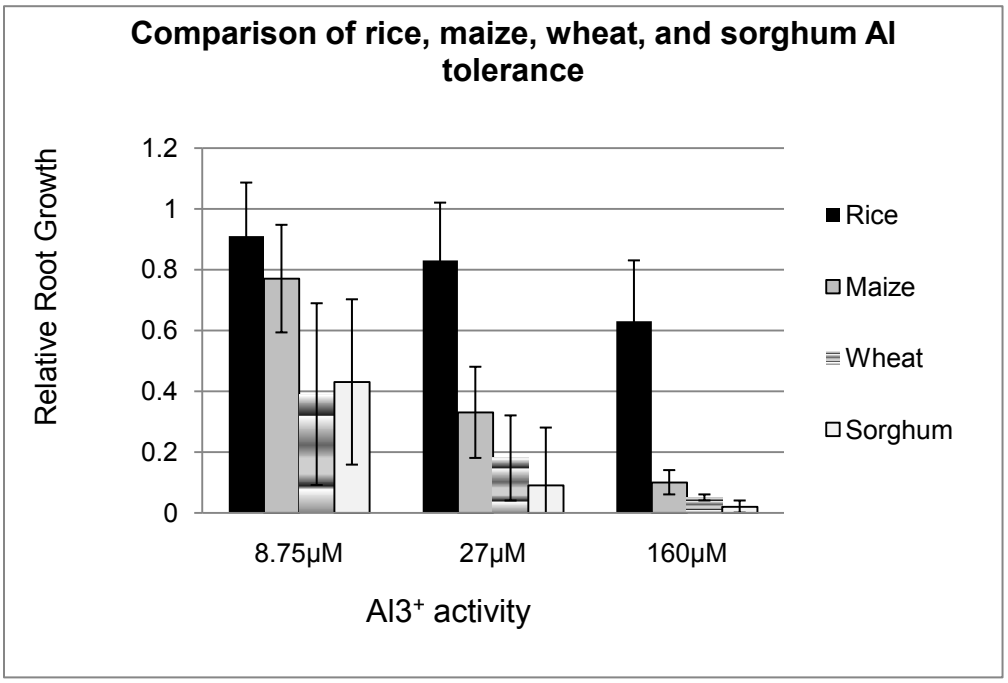


Figure 2.3. Comparison of Al tolerance across cereals. Average Al tolerance (RRG) of rice (n=8), maize (n=9), wheat (n=8), and sorghum (n=7) at three Al^{3+} activities +/- standard deviation.

To ensure that we were able to observe the full distribution of Al tolerance in each of the four cereal species, we evaluated RRG of the total root systems at three different concentrations of Al, based on previous studies. The use of sub- or super-optimal Al concentrations tends to mask the range of natural variation for Al tolerance that exists within each species (Supplemental Table S2.1). To avoid this, we used a free Al^{3+} activity of 8.75 μM that had been optimized for wheat (Sasaki et al., 2004), 27 μM as reported for sorghum and maize (Magalhaes, 2004), and, based on the results of this study, we used an Al^{3+} activity of 160 μM as optimal for rice.

Figure 2.4 displays the mean RRG and the standard deviation observed among the accessions of wheat, sorghum, maize, and rice at each of the Al^{3+} activities employed. An Al^{3+} activity of 8.75 μM produced the highest SD and widest range of Al tolerance within wheat and sorghum. The eight wheat genotypes and seven sorghum genotypes screened at 8.75 μM Al^{3+} displayed similar means and ranges of variation. Maize and rice were both significantly more Al tolerant ($p>0.007$) than wheat and sorghum at 8.75 μM Al^{3+}. The RRG in rice and maize was not significantly different and was close to 1 (little or no inhibition of root growth) (Figure 2.4). Two rice genotypes exhibited increased root growth (RRG > 1) at 8.75 μM Al^{3+} compared to their root growth under control conditions. Although, on average, rice and maize were both more Al tolerant than wheat and sorghum, there was overlap between the species, with the most tolerant genotype of wheat and sorghum being more tolerant than the most sensitive genotypes of maize and rice.

At 27 μM Al^{3+} root growth was severely inhibited in wheat and sorghum and significantly reduced in maize, while minimal root growth inhibition was observed in rice. The most tolerant wheat variety was Atlas 66 (RRG=0.58), which was the source

of the Al tolerance allele in *ALMT1* (Sasaki et al., 2004). In sorghum, root growth was inhibited by >90% in six of the varieties, while SC566 was considerably more tolerant than the other genotypes. SC566 is similar in Al tolerance to SC283, the donor of the tolerance allele for the sorghum Al tolerance gene, Alt_{SB} (Magalhaes et al., 2007; Caniato et al., 2007). The Al^{3+} activity of 27 μM produced a significant decrease in mean RRG in maize, but a clear distribution of Al tolerance was observed. At this activity it is clear that rice is more tolerant than the other cereals screened, as the mean RRG for rice was 2.5 times greater than that of maize and 4 times that of wheat. The two most tolerant rice varieties at 27 μM Al^{3+}, Cybonnet (RRG=1.09) and Nipponbare (RRG=1.16), demonstrated increased growth compared to the control. The least tolerant rice variety at 27 μM Al^{3+}, China 1039 (RRG=0.6), was more tolerant than nearly all the other genotypes of maize, sorghum, and wheat. Thus, at 27 μM Al^{3+} rice was significantly more tolerant than all the other species examined ($p<0.001$**).**

The differences in tolerance between rice and the other species became even more apparent at 160 μM Al^{3+}. Growth was essentially halted in all sorghum and wheat genotypes screened and severely inhibited in all maize genotypes. Of the nine maize genotypes screened, root growth in all but two genotypes was inhibited over 90% (RRG<0.1). The two most tolerant maize genotypes at 160 μM Al^{3+} were B57 (RRG=0.13) and Cateto (RRG=0.17). Cateto is a Brazilian variety bred for acid soils and is known to be one of the most Al tolerant maize varieties (Piñeros et al., 2005). In 160 μM Al^{3+}, the eight rice genotypes had a mean RRG of 0.63 +/- 0.2, and a range of 0.25 to 0.95. At this Al^{3+} activity the most susceptible rice variety, Kasalath (RRG=0.25), shows significantly higher relative root growth than that of the most tolerant maize variety, Cateto (RRG=0.17). These results clearly demonstrate that as a species, rice is significantly more Al tolerant than maize, sorghum, and wheat.

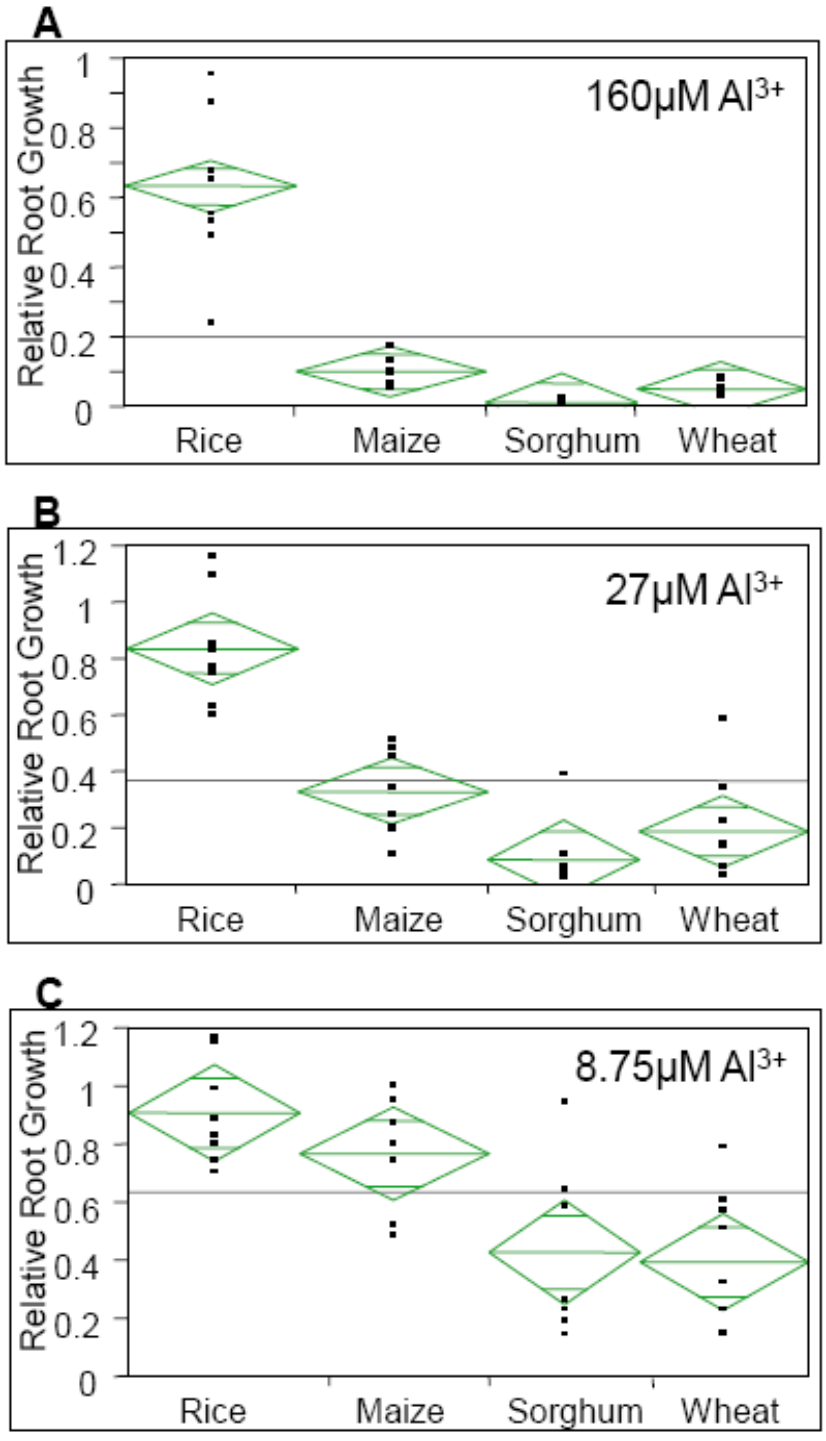


Figure 2.4. Phenotypic distribution of Al tolerance within rice, maize, sorghum, and maize at three Al^{3+} activities.

Investigation into the role of Al exclusion in rice Al tolerance

An Al tolerance diversity panel of 23 rice genotypes, representing the genetic and Al tolerance diversity of the *Indica* and *Japonica* varietal groups was evaluated to determine if rice Al tolerance involves root apex Al exclusion, as it does in other cereals. Al tolerance was screened at 160µM Al^{3+} and ranged from 0.15 to 0.97 RRG among all genotypes, with a mean value of 0.55 (SD=0.21) (Supplemental Table S2.2). The *Japonica* varietal group (n=11) had a mean tolerance value of 0.69 (SD=0.16) and RRG ranged from 0.34 to 0.97. The *Indica* varietal group (n=12) was generally more susceptible than *Japonica,* with a mean tolerance value of 0.42 (SD=0.18) and a range from 0.15 to 0.97.

The mean root apex (1cm) Al concentration among all rice genotypes was 3,027 µg Al/gram (SD=889) and ranged from 326 to 4,846. The mean root apex Al concentration was not significantly different between the *Indica* (3217 µg Al/gram, SD=724) and *Japonica* (2875 µg Al/gram, SD=1065) varietal groups (Supplemental Table S2). The correlation of Al tolerance by Al exclusion across all rice genotypes demonstrated that there is no relation between Al exclusion and Al tolerance in rice (R^2=0.002) (Figure 2.5). Similar results were obtained when each varietal group was analyzed separately; *Indica* (R^2=0.01) and *Japonica* (R^2=0.0).

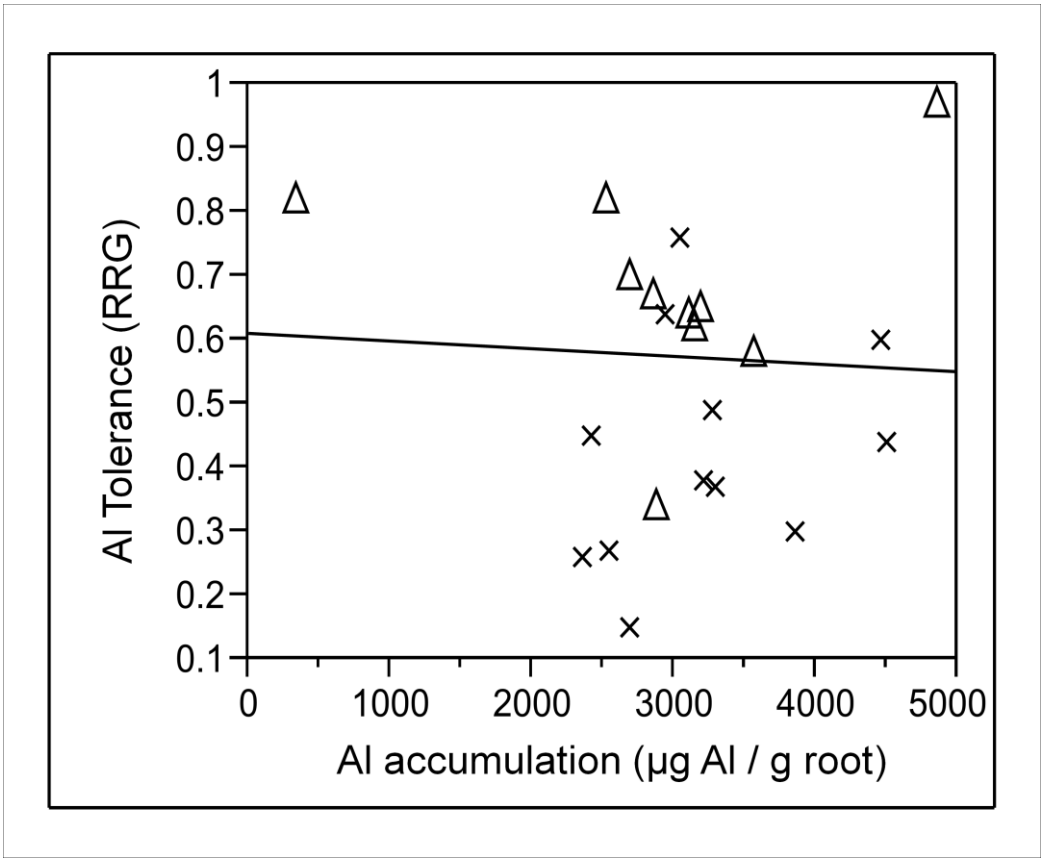


Figure 2.5. Relationship between Al accumulation and Al tolerance in rice. Correlation of root tip (1cm) Al accumulation and Al tolerance across 23 genetically diverse rice genotypes. Genotypes were selected to represent the genetic and Al tolerance variation across the Indica (x) and Japonica (triangles) varietal groups. Note that there is no correlation between root tip Al accumulation and Al tolerance (R^2=0.002).

Two maize controls were included in the above analysis for comparison between the two species, and each maize line was highly susceptible at 160µM Al^{3+}. The Al tolerant maize line, Cateto, and the Al sensitive maize line, B73, had RRG values of 0.17 and 0.13, respectively. These genotypes differed considerably with regards to Al exclusion; Cateto had a mean root apex Al concentration of 2,192 µg Al/gram (SE=74.4) compared to 4,062 µg Al/gram (SE=140) in B73. The mean root apex Al concentration of rice was >900 µg Al/g root higher than that of Cateto, which was

much more susceptible at 160μM Al^{3+} than any rice genotype. Three rice genotypes accumulated higher Al concentrations than the B73 genotype, yet were between 3-7 times more Al tolerant (Supplemental Table S2.2).

Investigation into the role of root exudates in rice Al tolerance

Root exudation of citrate, malate, and phosphate, the three root excreted Al-binding ligands implicated in cereal Al tolerance, was quantified in 21 rice genotypes (10 *Japonica* and 11 *Indica*) evaluated under control (-Al) and treatment (+Al) conditions (Supplemental Table S2.2). Across all rice genotypes, the mean citrate exudation (in Al+) was 37.1 pmole $plant^{-1}$ day^{-1} (SD=39.8), the mean malate exudation was 47.4 pmole $plant^{-1}$ day^{-1} (SD=47.4), and the mean phosphate exudation was 102.4 pmole $plant^{-1}$ day^{-1} (SD=94.2). Regression analysis of root exudate by Al tolerance score (RRG) revealed no significant relationship between root exudation and Al tolerance for citrate, malate or phosphate (Supplemental Figure S2.2).

When the relationship between root exudates and Al tolerance was analyzed independently in each rice varietal group, it revealed that within the *Indica* group, there was a small correlation between Al tolerance and malate (R^2=0.24) and phosphate exudation (R^2=0.13) (Figure 2.6A-2.6C). In the more Al tolerant *Japonica* group, there was a negative correlation between Al tolerance and phosphate exudation (R^2=-0.18) and no correlation between citrate and malate exudation (Data not shown).

In tolerant maize variety, Cateto, which has previously been reported to utilize an Al-activated citrate exudation Al tolerance mechanism (Piñeros et al., 2005), we observed Al-activated citrate exudation, and exudation rates were significantly higher in Cateto than in any rice variety. The citrate exudation rate of Cateto roots grown in treatment

solution (+Al) was 288.3 pmole plant^{-1} day^{-1} (SE=66.1), compared to 76.4 pmole plant^{-1} day^{-1} (SE=8.8) when seedlings were grown under control (-Al) conditions. Under Al stress, the citrate exudation rate of Cateto was over 6X that of any rice genotype; however, Cateto was more sensitive to Al than any rice variety (Supplemental Table S2.2).

Investigation into the role of root exudates in Al exclusion

When levels of organic acid exudation were compared with Al accumulation in root apices across all rice genotypes, we observed a slightly negative correlation between citrate exudation and root tip Al concentration (R^2=0.06), and no relationship between malate or phosphate exudation and root tip Al accumulation (Supplemental Figure. 2.2).

When exudation levels were compared within each varietal group independently, there was a significantly negative correlation in the *Indica* group between citrate exudation (R^2=0.47) and Al accumulation, and a slightly negative correlation between malate (R^2=0.07) and phosphate (R^2=0.075) exudation and root tip Al accumulation (Figure 2.6D-2.6F). Therefore, it appears that in the *Indica* varietal group, citrate exudation is associated with Al exclusion from the root apex, but this Al exclusion does not confer Al tolerance. In the *Japonica* varietal group, we observed a negative correlation between root exudation and Al accumulation, or between either parameter and Al tolerance.

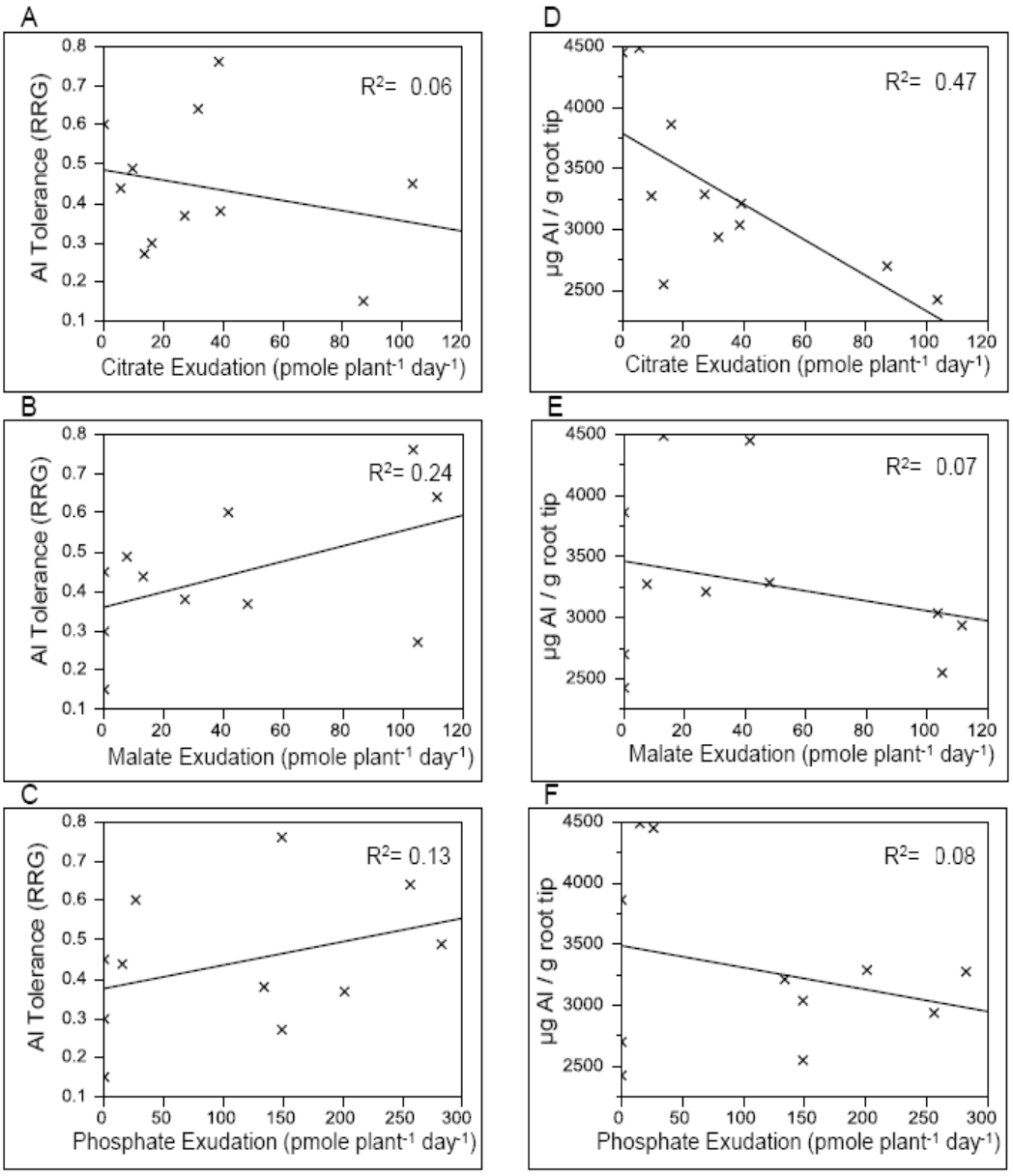


Figure 2.6. Correlation of root exudates and Al accumulation and Al tolerance within *Indica*. Correlation of root exudates of: A, D) citrate, B, E) malate, and C, F) phosphate with Al tolerance (RRG) in the left column and root tip Al content in the right column for 11 genetically diverse *Indica* varieties. A significant negative correlation is observed between root citrate exudation and root tip Al content (D). However, there is no correlation between root citrate exudation and Al tolerance. There is a slight correlation between malate exudation and Al tolerance, however there is no relation between malate exudation and root tip Al exclusion (A). For the rest of the parameters, there is either no or very weak correlations.

Gene expression analysis of the rice MATE homolog of SbMATE, the sorghum Al tolerance gene

To investigate if the rice homolog (LOC_Os01g69010) of the sorghum *MATE* Al tolerance gene (*SbMATE*) is involved in rice Al tolerance, quantitative RT-PCR was conducted in root tips (1cm) of four diverse genotypes; Azucena (RRG=0.82; *tropical japonica*), Nipponbare (RRG=0.82; *temperate japonica*), IR64 (RRG=0.45; *indica*), and Kasalath (RRG=0.15; *aus*). Under control conditions, Azucena had significantly higher levels of root *MATE* expression than any of the other varieties (Figure 2.7). Al treatment reduced *MATE* gene expression in Azucena, and no root citrate exudation was observed under either control or Al treatment, while Al treatment increased *MATE* gene expression and citrate exudation in Nipponbare, IR64 and Kasalath. Under Al stress, Nipponbare exhibited the highest *MATE* gene expression and Kasalath had the lowest. Nipponbare had significantly higher *MATE* gene expression than Azucena ($p=0.1$), though both varieties were equally Al tolerant. Additionally, Azucena accumulated less Al (326 μg Al/gram) in the roots than any of the other varieties, nearly 8 times less than that of Nipponbare (2523 μg Al/gram). These findings provide strong evidence that expression of the rice *MATE* homolog is not associated with either Al exclusion or Al tolerance in any of the genotypes evaluated here.

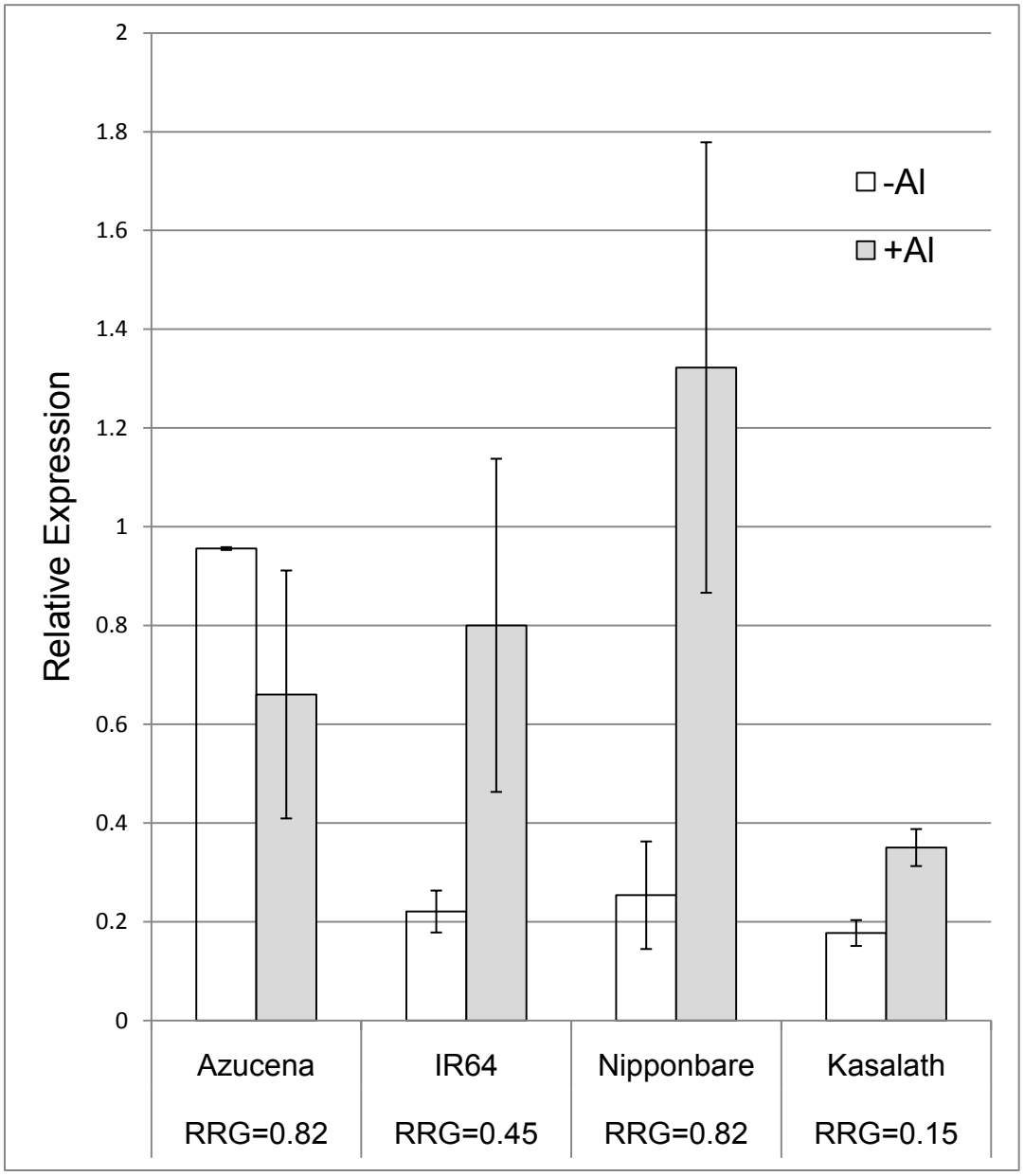


Figure 2.7. *OsMATE* gene expression in control and Al solutions. Relative gene expression determined using quantitative RT-PCR for the rice homolog (LOC_Os01g69010) of the sorghum Al tolerance gene, *SbMATE*, in root tips (1cm) of four rice genotypes that represent a wide range of Al tolerance. The Al tolerance (RRG) for each genotype is indicated below the name of each genotype.

DISCUSSION

Optimization of nutrient solution for rice Al tolerance screening

One objective of this study was to address the problems encountered by the high concentrations of Al required for rice Al tolerance studies. Because Magnavaca's nutrient solution had been used successfully to screen for Al tolerance in sorghum and maize at relatively low levels of Al^{3+}(Magnavaca, 1987), we modified this solution so that it would be appropriate for screening rice at elevated Al concentrations, while maintaining similar root growth as in Yoshida's solution under control (-Al) conditions.

Yoshida's rice solution is a complete and versatile hydroponic solution that was developed specifically for rice physiology experiments (Yoshida et al., 1976). It has been widely used to evaluate rice mineral nutrition, including toxicities to salt, iron, and aluminum, as well as deficiencies of phosphorus (Lin et al., 2004; Dufey et al., 2009; Nguyen VT et al., 2001; Shimizu et al., 2004). The long history and functionality of the Yoshida's solution make it a natural first choice for Al tolerance screening in rice. However, because of the high ionic strength and high concentrations of mineral ions that complex Al in Yoshida's solution, there are very serious problems when it is used to evaluate Al tolerance in rice. These are exemplified by the clearly visible precipitate that forms in the Yoshida's +Al treatment solution. Some studies have avoided the problem of Al precipitation by screening seedlings in calcium chloride ($CaCl_2$) solution, which does not precipitate with Al and allows for reproducible Al^{3+} concentrations (Ma et al., 2002; Xue et al., 2006; Yamaji et al., 2009). However, a simple $CaCl_2$ solution can only be used to screen very young seedlings, when the seed is still capable of providing all necessary mineral nutrients.

The Modified Magnavaca solution developed in this study can be used to screen Al tolerance in plants at all stages of development.

The precipitation issues confound the ability to quantify rice Al tolerance, as it is difficult to design a nutrient solution with reproducible levels of Al, as well as the essential elements P, S, and Fe which can also impact root growth. Al^{3+} is highly reactive and readily precipitates with other essential elements; in the Yoshida's + Al solution, both phosphorus (P) and iron (Fe) were reduced to such low levels that it was difficult to distinguish between root inhibition due to Al and that due to lack of P and Fe. Phosphorus and Fe are typically present in nutrient solutions as PO_4^- and Fe^{3+}, and it has been well documented that different concentrations of P and/or Fe can lead to alterations in root growth and architecture (Lynch and Brown, 2001; Williamson et al 2001; Lopez-Bucio et al, 2003; Ward et. al., 2008). Furthermore, the use of citrate as the Fe-chelate in the Yoshida's solution is problematic, as citrate preferentially binds Al over Fe, leading to differences in Fe availability between the control and +Al Yoshida's solutions. In the Modified Magnavaca's solution, soluble P concentrations were reduced seven-fold with respect to Yoshida's solution (from 322 μM to 45μM), which is still well within the range of sufficient P concentrations for continuously flowing hydroponic solutions (Jones, 1997). Sulfate concentrations were reduced sixteen-fold (from 3.33mM to 0.2mM), the ionic strength was significantly reduced, and a Fe-HEDTA chelate was utilized to prevent Fe precipitation and citrate interaction with Al. ICP-ES analysis confirmed that the Modified Magnavaca's solution has significantly reduced precipitation of P, Fe, and Al in the Al treatment solutions compared to the Yoshida's solution.

The differential root growth responses observed in +Al treatments between the two nutrient solutions were consistent with Geochem-EZ predictions. It is generally accepted that the primary rhizotoxic form of Al is Al^{3+}, thus when a large proportion of Al is precipitated in the Yoshida's solution, it becomes unavailable to affect root growth (Kochian, 2004a). The increased root growth inhibition in the Modified Magnavaca's Al solution can be attributed to one or a combination of three factors: 1) less of the added Al is precipitated with S and P compared to Yoshida's, leaving more Al in the active (rhizotoxic) form; 2) the citrate in Yoshida's solution added as an Fe chelate preferentially complexes with Al, whereas the Modified Magnavaca's uses an HEDTA chelate, which chelates Fe preferentially over Al; and 3) the Modified Magnavaca's solution has a lower overall ionic strength than the Yoshida's solution, which increases the activity coefficient (and hence the concentration of thermodynamically relevant ion in solution) of a trivalent ion. Also, as the nutrient solution ionic strength decreases, it prevents the roots from being protected from Al^{3+}, as the Al ions have less competition for negatively charged sites within the root cell wall and root plasma membrane by decreasing the concentrations of other cations that can shield Al^{3+} from these negative sites.

Importance of quantifying the whole root system in Al tolerance studies

Rice seedling root systems are fibrous and can have multiple primary, secondary, and tertiary roots within a few days after germination. There is also significant genetic variation in rice root architecture among varieties, ecotypes, and/or subpopulations. The phenotypic variation in root growth habit *per-se* among varieties and ecotypes must be taken into consideration when determining Al tolerance. To date, published results on Al tolerance in maize, sorghum, and rice have all used the growth of the longest root(s) as the assay for Al tolerance (Magalhaes et al., 2004; Piñeros et al,

2001; Wu et al., 2000; Nguyen VT et al., 2001, 2002; Nguyen BD et al., 2003; Xue et al., 2006, 2007). Although this approach has proven useful in assessing Al tolerance in other cereals, as demonstrated by the cloning of Al tolerance genes in wheat and sorghum, our results suggest that Al tolerance based on RRG of the longest root is not the best predictor of Al tolerance in rice.

Using a set of 225 diverse rice genotypes and the RootReader2D software, we determined that the correlation between the RRG of the longest root and the RRG of the total root system was weak (R^2=0.17) (Supplemental Figure S2.1). Furthermore, in two QTL mapping studies where Al tolerance was evaluated based on both assays, we identified some of the same, but also some novel, major effect Al tolerance QTL that were only detected by TRG-RRG (Famoso et al., personal communication). Our observations in this study are consistent with studies in maize, wheat, sorghum, soybean, sugarcane, and tobacco where all have reported severe inhibition of lateral roots in sensitive genotypes (Bushamuka and Zobel, 1998; Silva et al., 2001; Hetherington et al., 1988; Brichkova GG, 1998). We thus conclude that the RRG of the total root system is clearly a much better quantitative indicator of rice Al tolerance than RRG of the longest root, and our newly developed automated image capture and computational determination of growth of the total root system makes it feasible to use this parameter in large-scale genetic and physiological studies.

Comparison of Al tolerance between cereal species

In this study, we demonstrated that young rice seedlings (3 days old) tolerate significantly higher concentrations of Al^{3+} than maize, sorghum, or wheat, consistent with the superior Al tolerance of rice observed in previous hydroponic Al^{3+} concentrations and field studies (Foy, 1998; Wu et al., 2000; Nguyen VT et al., 2001,

2002; Nguyen BD et al., 2003; Magalhaes et al., 2004; Sasaki et al., 2004). Yet, we know little about the genes and physiological mechanisms responsible for the high levels of Al tolerance in rice. Other cereals, such as rye, have been reported to exhibit high levels of Al tolerance (Gallego and Benito, 1997). However, the Al concentrations in which rye has been screened are 4 times lower than those at which rice is screened (Gallego and Benito, 1997; Gallego et al., 1998; Li et al., 2000; Collins et al., 2008). This suggests that rice is a very useful model for characterizing the mechanisms conferring high levels of Al tolerance in cereals. Rice also has an abundance of genetic and genomic resources, including several sequenced genomes, high density of genotyping arrays, the availability of numerous immortal mapping populations, and extensive germplasm collections (www.gramene.org; www.irri.cgiar.org).

Despite the fact that sorghum has been previously demonstrated to exhibit higher Al tolerance than wheat (Sasaki et al., 2006; Caniato et al., 2007), in this study sorghum and wheat seedlings exhibited similar levels of Al tolerance after 3 days in Al solutions. A likely explanation for this discrepancy is the extended time in Al required to observe the Al tolerance response in sorghum (5-6 days) (Magalhaes et al, 2007). Thus, the degree of Al tolerance observed for sorghum in our study is less than would be predicted if the plants were grown in Al solution for up to 6 days (Magalhaes et al., 2007; Caniato et al., 2007).

Rice must employ a novel Al tolerance mechanism

Organic acid-mediated root tip Al exclusion has been reported in numerous plant species, explaining most of the phenotypic variation in wheat (Sasaki et al., 2004 and 2006), sorghum (Magalhaes et al., 2007), and *Arabidopsis* (Hoekenga et al., 2003),

and a portion of the variation in Al tolerance in maize (Piñeros et al., 2005). As a species, rice is 2-5 times more Al tolerant than wheat, sorghum, and maize, yet this study demonstrated that there is no significant correlation between Al exclusion from the root apex and root growth in Al. This indicates that the roots of tolerant rice varieties can continue to grow even with significant Al accumulation into the root tip. Thus, rice must employ unique mechanisms of Al tolerance not found in other cereal species.

Unlike *Japonica,* the more susceptible *Indica* varieties do exhibit a significant negative correlation between rates of citrate exudation and Al concentrations in the root tip (R^2=-0.47), though this response is not correlated with Al tolerance (RRG of the total root system). However, rates of root exudation of malate (R^2=0.24) and phosphate (R^2=0.13) showed a weak positive correlation with Al tolerance in *Indica* varieties, but not with Al exclusion. These findings suggest that malate and/or phosphate exudation may function at least in part to chelate Al^{3+} within the apoplast of the root tip, rather than exclude Al^{3+} from entering the root tip. The primary function of root exudates in Al tolerance is believed to be the exclusion of Al from the root apex, but this alone is not responsible for the high levels of Al tolerance in rice. The clearest evidence for this comes from experiments where the wheat Al tolerance gene (*ALMT1*) was transformed into rice, resulting in Al-induced gene expression and enhanced malate exudation, but no effect on Al tolerance (Sasaki et al., 2004). However, when the *ALMT1* gene was transformed into barley, an Al susceptible species, Al tolerance was increased by >100%. Multiple rice Al tolerance QTL studies have identified a region on chromosome one that is in close proximity to the rice MATE family member that is a homolog of the sorghum Al tolerance gene (*SbMATE*), leading to the hypothesis that this gene may be underlying these QTL.

SbMATE functions in sorghum Al tolerance as an Al-activated root citrate efflux transporter that excludes Al from the root tip, with differences in Al tolerance across sorghum genotypes directly related to gene expression (R^2=0.98) (Magalhaes et al., 2007).

Quantitative RT-PCR was conducted to determine if differences in rice MATE gene expression correlated with differences in rice Al tolerance in four genotypes with widely varying levels of Al tolerance. The highly susceptible Kasalath had the lowest MATE expression under Al stress, significantly less than Al tolerant Nipponbare. However, Nipponbare and Azucena exhibit a similar level of Al tolerance, but *MATE* expression was significantly higher in Nipponbare under Al stress. Furthermore, Azucena accumulated less Al (326 μg Al/gram) than any other variety, nearly 8 times less than that of Nipponbare (2523 μg Al/gram). Based on the lack of correlation between rice *MATE* gene expression and Al exclusion, citrate exudation, and Al tolerance, we conclude that the rice homolog of the sorghum Al tolerance gene is not involved in mediating rice Al tolerance through Al-activated gene expression and root exclusion of Al.

In this study, one Al tolerant maize line (Cateto) and one susceptible line (B73) were compared to rice in terms of Al accumulation, root exudation of organic acids and phosphate, and Al tolerance at 160μM Al^{3+}. Both maize genotypes were severely inhibited. However, our results were consistent with previously published results (Piñeros et al., 2005) reporting Al-activated citrate exudation and Al exclusion in the tolerant maize line, and increased Al accumulation in the susceptible parent. At the high Al concentrations used in this study, RRG of Cateto was severely inhibited, showing levels of RRG similar to that of the susceptible maize line, B73, though

Cateto accumulated less than half the Al of B73. This suggests that the level of Al accumulated by Cateto in 160μM Al^{3+} was above the threshold at which root growth can occur in maize, and that additional Al accumulation beyond this threshold does not further inhibit root growth. When previous studies of Al accumulation in wheat (Delhaize et al., 1993) and maize (Piñeros et al., 2005) are compared with results in this study, it appears that significant Al inhibition of root growth occurs at root tip Al concentrations around 1000 μg Al/g root tip in wheat and maize. Delhaize et al. (1993) quantified Al inhibition of root growth over time in one Al sensitive and one tolerant wheat variety and a significant difference in Al tolerance was not observed until the susceptible variety accumulated over 1000 μg Al/g in the root tip. Similarly, Piñeros et al. (2005) reported a non-linear relationship between root tip Al accumulation and Al tolerance in 6 maize genotypes, two tolerant and four susceptible. The two tolerant genotypes, Cateto and Pioneer 3355, accumulated significantly different amounts of Al (495 and 900 μg Al/g, respectively) and had Al tolerance values of 0.97 and 0.75 RRG, respectively. The four susceptible lines showed values ranging from 1,250 to 2,225 μg Al/g in root tip Al accumulation, yet Al tolerance only ranged from 0.48 to 0.38. When comparing tolerant and susceptible maize lines, the tolerant line Pioneer 3355 that accumulated 900 μg Al g was over 50% more Al tolerant than the susceptible line that accumulated 1,250 μg Al/g. These results suggest that the relation between Al accumulation and Al tolerance in maize is not linear and, similar to wheat, a threshold is reached at around root tip Al concentrations of 1000 μg Al/g, where growth is significantly inhibited.

Based on these observations, it appears that rice, as a species, is capable of withstanding significantly higher Al concentrations both in the soil solution and the root tip than other cereals. All but one rice line was more tolerant than the most

tolerant maize line (Cateto), yet all but one rice genotype accumulated more Al in the root apex. Some rice genotypes that accumulated 50-100% more Al in their root tips were 2-5 times more Al tolerant than other rice genotypes that accumulated less root-tip Al (Supplemental Table S2). Based on the lack of correlation between rice Al exclusion and Al tolerance, and the relatively high levels of Al accumulation in rice compared to maize, we conclude that rice utilizes one or more novel Al tolerance mechanisms. At this time we have little information regarding the nature of this new Al tolerance mechanism. Because the majority of the Al in the root tip resides in the apoplast (Kochian, 1999), it is logical to speculate that the root cell wall may play a role in the high level of Al tolerance observed in rice. Recent work from Jian Feng Ma's lab supports this speculation based on the map-based cloning of an Al sensitive knock-out mutant locus in rice (Huang et al, 2009). This resulted in the identification of two mutant genes, *STAR1* and *STAR2*, which encode two interacting proteins that form an ABC transporter complex. Transport studies via the STAR1/STAR2 transporter complex in oocytes showed that the transporter mediates the efflux of UDP-glucose, presumably into the root apoplast, leading the authors to speculate that cell wall modification may play a role in rice Al tolerance. Furthermore, a study conducted by Yang et al. (2008) provided evidence that cell wall polysaccharides may be involved in rice Al tolerance.

It is known that the growing root tip is the site of Al toxicity (Ryan et al, 1993), however the mechanism by which Al inhibits root growth in plants is still unclear. Based on observations that Al tolerance in wheat, sorghum, and maize is related to the plants' ability to exclude Al from the growing tip, but not from the mature root regions, researchers have inferred that Al in these species poisons proteins and/or structural components of the root tip that are critical to cell growth, elongation, and/or

division. In rice, where there is no significant correlation between Al accumulation in the root tip and Al tolerance, it appears that the mechanism of toxicity must be categorically different than in other species. We hypothesize that at some point in evolution, the lineage leading to modern species of *Oryza* experienced a dramatic shift in its position within the landscape of plant response to Al, demonstrating greatly enhanced ability to grow under high concentrations of Al. If this hypothesis is true, identifying the genes/alleles underlying Al tolerance or susceptibility among rice varieties will provide limited insight into novel plant Al tolerance mechanisms. To fully understand the novelty of the mechanism(s) of Al toxicity and tolerance found in rice, it will be necessary to undertake very specific physiological, biochemical, and molecular experiments in a phylogenetic context. Thus, rice appears to hold the key to understanding how and when, over the course of evolution, a lineage of plants experienced a dramatic genetic change that led to enhanced levels of Al tolerance, and will provide critical insights that are likely to help move this capability into other species that are critical to human survival.

MATERIALS AND METHODS

Plant material

A set of 7 to 9 genotypes each of rice, maize, wheat, and sorghum were used to compare Al tolerance between species. Rice seeds were obtained from S. McCouch and included the genotypes Azucena, BJ 1, China 1039, Cybonnet, IR64, Kasalath, Nipponbare, and Sabjaraj. Maize seeds were obtained from E. Buckler and included the genotypes B164, B57, Cateto, H84, NC264, NC290A, NC310, NC328, and R10. Wheat seeds were obtained from M. Sorrells and included the genotypes AC Reed, Atlas 66, Bob White, Caledonia, Cham 1, Opata, Roane, and Scout 66. Sorghum seeds were obtained from S. Kresovich and J. Magalhaes and included the genotypes BR007, BTX623, Cowley, IS3620C, SC452, SC566, and T309.

Plant growth conditions

Seeds were germinated in rolled germination paper at 26-30 C for 3-5 days under dark conditions. Wheat and sorghum seeds were surface sterilized with 10 % bleach and rice with 20% bleach for 15-20 minutes. Maize seeds were treated with a fungicide treatment of Captan400, Trilex, and Allegiance. Upon germination, seedlings were transferred to control (-Al) solutions for 24 hours, then 20 uniform seedlings were photographed and root length was quantified using RootReader 2D. Subsequently, 10 seedlings were transferred to fresh control solution and 10 seedlings to Al treatment solution. After three days in the respective treatments, roots were photographed and measured and mean root growth in control and +Al treatment was calculated for each genotype and RRG was determined: RRG=treatment root growth / control root growth. Plants were grown in 9 L tubs with 48 plants per tub, and the plants were supported with 8 foam strips (6 plants/strip) with a slit cut into the foam to anchor the stem. Aeration was provided in all experiments, except for experiments comparing

the nutrient solutions in which only rice lines were compared. Plant growth chamber conditions for the maize diversity screen and specie comparison experiments were 26^{o} C (day)/23^{o} C (night), while the rice diversity screening conditions were 30^{o} C (day)/26^{o} C (night). All experiments were conducted with 12 hour days with a light intensity of 450 mmol photons m^{-2} s^{-1}.

Nutrient solutions

The control (-Al) Yoshida's nutrient solution was prepared as described previously (Yoshida, 1978) and the pH was adjusted to 4.0 with 1N NaOH. The Yoshida treatment (+Al) solution was identical to the control, but contained 35ppm (1297µM) $AlCl_3$. The control (-Al) Modified Magnavaca's nutrient solution was modified from Magnavaca et al (1987). The treatment (+Al) Modified Magnavaca's solution contained 540µM $AlCl_3$, added after pH adjustment to 7.8 with KOH to prevent Al precipitation, the final pH was adjusted to 4.0 with 1N HCl.

ICP-ES analysis of nutrient solutions

ICP-ES elemental profiling was conducted on all elements, except N, in both nutrient solutions. To determine the available concentration of each element, one liter of each nutrient solution/treatment was made and analyzed to confirm elemental composition. Four 50ml samples of each nutrient solution/treatment were then collected and stored in the dark for three days under plant growth conditions to permit chemical equilibrium. Samples were then homogenized and a 10ml sample was collected for ICP analysis to calculate total elemental concentration after three days. The remaining 40ml was centrifuged at 3250 x g for 15 minutes and 10ml of supernatant was collected for ICP analysis to determine the amount of each element that was precipitated out of solution.

Chemical speciation analysis

Chemical speciation analysis was conducted according to Shaff et al. (2009) using the nutrient composition of Yoshida's nutrient solution (1978) and the Modified Magnavaca's solution presented here. The pH of all solutions was fixed at 4.0 and the predicted available activity of each element was determined through the primary distribution and case progress table output by the prediction of solid formation for each element.

Root imaging and measurements

A custom root imaging system was used to accurately quantify total root length parameters from rice. For description of this system in detail, see Clark et al. www.plantmineralnutrition.net. The system utilizes digital photography and Java based RootReader2D software (available at http://www.plantmineralnutrition.net/rootreader.htm). The imaging system consists of a Nikon D200 digital SLR camera with a 60mm Macro lens which was calibrated and aligned to a fixed focal plane scale of 120 pixels/cm. Each plant was imaged with its root system spread out in a clear, solution-filled trough that was illuminated from below with a light box. Once the plants were photographed, the root images were converted from the RAW NEF file format to a 32-bit RGB TIFF file format using Nikon Capture NX software. The TIFF images were then converted from the RGB format to an 8-bit grayscale format using Adobe Photoshop. The grayscale images were then batch processed with RootReader2D software with a fixed threshold adjusted between 15-25 to maximize contrast and an error criterion of 6.0 pixels to optimize measurement accuracy. Total root system lengths were automatically measured while individual roots were semi-automatically selected and measured in RootReader2D software.

Root tip Al content

Plants were grown as described above and after three days of Al treatment, the first centimeter of the primary roots were collected (~3/rice plant) and bulked to 50 roots/replicate, with five replicates/genotype. Root tips were then dried in an oven at 60ºC for two days. Dry weights were determined using a microgram balance (MT2, Mettler, Greifensee, Switzerland). Dry samples were digested with 100µL of 50/50 ddH(NO_3) and 70% percholoric acid, resuspended in 10.25mL 5% ddH(NO_3), and analyzed using an inductively coupled argon plasma model 51000 emission spectrometer (Perkin-Elmer/Sciex, Foster City, CA).

Root organic acid exudation

Seeds were germinated as described above, then 10 uniform seedlings were transferred to a plastic tube stopper with 2mm holes drilled throughout the bottom, and placed in 45mL of 2.7mM $CaCl_2$ with and without 160µM Al^{3+} treatment in a 50mL Falcon tube (n=3). After 48 hours growth control plants were transferred to fresh $CaCl_2$ solution and Al treated plants were transferred to a $CaCl_2$ solution containing 39µM Al^{3+} for 24 hours and these solutions were processed and analyzed by capillary electrophoresis as described by Piñeros et al., (2005). The $CaCl_2$ concentration was determined to best replicate the Modified Magnavaca's ionic strength. It was necessary to grow the plants in $CaCl_2$ and reduce the Al concentration in the sample collected to prevent noise and interference during capillary electrophoresis.

Quantitative RT-PCR

Plants were grown as described above and after three days of Al or control treatment, the first centimeter of the primary roots were collected (~3/rice plant) and bulked to 50

roots/replicate, with two replicates/genotype and flash frozen in liquid nitrogen and stored at -80C. RNA extraction, cDNA synthesis, and RT-PCR were conducted as described in Liu et al., (2009). Primer sequences for the MATE gene and the actin internal control were as follows: MATE 5' AGGAGATTCGTGCCGTCCGTGA3' 5'CGGCTTGACGCCCATGATGC3' ; Actin 5'ATCCTTGTATGCTAGCGGTCGA 3' 5'ATCCAACCGGAGGATAGCATG 3'.

Statistical analysis

Analysis of variance was conducted for comparison of Al tolerance between species and correlation coefficients were determined using JMP V7.0 (SAS institute Inc.). Microsoft excel 2007 was used to conduct regression analysis comparing longest root RRG and total root RRG. Mean Al tolerance of the accessions was used to estimate species Al tolerance.

Acknowledgements

We would like to thank following undergraduate interns for their assistance in various aspects of phenotyping and/or greenhouse maintenance of plants: Crayton Montei, Laura Pursel, Vivian Li, Misty Carlisle, Sarah Villareal, Mandy Kain, and Melissa Major. We are grateful to Ed Buckler, Mark Sorrells, David Benscher, Steve Kresovich, and Jurandir Magalhaes for supplying maize, wheat, and sorghum seeds. We appreciate the proofreading and editing contribution of Elliot Heffner. We would also like to thank Jennifer Thaler for use of her microbalance and Matthew Milner and Jiping Liu for assistance and advice on RT-PCR experiments.

APPENDIX

Supplemental Table 2.1. **Mean, standard deviation, and range of Al tolerance of rice, maize, sorghum, and wheat.** Mean, standard deviation, and range of Al tolerance for rice, maize, sorghum, and wheat. Numbers in parenthesis indicates number of genotypes screened.

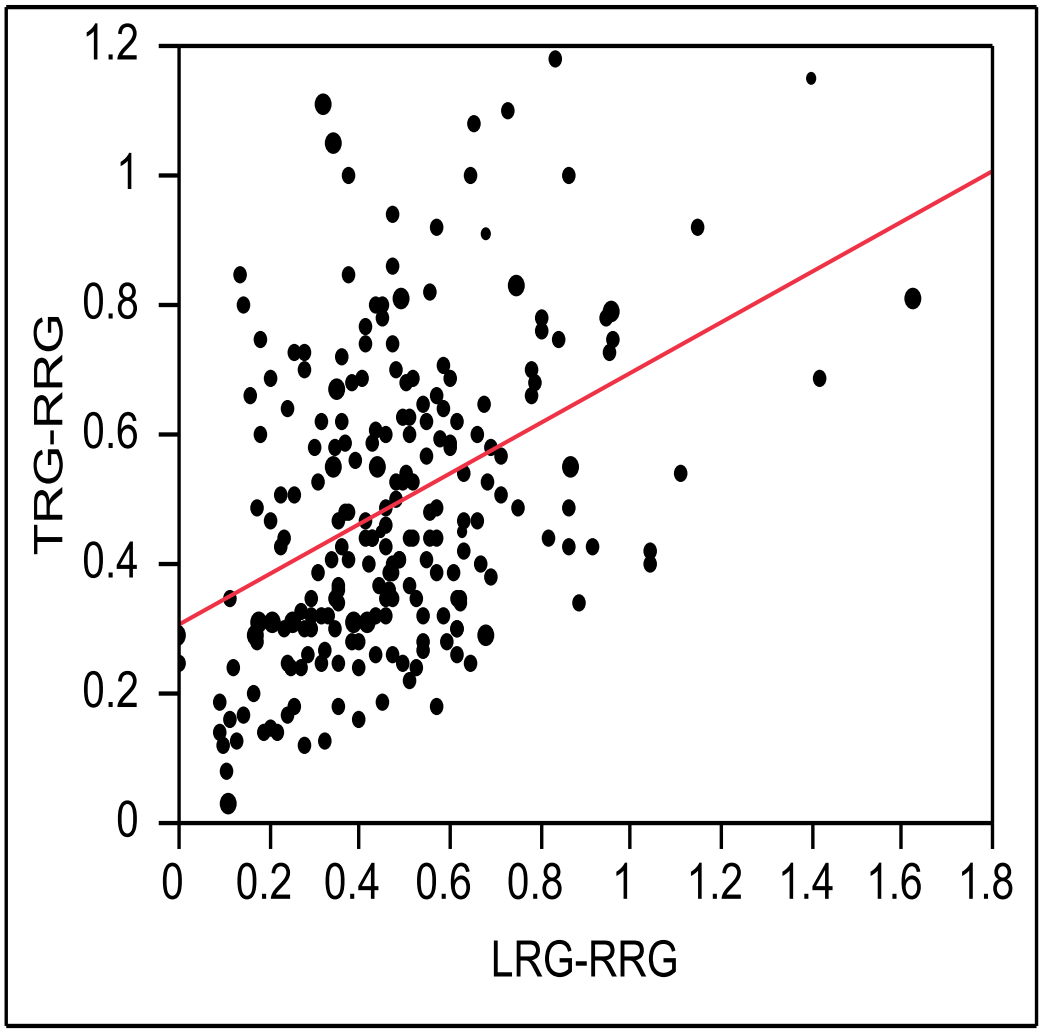
Species		**8.75 µM Al^{3+}**			**27 µM Al^{3+}**			**160 µM Al^{3+}**		
	n	Mean	SD	Range	Mean	SD	Range	Mean	SD	Range
Rice	**8**	0.91	0.18	0.70-1.16	0.84	0.20	0.6-1.13	0.63	0.20	0.25–0.95
Maize	**9**	0.77	0.18	0.48-1.0	0.32	0.18	0.1-0.51	0.10	0.04	0.05–0.17
Sorghum	**7**	0.43	0.30	0.14-0.94	0.09	0.13	0-0.39	0.01	0.004	0.0–0.02
Wheat	**8**	0.39	0.27	0-0.79	0.19	0.19	0-0.58	0.05	0.02	0.0–0.08

Supplemental Table 2.2. Summary of Al tolerance (RRG), root tip Al content (Al exclusion), and rates of root exudation of citrate (Cit), malate (Mal), and phosphate (PO_4^{3-}) for members of the rice Al tolerance diversity panel.

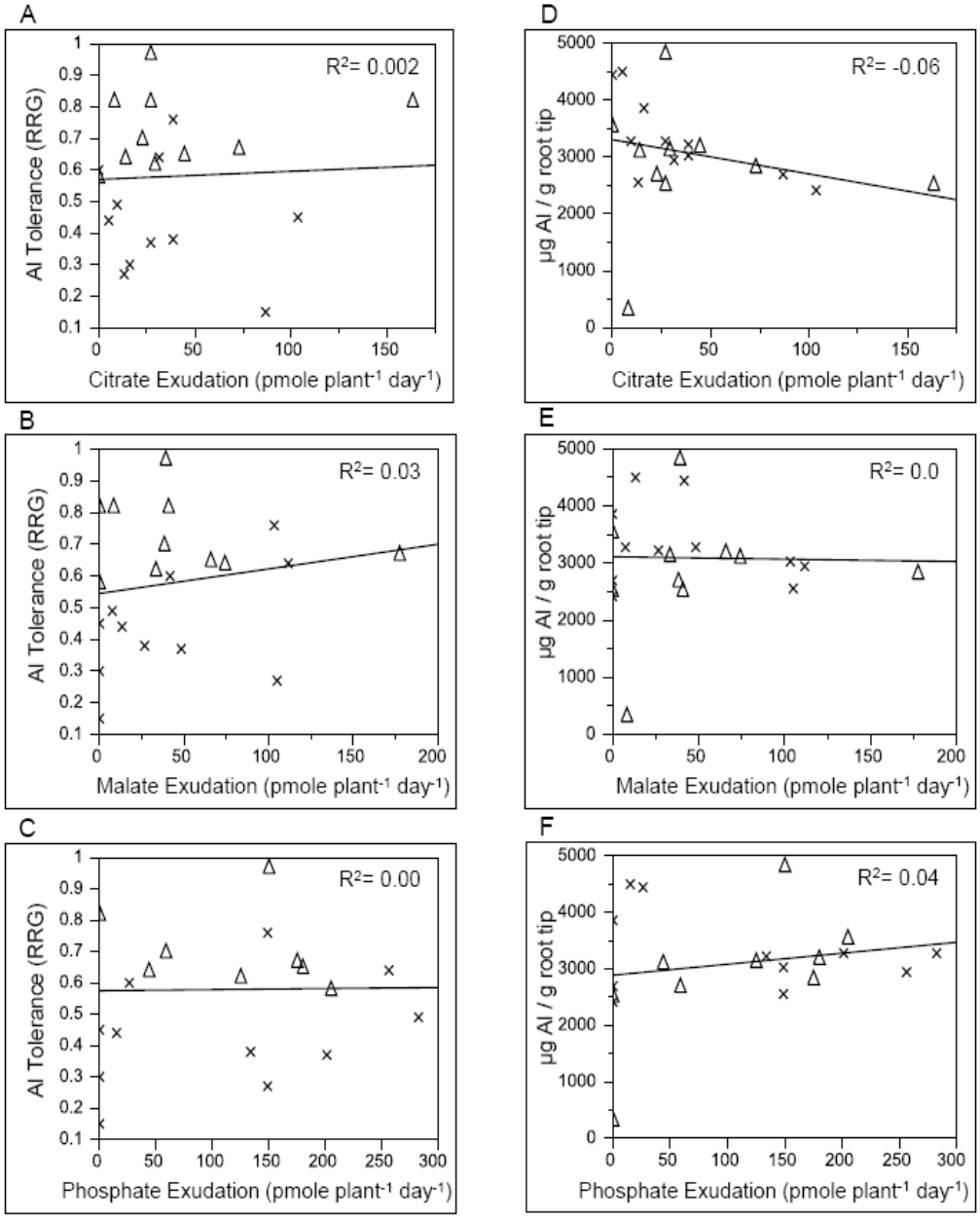
Genotype	**Rice Varietal Group/ species**	**Al Tol.** (RRG) 160µM Al^{3+}	**Root apex Al conc.**	**Root Exudation (pmole plant^{-1} day^{-1})**					
			µg Al / g root tip (+/- SE)	**Cit (- Al)** (+/- SE)	**Cit (+ Al)** (+/- SE)	**Mal (- Al)** (+/- SE)	**Mal (+ Al)** (+/- SE)	**PO_4^{3-} (- Al)** (+/- SE)	**PO_4^{3-} (+ Al)** (+/- SE)
9311	Indica	0.26	2361 (98)	x	x	x	x	x	x
Aijiaonante	Indica	0.30	3863 (70)	0	16 (16)	40 (34)	0	197 (50)	0
BR24	Indica	0.44	4490 (161)	0	5 (5)	0	13 (13)	216 (47)	15 (9)
CHANG CH'SANG HSU TAO	Indica	0.49	3280 (82)	56 (30)	10 (10)	0	7 (7)	218 (50)	282 (81)
CHIEM CHANH	Indica	0.76	3037 (76)	73(4)	38 (13)	160 (132)	103 (17)	x	149 (12)
ECIA76-S89-1	Indica	0.38	3219 (198)	0	39 (20)	8 (8)	27 (15)	199 (98)	134 (65)
IR64	Indica	0.45	2423 (92)	0	104 (20)	4 (5)	0	0	0
Kasalath	Indica	0.15	2695 (58)	0	87 (15)	10 (5)	0	0	0
Sabharaj	Indica	0.64	2934 (81)	174 (132)	32 (18)	9 (7)	112 (41)	x	256 (75)
SHAI-KUH	Indica	0.97	2804 (121)	x	x	x	x	x	x
SLO 17	Indica	0.60	4455 (149)	101 (19)	0	0	42 (25)	342 (27)	26 (21)
SML 242	Indica	0.37	3291 (142)	82 (26)	27 (14)	26 (23)	48 (5)	266 (36)	201 (69)
TOG 7178	Indica	0.27	2550 (53)	18 (18)	13 (8)	193 (193)	105 (34)	387 (74)	149 (17)
Azucena	Jap.	0.82	326 (14)	0	8 (8)	0	9 (5)	0	0
CHAMPA TONG 54	Jap.	0.34	2878 (46)	x	x	x	x	x	x

Supplemental Table 2.2 (continued)

Genotype	Rice Varietal Group/ species	Al Tol. (RRG) 160µM Al^{3+}	Root apex Al conc.	Root Exudation (pmole $plant^{-1}$ day^{-1})					
			µg Al / g root tip (+/- SE)	Cit (- Al) (+/- SE)	Cit (+ Al) (+/- SE)	Mal (- Al) (+/- SE)	Mal (+ Al) (+/- SE)	PO_4^{3-} (- Al) (+/- SE)	PO_4^{3-} (+ Al) (+/- SE)
CUBA 65	Jap.	0.67	2846 (192)	45 (17)	73 (31)	57 (34)	178 (106)	396 (66)	176 (64)
DA 5	Jap.	0.64	3103 (241)	23 (23)	14 (8)	4 (4)	74 (63)	378 (60)	44 (38)
IGUAPE CATETO	Jap.	0.62	3137 (102)	39 (20)	30 (9)	0	33 (17)	284 (5)	125 (10)
MIRITI	Jap.	0.97	4846 (184)	53 (8)	27 (14)	0	39 (10)	112 (12)	150 (8)
Nipponbare	Jap.	0.82	2523 (134)	1 (1)	164 (86)	3 (5)	0	0	0
SINAGUING	Jap.	0.82	2519 (107)	71 (56)	27 (4)	15 (15)	41 (9)	131 (38)	0
Tox782-20-1	Jap.	0.58	3569 (230)	25 (25)	0	0	0	314 (111)	205 (53)
WC 4419	Jap.	0.70	2694 (158)	61 (12)	23 (13)	13 (13)	39 (17)	306 (52)	59 (30)
WELLS	Jap.	0.65	3182 (55)	0	44 (2)	0	66 (49)	173 (21)	180 (29)



Supplemental Figure 2.1. Correlation of Al tolerance (RRG) based longest root growth (LRG) and total root growth (TRG). Analysis of 225 genetically diverse rice accessions demonstrates that the Al tolerance values obtained from using LRG-RRG is only weakly correlated to TRG-RRG (R^2=0.17). LRG-RRG is not an accurate predictor of Al tolerance of the entire root system.



Supplemental Figure 2.2. Correlation of root exudates and Al accumulation and Al tolerance in *Japonica* and *Indica*. Correlation of root exudation of: A, D) citrate, B, E) malate, C, F) phosphate with Al tolerance (RRG) in the left column and root tip Al content in the right column for 23 genetically diverse rice genotypes (*Indica* varieties=x, *Japonica* varieties=Δ). There were no significant correlations.

REFERENCES

Brichkova GG, Shishlova AM, Maneshina TV, Kartel' NA (2007) The tolerance of tobacco genetically modified plants to aluminium. Tsitol Genet 41: 23-28

Bushamuka V, Zobel R (1998) Tap, Basal, and Lateral Root Responses to a Stratified Acid, Aluminum-Toxic Soil. Crop Science 38: 416-421

Caicedo AL, Williamson SH, Hernandez RD, Boyko A, Fledel-Alon A, York TL, Polato NR, Olsen KM, Nielsen R, McCouch SR, Bustamante CD, Purugganan MD (2007) Genome-wide patterns of nucleotide polymorphism in domesticated rice. PLoS Genet **3:** 1745-1756

Caniato FF, Guimaraes CT, Schaffert RE, Alves VM, Kochian LV, Borem A, Klein PE, Magalhaes JV (2007) Genetic diversity for aluminum tolerance in sorghum. Theor. Appl. Genet. 114: 863-876

Champoux MC, Wang G, Sarkarung S, Mackill DJ, O'Toole JC, Huang N, McCouch S (1995) Locating genes associated with root morphology and drought avoidance in rice via linkage to molecular markers. Theor. Appl. Genet.90: 969-981

Collins NC, Shirley NJ, Saeed M, Pallotta M, Gustafson JP (2008) An *ALMT1* gene cluster controlling aluminum tolerance at the *Alt4* locus of rye (*Secale cereal* L.). Plant Physiol. 179: 669-682

Dally AM, Second G (1990) Chloroplast DNA diversity in wild and cultivated species of rice genus Oryza section Oryza cladisticmutation and genetic-distance analysis. Theor. Appl. Genet.209-222

Degenhardt J, Larsen PB, Howell SH, Kochian LV (1998) Aluminum resistance in the *Arabidopsis* mutant alr-104 is caused by an aluminum-induced increase in rhizosphere pH. Plant Physiol. 117: 19-27

Delhaize E, Craig S, Beaton CD, Bennet RJ, Jagadish VC, Randall PJ (1993) Aluminum Tolerance in Wheat (Triticum aestivum L.) (I. Uptake and Distribution of Aluminum in Root Apices). Plant Physiol. 103: 685-693

Delhaize E, Ryan PR (1995) Aluminum Toxicity and Tolerance in Plants. Plant Physiol. 107: 315-321

Devlin B, Roeder K (1999) Genomic control for association studies. Biometrics 997-1004

Dufey I, Hakizimana P, Draye X, Lutts S, Bertin P (2009) QTL mapping for biomass and physiological parameters linked to resistance mechanism to ferrous iron toxicity in rice. Euphytica 167: 143-160

Foy CD (1988) Plant adaptation to acid, aluminum-toxic soils. Communications in soil and plant analysis 19: 959-987

Gallego FJ, Benito C (1997) Genetic control of aluminum tolerance in rye (Secale cereale L.). Theor. Appl. Genet. 393-399

Gallego FJ, Calles B, Benito C (1998) Molecular markers linked to the aluminum tolerance gene alt1 in rye (Secale cereale L.). Theor. Appl. Genet. 97: 1104-1109

Garris AJ, McCouch SR, Kresovich S (2003) Population structure and its effect on haplotype diversity and linkage disequilibrium surrounding the xa5 locus of rice (Oryza sativa L.). Genetics 165: 759-769

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

Hetherington SJ, Asher CJ, Blamey FPC (1988) Comparative tolerance of sugarcane, navybean, soybean and maize to aluminum toxicity. Austrailian Journal of Agricultural Research 39: 171-176

Hoekenga OA, Vision TJ, Shaff JE, Monforte AJ, Lee GP, Howell SH, Kochian LV (2003) Identification and characterization of aluminum tolerance loci in *Arabidopsis* (Landsberg erecta x Columbia) by quantitative trait locus mapping. A physiologically simple but genetically complex trait. Plant Physiol. 132: 936-948

Hu H, Mu J, Zhang H, Tao Y, Han B (2006) Differentiation of a miniature inverted transposable element (MITE) system in Asian rice cultivars and its inference for a
diphyletic origin of two sub-species of Asian cultivated rice. J. Integ. Plant Biol. 260-267

Huang CF, Yamaji N, Mitani N, Yano M, Nagamura Y, Ma JF (2009) A bacterial-type ABC transporter is involved in aluminum tolerance in rice. Plant Cell 21: 655-667

Jones JB (1997) Hydroponics: A Practical Guide for the Soilless Grower., Ed 2. CRC Press, Boca Roton, Fl

Kochian LV, Piñeros MA, Hoekenga OA (2004b) The physiology, genetics and molecular biology of plant aluminum tolerance and toxicity. In H Lambers, ed, Root Physiology - From Gene to Function. Kluwer, Dordrecht

Kochian LV (1995) Cellular Mechanisms of Aluminum Toxicity and Resistance in Plants. Annu. Rev. Plant Physiol. Plant Mol Biol 46: 237-260

Kochian LV, Hoekenga OA, Piñeros MA (2004a) How do crop plants tolerate acid soils? Mechanisms of aluminum tolerance and phosphorous efficiency. Annu. Rev. Plant Biol. 55: 459-493

Li XF, Ma JF, Matsumoto H (2000) Pattern of aluminum-induced secretion of organic acids differs between rye and wheat. Plant Physiol. 123: 1537-1544

Lilley JM, Ludlow MM, McCouch S, O'Toole JC (1996) Locating QTL for osmotic adjustment and dehydration tolerance in rice. Journal of Experimental Botany 47: 1427-1436

Lin HX, Zhu MZ, Yano M, Gao JP, Liang ZW, Su WA, Hu XH, Ren ZH, Chao D (2004) QTLs for Na^+ and K^+ uptake of the shoots and roots controlling rice salt tolerance. Theor. Appl. Genet. 108: 143-160

Londo JP, Chiang YC, Hung KH, Chiang TY, Schaal BA (2006) Phylogeography of Asian wild rice, Oryza rufipogon, reveals multiple independent domestications of cultivated rice, Oryza sativa. Proc. Natl. Acad. Sci. U S A 103: 9578-9583

Lopez-Bucio J, Cruz-Ramirez A, Herrera-Estrella L (2003) The role of nutrient availability in regulating root architecture. Curr. Opin. Plant Biol. 6: 280-287

Lynch JP, Brown KM (2001) Topsoil foraging – an architectural adaptation of plants to low phosphorus availability. Plant and Soil 237: 225-237

Ma JF, Ryan PR, Delhaize E (2001) Aluminium tolerance in plants and the complexing role of organic acids. Trends Plant Sci. 6: 273-278

Ma JF, Shen R, Zhao Z, Wissuwa M, Takeuchi Y, Ebitani T, Yano M (2002) Response of rice to Al stress and identification of quantitative trait loci for Al tolerance. Plant Cell Physiol. 43: 652-659

Magalhaes JV, Garvin DF, Wang Y, Sorrells ME, Klein PE, Schaffert RE, Li L, Kochian LV (2004) Comparative mapping of a major aluminum tolerance gene in sorghum and other species in the poaceae. Genetics 167: 1905-1914

Magalhaes JV, Liu J, Guimaraes CT, Lana UG, Alves VM, Wang YH, Schaffert RE, Hoekenga OA, Piñeros MA, Shaff JE, Klein PE, Carneiro NP, Coelho CM, Trick HN, Kochian LV (2007) A gene in the multidrug and toxic compound extrusion (MATE) family confers aluminum tolerance in sorghum. Nat. Genet. 39: 1156-1161

Magnavaca R, Gardner CO, Clark RB (1987) Evaluation of inbred maize lines for aluminum tolerance in nutrient solution. Genetic Aspects of Plant Mineral Nutrition. Martinus Nijhoff, Dordrecht, Netherlands, pp 255-265

Maron L, Piñeros MA, Guimaraes CT, Magalhaes JV, Pleiman JK, Mao C, Shaff J, Belicuas SNJ, Kochian LV (2010) Two functionally distinct members of the MATE (multi-drug and toxic compound extrusion) family of transporters potentially underlie two major aluminum tolerance QTLs in maize. The Plant Journal 61: 728-740

Miyasaka SC, Buta JG, Howell RK (1991) Mechanism of Aluminum Tolerance in Snapbeans: Root Exudation of Citric Acid. Plant Physiol. 96: 737-743

Nguyen VT, Burow MD, Nguyen HT, Le BT, Le TD, Paterson AH (2001) Molecular mapping of genes conferring aluminum tolerance in rice (*Oryza sativa I.*). Theor. Appl. Genet. 102: 1002-1010

Nguyen BD, Brar DS, Bui BC, Nguyen TV, Pham LN, Nguyen HT (2003) Identification and mapping of the QTL for aluminum tolerance introgressed from the new source, Oryza Rufipogon Griff., into indica rice (Oryza sativa L.). Theor. Appl. Genet. 106: 583-593

Nguyen VT, Nguyen BD, Sarkarung S, Martinez C, Paterson AH, Nguyen HT (2002) Mapping of genes controlling aluminum tolerance in rice: comparison of different genetic backgrounds. Mol. Genet. Genomics 267: 772-780

Parker DR, Kinraide TB, Zelazny LW (1988) Aluminum speciation and phytotoxicity in dilute hydroxy-aluminum solutions. Soil Science Society of America 52: 438-444

Pellet DM, Papernik LA, Kochian LV (1996) Multiple Aluminum-Resistance Mechanisms in Wheat (Roles of Root Apical Phosphate and Malate Exudation). Plant Physiol. 112: 591-597

Piñeros MA, Kochian LV (2001) A patch-clamp study on the physiology of aluminum toxicity and aluminum tolerance in maize. Identification and characterization of Al(3+)-induced anion channels. Plant Physiol. 125: 292-305

Piñeros MA, Shaff JE, Manslank HS, Alves VM, Kochian LV (2005) Aluminum resistance in maize cannot be solely explained by root organic acid exudation. A comparative physiological study. Plant Physiol. 137: 231-241

Pritchard JK, Donnelly P (2001) Case-control for association in structured or admixed populations. Theoretical population biology 227-237

Ryan PR, DiTomaso JM, Kochian LV (1993) Aluminum toxicity in roots: an investigation of spatial sensitivity and the role of the root cap. Journal of Experimental Botany 44: 437-446

Ryan P, Delhaize E, Jones D (2001) Function and Mechanism of Organic Anion Exudation from Plant Roots. Annu Rev Plant Physiol Plant Mol. Biol. 52: 527-560

Sasaki T, Ryan PR, Delhaize E, Hebb DM, Ogihara Y, Kawaura K, Noda K, Kojima T, Toyoda A, Matsumoto H, Yamamoto Y (2006) Sequence upstream of the

wheat (Triticum aestivum L.) ALMT1 gene and its relationship to aluminum resistance. Plant Cell Physiol. 47: 1343-1354

Sasaki T, Yamamoto Y, Ezaki B, Katsuhara M, Ahn SJ, Ryan PR, Delhaize E, Matsumoto H (2004) A wheat gene encoding an aluminum-activated malate transporter. Plant J. 37: 645-653

Shaff JE, Schultz BA, Craft EJ, Clark RT, Kochian LV (2009) GEOCHEM-EZ: a chemical speciation program with greater power and flexibility. Plant and Soil DOI 10.1007/s11104-009-0193-9:

Shimizu A, Yanagihara S, Kawasaki S, Ikehashi H (2004) Phosphorus deficiency-induced root elongation and its QTL in rice (Oryza sativa L.). Theor. Appl. Genet. 109: 1361-1368

Silva IR, Smyth TJ, Raper CD, Carter TE, Rufty TW (2001) Differential aluminum tolerance in soybean: An evaluation of the role of organic acids. Physiologia Plantarum 112: 200-210

von Uexkull HR, Mutert E (1995) Global extent, development and economic impact of acid soils. In RA Date, NJ Grundon, GE Raymet, ME Probert, eds, Plant-Soil Interactions at Low pH: Principles and Management. Kluwer Academic, Dordrecht, Netherlands, pp 5-19

Ward JT, Lahner B, Yakubova E, Salt DE, Raghothama KG (2008) The effect of iron on the primary root elongation of *Arabidopsis* during phosphate deficiency. Plant Physiol. 147: 1181-1191

Williamson LC, Ribrioux SP, Fitter AH, Leyser HM (2001) Phosphate availability regulates root system architecture in *Arabidopsis*. Plant Physiol. 126: 875-882

Wolt J (1994) Soil Solution Chemistry: Applications to Environmental Science and Agriculture. John Wiley & Sons INC, New York, New York

Wu P, Liao CY, Hu B, Yi KK, Jin WZ, Ni J, He C (2000) QTLs and epistatis for aluminum tolerance in rice (*Oryza sativa* L.) at different seedling stages. Theor. Appl. Genet. 100: 1303

Xu Y, This D, Pausch RC, Vonhof WM, Coburn JR, Comstock JP, McCouch SR (2009) Leaf-level water use efficiency determined by carbon isotope discrimination in rice seedlings: genetic variation associated with population structure and QTL mapping. Theor. Appl. Genet. 18: 1065-1081

Xue Y, Wan J, Jiang L, Wang C, Liu L, Zhang YM, Zhai H (2006) Identification of quantitative trait loci associated with aluminum tolerance in rice. Euphytica 150: 37-45

Xue Y, Jiang L, Su N, Wang JK, Deng P, Ma JF, Zhai HQ, Wan JM (2007) The genetic basic and fine-mapping of a stable quantitative-trait loci for aluminium tolerance in rice. Planta 227: 255-262

Yamaji N, Huang CF, Nagao S, Yano M, Sato Y, Nagamura Y, Ma JF (2009) A Zinc Finger Transcription Factor ART1 Regulates Multiple Genes Implicated in Aluminum Tolerance in Rice. Plant Cell

Yang JL, Li YY, Zhang YJ, Zhang SS, Wu YR, Wu P, Zheng SJ (2008) Cell wall polysaccharides are specifically involved in the exclusion of aluminum from the rice root apex. Plant Physiol. 146: 602-611

Yoshida S, Forno DA, Cock JA, Gomez KA (1976) Laboratory manual for plant physiological studies of rice, Ed 3. International Rice Research Institute, Manila, Philippines

Yu J, Pressoir G, Briggs WH, Vroh Bi I, Yamasaki M, Doebley JF, McMullen MD, Gaut BS, Nielsen DM, Holland JB, Kresovich S, Buckler ES (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat Genet 38: 203-208

Zhang WH, Ryan PR, Tyerman SD (2001) Malate-permeable channels and cation channels activated by aluminum in the apical cells of wheat roots. Plant Physiol. 125: 1459-1472

Zhao K, Aranzana MJ, Kim S, Lister C, Shindo C, Tang C, Toomajian C, Zheng H, Dean C, Marjoram P, Nordborg M (2007) An *Arabidopsis* example of association mapping in structured samples. PLoS Genet 3: e4

CHAPTER 3:
GENETIC ARCHITECTURE OF ALUMINUM TOLERANCE IN RICE (*O. sativa)* DETERMINED THROUGH GENOME-WIDE ASSOCIATION ANALYSIS AND QTL MAPPING

ABSTRACT

Aluminum (Al) toxicity is a primary limitation to crop productivity on acid soils and rice has been demonstrated to be significantly more Al tolerant than other cereal crops. However, the mechanisms by which rice tolerates Al are largely unknown and no genes underlying natural variation have been identified. To investigate the genetic architecture of Al tolerance in rice we screened a diversity panel of 385 rice accessions as well as two bi-parental populations and conducted genome-wide association (GWA) analysis and QTL mapping. We determined that the *Japonica* varietal group is twice as Al tolerant as the *Indica* varietal group and identified loci that confer tolerance across and within these varietal groups. Among these loci were two that colocalized with previously identified Al sensitive mutant loci that were previously thought to only be involved in basal tolerance, as well as a 139Kb region identified by GWA and QTL analysis that contains an Nramp metal transporter gene previously demonstrated to have significantly reduced expression in response to Al in an Al sensitive mutant. Finally, five Indica outliers were identified that exhibited Al tolerance that is greater than the *Japonica* mean, and introgression analysis identified specific loci with *Japonica* ancestry introgressions that co-localize with GWA and/or QTL regions. The significant differences in Al tolerance between varietal groups and the presence of alleles with the ability to enhance Al tolerance across groups suggests that the *Indica* and *Japonica* varietal groups were domesticated from pre-differentiated gene pools that differed in Al tolerance. The loci identified in this study can be

utilized in breeding programs to exploit transgressive variation within and across rice varietal groups.

INTRODUCTION

Aluminum toxicity is the major constraint to crop productivity on acid soils, which comprise over 50% of the world's arable land (Von Uexkull and Mutert, 1995). Under highly acidic soil conditions (pH<5.0) Al is solubilized to Al^{3+}, which is highly phytotoxic, causing a rapid inhibition of root growth that leads to a reduced and stunted root system, thus having a direct effect on the ability of a plant to acquire both water and nutrients.

Cereal crops (*Poaceae*) have been a primary focus of Al tolerance research (Kochian et al., 2004). This research has demonstrated that levels of Al tolerance vary widely both within and between species (Foy, 1988, Sasaki et al., 2006, Piñeros et al., 2005, Furukawa et al., 2007; Caniato et al., 2007, Famoso et al., 2010). Of the major cereal species that have been extensively studied (rice, maize, wheat, barley and sorghum), rice demonstrates superior Al tolerance under both field and hydroponic conditions (Foy, 1988; Famoso et al., 2010). Although rice is 6-10 times more tolerant than other cereals, such as maize, wheat, and sorghum, very little is known about the genes underlying this tolerance. Based on its high level of Al tolerance and numerous genetic and genomic resources, rice provides a good model for studying the genetics and physiology of Al tolerance.

In wheat, sorghum, and barley, Al tolerance is inherited as a simple trait, controlled by one or a few genes (Sasaki et al, 2004; Magalhaes et al, 2004; Minella and Sorrells, 1992). However, in maize, rice, and *Arabidopsis,* tolerance is quantitatively inherited

(Ninamango-Cardenas et al., 2003; Hoekenga et al., 2003; Nguyen et al., 2001 and 2002). Al tolerance genes have been cloned in wheat and sorghum. The wheat resistance gene, *ALMT1*, encodes an Al-activated, malate transporter (Sasaki et al., 2004). The sorghum resistance gene, *SbMATE*, encodes a member of the multidrug and toxic compound- extrusion (MATE) family and is an Al-activated, root citrate efflux transporter (Magalhaes et al, 2007).

Three mutant genes that lead to Al sensitivity have recently been cloned in rice, *STAR1, STAR2* (Sensitive to Al rhizotoxicity) and *ART1* (Aluminum rhizotoxicity 1) (Huang et al, 2009, Yamaji et al, 2009). The products of *STAR1* and *STAR2* are expressed mainly in the roots and are components of a bacterial-type ATP binding cassette (ABC) transporter. Both are transcriptionally activated by exposure to Al and loss of function of either gene results in hypersensitivity to Al. *STAR1* and *STAR2* are similar to two Al sensitive mutants in *Arabidopsis*, *als1* and *als3*, also encoding ABC transporters (Larsen et al. 2005, 2007). *ART1* is a novel C2H2-type zinc finger-type transcription factor that interacts with the promoter region of *STAR1*. *ART1* is reported to regulate at least 30 down-stream transcripts, some of which are involved in Al detoxification and serve as strong candidate genes controlling rice Al tolerance (Yamaji et al., 2009). None of the three cloned rice genes map to previously reported Al tolerance QTL, suggesting that these genes may be involved in basal Al tolerance (Huang , 2009, Yamaji et al, 2009). A more thorough analysis is necessary to determine whether there might be natural variation associated with these loci that would help trace their evolutionary origins and clarify their contribution to the high levels of Al tolerance observed in rice.

Seven QTL studies on Al tolerance have been reported in rice using six different inter- and intra-specific mapping populations (Wu et al., 2000; Nguyen et al., 2001, 2002, 2003; Ma et al., 2002; Xue et al., 2006, 2007). Together, these studies report a total of 33 QTLs, located on all 12 chromosomes, with three intervals (on chromosomes 1, 3, and 9) being detected in multiple studies. In all of the QTL studies, Al tolerance was estimated based on relative root growth (RRG), and specifically on inhibition of the growth (elongation) of the longest root (elongation of the longest root in Al treatment / root growth of controls). A very weak correlation was recently demonstrated between RRG of the longest root and RRG of the total root system (R^2=0.17) (Famoso et al., 2010). This raises the question of whether mapping QTLs using the two RRG indices independently might identify novel loci, potentially leading to the discovery of QTL that co-localize with recently cloned genes underlying Al sensitive mutants.

Historically, *O. sativa* has been classified into two varietal groups, *Indica* and *Japonica*, based on morphological characteristics, ecological adaptation, crossing ability and geographic origin (Oka, 1988). These two varietal groups are believed to represent independent domestications from a pre-differentiated ancestral gene pool (*O. rufipogon*), followed by significant gene flow among and between subpopulations (Zhou et al., 2003; Barbier, 1989; Zhu and Ge, 2005; Ma and Bennetzen, 2004; Vitte et al., 2004, Londo et al., 2006; Sweeney et al., 2007; Kovach and McCouch, 2008). These two varietal groups (names are italicized with an upper case first letter, i.e., *Indica* and *Japonica*) have been further divided into five major subpopulations (names are italicized in all lower-case letters) (*indica, aus, tropical japonica, temperate japonica,* and *aromatic [group V]*) based on DNA markers (SSR, SNPs, indels, etc.) (Garris et al., 2005; Caicedo et al., 2007; Zhao et al., 2010). Genotypes that share <80% ancestry across subpopulations or varietal groups are classified as admixed

varieties (Zhao et al., 2010), while smaller groups adapted to specific ecosystems may be recognized as upland, deep water, or floating varieties (Glaszmann, 1987; Khush et al. 1997). Upland varieties, which are grown at high altitudes on dry (non-irrigated) soils, are those most commonly exposed to Al-toxic conditions. These varieties are almost invariably of *tropical japonica* origin, suggesting *a priori* that the *tropical japonica* subpopulation would be the most likely source of alleles for Al tolerance in rice.

O. sativa has elevated levels of linkage disequilibrium (LD) compared to species such as *Arabidopsis*, maize and human. The average extent of LD in rice has been reported to be between 50-500 kb (Garris et al., 2003; Olsen et al., 2006; Mather et al., 2007; Rakshit et al, 2007), compared to 100-250 kb in *Arabidopsis* and human (Nordborg et al., 2002; Daly et al., 2001; Jeffreys et al., 2001; Reich et al., 2001) and 1-2 kb in maize (Tenaillon et al., 2001; Remington et al., 2001). This is due to the inbreeding nature of rice, coupled with its demographic history. Strong selective pressure over the course of rice domestication has also led to deep population substructure (F_{st}=0.23 to 0.57) (Garris et al., 2005; Ali et al., 2010; Zhao et al., 2010), which sets it apart from *Arabidopsis* (Platt et al., 2010; Atwell et al., 2010). Population substructure can lead to false-positives in association mapping studies, and must be taken into account (Yu et al., 2005; Zhao et al., 2007; Atwell et al., 2010). The mixed-model has been demonstrated to work well in both maize and *Arabidopsis* to assess the contribution of population structure (Yu et al., 2005; Atwell et al., 2010), but its performance in rice is untested.

In rice, the small genome size (390Mb) and extensive LD facilitates genome-wide association (GWA) analysis by allowing genome-wide coverage with ~36,000 SNPs.

A diversity panel consisting of 400 *O. sativa* and 100 *O. rufipogon* accessions, representing the genetic diversity of the primary gene pool of domesticated rice, was recently genotyped with 44,000 SNPs (1-2 SNPs / LD block) (Tung et al., submitted; Zhao, K and McCouch, S., personal communication.). The slow rate of LD decay, while facilitating GWA analysis, limits the resolution of association mapping in rice. In maize, the resolution of association mapping is reported to be at the gene level (1-3 genes), whereas it is typically ~20-50 genes in the *indica* and *aus* subpopulations of rice (~100 kb; Garris et al., 2003; Mather et al., 2006; Rakshit et al, 2007), and up to ten times lower (> 1,000 genes) in the *japonica* subpopulations. When compared to the resolution of a QTL study consisting of 250 lines (~10-20 cM resolution, where 1cM = ~250kb), association mapping tends to provide between 10-20 times higher resolution for a population of similar size. In both cases, fine-mapping and/or mutant analysis is also generally required to identify the gene(s) underlying a region of interest. By doing association mapping for QTL discovery, the fine-mapping phase can generally be focused on a smaller target region.

To investigate the genetic architecture of Al tolerance in rice, bi-parental QTL mapping was conducted in two mapping populations using relative root growth of the primary root, the longest root, and the total root system as three related phenotypic measures of Al tolerance (Famoso et al., 2010). Additionally, genome wide association analysis for rice Al tolerance was also conducted using the 44,000 SNPs that had been recently genotyped on the rice diversity panel (Tung, CW and McCouch, S, personal communication, 2010). Regions of the rice genome associated with Al tolerance identified by GWA were compared with regions identified as QTLs in this and previous studies, as well as with Al sensitive mutants and/or candidate genes. Phenotypic outliers identified in the diversity panel were further investigated to

identify regions of subpopulation-admixture that accounted for extreme Al tolerance phenotypes.

RESULTS

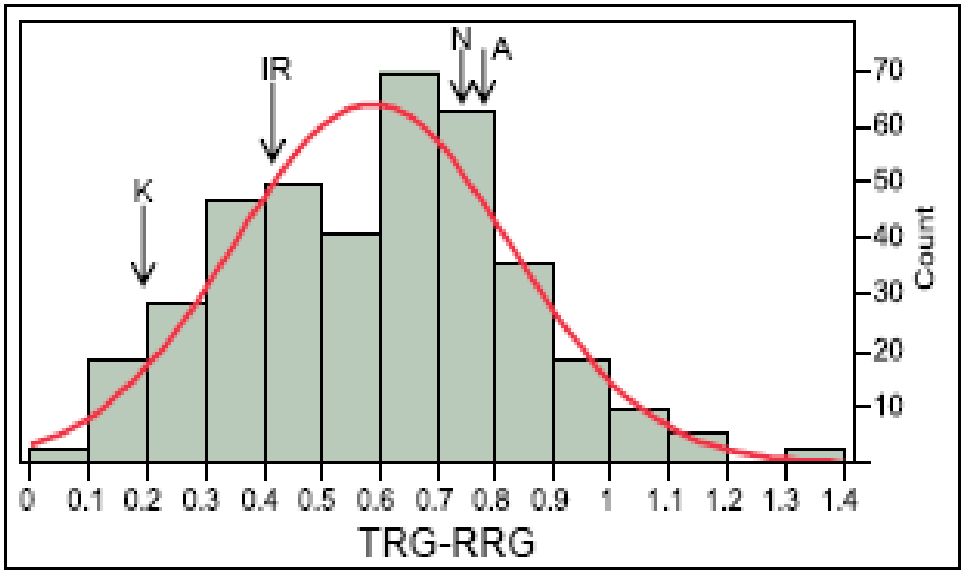
Al tolerance in rice

Three hundred eighty-five diverse accessions from the rice diversity panel (Ali et al., 2010; Tung et al., submitted; Zhao et al., 2010) (Supplemental Table 3.1) were evaluated for Al tolerance in our rice nutrient solution containing 160μM Al^{3+} activity using the phenotyping platform described by Famoso et al. (2010). This Al^{3+} activity had been previously identified as optimal for evaluating a wide range of Al tolerance in diverse rice germplasm (Famoso et al., 2010). In this diversity panel, Al tolerance, measured as relative root growth of the total root system (TRG-RRG), was normally distributed around a mean of 59% +/-24(SD) and ranged from 3-135% (Figure 3.1A). Some varieties were inhibited by as much as 97%, while 16 varieties (representing three subpopulations) showed greater root growth in the Al treatment (Supplemental Table 3.1).

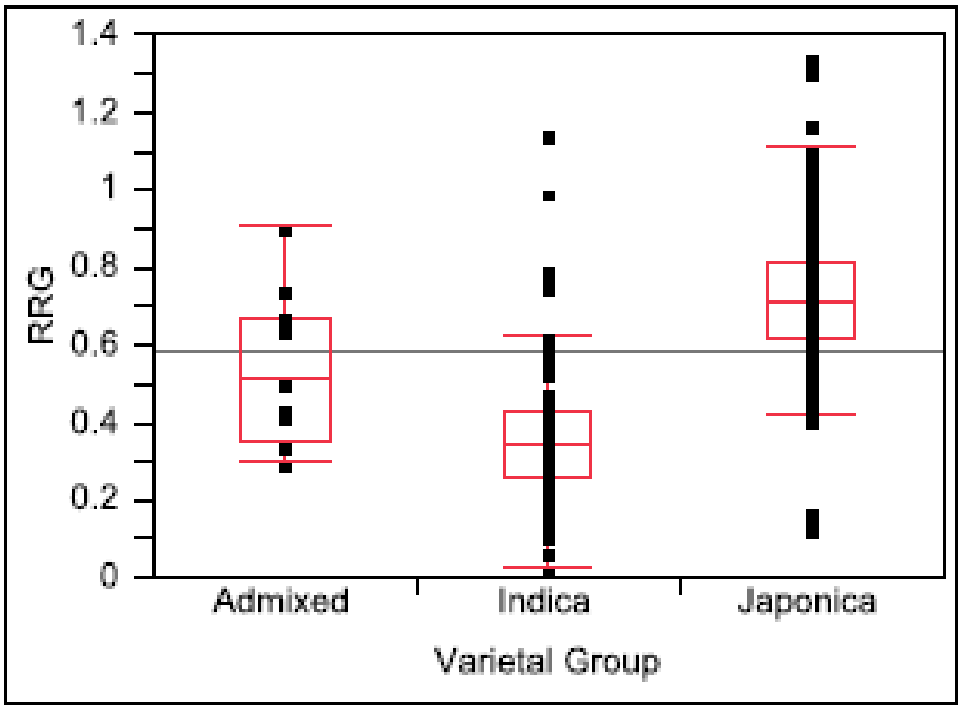
Figure 3.1. Al Tolerance Distribution and Subpopulation Variation
3.1A: Distribution of Al Tolerance. Distribution of total root growth Al tolerance (TRG-RRG) across 385 diverse accessions of *O. sativa* at 160µM Al^{3+}. Aluminum tolerance (TRG-RRG) was normally distributed around a mean of 0.59 +/-0.24(SD) and ranged from 0.03-1.35. The Al tolerance of the QTL mapping parents are indicated: K=Kasalath, IR=IR64, N=Nipponbare, A=Azucena.

3.1B: Al Tolerance by subpopulation. Variation of Al tolerance (RRG) within genetic varietal group (>80% identity). Admixed accessions contain <80% identity to either group. The *Japonica* varietal group (*temperate* and *tropical japonica* and *aromatic* subpopulations) is significantly more tolerant than the *Indica* varietal group (*indica* and *aus* subpopulations) ($p<0.0001$). Horizontal lines in boxplots represent 5^{th}, 25^{th}, 50^{th}, 75^{th}, and 95^{th} percentiles.

A


K
IR
N
A
70
60
50
40
30
20
10
Count
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1 1.1 1.2 1.3 1.4
TRG-RRG

B


1.4
1.2
1
0.8
0.6
0.4
0.2
0
RRG
Admixed
Indica
Japonica
Varietal Group

When accessions were grouped based on varietal group (>80% ancestry) the *Japonica* varietal group (consisting of the *temperate japonica, tropical japonica* and *aromatic* subpopulations) was significantly more tolerant than the *Indica* varietal group (*indica* and *aus* subpopulations) ($p<0.0001$) (Figure 3.1B). The *Japonica* varieties had a mean Al tolerance value of RRG = 0.72 and ranged from 0.13-1.35. The *Indica* varieties had a mean Al tolerance value of RRG = 0.36 and ranged from 0.03-1.15 (Figure 3.1B). Eleven accessions were classified as "admixed" between varietal groups, and these had a mean Al tolerance equal to the mean of all 385 accessions (TRG-RRG=0.59). A one-way ANOVA demonstrated that subpopulation explained 57% of the phenotypic variation observed for Al tolerance (TRG-RRG) among the 274 accessions that were classified into one of the five subpopulations of rice. Despite the differences in mean TRG-RRG between subpopulations, considerable variation was also detected within each subpopulation (Supplemental Figure 3.1).

QTL Analysis

Two immortalized QTL mapping populations were analyzed for Al tolerance. One consisted of 164 recombinant inbred lines (RIL) derived from the cross IR64/Azucena (Ahmadi et al., 2005), and the other was comprised of 78 backcross inbred lines (BIL) derived from the cross Nipponbare/Kasalath//Nipponbare (Lin et al., 1998). Both populations were evaluated for Al tolerance using three different indices of relative root growth (RRG): primary root relative root growth (PRG-RRG), longest root relative root growth (LRG-RRG), and total root relative root growth (TRG-RRG) (see Materials and Methods for details). The phenotypic distribution was approximately normal for each population, no matter which root screening index was used (illustrated for TRG-RRG in Supplemental Figure 3.2A and 3.2B). The Al^{3+} activity at which Al tolerance was screened was determined by identifying the Al^{3+} activity that provided

the greatest difference in tolerance between the parents. The tolerant parent of the RIL population, Azucena, and the tolerant parent of the BIL population, Nipponbare, are similar in Al tolerance, whereas the susceptible parent of the RIL population, IR64, is significantly more tolerant than the susceptible parent of the BIL population, Kasalath (Figure 3.1A). Based on the comparison between the mapping population parents, the RIL population was evaluated at 250μM Al^{3+} and the BIL population at 120μM Al^{3+}.

The RIL population (Azucena x IR64) had a mean TRG-RRG of 39% when assayed at 250μM Al^{3+}, with a range of 21–67% (Supplemental Figure 3.2A). Under control conditions the genetic component of phenotypic variation was 0.46, while in the Al^{3+} treatment the genetic component of phenotypic variation was 0.35. Transgressive segregation was observed in 20% of the RILs, with 10% of the population demonstrating greater Al tolerance than Azucena (the tolerant parent) and 10% demonstrating greater susceptibility than IR64 (the susceptible parent). The BIL population (Nipponbare x Kasalath) had a mean TRG-RRG value of 73% when assayed at 120μM Al^{3+}, with a range of 45–120%. In control conditions, the genetic component of phenotypic variation was 0.45 while in Al^{3+} treatment the genetic component of phenotypic variation was 0.55. Transgressive segregation was only observed for increased Al tolerance, as no BIL was more susceptible than the Kasalath parent.

The method of phenotyping, specifically, the RRG index used to estimate Al tolerance, directly impacted the significance of QTLs detected by composite interval mapping (Figure 3.2A-3.2C and 3.3A-3.3C). In the RIL population, three Al tolerance (*Alt*) QTL were detected using relative root growth of the total root system (TRG-RRG): *Alt*$_{TRG}$*1.1* on chromosome 1, *Alt*$_{TRG}$*2.1* on chromosome 2, and *Alt*$_{TRG}$*12.1* on

chromosome 12 (Figure 3.2A-3.2C, Table 3.1). The Azucena allele conferred increased tolerance at the loci on chromosomes 1 and 12 and reduced tolerance at the locus on chromosome 2. QTLs were detected in the same positions on chromosomes 1 and 12 using the PRG-RRG index, though with lower LOD scores (Figure 3.2A-3.2C, Table 3.1). Using the LRG-RRG index, a single QTL was detected on chromosome 9, $Alt_{LRG}9.1$, and this QTL was not detected using either the TRG-RRG or PRG-RRG indices. The major QTL on chromosome 12, located between 2.69 – 5.10 Mb, was detected only using TRG-RRG and PRG-RRG and encompasses the Al sensitive mutant *ART1*, which is located at 3.59 Mb.

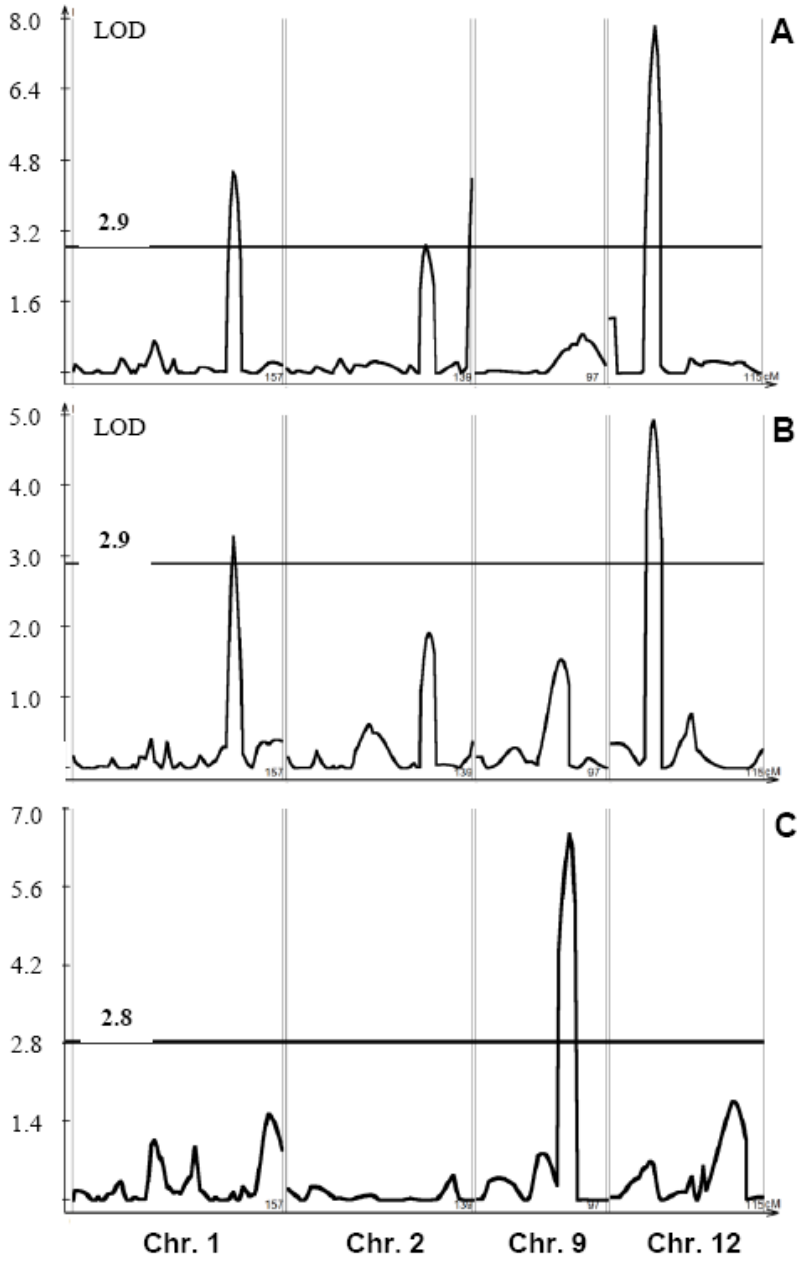


Figure 3.2. QTL Detected in RIL Population. Composite interval mapping output for significant loci detected in the RIL mapping population using three Al tolerance relative root growth (RRG) indexes. The Y-axis is the LOD score and the horizontal line is the significant LOD threshold based on 1000 permutations. 3.2A) Total root growth RRG; 3.2B) Primary root growth RRG; 3.2C) Longest root growth RRG.

Table 3.1. Summary of Al Tolerance QTLs. Summary of QTLs (1000 permutations) identified by composite interval mapping in the RIL and BIL populations. Al tolerance (RRG) QTLs were identified using three root growth parameters, total root growth (TRG), primary root growth (PRG), and longest root growth (LRG).

Trait Index	Pop.	Chr	QTL	Peak Marker	Mb Pos.	Flanking Markers	LOD	Additive effect	R^2
TRG-RRG	RIL	1	*Alt_{TRG} 1.1*	RM265	35.2	RM319/RM315	4.56	2.58 (Azu)	0.1
PRG-RRG	RIL	1	*Alt_{PRG} 1.1*	RM265	35.2	RM319/RM315	3.29	3.84 (Azu)	0.08
TRG-RRG	BIL	1	*Alt_{TRG} 1.2*	RM6333	38	RM5448/RM823	3.44	-10.58 (Nip)	0.12
TRG-RRG	RIL	2	*Alt_{TRG} 2.1*	RM221	27.6	RM526/RM318	2.9	-2.08 (IR64)	0.06
PRG-RRG	BIL	6	*Alt_{PRG} 6.1*	L688	5.81	R1954/G200	3.95	12.78 (Kas)	0.14
LRG-RRG	RIL	9	*Alt_{LRG} 9.1*	RM242	18.8	RM257/RM160	6.57	4.42 (Azu)	0.17
TRG-RRG	RIL	12	*Alt_{TRG} 12.1*	RM247	3.19	RM453/RM512	7.85	3.76 (Azu)	0.19
PRG-RRG	RIL	12	*Alt_{PRG} 12.1*	RM247	3.19	RM453/RM512	4.94	4.75 (Azu)	0.13
TRG-RRG	BIL	12	*Alt_{TRG} 12.2*	R2708	23.3	R1709/G2140	3.49	12.3 (Kas)	0.13

In the BIL population, two QTL were detected using the TRG index, *Alt_{TRG}1.2* on chromosome 1, which co-localized with the *Alt_{TRG}1.1* QTL identified in the RIL population, and *Alt_{TRG}12.2* on chromosome 12, which did not overlap with the *Alt_{TRG}12.1* identified in the RIL population (Figure 3.3A-3.3C, Table 3.1). The Nipponbare allele conferred resistance at the chromosome 1 locus and the Kasalath allele conferred resistance at the *Alt_{TRG}12.2* locus. Using the PRG-RRG index, one QTL was detected on chromosome 6, where the Kasalath allele conferred resistance. No QTLs were detected using the LRG-RRG index in the BIL population.

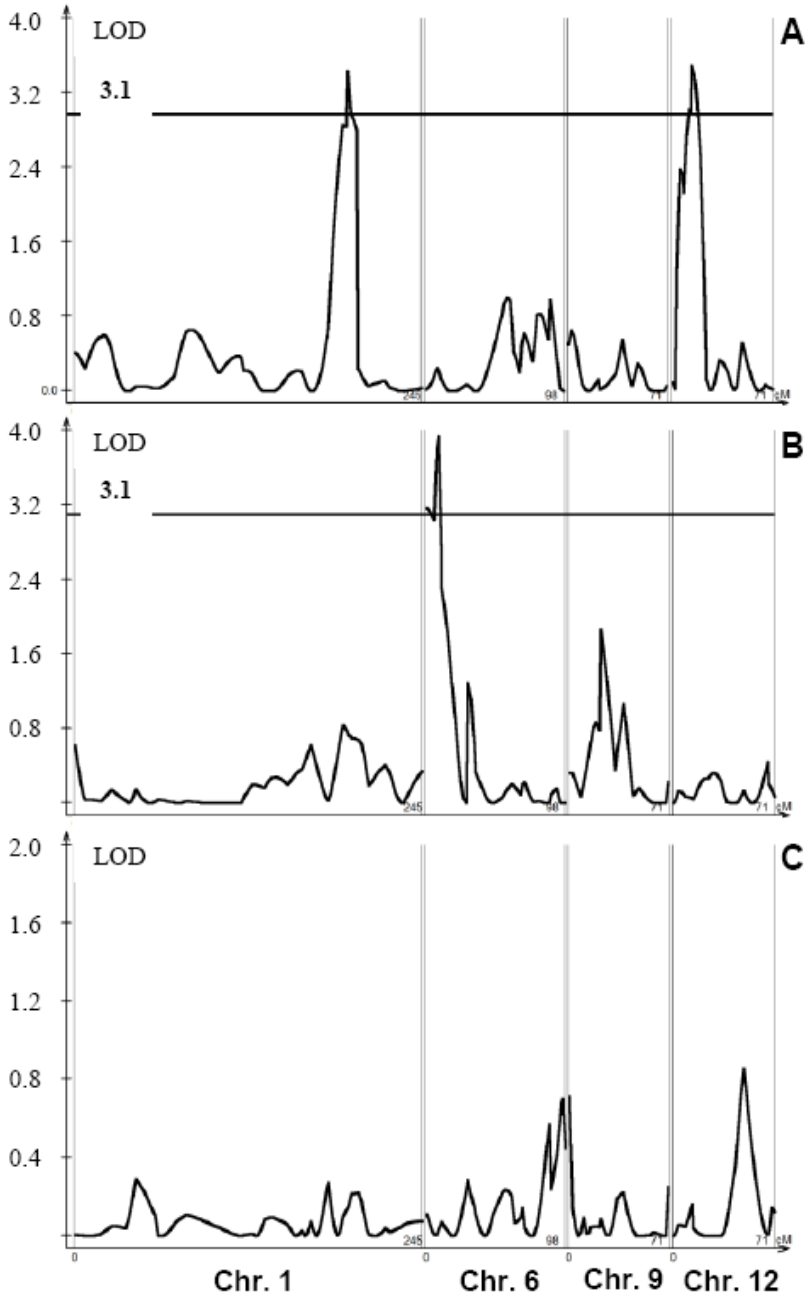


Figure 3.3. QTL Detected in the BIL Population. Composite interval mapping output for Al tolerance loci detected in the BIL mapping population using the three Al tolerance RRG indexes. The Y-axis is the LOD score and the horizontal line is the significant LOD threshold based on 1000 permutations. 3.2A) Total root growth; 3.2B) Primary root growth; 3.2C) Longest root growth.

Identification of Al tolerance loci through GWA mapping

To identify Al tolerance loci based on genome-wide association (GWA) mapping, we used a previously generated genotypic dataset consisting of 44,000 SNPs and Al tolerance phenotypes (TRG-RRG) generated on 385 *O. sativa* accessions in this study. GWA mapping was conducted in two steps; first using principle component analysis (PCA) to eliminate the confounding subpopulation effect across the diversity panel as a whole, and second using the Efficient Mixed-Model Association (EMMA) when association analysis was conducted independently within the *indica*, *aus*, *temperate japonica*, and *tropical japonica* subpopulations (Fig. 3.4A-3.4E). The *aromatic* subpopulation was only represented by 12 genotypes in our panel, and therefore we did not undertake GWA mapping within the *aromatic* subpopulation.

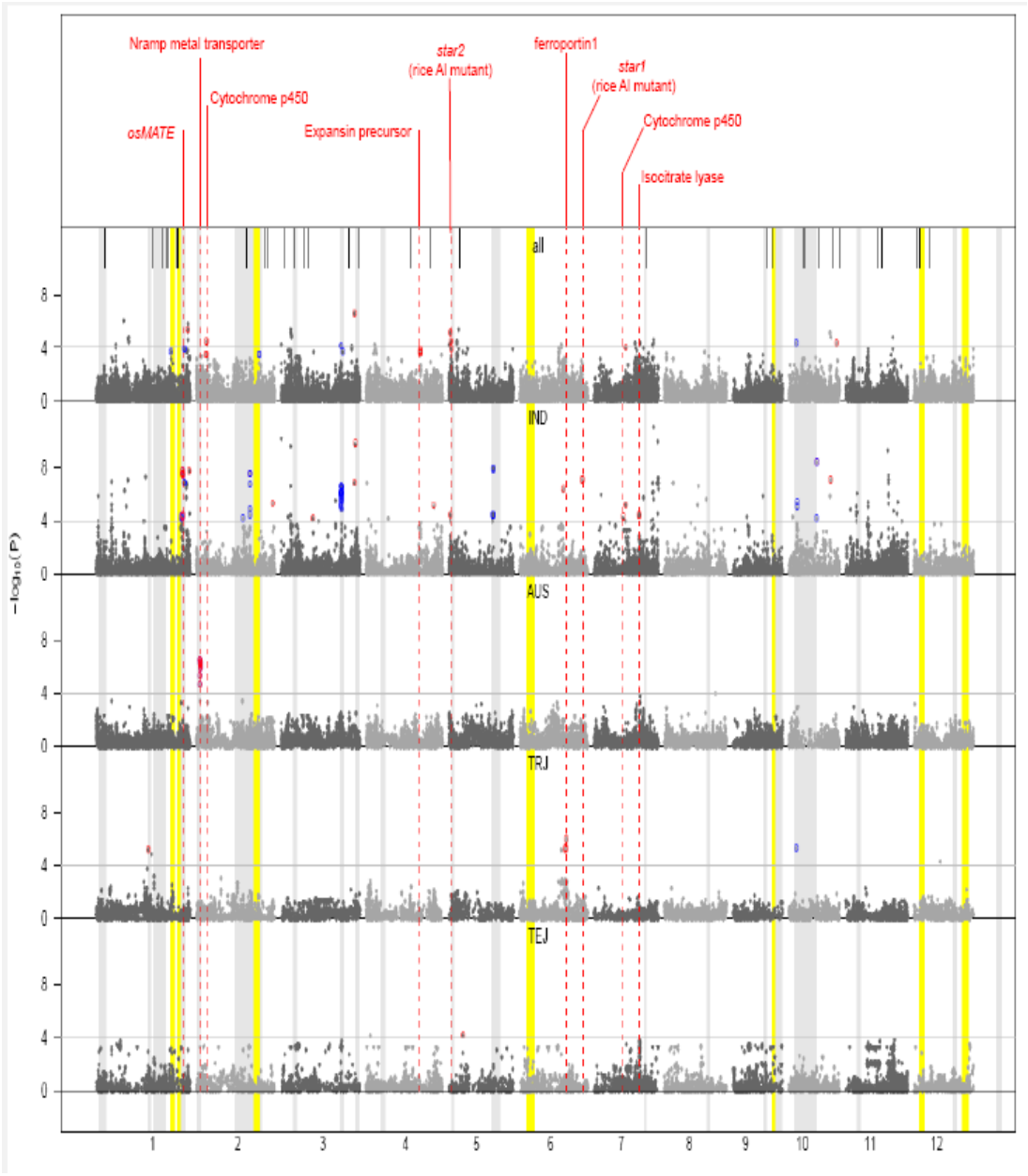


Figure 3.4. Plot of GWA Peaks Across and within Subpopulation. Negative log of p-value of SNPs (y-axis) plotted by chromosome and location (x-axis). Analysis was conducted across all 385 genotypes using Principle Component Analysis (PCA). Analysis was conducted within each subpopulation using the mixed-model (EMMA). *A priori* candidate genes identified within 500kb of significant SNPs are indicated along the top; all were identified to have altered expression in Al treatment in the *art1* Al sensitive transcription factor mutant (Yamaji, 2009). Color bands indicate QTL positions from previous reports (grey) or from this study (yellow). SNP color indicates co-localization with QTLs (blue) and/or candidate genes (red).

A total of 43 independent regions (>1Mb apart) were detected using GWA analysis ($p<1.8E^{-4}$); this includes loci detected across all genotypes using PCA (18 regions) as well as loci detected within the four subpopulations using the mixed-model (38 regions), with 13 loci identified in common by both methods. The 13 loci identified by both methods demonstrate that these loci segregate within two or more subpopulations (i.e., within both *indica* and *aus,* or within *tropical* and *temperate japonica,* etc.). The five loci that were associated with Al tolerance only when the entire diversity panel was analyzed together are not strongly associated with any particular subpopulation; the alleles conferring tolerance exist at low frequency or are not variable within several of the subpopulations (Fig. 3.4, Table 3.2).

When subpopulations were analyzed individually, the number and significance of the regions identified by GWA analysis varied considerably, due in large part to the small number of accessions belonging to each subpopulation within the panel (Table 3.2). Within the *indica* subpopulation (n=67), 34 significant regions associated with TRG-RRG were detected (21 with a MAF>0.05). Five of these regions co-localized with regions that were also detected within other subpopulations and 13 co-localized (<1 Mb) with regions identified across all genotypes (Figure 3.4, Table 3.2).

One clear cluster of SNPs associated with Al tolerance ($p=2.8E^{-7}$, MAF=0.19) identified within the *aus* subpopulation on chromosome 2 was unique to the *aus* subpopulation (Figure 3.4, Table 3.2). Another Al tolerance region was identified ($p=1.8E^{-4}$, MAF=0.24) within the *aus* subpopulation on chromosome 7 but it co-localized with peaks identified in the *indica* subpopulation, as well as across all genotypes (Table 3.1).

Four significant regions associated with Al tolerance were identified within the *tropical japonica* subpopulation, however all had a MAF <0.05 (Figure 3.4, Table 3.2). Two regions on chromosome 6 and 10 co-localized with peaks identified both within the *indica* subpopulation and across all genotypes.

Within the *temperate japonica* subpopulation three regions were associated with Al tolerance (1 with MAF>0.05) and two co-localized with regions identified within *indica*, across all genotypes, or with a-priori candidate genes (Figure 3.4, Table 3.2).

Table 3.2. Summary of GWA Analysis Results. Significant regions identified by GWA analysis across all 385 genotypes (PCA) and within subpopulations. SNPs within 1Mb are grouped in the same "GWA Region". QTL and a priori candidate genes within 1Mb of GWAS regions are identified.

GWA Region	Chr	Mb Pos.	SNPid	Subpop/ Model	p-value	MAF	QTL (<1Mb)	Candidates (ref.)	Pos.
1	1	1.06	id1000974	IND 1	1.4E-06	0.012	Nguyen V, 2002	no	
2	1	7.27	dd1000612	IND 2	6.3E-06	0.070	no	no	
3	1	8.99	id1006703	PCA 1	6.1E-05	0.094	no	no	
4	1	12.75	id1008716	PCA 2	9.3E-07	0.013	no	no	
4	1	14.53	id1009673	IND 3	7.8E-08	0.138	no	no	
4	1	15.13	id1009968	PCA 2	3.4E-05	0.051	no	no	
5	1	22.45	id1012735	IND 4	9.9E-05	0.089	no	no	
5	1	22.91	id1013116	IND 4	4.6E-08	0.036	no	no	
6	1	24.10	id1014178	TRJ 1	6.5E-06	0.021	Xue, 2006	LOC_Os01g46350 (Yamaji et al., 2009)	26.37
7	1	42.33	id1027542	PCA 3	4.2E-06	0.035	Ma, 2002; Wu, 2000	LOC_Os01g69020 (Yamaji et al., 2009)	40.10
7	1	42.98	id1028371	IND 5	1.8E-08	0.078	Wu, 2000	LOC_Os01g69010 (Magalhaes et al., 2007)	40.09
7	1	42.98	id1028371	PCA3	4.2E-04	0.405			
8	2	1.58	id2001198	Aus 1	2.8E-07	0.186	Ma, 2002	LOC_Os02g03900 (Yamaji et al., 2009)	1.65
9	2	4.58	id2002406	PCA 4	3.4E-05	0.324	no	LOC_Os02g09390 (Yamaji et al., 2009)	4.82
10	2	24.58	id2010437	IND 6	3.0E-08	0.034	Nguyen V, 2001	LOC_Os02g38200 (Krill et al., 2010)	23.10
11	2	35.13	id2016040	IND 7	5.2E-06	0.176	no	LOC_Os02g53130 (Yamaji et al., 2009)	32.51
12	3	0.16	id3000010	IND 8	6.2E-11	0.100	no	no	
13	3	4.63	id3002724	PCA 5	4.8E-06	0.374	Wu, 2000	no	
13	3	4.71	id3002777	IND 9	2.6E-10	0.074	Wu, 2000	no	
14	3	14.83	ud3000829	IND 10	8.0E-05	0.078	no	LOC_Os03g21950.1	14.38
15	3	27.78	id3012398	PCA 6	7.3E-05	0.362	Nguyen V, 2002	no	
15	3	27.86	id3012499	IND 11	3.0E-07	0.069	Nguyen V, 2002	no	
16	3	33.93	id3016592	PCA 7	2.6E-07	0.158	Nguyen V, 2001	Os03g0760800 (Yamaji et al., 2009)	35.66
16	3	34.39	id3016816	IND 12	1.7E-10	0.047	Nguyen V, 2001	Os03g0760800 (Yamaji et al., 2009)	35.66
17	4	0.97	id4000574	IND 13	8.6E-05	0.284	no	no	
17	4	1.86	id4000893	TEJ 1	5.6E-05	0.051	no	no	

Table 3.2 (continued)

GWA Region	Chr	Mb Pos.	SNPid	Subpop/ Model	p-value	MAF	QTL (<1Mb)	Candidates (ref.)	Pos.
18	4	4.17	id4001631	PCA 8	9.4E-05	0.043	no	no	
19	4	10.25	id4003335	IND 14	6.9E-05	0.063	no	no	
20	4	31.24	id4010688	IND 15	7.8E-06	0.139	no	LOC_Os04g49410 (Yamaji et al., 2009)	29.30
21	5	0.48	id5000230	PCA 9	6.8E-06	0.159	Xue, 2006	1) LOC_Os05g02750 (Yamaji et al., 2009 and Huang et al., 2009); 2) LOC_Os05g02780 (Yamaji et al., 2009)	0.989; 1.0
21	5	0.48	id5000230	IND 16	4.1E-05	0.296	Xue, 2006		
22	5	6.32	ud5000331	TEJ 2	6.5E-05	0.014	no	1) LOC_Os05g09440; 2) LOC_Os05g08810 (Krill et al., 2010)	5.29; 4.85
23	5	20.26	id5008491	IND 17	1.5E-08	0.023	Nguyen V, 2001	LOC_Os06g36450 (Krill et al., 2010)	21.14
24	6	19.02	id6010168	PCA 10	5.6E-05	0.205	no	no	
24	6	19.19	id6010206	TRJ 2	5.7E-06	0.010	no	no	
24	6	20.17	id6010607	IND 18	3.9E-07	0.037	no	LOC_Os06g36450 (Krill et al., 2010)	21.40
25	6	28.80	id6016397	IND 19	9.5E-08	0.038	no	LOC_Os06g48060 (Huang et al., 2009; Yamaji et al., 2009)	29.07
26	7	3.36	ud7000240	IND 20	9.8E-07	0.014	no	no	
27	7	14.65	ud7001005	IND 21	7.4E-06	0.038	no	LOC_Os07g23710 (Yamaji et al., 2009)	13.38
27	7	14.65	ud7001005	PCA 11	9.6E-05	0.308	no	no	
28	7	20.69	ud7001465	IND 22	4.5E-05	0.079	no	LOC_Os07g34520 (Krill et al., 2010)	20.69
28	7	20.72	ud7001467	PCA 12	4.4E-04	0.233			
28	7	21.25	id7003436	AUS 2	1.8E-04	0.237			
29	7	27.27	ud7002021	PCA 13	2.6E-05	0.479			
29	7	27.27	ud7002021	IND 23	9.3E-12	0.090	no		
30	8	3.68	id8001212	IND 24	1.5E-07	0.033	no	no	
31	8	15.44	id8004158	IND 25	2.6E-07	0.073	no		
32	8	19.21	id8005103	IND 26	2.6E-06	0.100	Nguyen B, 2003	no	
33	9	6.44	id9001787	IND 27	5.0E-05	0.031	no	no	
34	9	13.87	id9003762	IND 28	1.5E-06	0.080	Xue, 2006	no	

Table 3.2 (continued)

GWA Region	Chr	Mb Pos.	SNPid	Subpop/ Model	p-value	MAF	QTL (<1Mb)	Candidates (ref.)	Pos.
35	**10**	3.38	ud10000228	PCA 14	4.8E-05	0.031	Nguyen V, 2002	1) LOC_Os10g13940 (Yamaji et al., 2009); 2) LOC_Os10g12080 (Krill et al., 2010)	7.59; 6.73
35	**10**	3.38	ud10000228	TRJ 3	4.7E-06	0.010			
35	**10**	3.70	id10001120	IND 29	4.9E-06	0.017			
36	**10**	12.63	wd10002947	IND 30	3.8E-09	0.014	Nguyen V, 2002	LOC_Os10g26680 (Krill et al., 2010)	13.86
37	**10**	19.20	id10006009	IND 31	8.8E-08	0.045	no	LOC_Os10g38080 (Yamaji et al., 2009)	20.32
38	**10**	21.97	id10006959	PCA 15	4.7E-05	0.058	no	LOC_Os10g42780 (Yamaji et al., 2009)	23.00
39	**11**	4.99	id11001994	IND 32	9.5E-06	0.080	Xue, 2006	no	
40	**11**	10.25	id11003766	IND 33	6.0E-05	0.127	no	no	
40	**11**	10.45	id11003835	PCA 16	2.2E-04	0.375	no	no	
41	**11**	19.27	id11007381	IND 34	4.8E-10	0.080	no	no	
41	**11**	21.80	dd11000269	PCA 17	1.5E-05	0.012	no		
41	**11**	22.49	ud11001343	TEJ 3,	1.3E-04	0.023			
42	**12**	12.06	id12004429	TRJ 4	4.2E-05	0.019	no	no	
43	**12**	27.32	id12010071	PCA 18	1.1E-04	0.223	no	no	

Co-localization of GWAS with QTLs and candidate genes

A list of 49 a-priori Al tolerance candidate genes (Supplemental Table 3.2), was compiled based on Al sensitive mutants from rice and *Arabidopsis*, cloned Al tolerance genes from wheat and sorghum, expression profiles from Al treated maize and rice roots (Maron et al., 2008; Yamiji et al., 2009), and an association study on specific candidate genes of maize (Krill et al., 2010). These 49 genes were located in 37 different (1Mb) regions of the genome, as 12 of the genes were within 1Mb of another candidate gene. Twenty of the 37 (54%) 1Mb candidate regions containing candidate genes were significantly associated with Al tolerance based on GWA mapping in this study (Table 3.2). In addition, 40 rice QTLs associated with Al tolerance were identified from the literature, including 9 from this study, and 18 (40%) of these co-localized with regions identified by GWA mapping (>1Mb apart). In nine regions there was an intersection between a GWA mapping locus, a previously reported Al tolerance QTL and a candidate gene.

As illustrated in Figure 3.5, our GWA results identified eight significant SNPs (p-values=$2.3E^{-5}$-$2.8E^{-7}$) in a 139kb region containing 27 genes on chromosome 2 between 1.536 – 1.675 Mb. Previously, a QTL had been reported in the same location (0.536-1.9 Mb) where the susceptible parent was of *aus* origin and the tolerant parent was a *temperate japonica* variety (Ma et al. 2002). The strongest candidate gene in this region of chromosome 2, (LOC_Os02g03900 located at 1.66 Mb), encodes a Nramp6 metal transporter and was identified as a candidate in the Al tolerance gene network, based on altered expression patterns in Al-treated roots of a mutant containing a non-functional version of the *ART1* gene, located on chromosome 12 (Yamija et al., 2009).

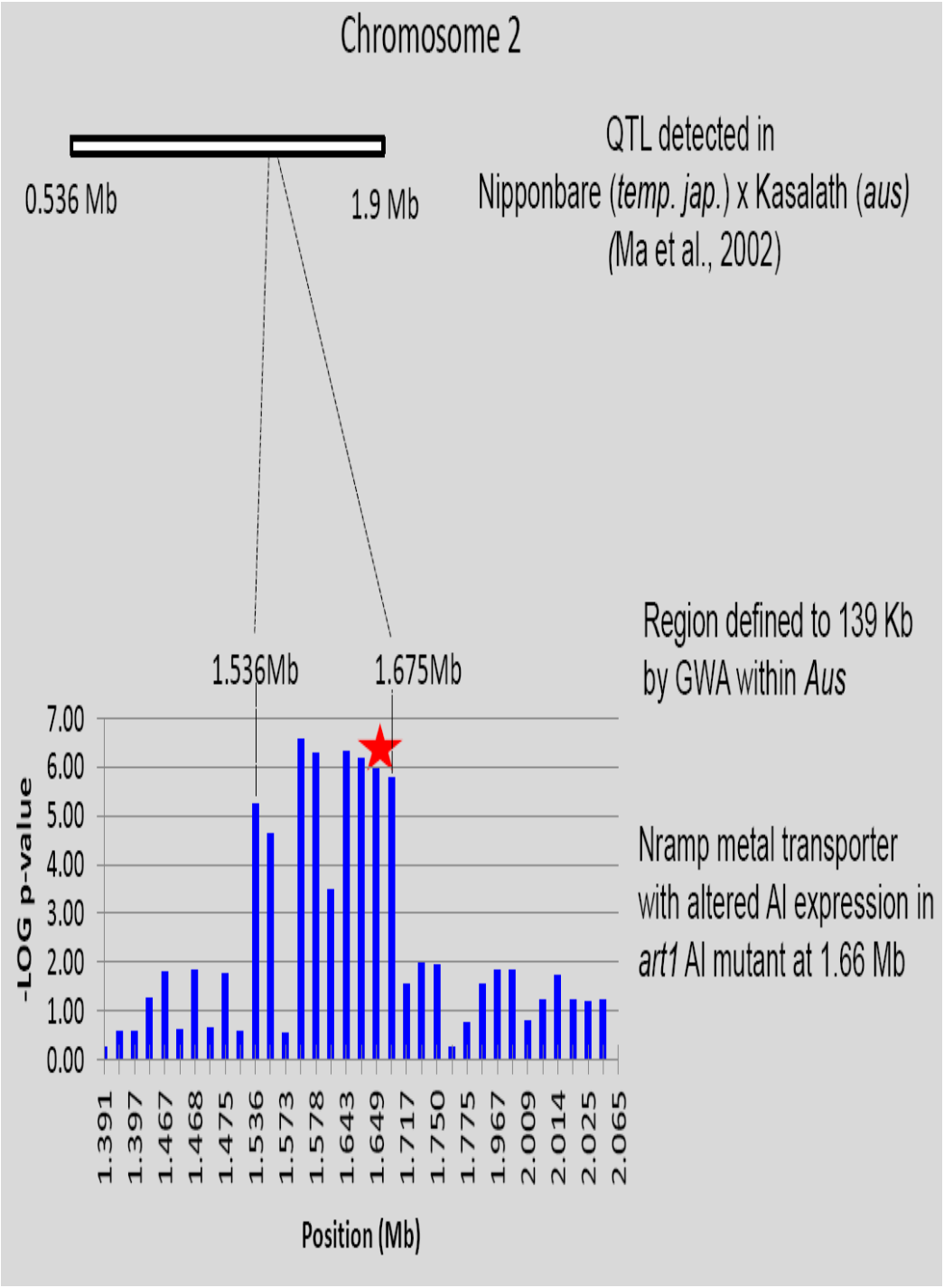


Figure 3.5. Chromosome 2 Region Delimited to Nramp Gene. This region on chromosome 2 was previously identified by QTL analysis (Ma et al., 2002) and was detected by GWAS within the *aus* subpopulation. The GWAS region encompassed a previously identified candidate Nramp gene.

Neither the Nramp6 metal transporter nor the *ART1* gene have previously been shown to associate with natural variation for Al tolerance prior to this study. Using the 44,000 SNP data to build haplotypes across the region, we demonstrated that 9 Al susceptible *aus* shared (including the six most Al susceptible *aus* varieties) in the association panel all contained a single haplotype at this locus which we termed the "susceptible haplotype", this Al susceptible haplotype was only found within the *aus* subpopulation, and the Kasalath parent used in the study by Ma et al. (2002) carried the susceptible, *aus*-specific haplotype, consistent with the population genetic analysis conducted here (Figure 3.6). The mean Al tolerance of the 9 accessions with the susceptible haplotype was RRG=0.16, while the mean of the 40 other *aus* accessions was RRG=0.38 ($p<0.001$). With the convergence of evidence from GWA, QTL, haplotype, and expression analysis, we hypothesize that LOC_Os02g03900 is the gene underlying Al tolerance in this region on chromosome 2 and demonstrate the power of whole genome association analysis to tie divergent pieces of evidence together to formulate a specific hypothesis.

Table 3.3. Haplotype analysis at the Nramp gene region on chromosome 2.

SNP ID (id200)	1179	1185	66	68	1196	1198	1200	1224	1231	1232	1234	1243	1244					
Position	1.536	1.542	1.544	1.564	1.573	1.575	1.577	1.642	1.648	1.648	1.649	1.671	1.675					
Hap #														AUS	IND	TEJ	TRJ	Total
1*	1	2	1	2	2	1	1	1	1	2	1	2	1	9				9
2	1	2	1	2	2	1	1	1	1	2	1	2	2	1				1
3	1	2	1	2	2	1	1	2	2	2	2	2	2	1	8			9
4	1	2	1	2	2	2	2	2	2	2	2	2	2	1	1			2
5	2	2	2	2	1	2	2	2	2	2	2	2	2	2	20			22
6	2	2	2	2	2	2	2	2	2	2	2	2	2	40	26	69		135
7	1	1	1	1	1	1	1	1	2	2	1	1	1				45	45
8	2	1	2	1	1	2	2	1	1	1	1	1	1				2	2
9	1	1	1	1	1	1	1	1	1	1	1	1	1				15	15
10	2	2	2	2	2	2	2	2	1	1	2	2	2		2	8		10
11	1	1	2	1	2	1	1	2	2	2	2	2	2		4			4
12	1	1	2	1	1	2	2	2	0	2	0	2	2		1			1
13	2	2	2	2	2	2	2	2	1	2	1	2	2		5			5

Similarly, we identified a significant phenotype-genotype association for SNPs located at 29.07 Mb on chromosome 6 near the *STAR1* (LOC_Os06g48060) locus, and at 0.989 Mb on chromosome 5 near the *STAR2/ALS3* (LOC_Os05g02750) locus. In both cases, significant SNPs ($p=9.5E^{-8}$ at 28.8 Mb on chromosome 6 near *STAR1* and $p= 6.4E^{-4} – 6.8E^{-6}$ at 0.42-0.93 Mb on chromosome 5 near *STAR2*) were found to be segregating within the *indica* subpopulation. This study provides the first evidence that there may be natural variation for Al tolerance in rice at *STAR1* and *STAR2*, loci that had previously been associated only with induced mutations.

Investigation of intersubpopulation introgressions to explain extreme phenotypes

There is a clear difference in the degree of Al tolerance found in the *Japonica* varietal group compared with the *Indica* varietal group, with the 10th percentile of Al tolerance of *Japonica* (0.53) being nearly equal to the 90th percentile of *Indica* (0.55) Al tolerance (Figure 3.1B). However, there are clear outliers within each varietal group. Five *Indica* accessions are highly Al tolerant (NSF 30, 66, 142, 163, 337), ranging from 0.76-1.15 TRG-RRG, and three *Japonica* accessions (NSF 12, 52, 112) are highly susceptible, ranging from 0.13-0.15 TRG-RRG (Figure 3.1B, Supplemental Table 3.1). To determine if these outliers are the result of introgressions across varietal groups, we calculated the allele ancestry of 5,467 SNPs distributed throughout the genome and identified specific genomic regions where historical *Indica* x *Japonica* admixture was only detected in the 8 outlier lines. To do this, *Japonica* introgressions identified in highly Al tolerant *Indica* lines were used to query all other *Indica* accessions and only those *Japonica* introgressions that were uniquely present in the outlier *Indicas* were considered as candidate regions underlying the outlier phenotype. When the five *Indica* outliers were used for this analysis, 2.4-4.9% of the genome corresponded to regions of *Japonica* introgression. NSF 52 contained *Indica* introgressions across 7.7% of the genome, while NSF 12 and NSF 112 were classified as *aus* (Group V) varieties and the small number of Group V varieties in this study and the high frequency of *aus* (Group V) alleles (>20%) common in this subpopulation precluded doing introgression analysis on the *Japonica* outliers. Therefore, the introgression analysis was only conducted within the five highly tolerant *indica* outliers. Across the five outlier accessions analyzed for admixture, 6 different introgressions (median size=780kb) were identified that were specific only to outlier genotypes of the respective varietal groups (Figure 3.6, Supplemental Table 3.3). Three of these introgressions were present in two genotypes, two of the introgressions

were present in three genotypes, and one introgression was present in four of the outlier genotypes. Five introgressions encompass SNPs identified by GWA, one co-localizes with a QTL, and one encompasses GWA SNPs and co-localizes with a QTL. The introgression that was present in four of the *indica* outlier genotypes was located on chromosome 7 between 27.72 – 28.29 Mb and contains 92 annotated genes. This introgression includes a cluster of highly significant SNPs identified by GWAS both within the *indica* subpopulation ($p=2.6e^{-5}$, MAF=0.10) and in the diversity panel as a whole (PCA) ($p=9.3e^{-12}$, MAF=0.25). The fact that this locus was identified by GWA within *indica* indicates that variation among *indica* varieties does exist for this trait, but because *indica* is relatively susceptible to Al, *indica* alleles are expected to confer low levels of tolerance. The fact that this locus was also identified by GWA across the entire diversity panel suggests that there may be an additional allele found outside of the *indica* subpopulation. Evidence from the outlier analysis described above, where a *Japonica* introgression confers superior Al tolerance in the *indica* genetic background, strongly suggests that there are novel alleles in *Japonica* varieties that can enhance the low levels of Al tolerance naturally found in *indica*.

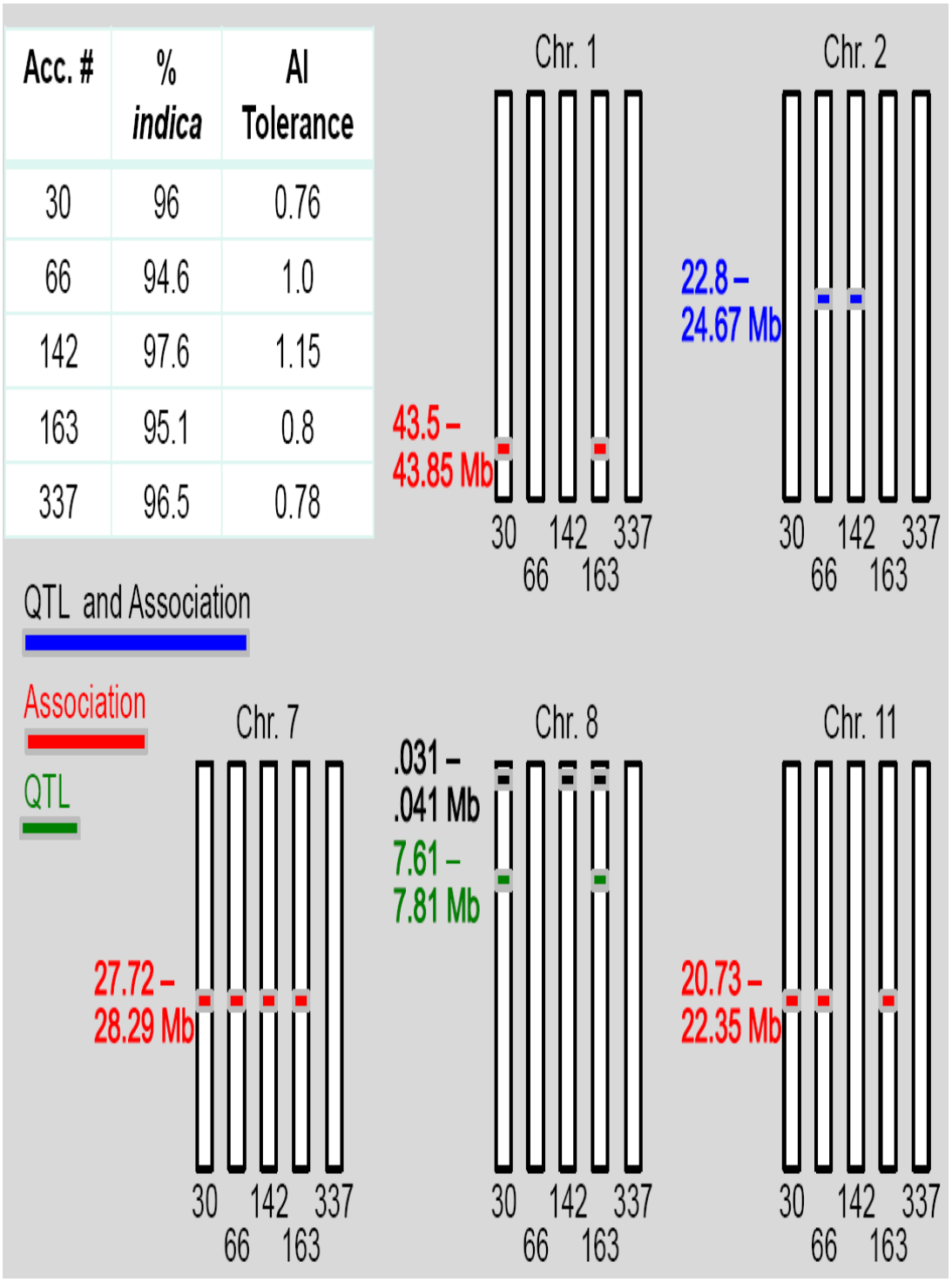


Figure 3.6. *Japonica* introgressions in Al tolerant *indica* outliers. Regions of unique *Japonica* introgressions into the highly Al tolerant *indica* outliers. The introgressions were only present in the five outlier accessions.

DISCUSSION

Utilization of GWAS and bi-parental QTL mapping

In this study, we utilized bi-parental QTL populations and GWA analysis to examine the genetic architecture of Al tolerance in rice and to identify novel Al tolerance loci and confirm previously reported loci. Phenotyping of the GWA diversity panel provided valuable information about the range and distribution of Al tolerance in *O. sativa* and offered new insights into the evolution of the trait. The range and distribution of Al tolerance in this study, evaluated at 120, 160, and 250 μM Al^{3+} activities, were similar to those reported for wheat screened at 8.75μM Al^{3+} activity (40μM Al) (Sasaki et al., 2006) or sorghum screened at 27μM Al^{3+} activity (148μM Al) (Caniato et al., 2007), supporting the conclusion that rice has superior Al tolerance compared to other cereals (Famoso et al., 2010).

Significant differences ($p<0.0001$) were observed among rice subpopulations, where 57% of the phenotypic variation was explained by subpopulation, and when subpopulations were clustered into varietal groups, mean Al tolerance in *Japonica* was twice that of *Indica*. The relative degree of Al tolerance in the five subpopulations (*temperate japonica >tropical japonica>aromatic>indica=aus*) is consistent with the level of genetic relatedness among them (Garris et al., 2005; Zhao et al., 2010) and suggests that *temperate* and *tropical japonica* germplasm contains alleles that would be useful sources of genetic variation for enhancing levels of Al tolerance within *indica* and *aus*. This is supported by the identification of highly tolerant *indica* varieties from the rice diversity panel that contain regions of admixture from *Japonica* that are also associated with significant SNPs identified by GWA analysis. This demonstrates the feasibility of using a targeted approach to increase Al tolerance in *Indica* varieties by introgressing genes from *Japonica*.

While less obvious, our QTL analysis also demonstrated the ability to increase Al tolerance in *Japonica* using targeted introgressions from *Indica*. This was demonstrated within both QTL populations by the identification of two loci in which alleles from the highly susceptible Kasalath parent conferred enhanced levels of Al tolerance in the Nipponbare genome (*temperate japonica*) and one locus where the moderately susceptible IR64 parent conferred tolerance in crosses with Azucena (*tropical japonica*) (Table 3.1). Further evidence of the value of this approach in the context of plant breeding comes from the transgressive variation observed in both QTL populations, where some RILs and BILs exceeded the Al tolerance observed in the tolerant *Japonica* parents, Azucena and Nipponbare, due to alleles derived from the susceptible *Indica* (IR64) or *Aus* (Kasalath) parents, respectively.

The significant differences in Al tolerance among varietal groups and subpopulations, and the growing evidence that different genes and/or alleles contribute to Al tolerance in the two major varietal groups, strongly suggests that *Indica* and *Japonica* were domesticated from pre-differentiated, wild *O. rufipogon* gene pools that differed in Al tolerance. Future experiments will test this hypothesis by comparing levels of Al tolerance found in wild populations of *O. rufipogon*. The inherently higher levels of Al tolerance found in the *Japonica* varietal group may help explain why *tropical japonica* varieties are so often found in the acid soils of upland environments.

The identification of highly tolerant *indica* outliers provides useful material for identifying loci that control Al tolerance differences between the *Indica* and *Japonica* varietal groups. It is notable that four of the highly tolerant *indica* accessions shared a *Japonica* introgression on chromosome 7, and this locus was also detected by GWA, both within the *indica* subpopulation (EMMA) and across all genotypes (PCA). On

the other hand, crossing a highly tolerant *Indica* genotype to a genetically similar, but susceptible, *Indica* would generate a population that segregates at a reduced number of loci, facilitating the identification of loci responsible for the different levels of Al tolerance in a common genetic background.

This study illustrates the power of using a joint approach to gene discovery based on QTL mapping, association mapping and candidate gene analysis. QTL and GWA analysis represent complementary strategies for trait dissection. Convergence of evidence from the two approaches provides support for the significance of a phenotype-genotype association. Of the 43 GWA regions identified in this study, 15 co-localized with QTLs identified using bi-parental mapping populations. However, there are cases in which QTLs discovered by bi-parental mapping are not detected by GWA, and vice-versa. One reason for this is that by choosing appropriate parents, QTL mapping can readily detect rare and/or subpopulation-specific alleles, while GWA has less power to do so. This is important in the case of Al tolerance, because of the significant evolutionary differences between the *Indica* and *Japonica* varietal groups, as discussed above. Thus, variation that is perfectly correlated with subpopulation structure is undetectable by GWA, leading to Type 2 error. The fact that these loci can be easily detected by QTL analysis using sub-population crosses is illustrated by the identification of the Al tolerance QTL $Alt_{TRG}12.1$, encompassing the *ART1* locus on chromosome 12. This large-effect QTL (LOD=7.85, R^2=0.193) was clearly detected in the RIL population but was not detected by GWA analysis. Another limitation of GWA is that it has little power to detect alleles that exist at low frequency (<5%) in the association mapping panel, but the allele frequency problem can be readily overcome by constructing QTL mapping population(s) derived from parents that contain the rare allele(s).

GWA significantly increases the range of natural variation that can be surveyed in a single experiment, and therefore increases the number of significant regions that are likely to be identified. As demonstrated in this study, two-thirds of the loci (28/43) detected by GWA were not identified in any of the nine QTL studies conducted to date (7 previous and 2 in this study). Among the loci associated with natural variation for the first time in this study are *STAR1* and *STAR2*, which do not segregate in any of the previously published QTL mapping studies. Furthermore, GWA also provides higher resolution than QTL mapping, facilitating fine-mapping and gene discovery without requiring the development of new QTL mapping populations. As demonstrated by the GWA locus identified within the *aus* subpopulation on chromosome 2 in this study, GWA narrowed the target region to 139kb containing only 27 genes, while the QTL interval was 1,360kb and contained 234 genes.

GWA mapping can generate targeted hypotheses about genotype-phenotype relationships and may also provide a critical link between natural variation in germplasm collections and functional genomics. As illustrated here for the GWA region on chromosome 2, we hypothesize that the *aus*-specific susceptibility haplotype in this region is functionally related to the *ART1* gene, previously identified only as an induced mutation with altered expression patterns in Al treated roots (Yamija et al., 2009). Similarly, we implicate the *STAR1* and *STAR2* genes, previously identified as induced mutations (Ma et al., 2002), as the genes underlying GWA loci on chromosomes 5 and 6 with susceptibility alleles coming from the *indica* subpopulation. Understanding how different subpopulations contribute to the spectrum of Al tolerance in rice will help us reconstruct the unique evolutionary trajectory by which rice has acquired superior levels of Al tolerance and provide clues as to how to enhance levels of Al tolerance in other cereals.

GWA analysis in rice requires the use of multiple statistical models in a two-step process. The mixed model (EMMA) is useful when GWA is conducted within each of the subpopulations individually, but PCA proved more effective when analyzing the rice diversity panel as a whole. When the mixed model is used to analyze the highly structured rice diversity panel, it suffers from Type 2 error, detecting only three Al tolerance loci, compared to 18 loci detected by PCA using the same genotypic and phenotypic data. The 18 regions detected by PCA include the three regions identified by the mixed model as well as 13 additional regions that were identified when the mixed model was used to analyze the subpopulations individually. Using both analysis methods iteratively to analyze the same dataset, we gained new insight into the genetic architecture of Al tolerance in rice, and identified two kinds of loci; those segregating only within a single subpopulation, and those segregating across subpopulations.

The convergence of candidate regions identified using different genetic populations and analytical techniques helps focus future research efforts aimed at identifying genes controlling Al tolerance in rice. Prior to undertaking this study, we identified 37 candidate genes from the literature and 20 (54%) of these were subsequently found in regions identified through GWA mapping using our diversity panel. Nine of the GWA regions contained both Al tolerance candidate genes and QTL identified in bi-parental mapping populations, making these the highest priority regions for further dissection. Five of these regions are worthy of note here. These include (1) a region on chromosome 3 (~34Mb) identified by GWA that co-localized with a bi-parental QTL and a candidate gene (GA-regulated protein) that had altered expression in Al in the *ART1* mutant background (Nguyen et al., 2001; Yamaji et al., 2009), (2 and 3) two cell wall-related candidate genes that co-localized with two GWA regions detected within the *indica* subpopulation, (4) an expansin precursor protein on chromosome 4

(SNPid id4010688) with altered expression in the *ART1* mutant that co-localized to a cluster of SNPs and (5) a pectinesterase gene on chromosome 10 (SNPid wd10002947) that was identified as a maize Al tolerance candidate and co-localized with a cluster of SNPs identified with *indica*. Expansins are of particular interest because they are involved in cell wall loosening and are required for cell expansion and growth, and Al has been demonstrated to be a strong inhibitor of expansin proteins in other plant species (Cosgrove, 1989), while a relationship between free pectin acid residues, pectin methyl esterase activity and Al tolerance has been reported by Yang et al. (2008).

A highly significant region on chromosome 1 was identified based on GWA analysis that co-localizes with QTL identified in both populations in this study and in four previously published rice Al tolerance QTL studies (Wu et al., 2000; Nguyen et al., 2001, 2002; Ma et al., 2002). This locus is in close proximity to the rice ortholog of the cloned sorghum Al tolerance gene, *SbMATE*. The co-localization of these QTLs with *SbMATE* leads to the hypothesis that Al tolerance genes are conserved in sorghum and rice. However, the QTLs identified on chromosome 1 in both the RIL and the BIL populations in our study do not overlap with the rice ortholog of the Alt_{SB} gene. Furthermore, Alt_{SB} confers Al tolerance in sorghum through an Al exclusion mechanism mediated by citrate exudation, and this mechanism has been demonstrated to be ineffective in conferring Al tolerance in rice (Ma et al., 2002; Famoso et al., 2010).

One of the objectives of this study was to determine if the three different Al tolerance relative root growth (RRG) indices, based on longest root growth RRG, primary root growth RRG, or total root growth RRG, influenced the detection and/or significance

of Al tolerance QTL. In our recent publication, we demonstrated that significantly different Al tolerance scores were obtained when the RRG of the total root system was calculated instead of the RRG of the longest root (Famoso et al., 2010). In all previous QTL studies, Al tolerance was determined based on relative root growth (RRG) of the longest root in treatment and control plants. However, in this study, all three QTLs detected in the BIL population were significant only when using one of the three specific Al tolerance indicies. Two QTLs were identified when using the TRG-RRG index (*Alt*$_{TRG}$*1.2* and *Alt*$_{TRG}$*12.2)*, one using the PRG-RRG index (*Alt*$_{TRG}$*12.2)*, while no QTLs were detected using the LRG-RRG index. The fact that no QTLs were detected with the LRG-RRG index in the BIL population can be largely explained by the small size of the BIL population (n=78) and the reduced statistical power due to population structure where Kasalath alleles were under-represented. On the other hand, in the RIL population, the large-effect locus identified on chromosome 9, *Alt*$_{LRG}$9.1, was only detected using the LRG-RRG index. However, three of the four QTL identified (Chr. 1, 2, and 12) were detected only when the TRG-RRG index was used, suggesting that these loci confer Al tolerance by increasing lateral and secondary root growth. The RIL locus on chromosome 2 (*Alt*$_{TRG}$*2.1*) was only detected using the TRG-RRG index; however a clear, but non-significant peak was observed using the PRG-RRG index. This suggests that this locus may function through both lateral and secondary root Al tolerance.

The strongest example of the importance of utilizing the TRG-RRG index is demonstrated by the identification of the *Alt*$_{TRG}$*12.1* QTL in the RIL mapping population. The *ART1* mutation was recently found to be in a C2H2-type zinc finger-type transcription factor, and this mutation results in Al sensitivity (Yamaji et al., 2009). Here we show that *ART1* is located close to the center of the *Alt12.1* QTL

peak. When this gene was identified, it was suggested that it was not involved in natural variation of Al tolerance in rice, as no QTL had ever been identified in the region (Yamaji et al, 2009). Based on our results, it is likely that this QTL was not previously identified because in those studies relative root growth was only measured based on growth of the longest root, rather than the total root system. If *ART1* expression differences underlie the natural variation for Al tolerance observed in rice, it would most likely be over-expressed in Al tolerant genotypes and under-expressed in Al sensitive genotypes. Further fine-mapping of this locus, along with sequence and expression analysis will be necessary to determine whether the *ART1* locus underlies this QTL and by which mechanism it contributes to natural variation for Al tolerance.

This study provides the most comprehensive analysis of the genetic architecture of Al tolerance in rice to date. It demonstrates the power of whole genome association analysis to identify phenotype-genotype relationships and to integrate disparate pieces of evidence from QTL studies, mutant analyses and candidate gene evaluation into a coherent set of hypotheses about the genes and genomic regions underlying quantitative variation. By tracing the origin of Al tolerance alleles within and between rice subpopulations, we provide new insights into the evolution and combinatorial potential of different alleles that will be invaluable in breeding new varieties for acid soil environments. Our study also lays the foundation for understanding the genetic basis of Al tolerance mechanisms that enable rice to withstand significantly higher levels of Al than do other cereals and to use this knowledge to enhance levels of Al tolerance in other plant species.

Materials and Methods

Plant Growth Conditions and Germplasm

Plants were grown hydroponically under growth chamber conditions as described by Famoso et al. (2010). Al tolerance was determined based on relative root growth (RRG) after three days of growth in Al (160µM Al^{3+}) or control solution. To obtain uniform seedlings, 80 seeds were germinated and the 30 most uniform seedlings were visually selected and transferred to a control hydroponic solution for a 24 hour adjustment. After the 24 hour adjustment period the root length was measured with a ruler and the 20 most uniform seedlings were selected and distributed to fresh control solution (10) or Al treatment solution (10). Plants were grown in their respective treatments for ~72 hours and the total root system growth was quantified using an imaging and root quantification system as described by Famoso et al. (2010). The mean total root growth was calculated for the Al treated plants and the control treated plants and RRG was calculated as mean growth (Al) / mean growth (control). The 385 genotypes screened for Al tolerance and used in the association analysis are part of a set of 400 *O. sativa* genotypes that have been genotyped with 44,000 SNPs as described by Tung et al. (2010).

QTL Analysis and Heritability

The QTL populations consisted of a population of 78 backcross introgression lines (BILs) derived from a cross between Nipponbare (tolerant *temperate japonica*) and Kasalath (susceptible *aus*) and backcrossed to Nipponbare (Lin et al., 1998) and a population of 134 recombinant inbred lines (RILs) derived from a cross between Azucena (tolerant *tropical japonica*) and IR64 (susceptible *indica*) (Ahmadi et al., 2005). To ensure a normal distribution was obtained in each mapping population the Al^{3+} concentration used to screen each population was based on the Al tolerance of the

parents of each population. The RIL population was screened at an extremely high Al^{3+} concentration (250μM Al^{3+}) because the Azucena parent is very Al tolerant and the IR64 parent is only moderately susceptible. The BIL population was screened at a relatively low Al^{3+} concentration (120μM Al^{3+}) because the Kasalath parent is extremely Al sensitive and the Nipponbare parent is very Al tolerant. The genetic component of the phenotypic variance was calculated as VarG = VarG + Var(GxE) + error. QTL analysis was conducted using the composite interval mapping (CIM) function in QTL Cartographer (Wang et al., 2010). The significance threshold was determined by 1000 permutations.

Genome Wide Association Analysis

We have performed the association using three approaches in all samples (# number count) with phenotypes. The first approach is the naïve approach, which is simply the linear regression of phenotype on the genotype for each SNP marker. The second approach is called PCA. We obtain the four main PCs (principle components) that reflect the global main 5 subpopulations in the sample to correct population structure estimated from software EIGENSOFT (Price et al., 2006). The first four PCs are included as cofactors in the regression model to correct population structure: $y = X\beta + C\gamma + e$.

Here β and γ are coefficient vectors for SNP effects and subpopulation PCs, respectively. X and C are the corresponding SNP vector and first four PC vectors, and e is the random error term. The third approach is the linear mixed model proposed by Yu et al. (2006) and Zhao et al. (2007) and implemented in R package EMMA (Kang et al. 2008). It models the different levels of population structure and relatedness. The model can be written in a matrix form as: $y = X\beta + C\gamma + Zu + e$,

where β and γ are the same as above and both are fixed effects, u is the random effect accounting for structures and relatedness, Z is the corresponding design matrices, and e is the random error term. Assume $u \sim N(0,\sigma_g^2 K)$ and $e \sim N(0,\sigma_e^2 I)$ and K is the IBS matrix, as in Zhao et al. (2007).

We also carried out both naïve approach and mixed model approach in each of subpopulations (IND, AUS, TEJ, TRJ). For the mixed model, the model is changed to $y = X\beta + Zu + e$, since there is no main subpopulation division within each subpopulation sample.

APPENDIX

Supplemental Table 3.1. Aluminum tolerance and subpopulation identity of the 385 rice genotypes used for GWAS analysis

NSF ID #	Al Tolerance (RRG)	Accession Name	Subpopulation (80% identity)	Varietal Group (80% idenity)
1	0.730	Agostano	TEJ	Japonica
2	0.784	AICHI ASAHI	N/A	Japonica
3	0.300	Ai-Chiao-Hong	IND	Indica
4	0.540	Arc 1N/A177	AUS	Indica
5	0.440	Arc 1N/A352	AROMATIC	Japonica
6	0.570	Arc 7229	AUS	Indica
7	0.860	Arias	TRJ	Japonica
8	0.664	Asse Y Pung	TRJ	Japonica
9	0.723	Baber	TEJ	Japonica
10	0.902	Baghlani Nangarhar	TEJ	Japonica
12	0.130	Basmati	AROMATIC	Japonica
13	0.390	Basmati 1	AUS	Indica
14	0.440	Basmati 217	TRJ	Japonica
15	0.829	Beonjo	TEJ	Japonica
16	0.680	Bico Branco	AROMATIC	Japonica
17	0.460	Binulawan	IND	Indica
18	0.420	Bj 1	AUS	Indica
19	0.410	Black Gora	AUS	Indica
20	0.550	Blue Rose	ADMIX	Japonica
21	0.310	N/A	IND	Indica
22	0.654	Caawa/Fortuna 6-1N/A3-15	TRJ	Japonica
23	0.460	Canella De Ferro	TRJ	Japonica
24	0.690	Carolina Gold	TRJ	Japonica
25	0.736	Carolina Gold	TRJ	Japonica
26	0.562	Carolina Gold Sel	TRJ	Japonica
27	0.490	Chahora 144	TRJ	Japonica
28	0.250	CHAMPA TONG 54	N/A	Indica
30	0.760	Chiem Chanh	IND	Indica
31	0.780	Chinese	TEJ	Japonica
32	1.050	Chodongji	TEJ	Japonica
33	0.240	Q33	AUS	Indica
35	0.451	Co18	IND	Indica
36	0.730	CS-M3	TEJ	Japonica
37	0.470	N/A	TRJ	Japonica
38	0.640	DA 5	N/A	N/A

Supplemental Table 3.1 (continued)

39	0.287	Da16	ADMIX	Indica
40	0.779	Dam	ADMIX	Japonica
43	0.320	Dee Geo Woo Gen	IND	Indica
44	0.440	Dhala Shaitta	AUS	Indica
45	0.600	Dom-Sofid	AROMATIC	Japonica
46	0.650	Dourado Agulha	TRJ	Japonica
48	0.642	DULAR	N/A	N/A
49	0.350	Dv85	AUS	Indica
50	0.120	Dz78	AUS	Indica
51	0.600	Early Wataribune	TEJ	Japonica
52	0.180	Eh Ia Chiu	TEJ	Japonica
53	0.590	Firooz	AROMATIC	Japonica
54	0.680	Fortuna	TRJ	Japonica
55	0.740	Gerdeh	ADMIX	Japonica
56	0.810	Geumobyeo	TEJ	Japonica
57	0.300	Gharib	IND	Indica
58	0.309	Ghati Kamma Nangarhar	AUS	Indica
59	0.944	Gogo Lempuk	TRJ	Japonica
60	0.600	Gotak Gatik	ADMIX	Japonica
61	0.430	Guan-Yin-Tsan	IND	Indica
62	0.617	Gyehwa 3	TEJ	Japonica
63	0.563	Haginomae Mochi	TEJ	Japonica
64	0.684	Heukgyeong	TEJ	Japonica
65	0.530	Honduras	TRJ	Japonica
66	1.000	66	IND	Indica
67	0.800	Hu Lo Tao	TEJ	Japonica
68	0.646	68	ADMIX	Admixed
69	0.690	Iac 25	TRJ	Japonica
70	0.630	Iguape Cateto	TRJ	Japonica
71	0.370	Ir 36	IND	Indica
72	0.280	Ir 8	IND	Indica
73	0.660	Irat 177	TRJ	Japonica
74	0.320	Irga 4N/A9	IND	Indica
75	0.650	Jambu	TRJ	Admixed
76	0.280	Jaya	IND	Indica
77	0.310	Jc149	IND	Indica
78	0.410	Jhona 349	AUS	Indica
79	0.815	Jouiku 393G	TEJ	Japonica
80	0.756	K 65	ADMIX	Admixed

Supplemental Table 3.1 (continued)

81	0.480	Kalamkati	AUS	Indica
82	0.440	KALUKANTHA	N/A	N/A
83	0.620	Kamenoo	TEJ	Japonica
84	0.774	Kaniranga	TRJ	Japonica
85	0.201	Kasalath	AUS	Indica
87	0.663	Keriting Tingii	ADMIX	Japonica
88	0.591	Khao Gaew	AUS	Indica
89	0.600	Khao Hawm	TRJ	Japonica
90	0.190	N/A	IND	Indica
91	0.932	Kibi	TEJ	Japonica
92	0.590	Kinastano	TRJ	Japonica
93	0.580	Kitrana 5N/A8	AROMATIC	Japonica
94	0.490	Koshihikari	TEJ	Japonica
96	0.604	Ku115	ADMIX	Japonica
97	0.448	Kun-Min-Tsieh-Hunan	IND	Indica
98	0.689	L-2N/A2	TRJ	Japonica
99	0.580	Lac 23	TRJ	Japonica
100	0.800	Lacrosse	ADMIX	Japonica
101	0.630	Lemont	TRJ	Japonica
102	0.440	1N/A2	IND	Indica
103	0.550	Luk Takhar	TEJ	Japonica
104	0.780	Mansaku	TEJ	Japonica
105	0.400	Mehr	AUS	Indica
106	0.450	Ming Hui	IND	Indica
107	1.180	Miriti	TRJ	Japonica
108	1.317	Moroberekan	TRJ	Japonica
109	0.333	Mtu9	IND	Indica
110	0.290	Mudgo	IND	Indica
111	0.510	Q32-111	TRJ	Japonica
112	0.150	N/A	AROMATIC	Japonica
113	0.640	Norin 2N/A	TEJ	Japonica
114	0.510	Nova	ADMIX	Japonica
116	0.700	Npe 844	TRJ	Japonica
117	0.320	O-Luen-Cheung	IND	Indica
118	0.530	Oro	TEJ	Japonica
119	0.300	Oryzica Llanos 5	IND	Admixed
120	0.450	Os6	TRJ	Japonica
121	0.830	Ostiglia	TEJ	Japonica
122	0.692	Padi Kasalle	TRJ	Japonica

Supplemental Table 3.1 (continued)

123	0.320	Pagaiyahan	IND	Indica
125	0.260	Pao-Tou-Hung	IND	Indica
126	0.578	Pappaku	IND	Indica
127	0.430	PATNAI 23	N/A	N/A
128	0.823	Pato De Gallinazo	ADMIX	Japonica
129	0.300	Peh-Kuh	IND	Indica
130	0.340	Peh-Kuh-Tsao-Tu	IND	Indica
131	0.160	Phudugey	AUS	Indica
132	0.350	Rathuwee	IND	Indica
133	1.000	N/A	TEJ	Japonica
134	0.480	Romeo	TEJ	Japonica
135	0.698	N/A	TRJ	Japonica
137	0.250	Rts14	IND	Indica
138	0.580	Rts4	IND	Indica
139	0.690	S4542A3-49B-2B12	TRJ	Japonica
140	0.740	Saturn	ADMIX	Japonica
141	0.250	Seratoes Hari	IND	Indica
142	1.150	Shai-Kuh	IND	Indica
143	0.810	Shinriki	TEJ	Japonica
145	0.030	Short Grain	IND	Indica
147	0.580	Sinampaga Selection	TRJ	Japonica
148	0.360	Sintane Diofor	IND	Indica
149	0.920	Sinaguing	TRJ	Japonica
150	0.530	Sultani	TRJ	Japonica
151	0.730	Suweon	TEJ	Japonica
152	0.246	T 1	AUS	Indica
153	0.260	T26	AUS	Indica
154	0.670	Ta Hung Ku	TEJ	Japonica
155	0.750	Ta Mao Tsao	TEJ	Japonica
156	0.290	Taichung Native 1	IND	Indica
157	0.540	Tainan Iku 487	TEJ	Japonica
158	0.710	Taipei 3N/A9	TEJ	Japonica
159	0.555	Tam Cau 9A	IND	Indica
160	0.500	Tchampa	AROMATIC	Japonica
161	0.320	Teqing	IND	Indica
162	0.434	Tkm6	IND	Indica
163	0.800	Taducan	IND	Indica
164	0.706	Tondok	TRJ	Japonica
165	0.700	Trembese	TRJ	Japonica

Supplemental Table 3.1 (continued)

166	0.410	Tsipala 421	ADMIX	Indica
167	0.570	N/A	TRJ	Japonica
168	0.240	Vary Vato 462	ADMIX	Indica
169	0.750	WC 6	TEJ	Japonica
170	0.910	Wells	TRJ	Japonica
171	0.280	Zhe 733	IND	Indica
172	0.240	Zhenshan 2	IND	Indica
173	0.750	Nipponbare	TEJ	Japonica
174	0.768	Azucena	TRJ	Japonica
175	0.785	1N/A21	TRJ	Japonica
176	0.779	583	TRJ	Japonica
177	0.926	68-2	TEJ	Japonica
178	0.180	Arc 6578	AUS	Indica
179	0.850	Bellardone	TEJ	Japonica
180	0.986	Benllok	TEJ	Japonica
181	0.470	Bergreis	TEJ	Japonica
182	0.788	Blue Rose Supreme	ADMIX	Japonica
183	0.713	Boa Vista	TRJ	Japonica
184	0.911	Bombon	TEJ	Japonica
185	0.768	185	TRJ	Japonica
186	1.030	Bul Zo	TEJ	Japonica
187	0.549	C57-5N/A43	TRJ	Japonica
188	0.530	Coppocina	TRJ	Japonica
189	0.440	Criollo La Fria	IND	Indica
190	0.680	Delrex	TRJ	Japonica
191	0.643	Dom Zard	AROMATIC	Japonica
192	0.750	Erythroceros Hokkaido	TEJ	Japonica
193	0.641	Fossa Av	TRJ	Japonica
195	0.704	Irat 13	TRJ	Japonica
196	0.377	Jm7N/A	IND	Indica
197	0.760	Kaukkyi Ani	ADMIX	Japonica
198	0.797	Leah	TRJ	Japonica
199	0.873	Mojito Colorado	TRJ	Japonica
200	0.441	P 737	AUS	Indica
201	0.676	Pate Blanc Mn 1	TRJ	Japonica
202	0.590	Pratao	TRJ	Japonica
203	0.270	Radin Ebos 33	IND	Indica
204	0.617	Razza 77	TEJ	Japonica
205	0.779	Rinaldo Bersani	ADMIX	Japonica

Supplemental Table 3.1 (continued)

206	0.350	Rojofotsy 738	ADMIX	Indica
207	0.469	Sigadis	IND	Indica
208	0.590	Slo 17	IND	Indica
209	0.250	Tchibanga	IND	Indica
211	0.875	Tokyo Shino Mochi	ADMIX	Japonica
212	0.765	Wc 281N/A	TRJ	Japonica
213	0.721	Wc 3397	TRJ	Japonica
214	0.650	Wc 4419	TRJ	Japonica
215	0.750	Wc 4443	TRJ	Japonica
216	0.790	N/A	TEJ	Japonica
217	0.690	Yrl-1	ADMIX	Japonica
218	0.960	Pi 298967-1	ADMIX	Japonica
219	0.809	Nucleoryza	TEJ	Japonica
221	0.680	Sadri Belyi	AROMATIC	Japonica
222	0.350	Paraiba Chines Nova	IND	Indica
223	0.614	Priano Guaira	TRJ	Japonica
224	0.620	Karabaschak	TEJ	Japonica
225	0.866	Biser 1	TEJ	Japonica
226	0.853	Irat 44	TRJ	Japonica
227	0.250	N/A	ADMIX	Indica
228	0.400	Ca 9N/A2/B/2/1	AUS	Indica
229	0.560	Niquen	TRJ	Japonica
231	0.427	N/A	IND	Indica
232	0.997	Shangyu 394	TEJ	Japonica
233	0.894	Sung Liao 2	TEJ	Japonica
234	0.310	Aijiaonante	IND	Indica
235	0.290	Sze Guen Zim	IND	Indica
236	0.713	Wc 521	ADMIX	Japonica
237	0.530	Estrela	ADMIX	Japonica
238	0.720	WAB56-1N/A4	N/A	Japonica
239	0.731	N/A	TRJ	Japonica
240	0.607	Wab 5N/A1-11-5-1	TRJ	Japonica
241	0.170	Ecia76-S89-1	IND	Indica
242	0.603	27	TRJ	Japonica
243	1.080	Tropical Rice	TEJ	Japonica
244	0.582	N/A	ADMIX	Japonica
245	0.729	Sab Ini	TEJ	Japonica
246	0.360	N/A	AUS	Indica
247	0.661	Desvauxii	TEJ	Japonica

Supplemental Table 3.1 (continued)

248	0.600	Caucasica	TEJ	Japonica
249	0.909	N/A	ADMIX	Admixed
250	0.833	Bulgare	TEJ	Japonica
251	0.530	H256-76-1-1-1	TRJ	Japonica
252	0.430	Djimoron	IND	Indica
253	0.575	N/A	ADMIX	Japonica
254	0.400	Hon Chim	IND	Indica
255	0.370	Pai Hok Glutinous	IND	Indica
256	0.510	N/A	TEJ	Japonica
257	1.026	Agusita	TEJ	Japonica
258	0.730	Tia Bura	TRJ	Japonica
259	0.300	Sadri Tor Misri	ADMIX	Indica
260	0.620	N/A	AROMATIC	Japonica
261	0.190	Shim Balte	AUS	Indica
262	0.470	Halwa Gose Red	AUS	Indica
263	0.660	Maratelli	TEJ	Japonica
264	0.552	Baldo	ADMIX	Japonica
265	0.910	Vialone	TEJ	Japonica
266	0.717	Hiderisirazu	ADMIX	Japonica
267	0.828	Hatsunishiki	TEJ	Japonica
268	0.670	Vavilovi	TEJ	Japonica
269	0.430	Sundensis	IND	Admixed
270	0.664	Osogovka	ADMIX	Japonica
271	0.798	M. Blatec	ADMIX	Japonica
272	0.517	923	ADMIX	Admixed
273	0.752	Varyla	ADMIX	Japonica
274	0.660	Padi Pagalong	TRJ	Japonica
275	0.723	Sri Malaysia Dua	TEJ	Japonica
276	0.400	Kaukau	AUS	Indica
277	0.690	Gambiaka Sebela	TEJ	Japonica
278	0.513	C1-6-5-3	ADMIX	Admixed
279	0.602	Kon Suito	TEJ	Japonica
280	0.744	Saku	ADMIX	Japonica
281	0.850	Patna	TEJ	Japonica
282	0.726	Triomphe Du Maroc	TEJ	Japonica
283	0.700	Chibica	TEJ	Japonica
284	0.390	IR-44595	IND	Indica
285	0.480	Tox 782-2N/A-1	TRJ	Japonica
286	0.575	Iita 135	TRJ	Japonica

Supplemental Table 3.1 (continued)

287	0.845	Zerawchanica Karatalski	TEJ	Japonica
288	0.695	Italica Carolina	TEJ	Japonica
289	1.110	Lusitano	TEJ	Japonica
290	1.100	Amposta	TEJ	Japonica
291	0.682	Toploea 7N/A/76	TEJ	Japonica
292	0.842	Stegaru 65	TEJ	Japonica
293	0.160	Tog 7178	ADMIX	Indica
294	0.440	SL 22-613	ADMIX	Admixed
295	0.639	Bombilla	TEJ	Japonica
296	1.000	Dosel	TEJ	Japonica
297	0.823	Bahia	TEJ	Japonica
298	0.310	Ld 24	IND	Indica
299	0.260	Sml 242	IND	Indica
300	1.066	Sml Kapuri	TEJ	Japonica
301	0.744	Melanotrix	TEJ	Japonica
302	0.662	Wir 3N/A39	TEJ	Japonica
303	0.896	Kihogo	TEJ	Japonica
304	0.140	519	IND	Indica
305	0.632	Doble Carolina Rinaldo Barsani	ADMIX	Japonica
306	0.849	Wir 3764	TEJ	Japonica
307	0.881	N/A	TEJ	Japonica
308	0.649	Llanero 5N/A1	TRJ	Japonica
309	0.716	Manzano	TRJ	Japonica
310	0.611	R 1N/A1	TRJ	Japonica
311	0.823	56-122-23	TEJ	Japonica
312	0.390	Aswina 33N/A	AUS	Indica
313	0.410	Br24	IND	Indica
314	0.180	Ctg 1516	AUS	Indica
315	0.330	Dawebyan	IND	Indica
316	0.490	Dd 62	AUS	Indica
317	0.080	Dj 123	AUS	Indica
318	0.390	Dj 24	AUS	Indica
319	0.280	N/A	AUS	Indica
320	0.630	Dm 43	AUS	Indica
321	0.440	Dm 56	AUS	Indica
322	0.540	N/A	AUS	Indica
323	0.410	Dnj 14N/A	AUS	Indica
324	0.350	Dv 123	AUS	Indica
325	0.220	Emata A 16-34	IND	Indica

Supplemental Table 3.1 (continued)

326	0.130	N/A	AUS	Indica
327	0.620	N/A	AUS	Indica
328	0.310	Jamir	AUS	Indica
329	0.340	Kachilon	AUS	Indica
330	0.300	Khao Pahk Maw	AUS	Admixed
331	0.340	Khao Tot Long 227	AUS	Indica
332	0.350	Kpf-16	ADMIX	Indica
333	0.656	Leuang Hawn	TEJ	Japonica
334	1.107	Lomello	TEJ	Japonica
335	0.550	Okshitmayin	ADMIX	Japonica
336	0.270	Paung Malaung	AUS	Indica
337	0.780	Sabharaj	IND	Indica
338	0.850	N/A	TEJ	Japonica
339	0.430	N/A	IND	Indica
340	0.350	Berenj	ADMIX	Admixed
341	0.320	Shirkati	AUS	Indica
342	0.622	N/A	TRJ	Japonica
343	0.712	Victoria F.A.	ADMIX	Japonica
344	0.350	Habiganj Boro 6	ADMIX	Admixed
345	0.140	Dz 193	AUS	Indica
346	0.370	Karkati 87	AUS	Indica
347	0.420	Creole	TRJ	Japonica
348	0.610	China 1N/A39	IND	Indica
349	0.470	Chang Ch'Sang Hsu Tao	IND	Indica
350	0.699	Ligerito	TRJ	Japonica
351	0.716	68-2	TEJ	Japonica
352	0.660	N/A	TRJ	Japonica
353	0.240	Arc 1N/A376	AUS	Indica
354	0.120	BALA	N/A	Indica
355	1.350	Asd 1	TEJ	Japonica
356	0.280	Jc 117	IND	Indica
357	0.380	9524	AUS	Indica
358	0.886	358	ADMIX	Japonica
359	0.320	Surjamkuhi	AUS	Indica
360	0.470	Ptb 3N/A	AUS	Indica
361	0.959	F.R. 13A	TEJ	Japonica
362	0.681	Jamaica 3	TRJ	Japonica
363	0.960	Edomen Scented	TEJ	Japonica
364	0.676	Rikuto Norin 21	ADMIX	Japonica

Supplemental Table 3.1 (continued)

365	0.982	N/A	TEJ	Japonica
366	0.809	Kiuki No. 46	TEJ	Japonica
367	0.749	Sanbyang-Daeme	ADMIX	Japonica
368	0.767	Deokjeokjodo	TEJ	Japonica
369	0.140	Sathi	AUS	Indica
370	0.120	Coarse	AUS	Indica
371	0.490	Santhi Sufaid	AUS	Indica
372	0.240	Sufaid	AUS	Indica
373	0.440	N/A	AROMATIC	Japonica
374	0.859	Benllok	TEJ	Japonica
375	0.799	Upland	TRJ	Japonica
376	0.749	Breviaristata	ADMIX	Japonica
377	0.751	Pr 3N/A4	TRJ	Japonica
378	0.170	Kalubala Vee	AUS	Indica
379	0.594	Wanica	TRJ	Japonica
380	0.891	Tainan-Iku No. 512	TEJ	Japonica
381	0.820	325	TRJ	Japonica
383	0.820	COLL 2712	N/A	N/A
384	0.940	318	TRJ	Japonica
385	0.390	Nira	IND	Indica
386	0.833	Palmyra	ADMIX	Japonica
387	0.924	M-2N/A2	ADMIX	Japonica
388	1.011	N/A	ADMIX	Japonica
389	0.628	Ci 11N/A11	ADMIX	Japonica
390	0.686	CI 11N/A26	ADMIX	Admixed
391	0.786	Della	TRJ	Japonica
392	0.714	Edith	TRJ	Japonica
393	0.260	LA 11N/A	N/A	Indica
394	0.688	Lady Wright Seln	TRJ	Japonica
395	0.785	Os 6 (Wc 1N/A296)	TRJ	Japonica
396	0.490	Cocodrie	TRJ	Japonica
397	0.920	N/A	TRJ	Japonica
398	0.317	9311	IND	Indica
399	0.755	Spring	TRJ	Japonica
400	0.339	Yang Dao 6	IND	Indica
644	0.412	IR_64	IND	Indica

Supplemental Table 3.2. A-priori Candidate Al Tolerance Genes

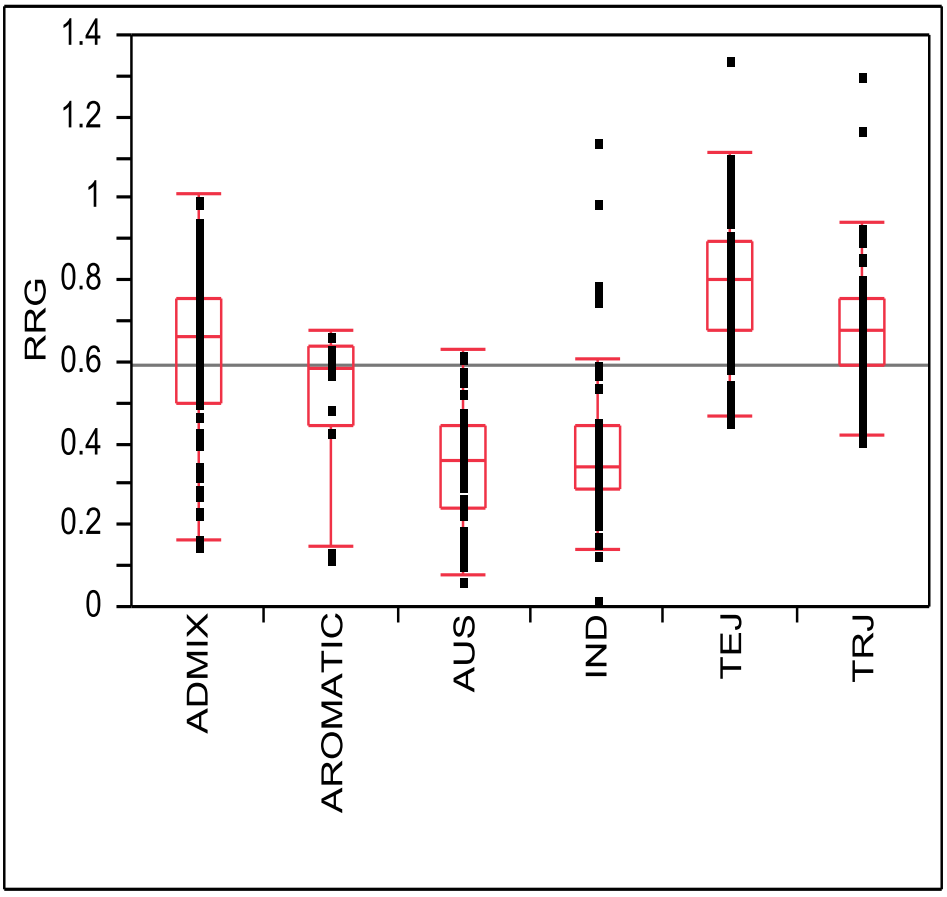
LOC	Reference	Chr	Mb	(Homolog) Description
LOC_Os01g178300	Yamaji et al., 2009	1	4.07	OSCDT3
LOC_Os01g46350	Yamaji et al., 2009	1	26.37	proteins of unknown function
LOC_Os01g53090	Yamaji et al., 2009	1	30.51	pathogen-related protein, putative
LOC_Os01g56080	Yamaji et al., 2009	1	32.28	expressed protein
LOC_Os01g64120	Yamaji et al., 2009	1	37.24	2Fe-2S iron-sulfur cluster binding domain
LOC_Os01g64890	Yamaji et al., 2009	1	37.66	CorA-like magnesium transporter protein
LOC_Os01g69010	Magalhaes et al., 2007	1	40.09	(AltSB) MATE efflux protein, putative
LOC_Os01g69020	Yamaji et al., 2009	1	40.10	retrotransposon protein, putative
NP_001044070	Yamaji et al., 2009	1	33.05	SAM-dependen methyltransferase
LOC_Os02g03900	Yamaji et al., 2009	2	1.66	metal transporter Nramp6, putative
LOC_Os02g09390	Yamaji et al., 2009	2	4.82	cytochrome P450, putative
LOC_Os02g38200	Krill et al., 2010	2	23.10	dehydrogenase, putative
LOC_Os02g51930	Yamaji et al., 2009	2	31.80	cytokinin-O-glucosyltransferase 2
LOC_Os02g53130	Yamaji et al., 2009	2	32.51	nitrate reductase, putative
LOC_Os03g11734	Krill et al., 2010	3	6.13	MATE efflux protein, putative
LOC_Os03g19170	Yamaji et al., 2009	3	10.75	GCRP7 - Glycine and cysteine rich
LOC_Os03g21950	Krill et al., 2010	3	12.54	fumarate hydratase, mitochondrial precursor
LOC_Os03g54790	Larsen et al., 2007; Yamaji et al., 2009; Huang et al., 2009	3	31.14	(ALS1) ABC transporter, ATP-binding protein
LOC_Os03g55290	Yamaji et al., 2009	3	31.46	GASR3 - Gibberellin-regulated
Os03g0760800	Yamaji et al., 2009	3	35.66	GA-regulated protein family
Os03g0126900	Yamaji et al., 2009	3	1.75	hypothetical protein
LOC_Os04g34010.1	Krill et al., 2010	4	20.42	(ALMT1) aluminum-activated malate transporter
LOC_Os04g41750	Yamaji et al., 2009	4	24.56	expressed protein
LOC_Os04g49410	Yamaji et al., 2009	4	29.30	expansin precursor, putative
LOC_Os05g02750	Larsen et al., 2005; Huang et al., 2009; Yamaji et al., 2009	5	0.99	(ALS3 and STAR2) ABC transporter
LOC_Os05g02780	Krill et al., 2010	5	1.00	glycine-rich protein A3, putative
LOC_Os05g08810	Krill et al., 2010	5	4.85	phosphatidylinositol 3-kinase, root isoform
LOC_Os05g09440	Krill et al., 2010	5	5.29	malic enzyme
LOC_Os06g36450	Krill et al., 2010	6	21.40	ferroportin1 domain containing protein
LOC_Os06g48060	Huang et al., 2009; Yamaji et al., 2009	6	29.07	(STAR1) ABC transporter, ATP-binding protein
LOC_Os07g23710	Yamaji et al., 2009	7	13.38	cytochrome P450, putative
LOC_Os07g34520	Krill et al., 2010	7	20.69	isocitrate lyase, putative
LOC_Os07g39860	Yamaji et al., 2009	7	23.90	expressed protein
LOC_Os09g25850	Yamaji et al., 2009	9	15.49	WAX2, putative; iron ion binding

Supplemental Table 3.2 (continued)

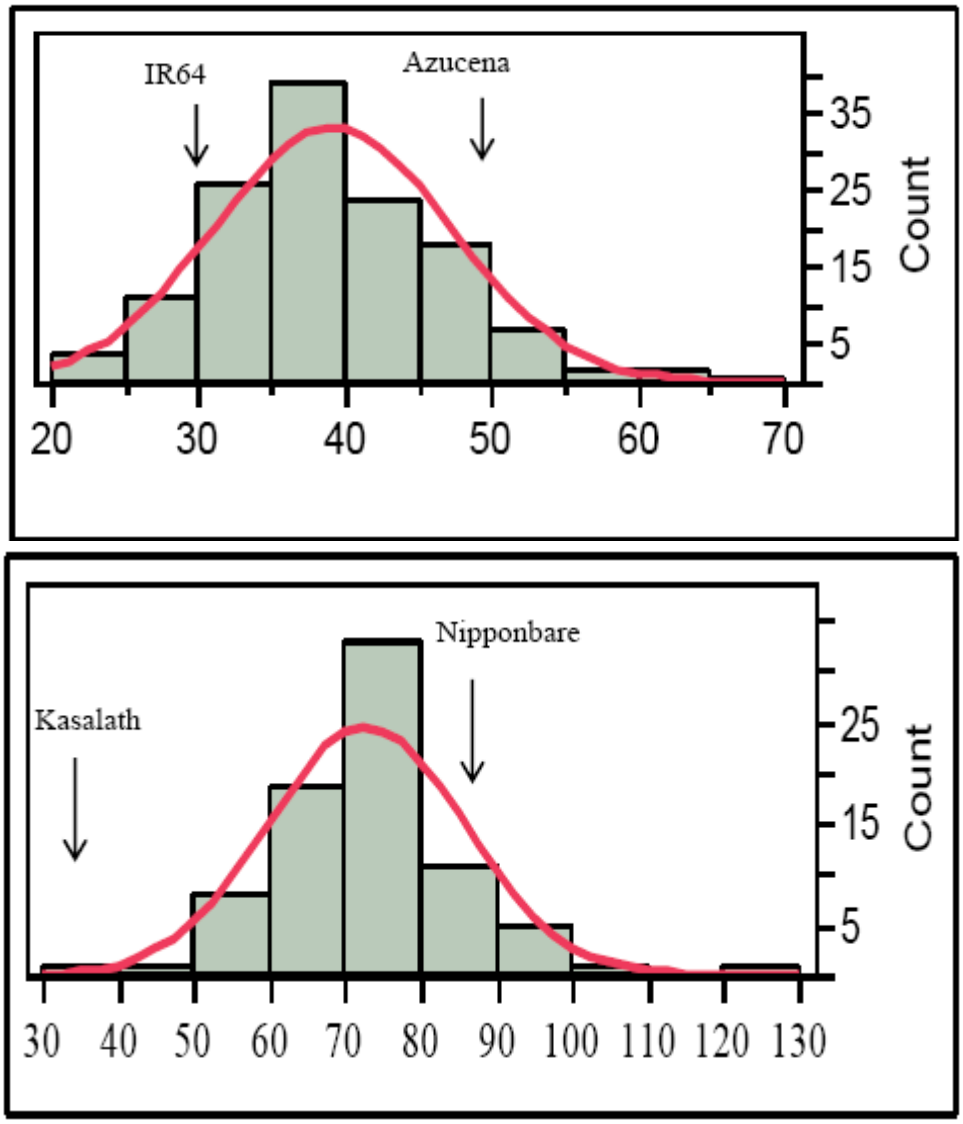
LOC_Os09g30250	Yamaji et al., 2009	9	18.41	OsSub58 - Putative Subtilisin homologue
LOC_Os10g12080	Krill et al., 2010	10	6.73	cytochrome P450, putative
LOC_Os10g13940	Yamaji et al., 2009	10	7.59	MATE efflux protein, putative
LOC_Os10g26680	Krill et al., 2010	10	13.86	pectinesterase, putative
LOC_Os10g38080	Yamaji et al., 2009	10	20.32	OsSub61 - Putative Subtilisin homologue
LOC_Os10g42780	Yamaji et al., 2009	10	23.00	lrgB-like family protein
LOC_Os11g26850	Krill et al., 2010	11	14.96	erythronate-4-phosphate dehydrogenase
LOC_Os11g29680	Yamaji et al., 2009	11	16.74	expressed protein
LOC_Os11g29780	Yamaji et al., 2009	11	16.82	plant-specific domain TIGR01627 family protein
LOC_Os12g03899	Krill et al., 2010	12	1.61	major facilitator superfamily antiporter
LOC_Os12g05860	Krill et al., 2010	12	2.69	Cupin domain containing protein
LOC_Os12g12590	Yamaji et al., 2009	12	6.93	NADP-dependent oxidoreductase

Supplemental Table 3.3. Admixture Introgressions in Al Tolerant *indica* Outliers
List of *Japonica* introgressions in the five highly Al tolerant *indica* outlier accessions, only those that are unique to the *indica* outliers were considered.

Chr.	NSF-TV ID	Left Flank (MSU6)	Right Flank (MSU6)	Introgression Size (Mb)	GWAS Peak	QTL
1	30	43.49	43.86	0.37	PCA3, IND5	no
1	163	43.49	43.86	0.37	PCA3, IND5	no
2	142	22.8	24.67	1.87	IND6	Nguyen V, 2001
2	66	22.8	23.97	1.17	IND6	Nguyen V, 2001
3	52	8.34	8.54	0.2	no	no
7	30	27.72	28.29	0.57	PCA13, IND23	no
7	66	27.72	28.29	0.57	PCA13, IND23	no
7	142	27.72	30.32	2.6	PCA13, IND23	no
7	163	27.72	30.32	2.6	PCA13, IND23	no
8	30	0.0317	0.417	0.3853	no	no
8	142	0.0317	0.417	0.3853	no	no
8	163	0.0317	0.417	0.3853	no	no
8	30	7.61	7.82	0.21	no	Nguyen V, 2002
8	163	7.61	10.14	2.53	no	Nguyen V, 2002
11	52	3.5	3.625	0.125	no	no
11	52	10.3	11.58	1.28	PCA15, IND31	no
11	52	14.52	17.97	3.45	no	no
11	163	20.73	22.35	1.62	PCA16, IND32, TEJ4	no
11	66	21.36	22.35	0.99	PCA16, IND32, TEJ4	no
11	30	21.36	22.35	0.99	PCA16, IND32, TEJ4	no



Supplemental Figure 3.1. Al tolerance distribtuion within genetic subpopulations. Distribution of Al tolerance (RRG) by subpopulation (>80% identity). Subpopulation explains 57% of phenotypic variation, however significant variation exist within each subpopulation. ADMIX represents all accessions with less than 80% identity to a specific subpopulation. IND represents the *indica* subpopulation, TEJ represents the *temperate japonica* subpopulation, and TRJ represents the *tropical japonica* subpopulation.



Supplemental Figure 3.2. Distribution of Al tolerance in the RIL and BIL QTL maping populations. Distribution of Al tolerance (TRG-RRG at 250µM Al^{3+}) observed in 134 RILs derived from Azucena (*tropical japonica*) and IR64 (*indica*) (Top). Figure 3.2 (bottom): Distribution of TRG-RRG Al tolerance at 120µM Al^{3+} observed in 78 BILs derived from Nipponbare (*temperate japonica*) and Kasalath (*aus*)

REFERENCES

Ahmadi N, Dubreuil C, Courtois B, Fonceka D, This D, McCouch S, Lorieux M, Glasmann JC, Ghesquiere A (2005) New resources and integrated maps for IR64 x Azucena: a reference population in rice. In IRRI, ed, 5th International Rice Genetics symposium. International Rice Research Institute, Manila

Atwell S, Huang YS, Vilhjalmsson BJ, Willems G, Horton M, Li Y, Meng D, Platt A, Tarone AM, Hu TT, Jiang R, Muliyati NW, Zhang X, Amer MA, Baxter I, Brachi B, Chory J, Dean C, Debieu M, de Meaux J, Ecker JR, Faure N, Kniskern JM, Jones JD, Michael T, Nemri A, Roux F, Salt DE, Tang C, Todesco M, Traw MB, Weigel D, Marjoram P, Borevitz JO, Bergelson J, Nordborg M (2010) Genome-wide association study of 107 phenotypes in *Arabidopsis* thaliana inbred lines. Nature 465: 627-631

Barbier P (1989) Genetic variation and ecotypic differentiation in the wild rice species O. rufipogon. II. Influence of the mating system and life history traits on the genetic structure of populations. Japanese Journal of Genetics 64: 273-285

Caicedo AL, Williamson SH, Hernandez RD, Boyko A, Fledel-Alon A, York TL, Polato NR, Olsen KM, Nielsen R, McCouch SR, Bustamante CD, Purugganan MD (2007) Genome-wide patterns of nucleotide polymorphism in domesticated rice. PLoS Genet 3: 1745-1756

Caniato FF, Guimaraes CT, Schaffert RE, Alves VM, Kochian LV, Borem A, Klein PE, Magalhaes JV (2007) Genetic diversity for aluminum tolerance in sorghum. TAG 114: 863-876

Cosgrove DJ (1989) Characterization of long-term extension of isolated cell walls from growing cucumber hypocotyls. Planta 177: 121-130

Daly MJ, Rioux JD, Schaffner SF, Hudson TJ, Lander ES (2001) High-resolution haplotype structure in the human genome. Nat Genet 29: 229-232

Famoso A, Clark R, Shaff J, Craft E, McCouch S, Kochian L (2010) Development of a novel aluminum tolerance phenotyping platform used for comparisons of cereal Al tolerance and investigations into rice Al tolerance mechanisms. Plant Physiology

Foy CD (1988) Plant adaptation to acid, aluminum-toxic soils. Communications in soil and plant analysis 19: 959-987

Furukawa J, Yamaji N, Wang H, Mitani N, Murata Y, Sato K, Katsuhara M, Takeda K, Ma J (2007) An Aluminum-Activated Citrate Transporter in Barley. Plant Cell Physiol 8: 1081-1091

Garris AJ, McCouch SR, Kresovich S (2003) Population structure and its effect on haplotype diversity and linkage disequilibrium surrounding the xa5 locus of rice (Oryza sativa L.). Genetics 165: 759-769

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

Glaszmann JC (1987) Isozymes and classification of Asian rice varieties. TAG 74: 21-30

Hoekenga OA, Vision TJ, Shaff JE, Monforte AJ, Lee GP, Howell SH, Kochian LV (2003) Identification and characterization of aluminum tolerance loci in *Arabidopsis* (Landsberg erecta x Columbia) by quantitative trait locus mapping. A physiologically simple but genetically complex trait. Plant Physiol 132: 936-948

Huang CF, Yamaji N, Mitani N, Yano M, Nagamura Y, Ma JF (2009) A bacterial-type ABC transporter is involved in aluminum tolerance in rice. Plant Cell 21: 655-667

Jeffreys AJ, Kauppi L, Neumann R (2001) Intensely punctate meiotic recombination in the class II region of the major histocompatibility complex. Nat Genet 29: 217-222

Khush GS (1997) Origin, dispersal, cultivation and variation of rice. Plant Mol Biol 35: 25-34

Kochian LV, Hoekenga OA, Piñeros MA (2004) How do crop plants tolerate acid soils? Mechanisms of aluminum tolerance and phosphorous efficiency. Annu Rev Plant Biol 55: 459-493

Kovach MJ, McCouch SR (2008) Leveraging natural diversity: back through the bottleneck. Curr Opin Plant Biol 11: 193-200

Krill AM, Kirst M, Kochian LV, Buckler ES, Hoekenga OA (2010) Association and linkage analysis of aluminum tolerance genes in maize. PLoS One 5: e9958

Larsen PB, Geisler MJ, Jones CA, Williams KM, Cancel JD (2005) als3encodes a phloem-localized abc transporter-like protein that is required for aluminum tolerance in *Arabidopsis*. Plant Journal 41: 353

Larsen PB, Cancel J, Rounds M, Ochoa V (2007) *Arabidopsis* ALS1 encodes a root tip and stele localized half type ABC transporter required for root growth in an aluminum toxic environment. Planta 225: 1447-1458

Lin S, Sasaki T, Yano M (1998) Mapping quantitative trait loci controlling seed dormancy and heading date using backcross inbred lines in rice, Oryza sativa L. Theoretical and Applied Genetics 96: 997-1003

Londo JP, Chiang YC, Hung KH, Chiang TY, Schaal BA (2006) Phylogeography of Asian wild rice, Oryza rufipogon, reveals multiple independent domestications of cultivated rice, Oryza sativa. Proc Natl Acad Sci U S A 103: 9578-9583

Ma J, Bennetzen J (2004) Rapid recent growth and divergence of rice nuclear genome. Proceedings of the National Academy of Sciences of the United States of America 101: 12404-12410

Ma JF, Shen R, Zhao Z, Wissuwa M, Takeuchi Y, Ebitani T, Yano M (2002) Response of rice to Al stress and identification of quantitative trait loci for Al tolerance. Plant Cell Physiol 43: 652-659

Magalhaes JV, Garvin DF, Wang Y, Sorrells ME, Klein PE, Schaffert RE, Li L, Kochian LV (2004) Comparative mapping of a major aluminum tolerance gene in sorghum and other species in the poaceae. Genetics 167: 1905-1914

Magalhaes JV, Liu J, Guimaraes CT, Lana UG, Alves VM, Wang YH, Schaffert RE, Hoekenga OA, Piñeros MA, Shaff JE, Klein PE, Carneiro NP, Coelho CM, Trick HN, Kochian LV (2007) A gene in the multidrug and toxic compound extrusion (MATE) family confers aluminum tolerance in sorghum. Nat Genet 39: 1156-1161

Maron LG, Piñeros MA, Guimaraes CT, Magalhaes JV, Pleiman JK, Mao C, Shaff J, Belicuas SN, Kochian LV (2010) Two functionally distinct members of the MATE (multi-drug and toxic compound extrusion) family of transporters potentially underlie two major aluminum tolerance QTLs in maize. Plant J 61: 728-740

Mather K, Caicedo A, Polato N, Olsen K, McCouch S, Purugganan M (2007) The Extent of Linkage Disequilibrium in Rice (Oryza sativa L.). Genetics 177: 2223-2232

Minella E, Sorrells ME (1992) Aluminum tolerance in barley: genetic relationships among genotypes of diverse origin. Crop Science 32: 593-598

Nguyen VT, Burow MD, Nguyen HT, Le BT, Le TD, Paterson AH (2001) Molecular mapping of genes conferring aluminum tolerance in rice (*Oryza sativa I.*). TAG 102: 1002-1010

Nguyen BD, Brar DS, Bui BC, Nguyen TV, Pham LN, Nguyen HT (2003) Identification and mapping of the QTL for aluminum tolerance introgressed from the new source, Oryza Rufipogon Griff., into indica rice (Oryza sativa L.). TAG 106: 583-593

Nguyen VT, Nguyen BD, Sarkarung S, Martinez C, Paterson AH, Nguyen HT (2002) Mapping of genes controlling aluminum tolerance in rice: comparison of different genetic backgrounds. Mol Genet Genomics 267: 772-780

Ninamango-Cardenas FE, Guimaraes CT, Martins PR, Parentoni SN, Carneiro NP, Lopes MA, Moro JR, Paiva E (2003) Mapping QTLs for aluminum tolerance in maize. Euphytica 130: 223-232

Nordborg M, Borevitz JO, Bergelson J, Berry CC, Chory J, Hagenblad J, Kreitman M, Maloof JN, Noyes T, Oefner PJ, Stahl EA, Weigel D (2002) The extent of linkage disequilibrium in *Arabidopsis* thaliana. Nat Genet 30: 190-193

Oka HI (1988) Origin of cultivated rice. Japan Scientific Societies Press ;$aAmsterdam; Elsevier, Tokyo; New York

Olsen KM, Caicedo AL, Polato N, McClung A, McCouch S, Purugganan MD (2006) Selection under domestication: evidence for a sweep in the rice waxy genomic region. Genetics 173: 975-983

Piñeros MA, Shaff JE, Manslank HS, Alves VM, Kochian LV (2005) Aluminum resistance in maize cannot be solely explained by root organic acid exudation. A comparative physiological study. Plant Physiol 137: 231-241

Platt A, Horton M, Huang YS, Li Y, Anastasio AE, Mulyati NW, Agren J, Bossdorf O, Byers D, Donohue K, Dunning M, Holub EB, Hudson A, Le Corre V, Loudet O, Roux F, Warthmann N, Weigel D, Rivero L, Scholl R, Nordborg M, Bergelson J, Borevitz JO (2010) The scale of population structure in *Arabidopsis* thaliana. PLoS Genet 6: e1000843

Rakshit S, Rakshit A, Matsumura H, Takahashi Y, Hasegawa Y, Ito A, Ishii T, Miyashita NT, Terauchi R (2007) Large-scale DNA polymorphism study of Oryza sativa and O. rufipogon reveals the origin and divergence of Asian rice. TAG 114: 731-743

Reich D, Cargill M, Bolk S, Ireland J, Sabeti P, Richter D, Lavery T, Kouyoumjian R, Farhadian S, Ward R, Lander E (2001) Linkage disequilibrium in the human genome. Letter to Nature 411: 199-204

Remington DL, Thornsberry JM, Matsuoka Y, Wilson LM, Whitt SR, Doebley J, Kresovich S, Goodman MM, Buckler ES,4th (2001) Structure of linkage disequilibrium and phenotypic associations in the maize genome. Proc Natl Acad Sci U S A 98: 11479-11484

Sasaki T, Ryan PR, Delhaize E, Hebb DM, Ogihara Y, Kawaura K, Noda K, Kojima T, Toyoda A, Matsumoto H, Yamamoto Y (2006) Sequence upstream of the wheat (Triticum aestivum L.) ALMT1 gene and its relationship to aluminum resistance. Plant Cell Physiol 47: 1343-1354

Sasaki T, Yamamoto Y, Ezaki B, Katsuhara M, Ahn SJ, Ryan PR, Delhaize E, Matsumoto H (2004) A wheat gene encoding an aluminum-activated malate transporter. Plant J 37: 645-653

Sweeney MT, Thomson MJ, Cho YG, Park YJ, Williamson SH, Bustamante CD, McCouch SR (2007) Global dissemination of a single mutation conferring white pericarp in rice. PLoS Genet 3: e133

Tenaillon M, Sawkins M, Long A, Gaut R, Doebley J, Gaut B (2001) Patterns of DNA sequence polymorphism along chromosome 1 of maize (Zea mays ssp. mays L.). 98: 9161-9166

Vitte C, Ishii T, Lamy F, Brar D, Panaud O (2004) Genomic paleontology provides evidence for two distinct origins of Asian rice (Oryza sativa L.). Mol Genet Genomics 272: 504-511

von Uexkull HR, Mutert E (1995) Global extent, development and economic impact of acid soils. In RA Date, NJ Grundon, GE Raymet, ME Probert, eds, Plant-Soil Interactions at Low pH: Principles and Management. Kluwer Academic, Dordrecht, Netherlands, pp 5-19

Wu P, Liao CY, Hu B, Yi KK, Jin WZ, Ni J, He C (2000) QTLs and epistatis for aluminum tolerance in rice (*Oryza sativa* L.) at different seedling stages. Theoretical and Applied Genetics 100: 1303

Xue Y, Wan J, Jiang L, Wang C, Liu L, Zhang YM, Zhai H (2006) Identification of quantitative trait loci associated with aluminum tolerance in rice. Euphytica 150: 37-45

Xue Y, Jiang L, Su N, Wang JK, Deng P, Ma JF, Zhai HQ, Wan JM (2007) The genetic basic and fine-mapping of a stable quantitative-trait loci for aluminium tolerance in rice. Planta 227: 255-262

Yamaji N, Huang CF, Nagao S, Yano M, Sato Y, Nagamura Y, Ma JF (2009) A Zinc Finger Transcription Factor ART1 Regulates Multiple Genes Implicated in Aluminum Tolerance in Rice. Plant Cell

Yang JL, Li YY, Zhang YJ, Zhang SS, Wu YR, Wu P, Zheng SJ (2008) Cell wall polysaccharides are specifically involved in the exclusion of aluminum from the rice root apex. Plant Physiol 146: 602-611

Yu J, Pressoir G, Briggs WH, Vroh Bi I, Yamasaki M, Doebley JF, McMullen MD, Gaut BS, Nielsen DM, Holland JB, Kresovich S, Buckler ES (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat Genet 38: 203-208

Zhao K, Aranzana MJ, Kim S, Lister C, Shindo C, Tang C, Toomajian C, Zheng H, Dean C, Marjoram P, Nordborg M (2007) An *Arabidopsis* example of association mapping in structured samples. PLoS Genet 3: e4

Zhou H, Xie Z, Ge S (2003) Microsatellite analysis of genetic diversity and population genetic structure of a wild rice (Oryza rufipogon Griff.) in China. Theoretical and Applied Genetics 107: 332-339

Zhu Q, Ge S (2005) Phylogenetic relationships among A-genome species of the genus Oryza revealed by intron sequences of four nuclear genes. New Phytologist 167: 249-265

CHAPTER 4:
OUTLINE OF FUTURE RESEARCH OBJECTIVES AND EXPERIMENTS TO IDENTIFY THE GENETIC AND PHYSIOLOGICAL MECHANISMS OF RICE ALUMINUM TOLERANCE

The long-term goal of this research is to enable the development of crops with superior performance in acidic soils based on a comprehensive understanding of the genetic and physiological mechanisms underlying plant response to Al stress. To achieve the full genetic potential in breeding Al^{3+} resistant crops, it is critical that we understand the genetic and physiological mechanisms underlying this trait. As a whole, cereal crops (Poaceae) provide one of the best models for studying Al^{3+} resistance. The extremely high levels of Al tolerance and the abundant genetic and genomic resources of rice are two critical components necessary to effectively identify the genetic and physiological basis for Al tolerance and apply this knowledge to improving crop productivity on acid soils. To fully understand the mechanisms of Al toxicity and tolerance within and across species, it will be necessary to take complimentary approaches of genetic and physiological approaches. The genetics approach will enable the identification of the genes conferring resistance within rice and physiological experiments will provide an understanding of how these genes function to confer tolerance. To determine the usefulness of the genes/loci that are identified it will also be necessary to identify the appropriate genetic background and to evaluate their effect under field conditions.

IDENTIFICATION OF RICE Al TOLERANCE GENES

Identifying the genes/alleles underlying the loci that confer tolerance within rice will be accomplished through fine-mapping of QTL, sequence and association analysis, expression analysis, and/or mutant analysis. The three strongest candidate regions for dissection are: 1) the major QTL *Alt*$_{TRG}$*12.1* identified in the IR64 x Azucena RIL population; 2) the region on chromosome 7 identified through GWA analysis and confirmed in the highly tolerant *indica* outlier accessions; 3) the strong GWA peak identified within the *aus* subpopulation that co-localizes with a bi-parental QTL and encompasses the Nramp gene identified to have altered expression in the *ART1* Al sensitive mutant (Yamaji et al., 2009).

For subsequent work on the major QTL (*Alt*$_{TRG}$*12.1*) identified in the IR64 x Azucena RIL population, this QTL will be fine-mapped using an advanced backcross population. To generate this population, an Al tolerant RIL (SSD48) with the Azucena allele at *Alt*$_{TRG}$*12.1* was crossed to IR64 and F2 plants were genotyped across the region to identify recombinants and heterozygotes across the target region. 4400 F3 plants, derived from informative and/or heterozygous F2 individuals, were genotyped with InDel and KASPar markers and >450 recombinant genotypes were identified. The recombinant plants were further divided into 10 recombinant classes. The F4 families derived from the recombinant plants will be phenotyped for Al tolerance and the phenotype will be assigned to the F3 genotype (recombinant class). This mapping is expected to delimit the QTL region to 100-500kb, at which point candidate genes will be identified and analyzed. The fine-mapping of this QTL will confirm whether or not the *ART1* mutant locus confers natural variation for Al tolerance in rice.

An Al tolerance region on chromosome 7 was identified by GWA analysis across all 385 genotypes and within the *indica* subpopulation. The importance of this region was confirmed by introgression analysis of the five highly tolerant *indica* outlier genotypes. Four of the *indica* outliers shared *Japonica* introgressions at this region between 27.72-28.29Mb on chromosome 7. These lines are a unique resource for the identification of genes/loci that can increase the Al tolerance of *indica* lines to the level observed in *Japonica*. The genetic background of the five outliers contains an average of 96% *Indica* ancestry and the most tolerant line (NSF 142) contains 97.6%, narrowing the portion of the genome likely controlling the high Al tolerance to 2.4% *Japonica* ancestry. To exploit this material to identify the gene(s) conferring the significantly increased tolerance, the genotype with the highest Al tolerance and lowest portion of *Japonica* ancestry, NSF 142, will be crossed to the most genetically similar *indica* accession with low Al tolerance. By crossing to a genetically similar genotype the proportion of the genome that segregates will be minimized, potentially allowing for the locus conferring tolerance to segregate as a single gene and conceptually functioning as a NIL. If high Al tolerance segregates with a 3:1 ratio in the F2 generation, it will confirm that only a single locus is segregating and the gene can be mapped as described above. The identification of the genes(s)/alleles from *Japonica* that confer significantly higher Al tolerance in *Indica* will be useful in breeding Al tolerant *Indica* varieties, which are currently not grown on acid soils.

A highly significant cluster of SNPs was identified by GWA within the *aus* subpopulation on chromosome 2 (1.536 – 1.675 Mb), providing higher resolution to a QTL detected at this region in which *aus* was the susceptible parent (Ma et al., 2002). This region also co-localizes with an Nramp metal transporter gene (1.66Mb) which has been demonstrated to have reduced expression in the *ART1* (transcription factor)

Al sensitive mutant, providing evidence that this gene confers tolerance at this locus. Furthermore, a unique haplotype was identified at this region in the highly susceptible *aus* accessions, providing information on which accessions can be used to dissect this locus. To confirm the importance of this gene, the gene sequence will be compared among a panel of diverse germplasm with varying degrees of Al tolerance, including a panel of *aus* genotypes with the various haplotypes at the region. As it was reported that Nramp expression was reduced in the *ART1* mutant and since our RIL QTL results identified a tolerance locus in the same region as the location of *ART1* in Azucena, expression analysis will be conducted in NILs with the Azucena allele at *ART1* in the IR64 background and compared with *ART1* expression in IR64. Nramp expression will also be investigated in a panel of germplasm similar to that in which the sequence analysis is conducted.

DEVELOPMENT AND FIELD EVALUATION OF Al NILs

To determine whether the Al tolerance loci identified in this thesis based on hydroponic screening confer tolerance under Al-toxic acid soil conditions it will be necessary to conduct controlled field evaluations on rice genotypes with a common genetic background. To control for the genetic background of the genotypes evaluated, reciprocal NILs will be developed at each of the QTL loci identified in the RIL population (*Alt*$_{TRG}$*1.1, Alt*$_{TRG}$*2.1, Alt*$_{LRG}$*9.1, Alt*$_{TRG}$*12.1*). Thus, an IR64 introgression will be backcrossed into the Azucena background as well as an Azucena introgression made into the IR64 background. The NILs and parents will be evaluated on Al toxic acid soils and on acid soils that have been treated with lime to raise the pH and thus ameliorate Al toxicity. The limed soils will allow a direct comparison of plant performance under similar environmental (temperature, humidity, disease/insect pressure) and soil conditions. These evaluations will determine which loci would be

most useful to employ in breeding programs and will determine how well Al tolerance in hydroponics correlates to Al tolerance under field conditions. The NILs will also provide materials to investigate the physiological mechanism by which tolerance is conferred by each locus.

INVESTIGATIONS INTO MECHANISMS OF Al TOXICITY AND TOLERANCE

Although extensive research has been conducted to understand the Al tolerance mechanisms of plants, little is known about the mechanisms by which Al is toxic to plants. To most efficiently identify novel Al tolerance mechanisms and/or genes, it will be helpful to identify the mechanisms by which Al inhibits plant growth and determine whether there are different toxicity mechanisms expressed in rice compared with less tolerant cereals species. Our preliminary physiological data demonstrates that the major Al tolerance mechanism utilized by rice is novel and is not facilitated by the known mechanism of Al exclusion from the root tip, which is mediated by organic acid exudation. The significantly higher levels of Al tolerance exhibited by rice, compared to maize, sorghum, and wheat, make it an interesting model to investigate novel Al tolerance mechanisms. Understanding the physiological mechanisms conferring rice Al tolerance will provide: 1) a better understanding of abiotic stress tolerance physiology, 2) valuable information for identifying novel Al tolerance genes through the selection of candidate genes, 3) insights into the mechanisms of Al toxicity in plants. Future experiments should aim to answer the following questions:

1) Do knockouts of rice Alt_{SB} homologs exhibit reduced Al tolerance?
2) Is Al toxicity due to a toxic effect of Al on proteins involved in cell wall proteins involved in cellular expansion/elongation?
3) Does Al cause modifications to the carbohydrate components of cell walls that inhibit root growth?

4) Do Al tolerant genotypes and/or species have unique modifications in cell wall carbohydrate components and/or structures that facilitate tolerance?

Mutant analysis of the role of rice Alt_{SB} homologs in rice Al tolerance

The work in this thesis has demonstrated that organic acid exudation and/or Al exclusion from the root tip does not account for differences in Al tolerance across 23 genetically diverse rice accessions. However, it is interesting that the MATE homolog for the sorghum Al tolerance gene (*SbMATE*) is in a region on chromosome 1 where Al tolerance QTL have been identified in multiple QTL studies. In addition, two other genes in the MATE family were identified that are in the vicinity of other previously identified rice Al tolerance QTL.

The multidrug and toxic compound extrusion (MATE) family is widespread in plants, mammals, bacteria, and fungi and members of the family share ~40% sequence similarity (Omote, 2007). There are >40 MATE genes in the rice genome, so the co-occurrence of MATE genes to Al tolerance QTL is not surprising and it may simply be a coincidence. To investigate whether MATE genes are involved in rice Al tolerance publicly available T-DNA knockout mutants will be screened to determine if they affect Al tolerance. Three T-DNA MATE mutants that coincide with Al tolerance QTL on chromosome 1, 3, and 9 (B09843-chr.1, B07031-chr.3, and A04767-chr.9) are available from the Postech Plant Functional Genomics Laboratory (www.postech.ac.kr/life/pfg). Mutant and wild-type plants can be compared for Al tolerance to determine if loss of function of these genes leads to a significant decrease in Al tolerance. If differences in Al tolerance are observed, it will be important to determine whether changes in Al tolerance are related to Al exclusion and/or organic

acid exudation by conducting experiments on mutants and wild type plants to quantify Al accumulation in the root tip and organic acid exudation as described in Chapter 2.

Investigation into Al toxicity and novel tolerance mechanisms

Although numerous mechanisms have been hypothesized, the actual mechanisms of Al toxicity in plants are still very poorly understood (Kochian, 2004). The majority of Al in a root resides in the cell wall, thus the cell wall has been frequently implicated in Al toxicity and tolerance in plants (Yang, 2008; Eticha, 2005; Mimmo, 2009). However, the role of the root cell wall in Al toxicity/tolerance has yet to be thoroughly investigated in multiple genotypes or across species. One possibility is that Al inhibition of root growth is due to Al-induced inhibition or damage to cell wall proteins involved in cell growth and/or cell expansion. This is consistent with the rapid inhibition (within minutes) of root growth observed when treated with Al (Jones and Kochian, 1995; Barcelo, 2002). As a trivalent cation, Al^{3+} is highly reactive and could theoretically bind to an enzyme active site and/or cause a conformational change to protein structure and inhibit a proteins function.

Expansins are one possible family of proteins whose inhibition by Al could explain the primary symptoms of Al toxicity associated with rapid inhibition of root growth. Expansins are the main regulators of wall extension during growth. They are encoded by two gene subfamililes, α- and β-expansins (Li, 2003; Cosgrove, 2005). The most compelling observation suggesting a role of expansins in Al toxicity/tolerance is that aluminum ions have been shown to be the most potent inhibitor of expansin activity (McQueen-Mason, 1992; Cosgrove, 1989). In addition, a β-expansin mRNA from barley roots was identified as being up-regulated by Al exposure, and when this

protein was expressed in yeast it conferred Al tolerance to the yeast, suggesting that Al may inhibit root growth by poisoning expansins (Cosgrove, 2000).

Experiments have been designed to determine if proteins involved in cell division/expansion are involved in differences in Al tolerance within and across cereal species. This research proposed here will focus on determining if proteins involved in cell growth/expansion are inhibited by Al and whether proteins from Al tolerant genotypes and/or species are capable of rescuing cell growth/expansion in Al sensitive genotypes and/or species. These experiments will provide evidence as to whether Al toxicity is due to Al interference with cell wall protein function.

In addition to the possibility that Al toxicity is due to toxic effects on cell division/expansion proteins, differences in cell wall properties/structure and/or cell wall responses to Al can also be involved in Al tolerance variation. To thoroughly investigate the role of the cell wall in Al toxicity and/or tolerance, we have developed a formal collaboration with Dr. Will York, Associate Professor, Complex Carbohydrate Research Center (University of Georgia). These experiments will be carried out on a small panel of genotypes representing tolerant and susceptible genotypes of rice, maize, and wheat. This approach will compare a very Al tolerant cereal (rice), an intermediately tolerant cereal (maize) and a fairly Al sensitive cereal species (wheat). Also, within rice, we will compare tolerant and sensitive genotypes and a NIL pair that are genetically similar but differ in Al tolerance. We will investigate whether Al tolerant genotypes and/or species have different cell wall compositions and/or if Al treatment leads to changes in cell wall structure or composition. If any of the analyses described here identify cell wall differences that explain why rice's basal Al tolerance is much higher than other cereals, or explain

within-species Al tolerance variation, these differences will be studied in a genetically diverse panel of 10 rice genotypes and 10 maize genotypes to determine whether they are predictive for Al tolerance and provides evidence for a novel mechanism of Al tolerance in rice.

Most of the research on the physiology of Al tolerance has focused on the root tip Al exclusion mechanism. Since rice is highly Al tolerant and does not utilize this mechanism, it is an excellent model to identify novel Al tolerance mechanisms. Based on previously published research, the cell wall is a likely target for Al rhizotoxicity. The physiological experiments proposed here are hypothesis driven and will provide useful insights as to whether the root cell wall is involved in Al toxicity and if differences in cell wall proteins or carbohydrate composition are involved in differences in Al tolerance. Understanding the physiological mechanisms of Al tolerance will be invaluable for developing Al tolerant crops and increasing food security.

REFERENCES

Barcelo J PC (2002) Fast root growth responses, root exudates, and internal detoxification as clues to the mechanisms of aluminium toxicity and resistance: a review. Env and Exp Botany 48: 75-92

Cosgrove DJ (2005) Growth of the plant cell wall. Nat Rev Mol Cell Biol 6: 850-861

Cosgrove DJ (2000) New genes and new biological roles for expansins. Curr Opin Plant Biol 3: 73-78

Cosgrove DJ (1989) Characterization of long-term extension of isolated cell walls from growing cucumber hypocotyls. Planta 177: 121-130

Eticha D, Stass A, Horst WJ (2005) Cell-wall pectin and its degree of methylation in the maize root-apex: significance for genotypic differences in aluminum resistance. Plant, Cell & Environment 28: 1410-1420

Jones DL, Kochian LV (1995) Aluminum Inhibition of the Inositol 1,4,5-Trisphosphate Signal Transduction Pathway in Wheat Roots: A Role in Aluminum Toxicity? Plant Cell 7: 1913-1922

Kochian LV, Hoekenga OA, Piñeros MA (2004) How do crop plants tolerate acid soils? Mechanisms of aluminum tolerance and phosphorous efficiency. Annu Rev Plant Biol 55: 459-493

Li Y, Jones L, McQueen-Mason S (2003) Expansins and cell growth. Curr Opin Plant Biol 6: 603-610

Ma JF, Shen R, Zhao Z, Wissuwa M, Takeuchi Y, Ebitani T, Yano M (2002) Response of rice to Al stress and identification of quantitative trait loci for Al tolerance. Plant Cell Physiol 43: 652-659

McQueen-Mason S, Durachko DM, Cosgrove DJ (1992) Two endogenous proteins that induce cell wall extension in plants. Plant Cell 4: 1425-1433

Mimmo T, Marzadori C, Gessa CE (2009) Does the degree of pectin esterification influence aluminum absorption by the root apoplast? Plant Soil 314: 159-168

Omote H, Hiasa M, Matsumoto T, Otsuka M, Moriyama Y (2006) The MATE proteins as fundamental transporters of metabolic and xenobiotic organic cations. Trends Pharmacol Sci 27: 587-593

Yang JL, Li YY, Zhang YJ, Zhang SS, Wu YR, Wu P, Zheng SJ (2008) Cell wall polysaccharides are specifically involved in the exclusion of aluminum from the rice root apex. Plant Physiol 146: 602-611