ANNUAL REPORT
COMPREHENSIVE RESEARCH ON RICE
January 1, 2014 – December 31, 2014

PROJECT TITLE: Application of Molecular Marker-Assisted Selection to Rice Improvement

PROJECT LEADER: Thomas H. Tai, Research Geneticist, USDA-ARS
Plant Sciences, UCD

PRINCIPAL UC INVESTIGATORS:

Thomas H. Tai, Research Geneticist, USDA-ARS, Plant Sciences, UCD

COOPERATORS:

Kent S. McKenzie, Director, Rice Experiment Station, Biggs
Virgilio C. Andaya, Plant Breeder, Rice Experiment Station, Biggs
Farman Jodari, Plant Breeder, Rice Experiment Station, Biggs
Stanley O. Samonte, Plant Breeder, Rice Experiment Station, Biggs
Sang-Ic Kim, Associate Specialist, Plant Sciences, UCD
Luca Comai, Professor, UC Davis Genome Center, UCD

LEVEL OF 2014 FUNDING: $45,000

OBJECTIVES AND EXPERIMENTS CONDUCTED, BY LOCATION, TO ACCOMPLISH OBJECTIVES:

The overall objective is to integrate molecular genetic approaches and conventional breeding methods to develop improved germplasm for the California rice industry. Primary emphasis is on the development of DNA markers that can be used to predict the presence or absence of traits of interest (e.g. disease resistance, cold tolerance, grain quality) and the application of these markers via molecular marker-assisted selection to expedite the identification of useful germplasm and streamline the breeding of improved varieties.

In order to employ DNA markers, marker-trait associations must be established (i.e. the value of a marker in predicting a trait must be determined). Basic genetic studies have resulted in the identification of markers for many important traits. Several of these markers are based on differences in the DNA of specific genes. Differences (or polymorphisms) that are directly responsible for the characteristic in question are sometimes referred to as perfect markers as they are always (perfectly) associated with the trait. Genes underlying important traits in rice such as grain quality, yield, grain size, fertility, etc. have been identified. The major objectives of our research in 2014 were to 1) continue to employ next-generation DNA sequencing-based methods to identify markers and generate highly detailed genetic fingerprints of rice germplasm and to evaluate rice genes present in California varieties; and 2) continue the development and evaluation of genetic mapping and mutant populations of rice to facilitate trait and gene

discovery. Work on this project was conducted in the USDA-ARS rice genetics lab, greenhouses, and other research facilities at UC Davis. In addition, mapping population development and evaluation work was conducted in the nursery and a cold screening greenhouse at the Rice Experiment Station.

Specific 2014 objectives included:

1) **Next-generation genotyping (NGG) of rice germplasm:** We will continue to employ next-generation sequencing-based approaches to identify markers and generate highly detailed genetic fingerprints of rice germplasm including varieties, breeding lines, mapping populations, and mutants. Emphasis will be on two genetic mapping populations (M2036 and RS629 recombinant inbred lines) and additional germplasm from the RES.
2) **Exome sequence analysis of California varieties and breeding germplasm:** We will employ exome sequencing to characterize the variation present in the genes of varieties developed for production in the California environment. Emphasis will be on identifying sequence variation in genes controlling yield components and stress tolerance.
3) **Development and evaluation of rice populations for genetic analysis of agriculturally important traits:** We will continue to develop genetic mapping and mutant populations for genetic studies of important traits including but not limited to improved yield (total and milling) and stress tolerance. Emphasis will be on advancing the RS629 mapping population, trait evaluation of the M2036 mapping population, and M-204 mutant population development and screening.

SUMMARY OF 2014 RESEARCH (MAJOR ACCOMPLISHMENTS) BY OBJECTIVES:

1) **Next-generation genotyping (NGG) of rice germplasm:** In 2014, a revised protocol for restriction enzyme site comparative analysis (RESCAN) genotyping, developed by our cooperator Dr. Luca Comai, was tested in cooperation with Dr. Areum Chun, a visiting scientist from the Rural Development Administration (RDA) of the Republic of Korea (South Korea), as part of a RDA-funded project. This revised protocol is more efficient and reliable in generating DNA sequencing libraries than the original RESCAN method. This improved protocol will be employed in 2015 to genotype the M2036 recombinant inbred mapping population (derived from a cross between M-203 and M-206 varieties) and the RS629 recombinant inbred mapping population (derived from a cross between Calmochi-101 and the long grain breeding line 94Y561).
2) **Exome sequence analysis of California varieties and breeding germplasm:** In 2014, we employed a custom exome capture reagent designed by Dr. Luca Comai's group and Nimblegen (Roche Diagnostics Corp.) in two experiments. Unfortunately, neither experiment was successful as there was no apparent enrichment (i.e. selectivity) for the coding sequence (i.e. exome) that the capture reagent was designed to target. Exome sequencing cannot be performed without sufficient enrichment of the coding sequences. The reason(s) for the failure of the capture (enrichment) process is unknown at this time and will be investigated. In addition, a different exome capture reagent will be developed with the company Mycroarray and this will be tested in 2015. Analysis of exome

sequence data from several rice accessions (Caloro, Colusa, Lady Wright, Dixiebelle, Cypress, M-204, M-205, and M-206) is ongoing and is expected to be completed in the first quarter of 2015.

3) **Development and evaluation of rice populations for genetic analysis of agriculturally important traits:** We continued our work on the development of mapping and mutant populations by advancing lines via single seed descent. Seed treatment policies established by UC Davis remain in place and continue to affect the speed and efficiency of this work. In addition population development, evaluation of a subset of the M-2036 mapping population for seedling vigor was performed and preliminary evaluation of seeds of the RS629 populations for morphology, pubescence, and waxy endosperm was conducted.

Mapping Populations: Germination of the M2036 population remained high and no generation advance (seed increase) was performed for this population. Emphasis was placed on evaluating the M2036 population for various traits of interest under field and controlled environment conditions. Generation advance was conducted on the RS629 populations (Table 1) and some preliminary evaluation of seeds was performed.

M2036 Population – In 2014, two objectives were pursued: 1) evaluation of milling quality-related traits and seedling vigor; and 2) DNA maker genotyping of the mapping lines using SNP markers generated from sequencing data obtained from the parental lines M-203 and M-206.

Field evaluation. A subset of the M2036 population (98 $F_{6:7}$ recombinant inbred lines; 3 replications) was drill-seeded (five 4' rows per line) in the stem rot evaluation nursery at the RES in 2014. Unfortunately, due to unusual weather and possible problems with the fertilizer application, the development and heading of the lines was very heterogeneous (even in some cases among adjacent rows of the same line) although these lines were evaluated in 2013 and selected based on expectations that they would head at about the same time. Due to these issues, it was not practical to collect data on grain moisture or to harvest seed for milling evaluation. If suitable space is available, field evaluation will be attempted in 2015.

Growth chamber/greenhouse-based evaluation. As reported in 2013, the M2036 mapping population lines show segregation for seedling vigor as the M-203 parent exhibits more vigorous germination and seedling growth under direct seeding in soil than M-206. In 2014, several experiments were performed to confirm the difference in seedling vigor and to evaluate the growth chamber environments. Due to the size of the population, multiple growth chambers are needed to conduct replicated trials and the consistency of the evaluations between experiments (and growth chambers; i.e. environments) is important for reliable results and analysis. After confirmation and growth chamber assessment, a total of 180 of the M2036 mapping lines were subjected to seedling vigor evaluation by assessing germination, seedling height (7 and 14 days after planting) and seedling fresh weight (14 days after planting) and examining growth rate (difference in seedling height between day 14 and day 7). To accommodate this set of lines, 90 lines were planted in a randomized complete block design with 3 replications in one growth chamber and a total of 2 chambers were employed. Both chambers were set

up at 28°C constant with a 12 hr photoperiod. Seeds (8 per line) were sown into Sunshine sterile soil mix #1in 50 cell plug flats. Seedlings were thinned to 5 per cell/line at 7 days after sowing and the height of each of these seedlings from the soil surface to the tip of the longest leaf was measured. Seedlings were fertilized (NPK plus iron) 10 days after sowing and final height and fresh weight of aboveground tissue was measured on day 14. This evaluation was performed twice and during the second experiment the chambers used for each set of 90 lines was switched as differences in the growth of the parental controls in each chamber were observed although the trend was consistent (i.e. M-203 more vigorous than M-206). Preliminary analysis indicated that the results of experiment #1 were in agreement with previous observations. Taking all the traits measured, the lines which showed better performance than the parents in experiment #1 are 14, 20, 50, 59, 125, 172, 175, 185, 187, 207 and 231. Results of these both experiments are currently being analyzed. In addition, some of the high vigor and low vigor lines are being evaluated in the greenhouse environment under more natural growth conditions.

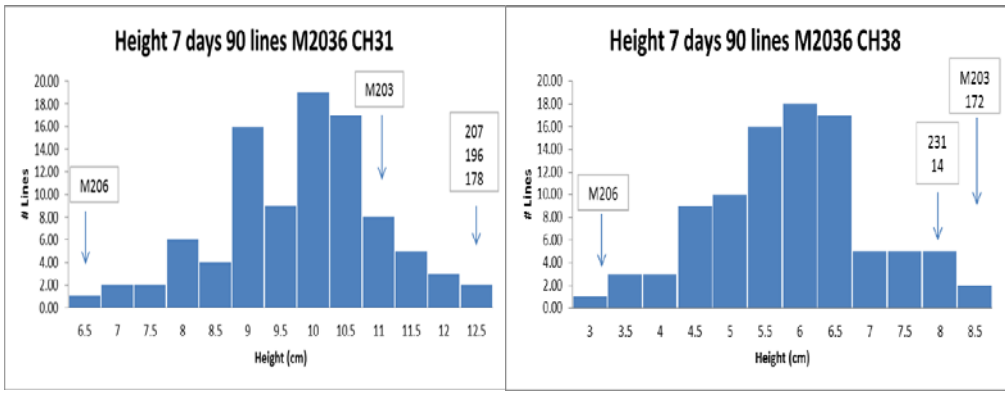

Fig. 1. Seedling height after 7 days in growth chamber (experiment #1). Seedlings in chamber 31 were in general taller than those in chamber 38 although the stronger vigor of M-203 compared to M-206 was consistent in both chambers.

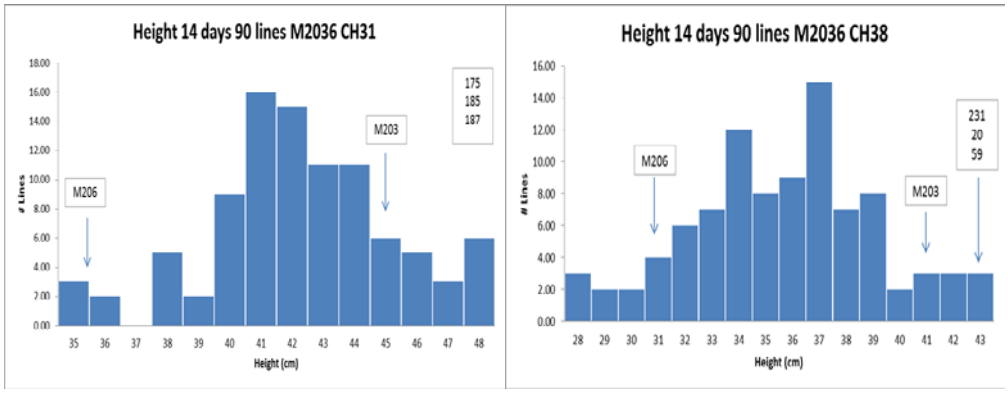

Fig. 2. Seedling height after 14 days in growth chamber (experiment #1).

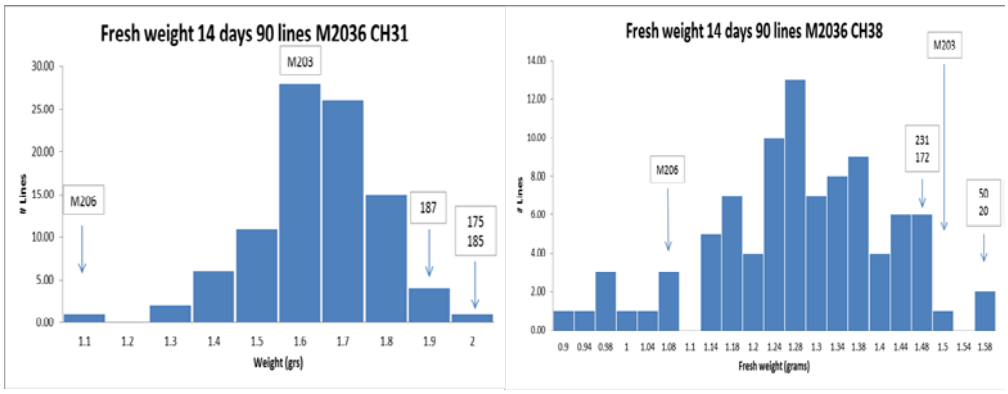

Fig. 3. Fresh weight of M2036 seedlings after 14 days.

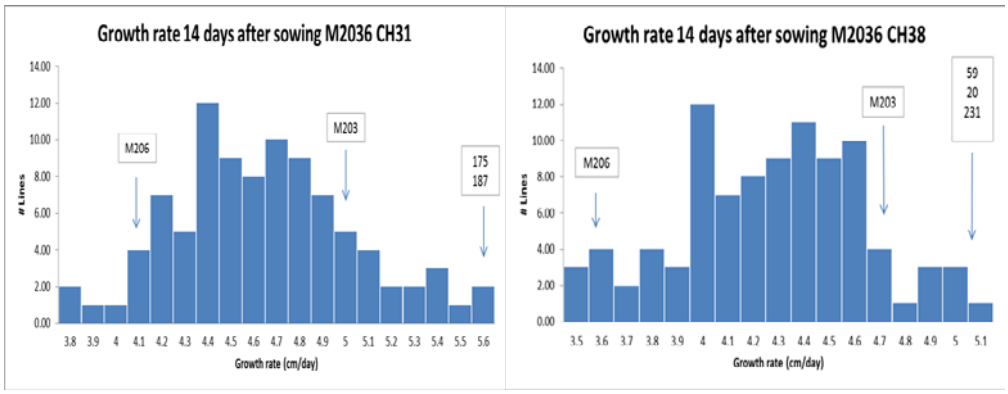

Fig. 4. Growth rate of M2036 seedlings based on the difference in seedling height between day 14 and day 7.

Fig. 5. Seedling vigor phenotypes at day 7 (left) and 14 (right). High vigor lines 172 and 231 and parents M-203 and M-206 as noted.

DNA marker genotyping. Previously, we attempted to genotype the M2036 RIL mapping population using RESCAN (i.e. selective genome sequencing). Problems with

the efficiency of generating the DNA sequencing libraries and with the UC Davis Genome Center sequencing service were encountered. In 2014, an improved protocol was tested (see objective 1 above) and during this time we elected to generate some marker information on the lines using an alternative approach based on single nucleotide polymorphism (SNP) markers identified in previous genome sequencing experiments involving the M-203 and M-206 parental lines (RB-3 project, 2012-2013). A total of 24 markers, one for each arm of each of the 12 rice chromosomes, were designed for detection of SNPs using the Fluidigm EP1 genotyping platform available at the UC Davis Genome Center DNA Technologies Core facility. Using these markers, 192 lines of the M2036 population were evaluated (Note: The number of lines evaluated was constrained by the genotyping platform which is based on DNA "chips" which are produced in limited formats). Of the 24 markers evaluated, 15 were considered to provide the most reliable scores. This 63% success rate of converting SNPs to markers from RESCAN sequence information (initial attempt) is comparable to our experiences on other projects to date. Some of the marker results are shown in Fig. 6 below.

Line	SNP2	SNP3	SNP4	SNP5	SNP6	SNP7	SNP8	SNP9	SNP10	SNP11	SNP12	SNP15	SNP21	SNP22	SNP23
172	A	A	A	B	A	B	B	B	A	B	B	B	A	B	A
231	A	A	A	A	B	A	A	B	B	A	B	A	B	B	B
59	A	B	A	B	B	B	B	A	A	B	A	A	A	A	A
20	A	B	A	A	B	A	A	A	A	A	A	B	B	A	A
201 (M203)	B	A	A	A	A	A	A	A	B	A	A	B	B	B	A
215 (M203)	B	A	A	A	A	A	A	A	B	A	A	B	B	B	A
235 (M203)	B	A	A	A	A	A	A	A	B	A	A	B	B	B	A
243 (M203)	B	A	A	A	A	A	A	A	B	A	A	B	B	B	A
M203	B	A	A	A	A	A	A	A	B	A	A	B	B	B	A
236 (M206)	A	B	B	B	B	B	B	B	A	B	B	A	A	A	B
M206	A	B	B	B	B	B	B	B	A	B	B	A	A	A	B
237 (M206)	A	B	B	B	A	B	B	B	A	B	B	A	A	A	B
241 (M206)	B	B	A	A	B	B	B	B	A	A	B	A	A	A	A
245 (M206)	A	B	B	B	B	B	B	B	A	B	B	A	A	A	B

Fig. 6. Genotypes of selected M2036 recombinant inbred lines and parents (M-203, M-206) using fifteen SNP markers derived from genome sequencing of M-203 and M-206. A and B denote the two different parental alleles for each SNP markers and have been color-coded for easier viewing. These results illustrate the different genetic fingerprints of the M-203 and M-206 parents. Lines thought to be M-203 (201, 215, 235, and 243) are consistent with the M-203 control while some of the lines thought to be M-206 (237 and 241) were revealed to be inconsistent with the M-206 control and are likely to be recombinant inbred lines that were misidentified. The patterns (fingerprints) of lines 172, 231, 59, and 20 are consistent with recombinant inbreds which differ in which parts of their genomes have inherited from M-203 and M-206. These four lines exhibit strong seedling vigor.

RS629 Population – In 2014, two populations derived from the cross RS629 (Calmochi-101/94Y561; Table 1) were advanced to the F_4 (RS629C) and F_5 (RS629AB) generations. Calmochi-101 is a short grain, early maturing, waxy endosperm cultivar known for its reproductive cold tolerance and pubescence. 94Y561 is a high yielding, long grain breeding line, which has stem rot resistance from 87Y550 and is glabrous (non-pubescent). During the greenhouse grow out, lines from each population were lost to rat damage, but sufficient numbers remain for genetic studies so the lost lines will not be re-planted.

Table 1. Status of California rice mapping populations in 2014

Population	Origin*	Original lines	2013
M2036	M-203/M-206 F_2 plants	294 F_2 plants, harvested single panicles of F_3 seeds	F_6 (~50 late/very late maturing) and $F_{6:7}$ (234 lines); evaluation stage
RS629AB	CM-101/94Y561 F_1 plants	2 F_1 plants with 400 to 500 F_2 seeds per plant	F_5 seeds harvested in 2014(~276 F_4 lines from 2 F_1)
RS629C	CM-101/94Y561 F_1 plants	1 F_1 plant with 400 to 500 F_2 seeds per plant	F_4 seeds harvested in 2014(~408 F_3 lines from 1 F_1)

*Seeds provided by Dr. Virgilio Andaya

Seeds from both populations were evaluated for hull pubescence and waxy endosperm. Examples of glabrous (non-pubescent) and waxy phenotypes are shown in Fig. 7. The waxy endosperm of Calmochi-101 is the result of a 23-base pair duplication in exon 2 of the *Waxy* gene (granule bound starch synthase) that results in its inactivation (Biselli et al. 2014). While the glabrous trait in U.S. rice varieties has been attributed to the *GL-1*, a WUSCHEL-like homeobox gene (*OsWOX3B*), which belongs to a family of transcription factors that regulate expression of genes involved in plant development (Angeles-Shim et al. 2012, Zhang et al. 2012). While the genes controlling these two traits are known, several important traits are exhibited by Calmochi-101 and 94Y561. Traits of particular interest are reproductive cold tolerance (Calmochi-101) and stem rot tolerance (94Y561). In addition, we have casually observed high yielding lines among the RS629 population. Interestingly, stem rot tolerant lines derived from 87Y550 (such as 94Y561) has at times been observed to have very high yields (F. Jodair, personal communication) and, recently, some derivatives of Calmochi-101 have shown very high yields as well (S. O. Samonte, personal communication). The RS629 populations should be useful in dissecting the genetic underpinnings of these traits.

Fig. 7. Grain traits of Calmochi-101, 94Y561, and various advanced generation progeny from the RS629 populations. A) Rough rice of both parental types and the F_1. Note the differences in seed type and pubescence. The F_1 shows intermediate seed length and pubescence when comparing it with the parents, B) Recombinant inbred line exhibiting short grain length, waxy endosperm, and glabrous hull, C) Recombinant inbred line exhibiting long grain, glabrous hull and intermediate endosperm type, D) Recombinant inbred line with long grain, pubescent hull, and waxy endosperm.

Mutant Populations: In 2014, due to limitations in space and personnel, emphasis was shifted away from the M-204 mutant population which has a low mutation density that would require increased resources to evaluate. Existing mutant populations in the genetic background of the temperate japonica variety Kitaake (~3-4,000 lines) were advanced from the M_3 to the M_4 generation. Kitaake is a short-duration variety (heading in 50-60 days) and grows well in greenhouse and field conditions. The M_4 seeds will be planted in the greenhouse during the winter and M_5 seeds will be available for 2015 field planting and evaluation. Generation advance of this population will be much more rapid, reducing the associated costs and producing fixed mutant lines (M_5 and greater) that can be screened under multiple conditions and for multiple traits. Emphasis in 2015 will be on screening the Kitaake mutants grow in the field and greenhouse environments. Resources permitting, re-mutagenesis of the M-204 mutant population to increase mutation density and seed increase of the low density mutation M-204 mutants will be considered in 2015.

Literature Cited:

Biselli C, Cavalluzzo D, Perrini R, Gianinetti A, et al. (2014) Improvement of marker-based predictability of Apparent Amylose Content in japonica rice through GBSSI allele mining. Rice 7:1.

Angeles-Shim R, Asano K, Takashi T, Shim J, et al. (2012) A WUSCHEL-related homeobox 3B gene, depilous (dep), confers glabrousness of rice leaves and glumes. Rice 5:28.

Zhang H, Wu K, Wang Y, Peng Y, et al. (2012) WUSCHEL-like homeobox gene, OsWOX3B responses to NUDA/GL-1 locus in rice. Rice 5:30.

PUBLICATIONS OR REPORTS:

Henry IM, Nagalakshmi U, Lieberman MC, Ngo KJ, Krasileva KV, Vasquez-Gross H, Akhunova A, Akhunov E, Dubcovsky J, Tai TH, Comai L (2014) Efficient genome-wide detection and cataloging of EMS-induced mutations using exome capture and next-generation sequencing. Plant Cell 26: 1382-1397

CONCISE GENERAL SUMMARY OF CURRENT YEAR'S RESULTS:

In 2014, an improved genotyping by DNA sequencing protocol based on the RESCAN method was tested and will be employed for future marker analysis of mapping and mutant populations. Exome sequencing has been delated due to technical issues although analysis of existing sequence data from previous experiments is ongoing and will be completed soon. The M2036 population was evaluated for seedling vigor traits and lines with greater vigor than the more vigorous parent (M-203) were identified. Genetic analysis of this trait is ongoing and single nucleotide polymorphism (SNP) marker genotyping has been successfully employed to characterize this population. Generation advancement of the RS629 populations was conducted and mapping populations for the evaluation of reproductive cold tolerance, stem rot tolerance, and yield are expected to be available by the end of 2015. Mutants in the short-duration variety Kitaake have been advanced and sufficient seeds (M_5 generation) will be available for evaluation of plants in the field and greenhouse as well as screening for seed-related traits during 2015.