ANNUAL REPORT
COMPREHENSIVE RESEARCH ON RICE
January 1, 2013 – December 31, 2013

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
Luca Comai, Professor, UC Davis Genome Center, UCD
Virgilio C. Andaya, Plant Breeder, Rice Experiment Station, Biggs
Farman Jodari, Plant Breeder, Rice Experiment Station, Biggs
Jeffrey J. Oster, Plant Pathologist, Rice Experiment Station, Biggs

LEVEL OF 2013 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 2013 were to 1) employ next-generation sequencing-based methods to generate highly detailed genetic fingerprints of rice germplasm and to evaluate rice genes present in California varieties; and 2) continue 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 2013 objectives included:

1) **Next-generation genotyping (NGG) of rice germplasm:** We will employ Restriction Enzyme Site Comparative Analysis (RESCAN), exome sequencing, and other NGS-based strategies to identify markers and generate highly detailed genetic fingerprints of rice germplasm including varieties, breeding lines, mapping populations, and mutants. Emphasis will be on high yielding lines and stem rot tolerance.
2) **Sequencing of expressed rice genes from California varieties:** 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 traits relating to yield and stress tolerance.
3) **Development and screening 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. The M2036 mapping and the M-204 mutant populations will be evaluated.

SUMMARY OF 2013 RESEARCH (MAJOR ACCOMPLISHMENTS) BY OBJECTIVES:

1) **Next-generation genotyping (NGG) of rice varieties and breeding germplasm:** We completed and published two studies on the application of the NGG method called Restriction Enzyme Site Comparative Analysis (RESCAN) to evaluating California varieties and to identifying regions of introgression which distinguish advanced backcross lines that exhibit differential recovery following seedling stage cold stress. In cooperation with RES breeders Dr. Virgilio Andaya and Dr. Farman Jodari, we initiated work on applying RESCAN to analyze breeding lines from the medium grain and long grain programs. RESCAN DNA libraries were prepared for a total of 38 accessions and sequenced using an Illumina HiSeq2000 (UC Davis Genome Center). A total of 160 million raw sequence reads were obtained and after processing about 125 million of the reads were used for identifying SNP markers. The number of sequence reads for the 38 entries ranged from 500,000 to 9 million with an average of 3.8 million. A total of ~7,000 SNP markers were detected over the 12 rice chromosomes, corresponding to an average marker density of 1 SNP per 53 kb.

   For the medium grain program, two small sets of germplasm were analyzed: 1) M-401 and M-401 mutants consisting of seven accessions, and 2) M-206 and M-206 backcross lines with stem rot resistance consisting of six accessions. A phylogenetic analysis of 10 of these germplasm (3 accessions did not yield sufficient sequence data to be included) and the stem rot resistant breeding line 87Y550 was conducted using a subset of 267 SNP markers (selected based on the availability of data for 70% of the lines, i.e. 8 out of 11). This analysis indicates that most of the M-401 germplasm clusters together with the exception of M-401 ES 2-7-2 and M-401 ES 2-1-2 (Fig. 1). Of the two M-206 backcross lines that were included in this analysis, RS 1396-P3 (40) is more closely-related to M-206 than RS 1396-P3 (14) which likely reflects more introgressions

from 87Y550. This is consistent with the analysis which shows RS 1396-P3 (14) grouping more closely to 87Y550 than M-206 and RS 1396-P3 (40).

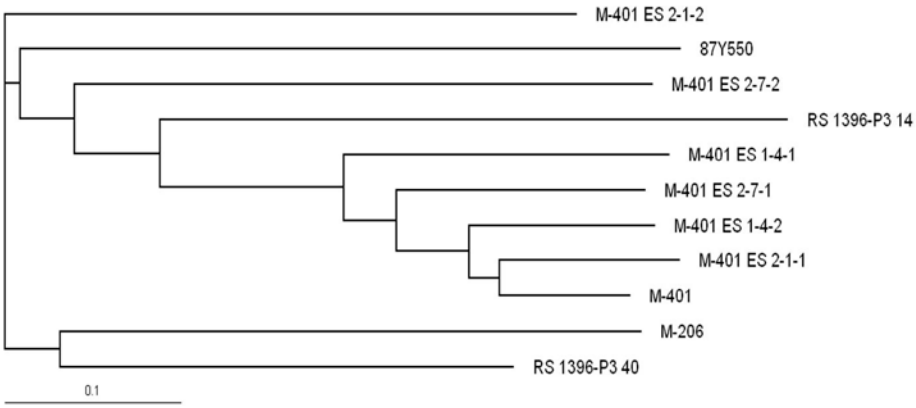


**Fig. 1 Genetic relationships of the medium grain germplasm.** Groupings based on unweighted pair group method with arithmetic mean algorithm. Genetic distance scale indicated below the dendrogram

For the long grain program, several lines which are high yielding (designated YLD) or high yielding and stem rot resistant (designated SR) were analyzed to determine if regions associated with either of those traits could be detected (Figs. 2-4). Although some potential regions of introgression were detected, sample size needs to be increased to confirm the association. Markers in the region detected on chromosome 1 (Fig. 4) may be applied to long grain lines that are being evaluated in the stem rot nursery to determine if these markers are associated with tolerance.

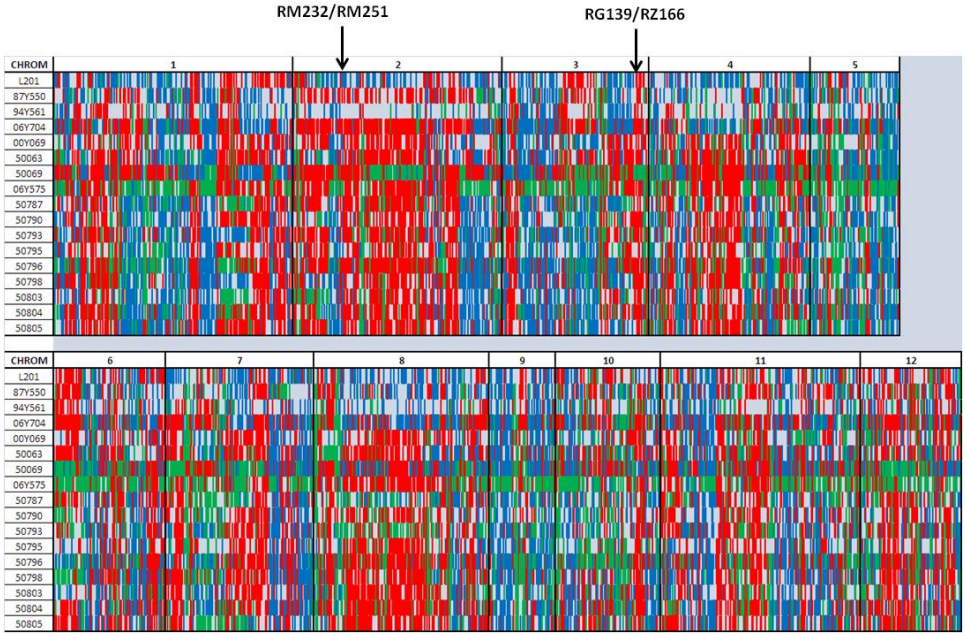


**Fig. 2. Genotypes of stem rot and high yielding lines from the RES long grain breeding program.** Twelve rice chromsomes are indicated at the top of each panel and

are separated by vertical black lines. Colored bars represent regions where the markers are identical to the Nipponbare reference genome (blue), different from Nipponbare (red), or heterozygous (green). Gray regions represent missing marker data. High levels of heterozygosity reflect the nature of the lines (not yet fixed) or may be admixtures. The markers indicated above chromsomes 2 and 3 define regions in 87Y550 reported to confer stem rot tolerance derived from *O. rufipogon* (Ni et al. 2001).

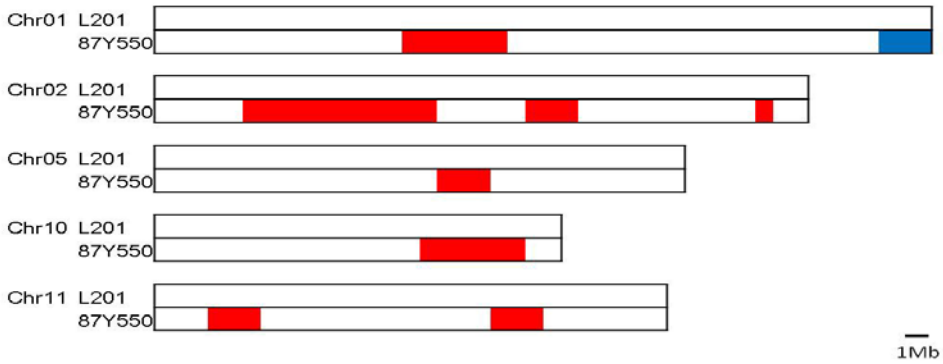


**Fig. 3. Simplified comparison of major genomic regions (≥ 1 Mb) differing between L-201 and 87Y550.** Regions differing are indicated in red or blue and are shown in reference to the Nipponbare genome as indicated in Fig. 1. Red boxes are potential introgressions in 87Y550 from *O. rufipogon*.

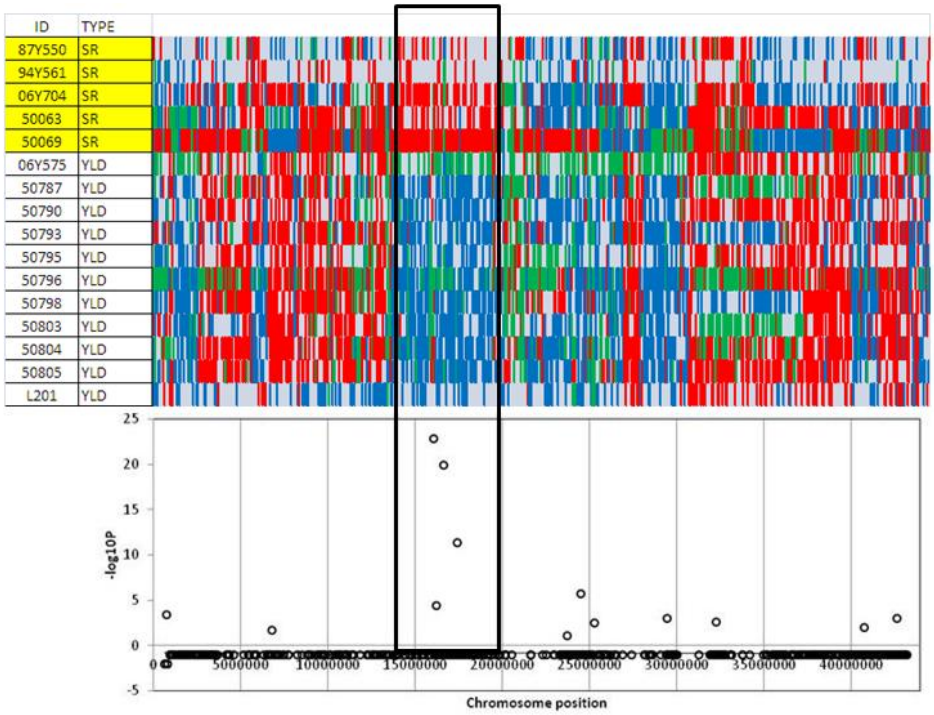


**Fig. 4. Linkage analysis of RESCAN marker data and stem rot identifies a potential region of interest in chromosome 1 (~ 3 Mb; position 15,883,807 to 18,869,430).** SR: stem rot tolerant and high yielding; YLD: high yielding

2) **Sequencing of expressed rice genes from California varieties:** As described in the 2012 Annual Report for RB-3, we are using a method called exome sequencing to conduct analysis of the expressed rice genes in California varieties. In 2013, the primary accomplishment was the completion of a study led by our UC Davis colleague and cooperator Prof. Luca Comai (UCD Genome Center) on the implementation of exome sequencing to detect and characterize mutations in rice genes from plants generated via chemical mutagenesis. This study demonstrated the efficacy of exome sequencing for detecting sequence differences in genes due to either induced (i.e. mutagenesis) or natural (i.e. evolution) variation and validated the exon capture reagent and computational analysis pipeline developed by the Comai lab to analyze the gene sequence data. This study has been submitted for publication. Based on this initial study, the Comai lab has developed a new rice exome capture reagent (i.e. designed new DNA probes for capture of gene sequences by solution hybridization) to facilitate more efficient capture of gene DNA for sequencing. This custom reagent is being ordered from Nimblegen (now part of Roche Diagnostics Corp.) for use in 2014. Analysis of the exome sequence data from Caloro, Colusa, Lady Wright, Cypress, and M-204 remains to be completed.

3) **Development of rice populations for genetic analysis of agriculturally important traits:** We continued work on the development of California mapping populations through generation advance (i.e. selfing to produce true breeding lines) and the development of mutant populations in M-204. Seed treatment policies established by UC Davis remain in place and continue to affect the speed and efficiency of this work. In addition to work on population development, preliminary evaluation of the M-2036 mapping population and screening of a subset of the M-204 mutant population for altered response to the herbicide clomazone was conducted in 2013.

   *Mapping Populations:* Work continued on developing mapping populations derived from crossing California rice varieties using the seed single-descent method to advance generations and ultimately produce fixed (true breeding) lines (Table 1). In 2013, the MS2041 and SM3014 populations were undergoing generation advance in the greenhouse. Seed set was very poor perhaps due to a major infestation of regular mites late in the season prior to and during flowering. SM3014 seeds were harvested but the MS2041 population was re-planted and moved to outdoor basins. These plants did not develop well and were not harvested. Due to insufficient resources and changing priorities, it was decided that the single seed advance of these populations and the SM3016 population should be discontinued at this time. No generation advance of the M2036 population was conducted in 2013 with research emphasis placed instead on genotyping and preliminary evaluations in the greenhouse and field.

   M2036 Population – In 2013, two objectives were pursued: 1) RESCAN genotyping of M2036 $F_6$ recombinant inbred lines, and 2) preliminary field, greenhouse, and growth chamber-based evaluation of M2036 $F_{6:7}$ recombinant inbred lines. For genotyping, three sequencing libraries covering 284 lines were constructed with a size selection of 600 bp (+/- 100 bp). Approximately 150 million reads were obtained from

each library; however, computational analysis of the sequence reads indicated that the sequence coverage was insufficient to identify SNPs across the 284 lines. Two libraries were prepared on a subset of the population (n = 74; 37 entries per library) using the size selection of 400 bp (+/- 50 bp). These libraries have been submitted for sequencing. The M2036 population was previously selected for development in order to facilitate genetic analysis of components of milling yield.

*Field evaluation.* Both parents (M-203 and M-206) have been bred for the California environment and have similar grain morphology and heading date (significant differences in these traits may complicate or confound evaluation of milling yield). To conduct preliminary evaluation, M2036 $F_{6:7}$ recombinant inbred lines (n = 234) were planted in 3 randomized complete blocks in a field at the RES (Fig. 5). Each entry was drill-seeded in three 5-ft. rows in mid-May and the field was flooded on May 30, 2013. Heading was observed around end of July with more 95% of the lines reaching 50% heading within ~1 week of each other (July 30- August 6). Observations of the entries indicated that some of the lines considered parents are likely to be either recombinant inbred lines or contaminants. This is consistent with casual observations made during the generation advance process. Genotyping will be used to establish the identity of each line prior to genetic analyses.

Table 1. Status of California rice mapping populations in 2013

| **Population** | **Origin*** | **Original lines** | **2013** |
|---|---|---|---|
| **MS2041** | M-204/S-301 $F_1$ plants | 7 plants ranging from 700 to 1000 $F_2$ seeds per plant | $F_5$ (485 lines); $F_6$ advance failed; discontinued |
| **SM3014** | S-301/M-204 $F_1$ plants | 8 plants ranging from 800 to 1000 $F_2$ seeds per plant | $F_5$ (487); $F_6$ advance failed; discontinued |
| **SM3016** | S-301/M-206 $F_2$ plants | 290 $F_2$ plants, harvested single panicles of $F_3$ seeds | $F_6$ (286); Generation advance discontinued |
| **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 | $F_4$ seeds harvested in 2013(285 $F_3$ 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 | $F_3$ seeds harvested in 2013(~450 $F_2$ lines from 1 $F_1$) |

*Seeds provided by Dr. Virgilio Andaya

**Fig. 5. Drill-seeded M2036 mapping population at the RES nursery (Photo courtesy of Dr. V. Andaya)**

In addition to heading date, a preliminary evaluation of grain moisture content was performed by sampling seeds from each row of each entry over 4 time points (Table 2). Initial sampling and moisture reading (using a hand-held, destructive moisture meter) were performed on all entries with 3 replications per entry (samples were harvested on 9-5 in small, sturdy brown paper bags and kept in plastic trash bags until reading the next day). Due to limited resources, subsequent sampling and reading was performed on a subset (n = 70) of the population on 9-13, 9-20, and 9-27-2013. Preliminary analysis indicates that there differences between lines in the M2036 mapping population with regard to the change in moisture content over time (Table 3, Fig. 6). Additional analysis of these data is needed and refinement of the methodology for sampling and moisture determination will be made for 2014 evaluation.

Table 2. Grain moisture content of M2036 $F_{6:7}$ mapping population (n = 70)

| Date | Moisture (%) range | Moisture (%) average | Moisture (%) median | Δ Moisture* (%) range |
|---|---|---|---|---|
| 9-6-2013 | 16.8 – 30.8 | 21.9±1.8 | 21.4±1.4 | - |
| 9-13-2013 | 19.1 – 26.2 | 21.3±1.4 | 21.0±1.4 | -3.6 - 7.3 |
| 9-20-213 | 16.2 – 23.3 | 18.2±1.7 | 18.2±1.3 | -0.3 - 7.1 |
| 9-27-2013 | 13.7 – 19.7 | 14.9±0.9 | 14.5±0.7 | 0.6 - 6.6 |

* Change in moisture (reduction) between previous and current date. A negative value indicates that the moisture increased for some of the lines. This may be the result of normal variation, sampling/handling error, or weather events.

Table 3. Distribution of M2036 mapping population lines (n = 70) based on change in moisture over three time periods

| | Time Periods | | |
|---|---|---|---|
| Range | 9/6 - 9/13 | 9/13 - 9/20 | 9/20 - 9/27 |
| < 0 | 26 | 1 | 0 |
| 0-2.0 | 33 | 12 | 8 |
| 2.1-4.0 | 8 | 38 | 46 |
| 4.1-6.0 | 2 | 18 | 14 |
| ≥6.1 | 1 | 1 | 2 |
| Total # Lines | 70 | 70 | 70 |

*Range: The change in average grain moisture content (%) of the population was sub-divided into five classes as shown.

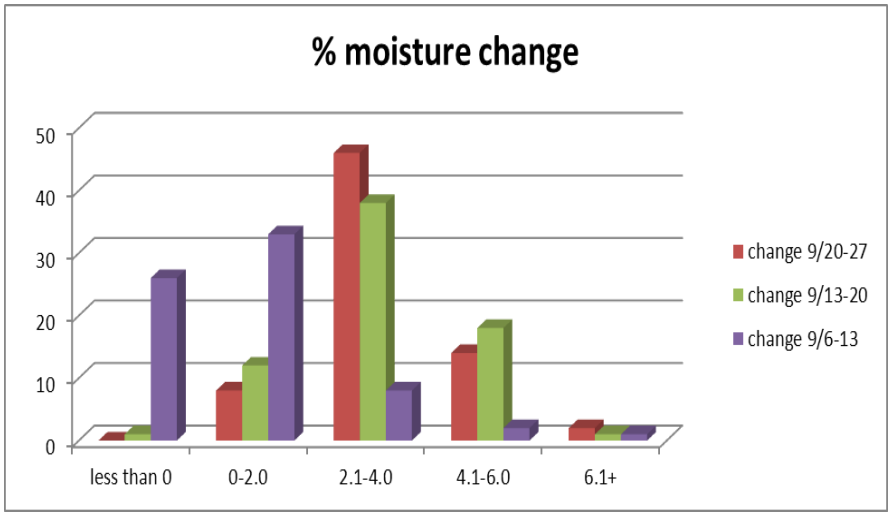


**Fig. 6. Distribution of M2036 mapping population (n = 70) based on change in moisture over three time periods.** The x-axis shows the range in changes in the moisture (%) over three 1-week time periods as indicated in the legend on the right. The y-axis is the number of lines in each group.

In addition to grain moisture evaluation, seeds of each entry (middle row, 3 replications) were harvested by the RES staff on October 23, 2013 and we will evaluate the brown and milled rice from these samples in 2014.

*Growth chamber/greenhouse-based evaluations.* A number of trait evaluations of the M2036 $F_{6:7}$ RILs were also performed under growth chamber or greenhouse conditions in 2013. Germination of the entries in soil under growth chamber conditions (25°C constant, 12 hour photoperiod) revealed significant differences between the parents M-203 and M-206 (Fig. 7). No significant differences were observed when germination and vigor were assessed under cold stress conditions (13°C constant, 12 hour photoperiod). Interestingly, both parents also exhibited similar vigor under slant board

and growth on culture media (K. Cordero Lara and S. Kim, personal communication). In addition to germination/seedling vigor, preliminary evaluation of reproductive cold tolerance in this population was performed using a subset of 108 entries (RILs plus parents). Three replications were planted in the cold screening greenhouse at the RES. Fertility data are currently being collected for analysis.

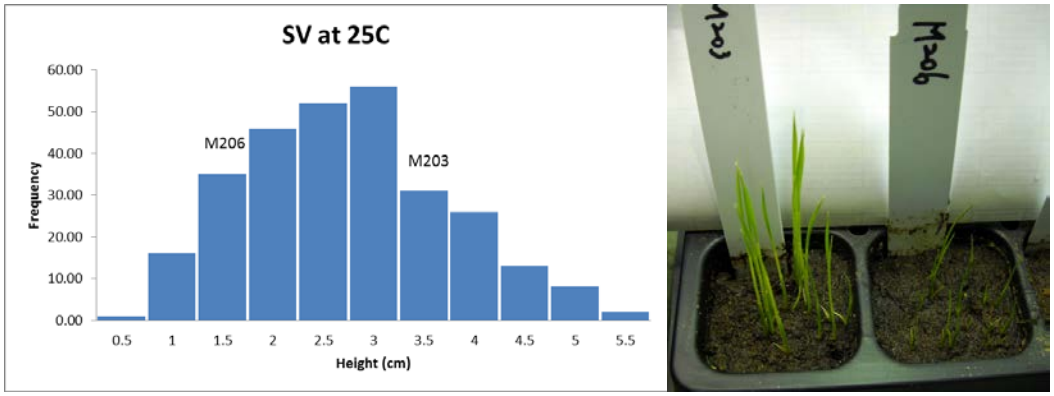


**Fig. 7. Evaluation of seedling vigor (SV)/emergence in direct seeded M2036 $F_{6:7}$ RILs grown at 25°C.** The height of seedlings (x-axis) is plotted against the frequency (number of RILs) in the y-axis (left). M-203 exhibits more vigor than M-206 as shown in the graph and pictured at the right

RS629 Population – In 2013, a second mapping population was selected for development as a priority over the MS2041, SM3014, and SM3016 populations. This population is derived from the cross RS629 (Calmochi-101/94Y561; Table 1). CM-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). We plan to conduct genetic analysis of reproductive cold tolerance, yield, and stem rot disease using this population and to introduce the glabrous trait into short grain breeding germplasm. Currently, $F_4$ seeds are being harvested from 285 $F_3$ plants (designated RS629AB population) and $F_3$ seeds are being harvested from ~450 $F_2$ plants (designated RS629C population).

*Mutant Populations:* In 2012, molecular analysis of a small number of M-204 mutants suggested a relatively low mutation density with the population (n = 3,941) of mutants derived from 3 mM sodium azide treatments exhibiting the highest density (0.8 mutations per Mb or ~300 mutations per M2 generation plant). These findings were consistent with casual observations of visual mutant phenotypes (e.g. albino, dwarf, hull color) in the populations derived from the various mutagenic treatments (see 2012 Annual Report). In 2013, focus was placed on the 3 mM sodium azide population. Generation advance of this population to the M3 was performed in 2012; however, evaluation of seed quantity indicated the need to re-grow the M2 lines to produce more M3 seeds for screening (i.e. trait evaluation). M2 seeds of the 3 mM sodium azide population (~3,900 entries) were

planted in the greenhouse and M3 seeds are currently being harvested. These M2 plants were grown in larger containers (3.5" pots vs. 50 cell plug trays in 2012) and substantially more M3 seeds have been produced and will be available for screening in 2014. In addition, a sampling of each of the ~3,900 M3 families (e.g., 5 seeds per line) will be collected and pooled for another round of sodium azide mutagenesis in order to increase the mutation density of the population.

In addition to the population development work, some screening of the 3mM sodium azide M-204 mutants was performed in 2013. For this work, two herbicides were examined: propanil and clomazone. Rice has some tolerance to each of these herbicides although there is genetic variation with some germplasm exhibiting more sensitivity than others. Experiments were conducted to determine an appropriate concentration of herbicide to use for screening. For propanil, the herbicide Stam® was applied to four rice varieties (M-204, Kitaake, Nipponbare and Terso) at five concentrations (0X, 1X, 2X, 5X, and 10X recommend label rate). The three other varieties were selected because of the availability of mutant populations in those genetic backgrounds. Results of this experiment suggest that there is genetic variation for tolerance to propanil. Terso exhibits greater tolerance than the other varieties at the higher rates applied (Fig. 8). If this differential response is confirmed, determining the basis for the response of Terso could be of interest.

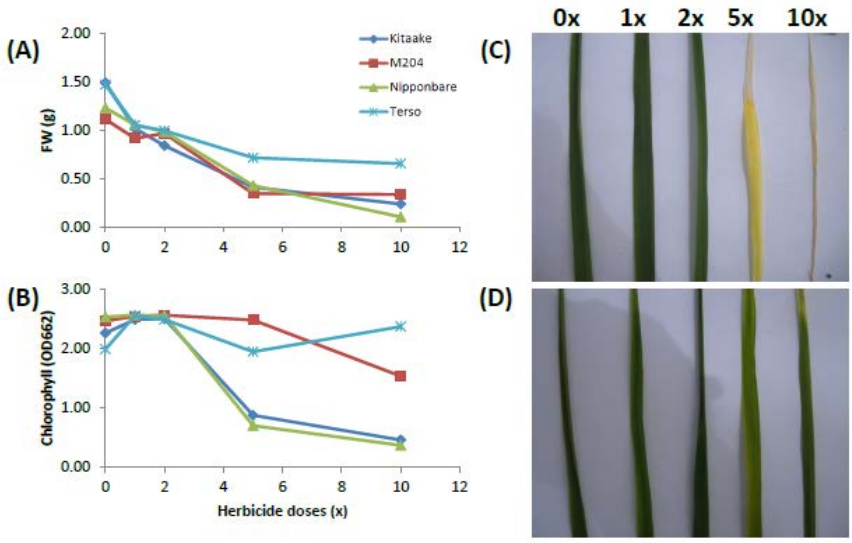


**Fig. 8. Differential response of rice varieties to propanil.** (A) fresh weight, (B) chlorophyll content – $OD_{662nm}$, (C) Nipponbare leaves, (D) Terso leaves

For clomazone, we decided to focus on M-204 and the herbicide Command® was applied at five concentrations (0X, 1X, 2X, 5X, and 10X recommend label rate). Results of this experiment indicated that M-204 exhibited chlorophyll loss and reduction of fresh weight at the 2X rate and significant damage was observed at the 5X rate (Fig. 9). The higher rate (5X) was selected for screening the M-204 mutants. Four M2 seeds from each of the mutant families (~3,900) were sown in single cells of a 50 cell plug flat containing UC

Soil Mix C and grown in the greenhouse until the 2-3 leaf stage at which time seedlings were subjected to herbicide treatment using a spray chamber. Tolerance scoring was conducted two weeks after spraying. No survivors were detected (Fig. 10).

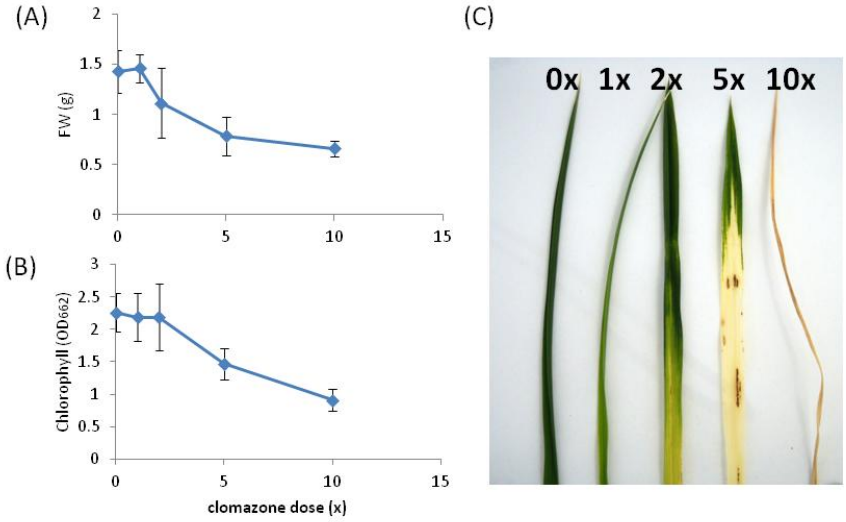


**Fig. 9. Dose reponse of M-204 seedlings to clomazone.** (A) fresh weight, (B) chlorophyll content – $OD_{662nm}$, and (C) visible damage to leaves

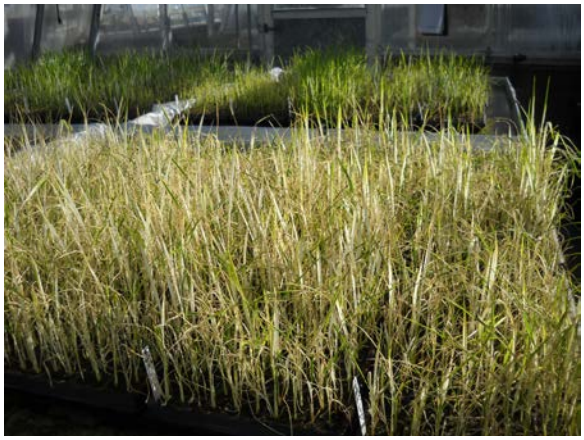

**Fig. 10. M-204 mutants two weeks after spraying with clomazone**

Several factors including mutation density of the population and the effectiveness of the bioassay may confound the identification of altered response to herbicides. While these factors may be addressed by further population development (e.g. developing populations with higher mutation density and or larger populations) and work on refining the bioassay, more targeted approaches should be considered.

Literature Cited:

Ni, J., Colowit, P.M., Oster, J.J., Xu, K., and Mackill, D.J. (2001) Molecular markers linked to stem rot resistance in rice. Theor. Appl. Genet. 102:511-516.

PUBLICATIONS OR REPORTS:

Kim, S.I. and Tai, T.H. (2013) Identification of SNPs in closely related temperate japonica rice cultivars using restriction enzyme-phased sequencing. PLoS ONE 8(3): e60176. doi:10.1371/journal.pone.0060176

Kim, S.I. and Tai, T.H. (2013) High resolution genotyping by restriction enzyme-phased se4quencing of advanced backcross lines of rice exhibiting differential cold stress recovery. Euphytica 192:107-115

CONCISE GENERAL SUMMARY OF CURRENT YEAR'S RESULTS:

RESCAN genotyping was applied to evaluate RES breeding lines and germplasm. The data generated were used to evaluate genetic relationships and in the case of the long grain breeding germplasm, regions that may be associated with stem rot tolerance or are otherwise derived from *O. rufipogon* were detected. Refinement of the next-generation genotyping pipeline including improvement of the quality and quantity of sequence data and increasing the capacity for computational analyses of these large datasets are goals for 2014. A cooperative study on the application and utility of exome sequencing for mutant analysis was completed with UC Davis colleagues. This work will serve as the basis for exome sequencing of California germplasm with the goal of developing markers based on DNA differences found within the protein-coding regions of genes. Such DNA variation may underlie differences in gene function and contribute to our understanding of important traits. Population development work was focused on a new mapping population derived from crossing Calmochi-101 and the long grain breeding line 94Y561. In addition to the well-characterized waxy endosperm and pubescence traits, the parental lines also differ in reproductive cold tolerance, stem rot resistance, and yield. Preliminary evaluation of the M2036 population in the field and in controlled environments was conducted in preparation for more detailed experiments in 2014. Advancement of one of the M-204 mutant populations was continued to generate seed for screening and additional mutagenesis experiments designed to increase the mutation density. Attempts to identify mutants with altered response to the herbicide clomazone were unsuccessful but are likely to have suffered from the relatively low mutation density of this M-204 mutant population. Other more targeted mutant approaches than direct herbicide screening of mutant populations are under consideration. Screening for useful mutants in other traits of value will be implemented in 2014.