ANNUAL REPORT
COMPREHENSIVE RESEARCH ON RICE
January 1, 2010 – December 31, 2010

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
Jeffrey J. Oster, Plant Pathologist, Rice Experiment Station, Biggs

LEVEL OF 2010 FUNDING: $37,500

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 molecular (DNA) markers that can be used to predict the presence or absence of a trait 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 DNA marker in predicting a trait must be determined). Basic genetic studies have resulted in the identification of DNA 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 objectives of our research in 2010 were to continue work to examine what genes are present in California varieties in order determine whether new genes may be useful for improving yield, quality, stress tolerance, or assist in the breeding process (e.g. wide-compatibility). In addition, work was performed towards the goal of understanding the genetic basis of key traits and the development of genetic materials to assist in studies and to serve as germplasm for the California breeding programs. All work on

this project was conducted in the USDA-ARS rice genetics lab, greenhouses, and rice research facility at UC Davis.

Specific 2010 objectives included:

1) Expansion of the assessment of gene-specific DNA markers: In 2009, ten genes (*Gn1a*, *Ghd7*, *GW2*, *GW5*, *Wx*, *SBE1*, *SBE3*, *SSI*, *SSIIa*, and *S5*) involved in yield, quality, and wide compatibility (fertility) traits were targeted for assessment by using previously characterized DNA markers and by DNA sequencing. The objective of this research was to establish a baseline of information on what alleles (forms of these genes) are present in California cultivars. This research will be continued and expanded as follows:
    a) *Assessment of breeding lines and germplasm*: Using the dataset generated in 2009, gene-specific markers will be selected to expand the analysis to breeding lines and germplasm from each of the RES breeding programs. Materials to be examined will be based on the priorities of the breeders.
    b) *Development of DNA markers and genotyping platform*: DNA sequence information from the genes analyzed in 2009 (e.g., *Gn1a*, *Ghd7*, *Gw2*, *SSIIa*) will be used to develop DNA markers. The most efficient platform for analyzing these markers will be identified in order to facilitate rapid, cost-effective analysis.
    c) *Trait evaluation and marker-trait association*: Trait data (e.g., yield, grain quality characteristics) from previous assessments by the RES and for materials grown in 2010 will be obtained/collected and examined together with the gene-specific DNA marker data to determine which markers are most strongly associated with desirable characteristics (i.e., best predictors).
    d) *Analysis of new target genes*: Due to time resource constraints, we were not able to develop a high priority gene list for future analysis in 2009. This year, a larger set of genes will be identified from published reports and subjected to analysis. Some of these targets will include but are not limited to genes controlling heading date and additional genes involved in starch synthesis. Genes selected for analysis will also be based on discussions with RES breeders.

2) Genetic studies of key traits: In addition to extending our molecular analysis of gene markers, genetic studies involving three important traits will be initiated/continued in cooperation with the RES.
    a) *Stem rot resistance*: Previous work by Dr. David Mackill (formerly USDA-ARS) and the RES (particularly Mr. Jeff Oster) resulted in the identification of potential genes from the long grain breeding line 87-Y-550 (originating from an *Oryza rufipogon* introgression). We have been unable to reproduce these findings, due in part to difficulties in trait evaluation. Dr. Virgilio Andaya and Mr. Jeff Oster have developed and evaluated a backcross inbred line population (BC_1F_5) derived from the backcross 87-Y-550/M-206*2. We will cooperate with Dr. Andaya and Mr. Oster to generate DNA marker data for this population. Another approach is to analyze the DNA of various resistant breeding lines developed in each breeding program using the *O. rufipogon* source of stem rot resistance. These lines have undergone several years of evaluation by Mr. Oster in his stem rot disease nursery. By identifying common

chromosomal regions derived from the *O. rufipogon* source, we should be able to identify markers for breeding resistance.

b) *Blanking tolerance*: In cooperation with Drs. McKenzie, Andaya, and Jodari, a set of cultivars, breeding lines, and germplasm accessions will be identified for genetic analysis of blanking tolerance. DNA marker data from blanking tolerant and susceptible accessions will be compared in an attempt to identify markers that are strongly associated with tolerance. Trait evaluation will be performed by the RES using the new cold tolerance screening greenhouse and these data along with historical information will be analyzed in conjunction with marker data.

c) *DNA marker discovery and analysis*: Single nucleotide polymorphism (SNP) and small insertion/deletion (Indel) markers are the most abundant types of DNA markers. As such, they provide the highest level of resolution (i.e., they can distinguish between the most closely-related individuals or lines). Advances in DNA sequencing and genotyping platforms (i.e., methods of detection) have reduced the cost of identifying and scoring these markers. We plan to test one method of identifying and scoring SNP and small indel markers using selective ultrahigh throughput DNA sequencing. This method will be applied to examine mapping populations that have been developed for analysis of milling quality, stem rot resistance, and blanking tolerance as described above. This work will determine the feasibility of this approach. Alternative markers and/or scoring methods will be applied to address the objectives of each study (a-c) if needed.

3) Population development and trait evaluation: Marker-trait association requires the development of genetic populations. These may be structured (i.e., developed from controlled matings such as F_2 populations, recombinant inbred lines, backcross inbred lines, etc. or unstructured (i.e., collections) such as cultivars, germplasm accessions, or breeding lines from derived from crosses of many different parental lines. As part of this project, we will continue to develop populations for current and future genetic studies. Where possible, evaluations of traits such as stem rot resistance and blanking tolerance will be conducted in cooperation with the RES.

   a) *Structured populations*: Several populations are currently under development for analysis of milling quality (derived from crosses made in Dr. Andaya's program, reported in 2009 Annual Report for RB-3). These populations will be advanced in 2010. Additional populations may be initiated in 2010 by identifying potentially valuable crosses from the RES breeding programs. Development of such populations will be done only in consultation with the RES breeders and director.

   b) *Unstructured populations*: Collections consisting of cultivars, germplasms, and breeding lines will be assembled for the analysis of milling yield/quality, stem rot resistance, and blanking tolerance. Each collection will represent the range of the targeted trait or traits and entries will be selected in consultation with the RES breeders and director.

SUMMARY OF 2010 RESEARCH (MAJOR ACCOMPLISHMENTS) BY OBJECTIVES:

1) Expansion of the assessment of gene-specific DNA markers:

a) *Assessment of breeding lines and germplasm*:

Gn1a (also known as cytokinin oxidase gene (*OsCKX2*), Loc_Os01g10110): The Gn1a gene encodes a cytokinin oxidase/dehydrogenase (OsCKX2) which breaks down the plant hormone cytokinin (Ashihikari et al., 2005). Eliminating or reducing the amount of the OsCKX2 protein causes cytokinin accumulation in inflorescence meristems and increases the number of reproductive organs, results in enhanced grain yield. Ashihikari and colleagues reported several differences in the *Gn1a* gene between Koshihikari (lower yielding) and Habataki (higher yielding), including a 16-bp deletion in the 5'-untranslated region, a 6-bp deletion in the first exon, and three nucleotide changes resulting in amino acid variation in the first and fourth exons of the Habataki gene (Fig. 1). They also found 11-bp deletion in the third exon in Chinese high yielding cultivar (5150) which created a premature stop codon (i.e. a null allele).

Fig. 1. Diagram of *Gn1a* gene (adopted from Ashikari et al. 2005). The yellow boxes represent the exons. The high yielding *Gn1a* alleles (or genes) have DNA changes that reduce or eliminate the function of the *Gn1a* (OsCKX2) protein.

We sequenced the *Gn1a* gene from four California cultivars (Calrose, S-102, M-206, and L-202). The results indicated that *Gn1a* gene in M-206 and L-202 is the same gene as Koshihikari while the Calrose and S-102 genes are the Habataki. These results indicate that the high yielding version of *Gn1a* is present in some, but not all California varieties. It would be useful to see the effect of moving the Calrose/S-102/Habataki allele into M-206. Also different versions of *Gn1a* (such as the one found in 5150) may have a positive effect in increasing yield.

GW2 (RING-type E3 ubiquitin ligase): The *GW2* gene has been shown to be associated with grain size. The hulls of some varieties with non-functional *GW2* genes have more cells, resulting in a larger (wider) spikelet hull, and accelerated the grain milk filling rate, resulting in enhanced grain width, weight and yield. Studies from other researchers have shown that the GW2 gene sequences from the variety FAZ1 (small grain) and WY3 (large grain) alleles differ and that the gene from WY3 (and the variety Oochikara) has a change that cause the creation of a non-functional protein (Fig. 2).

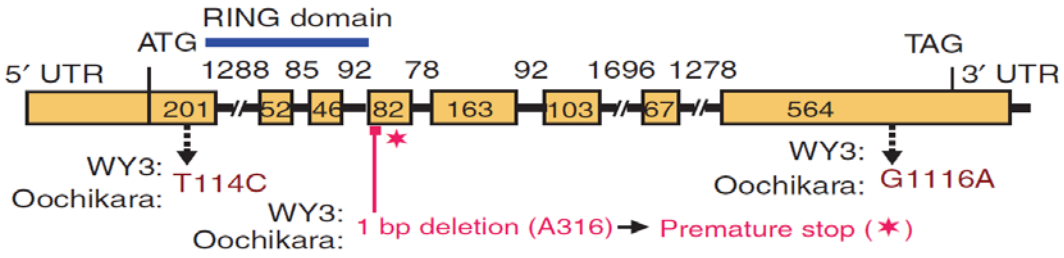

Fig. 2. Diagram of the *GW2* gene. The yellow boxes from ATG to TAG represent the exons that encode the protein. A change in the DNA of the gene in the varieties WY3 and Oochikara results in a premature stop in translation of the corresponding protein.

Sequencing of *GW2* from Calrose, S-102, M-206, and L-202 indicated that they do not have a version of the GW2 gene that might result in larger grains. This result suggests introduction of the gene from WY3, Oochikara, or some other germplasm with a similar version might result in larger grains.

GW5 gene/locus: *GW5* is a major locus controlling grain width and weight. Deletion of GW5 results in a significant increase in sink size due to an increase in cell number in the outer glume of the rice flower (Shomura et al. 2008, Weng et al., 2008).

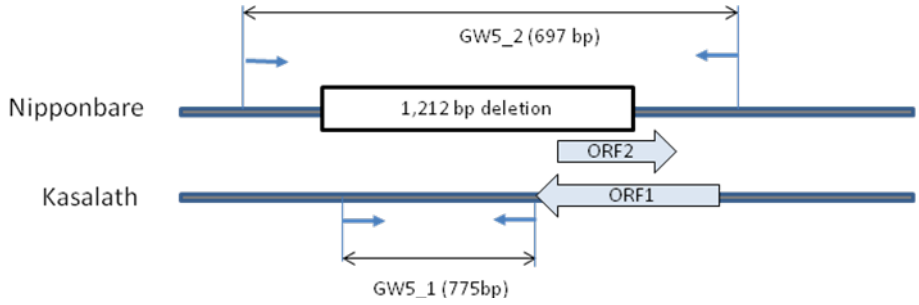

Fig. 3. Diagram of the GW5 locus from the japonica variety Nipponbare and the aus (indica) variety Kasalath. Nipponbare is missing a piece of DNA (also known as a deletion) that results in larger seed.

Out of 45 California cultivars, 12 cultivars have Nipponbare allele (Table 1). The *GW5* deletion allele is highly correlated with grain width. All the short grain cultivars have this version of the *GW5* locus except S-301. S-301 is a short grain whose pedigree is SD7/73-221//M7P-1.3.M7P-5. SD7 is medium grain from CS-M3/Calrose 76. 73-211 is a short grain from Colusa/Kitaminori. M7P-1 and M7P-5 are short grain made from an X_2 population of irradiated medium grain M7. PCR genotyping confirmed that S-301 has a wild-type allele.

Table 1. *GW5* locus in California cultivars

Cultivar	Grain type	L/W	GW5	Cultivar	Grain type	L/W	GW5	Cultivar	Grain type	L/W	GW5
Colusa	S	1.7	N	Calmochi-202	W	1.8	N	A-201	A	3.5	K
Caloro	S	1.4	N	M-302	M	2.2	K	L-204	L	3.6	K
Calrose	M	2.1	K	M-401	M	2.3	K	S-102	S	1.8	N
CS-M3	M	2.1	K	M-201	M	2.2	K	Calmati-201	B	3.9	K
CS-S4	S	2.0	N	L-202	L	3.6	K	Calhikari-201	S	1.8	N
S6	S	2.0	N	Calmochi-101	W	1.8	N	L-205	L	3.5	K
M5	M	2.2	K	M-202	M	2.1	K	M-402	M	2.3	K
Calrose 76	M	2.1	K	A-301	A	3.1	K	M-205	M	2.3	K
M7	M	2.1	K	M-102	M	2.3	K	M-104	M	2.2	K
M9	M	2.2	K	M-203	M	2.2	K	M-206	M	2.3	K
Calmochi-201	W	1.8	N	S-101	S	1.8	N	M-207	M	2.3	K
L-201	L	3.4	K	M-103	M	2.2	K	M-208	M	2.3	K
M-101	M	2.2	K	S-301	S	1.9	K	L-206	L	3.5	K
M-301	M	2.2	K	L-203	L	3.7	K	Calmati-202	B	3.9	K
S-201	S	1.7	N	M-204	M	2.2	K	Calamylow-201	LA	1.7	N

These results indicate that the version of GW5 associated with wider grain is already present in California germplasm and new introductions are not needed.

SSIIa (soluble starch synthase IIa or *alk* gene): According to Bao et al. (2006), peak gelatinization temperature is strongly associated with two adjacent SNPs (GC/TT) in the *SSIIa* gene, which can be detected by PCR amplification using specific primers. PCR based genotyping indicated that 13 cultivars have TT allele (Fig. 4, Table 2).

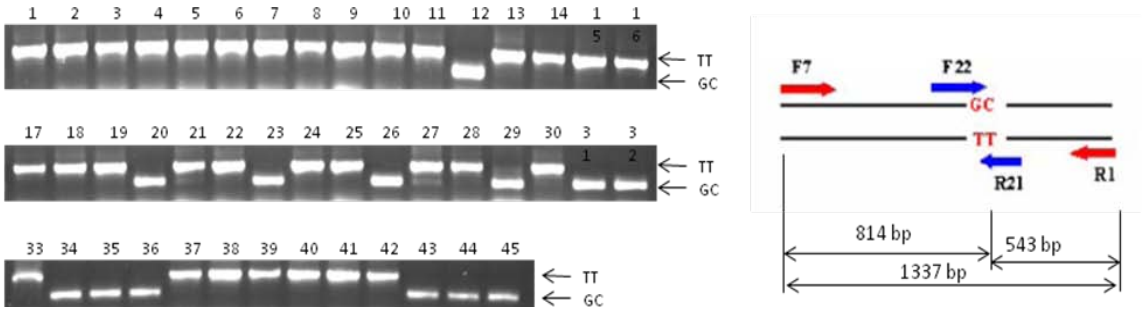

Fig. 4. PCR genotyping of SSIIa genes. PCR products analyzed by agarose gel electrophoresis (left) and diagram of the PCR products generated by different DNA primer (red and blue arrows) combinations with sizes of the products indicated in base pairs, bp (right).

Table 2. *SSIIa* alleles in California cultivars

	Cultivar	**SSIIa**		Cultivar	**SSIIa**		Cultivar	**SSIIa**
1	Colusa	**TT**	17	M-302	**TT**	33	S-102	**TT**
2	Caloro	**TT**	18	M-401	**TT**	34	Calmati-201	**GC**
3	Calrose	**TT**	19	M-201	**TT**	35	Calhikari-201	**GC**
4	CS-M3	**TT**	20	L-202	**GC**	36	L-205	**GC**
5	CS-S4	**TT**	21	Calmochi-101	**TT**	37	M-402	**TT**
6	S6	**TT**	22	M-202	**TT**	38	M-205	**TT**
7	M5	**TT**	23	A-301	**GC**	39	M-104	**TT**
8	Calrose 76	**TT**	24	M-102	**TT**	40	M-206	**TT**
9	M7	**TT**	25	M-203	**TT**	41	M-207	**TT**
10	M9	**TT**	26	S-101	**GC**	42	M-208	**TT**
11	Calmochi-201	**TT**	27	M-103	**TT**	43	L-206	**GC**
12	L-201	**GC**	28	S-301	**TT**	44	Calmati-202	**GC**
13	M-101	**TT**	29	L-203	**GC**	45	Calamylow-201	**GC**
14	M-301	**TT**	30	M-204	**TT**			
15	S-201	**TT**	31	A-201	**GC**			
16	Calmochi-202	**TT**	32	L-204	**GC**			

The TT allele is generally associated with low gelatinization temperature and the GC allele usually corresponds to high to intermediate gelatinization temperatures. Additional markers are needed to clarify grain quality characteristics.

The selection of additional varieties, germplasm, breeding lines for analysis is pending discussions with the RES breeders.

b) *Development of DNA markers and genotyping platform*:

During the course of gene sequencing for *Gn1a* and *GW2*, it was determined that the sequencing of a small number of genes was not a cost-effective approach relative to the information desired. In other words, determining the genes present in all 45 varieties instead of just 4 varieties would not be feasible. An alternative approach employs the use of ultra high capacity DNA sequencing instruments (referred to as next-generation sequencers) which are available at many sites including the UC Davis Genome Center. In this approach, gene sequences from many varieties and from many genes can be determined at a lower cost per unit of information. Such an approach is more cost-effective and the 2011 RB-3 proposal will incorporate this strategy.

c) *Trait evaluation and marker-trait association*:

Data are still being collected from the 2010 season. In addition, talks with the RES on the availability of their trait evaluation data have not been held.

d) *Analysis of new target genes*:

Recently, the group headed by Dr. Masahiro Yano (QTL Genomics Research Center, National Institute of Agrobiological Sciences, Tsukuba, Japan) reported the development of single nucleotide polymorphism (SNP) markers for distinguishing temperate japonica rice

varieties in Japan. Dr. Yano's group developed a core set of 768 and he has agreed to cooperate with us on the analysis of 16 short and medium grain California varieties to determine the usefulness of this core set of SNP markers in distinguishing California temperate japonicas. This work should be carried out in January 2011 and should enable us to determine what markers can be used for genetic studies (i.e. mapping of traits) and possibly for marker-assisted selection. In addition, we have identified about 30 genes (including some of the genes previously studied) that have SNP markers in them (also known as functional nucleotide polymorphisms) which we intend to analyze in early 2011.

2) Genetic studies of key traits:

a) *Stem rot resistance*:

Advanced backcross inbred lines (tissue and seeds) developed by Dr. Andaya and Jeff Oster (M-206 and 87-Y-550 parents) were obtained from Dr. Andaya. Initial studies by Dr. Andaya and Jeff Oster have identified 3 new QTLs (two from M-206). We are awaiting additional phenotypic data (disease scoring) before conducting genetic analysis.

b) *Blanking tolerance*:

A set of germplasm accessions were selected and seeds were amplified for use in testing for both blanking and vegetative cold tolerance. The accessions include: Silewah, Hayauki, Yoneshiro, Somewake, Mutsukogane, Leng Kwang, Thangone, and Pratao. Several of these varieties were identified previously in work performed at the International Rice Research Institute and in Hokkaido, Japan. In addition 102 additional lines grown in the 2009 UC Davis nursery were again grown in 2010 to examine heading date and fertility under the cool temperature environment at Davis. Testing of all germplasm accessions for blanking tolerance will be discussed with Dr. Kent McKenzie and the RES breeders.

c) *DNA marker discovery and analysis*:

As indicated above, a cooperative experiment has been initiated to test new markers developed in Japan. Work on next-generation sequencing and SNP marker genotyping is on-going in Davis.

3) Population development and trait evaluation:

Population	Description	Size	Use	2010
Terso mutants (unstructured)	Early flowering/maturing lines (earliest 78 days)	87 families (about 4 M3 generation lines per family)	Breeding germplasm	Amplified M4 generation (field nursery)
Terso mutants (unstructured)	Assorted M1 lines (M2 seed)	About 900-1,100 lines	Breeding germplasm	Amplified M3 generation (field nursery)
M-204/S-301 (structured)	F1 plants (seed from Dr. V.	7 plants ranging from	Mapping population	Generation advance to F3

	Andaya)	700 to 1000 F2 seeds per plant	development (milling yield)	(350-450 lines) (greenhouse)
S-301/M-204 (structured)	F1 plants (seed from Dr. V. Andaya)	8 plants ranging from 800 to 1000 F2 seeds per plant	Mapping population development (milling yield)	Generation advance to F3 (350-450 lines) (greenhouse)
S-301/M-206 (structured)	F2 plants (seeds from Dr. V. Andaya)	290 F2 plants, harvested single panicles of F3 seeds	Mapping population development (milling yield)	Generation advance to F4 (240+ lines) (greenhouse)
M-203/M-206 (structured)	F2 plants (seeds from Dr. V. Andaya)	294 F2 plants, harvested single panicles of F3 seeds	Mapping population development (milling yield)	Generation advance to F4 (240+ lines) (greenhouse)
Milling yield lines (unstructured)	Advanced medium grain breeding lines (provided by Dr. J. Lage)	135 lines (seeds from same source as used in pilot milling yield test; RF-1 project)	Mapping population development (milling yield)	Amplified seed (field nursery)

Literature cited:

Ashikari et al. (2005) Cytokinin oxidase regulates rice grain production. Science 309: 741-745

Bao et al. (2006) Nucleotide diversity in starch synthase IIa and validation of single nucleotide polymorphisms in relation to starch gelatinization temperature and other physicochemical properties in rice (*Oryza sativa* L.) Theor Appl Genet 113: 1171-1183

Shomura et al. (2008) Deletion in a gene associated with grain size increased yields during rice domestication. Nat Genet 40: 1023-1028

Weng et al. (2008) Isolation and initial characterization of *GW5*, a major QTL associated with rice grain width and weight. Cell Res 18: 1199-1209

PUBLICATIONS OR REPORTS:

None at this time.

CONCISE GENERAL SUMMARY OF CURRENT YEAR'S RESULTS:

We continued the analysis of genes involved in yield (*Gn1a*, *GW2*, and *GW5*) and quality (*SSIIa* *GW2*, and *GW5*) and based on some of these results recommendations on the introduction of new genetic materials have been made. In addition, a larger scale approach to characterization gene markers using high capacity DNA sequencing is recommended from an efficiency and cost-effectiveness standpoint. Additionally, DNA markers developed to distinguish Japanese varieties are being employed in a small scale study in cooperation with colleagues in Japan, which may provide another powerful platform for DNA marker fingerprinting of California varieties and breeding lines. Work continues on the genetic analysis of important traits for California rice production including the development and/or amplification of genetic materials (mapping and mutant populations, germplasm) for use in genetic studies and as potential donors of useful traits.