

RiceCAP Mid-Project Meeting Final Reports

**September 26-27, 2006
Stuttgart, Arkansas**

**Applied Plant Genomics
Coordinated Agricultural Project**

**A coordinated research,
education, and extension project
for the application of genomic discoveries
to improve rice in the United States**

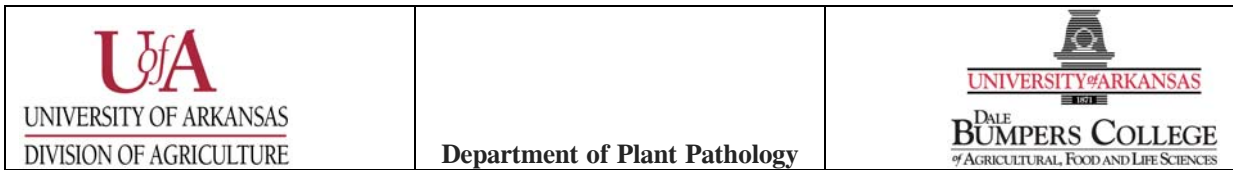


UofA UNIVERSITY OF ARKANSAS
DIVISION OF AGRICULTURE



TABLE OF CONTENTS

	<u>Page</u>
General response to reviewers from project director.....	3
Management Organization	5
Objective 1.....	6
<i>Table: The Strengths and Weaknesses of the RiceCAP Mapping Populations and the Expected Deliverables of Each Population.....</i>	<i>11</i>
Objective 2.....	18
Bioinformatics	19
Objective 3.....	23
Objective 4.....	24
Anonymous CSREES Reviews of Year 3 Continuation Award.....	26



217 Plant Science Building, University of Arkansas, Fayetteville, Arkansas 72701-1201
479-575-2445 · Fax 479-575-7601 · www.uark.edu/deps/plntpath

Dr. Ed Kalaikau

July 15, 2006

Dear Ed;

Thank you very much for your help in providing the anonymous reviews on the resubmission of our RiceCAP project for year 3. Overall, I think the comments are very pertinent and there were some very constructive suggestions. Based on our recent phone conversation, I would like to provide an abbreviated response how we can proceed with the RiceCAP effort and how making some alterations, based on the major points of the reviewer's comments, could greatly help in our continued effort. A more detailed response along these lines will be developed over the next few weeks with consultation from the executive committee. We will have the revised plan in place and distributed to the participating scientists by mid August, 2 weeks prior to the start of Year 3 of RiceCAP.

In reading the reviews, and based on our discussions of these reviews, although there are a number of areas that need to be addressed, there are two overriding concerns that are particularly germane. These areas are briefly summarized below followed by my response on the action that will be taken.

I believe one of the most salient and overriding comments in the reviews stated that RiceCAP needs to ask what is being accomplished that the individual scientists could not accomplish on their own with independent support. In other words, is the project greater than the sum of the parts. We are clearly at a critical stage in the effort. Although I believe the majority of the scientists involved in the effort are fully engaged in the project as a team effort, it is imperative that everyone receiving funding be fully engaged as we begin Year 3. To address this aspect of RiceCAP, each scientist requesting funding in Year 3 will be required to address this aspect of their efforts. In other words, "how are they contributing to the overall effort?" Furthermore, they will be asked to directly address how their efforts will impact the RiceCAP goals over the next 3 years. This impact will need to address how the work will improve U.S. rice germplasm in the near future, how the effort will impact the immediate rice community, how the effort will contribute to the biotechnology toolbox being assembled, or how the effort will continue to have an impact on educational and outreach activities. Documentation on the project "deliverables" through 2008 will also be requested from each PI. The RiceCAP project deliverables to the rice community include, but are not limited to, the development and release of

germplasm, the development and release of molecular markers, the identification and characterization of candidate genes that impact milling yield and sheath blight resistance in U.S. germplasm, publications in refereed scientific journals that will have an impact on U.S. rice germplasm, and a quantifiable assessment of how the educational and outreach activities have impacted the rice community.

A second issue that will be address is the concerns regarding the database, in the broader sense. The RiceCAP scientific advisory board highlighted some of the concerns regarding the database and these concerns were echoed in the reviews. It was anticipated that the database effort of the RiceCAP, in the broad sense, would expand as the project progressed, but this was clearly an oversight on my part and it is imperative for the project to progress, that this issue be addressed immediately.

To do so, considering the complexity of the database issues, myself along with Drs. Neil Rutger and Anna McClung, will outline and identify the most immediate database issues to be addressed by RiceCAP (queries, statistical analysis, etc.) and identify a team of personnel with the required expertise that can either directly or indirectly provide the needed guidance to address the issues that are outlined.

Although we have had several people provide input on this aspect of the RiceCAP in an informal basis, we will need to add personnel as PI's to directly help in the effort.

The personnel involved with the database from the WheatCAP and BarleyCAP will also be engaged in the effort to the extent they may be available. This critical aspect of the project will be supported in large part by the flexible funds earmarked for Year 3 as well as the redirection of a portion of Year 2 funds. I believe such an approach could be the most effective and efficient mechanism to address the concerns pointed out by the advisory boards and the reviewers.

If the above approach is satisfactory with you, I will distribute the reviews along with my response to all of the participants, begin developing the more detailed response, including putting together the database team. The short-term deliverables will also be obtained from each PI along with a statement of how each PI is contributing to the team effort, the Achilles' heel of RiceCAP.

Thanks again for your advice and guidance.

Sincerely,



James C. Correll

Management Organization

Objective 1 **Identify and use candidate genes and other molecular markers linked to quantitative trait loci which control milling quality and resistance to sheath blight disease.**

Lead Anna McClung, USDA-Beaumont, TX
Lead Jim Oard, Louisiana State University
 Steve Brooks, USDA-Stuttgart, AR
 Georgia Eizenga, USDA-Stuttgart, AR
 Bob Fjellstrom, USDA-Beaumont, TX
 Yulin Jia, USDA-Stuttgart, AR
 Sally Leong, USDA-ARS, CCRU, Wisconsin
 Steve Linscombe, LSU Ag Center, Rice Research Station
 Karen Moldenhauer, University of Arkansas, Stuttgart, AR
 Clare Nelson, Kansas State University
 Henry Nguyen, University of Missouri
 Shannon Pinson, USDA-Beaumont, TX
 Scott Hulbert, Washington State University
 Herry Utomo, Louisiana State University

Objective 2 **Validate the function of candidate genes associated with sheath blight resistance and milling quality**

Lead Yinong Yang, Pennsylvania State University
 Pamela Ronald, University of California at Davis
 Guoliang Wang, Ohio State University
 Jan Leach, Colorado State University
 Richard Nelson, Noble Foundation

Bioinformatics/Data Management

Lead Clare Nelson, Kansas State University
 Graduate Assistants

Objective 3 **Workshops**

Lead Anna McClung, USDA-Beaumont, TX
 Jim Correll, University of Arkansas, Fayetteville, AR
 Rick Nelson, The Nobel Foundation

Objective 4 **Outreach**

Lead: Rick Cartwright, UOA-Cooperative Extension Service
 Peggy Lemaux, University of California-Berkeley
 Ken Korth, University of Arkansas
 Sally Leong, USDA-ARS, CCRU, Wisconsin
 Richard Raid, Everglades-REC-Belle Glade

OBJECTIVE 1

Identify and use candidate genes and other molecular markers linked to quantitative trait loci which control milling quality and resistance to sheath blight disease

Mid-term assessment – Response to Specific Reviewer Comments (Reviewer Comments in bold)

The goals of the project need to be more clearly defined. Where are the gantt charts and critical path diagrams? Does everyone in the project see these regularly? ... Each project PI needs to be able to state precisely how their work fits in, to get to that goal. In the document I reviewed, very few of the PIs communicated the fit of their work to a specific overall goal...

Project goals for objective 1 are simply “to identify robust QTL for milling yield and sheath blight resistance in US adapted germplasm.” This goal has been understood from the beginning and activities of the Obj1 PI’s - development, phenotyping and genotyping of the mapping populations – obviously are fundamental to achieving these goals. Unfortunately reiteration of the goals and research tie-ins was not well communicated in this particular report. We have not used Gantt charts per se although we have discussed and laid out timeframes for data collection and analysis. The Gantt charts will be particularly helpful tool for integrating results from the different objectives of the RiceCAP program.

Why are four different locations doing genotyping? Which SSRs are being used and why? How many core/framework markers have been agreed upon and are in use? The cost-effectiveness of the genotyping needs to be analyzed and justified.

The purpose in having four different locations (now it is expected to be five TX, AR, LA, MS, and CA) doing the genotyping was to facilitate integration of MAS at each of the breeding locations. Having the equipment and expertise established at each of the breeding locations would allow breeders to become more familiar with the technology and learn how this can be used on a routine basis. Thus, the impact would go far beyond the scope of the RiceCAP grant. Unless breeding programs are forced to deal with various issues like tissue sampling, when to use or not use markers, what marker results mean, data management, turn around time, etc they will not truly adopt the technology. Although this may seem inefficient and not cost-effective, an underlying goal of RiceCAP was to build infrastructure and empower the individual US rice breeding programs.

SSR markers were selected from a large set of core markers that had been identified as polymorphic in US germplasm (Tai/McCouch) and others that were available through Gramene. Since an array of crosses are being used, markers used in relatively diverse crosses were not always informative in elite mapping populations. Although it took a year of genotyping the parents to determine a set of

some 200 polymorphic markers in each cross, there are still gaps in some of the more elite crosses (the markers are not well dispersed).

QTL mapping populations should have common parents (should be connected) to the extent possible, so that more than two alleles can be evaluated.

A pedigree chart is attached and shows how some of the parents in the crosses are interrelated. A series of crosses were developed that included relatively wide to relatively elite crosses. This arrangement allowed some assurance of finding polymorphisms in wide crosses that could at least guide us to the important chromosomal regions in more narrow crosses. The non-US germplasm, Jasmine 85 and Teqing have no connection with the US parents. RT0034 and MCR010277 are products of indica by US germplasm but still have little parentage in common with the parents they are crossed with. Pecos is a medium grain cultivar and is unrelated to any of the other US parents that are all long grains. In contrast, Lemont, Cypress, Cocodrie, and Rosemont have some common parentage that is believed to be enough to link the results of the mapping populations.

Germplasm in the southern US is quite distinct (and not well adapted) to what is used in California. Cypress, which was chosen for use in MY1 and MY2 because of its very unique, high milling quality, is a result of a cross between southern and California germplasm. However, to truly address milling quality in the California environment, a separate cross using germplasm well adapted to California (MY3) was developed. Three separate sheath blight populations (lacking common parents) were used for two reasons: 1) there are few, well adapted germplasm sources that have strong sheath blight resistance to choose from (sources from other countries appear to escape disease because they are tall and late maturing, not that they have a more resistant response) and 2) two of the populations (SB1 and SB4) were opportunistic choices because extensive phenotypic data had already been collected prior to the start of the RiceCAP effort. In all cases, mapping populations used in RiceCAP were ones that had already been initiated in various breeding programs (i.e. were not developed specifically for this project) which allowed us to go to the field at the onset of the project.

How 'big' and 'stable' does a marker association need to be to provide cost savings for MAS for SB and MY? This needs to be determined NOW, as this affects the population sizes and many other aspects of the project!

Additional support in analysis will help determine, using existing RiceCAP data, the population size and number of reps is necessary to make effective selection for these quantitative traits in the future. However, the population size and number of reps were agreed to be the maximum number of experimental units that it was feasible to phenotype and genotype with the resources that we have. Since milling quality and sheath blight are such difficult traits to measure, essentially any progress in developing QTLs for these traits will likely be utilized by breeders.

An additional concern, is the sense that the breeders are viewed as a phenotyping service for QTL studies, but are not being consulted about effective and efficient mating designs...This particular CAP is all about making selection more effective and efficient. All of the team members should be able to communicate how their efforts will contribute a tool to the breeder's toolbox. ...

All of the public rice breeding programs are full participants in this RiceCAP project, not just consultants. As the breeding programs learn how to use MAS technology the broad impact that it can have - from mating design, to selection, to quality control – will be realized. The impact of candidate genes on breeding programs will be best translated by demonstrating how these candidate genes are meaningful in the existing mapping populations.

The idea of a CAP was a good one. However I think this one has demonstrated how hard it is to develop and then execute a large project like this. I have considerable doubts that at the end of 3 years and \$5M there will be very much concrete to show to the community.... But the project seems to be somehow diffuse enough that there will not be the hoped-for deliverables

The group decided to focus on two very important issues to the industry, namely milling yield and sheath blight resistance. Sheath blight disease is a very important production constraint costing impacting yield and the cost of production due to the necessity of expensive fungicide applications. Thus, since there is a considerable knowledge base regarding disease resistance in rice, it was anticipated that the group could make significant progress on sheath blight resistance. Conversely, milling yield has a major economic impact on farmers and millers but was expected to be a very complex phenotype to address. Although an extremely complex phenotype, milling yield was chosen in addition to sheath blight resistance because of its potential impact on our stakeholders.

The group decided that we are most likely to have a demonstrated success in the sheath blight project than milling yield and, thus, the sheath blight mapping populations were ranked with the highest priority. However, results on SB2 need to be in hand by March 2007 (before planting) to determine if QTLs in this population are any different that what has been identified in SB1 (manuscript in co-author review) and SB4. If not, the 3 environments of data on SB2 may be adequate for publication. Resources for phenotyping SB2 in 2007 could then be re-directed elsewhere (other populations, candidate genes, etc). A new population SB5 (Lemont/Jasmine 85) is under development and selection is being performed against tall and late families that may confound sheath blight response. Because of the high level of resistance in Jasmine 85, it is possible that this population may identify novel SB QTL. In addition, over 200 induced genes have been observed in Jasmine in response to the sheath blight pathogen (Obj. 2). Enough information needs to be obtained to make the decision whether SB1, SB2, SB4, and SB5 are just validating the same SB QTLs or if novel genes are in some of these germplasms. SB4 Lemont/ Teqing is the only project that can achieve fine mapping of one or more SB

QTLs within the timeframe of this RiceCAP project. One concern is that isolated QTLs for SB may be so weak in themselves that current screening methods cannot detect their presence.

The coke bottle method has been developed and appears to be a robust, efficient method for screening SB. A manuscript has been submitted for publication. SB2 is being screened using this method and results will indicate whether the chromosomal regions for resistance using this method coincide with those identified using the field method.

Results from the toxin screening study of 17 cultivars have been weakly correlated with field resistance. Screening SB2 families showing extremes in resistance gave mixed results. Experiments need to be conducted to determine if this method identifies resistance response that is relevant. It is possible that the toxin screen could help identify isolated SB QTLs.

The wild species project will likely deliver new germplasm to the breeding and research community and identification of sheath blight resistant wild species. Continued funding will be required in 2008 to make progress on evaluating the Bengal/Nivara backcross population.

There are several candidate genes for disease resistance that need to be evaluated in the mapping populations. The sheath blight project will likely produce a complete and coherent story as compared to the milling yield effort. Resources need to be focused on several of these projects (but perhaps not all) to make sure that a clear accomplishment is achieved by the end of the grant period.

The milling yield effort is critical as far as the stakeholders are concerned. It is no surprise that it has been very difficult to assess this phenotypically and that there is a strong G x E affect. We need to put genotyping and mapping of MY2 at a high priority so that we know before planting in 2007 whether there appears to be QTL for milling or not. If there are strong QTL, we may need to consider if three locations are evaluated in 2007 (rather than the previous 2) to make sure we have enough good phenotypic data (i.e. don't lose a location due to weather) for publication of the result. If no strong QTL are identified we will need to reassess our phenotyping methods – perhaps they are not accurate enough or we are not measuring the right variables. Discussions with Terry Siebenmorgen will be helpful in this regard. We will also need to consider dropping MY2 and focusing on MY3 where we may have a better chance of detecting genotypic effects in the relatively stable California environment. By March 2007, we need to be clear about what we think can be achieved in the milling effort and make sure that is delivered by the end of the project in 2008.

Support of the establishment of MAS program in MS and MO is a clear technology transfer accomplishment that will be realized by RiceCAP which would not have been feasible without these funds. As a result, RiceCAP will be able to claim that all public US rice breeding programs will have the capability to perform MAS. If SB or milling QTLs are identified, all programs will be positioned to quickly utilize the

germplasm and markers that will have a dramatic impact on breeding efforts. This strategically places US rice breeding programs at a comparable level with many other US crops and international rice programs that have had funding for MAS technology for years.

Deliverables by December 2008

- Identification of QTL associated with sheath blight resistance, verified in more than one mapping population
- Fine mapped at least one SB QTL from SB4 Lemont/Teqing
- Identification of QTL associated with components of milling yield
- Identification of which sub-component traits are important for milling
- Mapping of candidate genes for disease resistance and factors associated with milling yield in at least one mapping population
- Micro-chamber method for evaluating for sheath blight resistance
- Determination of utility of RS toxin assay
- Winseedle method for evaluating grain dimension and other components of milling yield
- Several mapping populations genotyped and phenotyped available to the public
- Genotyped and phenotyped Lemont/Teqing backcrossed introgression lines
- New research partnerships with the US Rice Breeding community (USDA ARS Stoneville, MS; Jan Leach, Colorado State; Guoliang Wang, Ohio State, OMAP project, etc)
- Infusion of new germplasm into US breeding programs that were identified in RiceCAP mapping populations
- Establishment of knowledge base, infrastructure, and expertise to perform MAS in all US public rice breeding programs
- Development of new mapping population (for future work) containing multiple SB resistant QTLs from various germplasm sources.
- Protocols for sheath blight screening on website
- Summary of primers/markers for key traits useful for breeding on website
- New wild species of rice with sheath blight resistance available to public
- Identification of QTL associated with brown spot resistance

**The Strengths and Weaknesses of the RiceCAP Mapping Populations
And the Expected Deliverables of Each Population**

Population	Strengths	Weaknesses/Problems	Next Steps/ Bottlenecks	Expected Deliverables by Dec 2008
SB2 Cocodrie/ MC010277	51% polymorphism in 551 markers	Correlation of $r=0.6$ between 2005 and 2006 data	Need to fast track genotyping to complete ASAP. NEED TO VERIFY IF QTLs ARE SAME AS THOSE FOUND IN SB1 AND SB4 BY March 07. IF SO MAY NOT NEED TO GO FURTHER	Identify QTL associated with sheath blight resistance
	Have data from 3 environments and 1 coke bottle already	It is possible that these QTL will not be different than those identified in SB1 and SB4 already	Have run 10 markers, need more resources	Identify double haploid lines for use in US breeding programs as new sources of SB resistance
	Double haploid population, uniform, genetically stable	Possible that some of the QTLs may not be detectable phenotypically once isolated. May be a problem in other populations too.	Identify families for possible fine mapping	Genotyped & phenotyped double haploid mapping population
	Less confounding by height and heading than some pops		Incorporate candidate gene info in genotyping	Identified regions for future fine mapping of SB QTL
	Good range in SB res in population		Look for gap and redundancies in current markers	
	325 progeny		Convert sequence of cand genes to primers for mapping	
			Develop interaction with Scheffler for genotyping	

			SB2	
			2006 data was collected to determine if SB2 has potential for use in milling study	
SB5 - Lemont/Jasmine 85	JSMN has high and consistent SB resistance	Jasmine is tall and late, may confound SB ratings	Decision on whether to phenotype in field in 2007 as well as or in place of SB2	High level of resistance from Jasmine suggest it may have novel gene(s)
	Selecting against tall and late progeny to come up with subset population not confounded	Seed dormancy issue with Jasmine, causes field contamination	Genotyping of population	Identify QTL associated with sheath blight resistance
	Diverse cross, high polymorphism and widespread SB reaction	It is possible that these QTL will not be different than those identified in SB1 and SB4 already		Identify lines for use in US breeding programs as new sources of SB resistance
	Ready for field testing and genotyping in 2007	Possible that some of the QTLs may not be detectable phenotypically once isolated. May be a problem in other populations too.		Genotyped& phenotyped mapping population
	Currently being tested in CIAT with micro-chamber method	Population will be selected, not random		Test of candidate genes from SAGE, etc work
	Extensive SAGE work with Jasmine identified 200+ induced genes in response to SB pathogen			Identified regions for future fine mapping of novel SB QTL
	Detached laf method of SB screening works well with JSMN			
SB4 Lemont/Teqing TILS	Best chance for fine mapping within RiceCAP timeframe	Possible that some of the QTLs may not be detectable phenotypically	Screen TILs in coke bottle	Release of SB res germplasm

		once isolated in TILs. May be a problem in other populations too.		
	RIL population demonstrated putative SB QTL	This is a long grain/medium grain cross which has little value for milling studies	Screen TILs in toxin assay	Finely mapped SB QTL
	Have several environments documenting SB reaction		Phenotyping TILs in field 2007	Genotyped & phenotyped introgression lines
	Have genotyped 96 TILs to identify clean introgression lines		Extensive genotyping	
	Have identified 4 QTL for SB res independent of height and heading - 12a, 2, 4b, 9b			
	Have completed backcross to Lemont, BCF3 seed available for 2007 field trials			
	81% polymorphism in 564 markers			
	Mapping population used by several other research groups around the world			
MY2 Cypress/Lagrué	Cypress is the gold standard for milling quality 65%	Considering how difficult milling is to accurately measure, do we need additional field trials to get good data	Need about 60 more polymorphic markers well distributed	QTLs associated with factors affecting milling in US long grain cross
	Lagrué high yield potential, low milling quality 55%	In 2005 Cypress and Lagrué did not differ in milling at AR	Oard will ID which URN milling markers can be used	Possibly use MY2 for future QTL study on yield
	MY2 likely best chance at finding true milling markers		Translate Nguyen work - Develop snps for new markers	Identified regions for future fine mapping of milling QTL? However, may be

				very difficult to detect when isolated
	Classic population of interest to breeders		After mapping MY2, look for possible candidate genes from Ob2 group	
	37% polymorphism with 574 markers		NEED TO DETERMINE BY MARCH 07 IF QTL CAN BE DETECTED OR NOT	
	Expected less confounding factors of grain shape and maturity as compared to MY1		Send just MY2 milling extremes for fissuring assay first before doing whole population	
	325 F6 families planted		Current plans for field testing 2007 would overwhelm CA group that will also be phenotyping and genotyping MY3	
			Are we measuring the right traits?	
			Contact Siebenmorgen about other traits	
MY3 L204/01Y110	L204 high milling CA cultivar 66%	If there are problems finding QTL in other populations MY3 may be more difficult since it is more narrow	Identify gaps and redundancies in markers that need to be filled in. How does this compare with MY2?	Identified or verified milling QTL that are useful in both southern and CA germplasm
	01Y110 low milling CA line 57%	Will need to change work load or increase resources to accomplish 2007 CA field trial as well as fissuring evaluation of MY2 by CA	More markers will be needed	Demonstration that markers can be used in very narrow breeding crosses
	37% polymorphism using 530 markers		Genotyping by Tai - test run markers	Identified regions for future fine mapping of milling QTL? May be very difficult to detect when isolated

	256 F7 progeny to be planted in 2007		May need funds to finish analysis of project after end of RiceCAP grant	
	Population adapted to CA, MY2 is not well adapted to CA		Run markers ofn 14 random progeny by Bob - make sure all are working	
	The relatively stable CA environment may help remove environmental error so that genetic effects are more easily detected			
SB1 Rosemont/Pecos	Have identified 5 QTL for SB res	F3 population was used	Have planted families of lines segregating for SB QTL in BMT as possible source for fine mapping.	Have identified 5 QTL for SB res, 2 of which are independent of height and heading
	chr 1, 7 QTL assoc height, chr3 QTL assoc heading		Recover and advance population as RILS to make publicly available?	Publication in co-author review
	chr 2 and ch 9 SB res QTL independent of height and heading ALSO FOUND IN SB4		Screen purified extremes in coke bottle method	
MY1 RT0034/Cypress	Cypress is the gold standard for milling quality 65%	small population size of 120+ families	Reanalysis with dropped data from AR underway	Poster presentation at 07 PAG
	RT0034 low milling quality 40%	lack of correlation in milling between two locations may be due to G x E and post harvest handling	Reanalysis using milling as a categorical variable	Results may be used to support findings in MY2
	76% polymorphism out of 561 markers	Loss of some plots in AR	Assessment if machine dehulling caused breakage of long misshapen kernels, resulting in grain dimension not being sig in milling	Possible brown spot disease resistance QTL

			yield. However, grain dimension was not important in CPRS/Panda milling study at BMT either	
	F12 RIL population already developed, allowed field testing in 2005	Identified 2:1 segregation distortion for RT allele	Have planted families of lines segregating for milling and brown spot QTL in BMT as possible source for fine mapping	
		May not have enough CPRS introgressions to have an impact on milling	Testing of brown spot extremes in greenhouse test by Correll	
			Possible re-testing in 2007 at BMT if brown spot QTLs look sig	
OTHER PROJECTS				
SB Screening methods				Coke bottle manuscript submitted
				CIAT confident of both coke bottle and micro chamber method - but may reveal different res genes- Publication of comparison?
	Toxin screen much more controlled than other methods, less environmental noise	Loose correlation of toxin screen with field results of 17 cultivars	Set up experiment to verify toxin is a part of resistance	Protocols for screening methods on website
		Screening of SB2 extremes gave poor correlation with field results	Screen SB4 clean SB introgression lines to verify toxin screen is part of res	
			Screen wild species	
			Decision to proceed or redirect by May 2007	

Wild rice species	Identify novel sheath blight resistance genes	Difficult to produce seed, difficult to phenotype	Evaluate 50 species using coke bottle, toothpick, detached leaf, and toxin method	Summary of wild species response to SB disease
	Links to OMAP project		Identify best wild species having sheath blight res for crossing with Cocodrie SB sus parent	Identify SB QTL in Bengal/nivara cross
			Advance Bengal/nivara backcross population	Bridge to further research with OMAP group
			Bring additional species through quarantine in 2007	New wild species germplasm available to research community
MS & MO Marker Assisted Selection	Establish protocols and infrastructure for using MAS in MS and MO breeding projects	Need additional resources for 2007-2008	Screening of breeding materials for MS, MO is completed	MAS being used as a tool in all US public breeding programs
	Training of MS PhD student in MAS techniques		Decision on how to best use this technology in each program considering resource limitations	Bridge to additional funding for these state programs - MO, MS
				Establishment of ARS Stoneville lab as a genomics resource for rice community
				All US breeding programs will be positioned to use QTLs identified in RiceCAP

OBJECTIVE 2

Validate the function of candidate genes associated with sheath blight resistance and milling quality

Mid-term assessment:

With regard to Objective 2, the reviewers were mainly concerned with how the effort will be integrated into the overall mission and how the work will improve US rice germplasm. To complement the forward genetics effort by Objective 1 team, Objective 2 work was originally planned and focused on complementing the Objective 1 effort, primarily through a reverse genetics approach. Both genome-wide expression profiling and targeted approaches have been taken to identify candidate genes related to sheath blight resistance and milling yield. Candidate genes were and continue to be mapped for their association with sheath blight resistance and milling yield QTLs. Using RNAi, transgene overexpression and mutant analysis, a number of candidate genes have been or are being functionally validated for their role in sheath blight resistance. In addition, improved vectors and protocols for virus-induced gene silencing, RNAi and efficient rice transformation have been or are being developed that will contribute to the rice biotechnology toolbox. To better integrate Objective 1 work with the Objective 2 effort, the following list of priorities are proposed for the Year 3-4 research activities:

1. Convert validated genes (e.g., *OsOXL*, *OsPAL*, etc.) to molecular markers for QTL mapping and MAS.
2. Validate additional candidate genes associated with sheath blight resistance (ethylene pathway genes, *OsNH1*, *OsWRKY*, *OsChi*, *OsPR1*, etc.).
3. Work with the Objective 1 team to map and associate new candidate genes (e.g., those based on SAGE and MPSS analyses) with sheath blight resistance and milling yield QTLs.
4. Continue to improve biotechnology tools such as vectors for VIGS, RNAi and efficient rice transformation as well as PCR markers for QTL mapping and MAS in US germplasm.

Deliverables by December 2008

In collaboration with the mapping and bioinformatics groups, the Objective 2 team plans to provide the following short-term deliverables in Year 3-4.

1. Molecular markers for QTL mapping and MAS in US germplasm.
2. Additional candidate genes validated that are involved in sheath blight resistance.
3. Expression profiling databases related to sheath blight resistance and milling yield.
4. Improved biotechnology tools (e.g., improved vectors for VIGS, RNAi and efficient transformation of US cultivars).
5. Publication of 7 refereed papers.

BIOINFORMICS

Response to Reviewers

Analysis and recommendations for database/bioinformatics planning for second half of RiceCAP project

This document was produced by four external bioinformatics specialists and two external statisticians, meeting with PI Nelson and project manager Phelan at the DBNRRC in Stuttgart, AR on 9.27.06.

Introduction

External reviewers have raised a number of important issues for the informatics group to address, although we think the solutions should come from the project PIs rather than the reviewers. The reviewers' suggested solutions seem to be based on unstated and implicit assumptions and expectations regarding goals and objectives that do not exist in the RiceCAP project. To avoid similar future misunderstanding on the part of reviewers the PIs should develop and write a clear set of goals and objectives against which to evaluate the project informatics.

Recommendations

1. Define the consumers or clients for the Informatics. (Deadline: October 1, 2006)
 - a. During the discussion during the afternoon of 26 September, the consensus among project PIs was that the project members should be considered the clients for informatics for the short term.
 - i. The project "clients" include molecular biologists, breeders and quantitative geneticists/statisticians.
 - b. Discuss delivery of data and information to the broader rice community. Both content and timing needs to be decided.
2. After establishing the client base, develop and prioritize a clear set of goals and objectives for the Informatics of the RiceCAP project. (Deadline: October 1, 2006)
 - a. Project PIs need to meet with the Informatics PI to develop and prioritize a clear set of goals and objectives for project informatics. This should be accomplished at a meeting that all PIs are required to attend. It should be facilitated by the Project PI.
 - b. The objectives need to be written with clear, measurable and agreed upon deadlines.
 - c. Distinguish Data management goals from data analysis goals.
 - i. Examples of data management objectives include:
 1. Ability to visualize results across multiple QTL studies
 2. Ability to visualize location of candidate genes on the physical maps relative to the QTL/linkage maps.
 3. Ability to visualize recombinants relative to the linkage map across all segregating lines.

4. Reports on presence of desirable marker alleles linked to candidate genes in breeding lines and segregating progeny.
5. Flat-file format for raw data distribution seems to be sufficient.
- ii. Data analysis goals need to address:
 1. Responsibilities and protocols need to be established for data quality by the PIs
 2. Data generators need to agree to follow the data quality protocols
 3. Responsibilities need to be established for data analyses
- d. Goals should address the long term availability of the data and information generated by the project.
 - i. Discuss with Gramene, IRIS, GRIN, Oryzabase (others?) options and requirements to migrate data, viewers and other deliverables to these groups [having] committed funding.
3. Negotiate resources to meet the goals and objectives. (Deadline October 15, 2006)
 - a. Allow the designated PI for informatics to determine how the objectives can be met and what resources will be needed; i.e., costs, timelines and personnel (including outsourcing).
 - b. Needed resources will need to be communicated to the Project PI and Co-PIs for negotiation and approval.

Summary Statement

From the perspective of the informatics advisory group, the discussed objectives are reasonable and within the funding limits of the existing project.

Work Plan

Draft work plan for bioinformatics/data management (B&D) group for second half of RiceCAP project

Overview

The B&D tasks that RiceCAP PIs feel are needed can be placed into two classes, both centering on the discovery of DNA markers, candidate genes, and QTLs and their integration into rice physical and genetic maps. The first class of tasks, involving population, maintenance, and updating of an existing database/map viewing package (CMap) requires mostly computing skills along with some biological knowledge, and can be described as mostly data management with minor bioinformatics analysis. The second class, involving analysis of complex QTL mapping designs and interpreting and integrating results, requires expertise in statistical genetics and the computing that goes with it, plus considerable biological knowledge. These tasks are best described as bioinformatics analysis.

RiceCAP PIs do not include in their wish list the creation of an online relational database integrating all project data and results; particularly one that would persist

for the use of the broader community. While not opposed to the idea, they do not see its relevance to the immediate needs of a project with limited funds. The B&D group will therefore continue to treat such a database as a low priority for resources, while planning for the eventual contribution of RiceCAP data and results to existing public databases.

The two classes of task described above cannot be done by one person. The B&D group currently has one postdoc position (to be filled in late October) and one Ph.D. studentship, currently filled. The work plan described here is contingent on the availability of at least one further student position for the third year of the project.

Objectives

1. Maintain a WWW site giving access to RiceCAP data and analytical results
2. Receive, prepare, and make available for internal distribution all contributed RiceCAP data sets.
3. Develop and maintain CMap database and maps incorporating SSRs, candidate genes, and QTLs from literature, and expressed and tagged genes, SFPs, SNPs, and QTLs from RiceCAP and related experiments.
4. Analyze and interpret four RiceCAP QTL-mapping crosses. Reanalyze and advise on others as requested.

Integration into the overall project

Map displays maintained at the Data Center are the uniting feature of all RiceCAP experimental research, which aims to identify genes and markers for key rice traits. QTL analysis and interpretation of datasets with multiple-year replicate designs is essential to use of these data towards these aims.

Benchmarks for Year 3

Specific CMap additions

Replace already-mapped SB tags with new set assembled by Guo-liang Wang lab

Map Guo-liang MPSS differentially expressed tags for Cypress/LaGrue

Re-map microarray data from Yulin Jia

Place Nguyen/Kumar SFPs, when supplied, on MY2 polymorphism map

Place genes from Galbraith oligo chips on reference map to determine coverage

Perlegen SNP data at TIGR => reference map

Plot within-indica and within-japonica SNPs

M202/CPRS and 1 indica

General CMap display updates:

Show monomorphic markers in crosses, as well as polymorphic

Connector lines drawn between same marker on different crosses

Distinguish different QTL sources in CMap (*from RiceCAP crosses vs. from literature*)

Tai genotyping MY3. Use SFP to find more markers.

General data management

Provide integrated geno/pheno datasets ready for analysis by other RiceCAP participants

Bioinformatics analyses

Rerun QTL analysis for MY1 with cleaned data; interpret

Analyze SB2 geno + 3 yr pheno data

 help with pheno pre-analysis & quality check

Run QTL analyses and interpretation on MY2, MY3 (check MY2 pheno data before geno arrives). Assist G. Eizenga lab in analysis of BC₂F₂ (rice/rice wild relative) cross.

Lead or participate in publication development for these analyses

Personnel involved

One postdoc (Xiangqiang Sun, 100% time) will do all bioinformatics analyses including quality checks.

One Ph.D. student (Prashanth Boddhireddy, 50% time) will do all the CMap maintenance and additions, with assistance from student Joehanes, through 6/07
Ph.D. student Joehanes (50% time) will continue with QTL software development, system maintenance, and assisting in bioinformatics analyses.

OBJECTIVE 3 – EDUCATION

Objective 3 of RiceCAP involves the development of technical training programs and resources to ensure implementation of molecular marker and gene validation technologies to solve rice problems. Five educational workshops have been conducted since the project was initiated. The workshops are listed below. A general discussion was held with the RiceCAP PI's in Stuttgart and it is anticipated that another workshop will be held in 2007 but the details of the workshop will be developed during prior to February 2006.

June 2005	Marker Assisted Selection Workshop
October 2005	Virus-Induced Gene Silencing (VIGS) and Stable RNAi Workshop
February 2006	Workshop on Using Gramene
June 2006	DNA Marker Workshop: Markers, Mapping, and Beyond
June 2006	Laboratory Encounters in Plant Genomics Workshop, 5 th -12 th Grade Teachers of Science of Science and Agriculture

OBJECTIVE 4 - OUTREACH

Outreach Discussion

Stuttgart AR September 26, 2006

I. Measurable changes in attitudes or behavior (Questionnaires)

- 1) Best in winter meetings
- 2) Karen Ballard will help Rick with questionnaire
- 3) More grounded in evaluation techniques for these kind of audiences
- 4) Winter meetings in rice areas start in January; rough draft by end of October then distribute to outreach group
- 5) One meeting in early December to which Chris Greer went last year; Rice Outlook Conference
- 6) The focus is not on the numbers of exposures but did their attitudes change.
- 7) Determine knowledge gained on rice and genetics, attitudes toward use of gene mapping and breeding, value of project to end users. This will be difficult given the timeframe and resources.

II. Pens

- 1) Give to county agents, teachers, Ed Kaleikau (20 pens), rice industry leaders
- 2) Give pens to Jim Oard/Don Groth in LA (send 50/ea); Pam Ronald, Chris Greer in CA (send 25/ea); Shannon Pinson/ Anna McClung in TX (50/EA); Nathan Buehring in MS (25/ea); Donn Beighley in MO (25/ea); Rick Cartwright, Chuck Wilson, Ken Korth in AR (KK-20; CW/RC-100/ea)

III. Budget

- 1) \$15K for travel budgeted
- 2) Travel will come out of Berkeley for outreach; people will have to become aware of this.
- 3) Upon what basis do we decide who gets the funds and how do they apply?
- 4) Let people know about this through newsletter; submit proposal (1/2 page with name of meeting, what are your role, potential audience) Basis: legitimate outreach experience. Expect that we get back questionnaire data – or won't get the funds

IV. Efforts too diffuse

- 1) Focus is on agents, producers and rice consultants; once we get feedback from questionnaires from these groups we will have proof that we are focused. Made decision at beginning that this would be our focus. Korth effort was separately funded to focus on K-12
- 2) Two outreach audiences; one is the above the other is "selling CAP and CSREES" so they can get more money for CAP efforts. This is the purpose of sticky pads, hats, pens, etc.
- 3) Groups not targeted by outreach are marketers; should we do this and how? Rice Federation and U.S. Rice Producers represent marketing. How do you reach them? Crash their meetings? Personal interaction is the only way; not brochures

- 4) Do we address the GMO issues? Users want to know the basics: what is a GMO and a gene? But is this something that RiceCAP can do?
- 5) We need to stress that RiceCAP is not doing this but it is an opportunity to talk with people and likely the GMO issue will come up. And RiceCAP is the best source of science.
- 6) Can we develop a presentation that ties in the GMO issue in as a follow-up to talking about RiceCAP. Explain the GMO stuff. Start with Ken Korth presentation at miller's association last year. Ken will send to PGL to enlarge and modify and this will be made available to RiceCAPpers. Pick a person in each state and get them to use this resource. Korth/Cartwright AR, MO/ MS; Oard in LA. Good speaker and good science background; Greer/Lemaux in CA; Anna/Shannon in TX

V. Updating of brochure/poster

- 1) Give a pen/give a brochure/give a questionnaire
- 2) Update brochure: is there something to be updated? What are the accomplishments?
- 3) What about business-card type for passing out at the poster as an adjunct to the brochure?
- 4) FACT sheets? (Karen Moldenhauer suggestion) Rice industry is the end-user target. Create FACT sheets for general project or specific project. Is this for a specific project? Create FACT sheet template and this can be customized by individuals. It will be put up on the web.
- 5) Barbara and PGL create a template for FACT sheet with logo, outline for content, basic information and then others add customized information on their own project.
- 6) Connect selected ucbiotech.org FACT sheets to RiceCAP website – ask Terri to help.
- 7) Yinong send lay language summaries to outreach group for review and also as a source of information for outreach to draw on for updated research.
- 8) Please move all of the outreach materials from Downloads to Outreach section.
- 9) Basic format: Q&A; FACT sheet #1 use as a start; Title of person who is doing the customization; introduction; content with Q&A; graphics need to be able to be inserted in the center.
- 10) One sheet would also cover the Liberty Link issue

VI. Press releases, meetings with science writers

- 1) First year did some of this but may not be presented well to reviewers
- 2) Popular press is not willing to cover RiceCAP unless there is something that is of interest but we can tie it in with the GMO issue.
- 3) Giving a talk at a large meeting and get the press to come because of this issue and then also talk to them about RiceCAP

Anonymous CSREES Reviews of Year 3 Continuation Award

Review 1:

In summary, I am not impressed with the scientific progress to date or with the management of the project. The scientific and stakeholder advisory boards have pointed out the same issues that I see, but apparently the boards were not sufficiently emphatic to produce the desired changes.

At this point in the project there is no time for evolutionary change, so I recommend the following major alterations.

1. An informatics and data coordinator with industry experience (preferably someone who has managed breeders and molecular biologists through a merger) should be hired as a consulting PI reporting directly to Correll. This person will create a data handling system that will integrate every piece of data from RiceCAP, and will allow data mining, with appropriate and detailed metadata, quality control, and in a format that can be integrated with Gramene. There will be templates for data collection agreed upon and distributed BEFORE additional phenotype and genotype data are collected. Discretionary funds should be used for this; no new projects should be initiated in year 3. This person will be responsible for hardware as well. I note that tape backup, server upgrades, data curation, etc. are never mentioned or budgeted in the Year 3 proposed workplan.

2. The goals of the project need to be more clearly defined. Where are the gantt charts and critical path diagrams? Does everyone in the project see these regularly? If not, they should! For example, I would state the goals as “to develop ‘perfect markers’ for sheath blight resistance and milling yield components; these markers (genes, QTNs) will be validated by functional tests and by having strong correlations across environments and years”. Each project PI needs to be able to state precisely how their work fits in, to get to that goal. In the document I reviewed, very few of the PIs communicated the fit of their work to a specific overall goal. This problem needs to be fixed by better project management.

We have not used Gantt charts although we have discussed and laid out timeframes for data collection and analysis. The Gantt charts will be particularly helpful tool for integrating results from the different objectives of the RiceCAP program.

3. A project evaluator should be hired. This person will have expertise in design and administrator of grant progress evaluation and will regularly evaluate each part of the project and the fit to overall goals. This can be done using a carefully design web survey instrument. The results should be communicated first to Correll and then to the entire project team. I suggest that funds for this come from increasing the efficiency of genotyping (see below).

4. Why are four different locations doing genotyping? Which SSRs are being used and why? How many core/framework markers have been agreed upon and are in use? The cost-effectiveness of the genotyping needs to be analyzed and justified.

The purpose in having four different locations (now it will be five TX, AR, LA, MS, and CA) doing the genotyping was to facilitate integration of MAS at each of the breeding locations. Having the equipment and expertise established at each of the breeding locations would allow breeders to become more familiar with the technology and learn how this can be used on a routine basis. Thus, the impact would go far beyond the scope of the RiceCAP grant. Unless breeding programs are forced to deal with various issues like tissue sampling, when to use or not use marker, what marker results mean, data management, turn around time, etc they will not truly adopt the technology. Although this may seem inefficient and not cost-effective, the purpose was to build infrastructure and empower the individual US rice breeding programs.

SSR markers were selected from a large set that had been identified as polymorphic in US germplasm (Tai/McCouch) and others that were available through Gramene. Since an array of crosses are being used, markers used in relatively diverse crosses were not always informative in elite/elite mapping populations. Although it took a year of genotyping the parents to determine a set of some 200 polymorphic markers in each cross, there are still gaps in some of the more elite crosses (markers are not well dispersed).

5. The education effort is far too diffuse. For the small amount of money much more focused efforts need to be agreed upon and carried out. The amount of funding available is not probably not adequate for effective use in K-12 programs; I suggest efforts be focused on producers or extension agents. Whatever is done must include outcomes—MEASURABLE changes in attitude or behavior.

Background to my suggestions:

56The wheat and barley CAP projects have raised the bar—but rice needs to learn (i.e. take action and make changes) based on what has been learned. I summarize what I have learned as:

Applied molecular breeding = MAS (except in corn and soy where Bt and Round-up ready are already in use).

MAS is all about cost-effectiveness.

CAP projects must be able to demonstrate synergy, with some aspect of the work that clearly could not have been done by distributing the funds to the PIs separately.

Additional comments (noted while reading the proposal):

a. QTL mapping populations should have common parents (should be connected) to the extent possible, so that more than two alleles can be evaluated. There are published examples of use of these pedigrees now available and this has been implemented in industry for some time.

A pedigree chart is attached and shows how some of the parents in the crosses are interrelated. Germplasm in the southern US is quite distinct (and not well adapted) to what is used in California. Cypress, which was chosen for use in MY1 and MY2 because of its very unique, high milling quality, is a result of a cross between southern and California germplasm. However, to truly address milling quality in the California environment, a separate cross using germplasm well adapted to California (MY3) was developed. Three separate sheath blight populations (lacking common parents) were used for two reasons: 1) there are few, well adapted germplasm sources that have strong sheath blight resistance to choose from (sources from other countries appear to escape disease because they are tall and late maturing, not that they have a more resistant response) and 2) two of the populations (SB1 and SB4) were opportunistic choices because extensive phenotypic data had already been collected prior to the start of the RiceCAP project. In all cases, mapping populations used in this CAP grant were ones that had already been initiated in various breeding programs (i.e. were not developed specifically for this project) which allowed us to go to the field within months of receiving funding for the project.

b. One postdoc and possibly a PhD student in the Nelson lab will not be enough to do the large complex models needed for multiple populations/years/environments. I suggest collaborating with or hiring an experienced statistician, preferably one who can bring Bayesian expertise to the project. A data mining expert and a statistical modeling expert would provide two very useful and complementary perspectives on overall RiceCAP data analysis.

c. How 'big' and 'stable' does a marker association need to be to provide cost savings for MAS for SB and MY? This needs to be determined NOW, as this affects the population sizes and many other aspects of the project!

Additional support in data analysis will help determine, using existing data, what population size and number of reps is necessary to make effective selection for these quantitative traits. Since milling quality and sheath blight are such difficult traits to measure, essentially any progress in developing QTLs for these traits will likely be utilized by breeders.

d. Consider adding a Science Advisory Board member from the WheatCAP Science Advisory Board...in fact, just do this.

Review 2:

The SAB identified the lack of progress in meeting the Bioinformatics objectives as the most significant concern and made a number of specific recommendations. I'll address the specific recommendations below. I think that the lack of progress in Bioinformatics may be more symptomatic of a larger issue that needs to be clarified and addressed. The larger issue is that the project members do not yet share a common vision of how to implement their translational mission of improving milling quality and resistance to sheath blight in US rice germplasm.

The individual progress reports and work plans did little to convince me that most project members understand their roles as part of a team. In particular, labs working on objective 2 failed to communicate how their efforts will be integrated into the larger mission. Of what use are candidate genes to crop improvement? How will a breeder use this information? How will they communicate with breeders how to best use the information? Even though many of the labs working on objective 1 had statements about integration into the overall project, their statements were high-level and did not specify how their results were going to be integrated into the project so that the larger mission would be realized. Without such communication among the team members it is not surprising that the bioinformatics effort has floundered.

An additional concern, is the sense that the breeders are viewed as a phenotyping service for QTL studies, but are not being consulted about effective and efficient mating designs for deploying desirable alleles throughout the germplasm. In its most essential form, crop improvement is about making crosses among members of a breeding population and selecting segregating progeny from the crosses for another round of crossing and selection. This particular CAP is all about making selection more effective and efficient. All of the team members should be able to communicate how their efforts will contribute a tool to the breeder's toolbox. There were only a handful of reports (e.g., the Scheffler and Eizenga labs) that were able to clearly communicate how they would contribute such tools for use in crop improvement.

All of the public rice breeding programs are full participants in this RiceCAP project, not just consulted. As the breeding programs learn how to use MAS technology the broad impact that it can have - from mating design, to selection, to quality control - will be realized. The impact of candidate genes on breeding programs will be best translated by demonstrating how these candidate genes are meaningful in the existing mapping populations.

Bioinformatics:

The SAB expressed concern about the lack of progress in Bioinformatics, but as indicated above, this may be more symptomatic of a general lack of communication among project members. Regarding specific recommendations from the SAB and responses (*italicized*) from the RiceCAP executive committee see my comments (in **bold**):

1. The SAB was pleased to hear of developments at organizing an underlying database schema, but feels that this must be accompanied by interface development if the resource is to deliver maximum impact for the PIs and the project as a whole.
2. Beyond interface development, the Project has not yet shown significant progress in the planning or execution of systems that will contribute to data interoperability – either with existing public database resources (e.g. Gramene) or with other more specialized data repositories.
3. The Plan for future work in Year 2 is scant on detailed goals or timelines... In general, the Plan needs to articulate how and when the required personnel will be in place, how the past year's goals will be recovered and achieved, and how the coming year will better serve the needs of the PIs.
4. Finally, it is reasonable to expect a clear statement about how the project Informatics will interweave with existing public resources and interoperability protocols.

1) A collection of all the raw Microsoft Excel or text datafiles generated in the project will be made available to participants via hyperlinks on a WWW page.

While this is an important first step, it is insufficient and does not address the SAB comments and places an unnecessary burden on project PIs for independent investigations. The data should be stored in a relational database with a simple query interface for bulk downloads of (merged, concatenated and subsets) data. Such a simple resource could be developed by an undergraduate.

2) Ultimately the scientific community will expect the data generated by RiceCAP to be organized in a relational database for use by future researchers. Gramene principals stress that, for Gramene to consider serving as this repository, it will require (along with its own renewal of public funding in 2007) a commitment from RiceCAP to careful data preparation and interaction over at least several months. We will pursue this discussion; however we do not expect Gramene to serve as a working database for our data during the RiceCAP project, since these data may not be ready for public release until the late stages of the project.

By parsing project data into a relational database, particularly if the project uses previously developed schema, e.g., previously developed relational databases such as Panzea, most of the data preparation for eventual upload to Gramene will be simplified.

3) The RiceCAP personnel feel that there is a need for a focused effort on QTL mapping, and as a result, the data-management/bioinformatics group will shift its focus to QTL analysis. We will continue to maintain a set of CMap displays for showing maps with markers, candidate genes, and QTLs.

CMAP is clearly the best available tool for storing and comparing results from QTL analyses. I am not sure I understand the reasons for concentrating QTL

analyses in the hands of a few PIs. It would seem that all PIs would be interested in conducting QTL analyses: Numerous analysis packages are available, many students distributed throughout the project can be taught and results can be compared. The statisticians associated with the project should focus on harder problems such as calculating the power and expected bias of estimated genetic effects. I doubt that complex models with many parameters, e.g., QxE and QxQ interactions, will have sufficient power with the small sample and low heritability associated with the MY1 cross. It would be good to know this before considerable effort is wasted on complex models.

4) The Executive Committee feels that there has been a lack of communication among the RiceCAP participants to define the role of the data-management group. To rectify this, Leisha Vance will organize a conference call the first week of each month with Jim Correll, Neil, Clare, Anna, and Jim Oard to discuss efforts, updates, progress and plans on data analysis. These will be initiated in May 2006. Communication of the updates on data analysis will also be provided via e-mail by Clare on a monthly basis to the RiceCAP PI's.

Agreed (see general overview)

Suggested benchmarks for RiceCAP database management team until August 31, 2006

1. Plan for replacement of individual(s) on project to complete stated research goals in a timely manner. The postdoc position being advertised for database development will be replaced by one for QTL and expression/candidate-gene analysis.

This may be appropriate, if the post doc will do more than run existing software packages on these data. Power calculations and bias estimates need attention. Also, some thought needs to be devoted to meeting the data management needs of the group as per suggestion above.

2. Analyzing and summarizing the data in a timely manner that will benefit the entire RiceCAP effort. In mid-May, Ph.D. student Joehanes and another Ph.D student in the group, employed at half time on postdoc funds left from the first four months of 2006 will begin QTL mapping and presentation of results, depending on the supply of data by other RiceCAP participants. The postdoc to be engaged for these analyses will assume increasing responsibility for them probably in the summer. Following further discussions with Gramene, this person will also be asked to prepare data for integration into that database.

It is not clear that communication with Gramene is appropriate for a post-doc. Design and requirements-gathering should involve PIs, if for no other reason than to improve communication throughout the project and with Gramene. It will enable some serious discussion with Gramene about whether Gramene will be a breeders resource in the long-term. Implementation of data transfer

from a project db to Gramene can be accomplished with an undergraduate-level software developer.

3. Planning or execution of systems that will contribute to data interoperability – either with existing public database resources (e.g. Gramene) or with other more specialized data repositories. As noted above, RiceCAP will make raw data internally available and will negotiate with Gramene to receive data and results.

As noted above, a project db and web-site will better serve such a large group of PIs, particularly if they desire to combine data in unanticipated ways.

4. Justify plan for a Map Viewer that is already available via public resources (Gramene) as is a catalog of SNPs from an indica vs japonica comparison. A map viewer has not been proposed. Our installation of the publicly available CMap has been online for a full year. There will be no SNP catalog.

CMap should be sufficient for comparing results of QTL experiments.

5. Justify creation of an ontology database and interface against the backdrop of existing public resources in this area. We will do no further work with ontology resources. Those installed were created on a different project, as stated in the work plan, and have not cost RiceCAP any resources. This work was proposed and done to address concerns raised by the Boards at previous project meetings.

Agreed.

6. Create clear statement about how the project Informatics will interweave with existing public resources and interoperability protocols. See the response to point 3.

7. Participate in 2006 Marker Workshop by giving presentations on linkage mapping and basic genotype and phenotype data management.

More members (particularly the breeders and phenotypers) of the project need to participate in linkage mapping (see above). How can the project purport to teach others about data management if the project does not engage in serious data management through development of a project database?

8. Depending on the availability of QTL data for analysis, we will add genomic regions and candidate markers for sheath-blight resistance and milling traits to the CMap interface already available on the RiceCAP WWW site, beginning in July 2006.

Agreed.

Review 3:

RiceCAP year 3 review.

I have reviewed the Continuation Award, Year 3, July 1, 2006, for RiceCAP. The project is making excellent progress. Although I paid particular attention to the Outreach goals and reports, I commend the project utilizing their scientific and lay advisory boards to set priorities and for redirecting parts of the project based on advice given and progress made on scientific investigations.

Objective 4 – Outreach, is making excellent progress. I believe the project is on track with the development of curriculum and materials for initial delivery. I think that at this stage there has been significant participation in workshops and field days by the project. The efforts seem to be reaching the audiences that the project was designed to reach. It will be interesting to see what the level of recognition is for the RiceCAP project from the 2006 evaluations.

I would also commend the partnering with the outreach efforts of WheatCAP and BarleyCAP. Leveraging the work of those projects is an excellent way to extend scarce resources.

It is difficult to determine the exact amount of funding that is going into the outreach effort. I would recommend that the project director identify this in the overall budget narrative. I also recommend that the outreach team begin to identify outcomes and evaluations for those outcomes that they want to reach by the end of the project. Awareness of the RiceCAP project is a good short term goal, but longer term, they should try to determine knowledge gained about rice and genetics by targeted audiences, attitudes toward use of gene mapping and breeding techniques for rice and also the value of the project to the end users.

I recommend continued funding.

Review 4.

Progress made to date:

All-in-all the PIs have made relatively good progress, although clearly some more than others. In hind-sight the choice of the existing MY1 population was a poor one (outcrossing, etc.) but at least some QTL information was derived from it. The new populations should be more informative.

The outreach efforts seem OK. I have no experience with such things but I wonder if hats etc. really do much good. Press releases, meetings with science writers, etc. might have more payoff in the long run.

The almost complete lack of progress on the informatics was disappointing. This topic was given very short shrift in the original RiceCAP proposal and I severely dinged them for that. It looks like although they changed their stated intentions to get the funding they did not change their understanding of it's fundamental importance.

Year 3 work plan:

This seems reasonable given the current state of the project. The change in focus for the informatics group does little to convince me that there will be any way the raw data from the project will be available to the community for re-analysis, integration with other data, etc. The responses to the SAB's comments are good and clearly show that the PIs are aware of the problems and moving in the right directions to correct them.

Overall recommendation for funding and improvement:

A lot of money has been invested in RiceCAP and it would be foolish to not continue the work for a 3rd year. The changes in the research plan suggested by the SAB and accepted by the PIs are good and, in my opinion, will make the project's results more useful much sooner.

Some general comments:

The idea of a CAP was a good one. However I think this one has demonstrated how hard it is to develop and then execute a large project like this. I have considerable doubts that at the end of 3 years and \$5M there will be very much concrete to show to the community. Don't get me wrong: all of the participants are good and did good science. But the project seems to be somehow diffuse enough that there will not be the hoped-for deliverables. Perhaps because of my own biases, I am particularly distressed that there will be very little in the way of databases or raw data made available to other researchers.