Annual Technical Update Report

Date Report Submitted: October 15, 2008

Identification

GCP Project Number:
G3005.10 (Competitive grants 2005-2007 Project # 10)

Project Name:
Exploring Natural Genetic Variation: Developing Genomic Resources and Introgression Lines for Four AA Genome Rice Relatives

Principal Investigator:
Mathias Lorieux, IRD/CIAT, Colombia – mathias.lorieux@ird.fr, m.lorieux@cgiar.org
Joe Tohme, CIAT, Colombia – j.tohme@cgiar.org

Collaborators (name, affiliation):
Susan R. McCouch, Cornell University, USA
Claudio Brondani, CNPAF-Embrapa, Brazil
Baboucarr Manneh & Marie Noelle Ndjiondjop, WARDA, Côte d'Ivoire
César P. Martinez, CIAT, Colombia
Miguel Diago Ramirez, Fedearroz, Colombia

Total Project Budget:
1,074,900

Grant Period: (put the dates in original project contract)
(Start: 01_2005) (End: 12_2008)

Reporting period:
(Start: 01_2008) (End: 11_2008)

Report

I. Executive summary

Cereals provide the majority of calories consumed by humans. Cereal production faces growing challenges due to increasing human population, changing nutritional requirements and variable environmental conditions that require new approaches to crop production. Wild relatives of modern crop species have survived for millions of years using natural genetic defenses to endure biotic and abiotic aggressions. These wild relatives represent a valuable source of under-utilized genetic variation that is available to plant breeders and represent an invaluable source of genetic information for modern genomics research initiatives. A systematic approach is required to identify and characterize genes from wild species that can be used to enhance crop productivity in a range of environments and under diverse cultural conditions. Using rice as a model, we propose to (1) develop four libraries of interspecific lines called Chromosome Segment Substitution Lines (CSSLs), targeting chromosomal introgressions from different rice relatives, (2) develop a set of 140 molecular markers (called SNPs) identified in genes associated with tolerance to abiotic stress (drought, acid soils, mineral deficiencies or toxicities), (3) validate the utility of the SNPs by using them in the development of the CSSLs in this project and exploring their value in breeding programs for other cereals (4) analyze a set of advanced CSSLs generated from Asian x African rice crosses for their phenotypic response to drought stress. Generating such resources and knowledge will

contribute to the objectives of Subprograms 1 and 3 by (i) utilizing *natural genetic diversity* to develop whole-genome libraries of CSSLs as a permanent genetic resource for both breeding and genomics-based research (ii) producing high-throughput, cost-effective markers to *facilitate access to genetic diversity* in a range of different cereal species (iii) making the CSSLs available to breeders and geneticists so that the intersection of their efforts will continue to generate new knowledge.

Key words: *Chromosome Segment Substitution Lines* (CSSLs) – Natural Genetic diversity – *Oryza glaberrima – Oryza rufipogon – Oryza barthii – Oryza glumaepatula* – Single Nucleotide Polymorphisms (SNPs) – Simple Sequence Repeats (SSRs)

II. Introduction/Background

The future of crop improvement depends on the availability of genetic variation. Most modern crop varieties have undergone a genetic bottleneck associated with the process of domestication resulting in a restriction of the genetic options that are available to plant breeders. There is a larger pool of genetic variation available in landraces and wild relatives of crops. These resources are known to contain many interesting traits for breeding, including good to strong tolerance to abiotic and biotic stresses and various nutritional traits of interest (Sun et al 2001). However, it is often difficult to utilize these natural sources of genetic diversity because of fertility barriers, linkage drag, the time and resources required to recover useful recombinants. This project is designed to take advantage of the unexploited reservoir that exists in the wild relatives of cultivated rice (*Oryza sativa* L.) through the development of introgression lines that will be of immediate use to breeders and will simultaneously serve to enhance our understanding of the "wild alleles" that contribute favorably to plant performance under drought stress.

The cereal crops, rice, maize and wheat, are together the most important human food crops in the world. The demand for cereals is expected to increase dramatically, particularly in the developing world where population is increasing the fastest. Cereal production faces the difficult challenge of obtaining reliable yields under variable conditions, notably due to the prevalence of biotic and abiotic stresses. Among the abiotic stresses, water deficit is among the most prevalent and the most severe in terms of restricting yields in cereal crops. Fresh water is the most limiting resource on the planet and as it becomes more scarce, breeders will devote increasing energy to developing crops that are better able to withstand water deficits or to use water more efficiently.

Rice evolved as an aquatic plant and traditional rice cultivation makes use of standing irrigation water to control weeds. Upland rice represents an adaptation to dry land. There is enough genetic variation within the genus to enable rice to adapt to drought-prone conditions. Upland genotypes are widely grown throughout West Africa and large parts of Brazil. These areas represent important targets for our work to identify Chromosome Segment Substitution Lines (CSSLs) that perform well under drought. To improve current levels of drought tolerance in new varieties, we propose to search for "new" alleles associated with drought and acid soil tolerance in the reservoir of natural genetic diversity available in wild relatives.

It is well established that (1) a great deal of genetic diversity was left behind during rice domestication in Asia and (2) rice relatives belonging to the *O.* AA genome complex contain a high level of diversity that is sexually compatible with O. sativa and can be accessed via crossing and selection.

Plant geneticists have long recognized the value of exotic libraries. Interspecific crosses between *O. sativa* and two African species*, O. glaberrima* and *O. barthii*, an American species, *O. glumaepatula* and an Asian species*, O. rufipogon*, have demonstrated the utility of targeted introgressions as the basis for gene identification and plant improvement. These relatives are known to be a good source of genes for tolerance to

biotic or abiotic stress and the transgressive behavior of progeny derived from interspecific crosses has been demonstrated in several studies.

The evaluation of complex phenotypes relevant to agriculture remains a laborious activity that requires direct and sustained attention. To make the exploration of the phenotype-genotype relationship more efficient, a series of tools is required that will leverage the availability of genomic information and link it directly to field evaluation of germplasm resources.

We propose to develop a comprehensive toolkit of genetic and genomic resources that will allow breeders and geneticists to explore and more efficiently utilize wild relatives in crop improvement. Our approach involves the systematic introduction of foreign alleles from four different AA genome *O.* species into one elite, highly productive *O. sativa* varieties and to provide introgression lines (ILs) or sets of overlapping CSSLs as the basis for genetic analysis and applied plant breeding.

CSSLs are particularly valuable when complex, quantitatively inherited phenotypes are the breeding target. Because they represent permanent (inbred) genetic resources that can be easily replicated by seed and distributed to collaborators working in different environments. Each set of CSSLs consists of a relatively small number of lines that can be evaluated in replicated trials. They are constructed to provide maximum power of statistical analysis because each line can be compared to all others or may simply be compared to the recurrent parent, making it possible to extract a great deal of valuable information from a relatively small number of lines. For phenotypes that are difficult to measure, or require repeated evaluation over years and environments, the ability to focus quickly on a small number of lines is a critical component of success.

Complex phenotypes can be dissected genetically by evaluating and comparing CSSLs as a first pass. By linking the information about gene identity back to the Gramene database, we will be able to provide a dictionary of genes with known function that are contained in each of the CSSLs. This dictionary is also a key ingredient in enabling comparative approaches to the study of phenotype-genotype relationships. Once an association is established between a phenotype and a specific introgression line, it is possible to use advanced forward and reverse approaches to genetic analysis to mine down and identify the gene(s) that are directly involved (Yano, 2001; Jander et al., 2002), but it is also feasible to use the introgression line directly in a breeding program without knowing the identity of the gene(s) that are involved (Gur and Zamir, 2004). Genes known to be associated with a plant's response to water stress or adaptation to acid soils will be targeted for Single Nucleotide Polymorphisms (SNPs) development.

In addition to the targeted introgression of traits that can be identified phenotypically in the wild material, such as biotic or abiotic stress tolerance, it has been demonstrated that alleles hidden in low yielding, agronomically undesirable ancestors can enhance the productivity of many of the world's most important crop varieties. These yield-enhancing alleles are the basis of 'transgressive variation' and may confer an advantage in both favorable (irrigated) and unfavorable conditions (drought and weed competition) (Moncada et al., 2000; Gur and Zamir, 2004). Thus, the use of wild and exotic germplasm for CSSLs construction carries with it the possibility that favorable transgressive segregants will be identified, providing the basis for studies aimed at understanding the genetic basis of transgressive variation associated with resistance or tolerance to drought and acid soils.

III. Scientific activities (including tables and figures)

Universal Core Genetic Map of the rice genome

In order to facilitate the monitoring of the introgression process in the CSSL populations, we developed a *Universal Core Genetic Map* for rice. A total of 511 SSR markers

distributed as *anchors* were selected based on genomic sequence. Sixteen AA-genome accessions were selected to evaluate the polymorphism level for each anchor. A mean of 83.2% polymorphism was observed for the different interspecific combinations. The Universal Core Map is used for genetic mapping and genotype construction purposes. The URCGM Database was created and is available upon request.

Development of CSSL populations

O. glaberrima (Caiapo x MG12)

59 lines were chosen and backcrossed to the *O. sativa* parent Caiapo and selfed to obtain 59 BC_4F_2 families. 4200 individuals were planted in the field with the aim to identify plants bearing the target chromosomal fragment. The DNAs were bulked and they are currently being evaluated with SSRs. The BC3DH population is available to the scientific community upon request and has been already distributed to eleven partners. The BC4F3 lines will be available by 2009.

O. meridionalis

Foreground and background selection of 516 BC2F1 lines led to the selection of 60 lines that were subsequently backcrossed. Six seeds from each of the 60 lines were sown to generate a population of 360 BC3F1 plants. Foreground pre-selection is currently in progress to select candidate BC3F1 lines. We will develop double haploid lines from the pre-selected BC3F1 lines through anther culture at CIAT.

O. rufipogon

Similar results were obtained for this population, except that 600 BC2F1 plants were analyzed. The selected BC3F1s will be processed for anther culture during the second semester of 2008.

O. barthii

214 BC2F1 lines bearing 54 *O. barthii* segments were selected among the 600 genotyped plants. For backup purpose, the BC3F1 crosses were made from all BC2F1 plants. The number of BC_3F_1 seeds produced per plant varied from 5 to 62. They are currently being sowed and the plants will be checked for the presence of target *O. barthii* segments.

O. glumaepatula

153 BC_3F_1 plants were selected and the BC_4F_2 seeds are now available. 142 BC_2F_2 plants were evaluated for yield-related traits at an experimental field in Porangatu, Goias, Brazil, under two conditions, one fully-irrigated and one under water stress. A QTL analysis was performed and eight QTLs were detected in both conditions, from which four were detected in the first treatment and four under water stress. A scientific paper will be submitted with the results from this experiment.

Drought stress screenings

54 CSSLs from the cross IR64 x TOG5681 (*O. glaberrima*) was screened in hydromorphic soil during the dry seasons of 2006 and 2007 at WARDA trial fields in Cotonou, Benin. In 2006, percentage yield loss under drought varied from 3 to 88%. Yield loss due to drought in 2007 was more severe with a mean of 78%, ranging from 44 to 100%. Several CSSLs *yielded higher* than IR64 under both drought stress and continuous irrigation. This may mean that *O. glaberrima* has contributed several genes in this cross that either alone or through epistatic effects can increase grain yield of rice in these conditions. Nine transgressive genotypes were found to yield *consistently higher* than the average yield under drought stress in both years of screening. The SSR data available for this population will help in identifying the genomic regions associated to the drought tolerance.

Next steps

Molecular genotyping of the BC3DH of BC3F2 lines of the *O. barthii*, *O. meridionalis* and *O. rufipogon* will be done at Cornell University during the first semester of 2009. We may switch from the SSR to the SNP (Illumina technology available at Cornell), which offers a better marker density and higher throughput. BC4F2 seeds will also be prepared and made available to the scientific community. Additionally, a set of SNPs is currently being optimized at CIAT (Luminex technology), which will help in saturating the current SSR genetic map. The application and some alterations of the breeding strategies used during the course of this project will help breeders to improve future breeding programs. Further field trials for different important agronomic traits could permit the detection of several QTLs of significant relevance for cultivar improvement.

IV. Deviations from the work plan

There is no major deviation from the work plan. However, a few minor deviations need to be mentioned:

- The construction of the Universal Core Map was not planned in the original project. The construction of this tool was decided later on to facilitate the construction of CSSLs and the related activities were partially financed with external sources of funding. The results and outputs issued from this activity, we believe, represent a substantial added value to the project.
- Also, the decision was made to build a new *O. sativa* x *O. meridionalis* genetic map from a new population was taken due to the uncertainty in some genotypes identification stamps. We decided to redo the job in order to discard any even minor source of mistake in the subsequent selection of chromosomal segments.
- The same strategy was adopted for the two *O. sativa* x *O. glaberrima* populations in order to recover lost fragments and validate the genetic map.
- Overall, we estimate that we're about six month late on the CSSL development activity. This was mainly due to difficulties in handling some logistical aspects, i.e., re-directing some initial activities from WARDA to CIAT, or delays related to visa issues or other logistical issues.
- Drought stress screenings: it has not been possible to carry out drought stress screenings at Fedearroz experimental station in good conditions. However, a collaboration (external to this project) with Thaura Ghneim's group (IVIC, Venezuela) that aims to screen 93 lines for various physiological parameters under drought stress from the Caiapo x MG12 cross should adequately complete the field screening, although the results cannot be considered as a GCP product.

V. Data Availability

Various data sets have been produced so far:

- Genotyping data (SSR markers) of 6 interspecific BC1F1 populations,
- Partial genotypic data of 5 BC2F2 populations
- Genotypic data for two advanced pre-CSSL populations from *O. sativa* x *O. glaberrima* crosses
- Six interspecific genetic maps aligned to the Nipponbare sequence,
- Various sets of phenotypic data collected on two *O. sativa* x *O. glaberrima* ILs populations,
- Interspecific SNP locations and primers,
- SSR polymorphism for a collection of cultivated and wild rice accessions,
- SSR genomic location of the Universal Core Genetic Map.

All of these data will be posted on the public GCP database after their publication, if applicable.

Publications:

Orjuela J, Garavito A, Bouniol M, Tranchant C, Wilson G, McCouch S.R, Tohme J, Ghesquière A, Lorieux M. 2007. Universal Rice Genetic Core Map. Presentation given at the VI Encuentro Latinoamericano de Biotecnología Agropecuaria REDBIO 2007, Viña del Mar, Chile, October 22-26 2007.

Gutiérrez AG, Alvarez MF, Martínez CP, Giraldo OX, Tohme J and Lorieux M. 2007. Development of Chromosome Segment Substitution Lines in two interspecific *Oryza sativa* x *O. glaberrima* populations. Poster presented at the VI Encuentro Latinoamericano de Biotecnología Agropecuaria REDBIO 2007, Viña del Mar, Chile, October 22-26 2007.

VI. Conclusions

Although some deviations have occurred from the initial work plan, the main goals for this project have been achieved so far. As a summry:

· We pursued the development of six populations of CSSLs (Chromosome Segment Substitution Lines) that bear introgressions from the AA-genome species *O. glaberrima, O. barthii, O. meridionalis, O. rufipogon* and *O. glumaepatula*. These populations will constitute a valuable tool for genetic analyses and will allow us to identify key genomic regions that are associated to agronomically important traits. The populations are now at the BC3 stage and they should be ready for use by middle 2009.

· We are almost finished with the development of a *Universal Core Genetic Map* for rice. This map has already been demonstrated as a very useful tool to help at designing introgression populations, particularly in the case of interspecific crosses. It is based on microsatellite markers that we discovered and choose with the help of several bioinformatic packages, including some that we develop at CIAT.

· A database, *Rice Universal Core Map*, was developed and is available upon request to the author (contact: m.lorieux@cgiar.org). This database aims to provide means to easily and quickly choose a series of genetic markers to be used to genotype a population derived from a specific cross.

· Five first-generation backcross segregating populations were genotyped and five interspecific genetic maps were developed from these data. These maps will be useful to assess the recombination rates for every wild species we use and will facilitate the localization of important genes or QTLs. All the map we generated were based on the Universal Core Genetic Map.

· In order to fully exploit the information given by the genetic mapping analyses carried out using crosses that involve the *O. glaberrima* species, we collaborated with the Arizona Genomics Institute to develop a *library of Bacterial Artificial Chromosomes* (BAC) for this species. The library is available to the international community of plant genomicists, and it will constitute the basis of *positional cloning* approaches to identify and characterize important genes for *O. glaberrima* (this work was also supported by USAID funds).

· In parallel of this project, we designed a computer program that helps geneticists at creating CSSL populations. The program is called *CSSL Finder* at is available for download as a freeware at http://mapdisto.free.fr/.

· A bioinformatic tool to facilitate the discovery of single-nucleotide polymorphisms (SNPs) was set up.

· The first SNP validation tests were successful and are encouraging. They let us hope that we will soon count with a tool for fast and reliable genotyping system that will make our marker-aided breeding programs more efficient.

· Seven students and four research assistants were trained.
· Four students from Africa and Latin America do shuttle research between their respective centers and Cornell University.
· The international collaboration between several ARIs, CG centers and NARS was strengthened.
· Several publications are in preparation.

We expect the outputs of this project to provide very useful tools to the scientific community, for gene identification in wild species and pre-breeding purposes.

It is worth to mention that this project is now strongly connected to the OMAP project (Rod Wing, Arizona Genomics Institute). Indeed, BAC libraries for four of the parental wild accessions involved in this project have been or will be constructed soon at AGI. Also, LGDP (Olivier Panaud's group) have developed a BAC library for the TOG 5681 (*O. glaberrima*) accession. Having this genomic resource available will add a great value to the populations of CSSLs developed here, as rice geneticists will have access to complete kits for gene identification and positional cloning in wild rice species.

VII. References

Ahn SN, JP Suh, CS Oh, SJ Lee, and HS Suh (2002) Development of introgression lines of weedy rice in the background of Tongil-type rice. Rice Genetics Newsletter 19:14

Albar, L., M.-N. Ndjiondjop, Z. Essibak, A. Berger, A. Pinel, M. Jones, D. Fargette and A. Ghesquière, 2003. Fine genetic mapping of a gene required for rice yellow mottle virus cell to-cell movement. Theo. Appl. Genet 107: 371–378.

Aluko G., C. Martinez, J. Tohme, C. Castano, C. Bergman, J.H. Oard 2004. QTL Mapping of Grain Quality Traits From the Interspecific Cross Oryza sativa x O. glaberrima .Theor.Appl. Genet.109:630-639

Brondani, C, N. Rangel, V. Brondani, and E. Ferreira, 2002. QTL mapping and introgression of yield-related traits from Oryza glumaepatula to cultivated rice (Oryza sativa) using microsatellite markers. Theo. Appl. Genet 104: 1192–1203

Doi K, Sobrizal, K Ikeda PL Sanchez, T Kurakazu, Y Nagai and A Yoshimura (2003) Developing and evaluating rice chromosome segment substitution lines. Proc. IRRI Conference, Sept. 16-19, 2002. Beijing, China. p. 275-287

Enriquez, EC, DS Brar, TL Rosario, M Jones, GS Khush (2000) Production and characterization of doubled haploids from anther culture of the F1s of Oryza sativa L. x O. glaberrima Steud. Rice Genet News 17:26-27

Eshed, Y. and D. Zamir (1995). "An introgression line population of Lycopersicon pennellii in the cultivated tomato enables the identification and fine mapping of yield-associated QTL." Genetics 141: 1147-1162.

Frey KJ, EG Hammond and PK Lawrence (1975) Inheritance of oil percentage in interspecific crosses of hexaploid oats. Crop Sci 15:94-95

Ghesquière, A., J. Séquier, et al. (1997). "First steps towards a rational use of African rice, Oryza glaberrima, in rice breeding through a 'contig line' concept." Euphytica 96: 31-39.

Gur A and D Zamir (2004) Unused natural variation can lift yield barriers in plant breeding. Public Library of Science (in press)

Hash CT (1999) Contiguous segment substitution lines: New tool for elite pearl millet hybrid parental lines enhancement http://www.dfid-psp.org/projects/7382.htm

Jander G, SR Norris, SD Rounsley, DF Bush, IM Levin and RL Last (2002) Arabidopsis map-based cloning in the post-genomic era. Plant Phys 129:440-450

Koumproglou R, TM Wilkes, P Townson, XY Wang, J Beynon, HS Pooni, HJ Newbury (2002) STAIRS: A new genetic resource for functional genomic studies of Arabidopsis. The Pl. J. 31:355-364

M. Lorieux, G. Reversat, S. X. Garcia Diaz, C. Denance, N. Jouvenet, Y. Orieux, N. Bourger, A. Pando-Bahuon, A. Ghesquière. 2003. Linkage mapping of Hsa-1Og, a resistance gene of African rice to the cyst nematode, Heterodera sacchari. Theo. Appl. Genet 107:691-696

Moncada, M.P., C.P. Martínez, J. Tohme, E. Guimaraes, M. Chatel, J. Borrero, H. Gauch, Jr. and S.R. McCouch, 2001. Quantitative trait loci for yield and yield components in an Oryza sativa x Oryza rufipogon BC2F2 population evaluated in an upland environment. Theor Appl Genet 102:41–52

Monforte AJ, Tanksley SD (2000) Development of a set of near isogenic and backcross recombinant inbred lines containing most of the Lycopersicon hirsutum genome in a L. esculentum genetic background: A tool for gene mapping and gene discovery. Genome 43: 803–813

Pestsova EG, Borner A, Roder MS (2002) Development of wheat D-genome introgression lines assisted by

microsatellite markers. In: Hernández P, Moreno MT, Cubero JI, Martin A, editors. Proceedings of the Fourth International Triticeae Symposium; 2001 September; Cordoba, Spain. Published in Cordoba by Junta de Andalucia. Consejera de Agricultura y Pesca, pp. 207–210.

Pestsova EG, Borner A, Roder MS (2003) Application of microsatellite markers to develop Triticum aestivum–Aegilops tauschii defined introgression lines. In: Borner A., Snape JW, Law CN, editors. Proceedings of the Twelfth International EWAC (European Wheat Aneuploid Co-operative) Workshop; 2002 July 1–6; Norwich, United Kingdom, John Innes Centre, publisher. pp 32-35

Ramsay LD, Jennings DE, Bohuon EJR, Arthur AE, Lydiate DJ, et al. (1996) The construction of a substitution library of recombinant backcross lines in Brassica oleracea for the precision mapping of quantitative trait loci (QTL). Genome 39: 558–567

Rick CM (1976) Natural variability in wild species of Lycopersicon and its bearing on tomato breeding. Agraria 30: 249–510

Rick CM (1983) Conservation and use of exotic tomato germplasm. Conservation and Utilization of exotic germplasm to improve varieties. The 1983 Plant Breeding Research Forum: Conservation and utilization of exotic germplasm to improve varieties. Brown WL, editor. August 9-11, 1983, Heber Springs, Arkansas: Des Moines, Iowa, Pioneer Hybrid Int'l Inc., publisher. pp. 147–167.

Sears E (1956) The transfer of leaf-rust resistance from Aegilops umbellulata to wheat. Brookhaven Symp Biol 9: 1–22

Sobrizal K, Ikeda K, Sanchez PL, Doi K, Angeles ER, et al. (1996) Development of Oryza glumaepatula introgression lines in rice, O. sativa L. Rice Genet Newsl 16: 107

Sun et al 2001 Theo. Appl. Genet 102:157-162

Tanksley SD, McCouch SR (1997) Seed banks and molecular maps: Unlocking genetic potential from the wild. Science 277: 1063–1066

Thomson M, Tai T, McClung A, Xai X-H, Hinga M, Lobos K, Xu Y, Martinez P, McCouch S (2003) Mapping quantitative trait loci for yield, yield components and morphological traits in an advanced backcross population between Oryza rufipogon and the Oryza sativa cultivar Jefferson. Theor Appl Genet 107:479–493

Xiao J, Li J, Grandillo S, Ahn SN, Yuan L, et al. (1998) Identification of trait-improving quantitative trait loci alleles from a wild rice relative, Oryza rufipogon. Genetics 150: 899–909.

Yano M (2001) Genetic and molecular dissection of naturally occurring variation. Curr Opin Plant Biol 4: 130–135

Zamir D (2001) Improving plant breeding with exotic genetic libraries. Nat Rev Genet 2:983-9

Zhu, Y.-L, Q.-J. Song, D.L. Hyten, C.P. Van Tassell, L.K. Matukumalli, D.R. Grimm, S.M. Hyatt, E.W. Fickus, N.D. Young, and P.B. Cregan. 2003. Single nucleotide polymorphisms (SNPs) in soybean. Genetics 163: 1123-1134

Appendix A. Activities, Quantifiable Outputs, and Key Products

Project Title: Exploring Natural Genetic Variation: Developing Genomic Resources and Introgression Lines for Four AA Genome Rice Relatives	
Principal Investigator/Institute: M.Lorieux (IRD/CIAT); J. Tohme (CIAT)	
Objective 1: Creation of six CSSL populations using AA genome rice species as donors	
Tabla 1. Activities	**Quantifiable Outputs**
1. Develop CSSL populations using wild rice species as donors	1. Four CSSL populations of interspecific lines sharing the same genetic background (*O. sativa* tropical japonica cv. Curinga) and bearing introgressed genome segments from the AA genome wild species *O. meridionalis*, *O. barthii*, *O. rufipogon* and *O. glumaepatula*
2. Complete the development of two *O. sativa* x *O. glaberrima* CSSL populations	2. Two CSSL populations of interspecific lines and bearing introgressed genome segments from the AA genome cultivated species *O. glaberrima*
Objective 2: Development of a Universal Core Genetic Map for Rice	
Tabla 2. Activities	**Quantifiable Outputs**
1. Define and validate a series of *anchors* along the rice genome that provide wide polymorphism across the wild rice species (AA genome)	3. A set of 125 anchors represented by three SSR loci each. At least one SSR per anchor is polymorphic for any of the worked crosses of this project 4. A database of polymorphic markers eases the definition of the best set of SSRs to be used for a specific cross
Objective 3: Development of a SNP kit for genotyping	
Tabla 3. Activities	**Quantifiable Outputs**
1. Search for SNPs in stress-related genes, define primers and validate them for polymorphisms in the worked crosses	5. A set of 125 SNP markers based on the single base extension technology. These markers can be multiplexed in batches of 30 markers. The polymorphisms are validated for all the six worked populations
Objective 4: Analysis of *O. sativa* x *O. glaberrima* CSSLs for drought response	
Tabla 4. Activities	**Quantifiable Outputs**
1. To screen the *O. sativa* x *O. glaberrima* CSSLs for drought response in order to identify alleles from *O. glaberrima*	6. Lines with superior behavior under stress (drought) conditions 7. Chromosomal locations of QTLs for drought tolerance from *O. glaberrima*

Key Products Developed by the Project (those that you think have the biggest potential impact – please limit to 5):

1. Four interspecific rice CSSL populations using wild species as donors (tool for gene identification and pre-breeding)
2. Two interspecific rice CSSL populations using the african cultivated species as donor
3. Universal Core Genetic Map for Rice
4. Kit of SNPs for rice genotyping
5. Chromosomal location of drought response QTLs

Appendix B. Timeline

		Year 1												Year 2												Year 3											
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	3
Objective 1	Activity 1																																				
	Activity 2																																				
Objective 2	Activity 1																																				
Objective 3	Activity 1																																				
Objective 4	Activity 1																														6						

Year 4											
37	38	39	40	41	42	43	44	45	46	47	48
											1
					2						
	3	4									
					5						
					7						

Quantifiable Outputs:

1. Four CSSL populations of interspecific lines sharing the same genetic background (*O. sativa* tropical japonica cv. Curinga) and bearing introgressed genome segments from the AA genome wild species *O. meridionalis*, *O. barthii*, *O. rufipogon* and *O. glumaepatula*
2. Two CSSL populations of interspecific lines and bearing introgressed genome segments from the AA genome cultivated species O. glaberrima
3. A set of 125 anchors represented by three SSR loci each. At least one SSR per anchor is polymorphic for any of the worked crosses of this project
4. A database of polymorphic markers eases the definition of the best set of SSRs to be used for a specific cross
5. A set of 125 SNP markers based on the single base extension technology. These markers can be multiplexed in batches of 30 markers. The polymorphisms are validated for all the six worked populations
6. Lines with superior behavior under stress (drought) conditions
7. Chromosomal locations of QTLs for drought tolerance from O. glaberrima

Appendix E. Data Production and Availability Section Format

For each proposed project, please provide the information below:

Various data sets have been produced so far:
- Genotyping data (SSR markers) of 6 interspecific BC1F1 populations,
Format: Mapmaker/EXP compatible txt file
- Partial genotypic data of 5 BC2F2 populations
Format: CSSL Finder compatible Excel file
- Genotypic data for two advanced pre-CSSL populations from *O. sativa* x *O. glaberrima* crosses
Format: CSSL Finder compatible Excel file
- Six interspecific genetic maps aligned to the Nipponbare sequence
Format: MapDisto output txt file (Cmap-compatible format),
- Various sets of phenotypic data collected on two *O. sativa* x *O. glaberrima* ILs populations,
Format: Excel files
- Interspecific SNP locations and primers,
Format: Excel files
- SSR polymorphism for a collection of cultivated and wild rice accessions,
Format: Universal Core Map database (Excel-compatible)
- SSR genomic location of the Universal Core Genetic Map
Format: Universal Core Map database (Excel-compatible)

All of these data will be published on the GCP central database as soon as they have been valorised in publications.