



Generation Challenge Programme

CULTIVATING PLANT DIVERSITY FOR THE RESOURCE-POOR

2008

Project abstracts



## **Generation Challenge Programme**

### **Project abstracts**

**July 2008**

#### **Generation Challenge Programme (GCP)**

Hosted by CIMMYT

(Centro Internacional de Mejoramiento de Maíz y Trigo;  
the International Maize and Wheat Improvement Center)

Mailing address:  
Apdo Postal 6641  
06600 Mexico, DF Mexico

Physical address:  
Km 45 Carretera México-Veracruz  
El Batán, Texcoco, México, CP 56130

Tel: +52 55 5804 2004  
Fax: +52 55 5804 7558

Email: [generationcp@cgiar.org](mailto:generationcp@cgiar.org) or [info@generationcp.org](mailto:info@generationcp.org) [www.generationcp.org](http://www.generationcp.org)

© The Generation Challenge Programme of the Consultative Group on International Agricultural Research.  
Fair use: you are free to quote from, or reproduce, sections of this publication without further permission,  
provided the Generation Challenge Programme (GCP) of the Consultative Group on International  
Agricultural Research (CGIAR) is acknowledged as the source. We would also appreciate receiving a copy  
of the publication.

Correct Citation: Generation Challenge Programme, 2008. Generation Challenge  
Programme 2008 Project abstracts. Mexico D.F.: Generation Challenge Programme

*Editorial note: Abstracts are published as presented by the authors, with no further  
editing by the Generation Challenge Programme.*

# TABLE OF CONTENTS

Acronyms and abbreviations .....	vii
COMPETITIVE PROJECTS .....	1
<b>Subprogramme 1: Genetic diversity of global genetic resources</b> .....	1
1. G3005.10: Exploring natural genetic variation: developing genomic resources and introgression lines for four aa-genome rice relatives .....	1
2. G3005.13: Development of informative markers through association mapping in maize to improve drought tolerance in cereals .....	3
3. G3005.14: Characterisation of genetic diversity of maize populations: documenting global maize migration from the center of origin .....	5
4. G3005.17: Allele mining based on non-coding regulatory SNPs in barley germplasm .....	7
5. G3007.01: Ibridges: Interspecific bridges that give full access to the African rice allele pool for enhancing drought tolerance of Asian rice .....	9
6. G3007.02: Genomic dissection of tolerance to drought stress in wild barley .....	10
<b>Subprogramme 2: Genomics towards gene discovery</b> .....	13
7. G3005.01: Identifying genes responsible for failure of grain formation in rice and wheat under drought .....	13
8. G3005.02: Revitalizing marginal lands: discovery of genes for tolerance of saline and phosphorus deficient soils to enhance and sustain productivity .....	15
9. G3005.08: Targeted discovery of superior disease QTL alleles in the maize and rice genomes .....	17
10. G3005.11: Functional genomics of cross-species resistance to fungal diseases in rice and wheat (CEREALIMMUNITY) .....	19
11. G3005.15: Determination of a common genetic basis for tissue growth rate under water-limited conditions across plant organs and genomes .....	21
12. G3005.16: Isolation and characterisation of aluminum tolerance genes in the cereals: an integrated functional genomic, molecular genetic and physiological analysis .....	23
13. G3007.03: Development of genomics resources for molecular breeding of drought tolerance in cassava .....	26
14. G3007.06: Genetic dissection of drought adaptive mechanisms in bread and durum wheat through large scale phenotyping methodologies (SP2) .....	28
<b>Subprogramme 3: Trait capture for crop improvement</b> .....	30
15. G3005.03: Identifying the physiological and genetic traits that make cassava one of the most drought tolerant crops .....	30
16. G3005.05: Unlocking the genetic diversity in peanut's wild relatives with genomic and genetic tools .....	31
17. G3005.06: Marker development and marker-assisted selection for Striga resistance in cowpea .....	33
18. G3005.09: Development of low-cost technologies for pyramiding useful genes from wild relatives of cassava into elite progenitors .....	36
19. G3005.12: Drought tolerant rice cultivars for North China and South/Southeast Asia by highly efficient pyramiding of QTLs from diverse origins .....	37

20.	G3007.04: Tailoring superior alleles for abiotic stress genes for deployment into breeding programmes: A case study based on association analysis of AltSB, a major aluminum tolerance gene in sorghum (ALTSORGHUM).....	40
21.	G3007.05: Detecting and fine-mapping QTLs with major effects on rice yield under drought stress for deployment via marker-aided breeding .....	42
	<b>Subprogramme 4: Bioinformatics and crop information systems</b> .....	44
22.	G3005.04: An eco-physiological-statistical framework for the analysis of G×E and QTL×E as occurring in abiotic stress trials, with applications to the CIMMYT drought stress programmes in tropical maize and bread wheat .....	44
	COMMISSIONED PROJECTS .....	46
	<b>Subprogramme 1: Genetic diversity of global genetic resources</b> .....	46
23.	G4005.01.03 (1c): Completing genotyping of composite germplasm set of sorghum .....	46
24.	G4005.01.04 (1d): Completing genotyping of composite germplasm set of chickpea .....	47
25.	G4005.03.02 (3b): Molecular characterisation of pigeonpea ( <i>Cajanus cajan</i> ) composite collection.....	49
26.	G4005.03.06 (3f): Molecular characterisation of groundnut ( <i>Arachis hypogaea</i> L.) composite collection .....	51
27.	G4005.05 Assessing Eco-tilling as a methodology for targeted genotyping and SNP discovery.....	52
28.	G4005.06: Supporting emergence or reference drought tolerance phenotyping centers - Drought phenotyping network .....	54
29.	G4005.07: Whole plant physiology modelling (WPM).....	56
30.	G4006.01: Developing strategies for allele mining within large collections ABSTRACT NOT SUBMITTED.....	58
31.	G4006.02: A dataset on allele diversity at orthologous candidate genes in GCP crops (ADOC).....	58
32.	G4006.03: SNP analysis and the genetic diversity along the rice genome ( <i>HaplOryza</i> ).....	60
33.	G4006.05: Development of a composite collection and the genotyping of faba bean .....	62
34.	G4006.06: Genotyping of composite collection of finger millet [ <i>Eleusine coracana</i> (L.) Gaertn] .....	64
35.	G4006.30: Genotyping of composite collection of foxtail millet [ <i>Setaria italica</i> (L.) P. Beauv.] .....	66
36.	G4006.31: Development and genotyping of composite collection of pearl millet ( <i>Pennisetum glaucum</i> (L.)R. Br.) .....	67
37.	G4007.01: Genotyping validation of the GCP reference sets.....	69
38.	G4008.01: Population development through multiparent advanced generation inter-crosses (MAGIC) among diverse genotypes to facilitate gene discovery for various traits in rice .....	70
39.	G4008.02: Phenotyping sorghum reference set for drought tolerance .....	72
40.	G4008.03: Precision phenotyping of the GCP spring wheat reference sample for drought.....	73
41.	G4008.05 Connecting performance under drought with genotypes through phenotype associations .....	74
42.	G4008.33: Drought tolerance phenotyping of the GCP maize inbred line reference set.....	76

43.	G4008.42: Developing DArT markers for several crops in the GCP .....	77
<b>Subprogramme 2: Genomics towards gene discovery</b> .....		80
44.	G4007.02: Validation of drought-response/resistance pathway genes by phenotypic analysis of mutants .....	80
45.	G4008.06: Single nucleotide polymorphism discovery, validation, and mapping in groundnut.....	82
46.	G4008.07: Improving molecular tools for pearl millet.....	83
47.	G4008.08: Transcriptome analysis of near-isogenic rice lines to identify expression signatures and gene combinations conferring tolerance to drought stress.....	84
48.	G4008.09: Development of genetic and genomic resources for breeding improved sweetpotato .....	85
<b>Subprogramme 3: Trait capture for crop improvement</b> .....		87
49.	G4005.20: Optimising marker-assisted breeding systems for drought tolerance in cereals through linkage of physiological and genetic models .....	87
50.	G4007.04: Association mapping of downy mildew resistance in elite maize inbred lines in Thailand.....	89
51.	G4007.05: Bridging genomics, genetic resources and breeding to improve wheat and barley production in Morocco.....	91
52.	G4007.06: Integrating marker-assisted selection into the conventional breeding procedure for improvement of wheat ( <i>Triticum aestivum L.</i> ) in the drought-prone areas of Northern China .....	93
53.	G4007.07: Marker-assisted selection for Sweetpotato Virus Disease (SPVD) resistance in Sweetpotato germplasm and breeding populations.....	94
54.	G4007.08 Integration of genomic tools with conventional screening for developing NERICA rice cultivars for West Africa.....	95
55.	G4007.23: Field evaluation of wheat-barley introgression lines under different water regimes.....	97
56.	G4007.24: Seed smoke treatment to favour germination under water stressed conditions.....	99
57.	G4008.10: Assessment of the breeding value of superior haplotypes for AltSB, a major Al tolerance gene in sorghum: linking upstream genomics to acid soil breeding in Niger and Mali (ALTFIELD).....	101
58.	G4008.11 Dry bean improvement and marker assisted selection for diseases and abiotic stresses in Central America and the Caribbean.....	102
59.	G4008.12 Linking genetic diversity with phenotype for drought tolerance traits through molecular and physiological characterisation of a diverse reference collection of chickpea .....	104
60.	G4008.13: Improving drought tolerance phenotyping in cowpea .....	105
61.	G4008.14: Breeding for drought tolerance with known gene information.....	106
62.	G4008.15: Developing Potato cultivars adapted to Southern Africa countries.....	108
63.	G4008.16: Speeding the development of salt-tolerant rice varieties through marker-assisted selection and their dissemination in salt-affected areas of Bangladesh.....	109
64.	G4008.17: Application of marker-assisted selection for <i>Striga</i> resistance in cowpea.....	111
65.	G4008.19: Incorporation of an MSV resistance gene in Mozambican maize varieties, mediated by use of MAS .....	113

66.	G4008.30: Development of a GCP phenotyping network.....	114
67.	G4008.34: Environmental assessment for phenotyping network .....	114
68.	G4008.41: Application and validation of the major QTL phosphate uptake 1 ( <i>Pup1</i> ) .....	116
<b>Subprogramme 4: Bioinformatics and crop information systems .....</b>		<b>118</b>
69.	G4005.22: Development of Generation CP domain models and ontology .....	118
70.	G4005.23: Implementation of web services technology in the Generation Challenge Programme Consortium.....	120
71.	G4005.27: High performance computing facilities for the GCP platform .....	121
72.	G4006.08: Data analysis support for existing projects in SP2 with emphasis on integrating results from microarray and mapping experiments.....	122
73.	G4006.16: Integrated GCP Informatics Platform .....	123
74.	G4006.17: GCP quality management and data quality Improvement .....	125
75.	G4006.35: Statistical support for the design and data analysis of GCP projects... ..	126
76.	G4007.09: Design and analysis of marker-trait association studies, with special attention for genetically challenging crops.....	128
77.	G4007.10: Support to GCP scientists regarding issues related to bioinformatics and data handling.....	130
78.	G4007.11: Further development and support for use of iMAS by NARS and other user communities .....	132
79.	G4007.12: Development of tools and technology to increase the functionality of the GCP Information Platform.....	133
80.	G4008.20: Management of the GCP Central Registry and the creation and maintenance of templates for data storage in repositories.....	134
81.	G4008.21: Large scale phylogenomic analyses to gene function prediction for GCP crops .....	136
82.	G4008.22: Methodology development for reconstruction of Genealogies based on Haplotypes related to geographic patterns (HaploPhyle: graphical haplotype network in the light of external data) .....	138
83.	G4008.31: Upgrading the quality and utility of GCP phenotyping data through the development of a database template to facilitate the storage of data in a crop specific database .....	139
<b>Subprogramme 5: Capacity-building and enabling delivery .....</b>		<b>141</b>
84.	G4005.53: The use of molecular markers in efficient crop improvement: Marker-assisted breeding learning module.....	141
85.	G4005.63: Interactive Resource Center & Helpdesk.....	141
86.	G4006.13: Targeting and impact analysis of Generation Challenge Programme technologies .....	143
87.	G4006.14: <i>Ex Ante</i> impact analysis of marker-assisted selection technologies supported by the Generation Challenge Programme (GCP).....	145
88.	G4006.36: Capacity-building and research project .....	147
89.	G4007.03: The ‘Community of Practices’ concept applied to rice production in the Mekong Region: Quick conversion of popular rice varieties with emphasis on drought, salinity and grain quality improvement .....	148
90.	G4007.13: Capacity-building <i>à la carte</i> 2007 .....	150
90.01	G4007.13.01: Capacity-building <i>à la carte</i> 2007–Capacity building for characterising maize for water stress tolerance at KARI–Katumani .....	150

90.02	G4007.13.02: Capacity-building <i>à la carte</i> 2007–Marker-aided development of nutritionally-enhanced cassava for Nigeria .....	152
90.03	G4007.13.03: Capacity-building <i>à la carte</i> 2007–Application of molecular tools for controlled wild introgression into cultivated germplasm in Senegal .....	154
90.04	G4007.13.04: Capacity-building <i>à la carte</i> 2007–Characterisation of maize germplasm found in Ghana, using the bulking technique .....	155
90.5	G4007.13.05: Capacity-building <i>à la carte</i> 2007–An integrated proteomics and genomics approach to discover salt tolerance genes .....	156
90.6	G4007.13.06: Capacity-building <i>à la carte</i> 2007–Enhancing capacity of ICABIOGRAD in phenotyping and molecular analysis to develop elite rice lines suitable to Indonesian uplands .....	159
91.	G4007.14: Fellowships and travel grants 2007 .....	160
92.	G4007.20: Managing the Generation Challenge Programme in a post-International Treaty world .....	161
93.	G4007.21: Genotyping Support Service .....	162
94.	G4007.21.03.02: Provision of Genotyping Support Services (GSS) .....	163
95.	G4007.22: GCP Workflow and Repository System (WMS) .....	165
96.	G4008.23: Statistical rules for defining characteristic genotype and marker sets .....	167
97.	G4008.24: From attractiveness to feasibility: A strategic assessment of the capacity to develop and adopt GCP technologies .....	167
98.	G4008.25: Advanced course on “Applied statistical methods in plant genomics”, Zaragoza 18-29 February 2008. Request for GCP grants to finance participation of candidates from developing countries .....	169
99.	G4008.26: A cassava breeding Community of Practice in Africa for accelerated production and dissemination of farmer-preferred cassava varieties resistant to pests and diseases .....	171
100.	G4008.27: Phenotyping course for drought related traits across tropical legumes–Concepts and practices .....	173
101.	G4008.28: Characterisation of maize diversity in Central Europe .....	175
102.	G4008.29: Characterisation of bean diversity in Central Europe .....	177
103.	G4008.35: Toolbox of available molecular markers useful for marker assisted selection in GCP crops .....	178
104.	G4008.36: Getting the focus right: Food crops and smallholder constraints .....	179
105.	G4008.37: PhD in Plant Breeding at the West Africa Centre for Crop Improvement .....	181
106.	G4008.38: Fellowships and travel grants 2008 .....	182
107.	G4008.39: Capacity-building <i>à la carte</i> 2008 .....	184
107.01.	G4008.39.01: Capacity-building <i>à la carte</i> 2008–Enhancing MAS capacity for salt-stress rice breeding in Bangladesh .....	185
107.02.	G4008.39.02: Capacity-building <i>à la carte</i> 2008–Improving capacity for phenotyping for abiotic and biotic stress in Burkina Faso .....	186
107.03.	G4008.39.03: Capacity-building <i>à la carte</i> 2008–Improving capacity for phenotyping for abiotic and biotic stress in Senegal .....	187
108.	G4008.40: Workshop on “Reference sets of food crop germplasm for international collaboration” .....	187



FOCUS PROJECTS .....	189
109. GCP/Rockefeller project G4005.69.01 (CB19a/RF-FS022): Developing and disseminating resilient and productive rice varieties for drought-prone environments in India .....	189
110. GCP/Rockefeller project G5005.69.02 (CB19b/RF-FS029): Pathway dissection and candidate gene identification of drought tolerance in rice by a forward genetics approach .....	191
111. GCP/Rockefeller project G4005.69.03 (CB19c/RF-028): Innovative and integrated approaches to improve the tolerance of maize to water-limited environments ....	193
112. GCP/Rockefeller project G4005.70.01 (CB20a/RF-FS091): Tapping crop biodiversity for the resource poor in East and Central Africa .....	195
113. GCP/Rockefeller project G4005.70.02 (CB20b/RF-FS090): Tapping crop biodiversity for the resource-poor in East and Central Africa .....	
ABSTRACT NOT SUBMITTED .....	197
114. G6007: Tropical Legumes I (TLI): Improving tropical legume productivity for marginal environments in sub-Saharan Africa .....	197
114.01. G6007.01: TLI Objective 1--Improve groundnut productivity in marginal environments of sub-Saharan Africa .....	198
114.02 G6007.02: TLI Objective 2--Improve cowpea productivity for marginal environments in sub-Saharan Africa .....	200
114.03 G6007.03: TLI Objective 3--Improving common bean productivity for marginal environments in sub-Saharan Africa .....	202
114.04 G6007.04: TLI Objective 4--Improved chickpea productivity for marginal environments in sub-Saharan Africa .....	204
114.05 G6007.05: TLI Objective 5--Development of cross-species resources for comparative genomics in tropical crop legumes .....	206
114.06 G6007.06: TLI Objective 6--Provide training and capacity-building for sub-Saharan African scientists .....	207

## ACRONYMS AND ABBREVIATIONS

ABC	Agricultural Biotechnology Center, Gödöllő, Hungary
ABRII	Agriculture Biotechnology Research Institute of Iran
ACCI	African Centre for Crop Improvement, South Africa
ACGT	African Centre for Gene Technologies, South Africa
ACPFPG	Australian Centre for Plant Functional Genomics Pty Ltd
AGRA	Alliance for a Green Revolution in Africa
AICMIP	The All-India Coordinated Pearl Millet Improvement Project
ARI	Agricultural Research Institute, Tanzania
ARI–HAS	Agricultural Research Institute of the Hungarian Academy of Sciences, Hungary
ARI–Naliende	Agricultural Research Institute–Naliende Research Station, Tanzania
BAU	Birsa Agricultural University, Kanke, Ranchi, India
Beca	Biosciences Eastern and Central Africa, Kenya
BGBM	Botanic Garden and Botanical Museum Berlin–Dahlem, Germany
BIOSS	Biomathematics and Statistics Scotland Research Institution, UK
BIOTEC	National Center for Genetic Engineering and Biotechnology, Thailand
Bioversity	Bioversity International
BIRRI	Bangladesh Rice Research Institute
CAAS	Chinese Academy of Agricultural Sciences
CARDI	Cambodian Agricultural Research and Development Institute
CAZRI	Central Arid Zone Research Institute, India
CERAAS	Centre d'étude régional pour l'amélioration de l'adaptation à la sécheresse, Senegal
CGN–WUR	Centre for Genetic Resources–Wageningen University and Research Centre, The Netherlands
CIAT	Centro Internacional de Agricultura Tropical (International Center for Tropical Agriculture)
CIHEAM–IAMM	Centre International de Hautes Etudes Agronomiques Méditerranéennes–Institut Agronomique Méditerranéen de Montpellier, France
CIHEAM–IAMZ	Centro Internacional de Altos Estudios Agronómicos Mediterráneos–Instituto Agronómico Mediterráneo de Zaragoza, Spain
CIMMYT	Centro Internacional de Mejoramiento de Maíz y Trigo (the International Maize and Wheat Improvement Center)
CINVESTAV	Centro de Investigación y de Estudios Avanzados, Mexico
CIP	Centro Internacional de la Papa (International Potato Centre)
CIRAD	Centre de coopération internationale en recherche agronomique pour le développement, France
CNG	Centre National de Génotypage, France
CNPMF	Centro Nacional de Pesquisa de Mandioca e Fruticultura Tropical (Biotechnology Research Unit of EMBRAPA–Cassava and Tropical Fruits)
CRI	Crops Research Institute, Ghana
CRIL	Crop Research Informatics Laboratory
CRRI	Central Rice Research Institute, India
CRURRS	Central Rainfed Upland Rice Research Station, India
CSIR	Council for Scientific and Industrial Research, Ghana
CSIRO	Commonwealth Scientific and Industrial Research Organisation, Australia
CSU	Colorado State University

DArT P/L	Diversity Arrays Technology Pty Ltd
DARTS	Department of Agricultural Research & Technical Services, Malawi
DR4D	Department of Research for Development, Zimbabwe
DWR	Directorate of Wheat Research, India
Eger–Hungary	Department of Plant Sciences and Plant Physiology, Eszterházy Károly College, Eger, Hungary
EIAR	Ethiopian Institute of Agricultural Research, Ethiopia
EMBRAPA	Empresa Brasileira de Pesquisa Agropecuária (Brazilian Agricultural Research Corporation)
ETH	Eidgenössische Technische Hochschule, (Swiss Federal Institute of Technology), Zürich, Switzerland
GCP	Generation Challenge Programme
HAAS	Institute of Dry Farming, Hebei Academy of Agricultural Sciences, China
IAO	Istituto Agronomico per l'Oltremare, Italy
IARI	Indian Agricultural Research Institute
IBONE	Instituto de Botánica del Nordeste, Argentina
ICABIOGRAD	Indonesian Center for Agricultural Biotechnology and Genetic Resources Research and Development
ICAR	Indian Council of Agricultural Research
ICARDA	International Centre for Agricultural Research in the Dry Areas
IDA	Indonesian Department of Agriculture
IEB	Institute of Experimental Botany, Czech Republic
IER	Institut d'Economie Rurale, Mali
IFPRI	International Food Policy Research Institute
IFSSA	Indian Foundation Seed and Services Association
IGAU	Indira Gandhi Agricultural University, India
IGD	Institute for Genomic Diversity, Cornell University, USA
IGKV	Indira Gandhi Krishi Vidyalaya, India
IIAM	Institute of Agricultural Research of Mozambique
IIPR	Indian Institute of Pulse Research
IITA	International Institute of Tropical Agriculture
ILRI	International Livestock Research Institute
INCA	Instituto Nacional de Ciencias Agrícolas, Cuba
INERA	Institut de l'Environnement et de Recherches Agricoles, Burkina Faso
INIA	Instituto de Investigaciones Agropecuarias, Chile
INIA	Instituto Nacional de Investigación Agropecuaria, Uruguay
INIFAP	Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias, Mexico
INRA	Institut National de la Recherche Agronomique, France
INRA	Institut National de la Recherche Agronomique, Morocco
INRAN	Institut National de Recherches Agronomiques du Niger
INTA	Instituto Nacional de Tecnología Agropecuaria, Argentina
IPB	Institut Pertanian Bogor, Bogor Agriculture University, Indonesia
IPK	Institute for Plant Genetics and Crop Plant Research, Germany
IRAD	Institut de la recherche agronomique pour le développement, Cameroon
IRC	Interactive Resource Center
IRD	Institut de Recherche Pour le Développement, France
IRRI	International Rice Research Institute
ISRA	Institut Sénégalais de Recherches Agricoles, Senegal
JIC	John Innes Centre, UK
JIRCAS	Japan International Research Center for Agricultural Sciences
KARI	Kenya Agricultural Research Institute

LAAS	Luoyang Academy of Agricultural Sciences, China
LGDP	Laboratoire Génome et Développement des Plantes (Plant Genome and Development Laboratory), University of Perpignan, France
LZARDI	Lake Zone Agricultural Research and Development Institute, Tanzania
NAARI	Namulonge Agricultural Research Institute, Uganda
NaCRRI	National Crop Resources Research Institute, Uganda
NAFRI	National Agriculture and Forestry Research Institute, Laos
NAU	Nanjing Agricultural University, China
NCGR	National Center for Genome Resources, USA
NCSRC	National Corn and Sorghum Research Center, Thailand
NCSU	North Carolina State University, USA
NDUAT	Narendra Dev University of Agriculture and Technology, Faizabad, India
NIAS	National Institute of Agrobiological Sciences, Japan
NMRI	National Maize Research Institute, Vietnam
NRCPB	National Research Centre on Plant Biotechnology, India
NRCRI	National Root and Tuber Crops Research Institute, Nigeria
NSF	National Science Foundation, USA
NSFCRC	Nakhon Sawan Field Crops Research Center, Thailand
NU	Ningxia University, China
NWSUAF	Northwest Sci-tech University of Agriculture and Forestry, China
OSU	Oregon State University, USA
PAU	Punjab Agricultural University, India
PhilRice	Philippine Rice Research Institute
RAU	Rajasthan Agricultural University, India
RGDU	Rice Gene Discovery Unit, Thailand
RIKEN	Japanese abbreviation for ‘Rikagaku Kenkyūsho’: The Institute of Physical and Chemical Research
SAAS	Shanxi Academy of Agricultural Sciences, China
SABRN	Southern Africa Bean Research Network
SARI	Savannah Agricultural Research Institute, Ghana
SAU	Sichuan Agriculture University, China
SCRI	Scottish Crop Research Institute, UK
SIRDC	Scientific and Industrial Research and Development Centre, Zimbabwe
TNAU	Tamil Nadu Agricultural University, India
UAS	University of Agricultural Sciences, India
UBU	Ubon Ratchatani University, Thailand
UCB	Universidade Católica de Brasília, Brazil
UC–Berkley	University of California, Berkley, USA
UC–Davis	University of California, Davis, USA
UC–Riverside	University of California, Riverside, USA
UGA	University of Georgia, Athens, USA
UKZN	University of KwaZulu–Natal, South Africa
USDA–ARS	United States Department of Agriculture–Agricultural Research Service
UVA	University of Virginia, USA
WACCI	West Africa Centre for Crop Improvement, Ghana
WARDA	Africa Rice Center
WUR	Wageningen University and Research Centre, The Netherlands
YAAS	Yunnan Academy of Agricultural Sciences, China

**This publication has been published in black-and-white. For colour images, please contact the appropriate Principal Investigator.**

# COMPETITIVE PROJECTS

## ***Subprogramme 1: Genetic diversity of global genetic resources***

### **1. G3005.10: Exploring natural genetic variation: developing genomic resources and introgression lines for four aa-genome rice relatives**

*January 2005–December 2008*

#### **Principal Investigators**

Mathias Lorieux, Agropolis–IRD/CIAT; [mathias.lorieux@ird.fr](mailto:mathias.lorieux@ird.fr)

Joe Tohme, CIAT, Colombia; [j.tohme@cgiar.org](mailto:j.tohme@cgiar.org)

CIAT – AA 6713 – Cali – Colombia; Tel: +57 2 445 00 00; Fax: +57 2 445 00 73

#### **Collaborating institutions and scientists**

- Cornell University, USA: Susan R. McCouch
- CNPAF–EMBRAPA: Claudio Brondani,
- WARDA, Côte d’Ivoire: Baboucarr Manneh & Marie Noelle Ndjioudjop
- CIAT: César P. Martinez
- Fedearroz, Colombia: Miguel Diago Ramirez

#### **Context**

Wild relatives of modern crop species represent a valuable source of under-utilised genetic variation and represent an invaluable source of genetic information. We propose to develop libraries of interspecific lines called Chromosome Segment Substitution Lines (CSSLs), from different rice relatives. Generating such resources will allow utilizing *natural genetic diversity* as a permanent genetic resource for both breeding and genomics-based research.

#### **Findings and implications**

##### ***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<sub>4</sub>F<sub>2</sub> 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 BC<sub>2</sub>F<sub>1</sub> 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 BC<sub>3</sub>F<sub>1</sub> plants. Foreground pre-selection is currently in progress to select candidate BC<sub>3</sub>F<sub>1</sub> lines. We will develop double haploid lines from the pre-selected BC<sub>3</sub>F<sub>1</sub> lines through anther culture at CIAT.

### **O. rufipogon**

Similar results were obtained for this population, except that 600 BC<sub>2</sub>F<sub>1</sub> plants were analyzed. The selected BC<sub>3</sub>F<sub>1</sub>s will be processed for anther culture during the second semester of 2008.

### **O. barthii**

214 BC<sub>2</sub>F<sub>1</sub> lines bearing 54 *O. barthii* segments were selected among the 600 genotyped plants. For backup purpose, the BC<sub>3</sub>F<sub>1</sub> crosses were made from all BC<sub>2</sub>F<sub>1</sub> plants. The number of BC<sub>3</sub>F<sub>1</sub> 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<sub>3</sub>F<sub>1</sub> plants were selected and the BC<sub>4</sub>F<sub>2</sub> seeds are now available. 142 BC<sub>2</sub>F<sub>2</sub> 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 BC<sub>3</sub>DH of BC<sub>3</sub>F<sub>2</sub> 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. BC<sub>4</sub>F<sub>2</sub> seeds will also be prepared and made available to the scientific community. Additionally, a set of SNPs is currently being optimised 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 programmes. Further field trials for different important agronomic traits could permit the detection of several QTLs of significant relevance for cultivar improvement.

## **2. G3005.13: Development of informative markers through association mapping in maize to improve drought tolerance in cereals**

*January 2005–December 2007; no-cost extension to December 2008*

### **Principal Investigator**

Marilyn Warburton (Jan 2005–May 2008), Jianbing Yan from (effective June 2008), CIMMYT; j.yan@cgiar.org

Km. 45, Carretera Mexico-Veracruz El Batan, Texcoco, Edo. de México CP 56130  
México TEL: +52(55) 5804-2004-2206

### **Collaborating institutions and scientists**

- Cornell University, USA: Tim Setter, Edward Buckler
- CIMMYT: Yunbi Xu, Jonathan Crouch
- NSFCRC: Pichet Grudloyma
- KARI: James Gethi
- SIRDC: Ester Khosa
- SAU: Wanchen Li

### **Major research activities and progresses**

#### **1. Phenotype progresses**

- 1.1 Two years (2007&2008) field data for 15 agronomic traits both in well water and water stress conditions were collected from 5 locations (China, Kenya, Mexico, Thailand, and Zimbabwe).
- 1.2 Two years (2005&2006) metabolite data for 7 traits (3 organisms and 2 time points) both in well water and water stress conditions were collected from Mexico and measured at Cornell University. The extreme 100 lines based on 2005 and 2006 results were collected from Mexico, Thailand and Kenya and measured at Cornell University in 2008.

#### **2. Genotyping progresses**

- 2.1 Develop an Illumina Chip including 1536 SNPs from 732 amplicons (representing 582 unigenes, half of which were chosen to be drought candidate genes and half to provide high genetic information)
- 2.2 The 350 inbred lines were assayed using the developed chip.

#### **3. Association mapping progresses**

- 3.1 1200 SNPs could be used in the association mapping, after removing failed SNPs (very few), ambiguous calls (also few) or SNPs with one form present at very low frequencies (several of these; this is one of the weaknesses of association mapping since these cannot be predicted in advance nor avoided).



- 3.2 We have a total of 264 metabolite traits/treatment/environment combinations, and 108 SNPs were significantly associated with at least 1 of 183 metabolite combinations at the  $P=1e-4$  level using a GLM model.
- 3.3 We have 154 agronomic traits/treatment/environment combinations, and 115 SNPs were significantly associated with at least 1 of 154 agronomic combinations at the  $P=1e-4$  level using a GLM model.
- 3.4 36 SNPs (from 30 genes) were significantly associated with at least 6 and at most 39 related trait combinations (16 metabolic and 20 phenotypic, three of which were in common between the two). Most of them can be found biological evidences from the published references. Simple changes in other genes that are involved in the carotenoid pathway and ABA synthesis lead to a 4.3-10.3% variation in their respective trait. Two or more genes together explained up to 20% variation (Partial identified genes listed in Table 1).

#### 4. Outputs delivered

- 4.1 The developed drought chip has been used by other CGIAR and NARS scientists. It is the first drought Illumina chip and a good resource for future drought tolerance research.
- 4.2 The identified strong candidate genes can be used to develop markers for future MAS of drought tolerance in maize (Table 1).

**Table1 summary of identified candidate genes (Partial, Minor allele frequency>0.05)**

SNP	Candidate or nearest gene(s)	Species	location
PZB01403.4	zmAO(aldehyde oxidase)	Zea mays	1.11
PZD00056.3	mads2(MADS box protein 2)	Zea mays	5.05
PZB02194.1	ivr1(invertase gene)	Zea mays	2.03
PZD00027.3	zmm16(putative MADS-domain transcription factor)	Zea mays	3.05
PZB00137.1	pif3(Phytochrome Interacting Factor 3)	Arabidopsis	3.04
PZA03301.5	Harpin-induced 1 domain containing protein	Oryza sativa	1.08
PZB01400.2	zmAO (aldehyde oxidase)	Zea mays	1.11
PZB00728.1	acp (acyl carrier protein)	Zea mays	1.07
LYCE.4	lcyε(Lycopene epsilon-cyclase)	Zea mays	8.05
PZB01482.3	gn1 (homeobox transcription factor)	Zea mays	7.01
PZA03371.2	?	Zea mays	ctg460
PZB01389.1	abi1 (ABA insensitive 1)	Arabidopsis	8.05
PZA03637.3	set105 (SET domain-containing protein)	Zea mays	8.04
PZA03635.1	set104 (SET domain-containing protein)	Zea mays	2.03
PZB01186.1	mitochondrial phosphate transporter	Arabidopsis	5.07
PZA03573.4	zmet3 (DNA cytosine methyltransferase)	Zea mays	9.08
PZA03395.2	putative SF16 protein	Oryza sativa	8.07

### 3. **G3005.14: Characterisation of genetic diversity of maize populations: documenting global maize migration from the center of origin**

*January 2005–December 2007; no-cost extension to March 2009*

#### **Principal Investigator**

Marilyn Warburton [m.warburton@cgiar.org](mailto:m.warburton@cgiar.org)

USDA ARS Corn Host Plant Resistance Research Unit, Starkville, Mississippi

#### **Collaborating institutions and scientists**

- Agropolis–INRA: A Charcosset; C Mir\*; D Madur; F Dumas; V Combes
- CIMMYT: S Dreisigacker; S Taba; C Bedoya; J Franco
- IITA: S Hearne
- KARI: Z Muthamia
- CAAS: S Zhang; C.-X. Xie
- IARI: BM Prasanna
- IDA: M Yunus
- NSFCRC, Thailand: P Grudloyma; C Jampatong
- NMRI Hanoi, Vietnam: VT Nguyen; PX Hao

\*poster presentation

#### **Context**

Maize (*Zea mays* ssp. *mays*) originated from the domestication of a teosinte (*Zea mays* ssp. *parviglumis*) in Mexico ~ 9000 years ago (Matsuoka *et al.*, 2002). It slowly spread over the Americas, during which time gene flow from other teosintes increased the genetic base, and then it spread over all continents, through complicated patterns of introductions linked to trading and colonization. This wide genetic base and expansion into new environments, with the increasing use of maize globally, have favored the differentiation of thousands of local farmer's varieties (landraces) worldwide, adapted to local growing conditions and uses. Although modern maize hybrids now represent the most economically important portion of the species, those traditional landraces contain the majority of the diversity of the species, much of which has never been incorporated into breeding programs. This study aims to establish a global picture of maize landrace diversity, by genotyping more than 800 populations from America, Europe, Africa, and Asia, and representative teosintes with 28 SSR markers and 3 additional markers linked to flowering precocity. Our work will thus complete previous studies restricted to American and European landraces (Dubreuil *et al.*, 2006) and aid in effective maintenance and use of maize diversity. Our results should also help to better understand how maize has migrated globally over time.

#### **Results and perspectives**

One DNA-bulk of 15 individuals was genotyped per population. Analyses were performed by INRA (17 SSR + 3 flowering time markers) and CIMMYT (11 SSR markers) and were completed in May 2008. A deconvolution method was developed to limit the impact of stutters on bulked DNA genotyping (correction of allele frequency

according to the intensity of stutters and discrimination between true alleles and stutter bands). To do so, the initial approach of Dubreuil *et al.* (2006) was generalized to take into account stutters which do not follow the repeat unit of the SSR locus.

The landrace status of each population is being checked to exclude modern breeding material from the study. Preliminary analyses have been conducted on a partial population set of 800 populations based on PCA performed on the variance-covariance matrix obtained from SSR frequencies, hierarchical clustering and modeling of population structure. PCA represents the major lines of diversity organization and clearly discriminates i) African populations, ii) South American populations and iii) European + North American populations. This last group corresponds to the American Northern flints and their European derivatives, characterized by an early flowering adapted to cold temperate regions (Dubreuil *et al.*, 2006). Teosintes and Central American populations have a central position, which is congruent with their primary role in maize diffusion and landrace differentiation. PCA suggests that important differentiation processes occurred in Africa, America and Europe. On the contrary Asian populations are weakly discriminated from the Central American populations, suggesting a more limited differentiation of specific groups in this region.

Definitive data analysis will be completed by the end of 2008. The contribution of the four American ancestral groups previously identified by Camus-Kulandaivelu *et al.* (2006) (*ie* Northern flint, Mexican, Carribean and Andean) to the other groups will be assessed. The variation of frequency of alleles of interest (e.g. representative of the flowering precocity, Ducrocq *et al.* 2008) will be mapped according to coordinates of collecting sites. The results will be confronted to historical data about maize diffusion.

## References

- Camus-Kulandaivelu L, Veyrieras JB, Madur D, Combes V, Fourmann M, Barraud S, Dubreuil P, Gouesnard B, Manicacci D and Charcosset A (2006). Maize adaptation to temperate climate: relationship between population structure and polymorphism in the *Dwarf8* gene. *Genetics* 172: 2449-2463.
- Dubreuil P, Warburton M, Chastanet M, Hoisington D and Charcosset A (2006). More on the introduction of temperate maize into Europe: large-scale bulk SSR genotyping and new historical elements. *Maydica* 51: 281-291.
- Ducrocq S, Madur D, Veyrieras JB, Camus-Kulandaivelu L, Kloiber-Maitz M, Presterl T, Ouzunova M, Manicacci D and Charcosset A (2008). Key impact of Vgt 1 on flowering time adaptation in maize: evidence from association mapping and ecogeographical information. *Genetics* 178: 2433-2437.
- Matsuoka Y, Vigouroux Y, Goodman MM, Sanchez J, Buckler E and Doebley J (2002). A single domestication for maize shown by multilocus microsatellite genotyping. *PNAS* 99: 6080-6084.

#### **4. G3005.17: Allele mining based on non-coding regulatory SNPs in barley germplasm**

*January 2005–December 2007; no-cost extension to June 2008*

##### **Principal Investigator**

Michael Baum, ICARDA, [m.baum@cgiar.org](mailto:m.baum@cgiar.org)

ICARDA, P.O.Box 5466 Aleppo, Syrian Arab Republic, +96 3 2122 2551

##### **Collaborating institutions and scientists**

- NIAB: W Powell, K Stamati
- Università di Udine, Italy: M. Morgante
- ICARDA: S Ceccarelli, S Grando, SM Udupa, M Von Korff
- Tishreen University, Syria: W Choumane

##### **Selection of candidate genes for the allelic imbalance assays**

The selection of candidate genes was based on information from QTL mapping (Guo et al. 2007, Baum et al. 2003, Baum et al. 2007, von Korff et al. 2008), expression analyses (Guo et al. in prep) and from literature. Results of a previous microarray experiment (Guo et al. in prep.), testing gene expression under drought in three different cultivars, were used to identify candidate genes. Genes which were consistently up or down regulated at least 2-fold across all three cultivars and which matched the position of QTL ‘hotspots’ were prioritised for analysis. In addition, ESTs (expressed sequence tags) expressed under abiotic stress were selected for their polymorphism content. In order to identify SNPs in the selected candidate genes for haplotype tagging 70 genes were sequenced in the six parents. SNPs in the respective crosses were identified and amplification as well as extension primers designed.

##### **Allelic imbalance assays**

The imbalance assay was adopted from maize and optimised for barley. A total of 30 polymorphic genes have been tested for allelic imbalance on cDNAs extracted from the F1 and RF1 and derived from five different crosses (Hsp41-1/Alexis, Hsp41-1/Arta, Sloop/WI3408, Tadmor/Sloop and Tadmor/WI3408). The thirty different genes were tested for allelic imbalance at two developmental stages (vegetative, generative) and two treatments (control, drought). Relative allelic expression ratios were calibrated using parental DNA mixes and hybrid genomic DNA. For each cross the change of allelic expression ratios between developmental stages and control and drought treatment were calculated. In a second step we analysed the deviation of the allelic expression of the cDNA from the genomic F1/RF1 allele expression ratios. The number of genes analysed varied between crosses according to polymorphic SNPs detected.

Of the thirty genes assayed for allelic imbalance in five different crosses twenty-nine showed allelic imbalance in at least one cross. Altogether sixty-three (63%) of the eighty-two tested gene/cross combinations showed allelic imbalance. The percentage of genes with allelic imbalance varied between crosses. The cross Hsp41-1/Alexis showed with 60% the highest percentage of genes with imbalanced expression, followed by Tadmor/Sloop and Tadmor/WI3408 each with 56% imbalanced genes and then

Sloop/WI3408 and Hsp41-1/Arta with 45% allelic imbalance. The effect of imprinting on allele expression was tested by assaying the F1 and reciprocal F1 progeny simultaneously. However, no main effects of the cross direction was detected in any of the analysed gene/cross combinations.

### ***Developmental Stage***

The developmental stage caused changes in relative allele expression in ten different genes and all five crosses. In twelve out of eighty-two gene/cross combinations (15%) the relative allele expression differed significantly between the vegetative and generative stage. The majority of significant effects of the developmental stage on relative allele expression was detected in the cross Tadmor/Sloop. The examples indicate the effect of cis-regulatory units, which are responsive to developmental cues and thus cause a change of allelic expression between the vegetative and generative stage.

### ***Drought Treatment***

Seven genes and ten gene/cross combinations revealed a change in allelic expression upon drought treatment. The effect of drought on allelic expression was particularly pronounced in the cross Hsp41-1/Alexis with five out of fifteen tested genes illustrating a change in allelic expression between control and drought conditions. The ten gene/cross combinations with significant expression differences between control and drought indicate the effect of cis elements, which are regulated under drought.

The obtained data demonstrated that more than 50% of the assayed barley genes were regulated in cis. Some gene/cross combinations exhibited strong allelic imbalance, where sometimes one parental allele was exclusively expressed. These strong differences in allelic expression in barley as an inbreeding species, demonstrate the prominent role expression regulation plays in shaping a phenotype. The results indicated the presence of cis-elements responsive to drought and are thus a first step towards understanding regulatory gene networks in stress response cascades. Manipulation of cis-acting regulatory units in the breeding process will thus allow a targeted improvement of drought tolerance through the adaptation of gene expression.

### **References**

- Baum M, Grando S, Backes G, Jahoor A, Sabbagh A, Ceccarelli S (2003) QTLs for agronomic traits in the Mediterranean environment identified in recombinant inbred lines of the cross 'Arta' x *H. spontaneum* 41-1. *Theor Appl Genet.* 107:1215–1225
- Guo P, M. Baum, R.K. Varshney, A. Graner, S. Grando, S. Ceccarelli (2007). QTLs for chlorophyll and chlorophyll fluorescence parameters in barley under post-flowering drought. *Euphytica* DOI:10.1007/s10681-007-9629-6
- Baum M, M. van Korff, P. Guo, B. Lakew, S. M. Udupa, H. Sayed, W. Choumane, S. Grando, and S. Ceccarelli, 2007. Molecular Approaches and Breeding Strategies for Drought Tolerance in Barley. In R. Varshney and R. Tuberosa (eds.), *Genomic Assisted Crop Improvement: Vol 2: Genomics Applications in Crops*, 51–79.
- von Korff M., S. Grando, A. Del Greco, D. This, M. Baum, S. Ceccarelli (2008). Quantitative trait loci (QTL) associated with adaptation to Mediterranean dryland conditions in barley. TAG, in press.

## 5. **G3007.01: Ibridges: Interspecific bridges that give full access to the African rice allele pool for enhancing drought tolerance of Asian rice**

*August 2007–October 2009*

### **Principal Investigator**

Alain Ghesquière, Agropolis-IRD; alain.ghesquiere@ird.fr  
911, Av Agropolis BP 64501 34394 Montpellier Cedex 5 France; 33 (0)4-67-41-61 53; Fax 33 (0)4-67-41-62 22

### **Collaborating institutions and scientists**

- CIAT: J Tohme, C Martinez, M Lorieux (Agropolis-IRD out-posted)
- Fedearroz, Colombia: M Diago
- IER: F Cissé
- INERA: H Drissa
- LGDP, Agropolis-IRD/CNRS/Perpignan University, Montpellier and Perpignan, France: O Panaud; R Guyot
- PhilRice: A Alfonso
- PAU: JS Sidhu
- University of Arizona, USA: DW Galbraith
- WARDA, Benin: MN Ndjiondjop; M Semon; M Sié

### **Scientific context**

This project aims at overcoming the interspecific sterility barrier separating the two cultivated rice species, *O. sativa* and *O. glaberrima*. After interspecific hybridisation, a marker-assisted selection will be carried out on backcross progenies and will be focused on the  $S^1$  locus, which is the key-factor of the interspecific sterility. The fertility restoration will be monitored along three generations to derive fertile Backcross Inbred Lines (BILs) with improved crossability towards *O. sativa*. The genetic material will be scored (at Arizona University and later on at CIAT) for its *O. glaberrima* genome content in using Chip technology revealing Single feature Polymorphisms (SFPs) while the fine genetic and physical maps of the  $S^1$  locus will be done to bring this gene to positional cloning. The resulting interspecific bridges (*iBridges*) will be suitable for new evaluations and genetic studies to identify original genes/QTL with an accent on drought resistance and better water use efficiency.

### **1. Current activities**

#### **1.1 WP1: Development of interspecific bridges**

Interspecific crosses and backcrosses have been done in using IR64 (*indica*) and 2 upland *japonica* (WAB165 and Curinga) as recipient varieties. Twenty-five consensus accessions of *O. glaberrima* were selected to represent as best as possible the genetic diversity, interesting traits (drought resistance, African gall midge resistance, nematode, blast, rice yellow mottle resistance) as well as available molecular resources (BAC libraries and CSSLs). Twenty-four combinations are presently at the BC1 level with 100 BC1 seeds/combinations in average. Additional seeds have been obtained in using *O. glaberrima* as female parent.

### **1.2 WP2: Utilisation of SFP technology**

Preliminary activities have done during 2007 to check if the existing SFPs designed between *indica* and *japonica* genomic sequence can serve also directly to identify *O. glaberrima* specific markers. Bac End Sequences (BES) of CG14 were used to identify additional SFPs. First experiments were initiated at CIAT through collaboration with AGI in using the parents of *iBridges* to get a first insight on the effectiveness of this technology.

### **1.2 WP3: Physical mapping of S<sup>1</sup> locus**

A new genetic map based on the interspecific cross [IR64 x TOG5681] x IR64 has been established in using 140 SSR markers on 125 individuals. It confirmed the distortion observed at the extremity of chromosome 6 corresponding to locus S<sup>1</sup>. Additional markers deduced from the Nipponbare sequence were used to saturate the region of locus S<sup>1</sup>, which corresponds approximately to 130 Kbp. Genomic resources of OMAP (CG 14 BAC library) served to determine a Minimum Tilling Path (MTP) made of 8 BAC clones and verified by PCR assays on the BAC end sequences. Convenient markers have been derived for marker-assisted selection required for WP1.

### **2. Next steps**

Marker-assisted selection will be initiated on all available combinations while the BC1 seed-set will be increased accordingly on the remaining combinations. BC1F1 plants will be checked for fertility restoration and BC2F2 progeny will be shared between partners in charge of Back cross Inbred lines (BILs) development. Fine mapping of the S<sup>1</sup> locus will be done with a larger population and thanks to a better estimation of the self-fertility. BC1F1 populations will be mapped in using SSR core map.

### **Milestones and provisional outputs**

- 40 interspecific combinations determined for further genetic material development
- Two convenient SSR markers identified for S<sup>1</sup> marker-assisted selection
- S<sup>1</sup> locus identified in a 130 kbp interval of chr. 6

## **6. G3007.02: Genomic dissection of tolerance to drought stress in wild barley**

*August 2007–July 2009*

### **Principal Investigator**

Robbie Waugh, Head, Genetics, SCRI, Robbie.Waugh@scri.ac.uk  
Invergowrie, Dundee, DD2 5DA, Scotland, UK

### **Collaborating institutions and scientists**

- ICARDA: Michael Baum, Stefania Grando, Salvatore Ceccarelli
- OSU, Department of Crop and Soil Science: Patrick M Hayes
- SCRI: Joanne Russell, David Marshall
- INIA, Chile: Ivan Matus, Alehilejandro del Pozo
- UC–Riverside: Timothy J Close

## **Research activities and progresses**

Our approach is to combine three unique barley resources, a segregation population of 140 barley lines composed of an advanced elite genetic background containing introduced chromosomal segments from a wild barley accession from the Fertile Crescent, adapted and tolerant of drought and salt stresses (RCSLs see Matus and Hayes 2003); a collection of 480 lines (Syrian Jordanian Landrace collection- SJLC), sampled from a large barley collection (Ceccarelli et al., 1987) chosen from five geographical regions representing South Jordan, North Jordan-South Syria, West Syria, Central Syria and North-East Syria and a high throughput SNP genotyping platform for barley based on the Illumina GoldenGate and Oligo Pool Assays (OPA) (Oliphant et al., 2002, Fan et al., 2003).

### *1. Genotyping the RCSLs- detailed chromosomal introgression maps of the wild barley genome in an elite barley background*

DNA from the 140 RCSLs have been genotyped with 1536 SNPs. 765 of the 1536 SNPs were polymorphic with an average of 107 SNPs per chromosome and 1 SNP every 1.6cM. The average introgressed segment based on the initial data for chromosome 1H is approximately 28.9cM, with the size of the introgressed segments varying across the chromosomes from a few to 60 cMs.

### *2. Phenotyping the RCSLs - Quantitative phenotypic data from a suite of drought related trait components measured on the RCSLs at different locations.*

Seeds of the RCSLs were bulked in at INIA and sent from Chile to ICARDA in December. These have been grown in quarantine area in ICARDA's research station located near Tel Hadya. Data were collected on early growth vigour, growth habit, days to heading, and grain yield. The RCSLs will be planted next cropping season in at least two contrasting environments.

### *3. Regional allele mining in the SJLC- haplotype diversity across the introgression segments and estimates of LD across the barley landrace genome will be established and an assessment of functional variation will be initiated.*

A subset of 172 of the SJLCs has been genotyped with the same 1536 SNPs. A total of 1112 SNPs were polymorphic in this subset of the SJLC, with an average of 158.9 SNPs per chromosome (the map location of the majority of the SNPs has been determined previously). Using phenotypic data from previous trials at the ICARDA sites, we have some preliminary data that identifies several candidate regions for heading date and grain yield in the SJLR material. The remaining SJLCs are being prepared for genotyping in the next few months. Once this has been completed on the full collection we can begin to test whether these are potential novel candidates for drought tolerance in barley. The new data from Chile and ICARDA will be essential to this analysis.



## References

- Moragues M, Russell J, Ceccarelli S, Grando S, Flavell AJ (2008). High throughput SNP-based marker analysis of landrace barley. Poster presented at the *10<sup>th</sup> International Barley Genetics Symposium* (IBGS) 5-10 April 2008, Alexandria, Egypt.
- Russell J, Baum M, Ceccarelli S, Close TM, Grando S, Hayes PM, Matus I, Marshall D, Del Poza A, von Korff Schmising M, Waugh R. (2008). Genomic dissection of tolerance to drought stress in wild barley. Generation Challenge Programme (GCP) Workshop paper presented at the *10<sup>th</sup> International Barley Genetics Symposium* (IBGS) 5-10 April 2008, Alexandria, Egypt

---

## ***Subprogramme 2: Genomics towards gene discovery***

### **7. G3005.01: Identifying genes responsible for failure of grain formation in rice and wheat under drought**

*January 2005-December 2008*

#### **Principal Investigator**

John Bennett, IRRI, [jbennett@irrialumni.org](mailto:jbennett@irrialumni.org)

Plant Breeding, Genetics and Biotechnology Division, International Rice Research Institute; DAPO Box 7777, Metro Manila, Philippines; (63) 2-5805600; fax (63) 2-5805699

#### **Collaborating institutions and scientists**

- CSIRO: Rudy Dolferus
- NIAS: Shoshi Kikuchi; Kouji Satoh
- IRRI: Ken McNally; Rachid Serraj; Xuemei Ji; Muthurajan Raveendran
- TNAU: R. Chandra Babu
- NAU: Zhengqiang Ma

#### **Research activities**

The project aimed to use rice and wheat genotypes with contrasting behaviour under stress to identify the candidate genes that underlie differences in drought tolerance, using microarray analysis. Candidate genes would be validated genetically, and allele mining would be used to identify novel alleles contributing to improved physiological traits under stress. A key hypothesis under test was that yield-forming processes and stress-adaptive responses are in conflict (as illustrated by GA/ABA antagonism) under drought stress.

#### **1. IR64 and Apo differ in drought sensitivity at the reproductive stage**

The drought-sensitive indica cultivar IR64 was compared to Apo, an upland-adapted indica cultivar, under reproductive-stage drought stress measured in terms of the co-variable Fraction of Transpirable Soil Water (FTSW). Overall, our data revealed the existence of significant differences between Apo and IR64 for all spikelet sterility stages and grain development characters under drought at reproductive stages, including greater sensitivity of IR64 during grain filling and of Apo during pollen meiosis. In two field phenotyping experiments, the yield reduction was strongly related to reduction in spikelet fertility. Grain yield was positively correlated with total dry matter yield, panicle length, panicle exertion and peduncle length. One of the major causes of spikelet sterility seems to be related to the inhibition of peduncle elongation which in turn reduces the panicle exertion, causing sterility in the spikelets left inside the flag leaf. Molecular characterisation of drought-related QTLs in mapping populations developed from crosses of these parental lines is in progress.

## **2. Microarray analysis of transcription factor expression and ABA-GA antagonism in rice**

To elucidate global responses to drought stress in rice, a 60 mer oligomer microarray covering 22K unique genes based on the sequence of full-length cDNA clones was used to profile gene expression changes in shoots at the seedling stage and in peduncle at heading stage. Among 503 differentially expressed transcription factor-encoding genes, all the paralogous members of PHD and SNF2 families were up-regulated and those of Jumonji and TCP families were down-regulated by four drought stress treatments. AP2-EREBP, AUX/IAA, bZIP, C2C2-GATA, C3H, CPP, HB, HMG, HSF, MYB-related, NAC, SBP, SNF2 and Trihelix families were commonly up-regulated and Alfin-like, AUX/IAA, BES1, bHLH, bZIP, MYB, NAC, WRKY and ZIM were commonly down-regulated by four drought stress treatments. The metabolic pathway data in Rice Cyc showed that genes encoding many enzymes of sugar metabolism were down-regulated, along with genes encoding enzymes of cell-wall biosynthesis, while genes encoding enzymes of some amino acid biosynthetic pathways were up-regulated. Drought-induced ABA was involved in antagonizing GA-dependent events underlying peduncle elongation, but the biosynthetic genes related to these hormones were not clearly affected by the drought stress treatment.

## **3. ABA metabolism genes in drought-tolerant and drought-susceptible wheat cultivars**

We compared two Australian wheat cultivars at the reproductive stage: drought-sensitive Sundor, and drought-tolerant Sunstar. There was a dramatic decrease in grain number due to spikelet sterility when drought stress conditions coincided with the period around pollen meiosis. Spikelet fertility was close to unstressed levels when florets were stressed after meiosis. When stressed at meiosis, Sundor produced no seeds, while Sunstar still produced up to 20% of seeds. We screened wheat EST databases for genes of interest in this project: cell wall invertase, ABA biosynthetic genes (9-cis-epoxycarotenoid dioxygenase, NCED) and ABA catabolic genes (ABA-8' hydroxylase, ABA8OH). Our search was focused on ESTs that were identified in cDNA libraries from the reproductive parts of the plant. We identified 4 wheat cell wall invertases (TAINV), 3 NCED (TANCED) and 3 ABA 8'-hydroxylase (TAABA8OH). We confirmed using RT-PCR that these genes are expressed in wheat anthers and ovules.

## **4. Allele mining**

The goal of this work was to identify superior alleles of key genes for drought tolerance. Nominated genes such as OsVP1, OsABF1, and the cytP450 for ABA 8' hydroxylase were amplified from a panel of 1536 diverse accessions of *Oryza sativa* by the polymerase chain reaction (PCR) and the amplicons were screened by EcoTILLING to detect variant forms. The haplotypes of the variants were then compared statistically with phenotyping data obtained on reproductive-stage drought tolerance on over 1000 accessions.

## **Next steps and challenges**

Key issues are (i) how drought-stressed rice and wheat mobilise stored carbohydrates (starch and fructans, respectively) and allocate them to protection vs growth, (ii) how the

synthesis, function and degradation of ABA impact on drought tolerance, and (iii) how to identify the crucial members of multigene families, such as those encoding transcription factors, given that drought stress operates at many levels in addition to transcript abundance. QTL analysis will be crucial.

## References

- Ji XM, Raveendran M, Oane R, Ismail A, Lafitte R, Bruskiwich R, Cheng SH and Bennett J (2005). Tissue-specific expression and drought responsiveness of cell-wall invertase genes of rice at flowering. *Plant Molecular Biology* 59:945-964.
- Ji X, Van den Ende W, Schroeven L, Clerens S, Geuten K, Cheng S and Bennett J (2007). The rice genome encodes two vacuolar invertases with fructan exohydrolase activity but lacks the related fructan biosynthesis genes of the Pooideae. *New Phytologist* 173:50-62.
- Liu JX, Liao, DQ, Oane R, Estenor L, Yang XE, Li ZC and Bennett J (2006). Genetic variation in the sensitivity of anther dehiscence to drought stress in rice. *Field Crops Research* 97:87-100.
- Shobbar ZS, Oane R, Gamuyao R, De Palma J, Malboobi MA, Karimzadeh G, Javaran MJ and Bennett J (2008). Abscisic acid regulates gene expression in cortical fiber cells and silica cells of rice shoots. *New Phytologist* 178:68-79.
- Satoh K, Doi K, Nagata T, Hosaka A, Suzuki K, Ji X, Raveendran M, Leung H, Bennett J and Kikuchi S (2008). Gene expression analysis and data mining from the microarray analysis applied to the drought stress in rice. In: *Recent Progress in Understanding Drought Tolerance in Rice* (Serraj R, Bennett J and Hardy B, eds). International Rice Research Institute, Manila, Philippines. In press.

## 8. G3005.02: Revitalizing marginal lands: discovery of genes for tolerance of saline and phosphorus deficient soils to enhance and sustain productivity

*January 2005-December 2007; no-cost extension to June 2008*

### Principal Investigator

Abdelbagi M Ismail, IRRI, abdelbagi.ismail@cgiar.org  
International Rice Research Institute; Crop and Environmental Sciences Division,  
IRRI, DAPO Box 7777, Metro Manila, Philippines; +63 (2) 580-5600 ext. 2843;  
Fax: +63 (2) 580-5699

### Collaborating institutions and scientists

- IRRI: David Mackill; Michael J. Thomson; Sigrid Heuer; Glenn Gregorio; RK Singh
- JIRCAS: Matthias Wissuwa
- UC-Davis: Eduardo Blumwald
- Dhaka University, Bangladesh: Zeba I. Seraj
- ICABIOGRAD: Masdiar Bustamam
- UC-Riverside: Timothy J. Close
- ABRRI: Ghasem H. Salekdeh
- NIAS: Masahiro Yano

### Project context

Salt affected soils and soils deficient in essential plant nutrients have low productivity and are commonly associated with poverty. Problems of particular importance in these

soils are salinity and phosphorus deficiency, where amendments are too expensive but solutions through improved germplasm are affordable to farmers and are becoming more feasible with the developments in modern molecular breeding tools (Ismail et al. 2007). The project aimed to identify genes for tolerance of salinity and P-deficiency underlying two major QTLs: *Saltol*, associated with low salt uptake from saline soils, and *Pup1*, associated with enhanced P-uptake from soils with high P-fixing capacity.

### **Findings and implications**

Fine-mapping delimited the *Pup1* locus to a 272 kb region on chromosome 12, and gene discovery at the QTL proceeded to an advanced stage with two high priority candidate genes identified from the underlying DNA sequence of the tolerant Kasalath variety. Detailed phenotypic characterisation also found that *Pup1* may be more beneficial under drought-stressed conditions. These results led to a subsequent GCP project started this year to transfer the tolerant *Pup1* allele into several varieties using MAS and test them under different P-deficient and drought-prone environments. Further details are described in the abstract to the SP3 project “Application and validation of the major QTL *Phosphate Uptake 1*” (see G4008.41).

For *Saltol*, gene-based markers were developed and additional SSR markers were identified and genotyped across the *Saltol* region for fine-mapping, delimiting the QTL to a 1.2 Mb region between 11.4-12.5Mb on the physical map of chromosome 1. However, the data could not further pinpoint the QTL due to the complexity of several different Pokkali alleles at the *Saltol* locus, combined with the effects of several background introgressions, which complicated the analysis. To further investigate the underlying tolerance loci, several NILs were developed and functionally characterised by detailed phenotyping and expression of candidate genes identified through a microarray study between the sensitive IR29 and the tolerant RIL FL478. Furthermore, NILs containing at least two different Pokkali alleles (one allele found in FL478 and another represented by other RILs such as FL378), were developed and characterised for tolerance. While the NILs showed intermediate tolerance, higher levels were seen when *Saltol* was combined with other QTLs in the background. These results indicated that a QTL pyramiding strategy would be needed to obtain the highest levels of salinity tolerance in the field.

A precision marker-assisted backcrossing (MAB) strategy was applied towards *Saltol* to develop a package of tolerant donors and polymorphic SSR markers that could be used to transfer the QTL to popular varieties. Closely linked markers were tested across different pairs of donors and recurrent parents to identify the best markers for use in a molecular breeding programme, and crosses were initiated for several MAB populations. This progress led to a subsequent SP3 project to apply this strategy towards developing salt-tolerant varieties for Bangladesh (see G4008.08).

Progress was also made in identifying putative candidate genes associated with tolerance at *Saltol* region. About 17 genes were short-listed based on converging evidences including expression data and physiological role. Out of these, about 9 genes were identified with high priority, and some of them were being validated. Beside the *SKC1* gene identified in this region (Ren et al., 2005) a cation-chloride co-transporter was cloned and characterised, and its role in tolerance is being studied.

## Challenges

While the *Saltol* QTL has shown good levels of tolerance at seedling stage, additional QTLs are likely needed for higher tolerance in field conditions and at reproductive stage. Furthermore, multiple abiotic stresses as submergence and various problem soils often co-exist in farmer's fields, which present a challenge to breeders to simultaneously select for multiple traits. Subsequent steps should combine multiple QTLs for different traits into the same variety, to ensure more stable yields even under multiple stresses. While this project has laid the groundwork for optimizing a MAB system for several salinity tolerant alleles, more research is needed to add additional tolerance QTLs to the molecular breeders' toolbox to meet the challenges of breeding for more stable rice production in marginal environments and to improve farmers' livelihoods.

## Reference

- Ismail AM, Heuer S, Thomson MJ and Wissuwa M (2007). Genetic and genomic approaches to develop rice germplasm for problem soils. *Plant Molecular Biology* 65: 547-570.
- Ren ZH, Gao JP, Li LG, Cai XL, Huang W, Chao DY, Zhu MZ, Wang ZY, Luan S, Lin HX (2005) A rice quantitative trait locus for salt tolerance encodes a sodium transporter. *Nature Genet* 37:1141-1146

## 9. G3005.08: Targeted discovery of superior disease QTL alleles in the maize and rice genomes

*January 2005–December 2007; no-cost extension to December 2008*

### Principal Investigator

Rebecca Nelson, Cornell University, rjn7@cornell.edu  
Cornell University, Ithaca, NY 14853, USA

### Collaborating institutions and scientists

- NCSU: Peter Balint-Kurti
- IRRI: Darshan Brar, Hei Leung, Casiana Vera Cruz
- ICABIOGRAD: Masdiar Bustamam
- KARI: James Gethi, Jedidah Danson, Jane Ininda
- CSU: Jan Leach
- Cornell University, USA: Margaret Smith
- IPB: Utut Suharsono

### Near-isogenic lines (NILs) for quantitative resistance loci (QRLs)

- At IRRI, two rice NILs derived from SHZ2 x TXZ containing a QTL from chromosome 6 were confirmed. Line BC10, containing several QTL regions, is currently being used to develop additional NILs in the background of IR64-Sub1 and Swarna.
- Nine BC<sub>3</sub>F<sub>8</sub> rice lines derived from Vandana x Moroberekan were selected based on blast resistance in multilocation test, agronomic traits (duration, yield under low and high disease pressure, harvest index, drought tolerance), and defense response gene combinations.

- At NCSU, a set of maize NILs for southern leaf blight (SLB) QRLs has been produced. At CU, a set of maize NILs carrying QRLs for northern leaf blight has been produced using multiple strategies. The maize lines are being tested across field environments, with diverse pathogens in greenhouse tests. Histological analysis has identified distinct effects of different QRLs on pathogenesis.
- At KARI, a set of maize NILs is being produced for GLS resistance using the heterogeneous inbred family method. Inbreeding of 72 NIL pairs was advanced to the BC<sub>3</sub>F<sub>1</sub> generation. One more inbreeding step will be conducted before genotyping for introgression regions and phenotyping under artificial GLS infestation at Kakamega.
- A strong QRL in maize bin 8.06 has been fine-mapped. Significant progress has been made towards cloning this and several other maize QRLs.

### **QTL mapping, association mapping and nested association mapping**

#### **Datasets**

- At NCSU, three years of data were obtained for the maize nested association mapping population (NAM, a 5000-line resource). Five site-years of data were obtained for the association mapping panel for SLB.
- At Cornell, two site-years of data were obtained for the NAM for NLB. Three site-years of data were obtained for the maize association mapping panel for NLB. One site-year of data was obtained for the NAM for GLS in collaboration with Virginia Tech.

#### **QTL and association analyses**

- QTL were identified at high resolution for SLB and NLB using the NAM population. Several of these QTL have been confirmed and analyzed in NILs.
- Correlations among resistances to three foliar diseases were confirmed in an association mapping panel using multivariate mixed model association genetic analysis.
- A gene associated with resistance to SLB, GLS and NLB was identified using association analysis.

#### **Deletion mapping**

- The utility of the Affymetrix Rice GeneChip® was demonstrated for discovery of deleted genes in rice using the lesion mimic phenotype *sp11* (spotted leaf 1) for proof of concept.
- The positions of a total of 16 mutants onto the rice pseudomolecules were presented in a user-friendly browser.
- The density and distribution of the deletions suggests the feasibility of creating a database describing a collection of available deletions in the genome. This community resource can continue to grow with further hybridisations, allowing researchers to quickly identify mutants that harbor deletions in candidate genomic regions, such as QTL.

#### **Breeding**

- KARI: Seven three-way cross hybrids with superior performance and developed using the disease panel of the GCP are under National Performance Trials for the

Kenyan drylands and coastal lowlands. These the initial products from the GCP work and the testing is being supported under AGRA.

- To facilitate combining key traits for upland rice in Indonesia, QTL for tolerance to P-deficiency were mapped in the Way Rarem / OL5 population in which blast QTL were previously identified.

### **Selected publications**

- Wisser R.J., S.C. Murray, J.M. Kolkman, H. Ceballos and R.J. Nelson. Selection mapping of loci for quantitative disease resistance in a diverse maize population. In press, Genetics.
- Balint-Kurti, P.J., R.J. Wisser, and J. Zwonitzer. 2008. Use of the IBM Population for Precise Mapping of Quantitative Trait Loci for Gray Leaf Spot Resistance in Maize. In Press, Crop Science.
- Wisser, R.J., J.M. Kolkman, J.B. Holland, J. Yu, M. Krakowsky, R.J. Nelson and P.J. Balint-Kurti. Genes associated with multiple disease resistance identified through multivariate mixed model association genetic analysis. Submitted.
- Zwonitzer, J., Bubeck, D., Bhatramakki, D., Goodman, M., Arellano, C., Balint-Kurti, P. 2008. Use of backcross recurrent selection and QTL mapping to identify loci contributing to southern leaf blight resistance in a highly resistant maize line. Submitted.
- Poland, J., P.J. Balint-Kurti, R.J. Wisser, R. Pratt and R.J. Nelson. Shades of gray: The world of quantitative disease resistance. Submitted.
- Bruce, M., A. Hess, J. Bai, R. Mauleon, M. G. Diaz, N. Sugiyama, A. Bordeos, G. Wang, H. Leung, and J.E. Leach. Detection of genomic deletions in rice using oligonucleotide microarrays. Submitted.

## **10. G3005.11: Functional genomics of cross-species resistance to fungal diseases in rice and wheat (CEREALIMMUNITY)**

*January 2005–December 2007; no-cost extension to October 2008*

### **Principal Investigator**

Jean-Benoit Morel, Agropolis-INRA; jbmorel@cirad.fr  
Campus International de Baillarguet, UMR BGPI, INRA TA A-54/K ,  
34398 Montpellier, France; (33) 499 624 837; Fax: (33) 499 624 822

### **Collaborating institutions and scientists**

- EMBRAPA: M Ferreira, G de Capdeville, S Scagliusi, A Bonato, Y Mehta, J Maciel, P Scheeren, MS Chaves, S Brammer
- UC-Davis: P. Ronald, KH Jung
- CIMMYT: R Sing, M William
- NIAS: S Kikuchi, K Satoh
- JIC: L Boyd, H Tufan, G McGrann
- Agropolis-INRA: C Feuillet, P Sourdille, J Salse, L Miché, M Bousquet, C Michel

### **1. Project background**

Non-host resistance represents one promising defence mechanism in developing **durable resistance** against plant pathogens, namely due to its effectiveness against a broad range of pathogen species and its durability in nature. The proposed project will strengthen and extend ongoing research in rice and wheat and aims at defining the signalling and effector



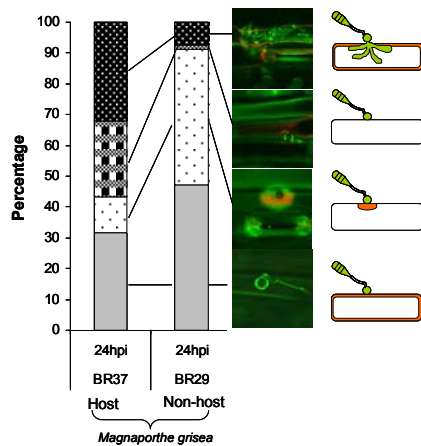
genetic components involved in non-host resistance in cereals to devise **novel defence strategies** which have the potential to **yield durable resistance against host pathogens in cereals**.

### 1. Latest progress and findings

Previous microscopical analysis (EMBRAPA) showed that pathogens non-adapted to rice were in close contact with rice surface as soon as 6 hours for *Magnaporthe* (from wheat) and 10 hours for *Puccinia* (from wheat). Non-host resistance of wheat to non-adapted strains of the fungus *M. grisea* was characterised at the cytological level (JIC- Figure 1). This was a prerequisite for further molecular studies.

Using a transcriptomics approach, we isolated more than 800 rice genes (NIAS and UC Davis), 140 of which seem to be specific to non-host resistance. These genes were positioned on the rice genome and this information is now available through the GCP bioinformatics platform. Twenty four of these genes were selected for further analysis (Agropolis). We have now isolated rice mutant lines corresponding to 11 potential rice non-host resistance genes (UCD and Agropolis). These mutant lines will soon be tested for host and non-host resistance.

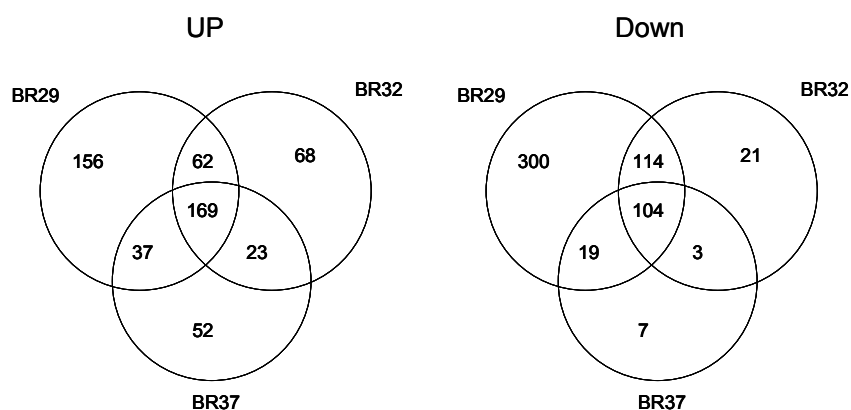
Preliminary results with a line mutated for a lipoxygenase (differentially expressed in wheat non-host resistance and rice host resistance) indicate that this gene is required for host resistance (Agropolis).



**Figure 1. Development of non-adapted *M. grisea* strain is halted by 24hpi.**

Wheat cultivar Renan was inoculated with host strain BR37 and non-host strain BR29 of the fungus *M. grisea* (transformed with the GFP reporter gene). Images were taken using confocal microscopy. Wheat shows an early response to the BR29 isolate

A transcriptomics approach has been done for wheat non-host resistance to *M. grisea* isolate BR29 from digitaria (JIC; Figure 2). The non-host preferential expression was validated by QRT-PCR for 17 genes.



**Figure 2. Commonalities and specificities identified between non-host and host resistance in wheat (Renan).** These diagrams show that non-host resistance and host resistance require overlapping sets of genes. Three isolates from *M. grisea* were used (isolate BR29 from digitaria and BR23 and BR37 isolates from wheat). Affymetrix experiments were done for RNA extracted 24 h post-inoculation.

From the 156 genes up-regulated in non-host resistance in wheat (Fig 2), 22 were mapped in wheat mapping population Renan X Recital (INRA-Clermont). These genes are now being mapped on other wheat mapping population and a set of 100 wheat accessions will be haplotyped (CYMMIT).

This work will provide a framework for functional analysis of rice and wheat resistance to non-adapted pathogens. It is one of the first large scale study of wheat defence systems. Combined to genetic information and diversity, this work could guide breeders in the choice of accessions and genomic regions for improving disease resistance of rice and wheat.

## 11. G3005.15: Determination of a common genetic basis for tissue growth rate under water-limited conditions across plant organs and genomes

*January 2004–December 2007; no-cost extension to December 2008*

### Principal Investigator

F Tardieu, Agropolis–INRA LEPSE Montpellier; [tardieu@supagro.inra.fr](mailto:tardieu@supagro.inra.fr)

### Collaborating institutions and scientists

- Agropolis–INRA: C Welcker, O Turc, B Parent
- IRRI: R Serraj, J Cairns, R Bruskiewich, R Mauleon
- CIMMYT: G Davenport, JL Araus, C Bencivenni, M Reynolds
- ETH: A Hund, P Stamp, M Liedgens, N Pa-In
- ACPFG: P Langridge, T Schnurbusch, U Baumann A Schreiber

- IARI: BM Prasana
- KARI: J Gethi
- Biogemma: P Lessard

### **Context**

We have undertaken a multiple-species, multiple-organ study on a key process of drought tolerance, growth maintenance under water deficit. We aimed to identify processes, QTLs and genes controlling tissue growth rate under water stress, in leaves across three cereals (maize, rice and wheat) and across three organs in maize (roots, leaves and silks, female reproductive organs). The project combines new approaches of phenotyping (controlled conditions and field), modelling, quantitative genetics and comparative genomics.

### **1. A "toolbox" for phenotypic analysis of several organs in 3 species (INRA, IRRI).**

- *Where and when sample ?* Growth is restricted to specific zones and times, with large differences in metabolites, RNA and protein contents between growing and non-growing zones. (i) Rice leaves elongate in the first 50 mm beyond the leaf insertion point in IR64 and Azucena, vs 80 mm in most maize genotypes. This is much longer than generally assumed for rice and wheat. (ii) The growth rate of rice leaves is maximum for the first two days (at 25°C), and declines afterwards (vs a stable growth rate in maize for 5-7 d). Sampling for transcript or metabolic patterns should therefore be performed 1 or 2 d after leaf appearance. (iii) Maize silks grow exponentially on their whole length until emergence. Then, emerged part stops growing. It is therefore essential to sample the hidden part of silks, before emergence.

- *Dealing with fluctuating temperatures.* Temperature is fluctuating in field experiments, thereby needing a correction via thermal time. Unfortunately, this calculation results in large errors in rice whose responses to temperature are clearly non-linear. The response of 15 traits has been studied in the range 10 - 40°C for 7 lines. Common response among traits and experiments were found, allowing a new method for calculation of thermal time based on a Johnson and Lewin law.

- *Growth response to drought.* A common temporal pattern of leaf growth was found in the same 7 rice genotypes in IRRI fields and at INRA. Each genotype can be characterised by a maximum rate for 2 d (25°C) and a common trend of decline afterwards. The responses to water deficit and evaporative demand were common to several experiments for each genotype. Growth maintenance can therefore be defined precisely for rice as it has been in maize.

### **2. Genomic regions involved in growth maintenance (all partners).**

*Maize.* Genomic zones controlling leaf, silks and root growths and growth maintenances under stress have been identified in a common mapping population (P1xP2, CIMMYT).

- *Leaves.* Seven genomic regions involved in leaf growth maintenance can be considered as reliable as they were found in different genetic materials and were positively tested in populations of introgression lines.

- Silks. Half of QTLs of leaf growth maintenance collocate with QTLs of silk growth (ASI) in dry fields in Mexico (CIMMYT).
- Roots. Major QTLs were constitutive one which affected most traits and one related to ramification, which collocates with a QTL of silk growth (ASI). A QTL applied to growth maintenances of both leaves and roots.
- Yield. 4 QTLs of leaf and silk growths collocate with QTLs of yield in CIMMYT fields with stress at flowering. This is tested in other 5 field experiments in Mexico, Kenya and India for stresses occurring earlier in the season. QTLs of growth maintenance in phenotyping platforms affect in most cases the yield simulated in 100 droughted fields.

*Rice.* Several regions have been identified for leaf growth maintenance in a series of experiments using the Vandana/Moroberekan advanced backcross population under field conditions. Two of them, involved in leaf growth maintenance are most reliable and are related to other traits of growth or yield under water deficit. These regions will be tested with NILs in controlled conditions.

*Wheat.* QTLs related to growth have been performed in the Kukri x Excalibur and in the Seri x Babax populations.

### **3. Transcript linked to growth maintenance in the three species (*all partners*)**

Growing tissues (see above) have been sampled for transcript analysis in roots, leaves and silks of maize, and leaves of rice and wheat.

- In maize, 2 parents and 2 contrasting RILs are compared at two stress levels (0.02 and 0.5 MPa). RNAs have been extracted in Europe and sent to Mexico where they are analysed. Data analysis (*CIMMYT*) will identify transcripts whose differential expression between genotypes and water status are associated with growth maintenance.
- In rice, 5 genotypes were sampled at 2 stress levels, and two at 3 stress levels (FTSW of 1, 0.5, 0.2). Transcripts have been analysed in Affymetrix chips. The meta analysis (*IRRI*) will include (i) identification of genomic regions with correlated expression, (ii) Aggregation of differentially expressed genes in contiguous genome regions, (iii) Association analysis between regions with correlated expression with drought response associated QTLs in the three studied species. A specific analysis of the expression of genes of the expansin family has also been carried out in the growing zone, and identifies expansins associated with local growth as this was the case in maize.
- In wheat, 3 genotypes have been analysed in scenarios of dehydration - rehydration. Results are under analysis (ACPFG).

## **12. G3005.16: Isolation and characterisation of aluminum tolerance genes in the cereals: an integrated functional genomic, molecular genetic and physiological analysis**

*January 2005–December 2007; no-cost extension to December 2008*

### **Principal Investigator**

Leon V Kochian; Robert W. Holley Center for Agriculture and Health;  
LVK1@cornell.edu

Cornell University, Ithaca, NY 14853, USA; (607) 255-2454; Fax: (607) 255-2459

### Collaborating institutions and scientists

- EMBRAPA Maize and Sorghum: Jurandir Magalhaes, Claudia Guimarães, Vera Alves, Newton Carneiro, Andrea Carneiro, Robert Schaffert
- EMBRAPA Wheat: Sandra Brammer, Luciano Consoli, José Pereira da Silva Junior, Euclides Minella, Eduardo Caierão, Alfredo Nascimento Junior
- EMBRAPA Rice and Bean: Flavio Breseghello, Pericles Neves
- Moi University, Kenya; Sam Gudu
- Robert W. Holley Center for Agriculture and Health, USDA-ARS: Lyza Maron; Owen Hoekenga, Jiping Liu, Ed Buckler

### Research activities and progress

**Sorghum and maize Al tolerance.** *Sorghum.* We have continued to characterise the major sorghum Al tolerance gene we cloned last year, *Alt<sub>SB</sub>*, which encodes a MATE transporter that mediates Al-activated root citrate exudation that is central to sorghum Al tolerance (Magalhães et al. 2007). In the course of developing NILs harboring different *Alt<sub>SB</sub>* alleles in the same genetic background to identify the best *Alt<sub>SB</sub>* alleles for molecular breeding for improved sorghum Al tolerance, we have obtained clear evidence for several novel Al tolerance loci in sorghum that contribute to the significant transgressive segregation for Al tolerance that was observed (Caniato et al. 2007). Furthermore, we have found strong evidence that one or more of these novel tolerance genes probably interacts with *Alt<sub>SB</sub>* to facilitate maximal *Alt<sub>SB</sub>* expression and tolerance. We are currently producing highly Al tolerant lines useful for breeding, for our recently funded GCP grant on advancing sorghum Al tolerance.

*Maize.* Al tolerance in maize is a typically complex, quantitative trait, unlike in sorghum where a single tolerance locus explains the majority of the variation in different mapping populations (Magalhaes et al, 2004). We have built upon our previous QTL mapping work for a population from a cross between a tolerant and sensitive tropical maize lines by developing RILs from the same cross and assembling a genetic map with 162 markers distributed across 1,657 cM of the maize genome. Six Al tolerance QTL were identified on chromosomes 3, 5, 6 and 8, explaining approximately 53% of the phenotypic variation. The two QTL on chromosomes 5 and 6 explained approximately 13 and 25% of the variation in tolerance respectively, while the other 4 QTL each explained around 5% of the variance. An analysis of changes in allele frequency was used to confirm the presence of the QTL controlling maize Al tolerance on Chr 5 and 6. We have also conducted a detailed analysis of root gene expression under Al stress using maize microarrays ([www.maizearray.org](http://www.maizearray.org)) with the Al-tolerant tropical maize inbred line C100-6, and Al-sensitive L53 (Maron et al., 2008). It was found that several maize members of the MATE gene family, for which our major sorghum Al tolerance gene, *Alt<sub>SB</sub>*, is also a member, exhibited much higher expression in the root tips of the tolerant line compared with the sensitive line. The most dramatic differences in expression are for the gene we have designated *ZmMATE1*. We also identified a second related MATE, *ZmMATE2*, that also differentially expressed in root tips of tolerant lines. Genetic mapping of *ZmMATE1* confirmed co-localisation of *ZmMATE1* to the major Al tolerance QTL detected on chromosome 6 described above). In addition, mapping of an indel in the first intron of *ZmMATE2* showed that this gene maps to the same location as the second major Al

tolerance QTL on chromosome 5. These expression patterns from microarray studies were confirmed with quantitative real time PCR analysis.

The MATE gene family is large and complex (56 members in Arabidopsis) and it is more complex in maize, as a number of duplications have occurred in the maize MATE family. Among the MATE genes found in maize, there are some that share significant sequence identity to our sorghum Al tolerance gene, *Alt<sub>SB</sub>*. However, it is interesting to note that neither *ZmMATE1* nor *ZmMATE2* are close homologs of *Alt<sub>SB</sub>*, and therefore would not have been identified without the integration of genetic and gene expression profiling approaches. The characterisation of this gene family and its role in Al tolerance across different cereals will provide insights into the relationship (or lack thereof) between sequence and functional homology. One of the future goals of our research will be to complete verification of *ZmMATE1* and *ZmMATE2* gene as *bona fide* maize Al tolerance genes, which will open up new avenues for improving maize Al tolerance via identification and tagging of the most tolerant alleles for these genes.

**Field testing in Africa.** At Moi University, together with the Kenya Agricultural Research Institute (KARI), in collaboration with Embrapa Maize and Sorghum and USDA-ARS at Cornell, we have been undertaking significant field-testing of both sorghum and maize germplasm from Kenya, ICRISAT and Tanzania (sorghum), CIMMYT and Brazil (maize), in the acid soils (pH<5.5) of Kenya both in the high (1600 – 2100 m) and medium (1200 m above sea level) altitude areas. We have confirmed that soil acidity reduces yield of maize and sorghum and have been able to begin to quantify this effect. Yield increases in response to lime and P application is greater than 28% and 35 %, respectively. Brazilian commercial hybrids and open pollinated materials including Kenyan elite materials, were tested at Bumala and Segla sites (medium altitudes) and at Kuinet and Moi University experimental sites (high altitude). In some cases where soils are rather poor, acidic and prone to drought, the grain yields of both sorghum and maize were doubled in response to P and/or lime application. Overall, Brazilian commercial hybrids out performed Kenyan commercial hybrids probably because they were selected for superior performance in acid soils. In order to improve acid soil adaptation, we are creating inbred lines from Brazilian single cross, top cross populations using Brazilian and Kenyan heterotic elite lines, and also producing inbred lines from local maize and sorghum landraces that have shown promise in our field tests. Using these approaches, we have identified recombinant lines that show good promise for yield on acid soils. These lines are being used for production of inbreds, hybrid seeds and synthetic varieties (in both maize and sorghum).

**Tangible outputs.** 1) Identification of highly aluminum tolerant alleles of our sorghum Al tolerance gene, *Alt<sub>SB</sub>*, for development of breeding lines. 2) Evidence for additional sorghum Al tolerance genes, which we are currently identifying. 3) Identification of candidates for the first maize Al tolerance genes, *ZmMATE1* and *ZmMATE2*. 4) Develop maize breeding lines improved for Al tolerance based on QTL and *ZmMATE* information. 5) Develop recombinant maize lines between crosses between Brazilian and Kenyan materials for ongoing field testing on acid soils in Africa.

### **13. G3007.03: Development of genomics resources for molecular breeding of drought tolerance in cassava**

*September 2007–February 2010*

#### **Principal Investigator**

Pablo Rabinowicz, Institute for Genome Sciences, and Department of Biochemistry and Molecular Biology, University of Maryland School of Medicine;  
prabinowicz@som.umaryland.edu

20 Penn Street, HSF-II Room S-445, Baltimore, MD 21201, USA. Tel: (1) (410) 706-6714; Fax: (1) (410) 706-1482

#### **Collaborating institutions and scientists**

- ACGT: Jane Morris
- ACGT, Department of Genetics, Forestry and Agricultural Biotechnology Institute (FABI): Alexander Myburg
- ACGT, School of Molecular and Cell Biology, University of the Witwatersrand, Johannesburg, South Africa: Chris Rey
- UC–Davis: Ming-Cheng Luo

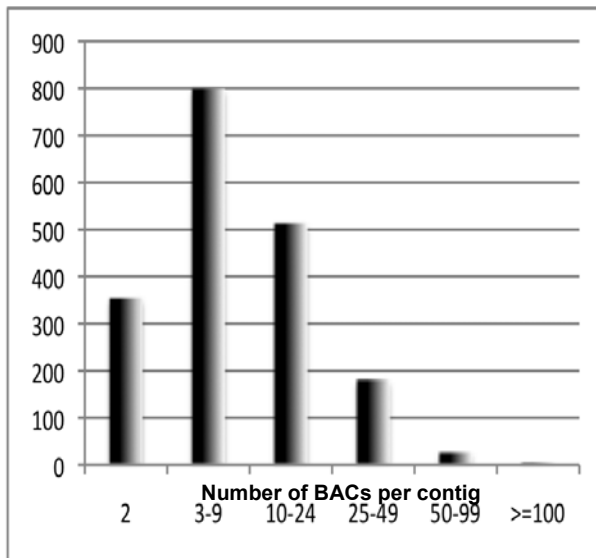
#### **Context**

Cassava is an important crop in unfavorable environments in poor areas of developing countries. As cassava is often cultivated in dry areas, research towards improvement of drought tolerance in cassava is needed. In order to accelerate the advance of genetic improvement of cassava, genomic resources must be made available to the community. A genome sequencing project for cassava has already been initiated at the U.S. D.O.E. Joint Genome Institute (JGI). A partially inbred cultivar generated at CIAT has been selected for genomic sequencing to avoid the problem of heterozygosity.

We have started a project aimed to developing a panel of single nucleotide polymorphism (SNP) markers on a genome-wide basis to localise favorable alleles in existing mapping populations, generated from contrasting drought tolerance genotypes. For this purpose, we will fingerprint a bacterial artificial chromosome (BAC) library from the same genotype being sequenced at JGI.

#### **Findings and implications**

We have had a BAC library made at Clemson University Genomics Institute using nuclear DNA from a partially inbred cassava genotype called AM560-2 that has been selfed for three generations by our collaborators Marin Fregene and Hernan Ceballos at CIAT (Cali, Colombia). CoPI Luo has completed the fingerprinting of 29,952 BAC clones, which corresponds to approximately one half of the BAC library. Of those 27,954 fingerprints were obtained and 23,441 were suitable for assembly. Using the FPC (fingerprint contig) software, a preliminary assembly of the BAC fingerprints was carried out. As a result, 21,551 BAC clones were merged into 1,877 contigs and 1,890 clones remained as singletons. The longest contig spans 1.9 Mbp and the distribution of numbers of contigs relative to the number of BAC clones per contig is shown in Figure 1.



**Figure 1: Distribution of numbers of BACs in each interval of number of clones per contig.**

A set of minimally overlapping clones or minimal tiling path (MTP) was selected and it contains 5,088 BAC clones.

#### **Next steps**

The second half of the BAC library is being delivered from Clemson to UC Davis for fingerprinting. Once the whole library is fingerprinted, we will run a final assembly that will represent approximately a 10X clone coverage of the cassava genome. We will then select a MTP and the ends of all BAC clones in the MTP will be sequenced for SNP discovery. In the mean time, we will test SSR markers in the parents of the mapping populations to estimate the level of polymorphism.

The genome-wide SNP marker set that this project will deliver will allow identification of quantitative trait loci (QTL) associated with drought tolerance by high-throughput genotyping of validated SNPs.

As cassava is a close relative of castor bean, whose genome has been sequenced, once the genome sequence and fingerprint map of cassava are released, it will be possible to conduct comparative analyses within the Euphorbiaceae family. The whole genome draft sequence of castor bean carried out in a previous project by the PI, shed light on the evolution of genome duplications in the related poplar genome, sequenced by the JGI. This kind of information will facilitate the elucidation of the evolutionary history and domestication of cassava.

Furthermore, the availability of a fingerprint map and the genome sequence for one genotype of cassava, will open the door for genome-wide diversity studies using other cultivated or wild cassava varieties in comparison with the reference genome. Next-generation sequencing technologies will make it affordable to re-sequence different cassava genomes and take full advantage of the broad genetic diversity of cassava for breeding and crop improvement.



**14. G3007.06: Genetic dissection of drought adaptive mechanisms in bread and durum wheat through large scale phenotyping methodologies (SP2)**

*January 2007–January 2010*

**Principal Investigator**

Matthew Reynolds, CIMMYT, m.reynolds@cgiar.org

Apdo. Postal 6-641, 06600 Mexico, D.F., Mexico, Phone: +52 55 5804 2004

**Collaborating institutions and scientists**

- CIMMYT: Daniel Mullan; Jose Crossa; Yann Manes
- ACPFG: Peter Langridge; Ali Izanloo; James Edwards; Dion Bennett
- DWR: Jagadish Rane; B. Mishra
- Agharkar Research Institute: Satish Mishra

**Context**

Research objectives focus on large scale phenotyping methodologies in wheat, towards the dissection of drought adaptive mechanisms. Three mapping populations (DH1=Kukri x RAC875, DH2=Excalibur x Kukri, DIC=Atil x *Triticum dicoccum*) are being genotyped and phenotyped under multiple environments in Mexico, Australia and India. QTL discovery will allow more detailed genetic analysis of specific loci, providing a foundation for positional cloning and germplasm development. The development of new mapping populations not segregating for key development traits and controlled environment studies are being used to further support these activities.

**Research activities at CIMMYT and ACPFG**

Extensive field trials have been conducted by CIMMYT at Cd. Obregon under irrigated, controlled drought and heat environments, and by ACPFG throughout southern Australia, during the past 2 years. Many traits have been phenotyped in the trials, including crop establishment, canopy temperature, NDVI, water use efficiency, stem carbohydrates and yield components, with data now ready for combined component trait analysis and further QTL discovery. Additional field trials are planned for the coming growing season, with the inclusion of environments in India. Genotyping of the three mapping populations has already begun, with the DH1 and DH2 population genetic maps completed and a map for the DIC population under construction. The DH1 population consists of 495 markers (199 SSRs and 296 DArTs), with an average marker spacing of 6.4 cM, while the DH2 population consists of 430 markers (170 SSRs and 260 DArTs) with an average marker spacing of 6.1 cM.

Genetic map construction has enabled QTL analysis and detected significant QTLs in the DH1 and DH2 populations. In the Excalibur x Kukri population a significant QTL located on the short arm of 7A was found to be collinear with a flowering time gene already mapped to the group 7 chromosome in both wheat and barley. In addition, other QTLs for yield, Na<sup>+</sup> exclusion and nematode resistance (*Pratylenchus neglectus*) have been mapped to chromosome 7A, and are independent of maturity. These QTLs, along

with other drought-specific physiological traits, will form the basis for fine mapping studies during the next year. Some of these QTLs are summarised below:

**Table 1. QTL discovery on wheat chromosome 7A in the Excalibur x Kukri population.**

Trait	H <sup>2</sup>	LOD	% Phenotypic variance explained
Maturity	0.76–0.98	41	20
Yield	0.26–0.73	13	13
Na <sup>+</sup> exclusion	0.4	4	7
Nematode resistance	0.62	58	51

Glasshouse experiments have been designed to assess drought adaptive traits in a controlled environment. Two subsets have been selected from the 11<sup>th</sup> SAWYT and Seri/Babax RILs for determining the most suitable traits for application in high throughput phenotyping. The first cycle of analysis was completed at CIMMYT during 2008, and another will be conducted in controlled environments at ACPFG in during 2009 to further characterise physiological responses of the genotypes. The subsets are in preparation for field evaluation in 2009 in Mexico, Australia and India for comparison with the controlled environment analyses. Additional activities also include the generation of mapping populations that are not confounded by phenology. From 20 crosses, involving parents with very similar phenology, 4 have been selected for further development due to the parents sharing the same alleles for *Ppd*, *Vrn* and *Rht* genes. The selected crosses represent the most likely candidates for new mapping populations that will enable the evaluation of drought adaptive mechanisms without the confounding affects of phenology.

---

### ***Subprogramme 3: Trait capture for crop improvement***

#### **15. G3005.03: Identifying the physiological and genetic traits that make cassava one of the most drought tolerant crops**

*January 2005–December 2007; no-cost extension to December 2008*

##### **Principal Investigator**

Alfredo Alves, EMBRAPA/CNPMPF; aalves@cnpmf.embrapa.br  
Caixa Postal 007, 44380-000, Cruz das Almas, BA, Brazil; +55-75-33128020;  
Fax: +55-75-33128096

##### **Collaborating institutions and scientists**

- Collaborators at EMBRAPA/CNPMPF: Miguel Dita; Antonio Souza
- EMBRAPA/CPATSA, Brazil: Alineaurea Florentino Silva
- CIAT: Martin Fregene; Hernán Ceballos
- Cornell University, USA: Tim Setter; Luis Duque
- IITA, Kenya: Morag Ferguson; Rosemary Mutegi
- ARI: Geoffrey Mkamilo
- SARI: Joseph Adjebeng-Danquah
- KARI: Joseph Kamau

##### **1. Evaluation of contrasting cassava varieties and identification of traits related to drought tolerance**

Drought tolerance contrasting cassava varieties have been screened and evaluated in four representative semi-arid environments in Brazil, Colombia, Tanzania and Ghana. Several trials have tested some hypotheses to characterise genotypes for drought tolerance traits, including: 1) genotypes that accumulate substantial stem starch reserves may be more able to sustain meristems and other respiring organs during a prolonged stress; 2) genotypes that maintain deep fibrous root growth would perform better in drought; 3) genotypes that maintain partitioning to storage roots and have high harvest index yield better in drought; and 4) genotypes that are more effective in closing stomata might be more able to maintain water status and retain leaves.

Leaf conductance, leaf retention, leaf size, harvest index, ABA, sugars and starch accumulation in the leaves and stems were differentially affected in the cassava varieties. A strong correlation ( $r=0.9$ ) was found between storage root yield under water stress (WS) and harvest index (HI). Genotypes with better HI accumulated more starch in stems, consistent with our hypothesis that better genotypes store and remobilise stem starch during stress.

Leaf abscisic acid (ABA) concentration at advanced stage of stress (60 d) was negatively correlated with HI, suggesting that it might be a useful phenotyping trait. ABA in stems was correlated with leaf carbohydrate, while ABA in leaves was correlated with stem carbohydrate. The negative correlations are consistent with the hypothesis that better genotypes (higher carbohydrate levels) are able to keep ABA levels lower.

In general, yield and other traits in WS were correlated with performance in well-watered conditions, indicating that improvement for these traits in WS environments is not likely to penalise yield in well-watered environments. The different responses to water deficit for the evaluated parameters has helped us to better discriminate varieties as tolerant or susceptible to drought and to define the best traits for selecting in breeding programmes.

## 2. Development and evaluation of segregating populations for drought tolerance

Contrasting cassava varieties were crossed and segregating populations of sexual seeds obtained at CIAT. The embryos of each seed were rescued and micropropagated *in vitro*. A total of 370 individuals from three populations (Table 1) were produced. Each individual was cloned and around five *in vitro* copies of each genotype were shipped to the target sites (Brazil, Colombia, Tanzania, Ghana, and Kenya) for multiplication and further phenotypic evaluation. Protocols for hardening and rapid micropropagation of *in vitro* cassava plantlets have been developed

**Table 1: Segregating populations developed for genetic mapping of drought tolerance in cassava.**

Family	Tolerant	Susceptible	# of individuals
F <sub>1</sub>	COL 1734	VEN 77	235
F <sub>1</sub>	BRA 255	COL 1468	61
S <sub>1</sub>	COL 1734	-	74

A genotypic parental screen has revealed 168 polymorphic SSRs in COL1734 x VEN77 and 147 in COL 1468 x BRA 255, from 307 SSRs screened. Genotyping of the mapping populations is underway.

## 3. Tangible outputs delivered

Drought tolerant and drought susceptible cassava varieties selected based on traits identified as highly correlated to yield performance under water deficit conditions.

Three mapping populations developed and ready to be field phenotyped and genotyped with molecular markers to identify QTL for drought tolerance.

## 16. G3005.05: Unlocking the genetic diversity in peanut's wild relatives with genomic and genetic tools

*January 2004–December 2007; no-cost extension to July 2008*

### Principal Investigator

José Valls and co-workers: Patrícia Guimarães; Soraya Leal-Bertioli; Marcio Moretzsohn.

EMBRAPA, valls@cenargen.embrapa.br

EMBRAPA, Recursos Genéticos e Biotecnologia, Parque Estação Ecológica Norte, Final Av W5 Norte, Brasília-DF, 70770-900, Brazil

### Collaborating institutions and scientists

- UCB: David Bertioli
- UCB, Goiás: Wellington Martins

- CERAAS: Ousmane Ndoye
- ICRISAT: Vincent Vadez: Rajeev Varshney: Aruna Rupakula
- UAS: Udaya Kumar
- Agropolis–CIRAD: Angelique d’Hont
- IBONE: Guillermo Seijo: Germán Robledo
- University of Aarhus, Denmark: Jens Stougaard: Lene Madsen: Niels Sandal
- Texas Agricultural Experiment Station, USA: Consultant; Charles Simpson

### **Context**

The limited allelic diversity present in cultivated germplasm, has constrained both breeders and the advances in genetics necessary for modern breeding. Wild *Arachis* species are a rich source of new alleles for peanut improvement, and have the polymorphism necessary for the genetic characterisation of these alleles. The project aimed to begin the process of incorporating wild alleles into cultivated peanut, and to build an initial genetic and genomics toolkit that would enable modern molecular breeding using wild alleles.

### **Findings and implications**

A genetic map for the AA genome component of peanut was constructed being mostly based on microsatellite markers, which are ideal for transference to other peanut mapping populations and universal legume anchor markers, which are ideal for the comparison of legume genomes. It also includes candidate disease resistance genes and overlapping QTLs for resistance to late leaf spot in wild diploids, giving a first glimpse of the genomic regions that control disease resistance in *Arachis*. The map has successfully provided a genetic framework for molecular breeding in peanut, especially, because of higher polymorphism levels, for lines that incorporate wild alleles, but also for breeding using cultivated x cultivated crosses. Using the universal legume anchor markers, and other sequence characterised markers, the *Arachis* genetic map was substantially aligned with the genomes of the model legumes *Lotus* and *Medicago*. This enables the use of the genome sequences of the model legumes to improve our understanding of the *Arachis* genome. The levels of macrosynteny between *Arachis* and the model legumes within ten synteny blocks appear sufficient to aid in gene cloning and candidate gene identification.

To improve genomic resources for peanut two large-insert libraries in Bacterial Artificial Chromosome (BAC) vector, one for each of the most probable diploid ancestral species of cultivated peanut were constructed. The libraries (AA and BB) are respectively c. 7.4 and c. 5.3 genome equivalents with low organelle contamination and average insert sizes of 110 and 100 kb. Both libraries were used for the isolation of clones containing genetically mapped legume anchor markers, and resistance gene analogues. The first links between genetic and cytogenetic maps were created by using BAC clones with single copy genes in fluorescent *in-situ* hybridisation (FISH). In addition, FISH was used to explore the repetitive elements within the AA and BB component genomes of peanut, and as a method for tracking the introgression of wild genome segments into cultivated peanut.

The responses to progressive water deficit in wild, synthetic and cultivated peanut were investigated. Although the transpiration behavior of synthetics was observed as being distinct from their wild parentals, transpiration efficiency was similar, showing that direct screening of wild species for desirable drought responses needs to be interpreted with caution. Large variations of transpiration response were found between different wild species, and a surprisingly high variation in cultivated peanut was observed. In general, wild accessions had a “conservative” behavior: transpiration decreasing dramatically when the fraction of transpirable soil water was high (0.8 – 0.6). On the other hand, the transpiration of cultivated peanut varieties declined at lower soil water content (FTSW *c.*0.2), showing a more “opportunistic” behavior regarding water use.

Three synthetic amphidiploids (previously produced in EMBRAPA and Texas A&M University) were transferred to ICRISAT-India and to CERAAS/ISRA in Senegal. Here, and in Brazil, synthetics are being used for the production of introgression lines for mapping and breeding, work that is being taken further forward by other GCP funded projects. The work has therefore provided an effective start for the incorporation of wild alleles, conferring, for instance, enhanced disease resistance, into peanut varieties in Asia, Africa and South America using modern molecular breeding.

#### **Links to previous projects**

European Union: INCO-DEV, Contract ICA4-CT-2001-10072 Project “ARAMAP”  
World Bank and EMBRAPA: PRODETAB project 004/01/01.

#### **Links to new projects**

GCP: Application of molecular tools for controlled wild introgression into Peanut cultivated germplasm in Senegal. A capacity building grant.  
GCP: Tropical Legume Initiative, TLI.

### **17. G3005.06: Marker development and marker-assisted selection for Striga resistance in cowpea**

*January 2005–December 2007; no-cost extension to December 2008*

#### **Principal Investigator**

Satoru Muranaka: IITA, [s.muranaka@cgiar.org](mailto:s.muranaka@cgiar.org)  
IITA Kano, Sabo Bakin, Zuwo Road, PMB3112, Kano, Nigeria, (+234 2) 241 2626

#### **Collaborating institutions and scientists**

##### **Project members**

- IITA: Christian Fatokun, Adebola Raji, Boukar Ousmane, Dong-Jin. Kim
- UVA: Michael Timko
- CERAAS: Ndiaga Cisse
- ISRA–Centre National de Recherches Agronomiques: Moctar Wade

#### ***Introduction/Background***

The present project seeks to develop a MAS strategy for cowpea that will allow rapid and reliable identification of race-specific *Striga* resistance genes in breeding lines and

integration of MAS for *Striga* resistance in cowpea breeding programmes. The ability of cowpea genotypes to resist *S. gesnerioides* parasitism depends on geographic origin of the parasite. Based on the differential resistance reaction exhibited by various cowpea genotypes, five different races (*SG1*, *SG2*, *SG3*, *SG4* and *SG5*) have been identified in West Africa (Lane et al., 1997a; 1997b). We also know that there is a sixth *Striga* race (designated *SG6*) predominant in Senegal. Five *Striga* resistance genes (Rsg1-1, Rsg2-1, Rsg3-1, and Rsg994-1, Rsg4-3) have been identified in cowpea genotypes (Ouédraogo et al., 2001). These genes map into two clusters in the cowpea genome with markers linked to the *SG1* and *SG3* resistance genes Rsg1-1, Rsg2-1, and Rsg4-3 present in the resistant lines B301, IT82D-849 and Tvu14676, respectively, mapping to LG1 (Ouédraogo et al., 2001), and *SG1* resistance genes Rsg3-1 and Rsg994-1 found in Suvita-2 and IT81D-994 mapping to LG6 (Ouédraogo et al., 2002).

#### ***Development of SCAR Markers for Striga resistance***

One of the major goals of our work is the development of a set of reliable markers for race-specific resistance to *S. gesnerioides*. We have generated one primer combination (dubbed 61R) that is effective for tagging resistance to races *SG3* (Nigeria/Niger) and *SG1* (Burkina Faso), and a derivative primer combination (MahSE2) that is equally, if not more, effective in selective amplification of a resistance marker. Studies were carried out to determine the correlation between the presence of the 61R and MahSE2 markers and the *Striga* resistance phenotype in materials previously reported to be either resistant or susceptible. We found about 77% (MahSE2) and 79% (61R) with correlation using 60 cowpea genotypes including local varieties and breeding lines for *Striga* race 3 (Nigeria/Niger) in pot experiment. Also, MahSE2 was tested on GCP cowpea mini-core collections (370 accessions) and approximately 17% (63 accessions) of them showed the presence of MahSE2.

Following BSA and analysis of individuals, we determined that *SG5* resistance is not effectively marked by 61R or MahSE2. While these two markers segregate weakly with resistance, the recombination frequency is high, giving approximate distances of 5-10 cM. The primer sequences that amplify ~200 bp band in IT-84S-2246-4 and in susceptible SCAR markers 61R and MahSE2 were also found to be weakly linked to resistance to *SG2*. The recombination frequency suggests that resistance to *SG2* is closer to *SG1* and *SG3* than is *SG5*.

In a separate approach we located one marker from a screen of potential TRAP markers. Marker AGG/CTT 200B showed polymorphic fragments of ~150 and 300 bp that were linked to susceptible phenotype of *SG5*. The PCR products of AGG/CTT 200B were isolated and their nucleotide sequence determined and new primers designed for use in additional rounds of BSA. One of the primer combinations AGG/CTT 200BR/50R showed a single band present in IT84S-2246-4 and segregating with the susceptible phenotype in bulks prepared from both *SG5* and *SG3*.

#### ***Screening of cowpea genotypes in Striga ‘hotspots’ in West Africa***

A total 37 genotypes were tested for 6 different *Striga* races (*SG1*: Burkina Faso, *SG2*: Mali, *SG3*: Nigeria/Niger, *SG4*: Zakpota in Benin, *SG5*: Cameroon and *SG6*: Senegal) in

two multi-location trials carried out in 2005 to 2007. Through the trials, we detected 3 multi-*Striga* resistant genotypes, TVu14676 (SG1, SG2, SG3, SG5 and SG6), B301 and IT98K-208-5 (SG1, SG2, SG3 and SG5) which is resist to more than 4 different *Striga* races. IT98K-205-8 is recommendable for *Striga* resistance breeding in Burkina Faso, Mali, Niger, Nigeria and Cameroon due to adequate seed characteristics (white colour, rough texture and medium size), early maturing and higher productivity. Several breeding lines, IT98K-216-44 (Extra late), IT81D-994 (Extra late) and IT98K-503-1(Extra early) showed resistance to the Zackpota race (SG4z) which two famous *Striga* resistant lines B301 and TVu14676 can not overcome.

#### ***Development of populations for MAS efficiency test for Striga resistance***

So far, 4 self-crossing populations from various combinations of resistant and susceptible genotypes have been created in Kano to test the efficacies of MAS and have been advanced two-three generations in 2005-2008. The leaf samples were taken from F<sub>2</sub> –F<sub>5</sub> plants of 4 populations for genotyping and all population will be tested for SG3 (Niger/Nigeria) resistance at F<sub>5</sub> generation in pot condition to estimate MAS efficiency of existing two SCAR markers 61R and MahSE2. One of the populations, IT06K-42 series (38 lines) from a cross of IT97K-499-35 x Dan Ila has just reached to F<sub>5</sub> stages and have been phenotyped for SG3 resistance.

#### ***Diversity analysis of Striga in Nigeria (SG3) and Senegal (SG6)***

To identify the diversity of *Striga* race existing in Senegal, selected 9 cowpea lines (IT81D-994, B301, 58-57, Mouride, CB5, Mougne, Suvita 2, UCR1115 and Tvx-3236) were tested on *Striga* seeds collected from different locations of Senegal. Also, Moctar Wade, (ISRA-CNRA) and a visiting graduate student from Senegal, Preliminary results based on AFLP profile analysis showed a clear separation between the populations of *Striga* adapted for growth on cowpea, with considerable inter-population variability. Genetic distances between isolates from Sindiene and other isolates were large. Clear differences were recognised between the races of the parasite present on cowpea in Senegal and those on wild legumes.

#### ***Capacity building and training***

In UVA, one of the collaborators of the project, Moctar Wade (ISRA-CNRA) spent 37 days from May 14 to June 20, 2007 as a visiting scientist. Francois Badiane, a PhD candidate from the University of Dakar, was trained in the laboratory in molecular marker techniques during August – October, 2007. Currently, Mr. Lucky Omoigui (IITA) and Mr. Gonne Sobda (IRAD) have been trained for basic molecular biology and MAS.

At ISRA-CNRA, Bambey, a CERAAS student, Ms. Charlotte Tonnesia involved in the characterisation of *Striga* races in Senegal is working with us in ISRA-CNRA Bambey. She spent 6 months at UVA (15 January-17 July 2006) working on the "characterisation of genetic diversity in the different sources of *Striga* using molecular marker". During this rainy season, she has involved in the layout and follow-up of the cowpea *Striga* resistant variety trials at Ngalbane, Ngoye and Sindienne.



## **18. G3005.09: Development of low-cost technologies for pyramiding useful genes from wild relatives of cassava into elite progenitors**

*January 2005–December 2007*

### **Principal Investigator**

Anthony Bellotti, CIAT; [A.Bellotti@cgiar.org](mailto:A.Bellotti@cgiar.org)

A.A.6713, Cali, Colombia

### **Collaborating institutions and scientists**

- CIAT: Martin Fregene
- EMBRAPA-CNMPF: Alfredo Alves
- CIAT/NRCRI: Emmanuel Okogbenin; Chiedozie Egesi
- CRI: Elizabeth Okai
- NaCRRI: Yona Baguma; Anthony Pariyo

Molecular markers associated with genes controlling traits of agronomic interest reduce by more than half the time required for transfer of genes between and within cassava gene pools. In this project development of molecular markers for resistance to the cassava green mites (CGM), the whiteflies, and delayed PPD found in wild *Manihot* species were completed or at an advanced stage.

## **1. Low cost technologies for pyramiding pest, disease and PPD genes**

### ***1.1 Development of low cost markers for CMD resistance***

To simplify assays, RME1, the polymorphic RAPD fragment for CMD resistance, was converted to a Sequence Characterized Amplified Region (SCAR) marker, a low cost and easy to use marker. The SCAR marker is now routinely being used for MAS at CIAT and primer sequences has been sent to NARs partners in preparation for its use in MAS in their breeding programs.

### ***1.2 Testing of the FTA paper for low cost MAS of CMD resistance in breeding populations***

To reduce the cost of marker genotyping, PCR amplification of leaf squashes on FTA paper (Whatmann Inc., UK) was tested. PCR amplification of FTA paper discs with leaf squashes from in vitro plants was 100% successful for both RAPD and PCR markers with or without the washing step. This result suggests that FTA paper leaf squashes could replace cumbersome DNA isolation step. However, a two-step DNA isolation method that was later developed showed lower costs compared to using FTA paper and the two-step method is the currently the method of choice for MAS in cassava breeding.

## **2: Molecular markers associated with genes for resistance to pests, diseases, and delayed PPD**

### **2.1 Post-harvest Physiological Deterioration (PPD) in Cassava**

Dramatically delayed PPD was earlier identified in an inter-specific hybrid (CW429-1) between cassava and *Manihot walkerae*. Multi-locational evaluation for delayed PPD of

CW 429-1 and 8 other elite genotypes indicate that results of mean PPD values at 10 DAH, ranged from 0% in CW429-1 to 58% in CM 523-7. The same trend was observed 14 DAH with CW 429-1 still displaying no visible sign of deterioration. Three linkage groups with eight putative markers for delayed PPD were identified accounting for 6.2 to 12.8% of phenotypic variance.

## 2.2 Cassava Green Mite

Good resistance to cassava green mites (CGM) was identified in 4 inter-specific hybrid families, CW68, CW65, CW67, and CW66, derived from a cross between cassava and an accession of *Manihot esculenta* sub spp *flabellifolia*. Three markers were found associated with resistance and they are SSRY11, NS1099 and NS346.

## 3: New collections of wild *Manihot* species for identification of useful genes

Over 1600 accessions from a total of 28 *Manihot* species were collected from: the CNPMF field collection at Cruz das Almas, natural populations of *Manihot* species around Bahia State, and the “cerrado” region of Brasília and surrounding areas. A total of 1740 sexual seeds from 175 families representing 5 wild *Manihot* species and their inter-specific hybrids with cassava were distributed to participating NARs of Brazil, Nigeria, Ghana, and Uganda for field establishment and evaluation for pest and diseases endemic in their own environments.

Other achievements of the project is the development of easily assayed markers for MAS and the building of capacity, both human and infrastructure, at 3 African NARS for molecular breeding. Cassava breeders from participating NARs, namely: Brazil, Uganda, Ghana, and Nigeria, along with cassava breeders from other African countries were trained in field-based and molecular breeding at CIAT for one month in 2005. Simple molecular marker labs consisting of DNA isolation equipment, PCR machines, and gel electrophoresis apparatus were also purchased and installed at the three participating African NARs (NAARI, CRI and NRCRI). Significant progress was also made by the African NARs in receiving, evaluating and submitting for release cassava germplasm coming out of the project at both CIAT and NARs.

## 19. G3005.12: Drought tolerant rice cultivars for North China and South/Southeast Asia by highly efficient pyramiding of QTLs from diverse origins

*January 2005–December 2008; no-cost extension to December 2008*

### Principal Investigator

Zhi-Kang Li, Institute of Crop Sciences, Chinese Academy of Agricultural Sciences (CAAS); [z.li@caas.net.cn](mailto:z.li@caas.net.cn), 12 South Zhong-Guan-Cun St., Beijing 100081, China; (86) 10-62136040; Fax: (86) 10-68918559

### Collaborating institutions and scientists

- Rice Research Institute, LAAS, China: Ze-Tian Hua
- Shenyang Agricultural University of China: Zheng-Jin Xu
- Institute of Crop Sciences, CAAS: Yongming Gao and Jianlong Xu
- IRRI: Arvind Kumar; Mary Smith; Hugo van Houten

To develop superior rice lines with high yield potential and significantly improved drought tolerance (DT) and/or water use efficiency (WUE) for North China and South/Southeast Asia, 3 large sets of DT introgression lines (ILs) and pyramiding lines (PDLs) in two elite japonica (Liaojing 454 and C418) and one indica (IR64) genetic backgrounds were developed using BC breeding and the approach of designed QTL pyramiding. More than 1000 DT ILs and PDLs were genotyped with polymorphic SSR markers and phenotyped for grain yield and related traits in replicated progeny tests under normal irrigated and drought stress to discover and characterise QTLs and QTL networks underlying DT in rice. The three important results have been obtained through these research activities:

First, a number of promising DT/WUE lines and hybrids have been developed. These included 6 promising Liaojing454 ILs (Table 1). Of these, two DT lines (HR94 and HR95) and a hybrid, Liaoyou 5224 have passed the multilocation yield trials in North China. HR94 and HR95 produced ~7.5 t/ha under the completely rainfed conditions in the farmer's field. Liaoyou 5224 was tested in the multilocation yield trial of North China with an average yield advantage of 16.6% over the check across all six locations. This hybrid was grown in 200 ha of demonstration under the rainfed conditions in 2007 and has been recommended to farmers in 2008. In addition, 4 new hybrid cultivars derived from DT C418 ILs (as restorers) and 6 inbred japonica DT lines are being tested in the multi-location yield trials in North China in 2008 summer. Many promising lines developed from pyramiding crosses between DT ILs are in the pipeline to be tested in the yield trials. From 8 2<sup>nd</sup> round pyramiding crosses derived from crosses between 1<sup>st</sup> round PDLs, 12 promising indica PDLs were identified, which have shown good DT and yield potential in two consecutive seasons (Table 2), which have been sent to NARES and farmers to be tested in the rainfed areas of South/Southeast Asia countries.

**Table 1. The yield performances and yield related traits for 5 promising DT Liaojing454 ILs**

Line designation	PH (cm)	PN	PL (cm)	GN	Fertility (%)	TGW (g)	Grain yield (t/hm <sup>2</sup> )	Over CK (%)
071D253	112.3	15.6	17.1	110	97.2	25.3	9.88	6.13
071D267	113.2	14.7	17.1	116	98.3	26.1	9.91	6.43
071D274	112.4	17.4	17.3	114	96.3	26.1	10.54	13.22
071D302	115.6	16.8	18.2	119	97.1	25.9	10.48	12.56
Liaonong 21 (CK)	113.2	14.6	18.2	110	95.2	25.3	9.31	
<b>HR94</b>	<b>104.2</b>		<b>17.1</b>	<b>104</b>	<b>95.3</b>	<b>25.3</b>	<b>7.41</b>	
<b>HR95</b>	<b>107.4</b>		<b>17.3</b>	<b>111</b>	<b>93.3</b>	<b>25.9</b>	<b>7.53</b>	

The data of HR94 and HR95 were from a farmer's field grown under rainfed conditions.

**Table 2. Grain yield performance of 12 promising 2<sup>nd</sup> round pyramiding IR64 lines under irrigated control and stress at reproductive stage during 2008 dryseason**

Line designation	Grain yield under control		Grain yield under stress		Mean	
	(kg/ha)	(%) over CK	(kg/ha)	(%) over CK	(kg/ha)	(%) over CK
IR 83140-B-11-B-B	5579.0	27.1	1216.1	115.7	3397.6	37.2
IR 83142-B-57-B-B	4876.9	11.1	1045.0	85.4	2960.9	19.5
IR 83142-B-7-B-B	4626.8	5.4	1184.8	110.2	2905.8	17.3
IR 83142-B-60-B-B	4150.1	-5.5	1576.9	179.7	2863.5	15.6
IR 83140-B-32-B-B	4596.4	4.7	1064.0	88.7	2830.2	14.3
IR 83141-B-11-B-B	4849.4	10.5	798.1	41.6	2823.7	14.0
IR 83142-B-79-B-B	4809.2	9.5	1110.4	97.0	2959.8	19.5
IR 83142-B-46-B-B	4881.2	11.2	916.8	62.6	2899.0	17.0
IR 83140-B-36-B-B	4551.3	3.7	1537.3	172.7	3044.3	22.9
IR 83140-B-6-B-B	5236.4	19.3	598.7	6.2	2917.6	17.8
IR 83143-B-39-B-B	4520.8	3.0	1189.3	111.0	2855.1	15.3
IR 83142-B-19-B-B	4624.8	5.3	1117.3	98.2	2871.0	15.9
<b>Mean</b>	<b>4775.2</b>	<b>8.8</b>	<b>1112.9</b>	<b>97.4</b>	<b>2944.0</b>	<b>18.9</b>
IR72 (CK)	4390.3		563.7		2477.0	

Second, large numbers of QTLs underlying DT or WUE were discovered, which are widely distributed across the rice genome. These QTLs in the DT ILs formed complex multilocus structure, or genetic networks of clear hierarchy, and also are under epigenetic control. This result provides new insights into the genetic basis of DT and other complex phenotypes in rice. In addition, the genetic and phenotypic information of DT QTLs and related phenotypes in 1000+ ILs and PDLs has been obtained, which will be deposited into the GCP databases soon.

Third, a new marker-assisted breeding strategy – trait improvement by designed QTL pyramiding has been developed and demonstrated to be a highly efficient breeding approach for genetic improvement of complex traits, particularly for abiotic stress tolerances in rice. Finally, 3 postdoctoral fellows, 12 PhD and 20 MS students have been trained in rice molecular genetics and breeding during the 3 years.

In conclusion, significant progress has been made in the GCP funded project in developing DT rice lines and identifying DT QTLs. We have demonstrated that designed QTL pyramiding using selected ILs and DNA markers is an efficient strategy to integrate the conventional breeding based on mass selection with gene/QTL discovery for complex traits. Some promising lines with significantly improved DT/WUE and high yield potential in both elite japonica and indica backgrounds have been developed, which will be released to rice farmers in the rainfed areas of South/Southeast Asia and North China in 1-2 years.

**20. G3007.04: Tailoring superior alleles for abiotic stress genes for deployment into breeding programmes: A case study based on association analysis of *Alt<sub>SB</sub>*, a major aluminum tolerance gene in sorghum (ALTSORGHUM)**

*December 2007–December 2009*

**Principal Investigator**

Jurandir Magalhaes, EMBRAPA Maize and Sorghum; jurandir@cnpms.embrapa.br  
Rd MG 424, Km 65, Sete Lagoas - MG 35701-970, Brazil; (55) 31-37791183;  
Fax: (55) 31-37791179

**Collaborating institutions and scientists**

- Robert W. Holley Center for Agriculture and Health, USDA-ARS: Leon Kochian; Owen Hoekenga; Jiping Liu
- Cornell University (Institute for Genomic Diversity, IGD): Stephen Kresovich; Sharon Mitchell; Theresa Fulton; Alexandra Casa
- EMBRAPA Maize and Sorghum: Claudia Guimaraes; Robert Schaffert; Antonio Marcos Coelho; Vera Alves
- INRAN: Soumana Souley; Maman Nouri; Magagi Abdou; Adam Kiari; Fatouma Beidari

**1. Rationale for the ALTSORGHUM project**

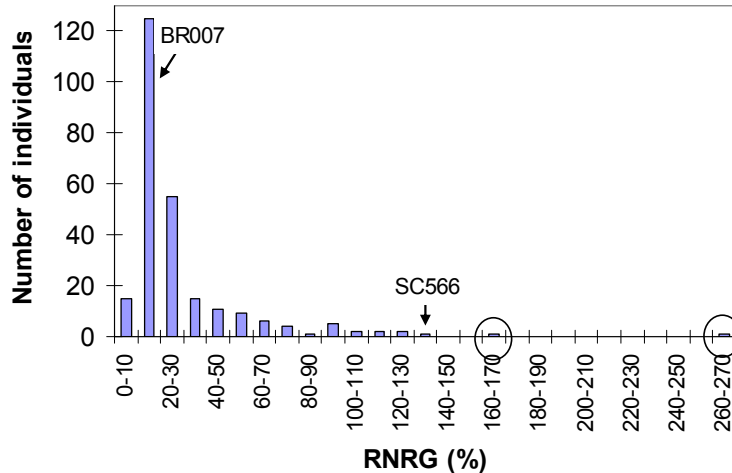
The project is based on the positional cloning and characterisation of *Alt<sub>SB</sub>*, a major gene conferring tolerance to aluminum (Al) toxicity in sorghum, which seriously constrains agriculture in several regions of acid soils throughout the world. Association analysis is being used to identify *Alt<sub>SB</sub>* haplotypes that confer superior tolerance to Al toxicity in sorghum, if possible to dissect the molecular nature of *Alt<sub>SB</sub>*, develop haplotype specific markers, and by molecular breeding develop near-isogenic lines (NILs) carrying elite haplotypes.

**2. Partial Results**

**2.1 Germplasm Exchange:** Three association panels in sorghum are being merged into a single, unified panel. To this end, the association panel developed by CIRAD (landraces CIRAD – LR-CIRAD) was transferred to Cornell University whereas the association panel being used by Cornell University (LR-Cornell) is now in quarantine in Brazil. Therefore, all the germplasm transfer processes were completed.

**2.2 Phenotyping for Aluminum Tolerance:** In our proposal we presented Al tolerance data for 140 accessions belonging to LR-CIRAD. We have completed phenotyping of the entire LR-CIRAD combined with the inbred lines that are used for Al tolerance studies at Embrapa (IL-Embrapa), totaling 256 sorghum accessions.

The panel was found to display extensive phenotypic variation for Al tolerance in sorghum. Seven new, highly Al tolerant accessions were identified in LR-CIRAD. Seventy five percent of the accessions showed mean RNRG < 30%, RNRG for 19% of the accessions were between 30% and 80% and 6% of the accessions in the panel were highly tolerant to Al toxicity (RNRG>80%).



**Figure 1.** Percent relative net root growth (RNRG) frequency distribution for 256 accessions belonging to LR-CIRAD and IL-Embrapa. Individuals were grown in nutrient solution containing  $\{27\} \mu\text{M Al}^{3+}$  for 5 days. The RNRG class position for the sensitive standard, BR007, and the tolerant elite line SC566 are shown. Circles indicate the position for two accessions showing extreme Al tolerance.

Accessions showing intermediate to high levels of Al tolerance are useful for breeding Al tolerance in sorghum. Most importantly, accessions showing extreme levels of Al tolerance are candidates for harboring elite *Alt<sub>SB</sub>* haplotypes. This will be confirmed by association analysis and NILs so that tag SNPs can be developed for MAS in diverse germplasm.

**2.3 Development of NILs carrying elite *Alt<sub>SB</sub>* haplotypes:** This activity aims at delivering a set of NILs carrying *Alt<sub>SB</sub>* haplotypes from Al tolerant accessions in the association panel, which will also allow allelic comparisons across a uniform genetic background. Twenty seven intermediately to highly Al tolerant accessions from LR-CIRAD were crossed to an Al sensitive breeding line from the Embrapa programme, BR007, and these activities are progressing towards BC<sub>3</sub>F<sub>2</sub> NILs using MAS for *Alt<sub>SB</sub>*.

**2.4 Association analysis of *Alt<sub>SB</sub>*:** We have used the SSR genotypic data obtained within SP1 to assess population structure and relatedness within the LR-CIRAD. In addition, we have genotyped the IL-Embrapa with the same SSR set used for developing the sorghum composite set and successfully used the control lines posted on ([http://sat.cirad.fr/sat/sorghum\\_SSR\\_kit/](http://sat.cirad.fr/sat/sorghum_SSR_kit/)) to merge the SSR data for both panels. Preliminary analysis with amplicons within the *Alt<sub>SB</sub>* locus highlighted SNPs and indels significantly associated with Al tolerance. In addition, analyzing different NILs we have found that regulatory factors are important for achieving high levels of Al tolerance in sorghum. The USDA, Embrapa and IGD labs have initiated efforts to develop a pilot Illumina platform in sorghum, which will be further expanded and used to identify these regulatory factors for breeding purposes. In addition, accessions within LR-Cornell are in the process of being phenotyped for Al tolerance.

**3. Tangible outputs delivered:** We have characterised for Al tolerance 256 sorghum accessions comprising the LR-CIRAD and IL-Embrapa. This data is available for sorghum breeders interested in developing highly Al tolerant sorghum cultivars.

## **21. G3007.05: Detecting and fine-mapping QTLs with major effects on rice yield under drought stress for deployment via marker-aided breeding**

*August 2007–July 2009*

### **Principal Investigator**

Arvind Kumar, International Rice Research Institute (IRRI); akumar@cgiar.org  
DAPO BOX 7777, Metro Manila, Philippines; (63) 2-580-5600; Fax: (63) 2-580-5699

### **Collaborating institutions and scientists**

- IRRI: Venuprasad Ramaiah; Jerome Bernier; Shalabh Dixit; Rowena Oane; R. Anitha; R Serraj; David J. Mackill
- IGAU: SB Verulkar
- CRRI: ON Singh; P Swain; LK Bose
- CRURRS: PK Sinha; N.P. Mandal
- NDUAT: JL Dwivedi
- UAS: S Hittalmani; Venkatesh Gandhi
- TNAU: R Chandrababu; A Senthil, S Robin
- BAU: BN Singh; RL Mahato
- Jawahar Lal Nehru Krishi Vishwavidyalaya, India: P. Perraju
- Barwale Foundation, Hyderabad, India: H.E. Shashidhar; Abhinav Jain
- YAAS: D Tao
- University of Alberta, Canada: Dean Spaner

### **1.1 Fine-mapping loci affecting yield under drought stress**

A large effect QTL (qtl12.1) for grain yield under drought located on a 10 cM region of chromosome 12 (Bernier et al. 2007) had been identified in Vandana/Way Rarem population. Using BC<sub>1</sub>F<sub>2</sub>-derived lines, this locus is now mapped to a distance of 1 Mbp (about 3 cM) in the region spanning 16.5 to 17.5 Mbp on chromosome 12 and is being fine mapped on BC<sub>3</sub>F<sub>2</sub>-derived population. In IR55419-04/Way Rarem, a large effect QTL between markers RM17403 and RM8220, qtl4.1 that explained 33% of the genotypic variation and was also found to affect drought-response index was detected. qtl4.1 is being fine mapped on BC<sub>3</sub>F<sub>2</sub> derived population. In Apo/IR64 population, fine mapping of RM404 locus that showed strong effect ( $P < 0.0001$ ; Genetic  $R^2 = 68\%$ ) on grain yield under stress but no effect under non-stress conditions on BC<sub>2</sub>F<sub>2</sub> derived population is also underway.

### **1.2 Identification of new QTLs affecting grain yield under drought stress**

In IR77298-14-1-2/IR64 population, an introgression on chromosome 2 (RM236-RM279) had shown highly significant effect ( $R^2 = 10-33\%$ ) on grain yield under stress but no effect of yield under non-stress. Another QTL affecting grain yield under drought was identified in Vandana/IR64 population on Chromosome 4 near RM131 and its effect was subsequently confirmed in Vandana/<sup>2</sup>\*IR72 and Vandana/<sup>3</sup>\*IR64 population. This region seems to be very close to the RM17403-RM8220 qtl4.1 region identified in IR55419-04/Way Rarem population. Populations derived from Apo/<sup>2</sup>\*Swarna, Apo/<sup>3</sup>\*IR72, and

Apo/<sup>3\*</sup>IR72875 were evaluated in DS2008 under upland drought conditions. The results indicate that QTL detected in one background are repeatable in other backgrounds although the magnitude of effect varied varies from population to population.

### **1.3 Physiological effects of drought yield QTLs**

The improvement in performance of lines with qtl12.1 for higher grain yield, harvest index and biomass accumulation over lines without qtl12.1 under drought conditions is likely to be due to an increase in plant water uptake resulting from more effective root architecture. The lines with the QTL had an 18% higher below 30 cm deep root length compared to lines without QTL, This appears to be the most important difference in explaining the increased water uptake in the lines with the favorable allele of the QTL. No differences between the two sets of lines were present under well-watered conditions.

### **2.1 QTLs effect validation in India and China; simulation studies in Canada**

At IRRI, the relative effect of the QTL on grain yield was shown to increase with increasing intensity of drought stress- from having no effect under well-watered conditions to having an additive effect of more than 40% of the trial mean under severe drought stress. Similar finding were observed in diverse environments in India (Bernier et al. 2008). High yielding drought tolerant breeding lines developed from different populations are being evaluated on partners' sites for dissemination of the promising lines. Computer simulation studies of bidirectional and unidirectional selective genotyping for QTL analyses, using population sizes and selection intensities relevant to cereal breeding revealed bidirectional selective genotyping to be more powerful than unidirectional. With genotyping of the best 30 out of 500 lines (6%), a QTL explaining 15% of the phenotypic variance could be detected with a power of 0.8 when tests were conducted at a marker 10 cM from the QTL. With bidirectional selective genotyping, QTL with smaller effects and (or) QTL farther from the nearest marker could be detected.

### **References**

- Bernier J, Kumar A, Ramaiah V, Spaner D, and Atlin G (2007). A large-effect QTL for grain yield under reproductive-stage drought stress in upland rice. *Crop Sci.* 47: 505-516.
- Bernier Jérôme, Kumar Arvind, Spaner Dean, Verulkar Satish, Mandal Nimai P, Sinha Pramod K, Perraju Puvvada, Dongre Praba R, Mahato RN and Atlin GN (2008). Characterisation of the effect of rice drought tolerance qtl12.1 over a range of environments in the Philippines and eastern India. *Euphytica* (Press)



---

## ***Subprogramme 4: Bioinformatics and crop information systems***

### **22. G3005.04: An eco-physiological-statistical framework for the analysis of G×E and QTL×E as occurring in abiotic stress trials, with applications to the CIMMYT drought stress programmes in tropical maize and bread wheat**

*January 2005–December 2007; no-cost extension to July 2008*

#### **Principal Investigator**

Fred van Eeuwijk, WUR, fred.vaneeuwijk@wur.nl

Biometris, Wageningen University UR; P.O. Box 100, 6700 AC, Wageningen, The Netherlands; (31) 317-482902; Fax: (31) 317- 483554

#### **Collaborating institutions and scientists**

- CIMMYT: Jean-Marcel Ribaut; Matthew Reynolds; José Crossa
- CSIRO: Scott Chapman
- Universidad Autónoma Chapingo, México: Mateo Vargas

#### **1. Introduction**

The identification and description of the genetic basis (QTLs) of drought tolerance and other abiotic stresses is difficult due to genotype by environment interactions (G×E) at the phenotypic level, following from QTL by environment interactions (QTL×E) at the genetic level. This project aimed at providing breeders with a powerful tool for QTL mapping in series of stress trials. The methodology focused on: i) single-trait QTL×E models in which the QTL expression can be modeled in direct dependence on stress indicators, and ii) multi-trait QTL×E models, in which the genetic correlations between traits can be modeled in their dependence on the environmental conditions. The methods were developed and illustrated with real life data from CIMMYT tropical maize and bread wheat programmes.

#### **2. Research activities and deliverables**

##### **2.1 Single-trait QTL×E models**

A mixed model QTL mapping strategy for single-trait QTL×E analysis has been developed and illustrated for drought stress in wheat. Attention was given to the modeling of spatial field trends as can be frequently encountered in stress trials. In addition, it was shown how to include explicit physiological information in the QTL model. Various QTLs were identified that explained observed patterns of G×E for a set of wheat genotypes that were grown with and without water stress in Mexico and Australia. Details on method and application can be found in Mathews et al. (2008).

##### **2.2 Multi-trait QTL×E modeling**

Realistic breeding for stress tolerance leads inevitably to questions about the genetic basis of correlations between traits within and across environments. To answer such questions, a mixed model multi-trait multi-environment QTL mapping approach was developed, with an application to drought stress in maize. The methodology is described in Malosetti et al. (2007, 2008a), where it is shown how the approach allows investigating issues

related to: a) the causes underlying G×E (QTL×E), b) the causes of genetic correlations between traits (pleiotropic and linked QTLs), and c) the causes of changing genetic correlations between traits across environments.

### 2.3 Courses on G×E and QTL×E

In the course of this project, five courses were given at different locations that reached around 150 researchers and advanced students around the world (Fig 1). The typical course lasted at least two days and contained presentations of the theory alongside extensive hands-on practical sessions. A manual has been compiled containing a brief overview of the theory, together with instructions for applying the methodology to breeding data collected in a series of trials with varying stress levels (Malosetti et al., 2008b).

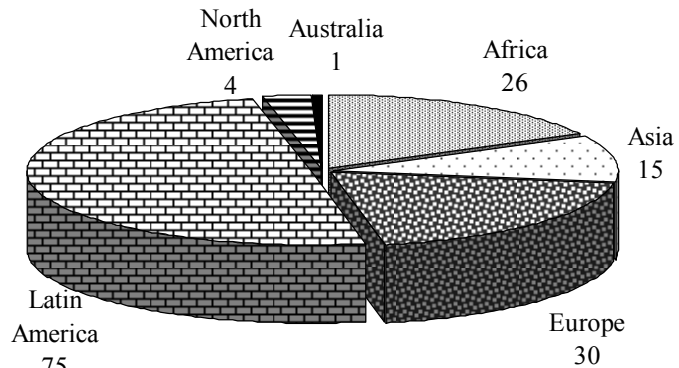


Fig 1. Number of course participants per region

### 3 Conclusions

This project has delivered a proven mixed model framework for QTL mapping in series of trials under different levels of abiotic stress. The main features of the methodology are: 1) detection of QTLs and QTL×E and estimation of the corresponding QTL effects; 2) modeling of QTL×E in relation to environmental covariables, where the latter may come from the application of crop physiology models to estimate stress levels; 3) extension of the methodology to multiple traits, allowing the investigation of genetic correlations between traits (pleiotropy and linkage). The methodology has been implemented in a set of programmes and procedures within the Discovery version of the statistical package Genstat. This version of Genstat is free for non profit organisations around the world. Various training courses were given, whose course notes can be obtained from the authors ([marcos.malosetti@wur.nl](mailto:marcos.malosetti@wur.nl); [fred.vaneeuwijk@wur.nl](mailto:fred.vaneeuwijk@wur.nl)).

### References

- Mathews KL, Malosetti M, Chapman SC, McIntyre L, Reynolds M, Shorter R and van Eeuwijk FA (2008). Multi-environment QTL mixed models for drought stress adaptation in wheat. *Theoretical and Applied Genetics* (in press).
- Malosetti M, Ribaut JM, Vargas M, Crossa J, Boer MP, van Eeuwijk FA (2007). Multi-trait multi-environment QTL modelling for drought-stress adaptation in maize. In: *Scale and Complexity in Plant Systems Research. Gene-Plant-Crop Relations* (Spiertz JHJ, Struik PC, van Laar HH ed). Springer, Dordrecht, The Netherlands, pp 25-36.
- Malosetti M, Ribaut JM, Vargas M, Crossa J, van Eeuwijk FA (2008). A multi-trait multi-environment QTL mixed model with an application to drought and nitrogen stress trials in maize (*Zea mays* L.). *Euphytica* 161:241-257
- Malosetti M, Ribaut JM and van Eeuwijk FA. The analysis of multi-environment data: modeling genotype by environment and QTL by environment interaction. In: *Drought phenotyping in crops: a practical manual* (Monneveux P and Ribaut JM eds.). In preparation.

---

## COMMISSIONED PROJECTS

### ***Subprogramme 1: Genetic diversity of global genetic resources***

#### **23. G4005.01.03 (1c): Completing genotyping of composite germplasm set of sorghum**

*January 2005–December 2006, no-cost extension to September 2008*

##### **Principal Investigator**

C Tom Hash, ICRISAT; c.hash@cgiar.org

ICRISAT-Patancheru, Hyderabad 502 324, AP, India; Tel: +91-40-3071-3322,

Fax: +91 40-3071-3074

##### **Collaborating institutions and scientists**

- ICRISAT: Hari D Upadhyaya; Rolf T. Folkertsma; Punna Ramu
- Agropolis: Claire Billot; Laetitia Gardes; Jean-François Rami; Ronan Rivallan; Monica Deu; Jacques Chantereau
- CAAS: Yu Li; Tianyu Wang; Ping Lu

**Context:** This project was designed to establish a composite germplasm set of circa 3000 accessions of wild and cultivated sorghum, determine the population structure of this using approximately 50 SSR marker loci distributed across all 10 linkage groups of the genome of cultivated *Sorghum bicolor*, and based on this information develop a reference germplasm set of sorghum for use in allele mining and linkage disequilibrium mapping. This was completed in 2007. Due to difficulties in getting the required genome coverage with publicly available SSR markers, additional markers were developed and an additional set of 40 of these have been used in 2008 to validate the population structure of the sorghum reference germplasm set.

**Findings and implications:** The population structure of the GCP sorghum reference germplasm set, which was originally determined based on allelic variation detected by 41 SSR primer pairs, was largely validated when independently assessed using a set of 40 additional EST-SSR primer pairs developed at ICRISAT from publicly available EST sequence information. This new set of EST-SSRs offers genome coverage comparable to the initial set of 41 SSRs (mostly genomic SSRs) that were used for genotyping the sorghum composite germplasm set. The new set of EST-SSRs detected a total of 362 alleles across the 384-entry sorghum reference germplasm set, with 2.1% missing data. Individual primer pairs detected 3 to 38 alleles, and had PIC values ranging from 0.13 to 0.94. Importantly, heterozygosity levels detected by these new EST-SSR primer pairs were generally low, averaging 0.038 and ranging from a low of 0.003 to a maximum of 0.106, suggesting that most detect single loci.

As in the earlier analysis completed in 2007, diversity analysis based on allelic variation across the 40 new EST-SSR markers indicates the margaritifera sub-group within the guinea race is distinct and more closely related to wild sorghums than to the other cultivated sorghums studied, suggesting once again that this group represents an independent domestication event. Landrace germplasm exhibited population substructure

based on geographic origin and this was further characterised within racial groups (five basic races and ten hybrid races). Race kafir (largely from Southern Africa) was distinct. Accessions of the durra, caudatum and guinea races each formed distinct geographic subgroups. Race bicolor was more structured than in the original analysis, with two clusters of East African origin, one of which grouped with bicolor accessions having passport data indicating a North American origin (which in turn suggests that the latter are originally from East Africa).

This additional marker genotyping has doubled the density of SSR marker genotyping of the GCP sorghum reference germplasm set, and largely validated the population structure of this set. The combined results will benefit sorghum research programmes globally, and ultimately sorghum producers and consumers around the world. The implications of this work are that the GCP sorghum reference germplasm set is sufficiently diverse that it can serve as a suitable panel for linkage disequilibrium mapping, and/or as an entry to global sorghum germplasm collections when seeking variation in any trait of interest, provided that the phenological diversity present in this germplasm set is not so great that it interferes with phenotyping of other traits of interest.

Work remaining: Submitting the final project report and drafting journal article manuscripts for publication.

#### **24. G4005.01.04 (1d): Completing genotyping of composite germplasm set of chickpea**

*January 2005–December 2005; no-cost extension to October 2007*

##### **Principal Investigator**

Hari Upadhyaya, ICRISAT, h.upadhyaya@cgiar.org

ICRISAT, Patancheru, 502324, India; Phone: 91-040-30713071

##### **Collaborating institutions and scientists**

- ICRISAT: SL Dwivedi; DA Hoisington; PM Gaur; CLL Gowda; RK Varshney; S Chandra
- ICARDA: SM Udupa; BJ Furman; M Baum

##### **Composite collection**

ICRISAT and ICARDA jointly developed a composite collection of 3000 accessions (Upadhyaya et al. 2006), which consists of 1956 accessions of the ICRISAT core collection (Upadhyaya et al. 2001), 709 accessions from ICARDA, 39 advanced lines/cultivars, 35 accession with distinct morphological variants, 20 accessions from wild *Cicer* species (*C. reticulatum* and *C. echinospermum*), and 241 trait-specific (resistance to biotic and abiotic stresses, early maturity, multi-seeded pods, double podded, large-seed, high seed protein, nodulation and responsive to high input conditions) accessions. Biologically, it represents 80% land races, 11% advanced lines/cultivars, 1% wild species, and 8% accessions of unknown biological status or geographic origin. This composite collection has been molecularly profiled using 50 SSRs in high throughput assay (ABI3700).

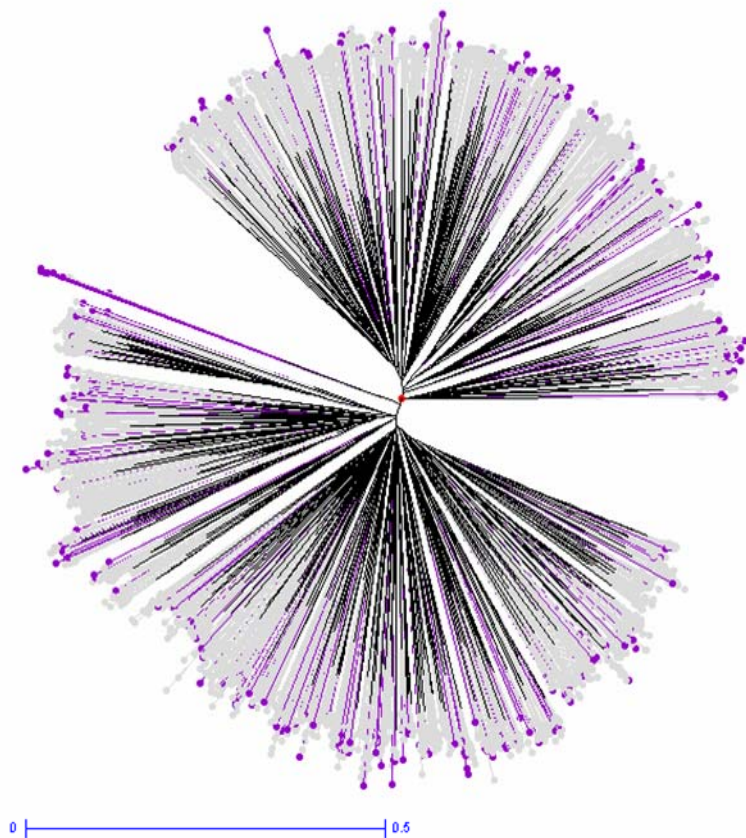
### Genetic structure of composite collection

The 48 SSR markers data on 2915 accessions were analyzed using PowerMarker V3.0 (Liu and Muse 2005) and DARwin 5.0 version (Perrier et al. 2003). This composite collection showed rich allelic diversity (1683 alleles, 35 alleles per locus, 748 most common alleles and 935 rare alleles at 1%), group-specific unique alleles, and common alleles sharing between the species and geographical groups.

Unique alleles are those detected in a group of accessions but absent in other groups. Group-specific unique alleles were 104 in Kabuli, 297 in desi, and 69 in wild *Cicer*; 114 each in Mediterranean and West Asia (WA), 117 in South and South East Asia (SSEA), and 10 in African region accessions. Desi and kabuli shared 436 alleles while wild *Cicer* with desi and kabuli, respectively, shared 17 and 16 alleles. The accessions from SSEA and WA shared 74 alleles while those from Mediterranean with WA and SSEA, respectively, 38 and 33 alleles. Desi's detected higher proportion of rare alleles (53%) than kabuli's (46%), while wild *Cicer* accessions were devoid of rare alleles.

### Reference set

A reference set consisting of 300 genetically most diverse accessions have been formed. This reference set captured 1315 (78%) of the 1683 composite collection alleles, representing diversity from the entire spectrum of composite collection (Figure). The usefulness of this reference set in genomics and breeding of chickpea needs to be investigated.



**Figure. Un-weighted neighbour-joining tree based on the simple matching dissimilarity matrix of 48 SSR markers across the 2915 accessions of chickpea composite collection (Grey colour) with proposed reference set (300 accessions) in purple colour**

## References

- Liu K, Muse SV (2005). Power marker: Integrated analysis environment for genetic marker data. *Bioinformatics* 21:2128-2129.
- Perrier X, Flori A, Bonnot F (2003). Data analysis methods. In P Hamon, M Seguin, X.Perrier, and JC Glaszmann (eds), *Genetic diversity of cultivated tropical plants*. Enfield, Science Publishers. Montpellier. pp 43-76.
- Upadhyaya HD, Bramel PJ, Singh S (2001). Development of a chickpea core subset using geographical distribution and quantitative traits. *Crop Sci* 41:206-210.
- Upadhyaya HD, Furman BJ, Dwivedi SL, Udupa SM, Gowda CLL, Baum M, Crouch JH, Buhariwalla HK, Singh S (2006). Development of a composite collection for mining germplasm possessing allelic variation for beneficial traits in chickpea. *Plant Genet Resour* 4:13-19.

## 25. G4005.03.02 (3b): Molecular characterisation of pigeonpea (*Cajanus cajan*) composite collection

*January 2005–December 2006; no-cost extension to December 2007*

### Principal Investigator

Hari Upadhyaya, ICRISAT, h.upadhyaya@cgiar.org  
International Crops Research Institute for the Semi-Arid Tropics (ICRISAT),  
Patancheru, 502324, India; Phone: 91-040-30713071 (extension 2333)

### Collaborating Institution and Scientist

ICRISAT: R Bhattacharjee; DA Hoisington; S Chandra; RK Varshney; KB Saxena

### Composite Collection

ICRISAT developed a composite collection of 1000 accessions, which consists of 146 mini core subset accessions (Upadhyaya et al. 2006), another 146 comparator mini core accessions, 236 of the 1290 core collection accessions (Reddy et al. 2005), 389 trait-based accessions (resistance to biotic and abiotic stresses and those with superior morpho-agronomic traits), 20 advanced lines/cultivars, and 63 accessions from seven wild *Cajanus* species. Geographically, it consists of 73% accessions from Asia, 13% from Africa, 5% from Caribbean, 3% each from America's and Oceania, and remaining 3% from other regions. This composite collection has been molecularly profiled using 20 SSRs in high throughput assay (ABI3700).

### Genetic structure of composite collection

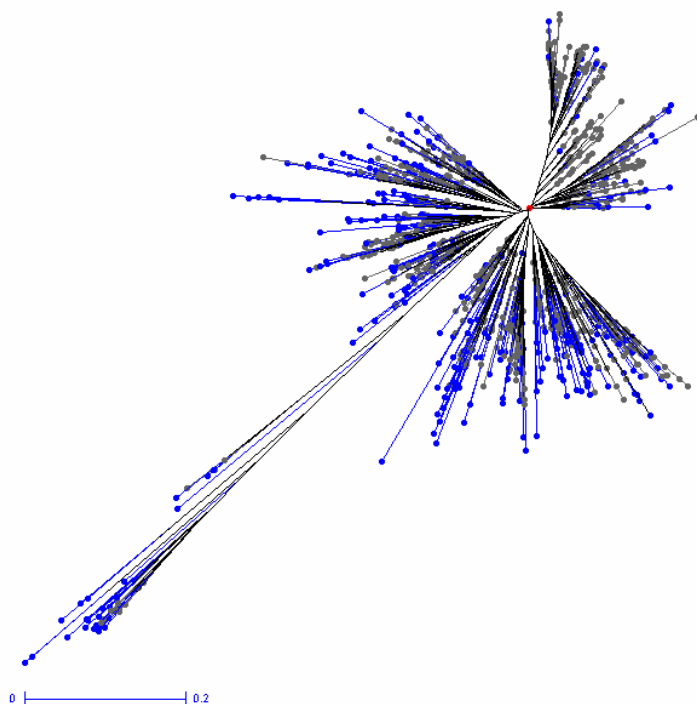
The 20 SSR markers data on 952 accessions were analyzed using PowerMarker V3.0 (Liu and Muse 2005) and DARwin 5.0 version (Perrier et al. 2003). This composite collection showed rich allelic diversity (197 alleles, 10 alleles per locus, 82 most common alleles and 115 rare alleles at 1%), group-specific unique alleles, and common alleles sharing between the species and geographical groups.

Unique alleles are those detected in a group of accessions but absent in other groups. Group-specific unique alleles were 60 in wild types and 64 in cultivated types. Accessions from Asia had 48 unique alleles while those from Africa had only two unique alleles. Non-determinate type (NDT) cultivated pigeonpea accessions were represented by 37 unique alleles while determinate types (DT) only one allele. Wild and cultivated

types shared 73 alleles, DT and NDT 10 alleles, DT and wild types 4 alleles, and NDT and wild types 20 alleles. A tree diagram using DARwin 5.0 revealed wild types as a group genetically more diverse than cultivated types, while NDT were more diverse than DT.

### Reference set

A reference set consisting of 300 genetically most diverse accessions have been formed. This reference set captured 187 (95%) of the 197 composite collection alleles, representing diversity from the entire spectrum of composite collection (Figure). The usefulness of this reference set in genomics and breeding of pigeonpea needs to be investigated.



**Figure. Un-weighted neighbour-joining tree based on the simple matching dissimilarity matrix of 20 SSR markers across the 952 accessions of pigeonpea composite collection (Grey colour) with proposed reference set (300 accessions) in blue colour**

### References

- Liu K, Muse SV (2005). Power marker: Integrated analysis environment for genetic marker data. *Bioinformatics* 21:2128-2129.
- Perrier X, Flori A, Bonnot F (2003). Data analysis methods. In P Hamon, M Seguin, X.Perrier, and JC Glaszmann (eds), *Genetic diversity of cultivated tropical plants*. Enfield, Science Publishers. Montpellier. pp 43-76.
- Reddy LJ, Upadhyaya HD, Gowda CLL, Singh S (2005). Development of core collection in pigeonpea [(*Cajanus cajan* (L.) Millspaugh)] using geographic and qualitative morphological descriptors. *Genet Resour Crop Evol* 52:1049-1056.
- Upadhyaya HD, Reddy LJ, Gowda CLL, Reddy KN, Singh S (2006). Development of a mini core subset for enhanced and diversified utilisation of pigeonpea germplasm resources. *Crop Sci* 46:2127-2132.

**26. G4005.03.06 (3f): Molecular characterisation of groundnut (*Arachis hypogaea* L.) composite collection**

*January 2005–December 2005; no-cost extension to October 2007*

**Principal Investigator**

Hari Upadhyaya, ICRISAT, h.upadhyaya@cgiar.org  
International Crops Research Institute for the Semi-Arid Tropics  
(ICRISAT), Patancheru, 502324, India, Phone: 91-040-30713071 (extension 2333)

**Collaborating institutions and scientists**

- ICRISAT: R Bhattacharjee; DA Hoisington; S Chandra; RK Varshney
- EMBRAPA: JFM Valls; MC Moretzsohn; S. Leal-Bertioli; P Guimaraes
- UCB: D Bertioli

**Composite collection**

ICRISAT and EMBRAPA jointly developed a global composite collection, consisting of 1000 diverse groundnut accessions, which included 184 groundnut mini core subset (Upadhyaya et al. 2002), another 184 mini core comparator, 110 accessions from Asia core and mini core, 408 elite germplasm/cultivars and trait-specific (resistance to biotic and abiotic stresses, early maturity and/or fresh seed dormancy, large-seed, high shelling percentage, high oil and/or protein content, and interspecific derivatives) accessions, and 114 wild *Arachis* accessions. This composite collection has been molecularly profiled using 21 SSRs in high throughput assay (ABI3700).

**Genetic structure of composite collection**

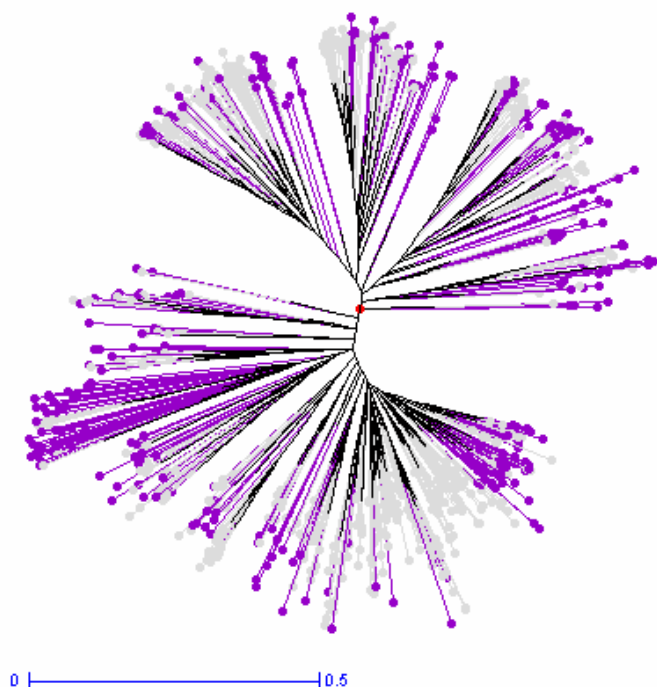
Twenty-one SSR markers data on 852 accessions were analyzed using PowerMarker V3.0 (Liu and Muse 2005) and DARwin 5.0 version (Perrier et al. 2003). This composite collection showed rich allelic diversity (490 alleles, 23 alleles per locus, 246 common alleles and 244 rare alleles at 1%), group-specific unique alleles, and common alleles sharing between subspecies and geographical regions.

Unique alleles are those detected in a group of accessions but absent in other groups. Group-specific unique alleles were 101 in wild *Arachis*, 50 in subsp. *fastigiata*, and only 11 in subsp. *hypogaea*. Accessions from America's revealed highest number of unique alleles (109) while Africa and Asia, respectively, had only six and nine unique alleles. The two subsp. *hypogaea* and *fastigiata* shared 70 alleles. The wild *Arachis* in contrast shared only 15 alleles with *hypogaea* and 32 alleles with *fastigiata*. A tree-diagram using DARwin 5.0 separated majority of the *hypogaea* from *fastigiata* accessions while wild *Arachis* accessions clustered with *hypogaea*.

**Reference set**

A reference set consisting of 300 genetically most diverse accessions have been formed. This reference set captured 466 (95%) of the 490 composite collection alleles, representing diversity from the entire spectrum of composite collection (Figure). The usefulness of this reference set in genomics and breeding of groundnut needs to be investigated.





**Figure.** Un-weighted neighbour-joining tree based on the simple matching dissimilarity matrix of 21 SSR markers across the 852 accessions of groundnut composite collection (Grey colour) with proposed reference set (300 accessions) in purple colour

## References

- Liu K, Muse SV (2005). Power marker: Integrated analysis environment for genetic marker data. *Bioinformatics* 21:2128-2129.
- Perrier X, Flori A, Bonnot F (2003). Data analysis methods. In P Hamon, M Seguin, X.Perrier, and JC Glaszmann (eds), *Genetic diversity of cultivated tropical plants*. Enfield, Science Publishers. Montpellier. pp 43-76.
- Upadhyaya HD, Cox PJB, Ortiz R, Singh S (2002). Developing a mini core of peanut for utilisation of genetic resources. *Crop Sci* 42:2150-2156.

## 27. G4005.05 Assessing Eco-tilling as a methodology for targeted genotyping and SNP discovery

January 2005– May 2007; no cost extension – May 2008

### Principle Investigator

Ken L McNally, TT Chang Genetic Resources Centre, IRRI, k.mcnally@cgiar.org  
IRRI, DAPO Box 7777, Metro Manila, Philippines; Telephone: +63 (2) 580-5621; Fax: +63 (2)580-5699

### Collaborating institutions and scientists

- Agropolis–CIRAD: C Billot, I Hippolyte, F-C Baurens, J-F Rami
- URGV, France: A Bendahmane

## 1. Research activities and developments at IRRI

### 1.1 EcoTILLING of selected *Oryza* germplasm for drought candidate genes

EcoTILLING, a tool that detects polymorphism in the form of SNPs or indels in natural populations, was employed to survey variation in a panel of *Oryza* germplasm consists of a mini-core collection of 1536 *O. sativa* accessions, 190 accessions of *O. glaberrima*, and

95 accessions representing six wild AA genome *Oryza* species. Each *O. sativa* or wild AA genome sample DNA was pooled against an indica (IR 64) or a japonica (Nipponbare) contrast while *O. glaberrima* samples were contrasted against Nipponbare or IRGC 96717 (CG-14). Primers based on the Nipponbare sequence for the drought candidate genes TPP, ERF3, DREB2, ADF2a, ADF2b, MAPk2, BZIP, SUC, and 14-3-3 worked in both cultivated and wild AA genome *Oryza* species producing a single amplicon of about 1 kb. Agarose-based ecotilling was developed as a simplified tool for visualizing digestion products generated from the endonuclease action of the CEL 1 enzyme. Haplotypes were generated based on the patterns from the indica or japonica contrasts and verified by confirmatory sequencing of representative accessions. The *O. sativa* samples showed an average of 10 haplotypes per gene, largely supported by sequence data with the number of SNPs ranging from 6 (ADF2a) to 17 (TPP) for areas covering the coding region. SNPs were mostly transitions (68%) or transversions (28%) while indels were observed only in the SUC and TPP loci. In *O. glaberrima*, putative haplotypes varied from 2 (DREB2 and TPP) to 7 (14-3-3). Confirmatory sequencing verified 4 SNPs for BZIP, 7 for ERF3, and 12 for MAPk2. Among the wild AA genome *Oryza* species, species-specific SNPs were identified, proving the utility of EcoTILLING in taxonomic authentication. Confirmatory sequencing showed an average of 30 SNPs for the BZIP, MAPk2, and ERF3 genes across the six AA genome species.

## 1.2 Association between molecular variation in drought responsive candidate genes for drought tolerance and performance under water deficit in the field

*O. sativa* accessions were screened over three consecutive dry seasons for performance under drought stress applied during vegetative growth. Large phenotypic variation was observed within each variety group. Our results indicate that several polymorphisms of key drought responsive genes are associated with differences in response to water deficits. Continuous phenotypic variation was observed within many haplotype groups for each candidate gene; however, several haplotypes were associated with lower tolerance to vegetative drought stress.

**Table 1. Haplotypes obtained in the mini-core collection of *O. sativa* for drought candidate genes.**

Gene	No. of putative haplotypes	No. of SNPs detected	Transition (%)	Transversion (%)	Indels (%)
ADF2a	10	6	70	30	0
ADF2b	7	8	75	25	0
DREB2	10	12	83	17	
ERF3(cds)/ ERF3upstream	11	7 (cds only)	57	43	0
MAPk2	9	11	82	18	0
BZIP	9	7	57	43	0
SUC	8	11	64	18	18
TPP	12	17	53	29	18

### **Deviations from workplan**

Sequence analysis of representative mismatches (putative SNPs) at candidate genes in rice required longer than expected to convert them to SNP identities. These results are now being analyzed, and these results compared to haplotype mismatch patterns. Further candidate genes for reproductive-stage stress tolerance are being screened.

### **Next steps**

This project validated the potential to detect SNPs in key candidate genes. Another output, not funded by this project, was the ability to link these polymorphisms to trait variability. A new commissioned project (G4008.5) is building on the vegetative-stage phenotyping for drought tolerance. Accessions that were used for EcoTILLING will be phenotyped for drought response at reproductive stage, the most sensitive period to stress. Two additional publications are in preparation – the association of SNPs in drought candidate genes with vegetative-stage phenotypes and the utility of EcoTILLING as a biosystematics tool for the AA genome species of rice.

### **References**

Raghavan C, Naredo MEB, Wang H, Atienza G, Liu B, Qiu F, McNally KL and Leung H (2007). Rapid method for detecting SNPs on agarose gels and its application in candidate gene mapping. *Mol Breeding*. 19:87-101.

## **28. G4005.06: Supporting emergence or reference drought tolerance phenotyping centers - Drought phenotyping network** *April 2005–December 2007; no-cost extension to June 2008*

### **Principal Investigator**

Reinaldo L Gomide, EMBRAPA Maize and Sorghum; [gomide@cnpms.embrapa.br](mailto:gomide@cnpms.embrapa.br), Rodovia MG 424 km 45, Sete Lagoas, MG 35701-970, Brazil; 55 (31) 3779-1228; Fax: 55 (31) 3779-1088

### **Collaborating institutions and scientists**

- EMBRAPA Rice and Beans: Cleber M Guimarães and others
- EMBRAPA Mid-North.: Edson A Bastos and Others
- EMBRAPA Wheat and Savannah: Walter Q. Ribeiro Júnior and Others
- EMBRAPA Semi-Arid: Luiz B. Morgado and others

### **Research activities and progresses**

The project developed and made useful phenotypic evaluation protocols for cereals (maize, sorghum, rice, and wheat) and legume crops (common bean and cowpea), and established phenotyping site specific experimental (SSE) areas of excellence (2) and reference (5) for drought tolerance (DT) studies according to specific climatic condition, soil physical and chemical properties, with laboratories, controlled environment target fields and greenhouses, training unit for researchers and assistants, with facilities and well defined dry season periods to assure total irrigation and soil moisture control during the drought phenotyping field trials. Overall, the project established a scientific and service network for drought tolerance phenotyping in Brazil.

**Irrigation Water Application, Control and Management:** The irrigation systems installed in the SSE areas are: conventional sprinkler, localised (drip), and continuously moving straight lateral or linear-move systems. These irrigation systems were tested and evaluated for water distribution uniformity and applied water depths by means of measuring and controlling water pressure, flow rate, radius of throw, and emitters or sprinklers spacing. The water depths applied in the irrigations were measured in collectors or catch cans in each genotype field plot. These collectors were placed transversally to the crop rows following a rectangular grid or a transect layout in the plots. The uniformity of the water distribution in the irrigated plot was set to be equal or greater than 95 % (Christiansen). The irrigation water application rate was set to be lower than basic soil saturated water infiltration rate in order to avoid surface runoff, which was not allowed. **Climatic Condition** was characterised and hydrological water balance (Thornthwaite & Mather) was determined with 15 to 50 years data series, obtained from standard weather stations. A standard procedure was established to calibrate and install the equipments and sensors of automatic weather stations in each SSE, configured to register automatically the main microclimatic surface parameters locally, with intervals of half to one hour. **Irrigation water management** was carried out by means of reference evapotranspiration (ET<sub>o</sub>) and crop evapotranspiration (ET<sub>c</sub>) computation, using both class A pan and modified Penman-Monteith equation methods, with the crop (k<sub>c</sub>) and pan (k<sub>p</sub>) coefficients. The ET<sub>c</sub> was determined by multiplying ET<sub>o</sub> for each genotype crop coefficient (K<sub>c</sub>). Irrigation management strategy and irrigation timing criteria were performed based on spread sheet (Excell) for ET<sub>o</sub> and ET<sub>c</sub> computation and soil water balance within the root system depth determination, associated with the measurements of soil water content in different layers. The irrigation was uniform after sowing, germinating, and stand formation with 100% replacement of the ET<sub>c</sub> and soil water availability (SWA) - non water stressed condition. Afterwards, the water stress treatments were obtained with different replacement level of the ET<sub>c</sub>, generating different application of water depths in the plots, and consequently different SWA, at pre-defined crop growth phases, defined for each genotype, according to breeder, physiologist, and irrigation engineers' indication in order to induce the water stress intensity. **Soil water content**, in different soil layers, was monitored by gravimetric method and other equipments and sensors (gypsum block, Diviner, tensiometer, neutron probe).

### **Tangible outputs delivered**

**Phenotyping SSE areas:** 2 Embrapa's Centres of Excellence (Sete Lagoas-MG & Santo A. de Goiás-GO) and 5 Embrapa's Sites of Reference (Janaúba-MG, Porangatu-GO, Teresina-PI, Planaltina-DF, Petrolina-PE) for DT studies. **Soil physical and chemical properties** characterised for each environment site. **Irrigation system schemes** (conventional sprinkler, linear moving sprayer lateral line, drip) installed and evaluated for each environment site with water flow rate and management monitoring devices (hydrometer, collectors kit, pressure meters). **One automatic weather station installed** in each environment site with microclimatic data registered. **Soil-water content & availability** controlled, measured & registered in each environment site in the soil root system profile. **Irrigation water application and soil-plant water stress** controlled and monitored for DT phenotyping in each cereal and vegetable genotype (at predefined crop growth stage) / environment site. **Plant-water content** measured and registered for some

genotypes in some environment site. **Genotypes tolerant and sensible**: a reasonable number of genotypes (access and selected materials for maize, sorghum, rice, wheat, common bean, cowpea) were phenotyped and the main contrasting genotypes (tolerant and sensible) to drought were identified and selected. **Database**: all the project information and data were transferred into database (Morpho) for each environment target site.

### Future research

Next steps and/or challenges: Mechanisms investigation for drought tolerance for each crop specie studied.

### References

Project Team (2008). Supporting Emergence or Reference Drought Tolerance Phenotyping Centers, Drought Phenotyping Network- DPN. Proceedings of the final project workshop, 17-18 June 2008, Gomide, R. L. ed., Embrapa Maize and Sorghum, Sete Lagoas, MG, Brazil, pp. 171.

## 29. G4005.07: Whole plant physiology modelling (WPM)

*May 2005–May 2008*

### Principal Investigator

Luquet Delphine, Agropolis–CIRAD, luquet@cirad.fr  
Av. Agropolis Lavalette 34398 Montpellier cedex 5, France, 00 33 4 67 61 55 71

### Collaborating institutions and scientists

- CSIRO: Scott Chapman
- University of Queensland, Australia: Graeme Hammer, Karine Chenu
- EMBRAPA: Camilla Andrade, Reinaldo Gomide, Cleber Guimaraes, Edson Bastos
- Agropolis–INRA: Claude Welcker, François Tardieu
- Agropolis–CIRAD: Delphine Luquet, Nourollah Ahmadi, Michael Dingkuhn

1. Context: WPM was initiated as a complementary project to Drought Phenotyping Network (DPN, Embrapa). Its concepts date back to 2004 Phenotyping meeting (Montpellier) aiming at improving phenotyping methods and capacities in GCP. While DPN was designed to physically develop a field drought phenotyping network across Brazil, WPM was conceived as a series of relevant case studies, with the objective to apply, prove or improve the potential of plant/crop modelling for: **Component 1 (C1)**: assisting in characterizing Target Population of Environments (TPE) met by a breeding programme; **(C2)**: assisting field phenotypic analysis, by extracting, from complex observed traits, elementary traits (model parameters), assumed to be less polygenic and less E dependent and thus more adapted to genetic studies; **(C3)**: assisting ideotype behavior analysis (in silico trait combination) or trait impact on plant performance in TPE.

### 2. Findings and implications

**C1**: The main C1 study was successfully conducted on upland rice and maize TPEs for the Brazilian Cerrados using SARRAH model. The main result (Heinemann et al. 2007) was that current breeders' screening sites are not representative of the TPE regarding drought intensity and probability, but that drought is generally not a severe constraint in

these TPEs. Results varied, however, between rice and maize, between short and medium duration rice, between maize as main or 2<sup>nd</sup> crop, and between deep and shallow soil situations. *Partners: Embrapa, Cirad, CSIRO.*

An additional study was conducted on West African sorghum cvs having different height and photosensitivity, based on drought prone environments on a N-S gradient (Mali). Results indicated that modern, early, photo-insensitive genotypes are only suited to the northern, dry environments whereas for long rainy season zones, traditional, photoperiod sensitive materials are advantageous (Kouressy et al. 2007). *Partners: Cirad, IER.*

**C2 and 3** used similar models simulating and combining specific genotypic adaptations. C2 was to apply models to multi-site field trials, while C3 explored model parameter relation with genetic information and virtual genotype behavior. Because of experimental issues in Brazil (DPN), research was also built on additional case studies.

**a- The morphogenetic model *Ecomeristem*** was developed and adapted to the study of dry-down drought experiments, with the objective of simulating phenotypic plasticity of genotypes in response to environment, and to phenotype genotypes using heuristics (optimisation of model parameters to be used as trait information, Luquet et al. accepted). The methodology is operational but remains to be further tested on large populations (genome wide association studies are now planned). Results obtained on two preliminary case studies (1) on a collection of 200 sativa rices under field optimal conditions at IRRI and (2) on a rice mapping population under P deficiency conditions, were quite promising. *Main partners: Cirad and IRRI, as closed partner of WPM.*

**b- An existing physiological model of leaf extension rates (LER) of maize** in response to soil and atmospheric drought parameters, previously confirmed to provide valuable QTLs, was further tested on maize silk growth. Results indicated common QTLs for model parameters for leaf and silk growth response to drought. The model was then implemented as a component of the APSIM maize model to measure QTL impacts on grain yield. Results showed that environmental and genetic control of leaf growth responses did flow through to affect grain yield in specific types of drought and provided a structure to test these concepts in a range of drought environments using known QTL effects as inputs (Chenu et al. 2008). *Partners: INRA/Lepse, CSIRO/UQ.*

**c- A new model of sorghum photoperiodism (*Impatience*)** was developed, validated, and implemented in SARRAH crop model. A field phenotyping methodology was developed to heuristically quantify genotypic parameters of photoperiod response (Dingkuhn et al. 2008). *Main partners: Cirad, IER.*

### 3. Next steps and/or challenges

WPM final meeting was held at Pioneer (Iowa, US, Feb.08); resulting in a set of recommendations. At short term, a few concrete breeding (genetic) activities should be targeted and associated with a complementary team of modelers: models should be applied for TPE, phenotyping for genetic trait analysis and ideotype exploration (2008 SP1 commissioned project: *GenePhene*, PI: J. Cairns (IRRI), involving CIRAD on the modeling side). To facilitate the integration of modelling approaches within breeding process, training should be more frequently organised (cf. two 2006 training activities

realised by or in collaboration with WPM: modelling in Brazil; phenotyping in Montpellier). Target groups should be breeders, physiologists, agronomists, geneticists participating as a team in a shared breeding process). An investment should be made by GCP in particular in cross-CP collaboration with the upcoming CCCP on TPE characterisation and environmental data set sharing.

## References

- Dingkuhn M, Kouressy M, Vaksman M, Clerge B, Chantreau (2008). Applying to sorghum photoperiodism the concept of threshold-lowering during prolonged appetite. *European Journal of Agronomy* 28, 74-89.
- Kouressy M, Dingkuhn M, Vaksman M, Heinemann AB (2008). Adaptation to diverse semi-arid environments of sorghum genotypes having different plant type and sensitivity to photoperiod. *Agricultural and Forest Meteorology* 148, 357-371.
- Chenu K, Chapman SC, Hammer GL, McLean G, Salah HBH, Tardieu F (2008) Short-term responses of leaf growth rate to water deficit scale up to whole-plant and crop levels: an integrated modelling approach in maize. *Plant, Cell Environ* 31:378-391
- Heinemann AB, Dingkuhn M, Luquet D, Chapman S (2008) Characterisation of drought stress environments for upland rice and maize in central Brazil. *Euphytica* - on line.
- Luquet D, Clément-Vidal A, This D, Fabre D, Sonderegger N, Dingkuhn M (2008) Orchestration of transpiration, growth and carbohydrate dynamics in rice during a dry-down cycle. *Functional Plant Biology* accepted with revision to *Func. Plant Biol*

## 30. G4006.01: Developing strategies for allele mining within large collections

### Principal Investigator

NR Sackville Hamilton, IRRI, [r.hamilton@cgiar.org](mailto:r.hamilton@cgiar.org)

ABSTRACT NOT SUBMITTED

## 31. G4006.02: A dataset on allele diversity at orthologous candidate genes in GCP crops (ADOC)

*January 2006–December 2007; no-cost extension to December 2008*

### Principal Investigator

Dominique This, Agropolis (Montpellier SupAgro–CIRAD),  
[dominique.this@supagro.inra.fr](mailto:dominique.this@supagro.inra.fr)

UMR DAP 1098, CIRAD TA A-96/03, Av. Agropolis, 34398 Montpellier cedex 5, France; (33) 4-67615829; Fax: (33) 4-67615605

### Collaborating institutions and scientists

- Agropolis–CIRAD: Brigitte Courtois, Claire Billot, Jean François Rami, Romain Philippe, Pierre Mournet
- CIP: Merideth Bonierbale, Roland Schaftleitner, Reinhart Simon, Percy Rojas
- ICRISAT: Rajeev Varshney, Tom Hash, Dave Hoisington, Spurthi Nayak, Hari Upadhyaya
- Agropolis–INRA/CNG: Dominique Brunel, Redouane El Malki
- IRRI: Ken McNally

- ICARDA: Michael Baum, Wafaa Choumane
- CIAT: Matthew Blair, Martin Fregene

### Context

The ADOC project aims to characterise allelic diversity at orthologous loci of candidate genes for drought tolerance in seven GCP crops (rice, barley, sorghum, bean, chickpea, cassava and potato), working on reference collections of around 300 accessions for each crop. Six gene families (ERECTA, DREB, SS, SPS, ASR and VIN) were selected as the initial subset of target genes. Except the DREB gene family, for which a specific focus has been given to DREB2A, and SPS gene family in cereals, for which only Os01g69030 orthology group will be studied, they represent a set of relatively small gene families acting at different levels of the drought stress response (transcriptional regulation, carbohydrate metabolism...) for which a comparative analysis of gene families was decided.

### Research activities and progresses by crop and gene family

Specific primers were designed and tested for each crop and each member of gene families, allowing theoretically the amplification and sequencing of either the whole sequence of the gene (ASR and DREB2A, due to their small size) or a representative segment of around 1000 to 3000 bp for larger genes.

From 7 (bean) to 20 (rice) different genes were targeted per crop, and from 6 (DREB) to 24 (SuSy) genes across species in each gene family (see table 1). Genomic fragments are currently sequenced in the reference collections and sequence alignments released to the ADOC partners through an ftp site. Completion of this work, or at least a representative subset of it, is expected within four months. Final release of SNP data and sequence alignments will be done through the GCP repository at the end of the project.

Ortholabs are analyzing sequence data by crop cluster and generate information on candidate gene diversity in relation with crop partners (for the structuration of reference collections and candidate gene mapping when available) and gene specialists (for establishing orthology relationships and functional inferences).

**Table 1. Step point on July 2<sup>nd</sup> 2008 of the sequencing work for the ADOC project ('5/6' means that 5 out of 6 genes targeted in this family have been at least partially sequenced)**

	ASR	VIN	ERECTA	DREB2A	SuSy	SPS	total/crop
rice	5/6	0/2	3/3 (2 fgts)	1/1 (1 fgt)	7/7 (1 to 3 fgts)	0/1	16/20
barley	3/3	0/3	2/2 (2fgts)	1/1 (1fgt)	2/2 (2 fgts)	0/1	8/12
sorghum	5/6	0/2	3/3 (1 to 2 fgts)	0/1	5/5 (2 fgts)	0/1	13/18
cassava	2/2	0/0	1/1 (3 fgts)	0/0	4/4 (2 fgts)	1/1 (1 fgt)	8/8
potato	0/3	0/0	1/2 (1 frgt)	0/0	0/2	0/5	1/12
chickpea	1/1	0/0	1/1 (2 fgts)	1/1 (2 parts)	0/2	1/2	4/7
bean	0/2	0/0	1/1(1 fgt)	2/2 (2A and 2	1/2 (1 fgt)	0/0	4/7
<b>total/gene</b>	<b>16/23</b>	<b>0/7</b>	<b>12/13</b>	<b>5/6</b>	<b>19/24</b>	<b>2/11</b>	<b>54/84</b>

### Findings and implications

In cereals, population structure influence partially haplotype patterns. Different patterns and intensity of sequence diversity have been found, within gene families, but generally orthologous genes presented similar diversity patterns. For some genes, computation of a



sequence-based neutrality test suggests selection events acting at the species and/or subgroup level.

In legumes, the initial survey of sequence diversity indicated presence of haplotypes based on population structure. The SNP diversity in DREB genes for chickpea was very low, whereas in common bean it was higher. Among all the genes, ERECTA showed highest SNP and haplotype diversity in legumes.

In potato, preliminary sequence analysis of an ERECTA gene in around 80 predominantly diploid accessions revealed good sequence quality and enabled to identify various SNPs. However, sequence quality in tetraploid accessions was strongly compromised due to indels in one or more of the 4 alleles, which caused that sequences after an indel resulted to be in-readable. In cassava, sequence analysis of ASR1 and ERECTA is accomplished, but heterozygosity induces difficulties in analyzing data.

Getting hold of allelic differences in candidate genes of our target crop, will allow researchers and breeders to investigate correlations between phenotypic data and the SNP data obtained by this project. This association genetics analysis approach will link drought tolerance phenotypes to candidate gene genotypes resulting in knowledge on alleles providing increased drought tolerance. Furthermore, the SNP database produced by this project will serve as resource for designing markers for favorable candidate gene alleles, and use relevant SNPs in selection programmes.

### **Next steps and challenges**

After completion of individual gene diversity analysis for each crop, comparison between different species will allow a critical analysis of “candidate gene” status for orthologous sequences, and influence of gene duplications and speciation in the evolution and eventual subfunctionalisation of genes under study. Eco-geographical information captured on the different populations will be used when available to analyze potential effect of agronomical constraints on gene evolution during domestication and breeding history.

## **32. G4006.03: SNP analysis and the genetic diversity along the rice genome (*HaplOryza*)**

*January 2006–December 2006; no-cost extension to October 2007*

### **Principal Investigator**

Kenneth McNally, IRRI; k.mcnally@cgiar.org

International Rice Research Institute, DAPO 7777, Metro Manila, Philippines

### **Collaborating institutions and scientists:**

- Agropolis–CIRAD: Claire Billot, Brigitte Courtois, Gaetan Droc
- Agropolis–IRD: Gérard Second
- CNG: Dominique Brunel, Mark Lathrop

Asian cultivated rice occurs as two major types, *indica* and *japonica* that appear to have arisen from independent domestication events. Even though rice is a predominantly self-pollinated crop, both types can frequently be found within the same region allowing the

prospect for genetic exchanges between them. Since whole genome sequences are available for each type, we have the opportunity to identify single nucleotide polymorphism (SNP) suitable for determining the extent of linkage disequilibrium and haplotype structure that is indicative of their differentiation. In this project, the high quality japonica Nipponbare sequence (IRGSP) was compared to the whole genome shotgun indica 93-11 (Beijing Genomics Institute) sequence to identify a set of 1536 SNPs suitable for undertaking genome scans. These SNPs were genotyped across 900 types predicted to cover the range of *indica/japonica* diversification and prospective natural hybrids between them.

A bioinformatics pipeline consolidated over 2.6 million SNPs *indica/japonica* from previous studies (Shen et al., Feltus et al., and Zang). These SNPs were filtered to limit them to unique polymorphisms with one SNP per 50 bp window, and 100% identical sequence for 20 bp on either side of the polymorphism. This filtering resulted in a set of 277,760 validated SNPs. The location, polymorphism, right and left flanking sequences were determined for the IRGSP Nipponbare sequence and loaded into a MySQL database indexed by pseudomolecule position. Illumina evaluated these SNPs on the basis of the definition and characteristic of primers to be used for assays and assigned each SNP a score of 0 to 1. The scores were obtained in September 2006 resulting in 261,562 SNPs with a score of >0.4 and the majority of scores >0.8.

The Illumina chip contained 1536 SNP assays. Two strategies have been taken into account for achieving estimates of LD, choosing SNPs for whole genome LD scans and for looking closer at specific regions. For the former, the level is 1 SNP for every 320 kb of the genome (covering 357 Mb). In the case of specific regions, the level is 1 SNP per 50 kb. The specific regions include:

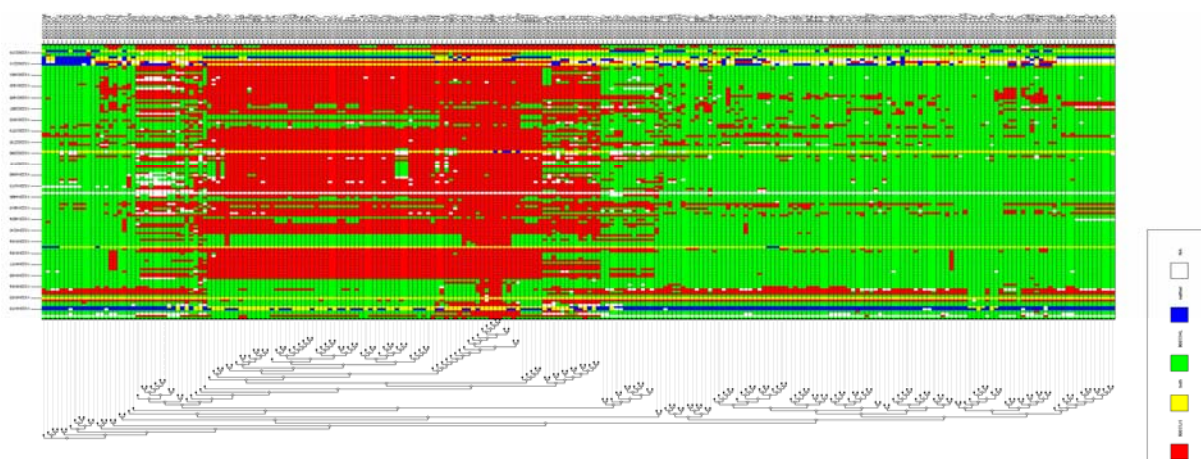
- a) 8 carrying candidate genes for drought of about 700 kb each centered on DREB2a on Chr. 1, OsCIN1, aquaporin, trehalose phosphatase, actin depolymerizing factor, and Erecta2 on Chr. 2, OsCIN2 on Chr. 4, auxin efflux carrier on Chr. 9, and ASR1 on Chr. 11 (totaling about 5.9 Mb);
- b) the short arm of Chr. 6 that includes the Wx, S6, hd3, WC, hd1, Erecta1, and Erecta-like genes (15.4 Mb);
- c) two regions of low SNP density on Chr. 7 and 2 regions of normal SNP density on Chr. 12 (about 4 Mb).

For the whole genome, 4 sets of SNPs having Illumina scores > 0.9 were chosen with 1 SNP/80 kb in each set. Correspondingly, 2 sets of SNPs for each specific region were chosen with 1 SNP/25 kb window and one set retaining a window of 50 kb. The prospective SNP sets include both synonymous and non-synonymous mutations. A unique set was chosen by Illumina, and developed at the beginning of 2007. The genotyping platform consisted of single base extension SNP assays implemented on the Illumina BeadArray platform at the National Genotyping Center at Evry, France.

The set of germplasm accessions to include in the study was finalised from the GCP composite collection (492 accessions) and from accessions at CIRAD that are likely to

contain introgressions between the indica and japonica groups as well as wild species (408 accessions). The materials from the GCP composite collection are a set of 273 representative accessions chosen to cover the different variety groups and 219 accessions deemed to be admixtures between *indica* and *japonica* (based on Structure analysis at  $K=2$  for 45 SSRs on 2339 sativa in the GCP composite collection where the contribution of alleles from one or the other group was  $< 80\%$ ). The materials from CIRAD include 16 species (281 accessions of *O. sativa*) from 31 countries where different introgressions are suspected. Quality and quantity prior to use on the Illumina chip was determined and ended into two sets of accessions tested, the last raw results were obtained in June 2008.

The first analyses enable to discriminate perfectly between *japonica* and *indica* accessions, as well as decipher the genetic structure of wild species in comparison to *indica* and *japonica*. Linkage disequilibrium and haplotype structure was also established and graphical representation was developed (see Figure 1 for an example).



**Figure 1: Graphical representation of haplotypes along a chromosome with reference to two accessions.**

### **33. G4006.05: Development of a composite collection and the genotyping of faba bean**

*January 2006–June 2008; no-cost extension to July 2008*

#### **Principal Investigator**

Bonnie Furman, Michael Baum, ICARDA; [m.baum@cgiar.org](mailto:m.baum@cgiar.org)  
 ICARDA, P.O.Box 5466 Aleppo, Syrian Arab Republic, +96 3 2122 2551

#### **Collaborating institutions and scientists**

- Agropolis–INRA/UR LEG: G. Duc
- Instituto de Agricultura Sostenible, Spain: MJ Suso

A global composite collection of 1000 accessions of faba bean was developed at ICARDA in collaboration with INRA-Dijon and IAS-Cordoba. The composite collection contains 505 accessions from the ICARDA global collection and 250 accessions from

IAS and 245 accessions from INRA. The ICARDA, INRA and IAS accessions have been grown and DNA extracted. Seeds from each plant were collected and stored at 10 °C.

Because of the availability of only few microsatellite primers for faba bean, 50 new microsatellite primer pairs were developed, and 33 functional microsatellite markers have been identified. The primers were labeled with three different colors FAM, NED and VIC (Primer details were submitted to GCP in the report of May 2007).

Unfortunately, not all the primers showed clear amplifications during the first test on ABI3100. Therefore, only 15 primers were selected according to their PCR product quality. Another four of microsatellite primers were used that were developed earlier (Pozarkova et al. (2002) (Table 1).

After completion of the first round of genotyping for 1000 accessions, unexpectedly, several problems had occurred (Although we had developed considerable expertise with the genotyping of barley, chickpea, wheat and lentil). Among these problems in faba bean genotyping are: 1) unspecific amplicons closed to the expected size 2) high frequency of missing data which were caused by ROX size marker therefore amplicons were not recognised by ABI3100 3) weak amplifications with a strong background. However, all these artifacts and variation were ignored and considered as a missing sample. The results of the current results were summarised in table (2).

We are planning to repeat the missing sample through a second round of genotyping; we expect to finish this work within the next three months July- September 2008.

**Table 1: Four additional primers were used to genotype 1000 faba bean accessions.**

Primer	Repeat	sequence	Expected size (bp)
JF1-AG2	CT7ATCT2	CAAAGCAGTTTTTGTACATGG TTGTTAAACTAAATAATGGCGTG	181
GA1154	CT2ATCT16	GGGTCACACATCAAGCATTGG GCATGGCGGTTTTAGGGTG	251
GA3	GA14	TTGGGAGTTTAAGGGAAGAA ATCATGACCTAGCGGATGAT	192
GATA2	GA3GATA4GA4	GACGCACCCTACTGCATCAC GGTGTCCCACCAGGTAAGTCTCG	152

**Table 2: Summary of genotyping 1000 faba bean accessions with 18 microsatellite primers.**

<b>Primer name</b>	<b>% Missing</b>	<b>Max (bp.)</b>	<b>Min (bp.)</b>	<b>Range (bp.)</b>	<b>% Heterozygosity</b>
A110-1	16.7	245	117	129	8.28
F112-1	18.2	308	250	57	2.12
E115-1	43.4	300	211	89	8.92
E114-1	44.4	306	219	86	14.12
C7-1	48.1	250	204	46	12.10
O25-JF1-AG2	48.7	217	145	71	28.13
A105-1	48.9	329	248	81	16.77
G114-1	49.8	137	92	44	11.15
A102-1	51.6	254	146	108	27.28
A9	51.9	301	250	60	17.3
O23-GA1154	52.1	252	176	76	12.21
O13-GA3	55.5	237	150	87	11.57
F117-1	62.0	250	197	53	9.98
F11-1	62.7	307	266	40	2.34
E109-1	64.9	282	194	88	7.01
A117-1	69.7	214	171	44	6.37
A116-1	71.5	300	239	61	2.76
O3-GATA2	79.3	198	128	70	6.48
A109-1	79.7	240	176	64	2.76
<b>Average</b>	<b>53.7</b>	<b>256.9</b>	<b>185.0</b>	<b>72.0</b>	<b>10.6</b>

### **34. G4006.06: Genotyping of composite collection of finger millet [Eleusine coracana (L.) Gaertn]**

*January 2006–December 2006; no-cost extension to December 2007*

#### **Principal Investigator**

Hari Upadhyaya, ICRISAT, h.upadhyaya@cgiar.org  
International Crops Research Institute for the Semi-Arid Tropics  
(ICRISAT), Patancheru, 502324, India; Phone: 91-040-30713071 (extension 2333)

#### **Collaborating institutions and scientists**

- ICRISAT: Ayyarappan Bharathi; RK Varshney; DA Hoisington; CLL Gowda; CT Hash, S Chandra

#### **Composite collection**

Finger millet germplasm pool remains largely uncharacterised, thus a constraint to finding genetically diverse germplasm with beneficial traits for use in breeding programmes. A composite collection of 1000 accessions, which also consists of 622 core collection accessions (Upadhyaya et al. 2006), has been developed and molecularly profiled using 20 SSRs (Dida et al. 2007) in high throughput assay (ABI3700).

#### **Genetic structure of composite collection**

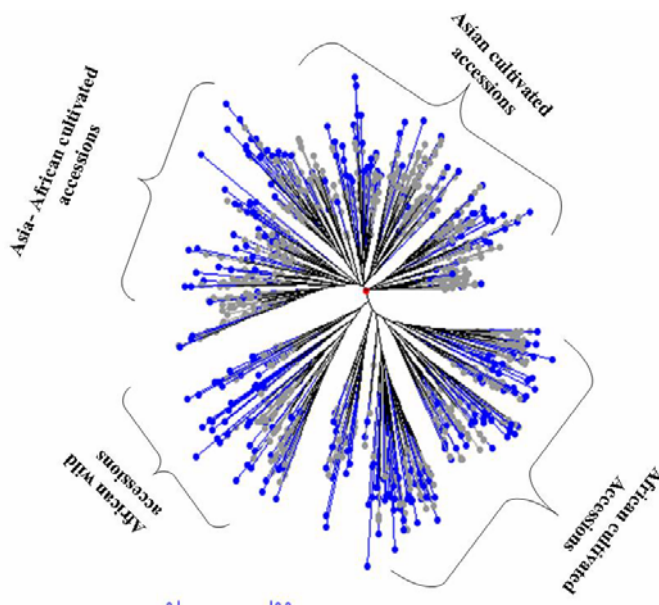
Twenty SSR markers data on 959 accessions were analyzed using PowerMarker V3.0 (Liu and Muse 2005) and DARwin 5.0 version (Perrier et al. 2003). This composite collection showed rich allelic diversity (231 alleles, 11.6 alleles per locus, 121 common

alleles and 110 rare alleles at 1%), group-specific unique alleles, and common alleles sharing between the races and geographical groups. Markers UGEP 81, UGEP 10, UGEP 102, UGEP 26, and UGEP 77 detected the large number of alleles (10-21).

Unique alleles are those detected in a group of accessions but absent in other groups. Race-specific unique alleles were 37 in *Vulgaris*, 5 in *Plana*, 4 in *Africana*, and 2 in *Compacta* while region-specific alleles were 29 in the accessions originating from East Africa, 12 in the accessions originating from South Asia, 11 in the accessions originating from Southern Africa, and one each in the accessions originating from Central Africa and Europe. The common alleles shared by two groups were 15 for East Africa and South Asia, 5 for East Africa and Southern Africa, and 3 for South Asia and Southern Africa.

### Reference set

A reference set consisting of 300 genetically most diverse accessions have been formed. This reference set captured 206 (89.2%) of the 231 composite collection alleles, representing diversity from the entire spectrum of composite collection (Figure). The usefulness of this reference set in genomics and breeding of finger millet needs to be investigated.



**Figure: Un-weighted neighbour-joining tree based on the simple matching dissimilarity matrix of 20 SSR markers across the 959 accessions of finger millet composite collection (Grey colour) with proposed reference set (300 accessions) in blue colour.**

### References

- Dida MM, Srinivasachary, Ramakrishna S, Bennetzen JL, Gale MD, Devos KM (2007). The genetic map of finger millet (*Eleusine coracana*). *Theor Appl Genet* 114:321-332.
- Liu K, Muse SV (2005). Power marker: Integrated analysis environment for genetic marker data. *Bioinformatics* 21:2128-2129.
- Perrier X, Flori A, Bonnot F (2003). Data analysis methods. In P Hamon, M Seguin, X Perrier, and JC Glaszmann (eds), *Genetic diversity of cultivated tropical plants*. Enfield, Science Publishers. Montpellier. pp 43-76.
- Upadhyaya HD, Gowda CLL, Pundir RPS, Reddy VG, Singh S (2006). Development of core subset of finger millet germplasm using geographical origin and data on 14 quantitative traits. *Genet Resour Crop Evol* 53:679-685.

**35. G4006.30: Genotyping of composite collection of foxtail millet [Setaria italica (L). P. Beauv.]**

*January 2006 –December 2006; no-cost extension to June 2008*

**Principal Investigator**

Hari Upadhyaya, ICRISAT, h.upadhyaya@cgiar.org

International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, 502324, India, Phone: 91-040-30713071 (extension 2333)

**Collaborating institutions and scientists**

ICRISAT: Ayyarappan Bharathi; RK Varshney; DA Hoisington; CLL Gowda; CT Hash, and S Chandra

**Composite collection**

Not much information is available about the usefulness of foxtail millet germplasm in breeding programmes. A composite collection of 500 accessions has been developed and molecularly profiled using 19 SSRs (Xia et al. 2007; Dida et al. 2007; ICRISAT unpublished data) in high throughput assay (ABI3700).

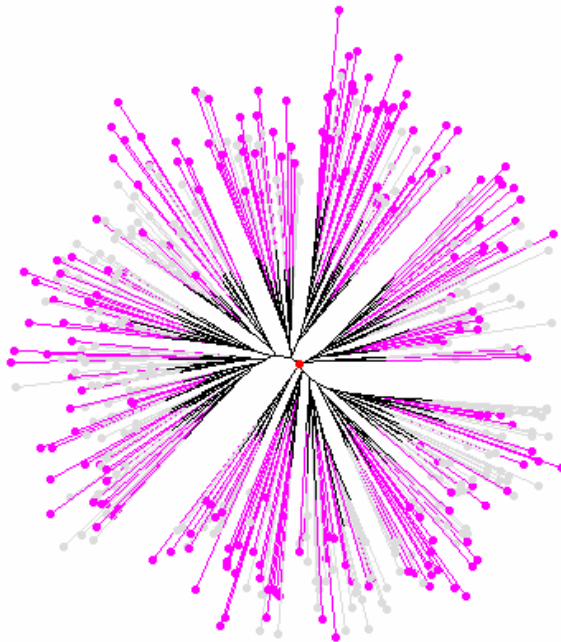
**Genetic structure of composite collection**

Nineteen SSR markers data on 452 accessions were analyzed using PowerMarker V3.0 (Liu and Muse 2005) and DARwin 5.0 version (Perrier et al. 2003). This composite collection showed rich allelic diversity (362 alleles, 19 alleles per locus, 196 common alleles and 166 rare alleles at 1%), group-specific unique alleles, and common alleles sharing between the races and geographical groups. Markers UGEP53, UGEP81, UGEP15, UGEP90, and UGEP33 detected large number of alleles (28-35).

Unique alleles are those detected in a group of accessions but absent in other groups. Group-specific unique alleles observed among the races were 40 in Indica, 21 in Moharia, 10 in Pumila, and 8 in Maxima, while region-wise unique alleles were 57 in South Asia, 17 in West Asia, 14 in East Asia, and 3 in Africa. The common alleles shared by two races were 28 between Moharia and Indica, 16 between Maxima and Indica, 9 between Maxima and Moharia, 4 between Moharia and Pumila, 3 each between Italica and Indica, and Pumila and Indica and one between Moharia and Italica. Region-wise shared alleles were 43 between East Asia and South Asia, 24 between South Asia and West Asia, 4 between Africa and South Asia, 3 each between East Asia and West Asia and Africa and West Asia, 2 between Africa and East Asia, and 1 between Europe and South Asia.

**Reference set**

A reference set consisting of 200 genetically most diverse accessions have been formed. This reference set captured 316 (87%) of the 362 composite collection alleles, representing diversity from the entire spectrum of composite collection (Figure). The usefulness of this reference set in genomics and breeding of foxtail millet needs to be investigated.



**Figure. Un-weighted neighbour-joining tree based on the simple matching dissimilarity matrix of 19 SSR markers across the 452 accessions of foxtail millet composite collection (Grey colour) with proposed reference set (200 accessions) in pink colour**

0 |—————| 0.2

## References

- Dida MM, Srinivasachary, Ramakrishna S, Bennetzen JL, Gale MD, Devos KM (2007). The genetic map of finger millet (*Eleusine coracana*). *Theor Appl Genet* 114:321-332.
- Liu K, Muse SV (2005). Power marker: Integrated analysis environment for genetic marker data. *Bioinformatics* 21:2128-2129.
- Perrier X, Flori A, Bonnot F (2003). Data analysis methods. In P Hamon, M Seguin, X.Perrier, and JC Glaszmann (eds), *Genetic diversity of cultivated tropical plants*. Enfield, Science Publishers. Montpellier. pp 43-76.
- Xia PJ, Yun SS, Yan CS, Wan GY, Wan TY, Li Y (2007). Development of EST-SSR in foxtail millet (*Sterea italica*). *Genet Resour Crop Evol* 54:233-236.

## **36. G4006.31: Development and genotyping of composite collection of pearl millet (*Pennisetum glaucum* (L.)R. Br.)**

*January 2006–December 2006; no-cost extension to Dec 2008*

### Principal Investigator

Hari Upadhyaya, ICRISAT, h.upadhyaya@cgiar.org  
International Crops Research Institute for the Semi-Arid Tropics  
(ICRISAT), Patancheru, 502324, India, Phone: 91-40-30713333

### Collaborating institutions and scientists

- ICRISAT: CT Hash; S Senthilvel; RK Varshney; DA Hoisington; KN Rai; RP Thakur; S Chandra

### Composite collection

A composite collection of pearl millet, consisting of 1021 accessions, has been developed from the world collection of 21,594 pearl millet germplasm held at ICRISAT genebank. This composite collection consists of 710 landraces, 251 advanced breeding lines, and 60



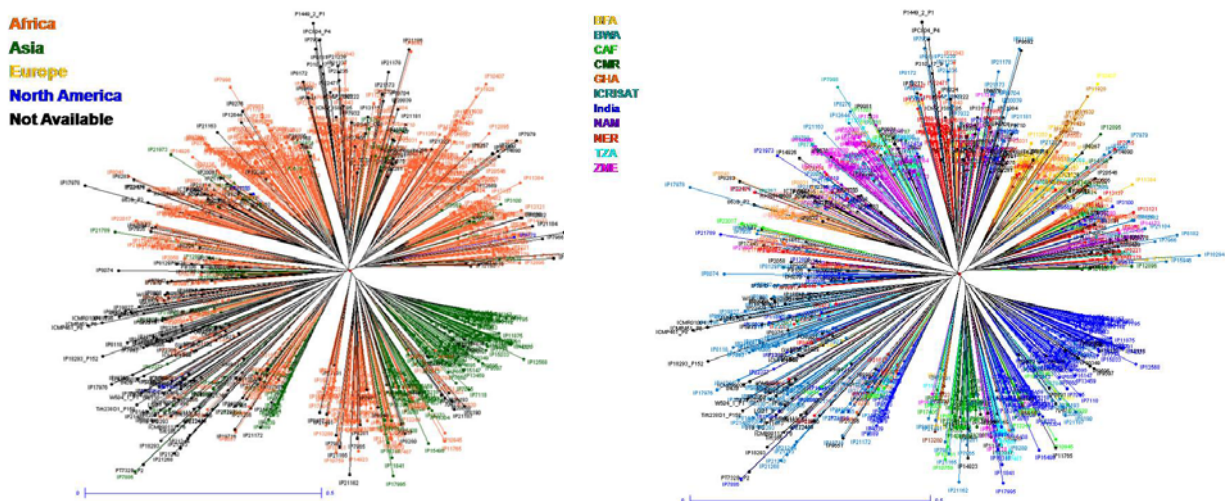
accessions from the seven wild species. Geographically, 441 accessions belong to Asia, 315 to West Africa, 147 to Southern Africa, 56 each to Eastern and Central Africa, five to America's, and one to Europe.

### Genetic structure of composite collection

Nineteen SSR markers (Allouis et al. 2001; Qi et al. 2001, 2004; Senthilvel et al. 2004) data on 1021 accessions were analyzed using PowerMarker V3.0 (Liu and Muse 2005) and DARwin 5.0 version (Perrier et al. 2003). This composite collection showed a total of 230 alleles, averaged 12 alleles per locus and 102 of them were rare alleles at 1%. The accessions were highly heterogeneous and up to 7 alleles were detected per locus. The accessions were grouped by geographical locations but not by biological status (Figure). Only seven alleles were unique to wild species whereas none were unique in landraces. The released cultivars and advanced lines were scattered across different groups. The allelic data of 19 SSR loci for 1021 accessions will be made available to the GCP central repository.

### Reference set

A reference set consisting of 300 accessions were chosen using 'max length subtree' option of DARwin5.0 which creates the subset of units minimizing the redundancy between units and limiting the loss of diversity. This reference set has captured 95% of the composite collection alleles (230), representing diversity from the entire spectrum of composite collection. The usefulness of this reference set in genomics and breeding of pearl millet needs to be investigated.



**Figure: Neighbour joining tree of composite collection based on Euclidean distance**

### References

- Allouis S, Qi X, Lindup S, Gale MD, Devos KM (2001). Construction of a BAC library of pearl millet, *Pennisetum glaucum*. *Theor Appl Genet* 102:1200-1205.
- Liu K, Muse SV (2005). Power marker: Integrated analysis environment for genetic marker data. *Bioinformatics* 21:2128-2129.

- Perrier X, Flori A, Bonnot F (2003). Data analysis methods. In P Hamon, M Seguin, X Perrier, and JC Glaszmann (eds), *Genetic diversity of cultivated tropical plants*. Enfield, Science Publishers. Montpellier. pp 43-76.
- Qi X, Lindup S, Pittaway TS, Allouis S, Gale MD, Devos KM (2001). Development of simple sequence repeat markers from bacterial artificial chromosomes without subcloning. *BioTechniques* 31:355-361.
- Qi X, Pittaway TS, Lindup S, Liu H, Waterman E, Padi FK, Hash CT, Zhu J, Gale MD, Devos KM (2004). An integrated genetic map and a new set of simple sequence repeat markers for pearl millet, *Pennisetum glaucum*. *Theor Appl Genet* 109:1485–1493.
- Senthilvel S, Mahalakshmi V, Sathish Kumar P, Reddy AR, Markandeya G, Reddy MK, Misra R, Hash CT (2004). New SSR markers for pearl millet from data mining of Expressed Sequence Tags. "New directions for a diverse planet". Proceedings of the 4th International Crop Science Congress, 26 Sep – 1 Oct 2004, Brisbane, Australia. [http://www.cropscience.org.au/icsc2004/poster/3/1/1223\\_hashct.htm](http://www.cropscience.org.au/icsc2004/poster/3/1/1223_hashct.htm).

### **37. G4007.01: Genotyping validation of the GCP reference sets**

*January 2008–December 2008*

#### **Principal Investigator**

Jean-Francois Rami, Agropolis–CIRAD

Agropolis–CIRAD - UMR DAP, TA A-96/03, Av Agropolis, 34398 Montpellier  
CEDEX 5, France, Tel:+33 4 67 61 44 65, Fax:+33 4 67 61 56 05

The scientific community involved in the SP1 sub-programme of the Generation Challenge Programme is about to deliver one of the biggest efforts of characterisation of genetic diversity on 21 crop species. This characterisation was based on the utilisation of microsatellite markers, which constitute a powerful marker system for such a purpose. However, this work was by nature composite, involving different species and different partners using different technologies.

For each crop, one of the main products of this exercise is a reference set of representative germplasm to serve as a material for international coordination in the future.

The present project proposes to assess the different microsatellite datasets produced in SP1 by having a subsample of germplasm accessions re-genotyped by an external genotyping facility (service provider). This subsample will be the reference set, so that the new data will also serve to validate and certify the genotypic information attached to the reference set.

This genotyping validation project will be connected to the management of the genetic material constituting the reference sets. As an output, stabilised materials specifically handled as genetic stocks by gene bank curators and associated to validated genetic diversity data will be available.

### **Research activities and progress at CIRAD**

The genotyping service provider has been identified and the technical specifications of the service have been specified. Special attention has been taken to the quality insurance part of the service.

Procedure for DNA shipment from each CG center to CIRAD has been setup, and DNA have been shipped or are ready to be sent for Sorghum, Coconut, Barley, Maize, Cowpea, Groundnut, Pigeon Pea, Chickpea and Finger Millet. A timetable for the availability of DNA of other species has been produced. Genotyping for Sorghum, Coconut and Barley is starting in July 2008.

### **38. G4008.01: Population development through multiparent advanced generation inter-crosses (MAGIC) among diverse genotypes to facilitate gene discovery for various traits in rice** *January 2008–December 2009*

#### **Principal Investigators**

Hei Leung, IRRI, h.leung@cgiar.org

Plant Breeding, Genetics and Biotechnology Division, International Rice Research Institute, Manila, Philippines

#### **Collaborating institutions and scientists**

- IRRI: E Redona, RK Singh

#### **Background**

MAGIC or multiparent advanced generation inter-crosses is an experimental method that increases the precision with which genetic markers are linked to quantitative trait loci (QTL). This method was first introduced by Mott et al. (2000, PNAS USA 97:12649-12654) in animals as an extension of the advanced intercrossing (AIC) approach suggested by Darvasi and Soller (1995, Genetics 141:1199-1207) for fine mapping multiple QTLs for multiple traits. Advanced Intercrossed Lines (AILs) are generated by randomly and sequentially intercrossing a population initially originating from a cross between two inbred lines. MAGIC involves multiple parents, called founder lines, rather than bi-parental control. AILs increase the recombination events in small chromosomal regions for the purpose of fine mapping. These lines are then cycled through multiple generations of outcrossing. Each generation of random mating reduces the extent of linkage disequilibrium (LD), allowing the QTL to be mapped more accurately. Lines derived from early generations can be used for QTL detection and coarse mapping while those derived from later generations will only detect marker-trait associations if markers are located very close to the QTL. This allows fine mapping of QTLs in a cost effective manner and high-resolution parallel analysis of different complex traits on the same population.

MAGIC populations have uses beyond trait mapping. The highly recombined germplasm can be used directly as source materials for the extraction and development of breeding lines and varieties adapted to different environments. The increased recombination in MAGIC populations leads to novel rearrangements of alleles and greater genetic

diversity. Thus, MAGIC populations provide an underpinning genetic resource that can facilitate the discovery, identification and manipulation of new forms of allelic variability (Cavanagh et al. 2008. *Current Opin. Plant Biol.* 11:215-221).

Rice is one of the four crops (along with sorghum, pearl millet, and cowpea) chosen for the development of MAGIC populations under the GCP. The goal of this project is to develop such MAGIC populations suitable for the localisation of QTL for multiple traits to regions of 3 centiMorgans (cM) or less in rice populations targeted for specific production environments.

### **Selection of founders and developing the populations**

Two MAGIC populations comprised of at least 2000 inbred lines are being developed at IRRI using progenitor or founder lines from two different eco-geographic races of *Oryza sativa* L. - *indica* and *japonica*. While both populations are targeted for eventual deployment in Asia, the *japonica* group (including a basmati type) may also be useful for Africa. Each population is comprised of eight diverse founder lines. This number is a compromise between the time required to establish the population and the diversity to be captured within the populations. The selection of founder lines was based on the knowledge of the available germplasm, their genetic relationship, and their relevance to breeding programmes underway at the collaborating institutes. The founder lines include traditional and modern varieties known to exhibit tolerance to a suite of biotic and abiotic stresses, wide-adaptation, high-yield potential, and good grain quality. Some lines are also part of the Rice SNP project (<http://www.oryzasnp.org/>; McNally et al. 2006. *Plant Physiology* 141:26–31).

In the first crossing cycle, 28 crosses in all probable combinations for each population have been made using a half-diallel approach. In the 2008 wet season, each single-cross population will be intermated to generate a second cycle of lines for finer mapping. Line purity and true hybridity in different mating cycles will be monitored with DNA markers. We expect to undertake 70 double or 4-way crosses out of all possible 210 crosses and 35 out of 105 possible eight-way crosses, with each founder line being represented in the genotype of each line derived for each eight-way cross. These MAGIC populations will be further intercrossed for at least an additional two more generations following schemes identified in a BBSRC (UK) funded project on wheat underway at NIAB, UK. This will lead to LD decay in a rapid and uniform manner across the whole genome. These advanced intercross populations can then be distributed freely to interested parties for use in research using the Standard Material Transfer Agreement of the International Treaty on Plant Genetic Resources for Food and Agriculture (<http://www.iris.irri.org/smta/>). Additional funding is required to advance the populations through additional rounds of intercrossing prior to the generation of a second set of inbred lines for fine mapping.

### **Phenotype and genotype lines and populations**

The generation and multiplication of fully inbred lines for use in conducting multi-location yield trials and QTL mapping would not be possible within the two-year duration of this project. However, sufficient seed of S1 or S2 progenies will be produced

to allow recording of easy-to-score characteristics. The rice founder lines will be phenotyped in the Philippines at one irrigated and one rainfed lowland trial site over the next two years (one wet and one dry season). S1 families will be available for selective phenotyping during Year 2. Within the duration of the project, we will concentrate on highly heritable traits that do not require extensive replicated trials and that are not influenced greatly by intra-plot heterogeneity.

The founder lines, samples from each population, and bulks of selfed progenies will be genotyped with a set of genome-wide markers to validate their population structure. It is planned to use a low-cost oligo-chip to generate whole-genome graphical genotypes anchored to the rice physical map, building on the rice SNP project (<http://www.oryzasnp.org/>). This will provide 300-500 well-spaced polymorphic markers for any pair of parental lines.

Though developing the diverse MAGIC populations can be cumbersome and challenging, once these populations are developed, they will become immortal resources for breeding and genetics research, in general, and for fine mapping of multiple QTLs for multiple traits in rice, in particular. The advanced recombinant lines may also be used in product development – either being used directly as released varieties or being used as parents for developing novel rice populations with wider genetic base.

**Acknowledgment:** We thank Ian Mackay and Collin Cavanagh for discussion and active collaboration in this project.

### **39. G4008.02: Phenotyping sorghum reference set for drought tolerance**

*January 2008–December 2010*

#### **Principal Investigator**

Hari Upadhyaya, ICRISAT, [h.upadhyaya@cgiar.org](mailto:h.upadhyaya@cgiar.org)  
International Crops Research Institute for the Semi-Arid Tropics  
(ICRISAT), Patancheru, 502324, India, Phone: 91-040-30713071 (extension 2333)

#### **Collaborating institutions and scientists**

- ICRISAT India: V Vadez, CT Hash, and L Krishnamurthy
- ICRISAT, Mali: F Rattunde and E Weltzein-Rattunde
- ICRISAT, Kenya: MA Mgonja
- UAS–Dharwad: PM Salimath
- KARI: CK Karari
- National Plant Genetic Resources Centre, Arusha, Tanzania: W Ntundu
- IER: M Diourte
- ISRA–CERAAS, Thiès, Senegal: N Cisse

Drought is one of the most important yields reducing abiotic constraint worldwide. It is proposed to evaluate sorghum reference germplasm set (360 of the 384 reference set accessions), selected based on the genotyping information of composite collection (41 SSR loci data on 3372

accessions), for post-flowering drought tolerance. In the first year, the reference set will be characterised for morpho-agronomic traits to classify accessions into distinct flowering and plant height groups at ICRISAT locations in India, Mali, and Kenya. In the second year, these subgroups will be evaluated for post-flowering drought tolerant traits at three ICRISAT locations (as above). In addition, they will also be evaluated at ICRISAT Patancheru, India for seed micronutrients (Zn and Fe) under varying water regimes (stressed vs unstressed conditions) to identify seed micronutrient dense lines. In third year, selected reference set accessions and stay-green QTL introgression lines will be evaluated for water uptake under stressed conditions in PVC tubes (2.0-m long and 25-cm diameter), and for the proportion of water used prior/after anthesis. In the same year, the most promising post-flowering drought tolerant reference set accessions and stay-green QTL introgression lines will be multilocally evaluated for post-flowering drought tolerance at ICRISAT and NARS locations in India and Africa. In addition to evaluating for post-flowering drought tolerance traits, additional data will be collected on grain/stover yield and component traits to identify lines that are better able to maintain normal growth/yield processes under stress. It is proposed to evaluate this select group of materials in the fourth year (subject to GCP provides funds) at NARS locations to generate additional data on the performance of post-flowering drought tolerant lines. At the completion of project, we will have a better understanding of post-flowering drought tolerance in sorghum, the traits associated with post-flowering drought tolerance, and a range of post-flowering drought tolerant sorghum lines for use in crop improvement programmes.

#### **40. G4008.03: Precision phenotyping of the GCP spring wheat reference sample for drought**

*January 2008–December 2010*

##### **Principal Investigator**

Susanne Dreisigacker, CIMMYT, [sdreisigacker@cgiar.org](mailto:sdreisigacker@cgiar.org)  
Apdo. Postal 6-641, 06600 Mexico, D.F., Mexico, Phone: +52 55 5804 2004,  
Fax: + 52 55 5804 7558

##### **Collaborating institutions and scientists:**

- CIMMYT: Matthew Reynolds, Yann Manes, Karim Ammar, Tom Payne, Hans-Joachim Braun, Jose Crossa, M Warburton, M Zaharieva
- INRA–Morocco: Rachid Dahan
- CIMMYT–Iran in collaboration with the Dryland Agricultural Research Institute: MR Jalal Kamali

Global genetic resources provide a fundamental source for further crop improvement. The GCP Subprogramme 1 aims to characterise the diversity of crop germplasm collections held by the CGIAR and its partners. This characterisation includes an assessment of the genetic structure of the collections as well as the phenotypic variation associated with that structure. The ultimate goal is to provide access to sources of genetic diversity that may supply genes and alleles involved in key agricultural traits, especially stress tolerance. During the last three years, 3000 wheat accessions provided by major germplasm banks were characterised by CIMMYT and collaborators with 50 SSR markers for the development of reference samples including accessions maximizing neutral genetic diversity. In the first year of this project we will build up a seed stock for

three developed international reference samples in wheat: the spring bread wheat, winter wheat and durum wheat reference samples. Seed will be stored in the CIMMYT wheat germplasm bank and made available for distribution. A drought specific spring bread wheat reference sample will be defined and characterised in multi-location trails for relevant agronomic traits, as well as physiological traits related to the main drivers of yield under drought. The same reference sample will be genotyped with high density DArT markers. This will allow associating the observed trait variation with the genotypic information in order to uncover QTL related to drought tolerance.

#### **41. G4008.05 Connecting performance under drought with genotypes through phenotype associations**

*January 2008–December 2010*

##### **Principal Investigator**

Jill E Cairns, IRRI, j.cairns@cgiar.org

Crop and Environmental Sciences Division, International Rice Research Institute (IRRI); IRRI, DAPO Box 7777, Metro Manila, Philippines; Telephone: +63 (2) 580-5621; Fax: +63 (2)580-5699

##### **Collaborating institutions and scientists**

- IRRI: KL McNally, A Kumar, R Serraj, H Leung
- Agropolis–CIRAD, France: M Dingkuhn, D Luquet, B Courtois
- WARDA, Nigeria: S Mande
- BIOTEC: T Theerayut
- CRRI: P Swain
- TNAU: S Robin, M Raveendran
- BAU: BN Singh

##### **Background**

Progress in developing rice varieties for water-limited environments has been slow but advances are now being made using direct selection for yield under drought stress (Bernier et al. 2007). However, this approach depends on favourable allelic combinations between parents, thus a large pool of potentially useful alleles in alternative donors remains unexploited.

Within the primary rice gene pool resides a large amount of genetic variation for abiotic stress tolerance, including drought tolerance. Indeed, drought-tolerant landraces are in the parentage of many of the mega-varieties. Rapid advances in molecular biology provide the opportunities to exploit this hidden diversity by linking allelic variation to whole-plant performance in the field through phenotype associations. Association studies at IRRI comparing observed polymorphisms for key drought responsive genes with performance under vegetative drought stress in the field allowed the identification and validation of key genes. This approach will allow the mining of favourable alleles however it was previously limited to predetermined candidate drought responsive genes. High density, whole genome, SNP genotyping has the potential to up-scaling association studies to hundreds of candidate genes/regions. Twenty rice accessions have been genotyped using this technology, with a further 2000 accessions to be genotyped

(McNally et al. 2006). To fully exploit this information for drought research, detailed physiological datasets on the performance of genotyped accessions under drought stress is required. This project will characterise a subset of accessions to be genotyped for performance under reproductive stage drought stress.

**Objective 1. Phenological characterisation of germplasm collection**

Phenological characterisation is ongoing at all six field sites. This information will be used to divide accessions into maturity groups to allow drought stress in experiments conducted in objectives 2 and 3 to be imposed at reproductive stage in all accessions.

**Objective 2. Field phenotyping for agronomic and physiological traits under control and reproductive-stage drought conditions**

**Objective 3. Model-assisted phenotyping of a subset (tropical japonicas) under semi-controlled conditions**

Phenotyping for drought response will be undertaken in project year 2. A subset of 300 tropical japonicas has been identified and all accessions are currently being increased through single seed descent at IRRI. Seeds will then be multiplied for all partners to characterise performance under reproductive stage drought stress.

**Objective 4. Genome-wide association analysis of traits associated with performance under drought environments**

To be undertaken in the last year of the project using data produced through this project.

**Deviations from work plan**

Accessions to be phenotyped under this project were to be selected from accessions which be used in OryzaSNP phase II for high-density SNP genotyping. This list is still not available thus a smaller set of accessions (approximately 700), which are expected to be included in OryzaSNP phase II, were chosen by the PI in collaboration with partners from IRRI and CIRAD. Selection was based on current genotyping projects and/or high tolerance or susceptibility to vegetative drought stress. All accessions were to be characterised at all six field locations. Specific isozyme groups will have poor agronomic performance at certain locations therefore each site will characterise sets of isozyme groups known to be suited to their environment.

This project builds on or compliments the following GCP projects:

2005-01g “Genotyping of composite germplasm sets”; 2006-03 “SNP analysis of the genetic diversity along the rice genome (HAPLORYZA)”; 2005-35 “Sequencing multiple and diverse rice varieties: connecting whole genome variation with phenotype”; SP2 Competitive Project 1 “Identifying genes responsible for failure of grain formation in rice and wheat under drought”



## References

- Bernier J, Kumar A, Ramaiah V, Spaner D and Atlin G (2007). A large-effect QTL for grain yield under reproductive-stage drought stress in upland rice. *Crop Science* 47,507-508
- McNally KL, Bruskiewich R, Mackill D, Buell RC, Leach JE and Leung H (2006). Sequencing multiple and diverse rice varieties: connecting whole-genome variation with phenotypes. *Plant Physiology* 141:26-31.

## 42. G4008.33: Drought tolerance phenotyping of the GCP maize inbred line reference set

*May 2008–February 2011*

### Principal Investigator

James G. Gethi, KARI; [jgethi@wananchi.com](mailto:jgethi@wananchi.com)  
Kenya Agricultural Research Institute (KARI), P.O. Box 340-90100 Machakos,  
Kenya +254 721831166 Fax +254 2 2338320

### Collaborating Institutions and scientists

- CIMMYT: M Warburton, M Zaharieva, S Taba, M Vargas, JL Araus, C Sanchez
- ETH: P Stamp; A Hund; R Messmer
- Agropolis–INRA: Francois Tardieu; Claude Welcker

Under GCP subprogramme 1, several projects have assessed the genetic structure of crop germplasm collections held by the CG centers and their partners, including maize in which a collection of 987 inbred lines, provided by CAAS, CIMMYT and IITA was characterised by CIMMYT and CAAS with 47 SSR markers. As a product of this study, a subset of 240 reference lines has been chosen to represent a majority of the neutral genetic diversity of the whole collection.

The overall objective of the project is to provide information about the variation, under well-characterised drought conditions, of phenotypic traits (particularly those associated with drought tolerance) in the GCP maize reference set, thus allowing the identification of accessions possessing specific drought tolerance traits and constituting a first step in further association analyses based on phenotypic and genotypic data.

### 2. Specific objectives are:

- Objective 1: Seed multiplication of the GCP reference set and complementary genotyping
- Objective 2: To develop testcross hybrids from the reference set
- Objective 3: To phenotype the reference and of the set of hybrids under field conditions
- Objective 4: To phenotype the reference set under controlled conditions: through the analysis of leaf elongation and root morphology

### 3. Activities

The reference set, as well as a set of hybrids generated by crossing the inbred lines by a common tester will be evaluated under drought (and well-watered) conditions in Mexico and Kenya. Inbred lines are from very different origins and most are not well

adapted to those environments. It is expected that yield will highly depend on adaptation patterns. Phenotyping will consequently focus more on drought tolerance traits than on yield *per se*.

Having the two phenotyping sites located in different hemispheres (but similar, subtropical, near mid equatorial latitudes) will allow two generations per year and will permit an evaluation of the two sets in different cycles/locations during the two years of the project.

In addition, the variation for leaf and root morphological traits, associated to early vigor and the leaf development under water deficit will be analyzed under controlled conditions, at INRA Montpellier and ETH Zürich. This analysis will permit examination of the variation, within the reference set for root system architecture and seedling vigor. Root system architecture, affects the ability of a root system to mobilise water and nutrients. In maize, the extent of axil root growth determines the ability to explore new, distant sources of water and nutrients. By contrast, the development of lateral roots ensures a more efficient exploitation of proximal sources.

#### **4. Expected outputs**

- The seed multiplication cycle is expected to provide 1kg of the test genotype for each line for evaluation in different locations.
- A set of at least 200 single hybrids, for testing in Mexico and Kenya.
- A set of phenotypic data, including drought secondary traits, for the lines of the reference set generated.
- An estimate and a ranking of the sensitivity of hybrid to drought during vegetative stage, and basic measurements which can be used in modelling.
- Information about variation in early vigor and its relationship with photosynthetic parameters will be generated.

#### **43. G4008.42: Developing DArT markers for several crops in the GCP** *January 2008–December 2008*

##### **Principal Investigator**

Jean-Christophe Glaszmann, Agropolis–CIRAD, jean-christophe.glaszmann@cirad.fr  
Agropolis–CIRAD, Montpellier, France, TA A-DIR PS3, Boulevard de la  
Lironde, 34398 Montpellier Cedex 5, France

##### **Collaborating institutions and scientists**

- DArT P/L, Canberra, Australia: A Kilian, subcontractor
- ICRISAT, Patancheru, India: D Hoisington, co-PI
- IITA, Agropolis–IRD, Agropolis–CIRAD, CRI, CIP: for each crop

This proposal aims at reinforcing the capacity to genotype large numbers of materials with large numbers of markers at a relatively low cost, one of the objectives of SP1 in order to facilitate the use of markers for monitoring genetic diversity. It builds on the successful commissioned project executed in 2005. It includes expanding arrays developed in the previous project for Musa (banana) and coconut, expanding arrays

developed by Diversity Arrays Technology Pty Ltd (chickpea, pigeonpea, potato) and developing new arrays for yams, groundnut and pearl millet (see details below) and testing those on an important set of germplasms (diversity and high density genetic mapping).

For each crop, a scientist or group of scientists involved in or connected to the GCP will contribute well chosen germplasm and corresponding DNA. The development of the arrays and the testing of the polymorphisms revealed, as well as further surveys will be handled by DArT P/L. In the case of pearl millet, ICRISAT will also be involved, being in the process of establishing a DArT service in its center in Hyderabad. This involvement of a CGIAR/GCP member will contribute to consolidation of a global capacity.

The outputs will be the arrays, the data, and the access to a service, either on a bilateral basis or through the GCP Genotyping Support Service. The libraries generated will be available to the GCP; their sequences will be provided when they are available.

#### **Musa array expansion and use for fine mapping**

Two arrays (6,500 clones) were developed and used for nearly 2,000 Musa accessions. We will expand by 20% the PstI/TaqI array to maximise the number of markers available to the >1000 range. It will include the Pahang genotype, the B genome germplasm (10%), with the rest (70%) of the library being a good representation of the various components within the acuminate (A genome) section.

#### **Coconut array expansion**

The available DArT markers based on worldwide coconut genotypic representation are useful for studying genetic diversity at the global level, but a low percentage are polymorphic in a local context. The aim is to obtain a set of 1000 markers for producing good genome coverage for surveying populations with high quality phenotypic information related to drought tolerance.

#### **Groundnut array development for mapping and introgression monitoring**

6 Icrisat's parental lines of mapping populations and 52 genotypes (diverse representation of *A. hypogaea* cultivated, diploid A and B genome lines, and AA x BB synthetic) are being used for proof-of-concept array development (2,500 clones). The frequency of markers is correct in wild/cultivated comparison but much lower in cultivated germplasm. A new PstI/TaqI library (4,500 clones) will be developed focussing on wide-basis backcross populations.

#### **Potato array expansion**

Initial technical tests of the array developed through the Potato consortium (including GCP and private sector) at DArT PL (over 6,000 clones) were positive, with marker frequency in the test analysis around 20%. We will contribute to the second phase, expending by 7000 clones the array, based on 48 new accessions.

### **Chickpea array expansion**

A narrow genetic background of cultivated chickpea warrants an urgent need to develop large number of polymorphic markers. A preliminary DArT array (PstI/BstNI, 9,000 clones) was developed with support from Australian Industry from cultivated (*Cicer arietinum*) and wild (*C. reticulatum*) accessions. It showed low intraspecific chickpea polymorphism. We will develop an expanded array of 18,000 clones targeting the diverse *C. arietinum* accessions, based on the reference collection and other *Cicer* species, used in introgression of wild into cultivated chickpea.

### **Pigeonpea array expansion**

An array was developed at DArT PL in collaboration with Shiyang Yang (China) (9000 clones) derived from cultivated and wild (primarily *Cajanus scarabaeoides*) accessions. We will double the number of clones from diverse cultivated pigeonpea lines (including the parental genotypes of mapping populations, accessions from the reference collection) and wild species of pigeonpea currently being used in generating backcross populations.

### **Yam array development and testing**

Array will be developed based on accessions from a collection of *Dioscorea alata* established from different countries of Asia-Pacific in the framework of an EU funded-project (South Pacific Yam Network) and on *D. rotundata* from Benin and wild relatives of these two species, already genotyped using SSR markers.

### **Pearl millet array development and testing**

Five complexity reduction methods will be evaluated based on bulked DNA samples from a diverse set of wild and cultivated *Pennisetum glaucum* and choose the most polymorphic one to develop an array of 7,000 clones. Germplasm will include parents of ICRISAT mapping populations, well-characterised inbreds from IRD and ICRISAT (Patancheru and Niamey) and wild forms of the primary gene pool.

## **Subprogramme 2: Genomics towards gene discovery**

### **44. G4007.02: Validation of drought-response/resistance pathway genes by phenotypic analysis of mutants**

*October 2007–September 2009*

#### **Principal Investigator**

Andy Pereira; Virginia Polytechnic Institute and State University, pereiraa@vbi.vt.edu  
Virginia Bioinformatics Institute, Washington Street, Blacksburg, VA 24061  
USA; Tel: (001) 540-231-3784, Fax: (001) 540-231-2606

#### **Collaborating institutions and scientists**

- IRRI: Hei Leung, Rachid Serraj, Jill Cairns
- Huazhong Agricultural University, Lizhong Xiong

#### **Research objectives**

- Screen rice mutant families for drought response phenotypes in the field at vegetative and reproductive stages.
- Characterise the drought response phenotypes of the mutant lines by a series of physiological analyses to test the putative drought resistance and water use efficiency component traits.
- Provide drought response phenotypes for Arabidopsis mutant lines of candidate genes including those not available in the rice population, both widening the gene number and encompassing comparative gene functions.

#### **Results**

##### **Drought screening of rice SAG mutants**

In the last year, a reverse genetics strategy was used to screen T-DNA mutants for candidate stress associated genes (SAGs). About 100 T-DNA mutants for SAGs were selected from the mutant database (<http://rmd.ncpgr.cn/>) and were screened for drought resistance at seedling stage in the green house.

Drought stress was initiated four weeks after germination. Visual scoring of leaf rolling and drying was carried out in the field during the development of drought stress for finding segregations of drought sensitivity/resistance within mutant family (20 plants each family).

A total of 12 mutant families showed segregation of drought sensitivity. For the mutant families showing segregations of drought sensitivity, PCR analysis was performed to check the genotypes of the inserted loci and therefore test the co-segregation of phenotype and genotype of T-DNA insertions. Among the 12 families, only 2 families showed likely co-segregation of phenotype and the T-DNA insertion. The two SAGs for the co-segregated families encode a heat shock factor and an unknown expressed protein, respectively. Co-segregation analysis at reproductive stage will be performed in rice grown season of 2008 to confirm the phenotype.

Over-expression transgenic lines were generated for three SAGs whose mutant lines showed drought sensitivity phenotypes identified before (project 2005-9). One of these

overexpression lines, with its gene encoding encoding a bZIP transcription factor, showed improved drought and salt tolerance and elevated sensitivity to ABA. The molecular mechanism of the improved stress tolerance conferred by this gene is under investigation.

### Drought response phenotyping of Arabidopsis regulatory SAGs

Microarray data analysis of Arabidopsis (Col) and rice (Nipponbare) drought induction revealed several significant differentially regulated genes using established methods (Krishnan & Pereira, 2008). The regulatory genes were identified and compared between Arabidopsis and rice (Fig 1).

Up			Down		
RtoA	A	R	R	A	RtoA
	7	7	*AP2-ERF	2	3
			ARF	2	2
			AUX-IAA	1	6
1	3	12	bHLH	2	7
4	6	5	*bZIP	3	5
		2	*C2C2-CO-like	6	
			C2C2-YABBY	1	3
	6	1	C2H2		4
	1	1	*C3H		
1	3	1	GARP-G2-like	1	
2	5	7	HB	2	3
			HMG	3	2
			HSF	2	
	1	5	LIM	1	
	2	1	MADS	5	4
1	11	16	*MYB	6	9
5	12	8	*NAC	1	1
	1	1	PHD		
	2	1	Trihelix		
2	9	3	WRKY		
	87	11	Kinases	5	79
	25	1	Phosphatases	1	18

**Fig 1: Drought responsive regulatory genes:**

Comparison of Arabidopsis (A) and rice (R) differentially expressed (Up and Down regulated) regulatory gene families (Transcription factors, kinases/phosphatases). The columns labeled RtoA denote the orthologous (conserved) genes between the two species.

From a total of around 300 Arabidopsis SAGs, 200 confirmed homozygote SALK collection mutant lines were obtained and are being screened for drought response phenotypes. In the first screen the drought response of replicated genotypes was measured in terms of reduction in biomass accumulation due to mild drought (constant soil water deficit stress maintained). Mutant genotypes have been identified which show significant differences to wild-type controls, showing higher or less reduction in growth indicating sensitivity and resistance/tolerance. The screening design allows estimation of water use efficiency and biomass changes, both of which are found among the mutants.

### Tangible outputs

1. Drought response phenotypes of around 100 stress associated genes (SAG), revealing altered phenotypes for 12 lines, for which 2 lines show cosegregation of T-DNA insertion and phenotype.
2. Drought response phenotypes of around 50 Arabidopsis regulatory SAG mutants, revealing significant differences to wild-type controls in a few.

## References

Krishnan A, Pereira A (2008) Integrative approaches for mining transcriptional regulatory programmes in Arabidopsis. *Functional Integrative Genomics and Proteomics* (in press)

### **45. G4008.06: Single nucleotide polymorphism discovery, validation, and mapping in groundnut**

*January 2008–December 2009*

#### **Principal Investigator**

Steven Knapp, UGA, [sjknapp@uga.edu](mailto:sjknapp@uga.edu)

Crop & Soil Sciences, University of Georgia, 0246 CAG T, Athens, GA, US

#### **Collaborating institutions and scientists**

- ICRISAT Groundnut Genomics and Breeding Programme: Rajeev Varshney, David Hoisington, Rupakula Aruna
- NCGR: Gregory May, Andrew Farmer
- USDA–ARS: Corley Holbrook, Peggy Ozias-Akins

DNA marker resources are currently inadequate for routine genomic and molecular breeding applications in cultivated groundnut (*Arachis hypogaea* L.;  $2n = 4x = 40$ ). The proposed research focuses on significantly enhancing the infrastructure for translational genomics and molecular breeding research in groundnut by testing the efficacy of massively parallel DNA sequencing and highly parallel single nucleotide polymorphism (SNP) genotyping strategies for SNP discovery, validation, and mapping. We are specifically proposing to: (i) develop protocols for reduced representation allele sequencing (RRS) in groundnut; (ii) enhance DNA sequence resources for groundnut using a combination of Sanger and Solexa sequencing; (iii) identify 2,000 or more common SNPs in elite lines and cultivars; (iv) develop a 1,536-SNP Illumina GoldenGate SNP genotyping array; and (v) complete the validation and genetic mapping of 1,536 SNPs in two elite recombinant inbred line (RIL) populations using an Illumina GoldenGate SNP genotyping array.

The proposed research will dramatically increase DNA sequence resources and the supply of mapped DNA markers in groundnut, should enable the identification and assembly of 20 linkage groups using elite mapping populations, particularly when coupled with genetic mapping of SSR markers, and should identify additional SNPs for genotyping assay development, validation, and mapping.

## 46. G4008.07: Improving molecular tools for pearl millet

February 2008–January 2010

### Principal Investigator

C Tom Hash, International, ICRISAT; c.hash@cgiar.org  
Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru,  
Hyderabad 502 324, AP, India, Tel: +91-40-3071-3322, Fax: +91-40-3071-3074

### Collaborating institutions and scientists

- ICRISAT: FR Bidinger, V Vadez, RK Varshney, T Nepolean and S Senthilvel
- AICPMIP: IS Khairwal
- CAZRI: OP Yadav
- RAU: PC Gupta
- ILRI: Michael Blümmel

Pearl millet (*Pennisetum glaucum*) is a dual-purpose grain and fodder crop that is an essential component of dryland crop-livestock production systems of sub-Saharan Africa (e.g., Nigeria, Niger, Burkina Faso, Mali, Senegal, Sudan, and Chad) and South Asia (e.g., India) in areas that are too hot, too dry, and/or have soils that are too acid or too infertile for reliable production of maize, sorghum or any other cereal crop. The crop is also increasingly used as the mulch component of sustainable minimum tillage crop production systems in the humid tropics (e.g., Brazil), where its acid soil tolerance, deep root system, and high vegetative growth rates under high temperature conditions often make it the best option for retrieving soil nutrients from depth, smothering weeds, and producing a mulch that protecting the soil surface from erosion by rain drop impact or surface water movement. There are limited genomic tools available for this orphan crop despite pearl millet being the 6<sup>th</sup> most important cereal crop globally and being likely to be, along with sorghum, an important source of genes and alleles that will enable plant breeders to engineer other crops (e.g., rice, wheat and maize) to better tolerate higher temperatures and increased frequencies of drought stress that are predicted to arise from on-going global warming.

This project proposes to strengthen genomic resources for pearl millet, developing EST libraries from the parents (841B-P3 and 863B-P2) of a well-characterised pearl millet drought tolerance mapping population, identifying EST sequence polymorphisms between the parents of this population, and mapping these polymorphisms using the 150 RIL progenies of this population. The augmented linkage map of this population, combined with information on the positions in the completed sorghum and rice genome sequences of homologues of the pearl millet ESTs from which these newly mapped markers are derived, be used to refine the rice-pearl millet comparative map and develop at sorghum-pearl millet comparative map. We will then use the additional markers mapping to pearl millet linkage group 2 to better define the position of a major drought tolerance QTL from 863B, using available segmental substitution lines (developed in a DBT-supported project) for this genomic region in the genetic background of elite seed parent maintainer line 841B (using funding from a BBSRC project that will start in April 2008).



In addition, we will use STS and SSR markers to skeleton linkage map two new conventional biparental pearl millet mapping populations of random inbred lines, and conduct initial testcross hybrid evaluations of these populations for terminal drought stress tolerance (measured in terms of grain and stover yield maintenance under stress conditions) and grain and stover nutritional value (measured in terms of digestibility and metabolizable energy content). Finally, we will advance eight additional pearl millet RIL populations to F7 inbred lines that will be ready for map saturation with DArT markers in a future project, which would permit development of a high density consensus linkage map for pearl millet.

**47. G4008.08: Transcriptome analysis of near-isogenic rice lines to identify expression signatures and gene combinations conferring tolerance to drought stress**

*January 2008–December 2009*

**Principal Investigator**

Shoshi Kikuchi; NIAS; [skikuchi@nias.affrc.go.jp](mailto:skikuchi@nias.affrc.go.jp)

2-1-2 Kannon-dai, Tsukuba Ibaraki 305-8602 Japan, Tel: (81)29-838-7007; Fax: (86) 29-838-7007

**Collaborating institutions and scientists**

- IRRI: Hei Leung (genetics); Venuprasad Ramaiah; Jerome Bernier; Arvind Kumar (breeding); Somayanda Impa (drought physiology); Rachid Serraj (drought physiology); Ramil Mauleon (bioinformatics) ; Violeta Bartolome (statistics)
- NIAB: Konstantina Stamati; Wayne Powell (non-funding collaborators)

By making use of two recent advances in gene expression analysis and drought-QTL mapping, the hypothesis that gene expression patterns in a chromosomal context are causally correlated with manifestation of drought tolerance as detected in near-isogenic lines will be tested. We will apply a new comprehensive 44K oligoarray platform to determine the transcriptomes of two pairs of near isogenic lines (NILs) exhibiting large difference in their yield response to drought stress at reproductive stage. Parallel to transcriptome analyses, we will determine the fine-scale genotypes of the NILs to determine whether expression signatures co-segregate with specific regions of the genome. Results from this series of studies will reveal genes or narrow chromosomal regions contributing to drought tolerance. Because the NILs are field-proven genetic stocks that are adapted to the rainfed and upland rice production environment, the results are likely to have high agronomic relevance. Experimental support to a causal relationship between gene expression patterns and QTL is of fundamental and practical interest in understanding the genetic control of a complex trait such as drought tolerance. The final output of the project will produce breeding-ready, well-characterised isogenic lines with specific chromosomal regions tagged for their contribution to drought tolerance. And also generate expression/QTL mapping datasets that can be further exploited by data mining.

### **Current status**

New 44k oligomicroarray were established by using the platform developed by the Agilent technologies and validation experiments were finished. Good reproducible gene expression data have been obtained in NAIS Group.

BC3F2 lines IR77298-14-1-2, IR77298-12-7, and IR77298-5-6 derived from IR64/Aday Sel for the transcriptome analysis are currently growing in the IRRI's greenhouse and waiting for the drought stress treatment and sampling. Gene expression analysis is planned in this August.

Based on our experience in using Affymetric chips to detect deletion mutations, we reason that the same approach can be used to detect regions of mismatches (contiguous stretches) in a pair of NILs. This can provide a high-resolution, genome-wide genotyping method to detect "heterogeneous" chromosomal regions between a pair of isogenic lines with clear phenotypic difference. We applied an algorithm (modified from that for deletion detection) to analyze Affymetrix data generated from one NIL pair. Chip hybridisation data revealed 11 regions distinct between the pair of NILs. These regions are putative introgressions from Aday Sel into IR64. We are using SSR markers within these 11 regions to confirm the prediction of chip data. We expect transcriptome and genotyping data together can help narrow down the regions accounting for different drought tolerance phenotype.

In addition, we have made F1 hybrids between the NILs to produce the materials suitable for assaying *cis*-regulation of interesting candidate genes detected by transcriptome analysis. This is in collaboration with Konstantina Stamati and Wayne Powell's group at National Institute of Agricultural Botany (NIAB), Cambridge, UK.

In ARM2008, the first gene expression data and their mining results will be presented.

## **48. G4008.09: Development of genetic and genomic resources for breeding improved sweetpotato**

*January 2008–December 2010*

### **Principal Investigator**

Roland Schafleitner, CIP, r.schafleitner@cgiar.org

International Potato Center (CIP), Avenida La Molina 1895, Apartado 1558, La Molina, Lima 12, Peru, Tel.: +51 (0) 1 349 6017, Fax: +51 (0) 1 317 5326

### **Collaborating institutions and scientists**

- IIAM: Nurbibi Cossa
- NAARI: Robert Mwanga
- INIA, Uruguay, Francisco Vilaro
- EMBRAPA: Andre Dusi
- DArT P/L: Andrzej Kilian

Production of sweetpotato, an important staple food in Sub-Saharan Africa, is limited by a number of constraints, such as low adaptability of available varieties and landraces, virus diseases, insect pests and drought. Consequently, yields achieved by resource-poor farmers in SSA are typically low and remain, on average, below 5 tons per hectare.

Improved and well adapted sweetpotato varieties with increased tolerance to biotic and abiotic stresses can significantly contribute to increasing productivity and will have a large positive impact on food and income security in Sub-Saharan Africa. However, breeding efforts are limited by the crop's genetic complexity and lack of information available about its genetic resources. The development of genetic tools, including populations and markers, and concerted efforts towards understanding the gene pools of sweetpotato would improve access to and targeted use of the allelic diversity for breeding improved varieties.

The basic tools needed to mobilise allelic diversity and to monitor introgression of desirable alleles in breeding populations consist of a well defined Composite Genotype Set and segregating populations for marker development and trait capture. Today, techniques such as DArT that yield a large number of markers for genetic studies and selection should be made accessible for sweetpotato. A diploid reference map will help to synthesize genetic information already available from independent hexaploid populations, and enable comparative genomics among sweetpotato and other crops.

This project aims at developing genetic and genomic resources for sweetpotato and will stimulate the use of these tools in ongoing breeding programmes in CG Centers and NARS.

---

## ***Subprogramme 3: Trait capture for crop improvement***

### **49. G4005.20: Optimising marker-assisted breeding systems for drought tolerance in cereals through linkage of physiological and genetic models**

*July 2005–June 2008*

#### **Principal Investigator**

Scott Chapman, CSIRO, scott.chapman@csiro.au

Queensland Bioscience Precinct, 306 Carmody Rd, St. Lucia 4067, Australia; (61) 7-32142254; Fax: (61) 7-32142950

#### **Collaborating institutions and scientists**

- The University of Queensland, Australia: Mark Dieters and Graeme Hammer
- CSIRO: David Bonnett and Greg Rebetzke
- Agropolis-INRA: Francois Tardieu, Claude Welcker
- Du-Pont Pioneer, USA: Mark Cooper (advisory role)

#### **Background**

The aim of this project is to build up the gene to trait information using data from QTL and physiological experiments and to further improve the ability of simulation models to capture the effects of traits and their integration to yield. Simulating molecular breeding programmes will assist in optimisation of MAS strategies and provide a platform for integrating a range of knowledge-based outputs from GCP into breeding programmes.

The project aims to contribute to an increase in the selection response for improved performance under drought tolerance, as well as to a better understanding of drought tolerance and its relationship with other plant breeding traits in both a physiological and plant breeding context. The main products are case studies, software tools and training.

This project is finishing in June 2008, with most of the software development being extended into the new project G4008.14. Please refer to that report for the current status of this area of work.

#### **Software development**

The main activity in the last year was vast improvement in the 2<sup>nd</sup> version of the software interface for loading genetic data and constructing breeding scenarios. The major attributes of the software are:

- accepts marker and gene information from common formats (e.g. XLS) and from the GCP iMAS project (Graphical Genotype Tool format)
- allows editing of genes and gene effects with graphical display
- reads and edits populations of genotypes to initiate breeding simulations
- produces flowcharts and allows editing of breeding programme descriptions
- runs simulations by distribution to virtual cluster network

This software system allows almost any breeding scenario (backcrossing, recurrent selection etc) to be evaluated for a given gene and marker system. The inputs are those from a typical QTL-type analysis.

More strategic analyses can be developed by including theoretical information about additional genes and traits. The system has the additional capability for dynamic linkage of crop modeling to the genetic/breeding simulations. This allows the evaluation of the pleiotropic effects of traits as mediated via their physiological linkages to yield, e.g. genes that influence leaf growth, will later impact crop yield via effects on accumulation of carbon and use of water and nitrogen resources.

### **Software delivery**

The software is now distributed as a self-installing package available from the University of Queensland website: <http://www.uq.edu.au/lcafs/qugene/>. By the time of the 2008 ARM meeting, the software links will have been made available on the GCP site.

The installation package includes:

- QUGeneUI – the graphical interface for gene and breeding simulation information
- QuGene V2 – version 2.0 of the QuGene engine to generate populations
- QuLine V2.0 – version 2.0 of the line breeding simulation model
- QuHybrid V10 – version 1.0 of the hybrid breeding simulation model
- Documentation for the above software tools

Example case studies that are packaged with the software include:

- wheat breeding programmes based on CIMMYT's programmes (modified pedigree and selected bulk) using disease, morphology and yield related genes
- single seed descent programme for designing crosses among wheat lines differing for glutenin alleles plus database of known alleles and lines
- recurrent selection for leaf elongation rate in maize using either MAS or phenotypic selection via a physiological model
- breeding strategies for selecting for increased rice quality
- hybrid maize breeding programme for recurrent selection using testcross evaluation

### **Tangible outputs**

This project has delivered the outputs (case studies and software tools) originally developed for the project to demonstrate how simulation can be used to improve the utilisation of MAS. Wang et al 2007 outlines many issues derived from these studies in designing optimal methods to use MAS in pyramiding major genes. Chapman et al 2008 describes an extension of this work to combine selection for diagnostic markers with that for minor QTL.

### **References**

- Chapman SC (2008) Use of crop models to understand genotype by environment interactions for drought in real-world and simulated plant breeding trials. *Euphytica* 161, 195-208.
- Chapman SC, Wang J, Rebetzke GJ, Bonnett DG 2008 Designing crossing and selection strategies to combine diagnostic markers and quantitative traits. *Int. Wheat Genetics Symp.* Brisbane Aug 22-25 2008.
- Wang, J, Chapman SC, Bonnett DB, Rebetzke GJ, and Crouch J 2007. Application of population genetic theory and simulation models to efficiently pyramid multiple genes via marker-assisted selection. *Crop Science* 47: 580-588.

**50. G4007.04: Association mapping of downy mildew resistance in elite maize inbred lines in Thailand**

**Reformed and postponed for 1 year, July 2008–April 2009; no cost extension**

**Principal Investigator**

Chalermpol Phumichai, BIOTEC, chalerm.pol.ph@gmail.com

BIOTEC, Center for Agriculture Biotechnology (CAB), Kasetsart University,  
Kamphaeng Saen Campus, Nakhon Pathom, 73140, Thailand

Julapark Chunwongse, Kasetsart University, julapark.c@ku.ac.th

Molecular biologist, Department of Horticulture, Kasetsart University,  
Kamphaengsaen campus, Nakhon Pathom 73140; Thailand

**Collaborating institutions and scientists**

NCSRC: Sansern Jampatong

NWSUAF: Pichet Grudloyma

Since the project started back in August 2007, the activities have been adjusted due to the late approval of proposal for the project. Because the maize planting season is during April or May, the optimum field inoculation is in June or July in Thailand. The activities for downy mildew inoculation, therefore, have to be postponed to May-July 2008. The activities are reformed and postponed for 1 year. There are 3 research activities; 1. Screening resistance/susceptible maize inbred line to artificial infection; 2. Identified grouping of genetic diversity within germplasm of maize inbred lines by using SSR markers; 3. Association analysis and SNPs development.

**1. Research activities for screening resistance/susceptible maize inbred line to artificial infection**

The data of planting and inoculating for downy mildew on National Corn and Sorghum Research Center are finished on July 2008. The other location on Nakhon Sawan Field Crop Research Center will be collected the data on August 2008.

**2. Research activities for identified grouping of genetic diversity within germplasm of maize inbred lines by using SSR markers**

Sixty maize inbred lines are genotyping. SSR primers were chosen from MaizeGDB ([www.maizegdb.org](http://www.maizegdb.org)) and tested in a preliminary experiment. A set of 40 SSR loci are in progress and will be completed by September 2008 (Figure 1).



**Fig 1. A silver-stained polyacrylamide gel (6%) at the umc1014 locus among 60 maize inbred lines.**

### **3. Research activities for association analysis and SNPs development**

DNA sequence data of 3 genes were obtained from coding regions and flanking sequence based on each chromosomal location near major QTLs: five genomic regions (chromosome 2, 3, 5, 6 and 9) (Jompatong, 2007). Primers specific to regions of the genes to be sequence were designed using software GeneFisher (<http://bibiserv.techfak.uni-bielefeld.de/genefisher/old.html>). These activities for association analysis are going to conducts.

#### **Tangible outputs delivered**

Phenotyping data of downy mildew screening of 60 maize inbred lines at National Corn and Sorghum Research Center

#### **Deviations from the workplan**

Haplotypes for selected candidate genes are concerned because there are a lot of routine works on this activity.

#### **Data availability**

Not ready yet

#### **References**

- Agrama HA, Moussa ME, Naser ME, Tarek MA, Ibrahim AH (1999) Mapping of QTL for downy mildew resistance in maize. *Theor Appl Genet* 99:519–523
- Buckler, E.S. and Thornsberry, J.M. (2002) Plant molecular diversity and applications to genomics. *Curr. Opin. Plant Biol.* 5, 107–111.
- Flint-Garcia, S.A., Thornsberry, J.M. and Buckler, E.S. (2003) Structure of linkage disequilibrium in plants. *Ann. Rev. Plant Biol.* 54, 357–374.
- Jompatong C, Jompatong S, Balla C, Grudloyma P, Jompuk C, Prodmatee N (2007) QTL Mapping for Downy Mildew (*Peronosclerospora sorghi*) Resistance in Maize. pp. 131-146. *In* Proceedings of The Thirty-Third National Corn and Sorghum Research Conference, 22-24 August 2007, Bangkok, Thailand.
- Raymundo AD (2000) Downy mildew of maize in Asia: new perspectives in resistance breeding. *In*: Vasal SK, Gonzalez Cenicerros F, Fan XM (eds) Proc 7th Asian Regional Maize Workshop. PCARRD, Los Baños, Philippines, pp 277–284

## **51. G4007.05: Bridging genomics, genetic resources and breeding to improve wheat and barley production in Morocco**

*January 2007–December 2009*

### **Principal Investigator**

Abbad Andaloussi Fouad, INRA, abbad@inra.org.ma

Institut National de la Recherche Agronomique (INRA), Division Scientifique,  
Département Protection des Plantes, Postal address 415 RB, Rabat, Morocco.

Phone: 212 37 72 72 74, Fax: 212 (0) 37 70 64 99

### **Collaborating institutions and scientists**

- INRA–Morocco: Nsarellah Nasserlehaq, Jlibene Mohammed, Lhaloui Saadia, Labhilili Mustapha, Saidi Seddik
- ICARDA: Sripada M Udupa
- University of Bologna, Italy: Roberto Tuberosa
- CIMMYT: Susanne Dreisigacker
- Cornell University, USA: Mark E. Sorrells
- University of Missouri, USA: J. Perry Gustafson

### **Research activities and progress report at INRA-Morocco and at its collaborators**

*1. Enhanced utilisation of wheat and barley germplasm using exotic/core collections and the reference sets developed by GCP to enable scientists to screen and use a broader genetic base*

Durum wheat (250 worldwide germplasm accessions and 96 Mediterranean collections), bread wheat (ME1, ME2, ME4, WWW, SYDERGRCP collections) and Moroccan barley collections were screened for biotic (Hessian fly, leaf rusts, septoria diseases) and abiotic stresses under field conditions. The screening resulted in identification of one durum wheat (DW) and 7 bread wheat (BW) lines resistant to the Hessian Fly (HF), 58 DW and 11 BW lines resistant to leaf rusts (LR) at Alla Tazi station and 13 BW lines resistant to Septoria diseases (under green house conditions).

During the current reporting period, new cross was made between some of the lines selected for leaf rusts and the Moroccan variety Aguilal (resistant to Hessian fly but susceptible to LR) to develop mapping populations, to improve the genetic base of the crop and improve agronomic performance.

*2. Development of segregating/mapping populations for tagging of loci involved in stress tolerance with molecular markers*

There was a continuation in the development of mapping populations of durum wheat and bread wheat. Parents involved presented polymorphism for yellow pigmentation, high protein content, resistance to Leaf rust, Hessian fly, cooking quality and tolerance to drought. These population are in their F<sub>2</sub>, F<sub>3</sub> for durum wheat, and in F<sub>1</sub> stage in bread wheat.

*3. Development of wheat breeding lines with improved yield, quality and adaptability and enhanced tolerance to stresses through application of genomics tools in conventional breeding programme*



Several new crosses were made in BW and DW involving parents with characteristics that are needed for improvement in dry land areas of Morocco. Some of the crosses will be also chosen for making new mapping populations for later molecular studies. Some of these crosses are being further advanced for marker-assisted selection and doubled haploid breeding. 1374 advanced durum wheat lines were also screened for the Hessian fly and saw fly and 216 lines were identified as resistant.

*4. Generation of new markers including candidate gene/allele-based molecular markers (functional markers) for enhancing the breeding strategies*

The concomitant development and phenotyping of the durum wheat mapping populations based on Telset and OMR crosses for drought, Hessian fly and Leaf rust and STB was partly made. Screening of mapping populations of DW for resistance to HF, STB was not possible in Marchouch Station due to extreme drought. Evaluation of LR in this same population was made in Sidi Allal Tazi. Segregating populations of BW were evaluated for LR and STB. Validation of markers for STB was made. Extraction of DNA from the bread wheat population was made.

*5. Mining the novel genes and alleles conferring tolerance to stresses and improving end use quality from the germplasm collection of Morocco*

The candidate genes for stresses such as Srg6, EDS1 and Coi1 were selected. The identified genes sequences were stacked and genetic polymorphism among the members of the same gene were studied.

Primers were designed to amplify the corresponding genes and genetic polymorphisms (genic-microsatellites, SNPs, and STS marker approach) in cereals using PCR techniques.

*6. Empower Moroccan researchers, young scientists and students to implement new tools of genomics in their breeding programmes*

During the current year, two undergraduate students completed their internship training, two master degree students from Moroccan Univ joined to conduct research for their thesis and one young INRA scientist started his research for his Ph.D. degree programme at University Bologna.

**References**

- Taghouti M, Rhrif K., Nsarellah N., Udupa SM. (2008) Quality and Productivity of Moroccan Durum Wheat Varieties. From seed to Pasta: The durum wheat chain, International Durum Wheat Symposium, Bologna Italy 1 - 4 July 2008, abstract, p. 214
- Ramchoun M, Amamou A, Nsarellah N., Taghouti M., M. Nachit, (2008) Genetic Variability In Agro-Morphological And Technological Traits In A Mediterranean Durum Wheat Populations. From seed to Pasta: The durum wheat chain, International Durum Wheat Symposium, Bologna Italy 1 - 4 July 2008, abstract, p. 138
- Nsarellah N., M. M. Nachit, A. Amri. (2008) Durum Wheat In Morocco, Constraints Changes, And Prospects In A Drought Prone Southern Mediterranean Country. From seed to Pasta: The durum wheat chain, International Durum Wheat Symposium, Bologna Italy 1 - 4 July 2008, abstract, p.151.
- Labhillili M. et al.. (2008) Study Of Genetic Diversity Of The Wheat Leaf Rust Fungus Puccinia Triticina In Morocco Using AFLP. From seed to Pasta: The durum wheat chain, International Durum Wheat Symposium, Bologna Italy 1 - 4 July 2008, abstract, p.188

**52. G4007.06: Integrating marker-assisted selection into the conventional breeding procedure for improvement of wheat (*Triticum aestivum* L.) in the drought-prone areas of Northern China**

*August 2007–July 2010*

**Principal Investigator**

Ruilian Jing, Institute of Crop Sciences and National Key Facility for Crop Gene Resources and Genetic Improvement, CAAS, jingrl@caas.net.cn  
12 Zhongguancun South Street, 100081 Beijing, P.R.China; Tel & Fax: ++86 (0)10 62186706

**Collaborating institutions and scientists**

- NU: Xing XU
- NWSUAF: Hui-Min XIE
- Institute of Crop Sciences, SAAS, China: Mei-Rong SUN
- LAAS: Can-Jun ZHANG
- HAAS: Xiu-Min CHEN

**The project objectives are as followings:**

1. To introduce molecular marker assisted selection (MAS) into conventional breeding programmes and establish wheat molecular breeding networks for drought tolerance (DT) improvement;
2. To develop a large-scale screening platform for drought tolerance and select DT accessions;
3. To establish a sharing platform of techniques and information to serve the wheat molecular breeding in DT cultivars.

1. Two training courses for MAS techniques and DT phenotyping in wheat were held in April, 2008. Total of 40 fellows come from 16 institutions in Northern China were involved in the training, including breeders, geneticists, postdoctoral fellows, PhD and MSc students (Fig. 1). Training content comprised (1) the National Standard “Technical Specification of Identification and Evaluation for Drought Resistance in Wheat”, which was compiled by us and first issued on 16th October 2007; (2) molecular marker detection; and (3) principium of QTL analysis and application in molecular breeding.
2. Round about 100 DT ILs at seedling stage were selected from 4500 ILs with the elite Chinese wheat genetic backgrounds using the method of repeated drought stress. Drought tolerance of about 5000 ILs has been evaluated in the fields of diverse environments. DT ILs will be primarily selected this summer.
3. Fourteen markers of large-effect QTLs controlling drought tolerance at germination and seedling stages, and DT associated traits chlorophyll fluorescence, accumulation and remobilisation of water-soluble carbohydrates in stems, ratio of spikelet setting and grain yield in drought environment have been primarily selected for DT selection in wheat. Confirmation and characterisation of target QTLs in diverse elite Chinese wheat backgrounds and environments are ongoing.

- DT accessions, techniques and methods of DT phenotyping, and information of DT QTL markers which is likely to be used in molecular breeding are exchanged among the molecular geneticists and breeders.



Fig. 1 The group photo of training course

In the next year, we are going to select 120-150 DT ILs based on the DT evaluation at seedling stage and in the fields, choose about 10 markers of DT QTLs to select DT ILs by MAS combining DT phenotype, continuously exchange information and technology among breeders and geneticists for building the capacity of wheat modern breeding in China.

### **53. G4007.07: Marker-assisted selection for Sweetpotato Virus Disease (SPVD) resistance in Sweetpotato germplasm and breeding populations**

*August 2007–July 2009*

#### **Principal Investigator**

WJ Grüneberg, CIP, [w.gruneberg@cgiar.org](mailto:w.gruneberg@cgiar.org)

International Potato Center (CIP), Apartado 1558, Lima, Peru

#### **Collaborating institutions and scientists**

- CIP: I Barker

#### **Supporting scientists**

- CIP: S Fuentes, K Huamani, J Espinoza
- NACRRI (National Crops Resources Research Institute, Uganda) R Mwangi

Sweetpotato virus disease (SPVD) is often causing serious yield losses, especially in orange fleshed sweetpotatoes (OFSPs) grown in high virus pressure zones of Sub-Saharan Africa. The disease occurs after infection of two viruses: the sweetpotato feathery mottle virus (SPFMV) and the sweetpotato chlorotic stunt virus (SPCSV). The

SPCSV is the more problematic component of SPVD, because yield losses due to SPFMV - without SPCSV infection – are low. There was no SPCSV resistance known until recently in the CIP germplasm one SPCSV resistant or highly tolerant clone was found (termed “Resistan”). The medium term objective of this study is to develop a marker assisted selection system for SPVD. The prerequisites for this objective is to have more SSR markers for sweetpotato, an appropriate mapping population, and an improved test system for SPVD. New SSR markers were developed by screening SSR sequences published between 2006 – 2008 in EST libraries for *Ipomoea*. As an appropriate mapping population a backcross population was chosen [(OFSPs x “Resistan”) x “Resistan”]. The improved SPVD test system was aiming at a combination of morphological, serological and molecular methods. In total 52 new SSR markers were developed for sweetpotato. Two populations OFSPs x “Resistan” were developed comprising more than 2 x 700 genotypes (105 OFSPs x “Resistan”). The first population was used to improve the SPVD test system and the second for field tests in a high SPVD pressure zone at NACCRI / Uganda. All genotypes of the first population were infected with SPVD by grafting on *Ipomoea batatas* cv. Costanero, which was infected by SPFMV strain RC and SPCSV strain M2-47. The developed SPVD test system comprises so far morphological and ELISA tests conducted in three replicated plants per genotype and three repeated measurements. This system will be extended by real time PCR. In total 719 (OFSPs x “Resistan”) genotypes were tested for SPVD. Twelve genotypes were found in which it was not possible to detect SPCSV and among these one genotype was found in which no SPCSV and no SPFMV was detected. The results are very promising to develop a large number of OFSPs with resistance to SPVD. It merits to develop the OFSPs x “Resistan”) x “Resistan backcross population to map SPVD resistance genes. However, it appears that the OFSPs x “Resistan” population is already segregating for SPVD, which was not expected. We conclude that SPCSV resistance is conferred one or two recessive loci and that the resistance allele already occurs at very low frequencies in OFSP populations.

#### **54. G4007.08 Integration of genomic tools with conventional screening for developing NERICA rice cultivars for West Africa** *September 2007–September 2009*

##### **Principal Investigator**

Marie Noelle Ndjiondjop WARDA, m.ndjiondjop@cgiar.org  
Africa Rice Center, 01 BP 2031 Cotonou, Benin, Tel: 00 299 21 35 01 88 , Fax:  
00 229 21 35 05 56

##### **Collaborating institutions and scientists**

- WARDA: Manneh baboucar/Drame K. Nani/Tsunematsu Hirochi
- Agropolis-IRD: A Ghesquiere/ V Verdier
- Agropolis-IRD/CIAT: Mathias Lorieux
- IER: Fousseyni Cisse

##### **Background of the Project**

The NERICAs (New Rice for Africa) are interspecific hybrids that bridge two cultivated rice species (*O.sativa* and *O. glaberrima*). They showed potential in African countries. However, additional effort to develop NERICAs with drought tolerance, resistance to rice yellow mottle virus (RYMV) and bacterial leaf blight (BLB) is needed.

## 1. Progress on research activities at WARDA

### 1a. Screening of *O. glaberrima* and interspecific lines for their tolerance to drought

Two field trials were conducted under upland conditions in 2007 using 23 accessions (18 *O. glaberrima* and 5 *O. sativa*). Two irrigation levels (control and drought stressed) were used. The drought stress was applied at vegetative stage, while the stress maintained from vegetative until flowering stage. Agronomic traits (tiller number, plant height, etc.), drought related characters (leaf rolling and drying), yield components were measured. Data were analyzed using SAS software (version 9.1).

The results showed that drought stress did not have significant effect in the first trial. It is considered that unusual rains during the stress period were the reason. Yield loss under drought stress resulted from 2.4 to 50.9% compared to control.

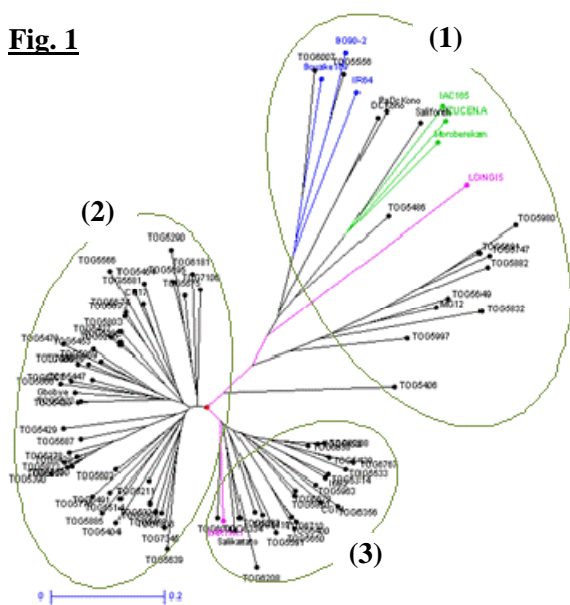
During the second field trial, drought stress effects were more pronounced than in the first trial. Average yield loss was 34.1%, ranging from 9.8% to 72.1%. Almost all varieties yielded more under control conditions than under drought stress.

As the conclusion of all the results, we selected 7 *O. glaberrima* and 2 *O. sativa* for candidates of drought tolerant donors.

### 1b. Screening of *O. glaberrima* for their resistance to RYMV

Two hundred and fifty five (255) representative accessions have been screened for resistance to RYMV through visual scoring and Elisa assays. Plants were inoculated by S1 strain. One hundred and four accessions showed high resistance (no symptom nor virus detection). PCR-specific markers of the *Rymv1* gene were tested and accessions having new alleles of *Rymv1-3* and *Rymv1-4* were identified. The lines showed complete resistance were tested in Montpellier using more aggressive S2 strain, and 21 accessions were confirmed as highly resistant. Sequencing of *Rymv1* showed that a majority of those resistant lines harbour the susceptible allele suggesting they have a resistant gene different from *Rymv1*.

**Fig. 1**



### 1c. Genetic diversity of a subset of *O. glaberrima* accessions

Ninety two accessions comprising 84 *O. glaberrima* accessions of this project, 6 *O. sativa* and 2 wild species (*O. barthii* and *O. longistaminata*) were used. PCR amplifications were performed using 31 labeled markers used in GCP genotyping world rice project. Analysis was done by using GeneMapper version 4.0 and DARwin 5.0.153. There were 3 main genetic groups consisted in (1) a mixture of *sativa*, *glaberrima* and *longistaminata* accession, (2) exclusively of *glaberrima* and (3) *glaberrima* plus *barthii* accession (Fig. 1).

## 2. Progress on research activities at IRD

A total of 47 *O. glaberrima* accessions were screened in greenhouse with four *Xoo* strains including three African strains BAI3 (race1) and BAI4 (race2) from Burkina Faso, MAI1 (race3) from Mali and one Asian's PXO86 (race2) from the Philippines as control. Nine accessions were resistant to MAI1 while PXO86 induced twelve resistance reactions.

## 3. Progress on research activities at IER

A set of 272 rice genotypes including *O. sativa*, 130 *O. glaberrima* and 120 interspecifics is being screened for drought tolerance in Mali. Data collected, included plant height, tiller number, leaf rolling, leaf drying, recovery ability, flowering date, the heading date, the sterility rate. The data will be analyzed using SAS (version 9.1).

### Tangible outputs delivered

- (1) Seven *O. glaberrima* accessions can be selected as potentially drought-tolerant:
- (2) Twenty one *O. glaberrima* accessions highly resistant to RYMV were identified
- (3) New RYMV resistant gene was suggested
- (4) Nine *O. glaberrima* accessions identified as resistant to African BLB strain
- (5) The 3 main genetic groups were identified by diversity test.

### Perspective

- (1) Data collected from drought screening of *O. glaberrima* and interspecific lines in Benin and Mali will be compile and analyze
- (2) The interspecific lines tolerant to drought will be identified and genotype with their parents using SSR markers
- (3) Screening for resistance to BLB and RYMV of the interspecific lines
- (4) Genetic analyses of the sub set of 92 accessions to pursued to include the drought tolerant data.

## 55. G4007.23: Field evaluation of wheat-barley introgression lines under different water regimes

*December 2007–November 2010*

### Principal Investigator

Márta Molnár-Láng, ARI-HAS, molnarm@mail.mgki.hu

Agricultural Research Institute of the Hungarian Academy of Sciences (ARI-HAS), 2462 Martonvásár, Brunszvik u. 2, Hungary; Tel: 36-22-569-505; Fax: 36-22-569-576

### Collaborating institutions and scientists

- CIMMYT: Maria Zaharieva
- Institute of Crop Sciences, CAAS: Ruilian Jing
- Eger-Hungary: Sándor Dulai
- ARI-HAS: Éva Darkó

## 1. Research activities and progress at ARI HAS, Martonvásár, Hungary

The present project aims to use the wheat/barley addition, substitution and translocation lines developed in Martonvásár to determine how the added barley chromosome (segments) influence various agronomic traits (drought, salt and Al tolerance) in wheat.

Objective 1. Multiplication and cytological control of the wheat/barley addition, substitution and translocation lines

**Activity 1.** The 2H, 3H, 4H Mv9kr1/Igri and 4H Asakaze komugi/Manas disomic addition lines, the 4H(4D) substitution line and the 3HS.3BL, 2DS.2DL-1HS and 7DL.7DS-5HS translocation lines were sown in the field in Martonvásár. Harvesting of the seeds will take place in July 2008.

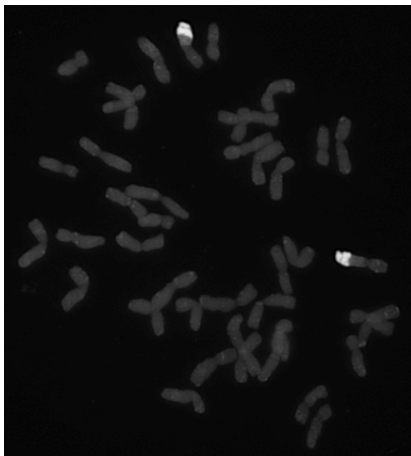
**Activity 2.** Ten plants selected by fluorescence *in situ* hybridisation from the 7H Mv9kr1/Igri and the 4H Asakaze komugi/Manas disomic addition lines and now are growing in the Martonvásár phytotron. Selection of 10 plants by GISH from the 1HS isochromosome Mv9kr1/Igri, the 6H and 7H Asakaze komugi/Manas disomic addition lines and the 4D-5HS and 6BS.6BL-4HL translocation lines is in progress. One 2H disomic addition plant has been selected from among the Asakaze komugi/Manas monosomic plants. The cytogenetic screening of the progenies of the 3H Asakaze komugi/Manas monosomic plants is under way.

Objective 2. Use of molecular markers to identify the introgressed barley chromosome segments in a wheat background and for the physical mapping of the wheat chromosomes in the translocated chromosomes

**Activity 2.** DNA samples isolated from the 3HS.3BL, 2DS.2DL-1HS, 6BS.6BL-4HL, 4D-5HS and 7DL.7DS-5HS wheat/barley translocation lines were sent to CIMMYT, Mexico for the physical mapping of the wheat chromosome arms involved in the translocated chromosomes. DNA samples from the 2H, 3H and 4H Mv9kr1/Igri disomic addition lines, wheat cultivars Chinese Spring and Mv9kr1, barley cultivars Betzes and Igri were sent simultaneously. Samples of 40 seeds from the 2D, 7D, 3B and 6B Chinese Spring ditelosomic addition lines were also delivered to the Mexican collaborator.

**Objective 3. Drought, salt and Al tolerance study of the wheat/barley addition, substitution and translocation lines**

The following genotypes were used for the drought test and the Al tolerance study:  
Wheat/barley addition lines: 2H, 3H, 4H Mv9kr1/Igri, 4H Asakaze komugi/Manas, 2H A. komugi/Betzes, 6H Mv9kr1/Betzes



Wheat/barley substitution line: 4H(4D) Mv9 kr1/Betzes  
Wheat/barley translocation lines: 3HS.3BL, 2DS.2DL-1HS, 6BS.6BL-4HL Mv9kr1/Betzes, 7DL. 7DS-5HS Mv9kr1/Igri (Figure 1)

Wheat parental lines: Chinese Spring, Mv9 kr1, Asakaze komugi

Barley parental lines: Betzes, Igri, Manas

**Figure 1.** 7DL.7DS-5HS Mv9kr1/Igri wheat/barley translocation line detected by GISH.

**Activity 1.** Ten seeds of each genotypes were sown in the field under a rain shelter, and under irrigated conditions as a control, in three replications in Martonvásár for drought tolerance test. There are three parallel plots per treatment with randomly placed genotypes. The heading time, plant height, tillering, fertility, thousand grain weight, etc. will be evaluated after harvest in July 2008. Samples of the wheat and barley parental lines of were delivered to the Chinese collaborator for phenotyping.

**Activity 2. Research activities and progress at Eszterházy Károly College, Eger, Hungary**

Tests on tolerance of polyethylene glycol(PEG)-induced drought have been made under laboratory conditions on 14 out of the 17 genotypes listed above until July 2008. CO<sub>2</sub> fixation, quenching analysis, biomass production and relative water content (RWC) have been measured.

**Activity 4. (ARI HAS, Martonvásár):** The wheat/barley addition, substitution and translocation lines listed above, together with the parental cultivars plus an Al-tolerant genotype Atlas 66 were used under laboratory conditions for aluminium-toxicity tests. The effect of toxicity triggered by Al<sub>2</sub>(SO<sub>4</sub>)<sub>3</sub> has been characterised by the extent to which inhibition in root regrowth and root mass rate are inhibited.

The results of the drought and Al tolerance tests will be evaluated in the second half of this year. The drought tolerance test under a rain shelter will be repeated in Martonvasar and China simultaneously next year. Wheat/barley addition lines which show good drought or Al tolerance will be used to induce new wheat/barley translocations which could be useful crossing partners in breeding. Some of the available wheat/barley translocation lines are genetically stable, so if these have good drought tolerance they can be used as crossing partners in breeding programmes in dry environments.

**56. G4007.24: Seed smoke treatment to favour germination under water stressed conditions**

*December 2007–September 2008*

**Principal Investigator**

Ervin Balazs, ARI–HAS, balazs@mail.mgki.hu

Department of Applied Genomics, Agricultural Research Institute of the Hungarian Academy of Sciences (ARI–HAS), 2462 Martonvasar, Brunszvik str 2.; Hungary, Tel.: +3622569500, Fax.: +3622569514

**Collaborating institutions and scientists**

- ARI–HAS: Vilmos Soos, Dr Angela Juhasz
- UKZN: Johannes van Staden, Marnie M Light



### **Objectives of the project**

The aims of the project are to investigate the physiological effect of smoke extract, the mode through which the active compound of smoke affects seed dormancy and germination under stressed conditions, using tools such as differential display and microarray and to characterise genes and regulatory networks involved in smoke action.

### **Research activities and progresses**

In the first months of the project the whole transcriptome of the smoke water - treated imbibed maize kernels and young seedlings were recorded using 48K microarray slides. Global changes in gene expression pattern were detected in 3h intervals of the germinating maize embryos and 12h intervals of maize embryo axes in the early postgermination phase, 24 and 42h after imbibition. The corresponding genes were GO annotated and the results were filtered against germination-responsive genes using rice and Arabidopsis microarray data. The up-, and downregulated genes were then clustered into functional groups. The promoter regions of the most pronounced genes were extracted *in silico* from maize BAC library databases.

### **Findings**

- smoke induced the rapid decay of mRNAs present in mature and desiccated embryos
- smoke accelerates germination
- the most highly affected genes are RING-domain containing E3 ligases, involved in the 26S proteasome regulation and LRR-receptor like kinase mediated signal transduction
- the expression of phytochrome P450 system and sulfiredoxin are also upregulated
- smoke induces the expression of many genes involved in ABA signaling; most of them are a positive regulators of the ABA response. This phenomena was particularly prominent in young seedlings.

We also conducted experiments on smoke-treated maize seedlings in the early postgerminative phase to obtain more information on the effect of smoke on seedling vigour. Smoke - treated seedlings had significantly higher dry mass, root and shoot length and their vigour was double that of control plants.

### **Next steps, expected results**

The validation of the expression pattern of the candidate genes is underway using Real-Time PCR. The possible regulatory plots will be drawn according to the transcriptome and new microarray experiments will be launched in the light of development. The microarray data will soon be placed in the ZEAMAGE and GEO database. Transgenic lines of tobacco harboring the candidate genes and their promoters will be applied for functionalisation tests and promoter-based reporter assays.

Vigour tests will be extended to stressed maize plants. Seedling vigour as a consequence of of chilling, flooding and drought stress will be determined in forthcoming months.

**57. G4008.10: Assessment of the breeding value of superior haplotypes for *AltSB*, a major Al tolerance gene in sorghum: linking upstream genomics to acid soil breeding in Niger and Mali (ALTFIELD)**

*January 2008–December 2010*

**Principal Investigator**

Robert Schaffert. EMBRAPA, schaffert@cnpms.embrapa.br and reschaffert@hotmail.com

EMBRAPA Maize and Sorghum, Caixa Postal 151, 357001-970 Sete Lagoas - MG, Brazil; tel: +55-31-3779-1076, cel: +55-31-9986-1698, fax:+55-31-3779-1076

**Collaborating institutions and scientists**

- INRAN: Maman Nouri: Soumana Souley: Magagi Abdou: Adam Kiari: Fatouma Beidari, Issoufou Kapran
- ICRISAT, Niger: Bettina Haussmann
- ICRISAT, Mali: Eva Weltzien Rattunde: Fred Rattunde
- EMBRAPA Maize and Sorghum: Jurandir Magalhães: Antônio Marcos Coelho

Aluminum (Al) toxicity is a major agricultural constraint on acid soils, which comprise over 50% of the world's potentially arable lands, particularly jeopardizing food security in the poorest regions of the globe. A major sorghum Al tolerance gene, *AltSB*, which is a membrane transporter that confers Al tolerance via Al-induced citrate release into the rhizosphere has been cloned by Embrapa and collaborating scientists. Evidences have been accumulated that a thorough scan into the sorghum genetic diversity can lead to the identification of improved versions of *AltSB* that may lead to significant agronomic advantages when used in crop production on acid soils. Thus, a research project was then designed and funded in the last competitive call from the Generation Challenge Programme to apply association genetics to identify superior haplotypes of *AltSB*, generate pre-breeding near-isogenic lines carrying these haplotypes, develop haplotype-specific markers and identify new Al tolerance genes in sorghum (ALTSORGHUM project). This project (ALTFIELD) aims at establishing a connection between the outputs of the ALTSORGHUM project and sorghum breeding programmes from Niger and Mali, ensuring that products will be properly validated in the specifically developed phenotyping sites and effectively used to attain higher and more stable yields in farmer's field on African soils where aluminum toxicity is a crop production limiting factor.

In the first months of this project concern has been raised by some that aluminum toxicity is not a crop production limiting factor in the Sahel. The impression by many is that Al toxicity is not expected to be a problem in soils that have evolved in an arid climate like the Sahel. However, the degradation of the sandy weakly buffered soils of the Sahel, can result in the acidification of the soil profile and replacement of essential nutrients such as Ca, Mg and K by exchangeable Al and H. The presence of exchangeable Al although low can be potentially toxic to crops directly and indirect through P-fixation, loss of organic carbon (OC) and decreased levels of plant available P. Toxic levels of aluminum in the soil reduces root growth, both cell division and cell elongation, thus limiting root development and increasing susceptibility to drought. Our impression is that this miss

understandingt may be due to the fact that most plant breeders are not versed in soil science and plant nutrition, and that most soil scientists are not versed in the art and science of plant improvement. Consequently, there is frequently an information void in the scientific community with respect to what the plant can do to cope with nutrient stresses, either toxic stresses like Al and deficiency stresses like P, N, and micronutrients. The problem does not have to be solved by the application of soil amendments and fertiliser alone as the plant has evolved with mechanisms to cope with these soil fertility stresses. We expect that the use of the Al tolerance gene *Alt<sub>SB</sub>* in sorghum landraces and improved cultivars can alleviate the toxic aluminum stress that is present in many soils of the African Sahel.

The ALTSORGHUM project has identified several haplotypes of *Alt<sub>SB</sub>* from several different sources with varying degrees of tolerance to Al toxicity. Our strategy is twofold; first we are initiating the process to screen a few hundred landraces from Mali and Niger for tolerance to aluminum toxicity. Those encountered with tolerance to Al toxicity can be recommended and used immediately by the farmer. Also by using Marker Assisted Selection (MAS) protocols we plan to introgress the best aluminum tolerance haplotypes into the best landraces and improved breeding improvement genotypes. In this process, we will establish aluminum tolerance field phenotyping sites for field validation of the selected and improved sorghum cultivars. We expect that the incorporation of the best alleles or haplotypes for tolerance to Al toxicity in the genetic resources from Niger and Mali will bring significant benefits to the farmers from this region. This is based on our experience of looking at the performance of near isogenic lines (NILs) and near isogenic hybrids (NIHs) on our Al tolerance phenotyping sites in Brazil. Even without noticeable moisture stress we see yield differences of over one ton of grain between the isogenic groups for tolerance to Al toxicity. Another positive aspect is the mechanism of tolerance to Al toxicity, exudation of citric acid in the rhizosphere. We have evidence in the soils of the cerrado of Brazil and from Western Kenya that citrate in the rhizosphere can increase the solubilisation of phosphate molecules, those that have been fixed by the Fe and Al oxides, in the rhizosphere and increas the availability of this nutrient.

Another positive outcome of this project will be the training and equipping of NARs scientists in implementing MAS protocols for improving sorghum cultivars. The collaborating scientists are initiating the implementation of phenotyping sites and field evaluations. Germplasm lists have been exchanged and the bureaucratic procedures for germplasm exchange have been initiated.

## **58. G4008.11 Dry bean improvement and marker assisted selection for diseases and abiotic stresses in Central America and the Caribbean**

*January 2008–December 2010*

### **Principal Investigator**

Jorge A Acosta Gallegos, INIFAP, jamk@prodigy.net.mx

Km 6.5 carretera Celaya – San Miguel de Allende. CP 38010 Celaya, Gto, Mexico; Phone 52-461-6115323 Ext. 200. Fax: 52-461-6115431

### **Collaborating institutions and scientists**

- CIAT: Steve Beebe and Matthew Blair
- INTA, Nicaragua: Aurelio del Llano; Julio Molina
- INCA: Humberto Rios Labrada and Orlando Chaveco
- CINVESTAV: Beatriz Xoconostle Cazares
- INIFAP: Ernesto Lopez Salinas; Raul Rodriguez Guerra; Alejandra Mora Aviles

Drought stress is a major constraint for common bean (*Phaseolus vulgaris* L.) production, particularly in developing countries, where the crop is popular. Depending on sowing time, the stress can be intermittent in areas of scarce rainfall, or terminal when the crop is sown at the end of the rainy season in tropical areas. Associated to drought resistance, root-rot resistance is a must for the cultivars to express whatever mechanism - trait they possess to cope with temporary drought. In addition, the bean golden yellow mosaic virus (BGYMV) transmitted by the sweet potato whitefly is an endemic disease in the tropical lowlands and tends to explode with vector populations that increase in dry years.

Previous findings have indicated that photoperiod sensitive cultivars are better adapted to intermittent stress, such as the one that occurs in the Mexican highlands; whereas neutral cultivars display wider adaptation and, if resistant, can do well in both intermittent and terminal drought. From studies conducted during the last decade in Mexico and at CIAT, bean genotypes have been developed that display outstanding drought resistance, among them cultivars Pinto Villa and Pinto Saltillo (Durango race) are still extensively used in the semiarid highlands of Mexico

### **Research activities and progresses at INIFAP**

Recently three black seeded genotypes ELS-15-55, NGO 17-99 and Negro INIFAP were identified as drought tolerant, whereas NGO 99038, NGO 99054 and NGO 99176 were susceptible; all are from the Mesoamerican race. These genotypes along with outstanding Durango race cultivars and parental stocks will be submitted to a macro array analysis. Dr. Xoconostle and her group previously carried out this macro array with a drought resistant and a susceptible cultivar (Zheng *et al.*, 2004; Barrera-Figueroa *et al.*, 2007). Several cDNAs displayed protein-coding genes associated with drought, cold and oxidative stress. The characterisation of selected protein-encoded genes demonstrated association of mRNA accumulation with drought-tolerance. The expression profiles of these mRNAs were quite contrasting between resistant and susceptible cultivars that would be interesting to screen different varieties and relate its spatial accumulation (above and below ground organs) with drought tolerance (Verdoy *et al.*, 2004). The genes identified will be used as markers for the selection of drought-tolerant genotypes in recombinant populations.

### **2. Research activities at CIAT**

A group of 200 genotypes representative of INIFAP holdings (>10,000 accessions) was analyzed with fluorescent microsatellites and data analysis is being carried out and two new markers for the bgm-1 effective recessive gene for resistance against BGMV are being tested. Several dozens of Bc2F5 lines developed at CIAT from root-rot and drought resistant cv. BAT 477 and Pinto Zapata and Mestizo, cultivars from the semiarid

highlands of Mexico, are being tested against *Fusarium* spp and *Rhizoctonia solani* isolates and under irrigated and rainfed conditions and parental cultivars will also be submitted to the macroarray analysis.

As for collaborators in Central America and the Caribbean Basin, adaptation nurseries, including only bred genotypes possessing seed of commercial value in the opaque black seeded market class, are being prepared for distribution. These nurseries will be planted at farmer's fields and evaluated in teams composed by a farmer and a researcher. The response of these genotypes to low fertility and acid soils will also be pursued in the lowlands of Mexico. Best genotypes will be increased and samples distributed among consumers to check for acceptance.

**Tangible outputs delivered:** None yet.

### References

- Barrera-Figueroa B.E., Pena-Castro J., Acosta-Gallegos J.A., Ruiz-Medrano R. & Xoconostle-Cazares B. 2007. Isolation of dehydration-responsive genes in a drought tolerant common bean cultivar and expression of a group 3 late embryogenesis abundant mRNA in tolerant and susceptible bean cultivars. *Func Plant Biol* 34: 368-381.
- Verdoy D., Lucas M.M., Manrique E., Covarrubias A.A., De Felipe M.R. & Pueyo J.J. 2004. Differential organ-specific response to salt stress and water deficit in nodulated bean (*Phaseolus vulgaris*). *Plant, Cell and Environment* 27, 757-767.
- Zheng J., Zhao J., Tao Y., Wang J., Liu Y., Fu J., Jin Y., Gao P., Zhang J., Bai Y. & Wang G. 2004. Isolation and analysis of water stress induced genes in maize seedlings by subtractive PCR and cDNA macroarray. *Plant Molecular Biology* 55, 807-823.

## **59. G4008.12 Linking genetic diversity with phenotype for drought tolerance traits through molecular and physiological characterisation of a diverse reference collection of chickpea** *January 2008–December 2009*

### Principal Investigator

Junichi Kashiwagi, Associate Scientist (Crop Physiology), ICRISAT  
International Crops Research Institute for the Semi-Arid Tropics (ICRISAT),  
Patancheru, India

### Collaborating institutions and scientists

- ICRISAT: Rajeev Varshney, Senior Scientist (Applied Genomics), Lekha Pazhamala, (Applied Genetics), Hari Upadhyaya, Principal Scientist (Genetic Resources), Subhash Chandra, Principal Scientist (Biometrics), David Hoisington, Principal Scientist and Leader- GT Biotechnology, L. Krishnamurthy, Scientist (Crop Physiology)
- JIRCAS: Satoshi Tobita, Senior Scientist (Crop Physiology), Osamu Ito, Director (Soil Chemistry)
- UAS: M.S. Sheshshayee, Associate Professor (Plant Physiology)

The objectives of this project are (a) characterisation of target drought environments (b) physiological characterisation of the chickpea reference collection (n=300) to aim at improving the drought tolerance through trait based breeding strategy, and (c)

identification of molecular markers using association genetics approach, to improve the drought tolerance in chickpea.

Till now, the seed multiplications of chickpea reference for coming field trials (at Patancheru-ICRISAT, and Bangalore-UAS) have been completed. In our recent collaborative study of ICRISAT and JIRCAS has shown the existence of a clear relationship between carbon isotope discrimination ( $\Delta^{13}\text{C}$ ) and TE in chickpea (Kashiwagi et al. 2006). It encouraged for facilitating to evaluate high throughput phenotyping on TE in chickpea reference collection.

To develop logistic workplan, planning meeting was held at ICRISAT with UAS scientist. We understand the role and time frame of each institute and scientist, sub-contract has been made with JIRCAS and UAS (in progress).

For coming field trials which will be started in next October, all necessary preparations are being progress.

## **60. G4008.13: Improving drought tolerance phenotyping in cowpea** *January 2008–December 2010*

### **Principal Investigator**

Jeff Ehlers, UC–R, jeff.ehlers@ucr.edu

Dept. of Botany and Plant Sciences, University of California, Riverside, 900 University Ave., Riverside, CA 92521-0124, USA; 951-323-5918; Fax 951-827-4436.

### **Collaborating Institutions and scientists**

- UC–R: Timothy Close, Philip Roberts
- Texas A&M, USA: William Payne, Bir B. Singh
- ISRA: Ndiaga Cisse
- INERA: Issa Drabo
- IITA: Nigeria: Satoru Muranaka, Ousmane Boukar

This project will: 1) provide baseline drought tolerance information for early and medium cycle cowpea varieties and assess the importance of genotype x environment interactions for grain yield under drought across a range of environments; (2) determine the relationship between grain yield under drought and various traits, and select applicable methodologies for practical and efficient indirect measures of drought tolerance, such as thermal imaging, that are relevant to the major cowpea production zones in Africa; and (3) determine the relationship between drought tolerance and shoot and root traits, and select potential drought tolerant genotypes with beneficial root characteristics which contribute higher productivity under drought conditions. Thirty early-maturing and thirty medium-maturing cowpea varieties are being compared for grain yield under terminal drought conditions using late plantings at two sites during the main growing season in West Africa and in four controlled irrigation and rain-free environments in West Africa and California. This will provide baseline drought tolerance information that will allow identification of drought tolerant and susceptible ‘checks’ for future drought studies and provide an estimate of genotype x environment interaction for grain yield under drought,

including the degree of correlation between the results of off-season controlled environment screening and results from main-season African growing environments. Information about the importance of genotype x environment interactions will guide future investigators on whether to breed for specific regions separately, or whether region-based and/or off-season drought-screening nurseries can be employed effectively to breed for improved drought tolerance. Identification of efficient indirect selection methods like thermal imaging allows screening of a large number of germplasm lines to help ensure capture of traits that exist in the cowpea germplasm pool, and may also help reveal important component characteristics contributing to grain yield under drought. Thermal imaging is a potentially powerful method for drought tolerance screening that has not been comprehensively evaluated for its ability to discriminate drought tolerant and susceptible cowpea genotypes and this proposal seeks to establish its usefulness in cowpea. To date, seed of the sixty cowpea entries have been multiplied at UCR for the series of multilocation trials being planned. Following USDA Phytosanitary Certification of the seeds (granted May 2008), 900 seeds per entry were distributed in June 2008 to the collaborating institutions in Senegal, Burkina Faso and to IITA-Kano. Planting of the first field trials will occur in July and August as scheduled. Also, work to validate several criteria for field/pot evaluation of drought tolerance has been initiated at IITA-Kano.

## **61. G4008.14: Breeding for drought tolerance with known gene information**

*January 2008–December 2009*

### **Principal Investigator**

Jiankang Wang, CAAS and CIMMYT–China; wangjk@caas.net.cn or jkwang@cgiar.org

Institute of Crop Science and CIMMYT China Office, No 12 Zhongguancun South Street, Beijing 100081, China; (86) 10-62133890; Fax: (86) 10-68975212

### **Collaborating institutions and scientists**

- CSIRO: Scott Chapman; Nick Hansen; David Bonnett
- University of Queensland, Australia: Mark Dieters
- CAAS: Ruilian Jing, Xianchun Xia; Huihui Li; Changbin Yin
- GCP: Jean-Marcel Ribaut
- CIMMYT: Gary Atlin; Matthew Reynolds; Yunbi Xu
- Agropolis–INRA: Francois Tardieu; Claude Welcker
- ICRISAT: Dave Hoisington; Shyam Nigam; Vincent Vadez

## **1. Simulation tools**

### *1.1 QuHybrid*

The existing QuLine module can simulate almost all breeding activities for selecting inbred lines. It can also be used when inbred parental lines are developed to be used in hybrids, such as in the case of maize. A preliminary hybrid breeding module QuHybrid has been built to implement of test crossing and hybrid performance prediction. The remaining development for QuHybrid is to implement reciprocal test crossing and tools to manage parental populations. Together, QuLine and QuHybrid will allow most breeding programmes for diploid CGIAR mandated crops to be simulated and optimised.

### *1.2 Desktop tools for creating and reviewing breeding strategies*

The new QU-Gene interface which was developed in the previous project has been developed further to facilitate simple loading of marker datasets to be used in simulation:

- gene and marker information from Qu-Gene can be read from or written to the GGT format which is used in the GCP-iMAS software;
- graphical diagrams of gene networks and breeding strategies can be drawn to facilitate accurate representation of real-world and simulation studies;
- results from breeding simulations can be graphically interrogated in a high-speed database format, allowing comparisons of genetic gain for any trait and changes in allele frequency for any gene or chromosome.

### **2. Improved statistical method for mapping quantitative trait genes**

Under the assumption of additivity of QTL effects on the phenotype of a trait in interest, we proved that the additive effect of a QTL can be completely absorbed by the flanking marker variables, and the epistatic effect between two QTL can be completely absorbed by the four marker-pair multiplication variables between the two pairs of flanking markers. Based on this property, we proposed an inclusive composite interval mapping (ICIM) by simultaneously considering marker variables and marker-pair multiplications in a linear model. Epistatic QTL can be identified by ICIM no matter whether the two interacting QTL have any additive effects. Simulated populations and one barley doubled haploids (DH) population were used to demonstrate the efficiency of ICIM in mapping both additive QTL and digenic interactions.

### **3. Gene information for drought and drought-related traits**

A number of maize mapping populations have been developed and phenotyped under various water-stressed conditions in the GCP partner project. Combined QTL analysis was conducted based on a maize RIL population derived from two inbred lines CML444 and SC-MALAWI, with the aim to identify the QTL distribution patterns in the maize genome. Additive and digenic interacting QTL for female flowering time, male flowering time and plant height were identified under three water stressed conditions in Mexico and Zimbabwe. A journal article is currently under preparation based on the results.

### **4. Designing efficient drought tolerance breeding strategies**

We have completed the case study using coleoptile length in wheat to show the simultaneous selection of major genes and QTL in plant breeding. One journal article is under preparation and will be submitted to Crop Science soon. In addition, we are collating marker and gene information in one selected wheat breeding programme in CAAS. This information is important for later simulation experiment design. To facilitate adoption of the simulation methods, we are planning to run a training course (“QTL mapping and breeding simulation”) in Sichuan Agricultural University in October 2008.

### **Tangible outputs delivered (refer to Appendix A of project description)**

*Outputs 2, 3, 4:*

- Prototypes of new simulation development and analysis tools developed
- Preliminary version of the hybrid breeding simulation tool: QuHybrid
- See <http://www.uq.edu.au/lcafs/qugene/> for latest simulation software release



*Output 5:*

- Preliminary version of the hybrid breeding simulation tool: QuHybrid
- Inclusive composite interval mapping (ICIM) for both additive and dominance QTL, and digenic interacting QTL
- Integrated QTL mapping software: QTL IciMapping v2.0, available from <http://www.isbreeding.net>
- Journal publication: Li, H, J.-M. Ribaut, Z. Li, and J. Wang. 2008. Inclusive composite interval mapping (ICIM) for digenic epistasis of quantitative traits in biparental populations. *Theor. Appl. Genet.* 116: 243-260.
- Journal publication: Zhang, L., H. Li, Z. Li, and J. Wang. 2008. Interactions between markers can be caused by the dominance effect of QTL. *Genetics* (accepted upon revision).

*Output 7:*

- Conference poster paper accepted: Chapman SC, Wang J, Rebetzke GJ, Bonnett DG Designing crossing and selection strategies to combine diagnostic markers and quantitative traits. *Int. Wheat Genetics Symp.* Brisbane Aug 22-25 2008.

**62. G4008.15: Developing Potato cultivars adapted to Southern Africa countries**

*February 2008–February 2010*

**Principal Investigator**

F Vilaró, INIA–Uruguay, [fvilaro@lb.inia.org.uy](mailto:fvilaro@lb.inia.org.uy)

INIA Las Brujas, Las Piedras, Canelones, Uruguay, phone: 00598-2-3677641, fax 00598-2-3677609

**Collaborating institutions and scientists**

- INIA, Chile: J. Kalazich
- INTA, Balcarce, Argentina: M. Huarte
- EMBRAPA: Arione Pereira
- IIAM: Carolino Martinho
- DARTS, Malawi: Obed J. Mwenye
- CIP, Malawi: Paul Demo
- CIP: Stef de Han

Potato is one of the highest value crops and provides high nutritious food in a very short growing period. Many developing countries including non Andean South American and in Southern Africa, grow long day adapted *Tuberosum* potatoes, almost year round. Breeding programmes in the northern hemisphere have developed varieties from this same Group, with high commercial quality. However, most of these varieties are mainly adapted to temperate climate and lack resistance to diseases and pests making potato highly dependent on external inputs. They also require well established seed programmes and are mainly adapted just to one crop per year. Adequate planting material is usually expensive and difficult to obtain in appropriate condition for most developing countries. Short day germplasm and landrace varieties from the Andes, have valuable traits but adapt poorly to long days and or high temperature. Genetic resistance sources for various diseases have been incorporated in advanced potato germplasm from participant non

Andean South American countries. These countries cover a wide region of environments, from southern temperate Chile to subtropical Brazil, possessing germplasm with a wide range of adaptation. In this region, with the exception of the most southern area, potatoes are grown on a two crop per year regime. Several varieties significantly improved on quality aspects have been released and are being grown in and out of the region. This project will evaluate advanced germplasm from this region, along with CIP improved germplasm on Southern Africa (Malawi and Mozambique). Microarray DaRt technology analysis will be employed to analyze population structure of germplasm from participating programmes. Secondly, easy to use molecular markers will be validated and applied in Latin America helping to characterise degree and stability of disease resistance. GIS site characterisation will be employed to determine potential variety deployment in given locations. It is anticipated that promising germplasm sources and very valuable genotypes adapted to various growing constraints, could be identified and multiplied for releasing new cultivars. This would promote a more sustainable crop for helping resource poor farmers in these countries.

**63. G4008.16: Speeding the development of salt-tolerant rice varieties through marker-assisted selection and their dissemination in salt-affected areas of Bangladesh**

*January 2008–December 2009*

**Principal Investigator**

Abdelbagi M. Ismail, IRRI, abdelbagi.ismail@cgiar.org

Crop and Environmental Sciences Division, International Rice Research Institute (IRRI), DAPO Box 7777, Metro Manila, Philippines; +63 (2) 580-5600 ext. 2843; Fax: +63 (2) 580-5699

**Collaborating institutions and scientists**

- IRRI: Michael J Thomson; David J Mackill; Thelma Paris
- BRRI: MA Salam
- Dhaka University: Zeba I Seraj
- Bangladesh Institute of Nuclear Agriculture: Mirza M Islam

Salt stress is a major constraint across many rice-producing ecosystems worldwide, because of the high sensitivity of modern rice varieties to salinity. This forces farmers to continue to grow their traditional landraces with low yield and low grain quality. In Bangladesh, salt-affected areas covers about 1 million ha across the southern parts of the country, and poses a serious problem for resource-poor farmers who depend on rice production for their livelihoods where other crops can barely grow during the monsoon season because of excess water, as well as during the dry season, because of high salinity. If modern high-yielding rice varieties were developed that were adapted to these saline conditions, there would be enormous scope for improving the lives of the resource-poor farmers living off these marginal lands.

This project aims to take advantage of modern breeding tools, such as marker-assisted backcrossing (MAB), to develop high-yielding salt-tolerant rice varieties adapted to the conditions of coastal areas of South Bangladesh. We are building upon the knowledge gained concerning the genetic control of salinity tolerance in rice to increase the speed



## References

Ismail AM, Heuer S, Thomson MJ and Wissuwa M (2007). Genetic and genomic approaches to develop rice germplasm for problem soils. *Plant Mol Biol* 65: 547-570.

### **64. G4008.17: Application of marker-assisted selection for *Striga* resistance in cowpea**

*January 2008–December 2009*

#### **Principal Investigator**

Jean Baptiste Tignegre, INERA, racinetignegre@yahoo.com or  
jb.tignegre@messrs.gov.bf

Institut de l'Environnement et des Recherches Agricoles (INERA), CREAM de  
Kamboinse 01 BP 476 Ouagadougou 01 Burkina Faso; Tel (Off.): +226 50 31 31  
92 02/08; (Hom): +226 50 38 27 70; Cell.: +226 76 02 49 50 Cell.:+226 70 04  
92 72

#### **Collaborating institutions and scientists**

- INERA: Jeremy T Ouedraogo, Issa Drabo
- IITA: Satoru Muranaka, Boukar Ousmane
- IER: Mamadou Toure

In West Africa, cowpea is a strategic edible crop due to its high protein and micronutrient contents, and therefore grown in a continuous fashion to alleviate poverty and achieve food security. However, biotic and abiotic constraints limit the production, resulting in severe yield reduction at smallholder farmer level (300-700 kg/ha), even though potential productivity of cowpea reaches 4t/ha under well managed field.

*Striga gesnerioides* (Willd.) is a parasite of cowpea and a major constraint of cowpea production in West and Central Africa. The cowpea infected by *Striga* causing severe chlorosis, wilting, and stunting of susceptible hosts and yield losses is estimated in millions of tons annually. Conventional breeding efforts have developed some varieties for the *Striga* problems as well as other important agronomic and resistance traits, but it is time-consuming and difficult pyramiding favorable traits. Marker assisted selection (MAS) is a modern and potential tool to fast track the breeding process and increase efficiency of breeding activities. Under GCP project “Marker development and marker-assisted selection for *Striga* resistance in cowpea”, MAS methodology for *Striga* resistance is now in the final stage of development. By using the MAS for *Striga* resistance, cowpea breeder can fasten the breeding process and reduce the size of population for field screening. The cooperative work proposed here, involving the “Institut de l’Environnement et des Recherches Agricoles” (INERA) of Burkina Faso and the International Institute of Tropical Agriculture (IITA), seeks to apply the MAS strategy into cowpea breeding activities for Burkina Faso and Niger to achieve rapid and reliable screening of *Striga* resistant cowpea lines. The outcome of this work will be well-adapted *Striga* resistant cowpea varieties available to farmers in Burkina Faso and Niger Rep. It is expected that farmers will achieve higher yields of better quality cowpea that would impact favorably on their general livelihoods.

### **1. Identification of farmers' constraints and preferences on cowpea production**

The objective of this work is to identify breeding priorities for cowpea in Striga-prone areas of Burkina Faso. Therefore, a participatory rural appraisal was conducted across 3 sites to identify the criteria upon which farmers adopt varieties with regard to Striga resistance and agronomic traits. Big grain size and white cowpea grain coupled with Striga and disease resistance are requested by farmers of central and Est Burkina Faso for food and trade. North Burkina farmers need brown colored grain for food and big and white grain for market. In Niger where another race prevails, farmers are kin for big size of cowpea grain.

### **2. Identification of new sources of Striga resistance for Burkina Faso (INERA)**

The objective of this study was to identify new sources of resistance with better adaptation to Striga races in Burkina Faso and Niger Republic, which will be utilised in the breeding programmes of both countries. A total of hundred and eight entries comprising lines from IITA core collection (including Niger lines or genetic back ground), Burkina Faso wild cowpea species, landraces with good characteristics and controls (High to moderate Striga-resistant and Striga-susceptible cultivars) were involved in a greenhouse-pot Striga-infested and field trials during the off-season 2007. An effective pot-screening method for Striga resistance developed by Musselman and Ayensu (1983) was used. As results, genetic materials, with Striga resistance were identified that enable Striga control (race 1) in Striga hot-spot zones of Burkina Faso. Two lines IT93K-693-2 and K VX 771-10 respectively improved lines from IITA and INERA-Burkina Faso were shown to have resistance to Striga, and met farmer's adoption criteria.

### **3. On-going activities**

#### **Study on the allelic relationship**

With Striga resistance genes Rsg1 and Rsg2, different nonallelic genes can confer a resistance behaving as isoepistatic (Fasoula and Fasoula, 1997. Parents, with good agronomic characteristics and disease resistance are therefore being crosses with Striga sources of resistance identified in the screening trials for the this purpose.

#### **Implementation of a MAS for Striga resistance in breeding programmes (IITA and INERA)**

An applied marker assisted selection study aims at (1) validating AFLP SCAR markers for Striga race 1 of Burkina Faso and (2) at determining how useful molecular markers can be in a plant breeding programme. The appropriate combination of pot-screening of backcross populations and MAS breeding, with a good exploitation of AFLP-SCAR aims at achieving an efficient and reliable MAS breeding technique (Boukar et al. 2004; Ouedraogo et al, in press). The top thirty resistant parents are being screened using AFLP SCAR markers. Back cross breeding assisted with MAS breeding are on-going.

### **References**

Boukar, O., Kong, L., Singh, B. B., Murdock, L. and Ohm, H.W., 2004. AFLP and AFLP-Derived SCAR markers associated with Striga gesnerioides resistance in cowpea. *Crop Science*, 44: 1259-1264.

Fasoula D.A. and Fasoula V.A., 1997. Gene action and plant breeding. *Plant breeding reviews*, 15:315-374.

Musselman, L.J. and Ayensu, E.S., 1983. Taxonomy and biosystematics of *Striga*. In: Ayensu E.S. and al. (eds.) *Striga Biology and Control*. Proceedings of a workshop, Dakar, Sénégal, 1983. ICSU Press: 37-43.

**65. G4008.19: Incorporation of an MSV resistance gene in Mozambican maize varieties, mediated by use of MAS**

*January 2008–December 2010*

**Principal Investigator**

Mark Laing, UKZN, [laing@ukzn.ac.za](mailto:laing@ukzn.ac.za)

ACCI, Plant Pathology, PBag 1 Scottsville, 3209, RSA, South Africa

**Collaborating institutions and scientists**

- IIAM: David Mariote, Pedro Fato, Calisto Bias
- UKZN: Tongo Tongoona, John Derera, Greg Watson, J Danson
- 

This a three way project between the GCP, UKZN (ACCI and the Molecular Biology Unit (MBU) and IIAM in Mozambique. The issue is that maize in Africa, including Mozambique, is often severely damaged by maize streak virus (MSV). However, CIMMYT has succeeded in identifying an excellent gene for MSV resistance and distinct flanking markers. However, IIAM in Mozambique does not have MAS equipment at the relevant research stations (Chimoio, Chokwe and Umbeluzi). The Mozambican plant breeders will undertake a backcross programme, to introgress MSV resistance into Mozambican maize lines, both for hybrids and popular OPs. They will then use Merck FTA cards to collect DNA from individual maize plants in the field. The FTA cards with their DNA samples will be shipped to the MBU at UKZN, where the DNA will be extracted and Real Time PCR will be used to determine the presence or absence of the MSV resistance marker. Once in the F2 generation, AFLP will be used to determine closeness of match to the recurrent parent, in order to accelerate the backcrossing process, again based on the DNA collected onto FTA cards.

Hence, we hope to develop a model of dispersed MAS, using a regional genotyping facility, where the genotyping is done as a routine assay in a service laboratory. The technical breakthrough that makes this really feasible and cost effective is the availability of Merck's FTA cards that allow for the dispersed collection of DNA. Similarly *RNAlater*® allows for the preservation of plant and animal tissue for the delayed extraction of RNA for RT-PCR. These two technologies transform the opportunity for plant breeders to use MAS and RT-PCR, irrespective of where they are located or what equipment and resources they may be limited to locally.

## **66. G4008.30: Development of a GCP phenotyping network**

*February 2008–February 2009*

### **Principal Investigator**

Abraham Blum and Greg Edmeades (consultants), [ablum@plantstress.com](mailto:ablum@plantstress.com);  
[greg\\_edmeades@msn.com](mailto:greg_edmeades@msn.com)

### **Collaborating institutions and scientists**

- John O'Toole (consultant)
- CIAT: Glenn Hyman
- University of Leuven, Belgium: Sam Geerts

This project aims to establish a strategic network of field phenotyping sites for GCP target crops. In year 1 the project identifies a total of some 10-12 field phenotyping platforms (FPP) that will become centres of excellence in phenotyping for drought tolerance, in environments to which the GCP target crops are well adapted. Methods used in identification will rely on analysis of georeferenced climate data, water balances, target crop distribution and G x E interaction of selected germplasm. These are combined with site visits and previous experience of requirements to conduct uniform managed stress field trials. Requirements in land, irrigation, field equipment and personnel needed to conduct precise managed stress drought trials were determined. Potential candidate local field phenotyping platforms (LPP) in national programmes possibly linked to FPPs are also being identified where possible.

At this time (July 2008) eight potential candidate FPP sites were visited and evaluated in China, three in total in Brazil and Colombia and as yet an undetermined final number in Africa and India (and possibly Thailand). Development of crop homology maps and georeferencing of these candidate sites is initiated based on data produced by site visits.

## **67. G4008.34: Environmental assessment for phenotyping network**

*January 2008–December 2009*

### **Principal Investigator**

Glenn Hyman, CIAT, [g.hyman@cgiar.org](mailto:g.hyman@cgiar.org)  
International Center for Tropical Agriculture (CIAT); AA 6713, Cali, Colombia  
(57-2) 445-0000; Fax: (57-2) 445-0073

### **Collaborating institutions and scientists**

- KU Leuven, Belgium: Dirk Raes, Sam Geerts and Nirman Shrestha
- EMBRAPA: Reinaldo Gomide
- CIAT: Peter Jones

### **1. Project summary**

The GCP's Phenotyping Network is comprised of people and experimental sites. The network addresses the phenotyping needs of breeding and research programmes through three components. First, sites are being selected where the GCP will support phenotyping activities. Second, the GCP will develop protocols and standards for the research conducted at the selected sites. Finally, our sub-project will characterise the sites in terms

of climate, soils and agriculture, and in relation to other sites and environments in GCP priority areas.

## 2. Activities and progress

### *Agro-climatic assessment*

We first worked with others in the network to build an initial list of proposed field phenotyping sites. The list of candidate sites was shared via the Internet and then converted to map files for GIS and Google Earth.

An agro-climatic assessment of the each site used LOCCLIM and other software to calculate climatic variables (FAO, 2005). Site profiles show climate data for 50 potential sites. Data for each site was interpolated from weather stations from an FAO database of nearly 30,000 stations around the world. The data developed included:

- Rainfall and reference evapotranspiration (ET<sub>o</sub>)
- Solar, photosynthetically active and net radiation
- Mean, maximum and minimum temperatures
- Day length
- Aridity index
- Sunshine hours
- Length of growing period

The variables listed above were derived from direct measurements from weather stations and calculated according to standard methods. Evapotranspiration was calculated using the Penman-Monteith method. The growing season as defined by the Agro-Ecological Zones Project (FAO, 1978) is the period during a year when precipitation (P) exceeds half the reference evapotranspiration. The Aridity Index (AI) is used to quantify precipitation deficit over atmospheric water demand as described by UNEP (1997). Radiation variables were calculated with the ‘ET<sub>o</sub> calculator’ – developed for calculating ET<sub>o</sub> using FAO Penman-Monteith equation ([www.iupware.be](http://www.iupware.be), select *downloads*, select *software and manuals*).

### *A Phenotyping Network Atlas*

We used a map viewer to overlay the network sites on agricultural and environmental information. These data were also used in the subsequent overlay analysis. The agricultural information included maps of key GCP crops from IFPRI’s SPAM – Spatial Allocation Model (You and Wood, 2006). Spatial information on soil constraints--developed by IFPRI and Colombia University--is based on the Fertility Capability Classification (Sanchez et al., 2003). Thus far, the Atlas contains only the FPP candidate sites, the crop data, the soil constraints data, and an irrigation map (see Thenkabail et al., 2006 for details on the irrigation map). Other layers will be added in a future version of the atlas.



### ***The Overlay analysis***

The FPP candidate sites were characterised at three levels – the site itself, a circular buffer of 50 km radius and another of 100 km radius. Variables include the crop area data and soil constraints. Additional variables can be added in the future, depending on recommendations of the phenotyping network team.

The buffer maps were made using Robinson projection and then reconverted back to latitude and longitude in order to reflect the area of the buffers at different latitudes. An overlay algorithm generated descriptive statistics for each of the FPP sites.

### **References**

- FAO, 1978. Report on the agro-ecological zones project. Vol 1: Results for Africa. World Soil Resources Report 48/1. FAO, Rome, 158pp.
- FAO, 2005. New\_LocClim: Local Climate Estimator, Environment and Natural Resources, Working paper No. 20 (CD-ROM)
- Sanchez, P. A., C. A. Palm, and S. W. Buol. 2003. Fertility capability soil classification: a tool to help assess soil quality in the tropics. *Geoderma* 114:157-185.
- Thenkabail, P. S. C. M. Biradar, et al. 2006. An irrigated area map of the world (1999) derived from remote sensing. IWMI Research Report # 105. Colombo: International Water Management Institute (IWMI).
- UNEP, 1997. World Atlas of desertification, 2nd edition, edited by N. J. Middleton and D. S. G. Thomas, Arnold, London.
- You, L. & Wood, S., 2006. An entropy approach to spatial disaggregation of agricultural production, *Agricultural Systems* 90:329-347.

## **68. G4008.41: Application and validation of the major QTL phosphate uptake 1 (*Pup1*)**

***January 2008–December 2009***

### **Principal Investigator**

Sigrid Heuer, IRRI, sheuer@cgiar.org  
Plant Breeding, Genetics and Biotechnology Division, International Rice Research Institute (IRRI), DAPO Box 7777, Metro Manila, Philippines  
Phone: +63 (2) 580 5600 -2761 (office) 2415 (home); Fax: +63 (2) 580 5699

### **Collaborating institutions and scientists**

- IRRI: Abdelbagi Ismail
- JIRCAS: Matthias Wissuwa
- ICABIOGRAD: Masdiar Bustamam and Joko Prasetyono

The *Pup1* locus, associated with high P-uptake efficiency, was originally identified in the tolerant *aus*-type variety Kasalath based on a screening of thirty rice varieties in a highly P-fixing soil in Japan under upland conditions. Within the GCP project “Revitalizing marginal lands: discovery of genes for tolerance of saline and phosphorus-deficient soils to enhance and sustain productivity”, the *Pup1* locus was fine-mapped and delineated to a 272 kb on Chr. 12. In order to assess the function of *Pup1*, extensive physiological and phenological analyses were conducted and the genes present in the corresponding region in the Nipponbare reference genome were assessed. Since none of these analyses revealed processes or genes that can be directly related to P uptake, the *Pup1* locus was

sequenced in the tolerant donor Kasalath to identify the tolerance gene(s) and gain insight into the function of *Pup1*. Comparative sequence analyses of the *Pup1* regions of Kasalath (*aus*), Nipponbare (*japonica*) and 93-11 (*indica*) revealed that *Pup1* is a hot-spot of transposon integration with a complex genetic structure. The large number of transposon-related elements (>50%) in conjunction with several inversions and deletions led to the truncation of genes, generation of novel genes, and to the integration of segments from other chromosomes. Consequently, the validation of the sixty-eight genes predicted in the Kasalath *Pup1* region required detailed analyses. In agreement with previous data, none of the short listed candidate genes can be directly related to P-uptake processes suggesting that *Pup1* probably improves P uptake via a novel mechanism. Two *Pup1* candidate genes are now being assessed in detail. Gene #46n codes for a putative protein kinase that might be part of a P-sensing and signaling pathway regulating plant responses to P (and drought) stress. Interestingly, this gene is conserved in the wild rice *Oryza rufipogon* and the African rice *O. glaberrima* but absent from the Nipponbare and 93-11 genomes. A germplasm survey additionally showed that gene #46n is present in many breeding lines developed for drought prone environments suggesting that breeders have unknowingly selected for *Pup1*. The other gene, #21n, codes for a dirigent-like gene putatively involved in the lignin biosynthesis pathway and/or scavenging of reactive oxygen species. Both processes are known to be important under drought and other stresses. Preliminary data at IRRI indeed showed a P- and drought induced increase in root lignification which was significantly less in *Pup1* NILs.

The above findings are directly relevant for the application of *Pup1*: (i) The detailed sequence information facilitates the development of highly specific diagnostic markers and repetitive sequences can be avoided. Furthermore, it is now possible to develop markers that target *Pup1* specific genes (e.g., gene #46n) and specific alleles (e.g., gene #21n). (ii) A germplasm survey with *Pup1* specific markers revealed that the locus is already present in many drought breeding lines providing further evidence for the beneficial effect of *Pup1*. With the availability of *Pup1* foreground markers, breeders now have the tool to specifically select for it. The germplasm survey also revealed a group of accessions with partial presence of *Pup1*. Based on the detailed molecular data and gene-specific markers, it is now possible to select those recipient parents from the latter group that lack regions putatively important for tolerance. Based on this survey, we have selected two irrigated rice varieties (IR64 and IR74) and three Indonesian upland varieties (Dodokan, Situ Bangenit, Batur) that are lacking the *Pup1* locus, for the development of breeding populations using MAB. (iii) The absence of known P-uptake related genes in *Pup1* and the finding that *Pup1* improves performance under drought suggest that this locus can be beneficial in drought prone environments. This might be directly related to the higher root growth rate of *Pup1* lines observed under stress. However, whether sustained root growth improves access to water and thereby P, or whether higher P uptake facilitates root growth remains to be clarified. Within the ongoing project we will therefore complete the development of *Pup1* introgression lines and test them in different P-deficient and drought-prone environments. A screening of diverse germplasm with contrasting *Pup1* haplotype under irrigated and upland conditions is ongoing at IRRI.

## **Subprogramme 4: Bioinformatics and crop information systems**

### **69. G4005.22: Development of Generation CP domain models and ontology**

*January 2005–December 2008*

#### **Principal Investigator**

Richard Bruskiewich, IRRI; r.bruskiewich@cgiar.org

International Rice Research Institute DAPO Box 7777, Metro Manila, The Philippines; +63-2-580-5600, Ext. 2872; Fax: +63-2-580-5699

#### **Collaborating institutions and scientists**

- IRRI: Thomas Metz; Martin Senger; Graham McLaren
- Bioversity: Elizabeth Arnaud; Tom Hazekamp; Adriana Alercia
- CIMMYT: Rosemary Shrestha; Guy Davenport
- CIP: Reinhard Simon
- ICRISAT: Jayashree Balaji

#### **Context**

This GCP SP4 commissioned research project was set up in 2005 to develop a shared public platform-independent set of scientific domain models, ontology, and data templates to crosslink all data types and analysis processes within the GCP software platform – middleware and network. Research and development so far has resulted in a mature “object oriented” domain model, heavily parameterised with ontology – dictionaries of defined domain concept terms, interrelated with each other (Bruskiewich *et al.* 2006). This domain model and ontology is being applied to implementations of a software framework of networked databases and tools (Bruskiewich *et. al.* 2008).

#### **Findings and implications**

In 2008, the following activities and outcomes were envisioned.

- **Domain Model Maintenance:** Maintain and validate GCP domain models by stress testing the models practical application to GCP platform development and data curation of sample target data sets.
- **Ontology Management embedded in the GCP Platform:** collaborate with “Crop Registry & Data Templates Development” and “Platform Development” tasks to incorporate pertinent end user access to GCP ontology dictionaries for data annotation, ontology term nomination and data queries.
- **GCP Ontology Curation:** develop general and crop-specific GCP ontology with involvement of GCP and external scientific experts, in particular, for crop-specific ontology relating to plant anatomy, developmental stage, trait and phenotype ontology, for selected priority GCP crops.
- **Validation of GCP Ontology:** by application to selected representative crop-specific GCP data sets.

- **Disseminate Knowledge about GCP Domain Model & Ontology:** by the project web site tutorial workshops and scientific publication.

Project progress toward these objectives to date as included the following:

- CIMMYT-hosted ontology postdoctoral fellow (Rosemary Shrestha) was recruited in January 2008 to guide both general and local (maize, wheat) ontology development.
- Descriptive and tutorial documentation for the GCP domain models and ontology project has been established at (<http://pantheon.generationcp.org> and project outputs published at <http://cropforge.org/projects/gcpontology/>). Use of a “Web 2.0” framework will promote further community engagement.
- Crop-specific ontology dictionaries are being elaborated and will be published in an online GCP database, for direct query by GCP platform tools and network protocols.
- Representative GCP project datasets are being annotated with ontology.
- The postdoctoral project liaison scientist, Dr. Rosemary Shrestha, visited the Plant Ontology Consortium team at Cornell University (<http://www.plantontology.org>) to coordinate GCP efforts in plant, trait and phenotype ontology such that they remain fully compliant with and leverage international public efforts.

GCP scientists will generally benefit from the resulting workbench of tools and data sources cross-linked using GCP semantic standards. GCP scientists and beneficiaries will also benefit from the enhanced accuracy and completeness of research documentation of experimental results using ontology standards.

#### **Next steps and challenges**

Validation and application of the GCP domain model and ontology remains the primary challenge. The full impact and value of these semantic standards has not yet been felt by the GCP research community. The project team hopes that as these standards become more fully applied to GCP platform and network implementations, and to the annotation of GCP data sets, that the full benefits of semantic integration will be achieved.

#### **References**

- Bruskiewich R, Senger M, Davenport G, Ruiz M, Rouard M, Hazekamp T, Takeya M, Doi K, Satoh K, Costa M, Simon R, Balaji J, Akintunde A, Mauleon T, Wanchana S, Shah T, Anacleto M, Portugal A, Ulat VJ, Thongjuea S, Braak K, Ritter S, Dereeper A, Skofic M, Rojas E, Martins N, Pappas G, Alamban T, Almodiel R, Barboza LH, Detras J, Manansala K, Mendoza MJ, Morales J, Peralta B, Valerio R, Zhang Y, Gregorio S, Hermocilla J, Echavez M, Yap JM, Farmer A, Schiltz G, Lee J, Casstevens T, Jaiswal P, Meintjes A, Wilkinson M, Good B, Wagner J, Morris J, Marshall D, Collins A, Kikuchi S, Metz T, McLaren G, and van Hintum T. 2008. “The Generation Challenge Programme Platform: Semantic Standards and Workbench for Crop Science,” *International Journal of Plant Genomics*, vol. 2008, Article ID 369601, 6 pages. doi:10.1155/2008/369601
- Bruskiewich R, Davenport G, Hazekamp T, Metz T, Ruiz M, Simon R, Takeya M, Lee J, Senger M, McLaren, G and van Hintum T. 2006. The Generation Challenge Programme (GCP)-Standards for Crop Data. *OMICS* 10(2):215-219

## **70. G4005.23: Implementation of web services technology in the Generation Challenge Programme Consortium**

*January 2008–December 2008*

### **Principal Investigator**

Milko A. Skofic, Bioversity International; m.skofic@cgiar.org

Via dei Tre Denari, 472/a, 00057 Maccarese (RM) – ITALY, Tel: +39 06 6118286

### **Collaborating institutions and scientists**

- Rice Research Institute: Martin Senger
- CIMMYT: Guy Davenport
- Bioversity International: Elisabeth Arnaud, Michael MacKay, Mathieu Rouard

### **1. Objective**

This project, rather than generating scientific data, is a service to the other sub-programmes: its goal is to enable access and sharing of scientific data by scientists.

The implementation of Web Service technology, complying with Generation CP data standards, will create a virtual network in which scientists from all over the world can access and share data from geographically distant locations.

Providing data access to researchers and computer analysis programmes will ease and speed up the research activities in the Generation Challenge Program.

### **2. Methodology**

The first concern of this project is to provide a bridge between researchers that produce and use data with the project “*Development of tools and technology to increase the functionality of the GCP information platform*”, which provides the framework for analysis and display tools.

The second concern is to provide user friendly instruments which can be used with minimal training by the people responsible of managing databases and datasets: this ensures that the GCP virtual network can be established in an efficient and sustainable way, maintaining the existing roles and responsibilities of the staff involved in scientific activities.

The final concern is to apply these technologies to the “*Management of the GCP Central Registry and the creation and maintenance of templates for data storage in repositories*”, so that all relevant datasets, either in database or dataset format, can be shared.

### **3 Background**

In the first phase of the project technology from GBIF, PyWrapper, has been selected because it provided a configuration user interface that enabled non-specialists to wrap databases to the Generation CP Domain Models.

The BioMOBY Web Services were selected by the Generation CP as the principal technology; efforts were made to provide PyWrapper with native BioMOBY support. Unfortunately, this solution proved not optimal and not sustainable because of insufficient flexibility and reliability in wrapping databases, and because the original developers were no more available to improve the software.

#### **4 Status report**

Work has undergone to develop a new solution in-house, specifically tailored to the GCP environment. In May, in collaboration with the project “*Development of tools and technology to increase the functionality of the GCP information platform*”, a series of prototype web services have been developed which open the road to a more generalised solution.

- By the end of July, four web services covering the Germplasm Passport domain will be established at the GCP Central Registry site, providing data covering the EURISCO and SINGER catalogues.
- In the following month the low-level software implementation will be generalised and optimised, to constitute the framework of the new mapping tool, MyWrapper.
- The GCP Template prototype parser, developed in April, will be integrated in this tool to provide the ability to serve as Web Services datasets provided as GCP templates.
- A web based user interface will be developed to create a package that can be distributed and installed at data provider sites and managed by local staff.
- Documentation on the usage of the tool and concrete examples of web service enabling of databases will constitute a knowledge resource that will provide sustainability to this solution.
- As new GCP Domain Models will be developed, the tool will be adapted to manage different kinds of data, providing access to a greater variety of information.

### **71. G4005.27: High performance computing facilities for the GCP platform**

*January 2005–December 2009*

#### **Principal Investigator**

Anthony Collins, CIP, a.collins@cgiar.org

1895 Avenida La Universidad, La Molina, Lima 12, Peru; (51) 1-3496017; Fax: (51) 1-3175268

#### **Collaborating institutions and scientists**

- CIP: Reinhard Simon
- ICRISAT: Jayashree B; Dave Hoisington; Rajeev Varshney
- IRRI: Richard Bruskievich; Ramil Mauleon
- NIAS: Shoshi Kikuchi; Koji Doi

The primary goal of this project is to provide high performance computing facilities for the GCP platform, where success is measured by impact on subprogrammes 1, 2 and 3, and reflected by a user community including more GCP collaborators beyond CGIAR.

Therefore the HPC support and maintenance program focuses on this goal to maximise use of HPC facilities, as globally supported by CIP, together with the bioinformatics support teams at each of CIP, ICRISAT, and IRRI working with NIAS. Significant SP1, 2 and 3 use case examples will be highlighted at the ARM in 2008.

As the HPC hardware funded by GCP is approaching the limit of processing capacity, a key new goal is to review and test sustainability options for the GCP Grid beyond 2008 with external Grid collaborators. Ongoing performance, user and load monitoring from CIP will enable a profile of future requirements to be defined.

Thus the primary outputs of this HPC task in 2008/9 will be

**CIP: reports targeting**

1. Usage and impact of the HPCs for GCP SPs
2. Collaborators identified for Grid computing capacity expansion experiments in 2008, leading to a review of sustainability options for the GCP Grid beyond 2008

**CIP, ICRISAT, IRRI, NIAS**

Highlights of specific application development presented as use cases in ARM 2008

**NOTE:** as the timeline for the outputs of this project year runs from June 2008 through May 2009, only preliminary outputs for these goals will be presented at the ARM 2008.

**72. G4006.08: Data analysis support for existing projects in SP2 with emphasis on integrating results from microarray and mapping experiments**

*August 2006–December 2008*

**Principal Investigator**

Guy Davenport, CIMMYT, g.davenport@cgiar.org

Centro Internacional Mejoramiento de Maíz y Trigo, CIMMYT Apdo. Postal 6-641, 06600 México, D.F.Mexico; Tel: +52(55) 5804-2004 Fax: +52(55) 5804-7558

**Collaborating institutions and scientists**

- CIMMYT: Jose Crossa; Trushar Shah
- IRRI: Richard Bruskiewich; Hei Leung; Ramil Mauleon
- ICRISAT Jayashree B; Rajeev Varshney
- CIP: Simon Reinhard
- NIAS Shoshi Kikuchi
- JIC: Andreas Magusin

**Context**

The major goal of this project is to further elucidate genes, alleles, mechanisms and other factors relating to abiotic and biotic stress response across multiple crops. This will be achieved through the analysis of available crop gene expression, genomic sequence, phenotype, genotype and QTL mapping data sets from GCP SP2 commissioned and competitive research projects.

### **Findings and implications**

Extensions to the maxd software for supporting genomic expression analysis have been made to allow this software to be fully Plant-MIAME compliant and to take advantage of the ontologies in development by the GCP. Other extensions include those for the EASE software for gene enrichment in gene expression analysis and MAPMAN for displaying gene expression data in relation to pathways. These plant specific extensions will allow GCP researchers to use these tools, which were not developed originally for crop species. The RiCES: Rice Cis-Element Searcher - RiCES is a *cis*-element searching tool optimised for rice genome and MANOVA software is available on the HPC (<http://hpc.irri.cgiar.org/>). QTL data, publicly available sequence data, microarray data, and genetic and physical maps have been integrated using CMAP, a comparative map viewer. Data analysis and management support has also been given to several other GCP projects. For example, *in silico* mapping, and haplotype analyses for the GCP#13 project and microarray analysis and curation for the GCP#15 project have been carried out. The results of these analyses will be available for GCP researchers to use.

### **Links**

Links to software and documentation developed and used by this project can be found on the GCP bioinformatics portal (<http://www.generationcp.org/bioinformatics>).

### **Next steps and/or challenges**

Much of the software envisioned by this project is dependant on the availability of a full genomic sequence and therefore has been developed for and used with rice data. However, the release of the draft maize sequence now means that these tools can be applied to maize. Software is in development for analysis of EST and Next Generation Sequencing (NGS) data, including software for SNP feature identification, haplotype prediction and integrated marker mining. Software will also be developed for quality control of microarray and automatically generating reports. Curation and integration of GCP QTL and gene expression datasets is ongoing.

## **73. G4006.16: Integrated GCP Informatics Platform**

*January 2008–December 2008*

### **Principal Investigator**

Graham McLaren; IRRI-CIMMYT Crop Research Informatics laboratory,  
DAPO Box 7777 Manila Philippines; [g.mclaren@cgiar.org](mailto:g.mclaren@cgiar.org)  
Tel (63) 2 8450563; Fax: (63) 2 5805699

### **Collaborating institutions and scientists**

- Agropolis–CIRAD: Manuel Ruiz; Pierre Larmande; Xavier Argout; Gautier Sarah
- CIMMYT: Guy Davenport; Luis Avila
- IRRI: Richard Bruskiewich; Martin Senger
- ICRISAT: Jayashree B; CT Hash
- EMBRAPA: Marcos Costa
- NCGR: Andrew Farmer



## Context

A key problem of biological scientists in general and GCP scientists in particular is integration of diverse and dispersed data sources and analysis of data via diverse analytical tools. The GCP Informatics Platform seeks to alleviate this problem by providing an informatics platform which allows data integration via an agreed domain model and a workbench of interoperable applications. The domain model and basic architecture are in place and the current stage of the Platform project is to implement biological use cases designed to facilitate analysis of genetic diversity, functional genomics and molecular breeding.

In 2008, the GCP Informatics platform task will implement the following use cases:

- SP1 Use Case: Develop a query, visualisation and analysis workbench for genetic diversity studies.
- SP2 Use Case: Develop a query, visualisation and analysis workbench for comparative functional genomics research.
- SP3 Use Case: Develop a query, visualisation and analysis tool for marker assisted breeding programmes.

The project will also support both GCP and non-GCP scientists in using the platform for the above use cases through training and documentation, and will continue to promote the GCP platform within and outside of the GCP, by providing adequate documentation and support to allow developers to integrate their data sources and applications.

## Progress

**1. General Platform Infrastructure:** GCP data source query specifications are being systematically developed and documented for all three use cases. The Koios web search workbench is being refined for application to all the use cases. GCP standards have been applied to the ICIS database schemata to serve GCP defined germplasm, passport, trait and study query use cases. General and specialist Koios web query and data visualisation interfaces were also refined this year.

### **2. Query, visualisation and analysis workbench for genetic diversity studies**

The Java GCP Datasources developed for the TropGene and ICIS databases have been made completely interoperable and a common Web portal for the SP1 and SP2 use cases has been defined. The Following new functionalities have been added to the GenDiversity application: filtering by passport data, computation of allelic frequencies, computation of LD, and checking for duplicates.

### **3. Query, visualisation and analysis workbench for functional genomics research**

This use case provides a simple workbench interface to query and cross-link germplasm and gene data, including gene annotation linked to phylogenomics (G4008.21, Matthieu Conte), gene expression, lists of germplasm with associated passport, trait and study data, and QTL mapping data. The interface will use GCP Koios technology in a Joomla! portal and the Taverna work flow engine (<http://taverna.sourceforge.net>) (G4007.12, Martin Senger) which will also cross-link with GenDiversity.

#### **4. Query, visualisation and analysis tool for marker assisted breeding programmes**

The Molecular Breeding Design Tool (MBDT), a component of the MAB information system is a stand-alone application to drive visual integration of data (genotype, marker, LG map and QTL data) to allow users to design a desired genotype and select promising parents. Progress so far includes being able to upload matrices of genotype data, create shortlists of potential parents, upload LG maps and display them horizontally, visualise allele data for short-listed individuals below the linkage map with color coded alleles.

The Molecular Selection Tool (MOSEL) is a stand-alone plug-in for GenoMedium which will obtain and display pedigree, molecular and phenotypic data for breeding lines through the GCP data source and facilitate selection decisions. GenoMedium has been upgraded to use the latest version of the ICIS data source and we are replicating the Koios functionality within GenoMedium to give a standalone query tool for breeders who do not have direct connection to the internet.

**Tangible outputs:** Documented on Pantheon web site (<http://pantheon.generationcp.org>). Enhanced GenDiversity interface using platform consistent portal technology connection to TropGene, Repository and ICIS data sources. New functionality for analysis of genetic diversity including integration of SNP data and linkage to R and Haploview Tools. Platform portal using Joomla! and integrating Koios query technology to give web based access to diverse and distributed data sources. Prototype functional genomics workbench within the portal linking gene expression results with gene annotation, phylogenomic and QTL data. Cross linkages to GenDiversity for genetic diversity applications. First prototype of MBDT operating on a stand-alone data source. The latest version of GenoMedium with diverse query capability and prototype molecular selection tool is available at <http://www.genomedium.org>.

#### **References**

Richard Bruskiwich, et al. (2008) The Generation Challenge Programme Platform: Semantic Standards and Workbench for Crop Science. *Int'l Journal Plant Genomics*, Article ID 369601, 6 pages, doi:10.1155/2008/369601.

### **74. G4006.17: GCP quality management and data quality Improvement**

#### **Principal Investigator**

Thomas Metz, IRRI

#### **Collaborating institutions and scientists**

- WUR-CGN: Theo Van Hintum
- ICRISAT: B Jayashree
- CIP: Reinhard Simon
- Agropolis-CIRAD: Claire Billot

This project is a continuation of the 2007 project Data Quality Improvement and Assurance. In 2008, the project incorporates the previous project GCP Software Engineering and Collaboration Platforms as an objective.

The project addresses the following issues that have strong implications on data quality and/or quality management in the GenerationCP:

- Providing support to institutions that consider the adoption and adaptation of the ICRISAT LIMS system. This is a continuation of a similar activity in 2007. The 2008 activities and outputs will focus on enhancing the LIMS functionality, development of instructional and training materials for easy adoption, and building of a community of users and developers.
- Development of a toolkit that consists of data quality indicators, best practice manuals, and a customised set of database/informatics and statistical tools applied to the main dataset types of the GCP. This toolkit will allow the routine quality assessment of GCP datasets. The major outputs in 2008 will be definition of data quality indicators, development of a data resolution methodology, and application of this methodology to the molecular marker data sets in the GCP Central Repository.
- Maintaining and supporting the collaboration systems CropForge and CGPWiki. This activity is a continuation of the former project GCP Software Engineering and Collaboration Platforms. The main activities in 2008 are systems maintenance and user support and training.
- Development of a white paper on Requirements for GCP Projects Producing Primary Data. The white paper will allow GCP management to specify service level agreements for data-producing projects.

## **75. G4006.35: Statistical support for the design and data analysis of GCP projects**

*January 2008–December 2008*

### **Principal Investigator**

Marco Bink, Biometris, WUR; marco.bink@wur.nl

P.O. Box 100, 6700 AC, Wageningen, The Netherlands; (31) 317-481072; Fax: (31) 317- 483554

### **Collaborating institutions and scientists**

- CIMMYT: José Crossa
- Agropolis–CIRAD: Xavier Perrier, Jean-Francois Rami
- WUR: Fred van Eeuwijk, Marcos Malosetti, Hans Jansen

### **1. Context and overview of the project**

The objective is to support scientists from all SP's to design, curate, and analyze the generated genotypic and phenotypic data in an optimal way, identifying relevant QTLs with appropriate and tailored statistical procedures. This project involves also expert scientists from CIRAD (diversity analysis) and CIMMYT (experimental design). The project will contain consultancy, communication and training components.

### **2. Activities**

#### **2.1 Workshop GSS – September 2007**

A one-week workshop was held in Zaragoza, Spain from 24<sup>th</sup> to 28<sup>th</sup> September 2007. The workshop was attended by seven researchers from Bolivia, Colombia, Brasil, Chile,

Nigeria, Tanzania and The Philippines, and assisted by three staff members of Biometris. The hands-on workshop started with participants presenting their research question and data, and was followed by the discussion about suitable data analysis methods on a person-to-person basis and in general group discussions. Significant progress in the analysis of the different data sets was achieved, in some cases resulting in highly advanced manuscript drafts.

## 2.2 Consultation to SP scientists

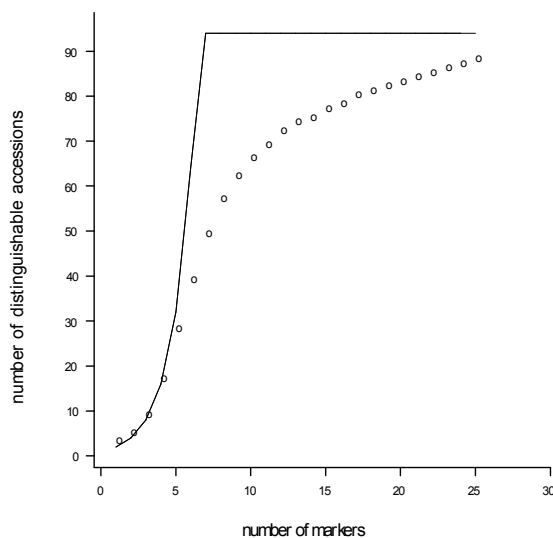
Elisa Mihovilovich, a researcher from CIP approached the Helpdesk with regard to an association mapping case in potato. The consultation related to optimal methods to test for the association of SSR markers potentially tagging candidate genes for late blight resistance.

While visiting Mexico in April 2008, two staff members of Biometris met with dr. Crossa (CIMMYT) to discuss strategies to provide consultancy on experimental design issues in SP projects.

## 2.3 GSS support

As follow-up of the 2007 GSS Zaragoza workshop, statistical support is given to the publication of a paper on the use of DArT markers for studying the genetic composition of the Musa genome. Many Musa varieties are interspecific hybrids between *M. Balbisiana* and *M. acuminata* [Collaborators: H. Gomez (GCP-GSS), E. Sales (USM-The Philippines), A. Vilarinhos(Embrapa-Brazil), J. Jansen (WUR-Netherlands)].

Following ideas developed during the 2007 GSS Zaragoza workshop, work is underway to develop procedures for selecting minimum sets of binary (DArT) markers for identification of accessions. The search for selecting minimum sets of binary markers is combined with the search for specific binary identification keys as used in taxonomical applications.



**Figure 1 Relationship between the number of distinguishable (groups) of accessions and the number of markers in the optimum identification set. Data: Musa - A. Vilarinhos (Embrapa-Brazil).**

In preparation of the 2009 GSS Cali workshop, collaboration has started between H. Gomez (GCP-GSS), WUR statisticians and several researchers: Dr. M. Shehabu (Ethiopia) on enset/musa, Dr. A. Alvez on cassave and Mr. C. Lung'áho on potato.

#### **2.4 Helpdesk website**

Information about Linkage Disequilibrium and Association Analysis is currently available on a website entitle Germplasm Data Analysis (<http://gda.biometris.wur.nl/>). After consultation with M.C. de Vicente (GCP) and Th. van Hintum (GCP) the website will be redesigned in order to facilitate researchers to enter the various topics at different levels (from introductions to advanced analyses).

#### **3 Tangible outputs**

- [1] Workshop, Zaragoza (Sp), Fall 2007
- [2] Website (<http://gda.biometris.wur.nl/>)
- [3] GSS consultation

### **76. G4007.09: Design and analysis of marker-trait association studies, with special attention for genetically challenging crops January 2008–December 2008**

#### **Principal Investigator**

Fred van Eeuwijk, Biometris, WUR; fred.vaneeuwijk@wur.nl  
P.O. Box 100, 6700 AC, Wageningen, The Netherlands; (31) 317-482902; Fax:  
(31) 317- 483554

#### **Collaborating institutions and scientists**

- University of Hohenheim, Germany: Hans Peter Piepho; Albrecht Melchinger
- Imperial College London, UK: David Balding
- NIAB: Ian Mackay; Wayne Powell
- SCRI/BIOSS: Christine Hackett; Dave Marshall
- Leiden University Medical Center, The Netherlands: Hans van Houwelingen; Jeanine Houwing-Duistermaat
- WUR: Marcos Malosetti; João Paulo; Marco Bink; Hans Jansen

#### **1. Motivation and back ground**

A first step in any marker assisted breeding strategy is the localisation of quantitative trait loci (QTLs). Since the 1990s, the standard methodology for the detection of QTLs in crops is based on a linkage analysis of offspring populations created from crossing two inbred parents. Although successful, weak points of such linkage analyses are: i) making crosses is time consuming, ii) crosses often are not representative of breeder's germplasm, iii) QTLs can be located at a relatively low precision.

A recent attractive alternative to pure linkage based QTL mapping is linkage disequilibrium (LD) mapping, or association mapping. LD approaches can be applied to any pool of selected or arbitrarily structured genotypes, allowing breeders to search for QTLs in relevant genetic back grounds. As LD methods assay the accumulated generative history in the germplasm / population under study, they are often more

powerful and precise than standard QTL mapping approaches. LD approaches are appealing within the Generation Challenge Programme (GCP) where inventories of genetic diversity are being made on the basis of molecular markers with the purpose of investigating that genetic diversity in relation to phenotypic variation.

While successful methodology for LD mapping has been proposed for major crops, for smaller crops and genetically more challenging crops, little knowledge and experience is available. For major crops, mixed models are a popular vehicle for LD mapping as they provide various ways to control for spurious associations caused by population structure, i.e., the phenomenon that the whole of the set of genotypes under study falls apart in genetically different groups with group specific allele frequencies. Also for other crops than major crops, mixed models seem a proper choice for LD mapping, but then the mixed models need to be attuned to the requirements of the specific crop.

The current project aims at defining a statistical protocol for the design and analysis of LD strategies in a variety of crop species of importance to the GCP. Design theory for association studies in smaller and genetically complex crops, like polyploids, requires study of the genetic mechanisms causing LD and a proper translation of those mechanisms in statistical parameters. For example, to quantify LD decay with genetic distance on the chromosome in polyploids, first relevant measures for LD need to be defined. This project will bundle the insights of specialists in LD mapping theory to arrive at statistical protocols for conducting LD feasibility studies in crops relevant to the GCP. Such feasibility studies should answer questions on the choice of marker system, marker density, and the type of population in relation to defined phenotypic traits.

For the analysis part of this project, we propose to adapt and develop special purpose mixed model strategies focusing on the genetic properties of small and challenging crops. Mixed models are highly suitable for modeling genotype by environment interaction in multi-environment data, data obtained from germplasm evaluations across multiple trials and stress gradients. In the context of the GCP work on stress tolerance, the modeling of genotype by environment interaction has high priority. Mixed models also have good facilities for representing relationships between genotypes, a feature that facilitates correction for pedigree and population structure in LD studies.

The statistical protocols we develop on design and analysis of LD studies in small and challenging crops will be accompanied by documented software and course material that should open up this methodology to the whole of the GCP.

## **2. Results of first research activities**

Within the project, the first major activity consisted in the assessment of the state of the art in LD mapping in general. Subsequently, the specific requirements for smaller and genetically challenging crops needed to be defined, based on the earlier general assessment. (Later, tasks with respect to special developments will be allocated and carried out.) To these purposes, a workshop entitled “Statistical methods for linkage disequilibrium analysis” was organised in Wageningen at the 19<sup>th</sup> and 20<sup>th</sup> of June, 2008 (see [www.biometris.nl](http://www.biometris.nl)). At the first day of the meeting presentations were given by researchers involved in the project. The target was to cover as good as possible all aspects

of importance in LD mapping and elicit discussion on where existing LD methodology requires extension or adaptation for smaller and genetically challenging crops within the GCP mandate. The first day attracted around 80 scientists from different institutes and organisations within, especially, Europe. The second day was used for a more in-depth discussion between the participating scientist in the project on the major points to be included in a document describing a protocol for design and analysis of LD studies, both paying attention to the state of the art in the larger crops and the required extensions and adaptations in the smaller crops. The expectation is to submit this protocol for publication later this year.

## **77. G4007.10: Support to GCP scientists regarding issues related to bioinformatics and data handling**

*January 2007–December 2008*

### **Principal Investigator**

Theo van Hintum, CGN–WUR, The Netherlands

Bornsesteeg 65, 6708PD Wageningen; theo.vanhintum@wur.nl, tel. (31)317-480913

### **Collaborating institutions and scientists**

- CGN–WUR: Elisabeth van Strien

### **Background**

Within SP4, the Subprogramme of the GCP dealing with bioinformatics and crop information systems, a wide array of facilities, expertise and products has been created. The array of product ranges from white papers, data models, software packages, templates, manuals, algorithms, scientific papers, to collaboration platforms and helpdesks. Many of these products were accompanied by websites, elaborating the activities behind it and offering access to the software, model, facility or whatever had been produced. However, from a user perspective this is not very accessible. If a user has a question (s)he will not know where to find the answer in the ‘jungle’ of SP4 products, or other resources available elsewhere in the world. Therefore an easy access to the available facilities, expertise and products needed to be created. Furthermore, a mechanism needed to be established that allows GCP management, and especially the SP4 leader, to monitor the needs in the GCP scientific user community. To address these issues the ‘SP4 helpdesk’ has been established, a one-stop-shop to support GCP scientists regarding issues related to bioinformatics and data handling.

The GCP-SP4 helpdesk will be the entry point for any GCP scientist who has questions regarding handling, storing, or analyzing GCP data. The helpdesk is responsible for creating transparency in the available expertise and resources in the field of biometry, bioinformatics, and software engineering relevant to GCP scientists, available in the GCP. It will pro-actively improve (or advise on the improvement of) GCP web-sites, create an expert network and act as a point of reference for GCP scientists.

## **Results**

The project started January 2008, by defining the rules of the game: a protocol was written, outlining how questions from researchers to the helpdesk would be handled. To make the helpdesk known to the GCP community a new webpage, outlining the goal, intended audience and contact possibilities was designed and published on the GCP internet pages ('The SP4 helpdesk on bioinformatics and biometrics', <http://www.generationcp.org/bioinformatics.php?da=08106902>, or Google on 'SP4 helpdesk'), an email account ([sp4helpdesk@generationcp.org](mailto:sp4helpdesk@generationcp.org)) was registered, and a brief contribution to the GCP eNewsletter was written announcing the launch of the helpdesk.

To create an overview of the existing expertise in the network, current and past SP4 PIs have been approached and an expert database, obviously only for internal use, was created. This database will allow easy identification of experts that could be approached in answering the questions received via the helpdesk.

When this first infrastructure was established, a start was made in documenting the SP4 products and rearranging the information on the SP4 webpages, in order to increase both the visibility and the accessibility of facilities and products, generated over the last years in the Generation Challenge Programme's Subprogramme 4.

## **Next steps**

The bioinformatics portal on the GCP website will continue to be improved, expanded and updated.

The attempts to capture the SP4 products (identify, describe, and make them accessible) will continue. Obviously these products must have reached a certain stage of maturity in terms of accessibility, stability, validation and documentation. Often it will require a small step from the SP4 scientist to push the product from the development stage to the production stage. These products will be promoted via the SP4 bioinformatics portal.

The SP4 helpdesk will not replace other GCP helpdesks, it will however help (potential) users find their way to these expertise specific helpdesks such as those giving support on IP issues, the use of data templates or experimental design.

Finally, the helpdesk will continue to increase its visibility, and thus bring the SP4 facilities, products and expertise to the attention of the GCP crop scientists and breeders. Feedback and suggestions are much appreciated.

We invite you to use SP4 helpdesk, when you think we might be of help!



## **78. G4007.11: Further development and support for use of iMAS by NARS and other user communities**

*January 2008–December 2008*

### **Principal Investigator**

Dave Hoisington, DDG Research, ICRISAT  
Patancheru 502324, Andhra Pradesh, India; d.hoisington@cgiar.org  
Tel 91 40 3071 3071; Fax: 91 40 3071 3074

### **Collaborating institutions and scientists**

- ICRISAT: Tom Hash, Jayashree B
- IRRI: Richard Bruskiewich
- CIMMYT: Guy Davenport

### **Context**

The iMAS system integrates a number of open source quality computing tools for marker aided selection and breeding. The system helps free the user from the manual input data files preparation process and provides the user with simple to use online decision guidelines to correctly interpret and use outputs from the different computing tools. The beta version of the software was released at the ARM in South Africa in September 2007. During the course of testing workshops, user feedback indicated further development of the software to include multi environment QTL analyses, comparative QTL mapping through integration of CMTV, linkage with Qu-gene for modeling MABC and a means of integrating the application with the GCP platform. The system consists of six modules – Data Validation, Phenotyping, Linkage Map building, QTL analysis, Genome Display and MABC sample size.

### **Progress**

During the course of the year 2008, we redesigned a number of user interfaces within the data validation, phenotyping, Linkage map building and QTL analysis modules. In the data validation module we have added new functionality in terms of validation for multiple locations for phenotypic data and multiple populations for mapping data. The phenotyping module has been extended to test for multi-environment phenotyping data analysis in IRRISTAT. In the linkage map and QTL analysis modules, a number of GUI changes have been implemented to make things easier and more logical. The user interfaces allows the upload of multiple data files/sheets for the Gmendel and PlabQTL software, selective viewing and printing options to enhance ease of use.

Earlier problems relating to installation of the iMAS executable have now been fixed and the software is now being distributed as an easy to install .msi (microsoft installer file). The software has been used in two training courses for breeders in April and will be used again for two more courses in October this year. All of those who used the software have provided vital feedback for the improvement of the software.

### **Next steps**

Towards the integration of CMTV with iMAS, iMAS now exports linkage map outputs in formats that are compatible for viewing with CMTV. This facility has to be extended to QTL maps also; this we hope will allow the CMTV tool to be distributed with iMAS as a single package.

The integration of iMAS with the GCP platform has been initiated. As early steps towards the process, iMAS input requirements have been reviewed in terms of domain model data objects and available GCP data source use cases have been studied to check if use case outputs match iMAS input requirements. A use case is currently in development to pull out iMAS genotypic data input files from existing databases (like ICIS at first, and later ICRIS). The integration with the GCP platform is expected to be accomplished through tying up the iMAS application with the Koios query engine.

### References

Holloway JL and Knapp SJ, 1994. GMendel reference manual.

## **79. G4007.12: Development of tools and technology to increase the functionality of the GCP Information Platform** *February 2008–December 2008*

### Principal Investigator

Martin Senger, IRRI-CRIL, m.senger@cgiar.org

### Collaborating institutions and scientists

- IRRI-CRIL: Graham McLaren, Richard Bruskiwich
- Bioversity International: Milko Skovic

### Context

The GCP Platform is a set of collaborating software tools constructed using shared GCP-developed semantic and informatic standards. This project helps to manage and to enhance the software development. These efforts include the continued development of the core framework for GCP platform and specific implementations of GCP-compliant platformsoftware tools, internet protocols and data resource wrappers.

### Findings and implications

Findings: The GCP Platform infrastructure is well designed. The GCP Domain Model covers correctly the areas of interest for GCP users.

### Products

- 19 formalised use cases for searching and retrieving GCP data;
- Validator for GCP software components (data sources);
- Software libraries allowing access to GCP data;
- Well updated API documentation and several Tutorials for software developers;
- Released version of the Soaplab, a toolkit for creating web service on top of existing analysis tools

**Benefactors:** GCP software developers (and via them, indirectly, all GCP software end-users).

### Next steps and/or challenges

- To create definitions of biologically-relevant workflows representing useful pipelines of software tools. Use these definitions in the software tool Taverna, including its web interface.
- To enhance software tools (GCP data sources) that are accessing data in the ICIS database.

## 80. G4008.20: Management of the GCP Central Registry and the creation and maintenance of templates for data storage in repositories

May 2008– December 2008

### Principal Investigator

Elizabeth Arnaud, Bioversity International, e.arnaud@cgiar.org

### Collaborating institutions and scientists

- CIMMYT: Guy Davenport
- CRIL–CIMMYT: Genevieve Mae Aquino
- Bioversity International: Milko Skofic, Michael MacKay

### 1. Objective

The project aims at increasing the availability of Generation CP data templates; enhancing the data validation results; extending the collection of resources managed by the Central Registry (CR) and improving access to the data. The project, temporarily suspended, resumed in May but was revised to fit in 8 months.

### 2. Status report on use of GCP data templates and data sets validation

A table<sup>1</sup> summarizing the CR content was produced. The project “G4008.31- Upgrading the quality and utility of GCP phenotyping data through the development of a data input template to facilitate the storage of data in crop-specific databases”, produced new templates available in GCP website in July.

**Table 1: Present status of the templates and their use on the Central Registry**

Status	GCP Template	version	# of files in CR
to be released for test	General Template	2.0 RC2*	-
	GCP Passport Template	2.0 RC2	-
	SSR Genotyping Template	2.0 RC2	-
	Mapping Template	2.0 RC2	-
	QTL Template	2.0 RC2	-
	Phenotyping Template	2.0 RC2	-
	DArT Genotyping Template	2.0 RC2	-
	SNP Genotyping Template	2.0 RC2	-
	ECO-TILLING Genotyping Template	2.0 RC2	-
	Gene Expression (Microarray) Template	2.0 RC2	-
EST Template	2.0 RC2	-	
Available for test	General Template	1.0 RC1	-
	GCP Passport Template	1.2 RC1	2
	SSR Genotyping Template	1.1 RC1	13
	Mapping Template	1.0 RC1	-
	QTL Template	1.0 RC1	-
	Phenotyping Template	1.0 RC1	-
Approved	EURISCO Passport Template	1.0	-
	GCP Passport Template	1.1	10
	SSR Genotyping Template	1.0	18

\*RC= Release Candidate

The format validation of existing datasets started using GenoMedium, developed in project “G4005.25 - Creation and maintenance of templates for GenerationCP data storage in repositories”. Validation progress has been monitored using CRIL JIRA<sup>2</sup>

Data sets will be corrected with data providers using reported validation errors. Presently, 51 files posted in the Central Registry do not use the templates.

**Table 2. Validation status of datasets in the Central Registry**

Validation Status	Number of Files
Validation Not Started	13 (22%)
Validation Failed	3 (5%)
Validation Successful	3 (5%)
Correction In Progress	2 (3%)
Feedback Required	18 (31%)
Validator Error	20 (34%)

An online survey will help gauge the extent to which researchers used or will use the templates, establish if they satisfy the format of their data sets and assess the use made of Central registry and helpdesk.

### 3. Development of the Central Registry and access to data sets

Web services developed in the project “G4005.23 -Implementation of Web Services technology in the Generation CP Consortium”, once validated, will be applied to the data sets posted on the Central registry.

### 4. Outputs of the project

The GCP CR<sup>3</sup> provides a facility where GCP datasets can be registered. In one year, uploaded files increased from 45 to 95 and registered data sets from 106 to 152. Once the new templates are fully validated, their use will be promoted via the proactive helpdesk.

### 5. Quality data uploading in the Central Registry

The status report on the availability, quality and use of the data templates will be available on GCP site. A live helpdesk will be held in the next ARM to assist PIs in correcting their data and uploading their file.

SP1 Data sets are downloaded for genetic diversity analysis and the high quality of the data sets is a fundamental prerequisite. The current double checking of SP1 genotyping data will bring results to be loaded in the CR. SSR data sets posted in the CR are being studied using a new statistic method called the Data Resolution (Hintum TLJ van, 2007). Data sets provided as valid templates will be made available via web-services.

### References

- Hintum, Th.J.L. van, 2007. Data resolution: a jackknife procedure for determining the consistency of molecular marker datasets. *Theoretical and Applied Genetics* 115: 3, 343-349
- Hazekamp T., Davenport G., Gaiji S., Sood R., Skofic M., 2007- Poster-Management of the Generation CP Central Registry, GCP ARM, Bioversity International
- Hazekamp T.\* and Davenport G. 2007- Poster-Generation Challenge Programme: Making data accessible, GCP ARM, \*Bioversity International, CIMMYT

<sup>1</sup> <http://spreadsheets.google.com/ccc?key=p38Ked4R4my3mP6tM6u3jBQ&hl=en>

<sup>2</sup> <http://cril.cimmyt.org/jira/browse/GCPDATASETS>

<sup>3</sup> <http://gpcpr.grinfo.net>.

## 81. G4008.21: Large scale phylogenomic analyses to gene function prediction for GCP crops

January 2008–December 2009

### Principal Investigator

Mathieu Rouard, Bioersity International; m.rouard@cgiar.org  
 Parc Scientifique Agropolis II 34397 Montpellier - Cedex 5 - France;  
 Tel: (+33) 4-67612908; Fax: (+33) 4-67613034

### Collaborating institutions and scientists

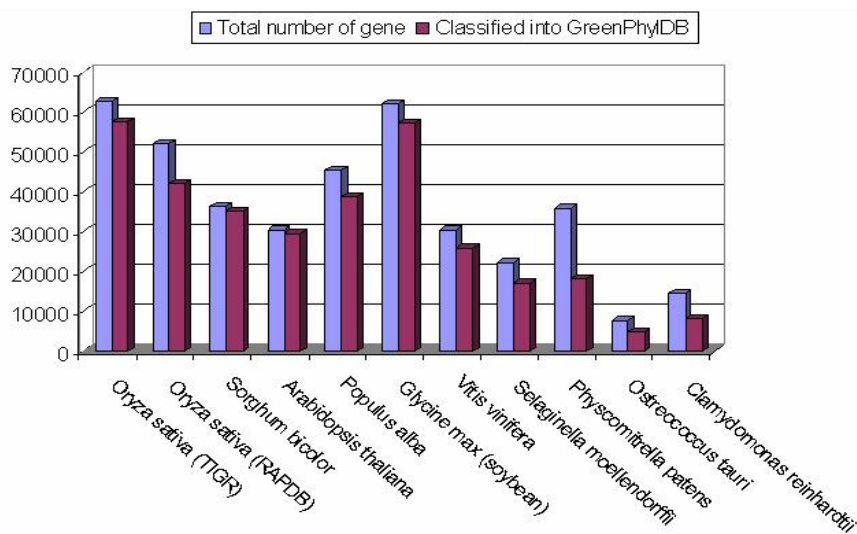
- Agropolis–CIRAD: Christophe Périn
- IRRI: Matthieu Conte; Martin Senger; Richard Bruskiewich

### 1. Research activities and progresses at Bioersity and Cirad

Clustering of 10 plants' genomes was performed to consolidate and improve the previously-computed clustering on the 2 model plants currently featured in GreenPhylDB (<http://greenphyl.cines.fr>). By integrating genomes based on a broad taxonomy (from algae to angiosperms), we expect to define consistent and comprehensive set of homeomorphic plant families<sup>1</sup>

#### 1.1. Genome Clustering

The database currently contains 21000 clusters, among which 2500 clusters were created by the integration of full genomes (cf. figure 1). Lists of gene families were created and some already show a degree of specificity (e.g. monocotyledon- or dicotyledon-specific)



**Figure 1:** Total number of sequences (in blue) and number of sequences allotted to a clusters by species. This diagram shows that the clustering was efficient with the exception of *Physcomitrella patens*.

#### 1.2. Protein pattern search and statistics

Protein signature has become very useful for the classification of protein sequences and for inferring their function. Therefore, we scanned the protein sequences of the 10 genomes (~400,000) against protein signature databases using **Interpro** (<http://www.ebi.ac.uk/interpro/>) motifs (IPR) to define their domain architecture and to provide statistics about clusters.

**Table 1: Proportion of clusters at the first level of clustering<sup>2</sup> (6963 clusters) sharing a common domain architecture (i.e. present at least in 60% of sequences).**

Clusters	Percentage
with IPR specific motif (or combination)	28.7%
without IPR specific motif (or combination)	36%
without IPR	35.3%

### 1.3. Guidelines for protein family annotation

As far as we know, there is no clear standard to annotate protein families. We decided to follow some definitions established by the **PIRSF classification system** (<http://pir.georgetown.edu/pirwww/about/doc/PIRSF.pdf>) proposed by PIR and the **Protein naming guidelines** (<http://www.expasy.org/cgi-bin/lists?nameprot.txt>) proposed by Uniprot. We are currently writing guidelines explaining the annotation steps that annotators will need to follow in order to properly name unannotated clusters.

### 1.4. Web interface for data curators

A web interface has been designed to display relevant information (Statistics, Domain shuffling, cross-references, publications etc.). This is intended to facilitate the manual curation of the gene families. Anyone is free to visit these pages, although in order to participate in the cluster curation it is necessary to log-in.

## 2. Research activities and progresses at IRRI

### 2.1. Pipeline

In order to cope with the increase of sequences contained in clusters, a more efficient and optimised pipeline for phylogenomic analyses has been designed and is being implemented at the **CINES** (<http://www.cines.fr/?lang=en>) supercomputer center.

### 2.2. Website

The website runs now the latest version of the ATV software to display phylogenetic trees.

## References

- Conte MG, Gaillard S, Droc G, Périn C (2008). Phylogenomics of plant genomes: a methodology for genome-wide searches for orthologs in plants. *BMC Genomics*, 2008, 183
- Conte MG, Gaillard S, Lanau N, Rouard M, Périn C (2008). GreenPhylDB: a database for plant comparative genomics. *Nucleic Acids Research*. 2008 January; 36 D991–D998
- Wanchana S, Thongjuea S, Ulat VJ, Anacleto M, Mauleon R, Conte M, Rouard M, Ruiz M, Krishnamurthy N, Sjolander K, van Hintum T and Bruskiewich RM (2008). The Generation Challenge Programme comparative plant stress-responsive gene catalogue. *Nucleic Acids Research*; 2008 January; 36: D943–D946

<sup>1</sup>Homeomorphic families consist of proteins that are both homologous (evolved from a common ancestor) and homeomorphic (sharing full-length sequence similarity and a common domain architecture)

<sup>2</sup>with the least stringent parameters

**82. G4008.22: Methodology development for reconstruction of Genealogies based on Haplotypes related to geographic patterns (HaploPhyle: graphical haplotype network in the light of external data)**

*January 2008–December 2008*

**Principal Investigator**

Claire Billot, Agropolis–CIRAD, [claire.billot@cirad.fr](mailto:claire.billot@cirad.fr)

TA A96/03, Avenue d'Agropolis, 34398 Montpellier Cedex 5, France

**Collaborating institutions and scientists**

- Agropolis–CIRAD: Xavier Perrier, Manuel Ruiz, Jean-François Rami
- CIP: Reinhard Simon

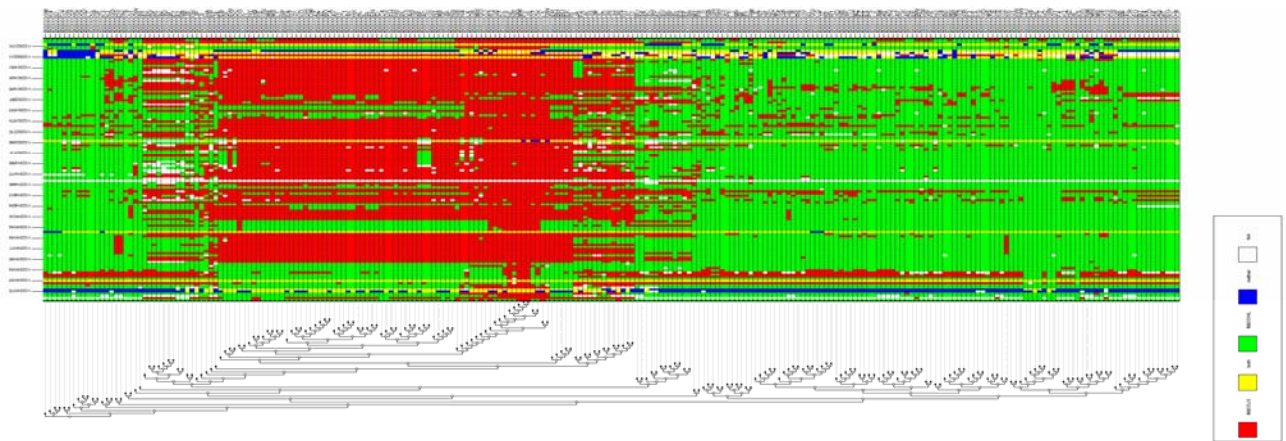
Genetic diversity assessment through genotyping is gaining much resolution and currently produces now large quantity of DNA sequences or chains of tightly linked markers. The close linkage between the polymorphisms observed is likely to leave the patterns little affected by recombination. Therefore the patterns evolve less quickly (than with unlinked or loosely linked makers) and better reflect past situations. In the particular context of crops, the past comprises a major event, domestication, which generally involved strong genetic bottlenecks. Pattern analysis of polymorphisms can lead to identification of predominant haplotypes, and inference of ancestral haplotypes vs recombinant haplotypes. These can generally be organised into networks and series of derivations from ancient to recent, both through mutation and recombination. One illustration, of immediate practical analytical use, is the realisation that the structure of polymorphisms into haplotypes is an important feature when using them test for association with complex traits. Instead of focussing on one marker at a time, it has been suggested that haplotypes defined as regions of strong inter-marker linkage disequilibrium (LD) (i.e., haplotype blocks), would be more powerful at detecting the role of a given genomic region.

On a more global line, haplotype analysis is likely to enable breakthroughs in crop germplasm analysis. Haplotype networks enable polarisation in time and investigation of history. The major food crops are those crops that were very successful and expanded widely, generally throughout the world. The geographic patterns of diversity are another result of a process in time. Altogether, analyzing haplotype networks and relating them to geographic patterns allows development of phylogeographic analyses and has a great power for resolving crop domestication and understanding further crop adaptation.

Quick and user-friendly methods are required for conducting these analyses, especially in the context of the GCP, which deals with worldwide diversity of numerous crops with a wealth of genotypic information.

In the frame of this project, progress has been made in different directions for now:

- Looking for existing haplotyping software that can be used with Linux, with command line. Three softwares have been chosen for now: Phase (with the need to verify the possibility of using it freely), Ishape and Gevalt.
- Looking for software able to built a network from haplotypes, freely available, under Linux and with command lines. Nothing found yet... Thus analysis of the algorithms used to built a network: Median Joining Network and Reduced Median Network in order to built ourselves a package.
- Looking for software able graphically represent the network, freely available, under Linux and with command lines. The choice would be GraphViz, but there is no possibility for the user to work with the graphic. There is a need to build an API based on API Touchgraph (one of the possibilities).
- Development in R of a function able to combine haplotypes and phylogenic trees (see Figure 1 for an illustration)



**Figure 1: Graphical representation of haplotypes or rice along a chromosome with reference to two accessions (see the research development of project G4006.03: SNP analysis and the genetic diversity along the rice genome (HaplOryza)**

**83. G4008.31: Upgrading the quality and utility of GCP phenotyping data through the development of a database template to facilitate the storage of data in a crop specific database**

*February 2008–February 2009*

**Principal Investigator**

Robert Koeber, CropGen International, mockbeggars@gmail.com  
 Mockbeggars, Townhouse Rd, Old Costessey, Norwich, NR8 5BX , UK, +44 (0)  
 1603 745 036; [www.cropgeninternational.com](http://www.cropgeninternational.com)

**Collaborating institutions and scientists**

- CIMMYT: Guy Davenport
- CRIL/IRRI: Warren Vincent E Constantino

The Generation Challenge Program is seeking to establish a phenotypic data storage system, which will also allow for protocol documentation and the definition of the



environmental conditions associated with each experiment. The intention is both to encourage the standardisation of phenotyping in GCP projects (without seeking to restrict the initiative of any individual project team), and to house the resulting data in a way which will facilitate meta-analyses. Our immediate goal is to create a wizard-driven template (“first generation template”) design to capture phenotypic observations and all associated environmental and experimental data required to make them interpretable, whilst assuring compatibility with the GCP domain models and crop information systems such as ICIS. This template is directed at GCP projects which will deliver data during 2008. We will follow this up by developing a “second generation template”, which we anticipate will be more crop-specific and prescriptive, and will incorporate mandatory traits and fields (including drought tolerance indicator traits, experimental designs, environmental indicators etc.). It will be accompanied by a user manual. We seek also to be able to export, as far as possible, the data presently lodged in the GCP Central Registry into the ‘first generation’ template. PI returns of the templates will be quality checked by the project team to ensure compliance. Finally, we wish to explore the possibility of establishing electronic field data capture technology for the GCP community, as a tool to improve the accuracy of phenotyping. The project has three components. The bulk of the activity in template development and the planning phase for electronic data capture will take place during calendar year 2008; some of the compliance activity, and much of the anticipated implementation of data capture are expected to take place during calendar year 2009. An outline of the (“first generation template”) is available in poster and/or pamphlet form.

---

## ***Subprogramme 5: Capacity-building and enabling delivery***

### **84. G4005.53: The use of molecular markers in efficient crop improvement: Marker-assisted breeding learning module**

*August 2007–July 2008*

#### **Principal Invtesigator**

Theresa Fulton, IGD, Tf12@cornell.edu

Institute for Genomic Diversity, Cornell University, 130C Biotechnology Building, Ithaca, NY, 14853 USA, Phone:1-607-255-4323, Fax: 1-607-254-6379

An IPGRI-IGD collaboration in 2004 of a learning module CD on molecular marker technologies was very favorably received, generating hundreds of requests from more than 30 countries worldwide (de Vicente and Fulton, 2004). However, while this module describes markers and their uses, it stops short of describing how they are used specifically in crop improvement and marker-assisted breeding. This next step will be the focus of this new learning module.

This learning module will also be a good complement to the other modules in progress and supported by the GCP, including those on molecular markers in plant diversity, genomics, phenotyping, and bioinformatics, creating a “bookshelf” of material available for scientists world-wide.

The topics in this module will include (but not be limited to) genetic linkage, mapping, QTL identification, population types and development, targeted introgression, positive and negative marker-assisted selection, development of near-isogenic lines, etc. Real-life applications will be given as examples, along with limitations and considerations.

The targeted audience is plant breeders in developing countries; the modules will be developed in such a way as to be useful either as a self-tutorial, or as the basis of a training course. Users will be apprised that the Interactive Resource Center & Helpdesk can answer any follow-up questions they may have.

The slides are in the final stages of finishing and then will be sent out for review.

### **85. G4005.63: Interactive Resource Center & Helpdesk**

*January 2005–July 2009*

#### **Principal Invtesigator**

Theresa Fulton, IGD, Tf12@cornell.edu

Institute for Genomic Diversity, Cornell University, 130C Biotechnology Building, Ithaca, NY, 14853 USA, Phone:1-607-255-4323, Fax: 1-607-254-6379

The Interactive Resource Center & Helpdesk (<http://irc.igd.cornell.edu/>) was developed in 2005 by the Cornell Institute for Genomic Diversity as a support tool for scientists worldwide, with a particular focus on those implementing molecular marker assisted plant breeding and plant genetic diversity assessment programmes.

The IRC now includes a large number of resources, including protocols, tutorials, lists of funding opportunities, learning modules, literature, general resources (such as information on writing proposals), and news. Freely available data is also available for download. A contact list of sorghum researchers was added last year (others underway). Also posted are key links, including journals, the African Molecular Marker Network, and GCP resources. A ‘helpdesk’ answers scientists’ questions, with the help of a volunteer team of scientists from various fields (specializing in molecular markers, population genetics, plant breeding, genetic diversity, etc.).

Since inception, the number of visitors to the site has increased every year, with more than 2,500 unique visitors to the site in the first part of 2008. Site visitors have been from Senegal, Ghana, South Africa, Kenya, Ethiopia, Cote D’ivoire, Uganda, Qatar, Iran, Egypt, Turkey, Thailand, Malaysia, Japan, China, Georgia, India, Pakistan, Europe and the Americas. Recently added is a list of sources of laboratory supplies (with links to regional representatives), FAQ, info about the Helpdesk Team, and a DNA extraction troubleshooting page. Recent questions to the Helpdesk have included queries about DNA extractions, software use (such as Mapmaker and QGene), and data analysis.

Upcoming additions to the IRC will include several highly requested items: additional contact lists (millet and cassava), results of the recent survey, new tutorials and protocols and more.



**Figure 1. A map of the locations of the 500 most recent visitors to the Resource Center.**

## 86. G4006.13: Targeting and impact analysis of Generation Challenge Programme technologies

January 2007–December 2008

### Principal Investigator

Glenn Hyman, CIAT, g.hyman@cgiar.org

International Center for Tropical Agriculture (CIAT); AA 6713, Cali, Colombia

(57-2) 445-0000; Fax: (57-2) 445-0073

### Collaborating institutions and scientists

- CIAT: Sam Fujisaka, Peter Jones
- CIMMYT: John Dixon
- IFPRI: Stanley Wood

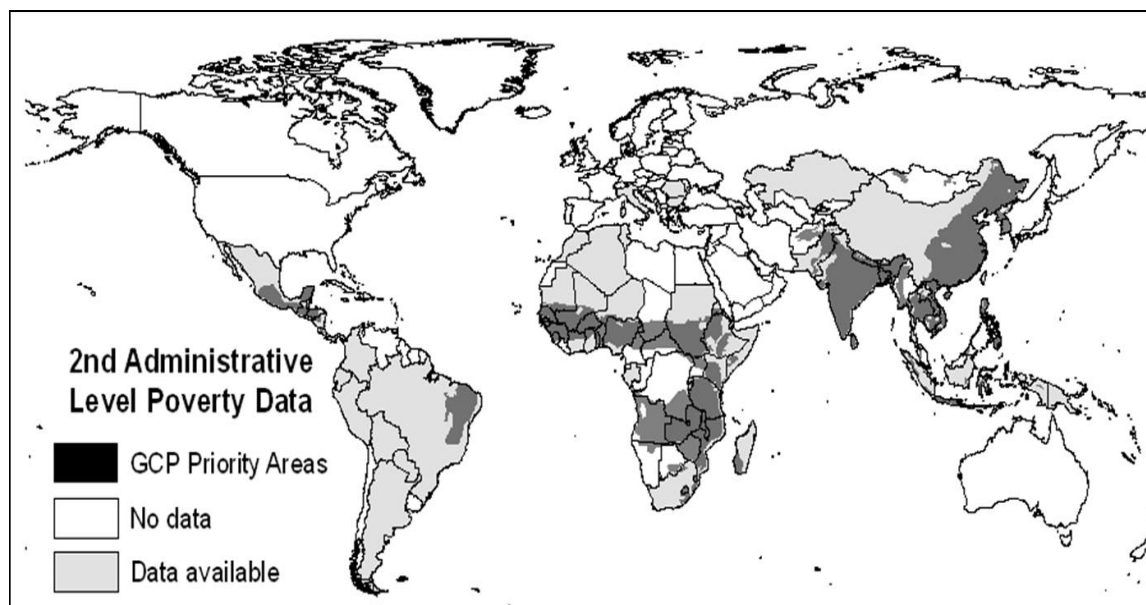
### 1. Project summary

This project prioritizes crops and regions for staple crop improvement that addresses drought problems in marginal environments in developing countries. The research identifies the spatial coincidence of the poor, the production staple crops on which the poor depend, and drought-prone production environments.

### 2. Activities and progress

#### 1) Geographic distribution and the magnitude of poverty in 14 GCP priority farming systems.

Our initial research suggested 14 farming systems should be given priority in crop improvement research for the poor in marginal environments (Hyman et al., 2008). An inventory of poverty maps was carried out for the 14 priority systems. For each system, a poverty profile was written that included infant mortality, stunted and underweight children and the proportion of the population below the poverty line. The profiles took into account the inconsistency of some of the data sets across countries.

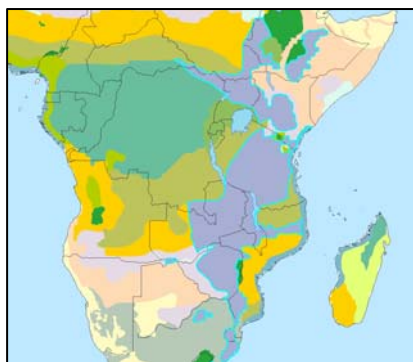


### 2) Crop-specific modeling of drought severity and type

The failed season modeling approach is being extended to study crop-specific drought (Hyman et al., 2008). Input parameters were taken from the DSSAT family of crop models. The detail of the modeling will vary because some of the GCP crops have no (e.g., sweet potato) or less well developed (e.g., cassava) DSSAT models. We are initiating work to develop computer algorithms to calculate drought susceptibility indices.

### 3) Analysis of GCP priority farming systems and their patterns of technology adoption at disaggregated levels within each region

We collected data sets to examine technology adoption and constraints to adoption at the level of farming system. Included are the adoption study of the CGIAR system-wide panel on impact assessment (SPIA), and data on infrastructure for agricultural research and development. A key task of this activity is to examine which farming systems occur in each country and estimate the cross-country differences in capacity to adopt technology. For example, the map and table below shows the 12 countries that share the maize-mixed farming system in East Africa. Brief assessments of four farming systems were carried out and written into a draft report. The systems are Mesoamerican maize-beans, Southeast Asia lowland rice, South Asia rainfed mixed and South Asia rice-wheat. These farming system profiles include suggestions on which countries within the system present opportunities and/or constraints for successful delivery of GCP technology.



COUNTRY	SYSTEM	REGION	BEAN_HA	CASS_HA	MAIZ_HA	RICE_HA	POP_TOTAL	POP_RUR	POP_URB	STUNT	PREV
Central African Republic	Maize mixed	SSA	0	0	0	0	16,646	1,038	15,608	2,161	38
Ethiopia	Maize mixed	SSA	3,008	2	48,452	1	8,336,471	7,841,053	696,418	1,223,742	46
Kenya	Maize mixed	SSA	0	34,007	1,130,941	8,436	23,307,068	17,814,058	5,492,993	859,591	38
Malawi	Maize mixed	SSA	123,407	144,914	1,154,800	31,807	7,639,565	6,656,960	982,596	690,579	47
Mozambique	Maize mixed	SSA	30,843	276,214	530,940	5,466	2,371,927	2,140,448	231,479	296,102	54
Namibia	Maize mixed	SSA	40	0	5,481	0	57,502	37,940	19,562	1,682	42
South Africa	Maize mixed	SSA	10,417	0	40,927	829	9,267,621	5,402,134	3,865,487	159,127	24
Swaziland	Maize mixed	SSA	4,665	17	53,445	51	894,062	678,241	206,821	49,698	29
Tanzania, United Republic of	Maize mixed	SSA	67,713	335,759	1,056,918	261,512	15,330,089	10,984,012	4,346,086	1,296,633	47
Uganda	Maize mixed	SSA	455,387	291,057	454,521	42,538	9,391,764	8,893,310	498,452	795,642	27
Zaire	Maize mixed	SSA	186	1,428	524	2,305	3,456,538	1,838,302	1,657,236	116,710	38
Zambia	Maize mixed	SSA	100	87,209	415,327	3,921	7,391,041	3,935,157	3,445,894	539,205	51
Zimbabwe	Maize mixed	SSA	48,043	29,885	1,184,655	209	8,073,255	6,212,585	1,860,670	319,674	28
TOTAL			743,749	1,200,582	6,077,941	357,075	95,562,559	72,225,228	23,337,292	6,309,634	

### 4) Economic assessment of the potential market-scale impacts of a variety of household-level adoption scenarios

Economic assessment is being conducted by the HarvestChoice project funded to IFPRI by the Gates Foundation. That project is developing a global data set and crop simulation models that will permit us to estimate the impact of specific GCP technologies.

### References

Hyman G, Fujisaka S, Jones P, Wood S, de Vicente S and Dixon J (2008). Strategic approaches to targeting technology generation: Assessing the coincidence of poverty and drought-prone crop production. *Agricultural Systems*. 98:50-61.

**87. G4006.14: *Ex Ante* impact analysis of marker-assisted selection technologies supported by the Generation Challenge Programme (GCP)**

*December 2006–December 2008*

**Principal Investigator**

George W Norton, Virginia Tech; gnorton@vt.edu

Department of Agricultural and Applied Economics, 205B Hutcheson Hall,  
Blacksburg, VA, USA (1) 231 7731; Fax: (1) 540 231 7417

**Collaborating institutions and scientists**

- Virginia Tech, USA: Jeffrey Alwang
- IRRI: Abdelbagi Ismail
- CIAT: Martin Fregene

**1. Context**

Ex ante impact analysis is being used to estimate benefits of GCP investments and to validate an approach to impact assessment that might be used broadly in the GCP to document progress. Two GCP projects: “Revitalizing marginal lands: Discovery of Genes for Tolerance of Saline and Phosphorus Deficient Soils to Enhance and Sustain Productivity,” and “Development of Low-Cost Technologies for Pyramiding Useful Genes From Wild Relatives of Cassava into Elite Progenitors” were chosen as cases studies for the impact assessment. Total economic benefits have been projected with and without the new technologies (traits). The benefits consider (a) area planted to crops currently affected by target stresses, projected changes in area under cultivation, and production of the crops in specific countries, (b) the nature of the markets for the crops, (c) projected yield and cost changes due to the new technologies, (d) estimated time for discovery, development, and deployment of the DNA marker technologies and associated germplasm, (e) estimated time required to breed, test and disseminate superior new cultivars, including rates of adoption by farmers, and (f) the discount rate for benefits and costs that occur in the future. Three graduate students have been engaged in the project. One completed her MS thesis in June 2008, a second is scheduled to complete in July 2008, and a third in December 2008. The third student is developing and applying methods to assess gender impacts of improved cassava varieties in Nigeria.

**2. Findings and implications**

Marker-assisted breeding (MAB) in rice is estimated to save at least 3-6 years compared to conventional breeding (CB) and result in significant incremental benefits in the range of \$50 to \$500 million depending on the country, abiotic stress, and lag for CB under base assumptions (table 1). Benefits almost double (\$930 million for Bangladesh) if the 3 year time advantage for MAB over CB is assumed to be 6 years. For cassava, benefits for MAB to incorporate resistance to cassava mosaic disease, green mites, white flies, and post harvest deterioration vary from \$34 to \$817 million depending on the country.

**Table 1. Estimated economic benefits of marker-assisted breeding (MAB) for salinity and P-deficiency tolerance as compared to current varieties and compared to conventional breeding (CB) for the same traits (base assumptions and 3 year advantage of MAB over CB)**

Country and constraint	Net present value <sup>1</sup> (\$million) (marker-assisted breeding)	Net present value <sup>1</sup> (\$million) (conventional breeding)	Incremental benefits of MAB over CB (\$million)
<b>Salinity</b>			
Philippines	226.9	177.7	49.2
Bangladesh	3,666.4	3,167.5	498.9
India	6,847.9	6,400.9	447.0
Indonesia	895.8	701.8	194.0
<b>P-deficiency</b>			
Indonesia	2,069.7	1,788.1	281.6

1. NPV at 5% discount rate 2. Source: Alpuerto, 2008

**Table 2. Estimated economic benefits of marker-assisted breeding (MAB) for resistance to mosaic disease, green mites, and whiteflies, and comparison of MAB with conventional breeding (CB) (base assumptions)**

Country	Net present value <sup>1</sup> (\$million) (marker-assisted breeding)	Net present value <sup>1</sup> (\$million) (conventional breeding)	Incremental benefits of MAB over CB (\$million)
Nigeria	1,493.4	676.0	817.4
Ghana	675.9	304.7	371.1
Uganda	52.6	18.5	34.1

1. NPV at 5% discount rate 2. Source: Rudi, 2008

### 3. Next Steps

Finish the analysis of gender impacts of marker-assisted breeding for cassava in Nigeria and continue writing up and publishing results of the project in various formats.

### 4. References

- Alpuerto, V.L (2008). Economic Analysis of Marker Assisted Breeding in Rice. MS thesis, Virginia Tech, June.
- Rudi, N. (2008). An Ex Ante Economic Analysis of Developing Low Cost Technologies for Pyramiding Useful Genes from Wild Relatives into Progenitors of Cassava, MS thesis, Virginia Tech (expected completion, July 2008).

## **88. G4006.36: Capacity-building and research project**

*January 2007–December 2011*

### **Principal Investigator**

Mark Laing, UKZN, [laing@ukzn.ac.za](mailto:laing@ukzn.ac.za)

ACCI, Plant Pathology, PBag 1 Scottsville, 3209, RSA, South Africa

### **Collaborating institutions and scientists**

- UKZN
- AGRA

Project G4006.36 was established between the GCP and the ACCI of the University of KwaZulu-Natal in October 2006. The goal was to hire a molecular plant breeder to work in the ACCI to assist in the training of the 80 Plant Breeding PhD students from 12 countries in Africa where the ACCI currently draws its students. As secondary goals, the incumbent would have their own research programme, in particular to develop novel molecular markers for important traits in the food security on which the ACCI works. Furthermore, there would be opportunities to develop joint projects with collaborating scientists, graduates and current students in their home countries. A first round of advertising (which took 6 months to arrange) identified a suitable candidate, but due to excessive delays in his getting a work permit, and the relatively low salaries offered in South Africa (when translated into dollars), the candidate took another job. A subsequent round of advertising has identified a new candidate, who has accepted the position. A work permit is being sought, using the services of a consultancy, Global Migration, to accelerate the process. Generic delays in appointment of staff in South Africa, and UKZN specifically, include: (1) The Research Office at UKZN has very slow evaluation and approval processes of projects, even when they are relatively simple, such as this one with the GCP; (2) South African institutions effectively close down from 1<sup>st</sup> December to 20 January in terms of hiring processes, losing 6-8 weeks; (3) UKZN has complex advertising and hiring processes. This position needed a selection committee that included the Deputy Vice Chancellor, the Dean, the Head of School, the Director of the ACCI, 3 faculty members and a student member. Getting such an unwieldy committee together is extremely difficult; (4) The labour laws in South Africa are complex and xenophobic, delaying the process considerably; e.g., no foreigner may be interviewed until all South African applicants have been reviewed, interviewed and rejected with detailed, documented reasons; (5) If a foreigner is appointed, then a process has to be undertaken to get the person a work permit. Numerous documents are required, including: police clearance from all countries lived in for the past 8 years, accreditation of degrees by SAQA (this is slow due to limited capacity at SAQA), documentation of the appointment process, evidence of Scarce Skills and an appointment contract. However, all the barriers and delays have been overcome and the appointee, Dr Jedidah Danson, starts in the position in August 2008.



**89. G4007.03: The ‘Community of Practices’ concept applied to rice production in the Mekong Region: Quick conversion of popular rice varieties with emphasis on drought, salinity and grain quality improvement**

*January 2007–December 2008*

**Principal Investigator**

Theerayut Toojinda; BIOTEC, theerayut@dna.kps.ku.ac.th  
113 Thailand Science Park, Phahonyothin Road, Klong 1, Klong Luang,  
Pathumthani 12120 THAILAND, Tel: (66-2) 564 6700, Fax : (66-2) 564 6701-5

**Collaborating institutions and scientists**

- RGDU: Jonaliza L. Siangliw
- UBU: Sureeporn Kate-ngam, Watcharapong Wattanakul
- NAFRI: Monthathip Chanphengsay
- CARDI: Men Sarom
- DAR: Toe Aung

The results presented include lines development using MAS which took place 3 times (May 2007, November 2007 and May 2008) in RGDU. The backcross populations were advanced to F2 or F3 and are now ready for trait validations and field testing.

**1. Research activities of DAR and progress in line conversion**

DAR screened 188 BC3F4 lines using salT marker for salinity tolerance and selected 77 lines. Ten lines were crossed to IR53936 and to Pokkali to produce 150 BC4F1. Using the same marker, 58 BC4F1 lines were selected. Fifty families of BC4F1 and BC3F4 were selfed to get 582 BC4F2/BC3F5 and resulted to 303 lines with Pokkali allele at salT marker. Yoshida and pond methods will be used for salinity screening and field testing will be in Kyaukse and Myaungmya provinces. DAR also improved the grain quality of Manawthuka. 200 BC3F1 plants were screened selecting 4 lines. All plants were crossed to Manawthuka producing 200 BC4F1. Two plants were selected and were selfed to get BC4F2 and only 12 were carrying Basmati alleles in 200 F2s. The 12 lines were planted for yield trial. To identify more desirable plant type of Manawthuka, 460 BC4F3 were obtained from the segregating BC4F2 and 45 lines were homozygous. Another set of 41 lines from BC4F3 which were not genotyped were selected using agronomic characters. Both 45 and 41 lines will be in observation trials this wet season.

**2. Research activities of CARDI and progress in line conversion**

CARDI developed aromatic CAR3, which is a drought tolerant line. Ten out of 125 BC2F4 were selected using MAS and five BC2F4 were crossed to CAR3 generating 94 BC3F1 where 4 lines were selected containing the target. From the 4 BC3F1, 484 BC3F2 were produced and screened with BADH, RM587 and RM589 resulting to 15 lines carrying alleles of Phka Rumduol. Field trials and quality evaluations will be conducted in CARDI for seed increase and will be tested in drought prone areas including 8 provinces in Cambodia.

### 3. Research activities of NAFRI and progress in line conversion

NAFRI improved the glutinous rice TDK1 to be aromatic by transferring aroma gene from Homnangnouane (HMN). Seven lines were selected from 201 BC2F3 and BC3F1 were produced. 94 BC3F1 were screened using BADH and 45 lines carrying the HMN allele. Five hundred BC3F2 were produced from 9 BC3F1 and were screened with markers and 26 BC3F2 were carrying HMN allele. The selected BC3F2 will be planted in NAFRI this wet season for seed increase before testing in target location trials.

### 4. Research activities of RGDU and UBN and progress in line conversion

RGDU and UBN work together in developing aromatic IR57514 with good grain qualities, tolerance to submergence and drought and resistance to BLB. From 352 BC3F1 lines, 6 were selected but were not able to cross because the plants were weak. Thus 436 BC3F2 were screened resulting to 21 lines carrying the target QTL/genes and crossed to IR57514. From 96 BC4F1 and 12 lines were selected and 2595 BC3F3 from 21 BC3F2 were also screened and 236 were selected. 236 lines are now screened for submergence and drought tolerance as well as aroma and grain qualities at RGDU. Observation trial will also be conducted at RGDU this season.

**Table 1. Progress in line development by DAR, CARDI, NAFRI and RGDU/UBN using marker-assisted selection at RGDU**

Workshop	Institute	Trait	Generation	No. of lines for MAS	No. of selected lines	Markers
21-30 May 2007	DAR	Salinity tolerance	BC3F4	188	77	salT
		Aroma/Quality	BC3F1	200	4	BADH, RM190, RM204
	CARDI	Aroma and GC	BC2F4	125	10	BADH, RM587, RM589
	NAFRI	Aroma	BC2F3	201	7	BADH, RM21, RM5349
	UBU/RGDU	Aroma/Quality Xa21	BC3F1	352	6	BADH, Waxy SNP2341, SNP2209, PB7/PB8
10 - 24 Nov. 2007	DAR	Salinity tolerance	BC4F1	150	58	salT
		Aroma/Quality	BC4F1	200	2	BADH, RM190, RM204
	CARDI	Aroma and GC	BC3F1	94	4	BADH, RM587, RM589
	NAFRI	Aroma	BC3F1	94	45	BADH
	UBU/RGDU	Aroma/Quality/Xa21	BC3F2	436	21	BADH, Waxy, SNP2341, PB7/PB8
1 -15 May 2008	DAR	Salinity tolerance	BC4F2	170	60	*salT, RM10720
			BC3F5	412	243	RM3412, RM1287, RM1072, RM411 ^Waxy, RM5349, RM457, RM21
		Aroma/Quality	BC4F2	200	12	BADH, RM190, RM204
			BC4F3	460	45	BADH, RM190, RM204
			BC4F3		41	selected only by plant type
	CARDI	Aroma and GC	BC3F2	484	15	BADH, RM587, RM589
	NAFRI	Aroma	BC3F2	500	26	BADH, RM21, RM5349
	UBU/RGDU	Aroma/Quality/Xa21	BC4F1	96	12	BADH, Waxy, SNP2341, PB7/PB8
			BC3F3	2595	236	BADH, Waxy, SNP2341, PB7/PB8

\* - markers used to select for salinity tolerance

^ - markers used to select for quality of IR53936

At present, DAR, NAFRI, CARDI and RGDU/UBN are preparing for trait validations and yield trials to be conducted in and outside the institute. At the end of the year to first quarter of 2009, the data for all trait screening and field trials may be available.

## 90. G4007.13: Capacity-building à la carte 2007

### Principal Investigator

Carmen de Vicente, GCP, [c.devicente@cgiar.org](mailto:c.devicente@cgiar.org)

Apdo. Postal 6-641, 06600 Mexico, D.F., Mexico, Phone: +52 55 5804 2004

The following six teams were selected for 2007's Capacity-building a la carte programme, and continue to carry out activities in 2008:

- *Capacity building for characterising maize for water stress tolerance at KARI-Katumani* (Lead Institution: KARI)
- *Marker-aided development of Nutritionally-enhanced cassava for Nigeria* (Lead Institution: NRCRI)
- *Application of molecular tools for controlled wild introgression into cultivated germplasm in Senegal* (Lead Institution: ISRA)
- *Characterisation of maize germplasm found in Ghana, using the bulking technique* (Lead Institution: CSIR, Ghana)
- *An integrated proteomics and genomics approach to discover salt tolerance genes* (Lead Institution: ABRII)
- *Enhancing capacity of ICABIOGRAD in phenotyping and molecular analysis to develop elite rice lines suitable to Indonesian uplands* (Lead Institution: ICABIOGRAD)

Further details on the activities carried out by these 6 teams follow.

For more details on the concept of Capacity-buidling à la carte, see abstract for project G4008.39: Capacity-building à la carte 2008.

### 90.01 G4007.13.01: Capacity-building à la carte 2007–Capacity building for characterising maize for water stress tolerance at KARI–Katumani

May 2007–July 2009

#### Team Leader

James G Gethi, KARI; [jgethi@wananchi.com](mailto:jgethi@wananchi.com)

Kenya Agricultural Research Institute (KARI), P.O. Box 340-90100 Machakos, Kenya +254 721831166 Fax +254 2 2338320

#### Collaborating Institution and scientists

- Agropolis–INRA: Francois Tardieu; Claude Welcker

### 1. Project activities and progress

The grant was aimed at improving the phenotyping capacity of the KARI-Katumani phenotyping site at Kiboko through improvement of physical facilities and training of the staff involved in the research activities.

A weather station compatible with data collection required by the projects being implemented by Katumani was borrowed from INRA France. This was pending resolution of purchasing procedures of the weather station by KARI.

A printer/copier/fax/scanner was purchased from the funds for the office. Two batteries to run the weather station at Kiboko were purchased and installed.

A visit before planting the experiments was made towards end of November. During this visit, detailed plans for drought screening and characterisation trials were made.

Twenty surface pipes and their accessories were purchased to supplement the irrigation system already present at Katumani-Kiboko. Fourteen sprinklers (14) and their stands were purchased and installed at Kiboko. In addition repairs were made on several hydrants to allow for a larger irrigated field

## 2. Research activities

A total of 60 inbred lines were phenotyped. All materials were treated the same until three leaf stage when the water stress treatment was imposed. Plant growth parameters such as plant height, ear placement, leaf number and yield were taken.

**Table 1. Germplasm phenotyped at KARI-Kiboko**

Group	Type of material	Area collected	No. of lines
1	Recombinant Inbred lines	CIMMYT, KARI and IITA	10
2	Inbred lines	Coastal lowlands	20
3	Inbred lines	Mid altitude drylands	20
4	Recombinant Inbred lines	Eastern and Southern Africa	10

The three groups of materials were all significantly different from each other when subjected to water stress in number of ears harvested, ear length, shelled and field weight and hundred kernel weight.

Leaf emergence, and ligulated leaf counts was higher in well watered conditions compared to stressed conditions in all the groups, and the highest leaf maintenance was obtained from RILs developed from inbred lines.

**Table 2. Traits phenotyped at KARI Kiboko in the dry season, 2008**

TRT	Group	Fwt (g)	Plant ht (cm)	Ear ht (cm)	Tassel (cm)	DTA	DTS	ASI
Stressed	1	0.672	119.1	57.3	27.1	58.2	60.8	2.7
Watered	1	1.120	141.1	69.5	36.1	61.5	62.4	0.9
Stressed	2	0.308	89.5	44.4	25.6	47.7	50.0	2.3
Watered	2	0.549	112.7	55.2	31.4	47.5	49.2	1.7
Stressed	3	0.287	92.5	41.4	28.7	45.8	49.3	3.4
Watered	3	0.388	107.4	46.6	33.4	47.0	48.7	1.8
Stressed	4	0.267	92.8	43.2	27.7	48.1	48.3	0.2
Watered	4	0.533	106.3	48.6	38.6	47.4	49.2	1.8

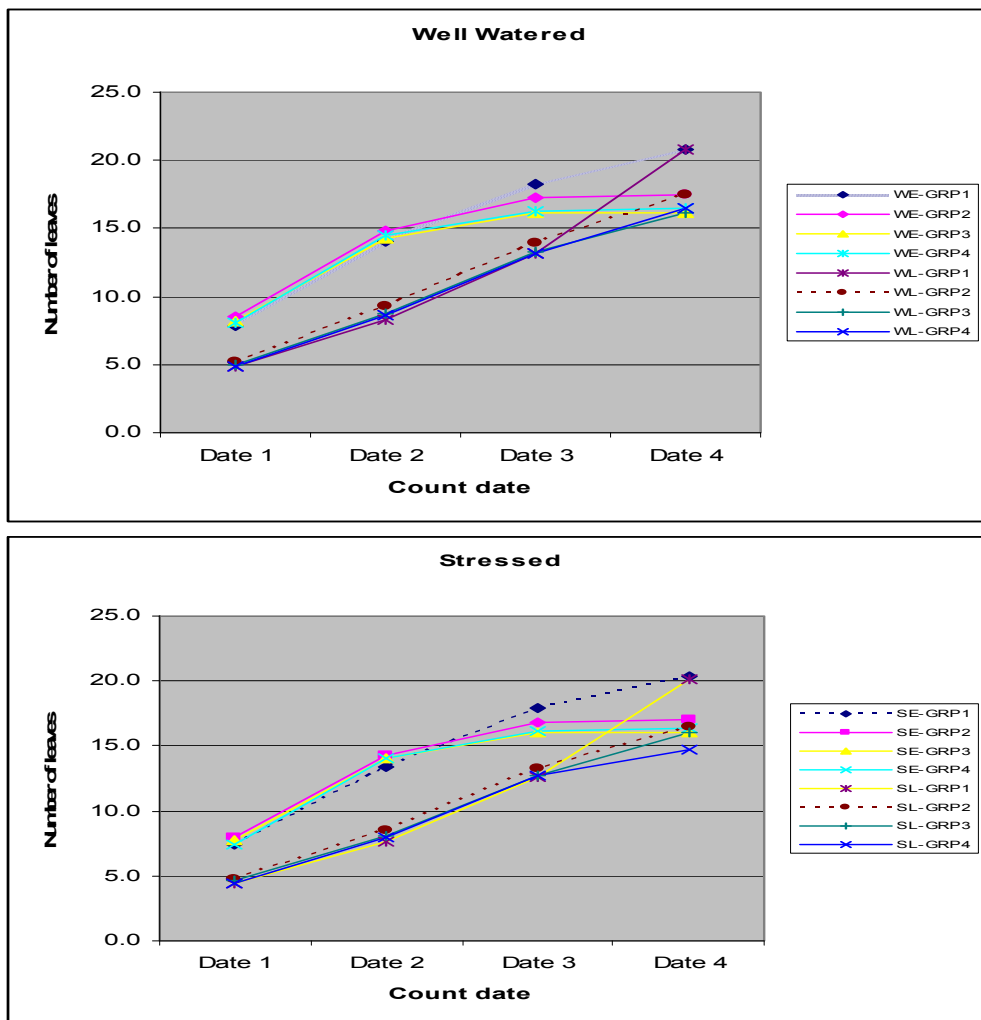


Figure 1. Leaf emergence under stress and well watered conditions

### Tangible outputs delivered

Irrigation facilities, equipment and some training were accomplished. Lines possessing better leaf growth identified for further breeding activities. The data will be used to select lines for further development of drought tolerant varieties.

90.02 G4007.13.02: Capacity-building à la carte 2007–Marker-aided development of nutritionally-enhanced cassava for Nigeria

July 2007–July 2009

### Team Leader

Chiedozie Egesi, NRCRI, cegesi@yahoo.com

National Root Crops Research Institute (NRCRI), Umudike, Nigeria; Tel: +2347034962100

### Collaborating institutions and scientists

- NRCRI: Emmanuel Okogbenin; Ada Mbanaso; Nnamdi Eke-Okoro; Khaya Shuaibu; Oluwakemi Ogundapo; Samuel Baiyeri
- CIAT: Martin Fregene

## 1. Research activities and progress at NRCRI

Two NRCRI research staff visited CIAT's Cassava Genetics Laboratory for training in marker-assisted selection and post-flask management of *in vitro* plants. Nnamdi Eke-Okoro and Khaya Shuaibu both spent one month learning all relevant techniques for molecular breeding with different scientists at CIAT. The training activities included:

- Cassava embryo rescue
- DNA extraction, quantification, dilutions, agarose gel and polymerised chain reaction (PCR)
- Estimation of quality and quantity of DNA extracted
- Cassava micro-propagation
- Hardening of *in vitro* cassava plants in greenhouse
- Mapping/QTL analysis
- Hand pollination for generation of seedlings
- *In vitro* conservation and inventory management
- Germplasm health and quarantine techniques

The enhancement of manpower for molecular breeding in relevant breeding programmes in African NARs is very critical for the application of modern tools in agriculture for food security, poverty alleviation and income generation.

In addition to the on-hand trainings, two NRCRI research fellows are being supported by the project for M.Sc. degrees in Plant Breeding at 2 Nigerian universities: the University of Nigeria, Nsukka and the Michael Okpara University of Agriculture, Umudike. Ms. Kemi Ogundapo and Mr Samuel Baiyeri are carrying out degree – research projects in cassava molecular breeding related to project activities. They have both submitted progress reports. While one of the students is working on “SSR analysis of the Genomic region of the CMD3 Gene in Cassava” the second is working on “Evaluation of Latin American Germplasm for improved adaptation, high yield and resistance to pests and disease in Nigeria.” They are both supposed to submit their dissertations in the second year of the project.

## 2. Research activities at CIAT

Crosses were made at CIAT to generate populations with combined enhanced beta carotene and high protein contents in cassava roots as well as resistance to cassava mosaic disease (CMD). The parents were known to have cream or deep yellow coloured-roots and crude protein contents of 3-8%. A total of 1,555 seeds were derived from 17 parents in different combinations. The seedlings were established from embryo axes and micro-propagated to generate 1-5 plants per genotype. The number of plants successfully established *in vitro* from embryo axes was 981 (68%). DNA was extracted from each of the genotypes and used for marker-assisted selection (MAS) for resistance to CMD using one SSR marker, NS158 and a SCAR marker, RME 1. Both markers have shown fidelity for field resistance to CMD by 70% in Nigeria. Fourteen percent (14%) of the plants (138) were selected for CMD resistance from 16 families and further micro-propagated to generate 8-10 plants each. The selected CMD-resistant genotypes with enhanced

nutritional attributes were shipped as in-vitro materials to NRCRI in April 2008. A total of 687 plants were shipped.

### **Tangible outputs delivered**

138 CMD resistant, high beta carotene and protein Latin American cassava genotypes developed and shipped to Nigeria for evaluations by farmers in different agro-ecologies. These materials are undergoing hardening and micro-propagation for another related project in 3 other African countries. Two NRCRI research staff were trained on-hand in molecular breeding in an advanced laboratory. Also two other NRCRI research fellows are undergoing MSc - degree training in molecular breeding in 2 Nigerian universities.

### **Next steps/challenges**

The hardened in vitro plants shall be transferred to the field for evaluation for the traits of interest (beta carotene and protein content) as well as other important agronomic characters. There shall be further training for another research assistant at an advanced laboratory. Genotyping for resistance CMD, and marker assisted selection for beta carotene and protein contents shall be done. Participatory evaluation of selections by farmers shall be conducted.

*90.03 G4007.13.03: Capacity-building à la carte 2007–Application of molecular tools for controlled wild introgression into cultivated germplasm in Senegal  
July 2007–July2009*

### **Team Leader**

Ousmane Ndoye, ISRA, ousndoye@refer.sn  
Institut Sénégalais de Recherches Agricoles (ISRA–CNRA), B.P. 53 Bambey,  
SENEGAL (221)33973-6348, Fax (221)33973-6348

### **Collaborating institutions and scientists**

- Agropolis–CIRAD: Jean-François Rami
- UCB: David John Bertioli
- EMBRAPA: Marcio Moretzsohn

Groundnut cultivation in the Sahel zone of Africa faces important constraints, particularly drought stress and diseases, but the narrow genetic basis of the cultivated peanut *Arachis hypogaea* L. hampers the development of improved varieties through conventional breeding. The main objective of the project is to explore and exploit the up to now limitedly used variability of cultivated peanut's wild relatives through the utilisation of amphidiploids together with molecular tools.

In this project, work was carried out from crossing to building a genetic map. In fact various techniques, such as the use of moisture hat over the pollinated flower, improve the moisture level in the pots containing the plants, have been demonstrated and are now being used routinely in the crossings. The rate of success is high up to 80% in some crosses. From this work a BC1 population was developed using 73-30 as dormant cultivated parent. This population showed some level of heterogeneity when checked with SSR markers, due to probable heterogeneity of the seed lot used for the crosses. It was decided to develop a new F2 population from the cross Fleur11 x 73-30, using

controlled seed lots of the parents. In the course of developing the F2 generation, individual F1 plants will be controlled to select plants to be self-pollinated, based on their conformity to the parental lines.

The cross Fleur11 x (*A. ipaensis* x *A. duranensis*) was used as a first population for the development of chromosome segment substitution lines. Eighty eight (88) BC1s were grown and DNA was extracted for all BC1 individuals. At the same time, forty (40) Fleur 11 plants, the recurrent parent, were grown and DNA was also extracted. Genotyping of all individuals was done for 10 SSR selected as being polymorphic in a panel of cultivated varieties from Senegal. This first step allowed:

- i) to distinguish true BC1 plants from plants originating from selfed F1
- ii) to identify BC1 plants with unexpected alleles on the cultivated parents (coming from heterogeneity in the cultivated parent used in earlier generations) and;
- iii) to identify a homogeneous set of recurrent parents conform to be used for the crosses to BC2 generation.

After this first step of conformity control, 46 BC1 plants were selected as true BC1 to be advanced to the BC2 generation and crossed with the verified recurrent parents. All 88 BC1 plants (including those showing heterogeneity on the cultivated side) were used to build a SSR genetic map of groundnut. Two hundred and thirty one (231) SSR loci were scored on this population and a genetic map was build spanning a total distance of 1793 cM with 21 linkage groups. Using the genetic map obtained a set of 22 BC2 families representing optimal genome coverage of chromosome segments from the donor genome was selected to be advanced to the BC3 generation.

#### **Tangible outputs delivered**

- Improvement of the skill for crossing peanut with its related wild species
- DNA extraction protocol for peanut was refined
- Level of molecular polymorphism between the different parents was revealed using microsatellites markers.
- A genetic map was constructed with the BC1 population developed from the cross Fleur 11 x (*A. ipaensis* x *A. duranensis*)

#### **90.04 G4007.13.04: Capacity-building à la carte 2007–Characterisation of maize germplasm found in Ghana, using the bulking technique**

**July 2007–July 2008; extension July 2009**

#### **Team Leader**

Allen Oppong, CSIR, a.oppo<sup>ng</sup>@cropsresearch.org; alnopp@yahoo.co.uk  
Crops Research Institute (CSIR), Kumasi, Ghana, Tel. No.: 23305160391 Fax:  
233051-60142

#### **Collaborating institutions and scientists**

- USDA–ARS–CHPRRU: Marilyn Warburton
- CIMMYT: Yunbi Xu
- Universidad de la Republica, Uruguay: Jorge D. Franco



### **Supporting scientists**

- Manfred Ewool, CSIR-CRI, Kumasi, Ghana
- Maxwell Darko Asante: CSIR-CRI, Kumasi, Ghana
- Ruth Thompson, CSIR-CRI, Kumasi, Ghana

Maize is a very important staple crop in Ghana. Research efforts have been made to improve the yield and quality of the crop. Ghanaian maize landraces are not adequately characterised and it is believed that some of these landraces have very good traits such as disease resistance, tolerance to drought, etc. which when exploited could be used to produce hybrid maize, synthetic lines etc. The CSIR-CRI with the support of the GCP and CIMMYT is making efforts to characterise both genotypically and phenotypically Ghanaian landraces to identify lines for varietal improvement. To realise these objectives a nationwide collection of Ghanaian maize landraces was carried out from October to December 2007. Collections were made from all the agro-ecological zones; comprising the Coastal savanna in the south-east, high forest in the south-west and the middle zone, forest transition in mid-north and the Guinea savanna in the northern parts of Ghana with some few collections from the neighboring Republic of Togo. In all, about 700 accessions of varied diversity were collected. Three to four cobs were collected from each farmer in a random manner. The means of possession by which these farmers came by these landraces were ascertained. These were by through inheritance from their fathers, grand fathers and friends with some attributing source to their ancestors association with early European Missionaries and also from neighboring countries (especially those from the border towns and villages).

Also to improve the human resource capacity and the equipment needs of the CSIR-CRI to cope with its research drive, training in Marker assisted selection was undertaken at the ABC laboratory CIMMYT, Mexico from 22nd January to 15<sup>th</sup> May, 2008. The training at CIMMYT covered as much as possible all areas of Marker assisted selection (MAS) as well as an introduction into programming used in the analysis genetic data, such as R, Structure and Powermaker and the creation of core subsets. The training in MAS involved techniques in plant (seed and leaf) DNA extraction, DNA amplification using PCR and the use of the gel electrophoresis as well as scoring and analysis of gels. SSR markers linked to MSV resistant alleles were used in this training. Tuition was provided by the Co-Principal Investigators of this project listed above. At the end of the training a modified implementation plan of the original proposal, was developed which emphasises on the development of maize streak resistant varieties from local landraces.

### ***90.5 G4007.13.05: Capacity-building à la carte 2007–An integrated proteomics and genomics approach to discover salt tolerance genes***

***July 2007–July 2008***

#### **Principal Investigator**

Ghasem Hosseini Salekdeh, ABRII, h\_salekdeh@abrii.ac.ir

Agricultural Biotechnology Research Institute of Iran (ABRII), Karaj, Iran,  
Mahdasht Road, P.O.Box: 31535-1897, Karaj, Iran. Tel: +98-261-2702893; Fax:  
+98-261-2704539

## **Collaborating institutions and scientists**

- IRRI: Abdelbagi M. Ismail
- IPK: Mohammad-Reza Hajirezaei

Proteomics proved to be a powerful approach to identify proteins involved in plant tolerance to abiotic stresses. In past few years several groups have applied this approach to discover salt responsive genes in crops particularly rice. However, it has been observed that many important proteins including transcription factors are masked by high abundant proteins and cannot be detected on two dimensional electrophoresis gels (2-DE). On the other hand, function of most of stress responsive genes have not be validated using approaches such as RNAi before introducing genes in marker assisted breeding (MAB) and/or transformation programme. To address these two important issues, we initiated total and nuclear proteome analysis of rice tolerant (FL478) and sensitive (IR29) lines backed up with metabolome analysis. The candidate genes will then be validated using RNAi approach. Tolerant genes can be introduced to MAB and transformation programmes. The proteome, metabolome are being performed at ABRRI and IPK, respectively and RNAi analyses will be performed at IRRI.

## **1: Research activities at ABRRI**

### **1.1: Samples preparation**

Seeds of salt-tolerant genotype, FL478, and salt sensitive genotype, IR29 were obtained from the International Rice Research Institute (Manila, Philippines). Surface sterilized seeds were grown on rectangular plastic trays filled with complete strength Yoshida nutrient solution. After two weeks, plants were treated by 50 mM NaCl for 2 days followed by 100 mM NaCl. After 15 days of salt stress, leaves and root were collected from stressed and normal FL478 and IR29. The Na<sup>+</sup> and K<sup>+</sup> contents were measured using flame photometrically. Morphological characters including root and leaf length, fresh and dry weight were measured in normal and stressed plants. All analysis was done on a completely randomised design with 3 replicates. The overall effect of salinity was significant as determined by shoot dry matter (Fig. 1) and Na<sup>+</sup> content.

### **1.2. Proteome analysis**

The proteome analysis of samples are being performed at ABRRI. Total root and leaf protein extraction was performed using the TCA/acetone precipitation methods. The protein concentrations were measured by Bradford assay with BSA as standard. The isoelectric focusing was performed using immobilised pH gradient (IPG) 24 cm 4-7 L strips and the second dimension was performed on 12.5% SDS-PAGE gels. Analytical gels were stained with silver nitrate and gels are being analyzed using Melanie 4 software (Fig 2). The nuclear proteome has been extracted from similar leaf samples and is being analyzed using 2-DE.

## **3. Research activities at IPK**

Root metabolites were extracted and analyzed at IPK, Germany. Soluble and insoluble sugars (glucose, fructose, sucrose, and starch), 20 amino acids, nucleotides (AMP, ADPGLC, ADP, ATP, ATP/ADP), sugar alcohols (inositol, glycerol, erythrol, sorbitol, and manitol), antioxidant metabolites (total ascorbate, reduced ascorbate, total glutathione, oxidatised glutathion, reduced glutathione, total cysteine, oxidatised cysteine, reduced cysteine), and several other metabolites (malate, PEP, 3PGA, citrate,

isocitrate, Glc-1-P, Fru6P, Fru1,6bisP, and UDPGlc) were measured. The amount of glucose, fructose, inositol, glycerol, alanine, and proline increased under stress in tolerant genotype, FL478, but decrease in sensitive genotype, IR29 (Fig 3 and 4). Metabolome analysis of leaf blade is in progress.

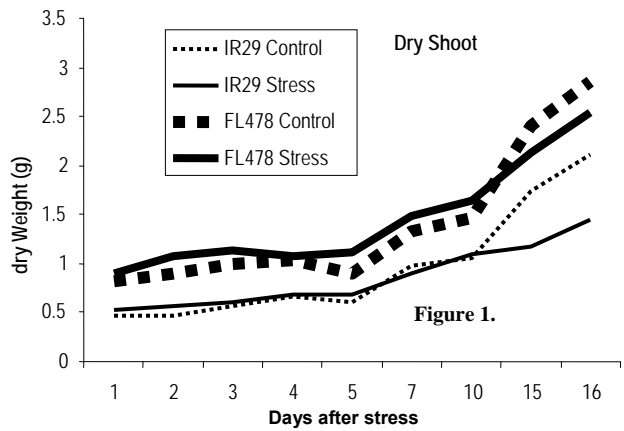


Figure 1

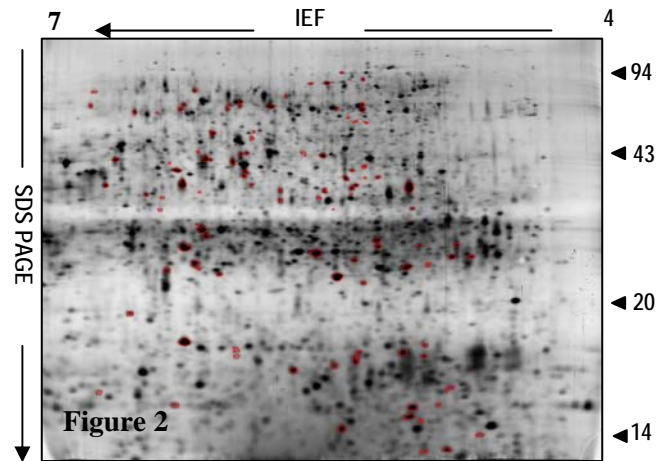


Figure 2

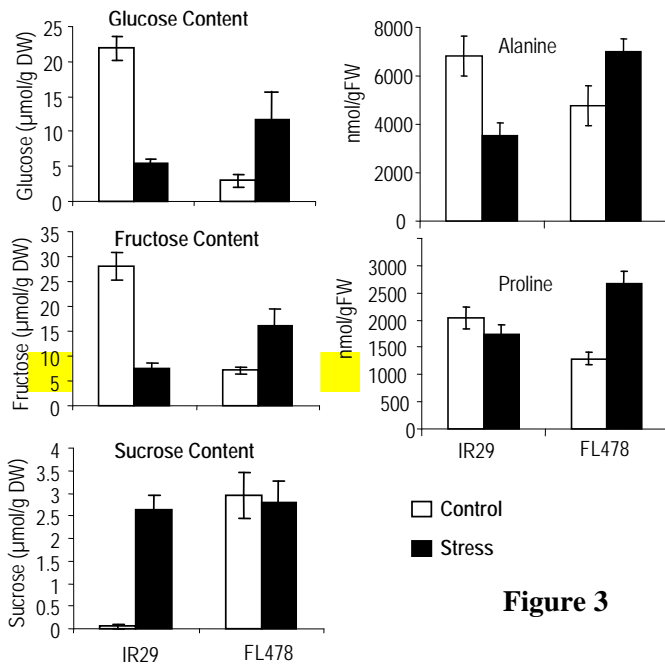


Figure 3

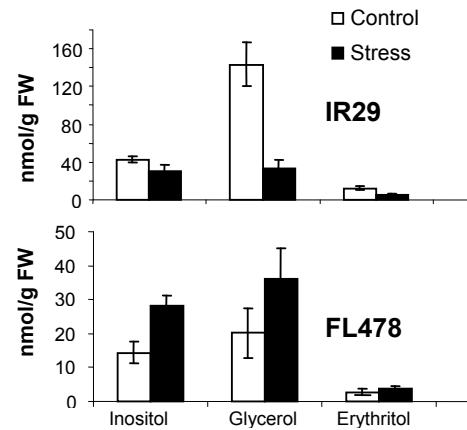


Figure 4

**90.6 G4007.13.06: Capacity-building à la carte 2007–Enhancing capacity of ICABIOGRAD in phenotyping and molecular analysis to develop elite rice lines suitable to Indonesian uplands**

**July 2007–July 2008**

**Team Leader**

Masdiar Bustamam, ICABIOGRAD, [masdiarb@indo.net.id](mailto:masdiarb@indo.net.id)

Jalan Tentara Pelajan 3A Bogor 16111, Republic of Indonesia, 62251338820

**Collaborating institutions and scientists**

- ICABIOGRAD: Kurniawan Rudi Trijatmiko (Research Scientist), Wening Enggarini (Research Scientist)
- IRRI: Casiana Vera Cruz (Plant Pathologist)

In this GCP CB a la Carte 2007, ICABIOGRAD conducted five solutions, i.e. nine months hands-on research on marker-assisted selection and phenotyping at IRRI, partial support for ongoing long-degree programme, minigrants for two nondenaturing PAGE apparatus, minigrants for computers, minigrants for building blast inoculation room and moist room. By the end of this 2007 project, we have finished all solutions.

For the hands-on research at IRRI, we have sent one of ICABIOGRAD's staff (Wening Enggarini) to IRRI for learning how to do marker-assisted selection and phenotyping, by using the ongoing research population material (BC2F5 derived from Way Rarem x Oryzica Llanos-5). We enclose the result of this hands-on research in the appendixes. Wening is an ongoing PhD student at Bogor Agricultural University (BAU). Due to the delay on immigration administration arrangement and adjustment to her course schedule at BAU, Wening did her hands-on research at IRRI for four months, instead of nine months.

Partial support for ongoing long-degree programme, have been used to support the PhD programme of Wening. This includes tuition fee for three semesters, living expenses and travel.

Minigrants for two nondenaturing PAGE apparatus, have been used for purchasing two sets of Dual Triple-Wide Mini-Vertical Kit, CE (CBS Scientific – USA), including Mini Power Supply.

Minigrants for computers have been used for purchasing a Toshiba Satellite Notebook (L40-N502). We have installed some useful software, i.e. QGene for Windows and GGT 2.0 (Graphical GenoTypes) in this notebook to support our work.

Minigrants for building blast inoculation room and moist room, were used for the purpose as mentioned in the proposal. Recently ICABIOGRAD have built a new greenhouse before the GCP CB a la Carte 2007 started. We built the moist room inside the greenhouse. This moist room was framed by hollow iron and covered by polycarbonate wall. The area of the room is 6 m x 3.5 m. This room has double-door. We supplemented

this room with two sets of air conditioner, programmable water sprinkle system, thermometer, and hygrometer.

We put up a building attached to the greenhouse containing the moist room. This building has single entry door, and there is no other entry to greenhouse except through this building. This building has a room for blast inoculation. The area of the inoculation room is 4.5 m x 3 m. We put an inoculation box inside this inoculation room.

## **91. G4007.14: Fellowships and travel grants 2007**

### **Principal Investigator**

Carmen de Vicente, GCP, [c.devicente@cgiar.org](mailto:c.devicente@cgiar.org)

Apdo. Postal 6-641, 06600 Mexico, D.F., Mexico, Phone: +52 55 5804 2004

The following are still ongoing in 2008:

### ***Fellowships***

**1) Name:** Habibul Bari

**Research Title:** Use and application of mapping software to analyse polymorphism data in salt sensitive and tolerant BC2F2 progeny from a cross between salt tolerant rice landrace, Boilam and Farmer popular rice, BR26 and BR27, in an effort to find linkage between tolerance and specific DNA markers"

**Home institution:** University of Dhaka

**Host institution:** International Rice Research Institute (IRRI), Philippines

**2) Name:** Gustave Djedatin

**Research Title:** Characterisation and molecular introgression of bacterial leaf blight resistance gene in rice

**Home institution:** University of Abomey-Calavi

**Host institution:** Institut de Recherche pour le Développement (IRD), France

**3) Name:** Dongcheng Liu

**Research Title:** Analysis of leaf and root growth kinetics and related gene expression in rice during progressive soil drying

**Home institution:** Institute of Genetics & Developmental Biology, Chinese Academy of Sciences

**Host institution:** International Rice Research Institute (IRRI), Philippines

**4) Name:** Asrat Asfaw Amele

**Research Title:** Genetic investigation for drought tolerance in common bean (*Phaseolus vulgaris* L.)

**Home institution:** AWASSA AGRICULTURAL RESEARCH CENTER

**Host institution:** International Center for Tropical Agriculture (CIAT), Colombia

### ***Travel grants***

**1) Name:** Soraya de Macedo, Embrapa Genetic Resources and Biotechnology

**Traveled to:** International Crops Research Institute for the Semi-Arid Tropics

(ICRISAT), India

**Purpose:** Hands-on training

**2) Name:** Jude Obidiegwu, National Root Crops Research Institute (NRCRI)

**Traveled to:** Lancaster Environment Centre, UK

**Purpose:** Hands-on training

For further details on the concept behind the Fellowships and Travel Grante programme, see abstract for project G4008.38: Fellowships and travel grants 2008.

## **92. G4007.20: Managing the Generation Challenge Programme in a post-International Treaty world**

*August 2007–July 2008 (No-cost extension requested until 31 December 2008)*

### **Principal Investigator**

Michael Halewood, Senior Scientist, Head, Policy Research and Support Unit, Bioversity International, m.halewood@cgiar.org

Via dei Tre Denari, 472.a, 00057 Maccarese, Rome, Italy

### **Collaborating institutions and scientists**

- Bioversity International: Gerald Moore, Honorary Fellow
- Central Advisory Service on Intellectual Property (CAS-IP), Bioversity International: Victoria Henson-Apollonio, Senior Scientist, Project Manager, (member of the GCP informal policy group)
- Generation Challenge Programme: M Carmen de Vicente, Subprogramme 5 Leader
- EMBRAPA: Maria José Sampaio (member of the GCP informal policy group);
- IRRI: Ruairaidh Sackville-Hamilton, Head, Genetic Resources Centre

The project aims to raise awareness and provide training for GCP partners regarding the International Treaty on Plant Genetic Resources for Food and Agriculture and to help them understand their options when acquiring and distributing PGRFA whilst conforming to international legal standards and obligations. Outputs include:

- Twenty-three GCP researcher partners trained on the implications of the Treaty and how to comply with it;
- Best practices guidelines and a package of related reference materials for use by GCP partners;
- Training materials regarding the multilateral system of access and benefit sharing (MLS) created by the Treaty, and policy options concerning movement of materials not included in the MLS; and
- A network established through which GCP members can access Treaty-related advice and troubleshooting.

A technical training workshop was held on 17-18 September 2007, in Benoni, South Africa, immediately following the GCP Annual Research Meeting. The 23 participants represented the diversity of GCP partners: NARS, ARIs, CGIAR centres, the private

sector and universities. Participants reviewed existing policies and practices concerning the pooling and transfer of genetic resources in support of GCP research, and the complex variables that can affect the final distribution of GCP research products. Topics included: the Treaty itself; the use of the Standard Material Transfer Agreement (SMTA) for Annex 1 materials; the use by CGIAR Centres of the MTA to be adopted by the Governing Body in October/November 2007<sup>1</sup>; the use of an MTA by state parties who are members of the GCP for non-Annex 1 materials; and legal instruments for acquiring materials from and distributing to GCP partners located in countries that are not parties to the Treaty. Ultimately, the participants agreed upon several issues including when to use and not use the SMTA, how to deal with 'borderline PGRFA' and contracting and non-contracting parties to the Treaty. A summary report including recommendations to GCP management team was prepared and submitted to the Sub-Programme 5 Leader. The full workshop report has been finalised. That report, and reference materials, are being assembled for uploading onto the GCP website.

The content of the best practices guidelines (and related legal instruments) will depend largely upon the responses of the GCP management to the recommendations of the workshop. When a formal response is provided, guidelines will be developed to assist GCP researchers in understanding the legal instruments required for transfers of materials in support of product delivery. If requested by GCP management, Bioversity will also draft these legal instruments.

It was agreed that all workshop participants could contact the resource persons for assistance in addressing the policy issues raised during the workshop. The reference materials and guidelines will emphasise the availability of this informal network of specialists for consultation.

### **93. G4007.21: Genotyping Support Service**

#### **Principal Investigator**

Humberto Gómez Paniagua, GCP, [h.gomez@cgiar.org](mailto:h.gomez@cgiar.org)

Apdo. Postal 6-641, 06600 Mexico, D.F., Mexico, Phone: +52 55 5804 2004,

Fax: + 52 55 5804 7558

#### **Collaborating institutions and scientists**

GCP Subprogramme Leaders

Generation Challenge Programme (GCP) studies the genetic diversity of germplasm using genomics to discover the genes and alleles controlling the expression of complex agronomic traits. The results are useful for the plant sciences in general but especially for crop breeding because they contribute to the better understanding of traits controlling plant performance, and at the same time they facilitate the job of breeders, i.e. to create varieties faster and more suited to the users' needs. The GCP strives to transfer these results to crop scientists in the developing countries.

The Genotyping Support Service (GSS) facilitates the access of national agricultural research systems (NARS) to technologies developed by the GCP, bridging the gap between research in advanced facilities and that closer to the fields of developing

countries. With these services, GCP offers cost-efficient genotyping services worldwide, access to data, support and training in statistics for proper interpretation of genotype and phenotype data. The aim is to raise the capacity of developing country researchers to access and use modern and more efficient tools and technologies.

In its first call, GSS received 33 applications, out of which 25 were selected (then 6 ceased). The 19 requests in process are as follows: Potato 5, Cassava 4, Rice 3, Sweetpotato 2, Maize 1, Chickpea 1, Coconut 1, Yam 1, *Musa Ensete* 1. These applications were received from Tanzania 4, Bolivia 1, Kenya 3, India 1, Ghana 3, Brazil 2, Iran 1, Sri Lanka 1, Chile 1, Ethiopia 1 and China 1. Fourteen of these applications aim to conduct molecular characterisation of germplasm and five address plant breeding issues such as marker assisted selection, gene tagging or genetic map development. The workshop for this call will take place on the last week of January 2009.

The GSS opened its second call between June 1<sup>st</sup> and August 10<sup>th</sup>.

#### **94. G4007.21.03.02: Provision of Genotyping Support Services (GSS)**

*March 2008–March 2009*

##### **Principal Investigator**

Casiana M Vera Cruz, International Rice Research Institute (IRRI), [c.veracruz@cgiar.org](mailto:c.veracruz@cgiar.org)  
DAPO Box 7777, Metro Manila, Philippines, Telephone: +63 (2) 580-5600 ext. 2727, Fax: +63 (2) 580-5699

##### **Collaborating institutions and scientists**

- Generation Challenge Programme, CIMMYT: H Gomez
- China National Rice Research Institute, Hangzhou, China: Jian-Li Wu
- CRI: Maxwell Darko Asante
- TNAU: S Manonmani

As part of its capacity building activities, GCP established the Genotyping Support Services (GSS) to assist developing country institutes to learn how to apply molecular marker technology in a plant breeding program or to study the genetic diversity of important plant species and varieties that are indigenous to their area, and help them in genotyping breeding lines for crop improvement. In this undertaking, GCP has engaged IRRI as an expert genotyping services provider.

The details and scope of the responsibilities of each party in each genotyping service is contained in the Work Order for each assignment. The general responsibility of CIMMYT (as host agent for GCP) is to provide the information of the institution whose DNA samples are to be genotyped as well as the Marker Data for use in the genotypic analysis. IRRI's responsibilities are to provide all the relevant information regarding the legal requirements for shipment of DNA to comply with customs, phytosanitary and other requirements, to conduct the genotypic analysis, and to provide the client institution and GCP the data set generated from the services described in each work order.



### 1. Work Orders: scope of work, timing, and payment

Two work orders are now in place: (1) screening a population for bacterial blight resistance genes *Xa7* and *Xa21* and for fertility restorer genes *Rf3* and *Rf4* (requested by the China National Rice Research Institute (CNRI) and, (2) analyzing the genetic diversity of African rice germplasm (requested by the Crops Research Institute, Ghana). In the pipeline is a project from India to genotype CBSN-derived indica x indica lines.

In general, the terms of the work commence when the work orders are signed until the appropriate timeframe of genotyping services is completed. The requesting institution has the right to declare confidentiality of the genotyping data.

**Table 1. Current and incoming GSS requests from developing country institutes.**

Work Order	Genetic Materials	Samples (no.)	Services required
J. Wu – China	Parental and F <sub>n</sub> progeny lines	625 DNA in 5 replicates	Genotyping for <i>Xa7</i> , <i>Xa21</i> , <i>Rf3</i> and <i>Rf4</i>
M. D. Asante – Africa	Rice germplasm collection of CRI, Kumasi, Ghana	122	Genotyping for genetic diversity using 40 SSR type markers
S. Manonmani – India	CBSN derived indica x indica lines	500	Genotyping for <i>Rf3</i> and <i>Rf4</i> and Waxy genes (oligos Wx484 and Wx485)

The total compensation is based on the number of samples and markers analyzed for each Work Order, which have been obtained through competitive cost estimate.

### 2. Requirements: paper work, DNA quality, shipping DNA

In order to facilitate the importation of DNA into IRRI's host country (Philippines), an Import permit needs to be secured by IRRI from the Philippines' Bureau of Plant Industry Quarantine Office containing the required details of the samples to be imported.

Processing of the permit should be initiated 2-3 weeks before the scheduled arrival of the DNA to ensure that the two-month effectivity of the permit will not lapse. The client institution should provide a Phytosanitary Certificate (PC) issued by the authorized Plant Quarantine Officer of the country of origin which is enclosed with the DNA package during transport.

IRRI provides a protocol for the extraction of DNA suitable for Polymerase Chain Reaction (PCR) as well as a procedure for checking the quality of the DNA. This is to ensure that IRRI receives uniform concentrations of the DNA and is PCR-ready. This will prevent any delays in the genotypic analysis. IRRI recommends further that the DNA samples be dried prior to shipment to protect DNA from degradation during transport. Instructions on re-resolution of the DNA should then be provided by the client institution to IRRI.

### 3. Genotyping activities and reporting

The genotypic analysis to be used in each work order is defined by the client institution in consultation with IRRI and GCP. A GSS contract can be used for marker-assisted selection such as the work order from CNRI (China) which involves genotyping for

genes *Xa7*, *Xa21*, *Rf3* and *Rf4*. GSS can also provide support for genetic diversity studies as in the case for the work order from the Crops Research Institute (Ghana) to study diversity of African rice germplasm. Data gathered by IRRI will be submitted in progress and final reports with a 100-word abstract describing the genotyping data attached to the report.

The initial GSS project from CNRRI includes 625 DNA samples in 5 replicates. Upon arrival at IRRI, the DNA quality and concentration of each sample are being checked. The genetic materials from Ghana are expected to arrive at IRRI on July 15 while that from India is still being set.

## **95. G4007.22: GCP Workflow and Repository System (WMS)**

*Phase one: July 2007–March 2008; Phase two: April–August 2008*

### **Principal Investigator**

Norbert Niederhauser, Cropster GmbH, [norbert.niederhauser@cropster.org](mailto:norbert.niederhauser@cropster.org)  
Tellenweg 10, 6850 Dornbirn, Austria, [www.cropster.org](http://www.cropster.org)

### **Context**

GCP's project management is a challenge due to its dynamic nature and international distribution of projects, partners and management. Donors demand that management processes are transparent and traceable and outcomes available.

To keep track of ongoing proposals, contracting and reporting processes and having project status and outcomes timely available to different audiences including the management team located in different parts of the world, a central online information management system is needed.

Access to data including project location, methods, output and budget overview from ongoing and past projects is needed for preparing new proposal calls, plan future project strategies or simply prepare a comprehensive report about ongoing and past research in certain fields. Central contact and event management was another demand to be covered by this project.

### **Implementation**

The implementation of the above briefly described system was structured in the following modules: Contacts, events, projects including project proposal workflow and GSS workflow, publications, products, repository, reports. All modules and their functions interact closely. This integration brings flexibility and efficiency to the users. The whole system is based on Internet technology, which enables decentralised access to the system while maintenance, monitoring and data storage is organised centrally. A short introduction to each module follows:

#### *Contact management system*

Besides the storage of contact information like address and institution, this system enables the storage of contact related documents and the management of standard GCP keywords as well as customised keywords. Each contact can be linked to any other module in the system. This enables the system operators to later relate projects, events or other activities to users and their institutions.

### *Event management module*

All GCP events and their attendees are registered in this module. An event is described like event name, date and place. For the AGM a special extension was created: AGM members were invited through the system and registered their participation details including flight details and hotel reservation.

Projects module including project proposal workflow and GSS workflow Genotype Support Service (GSS):

All GCP projects can be managed through this module. Project documents are stores to each project as well as timeline, outputs, activities, publications and products. Stored information is managed partly through project management team. The projects budget information and actual balance is retrieved from an third party software. The proposal workflow is part of the projects module. It monitors open tasks and sends automatic notification e-mail messages to different users depending on milestones. The GSS workflow is built into that module as a specific functionality, user tailored to the Genotype Support Service (GSS). It provides close online interaction between the GSS management and the GSS users, the PI's. Document submission is done via an online interface where also certain data input fields allow a better structuring of content and a more efficient, in some cases immediate, evaluation.

### *Publications module*

This module enables to keep track of publications, which were agreed upon projects acceptance or emerged through some projects. It additionally allows the GCP team to upload and register any GCP publication. This module forms part of the product inventory, publicly available (restrictions may apply for certain documents) through the GCP website.

### *Products*

Any product, which will be developed in some GCP project is registered immediately during the proposal and evaluation process. From that point on, the GCP product manager can keep track of the product and its development, while having results immediately available to the public and the GCP managers (restrictions apply).

### *Repository*

Any document uploaded to the WMS repository system. There all text-readable files (DOC,XLS,PDF,TXT, ...) are indexed an made available for full text search. The central system allows GCP users to perform a full text or structured search and make so more use out of the already collected information.

### *Reports*

The reports module allows advanced querying of the above described modules for compiling quick reports and online-publication of certain information on the GCP website. This powerful module supports strategic planning processes and donor communication.

The whole system is programmed in open source software, fostering free software development activities. A robust and state of the art security system protects the system against intruders or data misuse.

**96. G4008.23: Statistical rules for defining characteristic genotype and marker sets**

*January 2008–December 2008*

**Principal Investigator**

Fred van Eeuwijk, Biometris, Wageningen University UR; fred.vaneeuwijk@wur.nl  
P.O. Box 100, 6700 AC, Wageningen, The Netherlands; (31) 317-482902; Fax:  
(31) 317- 483554

Large sets of genotypic score tables for molecular markers have been and are constructed for a wide series of crops within various activities within GCP-SP1. This marker information forms the basis for population genetic studies on domestication, drift, selection, linkage disequilibrium, etc. A next step will consist in using this genotypic information for association mapping purposes.

For easy reference in future population genetic and mapping studies, it is valuable to define crop specific patterns of molecular marker variation. The main objective of this project is to investigate various strategies to arrive at small but informative sets of genotypes and markers. Such reference sets of genotypes and markers should regenerate large part of the relevant genetic variation with relatively little effort. The reference information should be used to connect different population genetic and quantitative genetic studies (including association studies) within the same crop.

The choice of the genotypic and marker reference sets should be driven by both statistical and molecular genetic principles. It is obvious that statistical dimension reduction techniques provide guidelines for choosing genotypes and markers. However, for easy use in future, also molecular genetic requirements should be involved in the decision criteria, especially those that determine the ease with which markers can be generated and the quality with which they can be read.

The output of the project will consist in 1) rules and protocols to be used in general for all GCP crops and 2) defined reference sets for a number of crops.

**97. G4008.24: From attractiveness to feasibility: A strategic assessment of the capacity to develop and adopt GCP technologies**

*January 2008–December 2009*

**Principal Investigator**

Stanley Wood, International Food Policy Research Institute, IFPRI; s.wood@cgiar.org,  
2033 K Street, Washington D.C. 20006-1002, USA; (1) 202-862-5600; Fax: (1)  
202-467-4439

## Collaborating institutions and scientists

- CIAT: Glenn Hyman
- Consultant/Analyst: Carlos Tovar
- Virginia Tech, IPM/CRSP Impact Coordinator: George Norton
- CGIAR Agricultural Science and Technology Indicators (ASTI): Nienke Beintema
- Country case study collaborators

### 1. Project summary

A previous GCP study (Hyman *et al.*, 2008) identified high priority geographic areas and specific cropping systems for which GCP technologies could potentially have the greatest human welfare impacts (e.g., highest levels of poverty, large extents of food staple crop production, and high incidence of drought). Implicitly, however, this study assumed that GCP technologies targeted for and delivered to such priority areas would be successfully adapted by local breeding programmes, and delivered to and adopted by smallholders. This project extends the prior assessment by developing and applying methods to assess the local capacity of crop development and delivery programmes as well as farming communities to actually realise the projected benefits of GCP technologies released to national programmes. We will, thus, assess the “feasibility” of achieving the potential benefits of GCP technologies in areas where their attractiveness has already been established.

The four main activities of the project can be summarised as:

- **Preparation of a concept/method paper:** setting out an analytical approach that integrates the existing assessment of the potential benefits of GCP technologies (attractiveness) with measures of local technology adaptation, delivery and adoption capacity (feasibility).
- **Country case studies:** Three or four specific priority trait and crop combinations will be examined to test and refine the improved method for assessing potential GCP impacts. For each case study, the project will undertake empirical work to develop and evaluate specific indicators of capacity and integrate them into an improved GCP priority-setting framework.
- **Two stakeholder/focus group review meetings:** The first will assess how best to extend the current GCP priority-setting approach to encompass local capacity indicators. A second meeting will evaluate the results of the case studies, assess the relevance and reliability of their findings, review the workplan for reporting, packaging, disseminating results, and propose GCP management processes for best utilizing them.
- **Extrapolation of the case study findings** for specific trait/crop technologies across all GCP focus regions/countries. The project researchers will prepare and present a draft set of more finely-tuned GCP priorities.

Analysis of capacity at the country scale is divided into three components: i) the technology adaptation capacity of national breeding programmes, ii) the technology delivery and support systems for smallholder agriculture, and iii) the ability of local farmers and communities to utilise locally-adapted GCP technologies. For the first two components, we rely on datasets from different sources including; the Agricultural

Science and Technology Indicators (ASTI)<sup>1</sup> database, government investment in rural sectors such as infrastructure (roads, irrigation, and electrification), and agricultural extension and education. Obtaining and compiling data for the third component is also challenging, and the project will begin that search through evidence compiled within and across CGIAR centers and, through local collaborators in case study countries. Furthermore, established linkages with the HarvestChoice initiative<sup>2</sup> will help to enrich the compilation of adoption datasets for many countries in sub-Saharan Africa and South Asia.

## 2. Activities and progress

The first review meeting was held on March 24-26, 2008 at IFPRI, Washington D.C., USA. The core project team (Wood, Hyman, and Tovar) was joined by Dr. Carmen de Vicente, Leader of the GCP Subprogramme 5, George Norton, and Nienke Beintema, as well as IFPRI staff (Falk-Zepeda – genetic resources, Nagarajan – seed systems) who provided specialised insights into the topics discussed. The meeting reviewed the goals and objectives of a feasibility assessment that accounts for local capacity building on the prior assessment of potential benefits (Hyman *et al.*, 2008.)

We are currently compiling estimates of investments in and human resources devoted to agricultural research and extension, and compiling national metrics related to the performance of seed systems and technology adoption. This work is being targeted only to existing GCP focus countries, and will be further refined when Phase II priorities have been established by the GCP Management Team. The country level information will add much richness to the previous analysis undertaken using global scale spatial datasets.

## References

Hyman G, Fujisaka S, Jones P, Wood S, de Vicente S and Dixon J (2008). Strategic approaches to targeting technology generation: Assessing the coincidence of poverty and drought-prone crop production. *Agricultural Systems*. 98:50-61.

<sup>1</sup> More about the ASTI initiative can be found at <http://www.asti.cgiar.org/>

<sup>2</sup> [www.harvestchoice.org](http://www.harvestchoice.org)

## 98. **G4008.25: Advanced course on “Applied statistical methods in plant genomics”, Zaragoza 18-29 February 2008. Request for GCP grants to finance participation of candidates from developing countries** *October 2007–May 2008*

### Principal Investigator

Mediterranean Agronomic Institute of Zaragoza (CIHEAM–IAMZ);

[Niamz@iamz.ciheam.org](mailto:Niamz@iamz.ciheam.org)

Avenida de Montañana 1005, 50059 Zaragoza, Spain; Phone +34 976 716000;

Fax +34 976 716001

### Collaborating institution and scientist

- Wageningen University and Research Centre, the Netherlands: Fred van Eeuwijk

## **Context**

An advanced course on ‘Applied statistical methods in plant genomics’ has been conducted in Zaragoza, 18-29 February 2008. The targeted audience consists of plant breeders that want to update their statistical skills to be able to take full benefit of new genomic tools that have been developed over the last decennium. For researchers from NARS within the GCP these genomic tools are gradually becoming available too. For these researchers there is a need not only to update their molecular genetic skills, but also their statistical skills. This proposal was to benefit a limited number of such researchers working in collaboration or close to the GCP goals by attending the course. The project covered grants for travel, room and board.

## **Products**

The course was organised by CIHEAM, through the Mediterranean Agronomic Institute of Zaragoza (IAMZ), in collaboration with the GCP, and took place at IAMZ. The course was held over a period of 2 weeks, from 18-29 February 2008, in morning and afternoon sessions, with a total of 74 hours. The working language of the course was English.

One of the main outputs of the course was that at the end of training participants had acquired the knowledge and the skills to:

- Manage questions related to general statistical theory involving parameter estimation and hypothesis testing.
- Verify quality of various types of data and prepare data in files amenable to statistical analysis.
- Apply mixed models to their own data and interpret the outcome in terms relevant to them.
- Construct genetic maps from genotypic scores on various types of population.
- Find patterns in gene expression and metabolic data.
- Quantify genetic diversity from molecular data.
- Perform QTL analyses and association mapping analyses for a wide variety of genomic and classical phenotypic traits.

The other main output was to make the documents provided by lecturers to support their presentations available to other professionals not participating in the course.

A total number of 28 participants attended the full course, coming from 18 countries: Algeria (1), Brazil (1), Chile (1), Croatia (1), Denmark (1), Egypt (2), France (1), Germany (2), Greece (1), India (1), Morocco (1), Nigeria (1), Portugal (1), South Africa (1), Spain (7), Syria (1), Tunisia (2) and Turkey (2). All the participants were already involved in the subject matter of the course, with previous experience in basic or applied research, working in universities, research centres or private companies.

13 full scholarships were awarded to participants from developing countries, covering registration fee payment, travel, full board accommodation and medical insurance. The Generation Challenge Programme granted 8 of these full scholarships and CIHEAM awarded 5, plus two other scholarships for fee payment.

## **Impacts**

The results of the course were assessed through surveys conducted among participants and lecturers.

In the survey conducted among participants, the global score for the course was high (4.3 over 5) and comments about the benefit taken from the course for their professional activities were extremely positive. Participants considered all topics included in the programme to be interesting and the applicability of all of them to their personal work was considered high. Computer practicals were well valued and the objective of acquainting participants with the basics of the use of statistical analysis programmes was fulfilled, nevertheless they asked for the time devoted to them to be extended. Participants gave satisfactory appraisal of the lecturers; their presentations were considered to be of appropriate quality and their interaction with participants was positively evaluated.

In the survey conducted among lecturers, they considered that, although the group was quite heterogeneous, the global quality of participants was good, and their motivation was very high, which had a direct impact on the success of the course. The degree of interaction between lecturers and participants and among participants was very good, particularly during practicals, which made up a large part of the course.

The documents provided by lecturers to be distributed among participants have been prepared for their incorporation to GCP web page to be consulted by other interested non-participant professionals working with genomic data.

## **99. G4008.26: A cassava breeding Community of Practice in Africa for accelerated production and dissemination of farmer-preferred cassava varieties resistant to pests and diseases**

*January 2008–December 2010*

### **Principal Investigators**

Chiedozie Egesi, NRCRI, cegesi@yahoo.com

Cassava Programme, National Root and Tuber Crop Improvement Institute (NRCRI), Umudike, PMB 7006, Umuahia, Abia State, Nigeria.

Emmanuel Okogbenin, CIAT, E.okogbenin@cgiar.org

International Center for Tropical Agriculture (CIAT), Cali, Colombia;  
Cassava Genetics Unit, A.A. 6713, Cali, Colombia.

### **Collaborating institutions and scientists**

- CRI: Joe Manu
- NaCRRI: Yona Baguma
- National Root and Tuber Programme, ARI-Naliendele: Geoffrey Mkamilo
- IITA: Alfred Dixon
- CIAT: Martin Fregene



Cassava has become a major staple and food security crop in Africa. However, there is an urgent need for improved varieties to stop the rapid spread of pest and diseases, especially the cassava brown streak disease (CBSD) and the cassava mosaic disease (CMD), two resurgent crop diseases that have already caused low-grade famine in parts of Africa. Although National Agricultural Research Systems (NARS) are best suited to breed cassava for local needs, NARS in the past limited their activities mainly to testing and selection of improved germplasm, but more recently NARS breeders have begun formal cassava breeding. Several donor funded projects, including a Generation Challenge Programme (GCP) competitive grant to CIAT, Brazil, and 3 African countries – Ghana, Nigeria, and Uganda, and a Rockefeller foundation grant to Tanzania, IITA, and CIAT, are now conducting field-based, Marker Assisted Selection (MAS), and participatory cassava breeding. There is a need to build synergies between these 4 NARS breeding programmes, IITA, and CIAT through exchange of experience and improved germplasm to ensure rapid production of improved varieties and delivery to farmers. We propose setting up of a community of practice (CoP) involving cassava breeders in the 4 target countries that will permit a free-flow of experiences and information on breeding methods, best field practices, and improved varieties amongst the 4 countries. A primary activity of the CoP will be integration of MAS with field-based breeding and pre-breeding strategies. The project will also provide training in MAS as well as field-based and participatory plant breeding for current and a new generation of breeders. MAS is a specialised form of cassava breeding complementary to traditional field-based breeding. The CoP will therefore create and maintain close links with International Institute Tropical Agriculture (IITA) and CIAT, and NARS breeding programmes via sharing of germplasm/information and training that are outcomes of this project. In addition, linkages will be built with primary, secondary, and tertiary users of improved cassava varieties to ensure prompt uptake of improved varieties. Lastly, the CoP will be proactive in developing linkages with existing cassava breeding networks, International breeding programmes, and related GCP projects, including the genotyping support services (GSS), to bring the best expertise and experiences to bear on the breeding goals.

#### *Objectives*

- 1) Develop a community of practice through breeder-to-breeder training in traditional and new methods of breeding, germplasm resource exchange, and web-based information sharing as ways of achieving set breeding goals and addressing common problems.
- 2) Facilitate the integration of MAS with field-based breeding by the transfer and implementation of effective strategies for breeding primarily for disease and pest resistance in target countries
- 3) Strengthen the capacity of National programme breeders through formal (M.Sc.) and informal (exchange visits and workshops) training programmes
- 4) Build linkages with primary, secondary, and tertiary users of improved cassava varieties to ensure quick uptake of technology

#### *Expected outputs*

- 1) Capacity of national breeding programmes in Ghana, Nigeria, Tanzania, and Uganda to conduct field-based and molecular breeding strengthened

- 2) A new generation of cassava varieties preferred by farmers and resistant to the major pests and diseases prevailing in the target countries
- 3) A cassava breeding community of practice in Africa that involves cassava breeders, primary, secondary and tertiary users

## **100. G4008.27: Phenotyping course for drought related traits across tropical legumes–Concepts and practices**

*December 2007–October 2008*

### **Principal Investigator**

Vincent Vadez, ICRISAT, v.vadez@cgiar.org

GT Biotechnology, ICRISAT, Patancheru 502324, Andhra Pradesh, India –(91)  
40 30 71 34 63 – (91) 40 30 71 30 74

### **Collaborating institutions and scientists**

- ICRISAT: Fran Bidinger; L Krishnamurthy, Junichi Kashiwagi, Kesava Rao
- Agropolis–INRA, France: Francois Tardieu
- UAS: MS Sheshshayee
- UC–Riverside, USA: Jeff Ehlers
- CIAT: IN Rao
- IITA: Satoru Muranaka

### **Context**

Traits putatively involved in the tolerance to drought are difficult to deal with because the environment in which they are measured is variable, and their value for the crop adaptation to a given environment varies accordingly. So both a good understanding and characterisation of the environment, and established protocols are needed to measure traits with sufficient precision to have a value in research and breeding. Moreover, few traits are known to play a role in drought adaptation, and some are common across crops. This workshop was set to serve the TLI community involved in the phenotyping for drought in groundnut, cowpea, bean, and chickpea. The workshop provided participants: (i) a practical hands-on training in the measurement of drought-related traits and data management, (ii) the key principles about phenotyping, and (iii) opportunities for cross-legume (groundnut, cowpea, bean and chickpea) discussions on key traits involved in their adaptation to drought.

### **Course development**

The training course took 4 weeks and use the phenotyping experiments of TLI - Objective 1 as framework of action, along with several experiments with genotypes of each of the crops of the TLI project, so that technicians/scientists could have a hands-on training on crops they know. The training provided an opportunity for a large scale, hands-on training to phenotype traits that are part of the yield architecture in crops (namely roots and transpiration efficiency), and to manage drought stress in field conditions. The training course was organised in two phases and for two groups of people, technicians and scientists, with overlap between them: (i) a phase 1 of about 3 weeks, concerning the technicians, focused on running experiments, collecting data on a

large scale, treating the data as the experiment went, managing irrigation in the field experiment; and (ii) a phase 2 of about 2.5 weeks, concerning mostly the scientists and focusing on also having some hands-on training, analyzing and interpretation of the data generated during the course, learning or revising theoretical considerations behind phenotyping, and discussion of key traits related to drought that appear to have a common importance across the 4 crops of TLI. We also insisted on the logistic aspects related to large scale phenotyping experiments, and that often are an important component of data quality. Technicians will still be involved in the beginning of phase 2, to have an overall view of the data and of their treatment.

### **Findings and implications**

The outputs were:

- a set of technicians from different locations in Africa trained and tuned to the traits they are expected to measure in their respective location;
- a community of practice of scientists involved in similar traits across legumes species, equipped with a common background of essential concepts to produce “good” phenotyping, and that share similar or comparable protocols to evaluate their respective crops. The course was a way of establishing ties with the African scientists involved in TLI, know them better, assess weaknesses and strength.
- beginning of cross-crop linkages in relation to the phenotyping activities giving ground to a platform for phenotyping across crops (service offered across crops, comparison of genotypes across crops in different environments).

We also produced course materials that were recollected in a CD provided to each participant, a few datasets comparing TE and water uptake under stress across several genotypes of groundnut, cowpea and bean. Main beneficiaries would be the national programme scientist and technicians that took part to the course.

**Links:** This course followed a phenotyping workshop held in Montpellier in 2006, and used the same principles and some of its materials, after updating with data generated during the practical sessions

### **Next steps and/or challenges**

Several sessions during the course dealt with comparing crop response to stress across legumes. We explored the possibility to have same traits measured across key genotypes of the 4 legumes crops, under same environments. For instance, we found that the water uptake was higher in cowpea and groundnut than in bean, and that all crops showed genotypic variation for water uptake in lysimeters (long and large PVC tubes filled with soil, where plants are grown and their water uptake assessed by regular weighing). Next step would be to exchange genetic materials across institution for a cross comparison of several traits.

## **101. G4008.28: Characterisation of maize diversity in Central Europe**

*March 2008–June 2008*

### **Principal Investigator**

Peter P Papp, Agricultural Biotechnology Center; ppapp@abc.hu  
No 4 Szent-Györgyi A., Gödöllő, H-2100, Hungary; +36 (28) 526-100; Fax: +36  
(28) 526-145

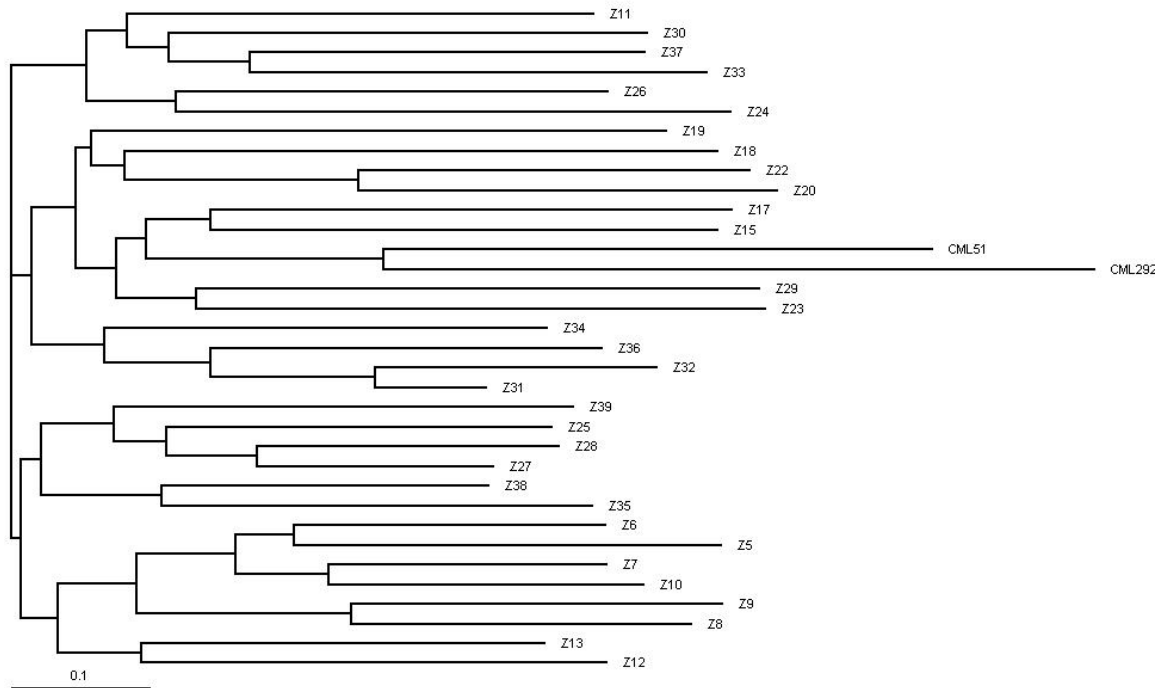
### **Collaborating institutions and scientists**

- ABC, Hungary: Tünde Golenar, György Botond Kiss
- Institute for Agrobotany, Tápiószele, Hungary: Laszlo Holly
- CIMMYT: Marilyn L Warburton

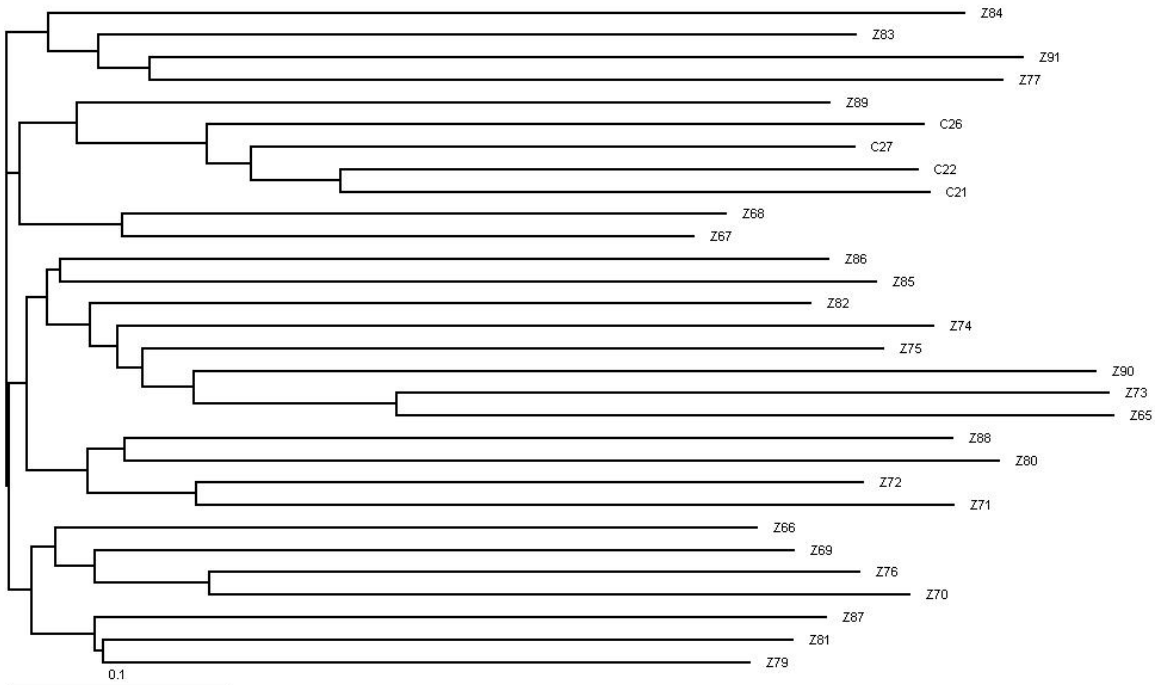
The Hungarian maize collection is comprised of over 4000 accessions managed by the Institute for Agrobotany at Tapioszele. Characterisation of the Hungarian maize germplasm has been planned in the context of global germplasm diversity. Collaborations between the Institutes of Hungarian Ministry of Agriculture (Agricultural Biotechnology Center and Institute for Agrobotany) and GCP affiliated Institute (CIMMYT) have been planned to compare the allelic diversity of the Hungarian maize collection with the diversity present in GCP's reference set as well as to provide training for a Hungarian scientist on molecular characterisation at CIMMYT.

Based on known pedigree, phenotypic traits and geographic locations, 32 inbred lines and 26 landraces were selected to represent a broad range of diversity of germplasm available in the Hungarian collection. Hungarian inbred lines and populations were scored with the same SSR markers (27 for inbred lines and 45 for populations) as has been used in the GCP for the characterisation of over 1500 maize inbred lines and 500 landraces. Control maize lines and populations were also included in the characterisation to ensure comparability of new and previously obtained data. Further comparison with the reference set of inbred lines and populations, generated by the GCP, will allow the Hungarian samples to be compared to the global diversity in maize.

Preliminary analyses of the marker data revealed that both the Hungarian inbred lines and the populations contain substantial diversity in maize, which will support national programme's breeding objectives as well as allow the participation in international programmes aiming at maize improvement and conservation. Further analyses, combining the SSR data of both the Hungarian inbred lines and populations with the data from the inbred lines and populations of the GCP reference set, respectively, are in progress. This will allow the Hungarian maize diversity to be directly compared to the global analyses of maize diversity; this in turn will allow an analysis on where maize in the world is most similar to Hungarian maize, where new sources of diversity can be found, an estimation of how unique the Hungarian maize is.



**Fig. 1.** Relationship between maize genotypes of 32 Hungarian inbred lines (Z5-Z13, Z15, Z17-Z20 and Z22-Z39) and 2 GCP control lines (CML51 and CML292) based on Euclidean distance analysis of data from 27 SSR markers generated by PowerMarker.



**Fig. 2.** Relationship between maize genotypes of 26 Hungarian landraces (Z65-Z77 and Z79-Z91) and 4 GCP control populations (C21-C22 and C26-C27) based on Euclidean distance analysis of data from 45 SSR markers generated by PowerMarker.

## **102. G4008.29: Characterisation of bean diversity in Central Europe**

*March 2008–June 2008*

### **Principal Investigator**

Peter P. Papp, Agricultural Biotechnology Center; ppapp@abc.hu

No 4 Szent-Györgyi A., Gödöllő, H-2100, Hungary; +36 (28) 526-100; Fax: +36 (28) 526-145

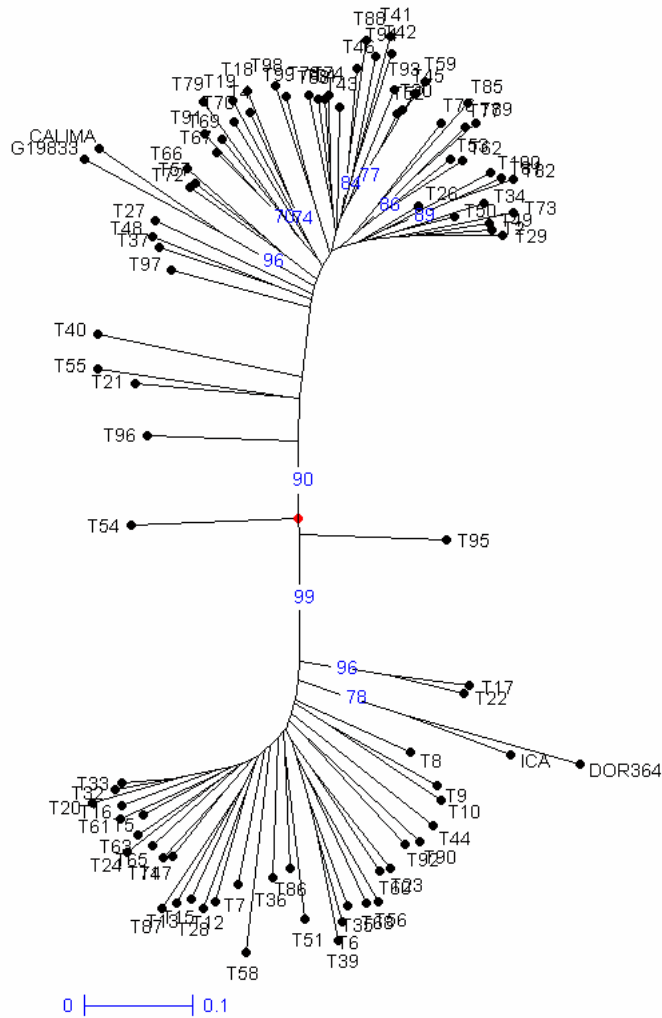
### **Collaborating institutions and scientists**

- ABC, Hungary: György Botond Kiss
- Institute for Agrobotany, Hungary: Laszlo Holly
- CIAT: Matthew W. Blair

The Hungarian bean collection is comprised of over 3000 accessions managed by the Institute for Agrobotany at Taposzele. Characterisation of the Hungarian bean germplasm has been planned in the context of global germplasm diversity. Collaborations between the Institutes of Hungarian Ministry of Agriculture (Agricultural Biotechnology Center and Institute for Agrobotany) and GCP affiliated Institute (CIAT) have been planned to compare the allelic diversity of the Hungarian bean collection with the diversity present in GCP's reference set as well as to provide training for a Hungarian scientist on molecular characterisation at CIAT.

Based on known pedigree, phenotypic traits and geographic locations, 100 Hungarian bean accessions were selected to represent a broad range of diversity of germplasm available in the Hungarian collection. Hungarian bean samples were scored with the same SSR markers as has been used in the GCP for the characterisation of beans from different origin, and the allelic diversity in the Hungarian bean germplasm will be compared to the global diversity in beans already characterised by the GCP.

In addition the 100 Hungarian samples, 2-2 controls were also included in the characterisation to ensure comparability of the data as well as to identify both the Mesoamerican and Andean gene pools. Out of 52 SSR markers, 51 were polymorphic and they identified 6 alleles in an average, spanning from 2 to 34 alleles/marker. To maintain strict conditions in pair wise comparisons, 90 samples were kept in the final analysis. 36 genotypes were clustered with the Mesoamerican controls (DOR364 and ICA Pijao) and 48 genotypes with the Andean controls (G19833 and CALIMA), while 6 samples represented hybrid genotypes and spread between the two clusters. Although the controls clearly marked the two kinds of gene pools, they stood out the clustered genotypes within both clusters indicating diverged allelic compositions in the Hungarian genotypes (Fig1). Preliminary analyses of the marker data also revealed that the Hungarian bean germplasm contains substantial diversity, which will support national programme's breeding objectives as well as allow the participation in international programmes aiming at bean improvement and conservation. Further analyses, combining the SSR data of the Hungarian bean accessions with the data from the GCP's bean reference set, are in progress. This will allow the Hungarian bean diversity to be directly compared to the global analyses of bean diversity.



**Fig 1. Radial representation of relationship between 90 Hungarian bean genotypes.**

2 Mesoamerican (DOR364 and ICA Pijao) and 2 Andean (G19833 and CALIMA) controls represent the two major gene pools. Dendrogram was created by using DARwin software package: Dissimilarities were calculated with bootstrap analysis (1000x) and the UnWeighted Neighbor-Joining method was applied to draw the tree.

### 103. G4008.35: Toolbox of available molecular markers useful for marker assisted selection in GCP crops

*March 2008–March 2010*

#### Principal Investigator

Veerle Van Damme, Consultant; veerle.vandamme@gmail.com  
Cali, Colombia; (57) 2-5552650

#### Collaborating institutions and scientists

- GCP: M Carmen de Vicente; Humberto Gómez Paniagua

Developing countries harbor the majority of the plant genetic resources for food and agriculture. These genetic resources contain numerous genes and alleles possibly useful to overcome most of the challenges of modern agriculture. Genomics has helped in identifying, targeting and deploying useful genes. Molecular markers greatly facilitate the selection of traits that are often difficult and time-consuming to detect based on phenotype. As such, marker assisted selection (MAS) enables speeding up the incorporation of these valuable traits.

Agricultural researchers and plant breeders, in particular in developing countries, face difficulties concerning access to up to date scientific information on useful molecular markers, as the latest discoveries are often scattered in numerous, expensive peer-reviewed journals or in databases of unknown existence to many. If access to information is not a problem, the avalanche of information can be one, as the information offered through digital resources is not always reliable, can be overwhelming and does not provide guidance for its appropriate use.

This project deals with the development of a toolbox providing free and easy access to information of all publicly available molecular markers (SSR, SCAR, SNP, STS) for simply inherited traits and QTLs useful in marker assisted selection in 19 food security crops, i.e. *Musa* spp., barley, bean, cassava, chickpea, coconut, cowpea, faba bean, groundnut, lentil, maize, millet, pigeonpea, potato, rice, sorghum, sweet potato, wheat and yam. The toolbox will include validated markers as well as markers in process of validation, the latter to avoid duplicated research among scientists. The activity will compile information available in internet sources, public databases, papers and through consultation of experts whereby the compilation will be validated again by actual crop breeding experts. Results will be made available via Internet as a global public good by the way of a searchable database. Its features will be described in a peer-reviewed publication. By sharing the latest advances in molecular plant breeding, the toolbox is an important step into supporting modern agriculture for the benefit of the poor in developing countries.

#### **104. G4008.36: Getting the focus right: Food crops and smallholder constraints**

*April 2008–February 2009*

##### **Principal Investigator**

John Dixon, CIMMYT Impacts Targeting and Assessment Unit; j.dixon@cgiar.org  
Apdo. Postal 6-641, 06600 Mexico, D.F., Mexico; (52) 595 952 1900; Fax: (52) 595 952 1983

##### **Collaborating institutions and scientists**

- CIAT: Glenn Hyman
- Food crop breeders, crop management, economics and GIS specialists in several CG Centers including CIAT, IRRI, CIMMYT, ICRISAT, IITA, WARDA
- Numerous NARES institutions and staff in South and East Asia, Sub-Saharan Africa, Latin America

##### **Objectives**

- Participatory assessment of smallholder production constraints for six food crops in 15 farming systems that exhibit extreme poverty
- Consultative identification of potential solutions to the more important constraints, especially those with a genetic/germplasm component.



## **Context**

Drought has been identified as a major priority for food crop improvement programmes in international agricultural research. However, it is generally accepted that a variety of other “secondary” constraints limit productivity in good seasons, as well as in drought years. The well known CABI data base contains comprehensive but rather general information on losses and distribution. However, few of these studies provide sufficient contextual information to extrapolate the results across zones, seasons and years.

In recognition of the complexity of factors which affect the improvement of food crop yields and productivity under smallholder conditions in different farming systems throughout the developing world, the relative importance of abiotic, biotic, crop management and socioeconomic constraints needs to be assessed in physical and economic terms. A systematic tapping of the tacit knowledge of experienced research and development practitioners is needed to provide valuable information on the relative importance of different production constraints and traits. The results of this study can be a checklist and guide to those involved in food crop breeding and related research by prioritizing key traits for their improvement in each of the systems, and by eliciting ideas on solutions to key constraints, especially those that may offer opportunities for the GCP.

## **Focus**

The study covers six high priority food crops for GCP – wheat, rice, sorghum, cassava, cowpea and chickpea – in 15 high priority farming systems for the GCP, featuring high levels of poverty. Five of these systems are in Sub-Saharan Africa, 5 in South Asia, 4 in East Asia Pacific and 1 in Latin America. The emphasis is on identifying and quantifying important constraints that contribute to the smallholder farm yield gap defined as: Best achieved yield on farm - Average yield on farm. Important constraints are being identified within four broad categories: Abiotic, Biotic, Management-related and Socio-economic.

## **Methods**

The study uses a modified “Delphi” methodology. Three rounds of interaction take place with panels of experts on the crops in the farming systems, primarily using a series of questionnaires administered through regional focal points and Email contacts. This allows convergence on important constraints, issues and solutions for the crops in the farming systems. Panels of experts include members from agricultural research (plant breeders, agronomists, agric economists), extension services, NGOs, seed houses, etc.

Round 1 questionnaires were sent for all six crops and information returned by panelists on yield gaps, important constraints, yield losses, their spatial and temporal distribution and effects on income. Round 2 has been developed and it will get confirmation of important constraints and magnitudes of losses, and examine constraint interactions. Round 3 will elicit ideas on solutions and opportunities to alleviate the constraints, particularly identifying those that can be addressed by the GCP. This could be an avenue to be further developed to address the needs of GCP in the future.

## **Links**

Our current project, *Getting the focus right: Food crops and smallholder constraints*, is developed from an initial study on constraints with maize that was completed by

CIMMYT for the GCP in 2007. This work also builds from the previous research funded by the GCP on identifying farming systems with high degrees of poverty (child stunting) and drought incidence undertaken by CIAT (Hyman et al 2008).

### Reference

Hyman G, Fujisaka S, Jones P, Wood S, de Vicente C and Dixon J (2008). Strategic approaches to targeting technology generation: Assessing the coincidence of poverty and drought-prone crop production. *Agricultural Systems* 98:50-61.

## **105. G4008.37: PhD in Plant Breeding at the West Africa Centre for Crop Improvement**

*January 2009–December 2012*

### Principal Investigator

Eric Danquah, WACCI; edanquah@wacci.edu.gh  
University of Ghana, Box LG 68, Legon, Ghana

### Collaborating institutions and scientists

- Biotechnology Centre, University of Ghana; Kwame Offei
- College of Agriculture and Life Sciences, Cornell University; Vern Gracen

It has long been recognised that capacities in plant breeding, including both conventional and modern technologies, in most developing countries are neither sufficient nor properly integrated to fully capture the benefits of the plant genetic resources that are conserved. Today, sub-Saharan Africa remains the only region that may not meet the millennium development goal of eradicating extreme poverty and hunger by 2015. New high-yielding varieties of staple crops with tolerance to biotic and abiotic stresses can help provide food security for increasing populations in the sub-Saharan Africa. A critical mass of a new generation of plant breeders with knowledge in both traditional field based selection methods and emerging laboratory based tools and techniques is needed to develop and provide the necessary high yielding varieties to farmers.

The University of Ghana has received a project support grant of \$5.78 from the Alliance for a Green Revolution in Africa to establish a West Africa Centre for Crop Improvement (WACCI). WACCI, a collaboration between the University of Ghana and Cornell University, started operating in the University of Ghana in June 2007 as an autonomous institution in the College of Agriculture and Consumer Sciences. WACCI is committed to the training of plant breeders with skills in genetic improvement of the staple crops of the west and central Africa sub-region. Plant breeding is an integrative science that combines the knowledge, information and expertise from a range of disciplines to produce scientists with the capacity to undertake research for germplasm enhancement and development of improved cultivars of the staple crops. The first cohort of eight students enrolled in February 2008. The students will undertake two years of course work in the University of Ghana and three years of field research in their research institutions. WACCI would increase its enrollment to ten students a year to accommodate two additional students who would be supported by the Generation Challenge Programme in 2009 and 2010.

## **Project description**

### **Mission**

To train plant breeders with capacity to lead the conversion of genetic and molecular discoveries into innovative solutions that result in improved varieties to benefit agriculture in West Africa.

### **Vision**

Agriculture in West Africa underpinned by innovation in conventional and molecular plant breeding.

### **Overall goal**

The overall goal of the WACCI is to train the next generation of West and Central African plant breeders in an African university, to breed crops in national agricultural research stations, for production systems in the two sub-regions.

### **Participating countries**

Currently, WACCI is focussing on five countries, Burkina Faso, Ghana, Mali, Niger and Nigeria. WACCI may consider candidates from other West and Central African countries subject to an agreement between WACCI and the GCP.

### **Focus**

The coursework and research will focus on the following food crops: Cereals: maize, rice, sorghum, millet, Root and tubers: cassava and sweet potato, Legumes: cowpea & groundnut, Vegetables: tomato & pepper.

The students will undertake a five-year PhD programme. The first two years will involve coursework at the University of Ghana. Students will then proceed to their home institutions to undertake their PhD research work over a three year period. They will return to the University of Ghana in the last quarter of the fifth year to submit their PhD thesis for examination. Successful students will be awarded a PhD degree in Plant Breeding by the University of Ghana.

### **Output**

Well-trained PhD graduates with the knowledge and skills to develop improved varieties

## **106. G4008.38: Fellowships and travel grants 2008**

### **Principal Investigator**

Carmen de Vicente, GCP, [c.devicente@cgiar.org](mailto:c.devicente@cgiar.org)

Apdo. Postal 6-641, 06600 Mexico, D.F., Mexico, Phone: +52 55 5804 2004

Eight **Fellowships** were offered. The maximum award per fellow was up to US\$25,000 (travel, living expenses, accommodation, laboratory consumables, and conference participation).

The Fellowship Programme started in 2005 and was based on a call for proposals with the following principles:

- 1) Proposals should deal with one of the GCP crops
- 2) They must be linked with ongoing research supported by the GCP, either by competitive or commissioned grants.
- 3) The proposal should present evidence that the fellowship will be oriented towards training of the candidate and improving capacity at the home institution, *rather than to provide extra funding for ongoing projects (\*)*.
- 4) The majority of the proposed research must be done at one of the GCP Consortium centers, or participating institutions in a GCP supported research project.

Applications were invited from crop science researchers from developing country research institutions (National Agricultural Research Systems at large). Applicants were to hold at least a Master of Science degree (MSc), or equivalent, in a relevant subject area. Applicants also had to demonstrate that they were engaged in a related ongoing research activity in their home country, and they were expected to return to their home institution and contribute to its research and education programmes.

Priority was given to scientists from National Agricultural Research Systems **already** involved in GCP research projects.

(\*) For 2008, a small twist was added. Principal Investigators of ongoing GCP projects were contacted with the request to propose a research subject, already part of the GCP project or complementary to it, for which they were willing to host a fellow for a training experience. As sufficient subjects were not received from PIs by November 15<sup>th</sup> 2007, the Call for Applications to the Fellowship Programme was extended in January 2008 to 29<sup>th</sup> February 2008, this time targeting a selection of research subjects made by the Management Team. The Call included a description of research subjects, the minimum desired qualifications of the candidate(s), the and proposed duration of the fellowship (depending on each subject), among other details. As customary with past calls, an application form plus other supporting documents was required for the selection of candidates. The applicants were asked to present evidence that their ongoing work was related to the subject of choice and that the learning would be used to benefit his research. Once the selection of winners was made, fellows were requested to prepare a work plan in collaboration with the PI.

The **Travel Grant Programme** is meant to foster linkages within current GCP projects to advance research while providing training opportunities for developing country scientists.

Travel grants and participation in conferences offer new occasions to start collaboration or trigger an interest on a GCP-related research project. As a consequence, the community of skilled and knowledgeable collaborators of the GCP in developing countries increases.

Three types of grants were offered:

**a. Hands-on training opportunities**

This grant may be requested to visit a GCP Consortium Institution, a collaborating institution, or an independent advanced research institution to have a **hands-on training** experience related to concepts and/or techniques useful or necessary for the advancement of the GCP research. It is not oriented to support conference participation. The applicant should belong to an institution from a developing country (NARS or Academia) that is either a member of the GCP Consortium or is working in collaboration with a GCP Consortium Institution.

**Eight** grants were available and the maximum grant award was 5,000 USD, which was intended to cover travel, accommodation, living expenses, and laboratory consumables, if needed.

**b. Participation in GCP organised workshops**

The GCP may take advantage of conferences or scientific events to organise workshops for specific purposes, mainly to bring together researchers working in similar subjects, in similar crops or in the same region. The purpose is to promote linkages among researchers at all levels to disseminate the benefits of the science being conducted and simultaneously enhance the number of potential users of GCP products.

**c. Participation in the GCP Annual Research Meeting**

These grants are meant to invite selected researchers from National Programmes working in the region where the ARM takes place. Participants should be already engaged with the GCP and may be requested to present results of the research being conducted in the home institution.

In total there were 5 Fellowship winners and 5 Travel grant winners in 2008. To view details on the winning awardees, please visit the SP5 website at <http://www.generationcp.org/sp5/>

## **107. G4008.39: Capacity-building à la carte 2008**

### **Principal Investigator**

Carmen de Vicente, GCP, [c.devicente@cgiar.org](mailto:c.devicente@cgiar.org)

Apdo. Postal 6-641, 06600 Mexico, D.F., Mexico, Phone: +52 55 5804 2004

This project relates to a capacity building concept, *à la Carte*, that seeks to identify and provide tailored capacity building to a select group of teams of applied researchers at developing country research programmes who will benefit significantly from short-term, personalised training and support. For each team selected to participate in this programme, a customised plan is proposed comprised of training events in the form of formal training at academic institutions or at events organised by the GCP, mini-grants for small equipment, hands-on research opportunities in advanced research institutions, and the in-situ assistance of technical experts.

This scheme provides opportunities for researchers to obtain high-quality training and follow-up support, and thereby mobilises a community of well-trained and well-prepared researchers to carry on GCP research.

In practice, the project targets short- to medium-term support, providing guidance to entice researchers to stay in their countries, hoping they become self-sufficient to attract further support in the long term.

The programme is linked to current GCP research projects and complementary to GCP established activities to strengthen national research institutions.

In 2008, a Call opened from from 1<sup>st</sup> December 2007 to January 31<sup>st</sup> 2008 with capability to accommodate 10 grants. In the end three teams were selected, with more details given below.

***107.01. G4008.39.01: Capacity-building à la carte 2008–Enhancing MAS capacity for salt-stress rice breeding in Bangladesh  
April 2008–April 2009***

**Team Leader**

MA Salam, BRRI, Bangladesh

**Collaborating institutions and scientists**

- BRRI: M Alamgir Hossain, M Rafiqul Islam (Plant Breeding Division), M Sazzadur Rahman (Plant Physiology Division)
- University of Dhaka, Bangladesh: Zeba I Seraj (Department of Biochemistry and Molecular Biology)
- IRRI: Abdelbagi Ismail (Senior Plant Physiologist), Michael Thomson (Post-doctoral fellow)

The application of molecular markers to increase the efficiency of breeding for varietal improvement targeted to problem soils is of vital importance for Bangladesh. Gradually increasing salinity levels in the south of Bangladesh is a major concern, particularly because it affects resource poor farmers living in those areas. About one million hectares of land is affected by different levels of salinity in the coastal areas of Bangladesh. BR11 and BRRI dhan28 are two popular varieties cultivated in Bangladesh for rainfed lowland and irrigated ecosystems, respectively, but those are sensitive to salinity. FL378 is an RIL having the Saltol QTL for salinity tolerance but is not well adapted to Bangladesh conditions. To introgress Saltol from FL378 into BR11 and BRRI dhan28, we made backcrosses using BR11 and BRRI dhan28 with FL378. Marker-assisted backcrossing activities are being pursued through a competitive (Project 2) project which is now coming to an end, and a commissioned GCP project just started with collaboration of IRRI and Dhaka University: population development and salinity screening are done at BRRI and molecular selection is being performed at Dhaka University. BRRI has good facilities for population development and phenotyping for salinity tolerance but only partial facilities for molecular analysis and application of MAS. At the same time, BRRI has experienced scientists trained in molecular marker techniques at IRRI, but due to the

lack of adequate facilities they cannot contribute their expertise in the current GCP activities. Strengthening BRRRI molecular research facilities through acquiring the additional equipments that are currently missing (PCR machine, electrophoresis unit with power pack, centrifuge and electronic pipette) will help equip the laboratory of BRRRI to undertake an effective MAB system and deliver the outputs of the ongoing GCP projects more efficiently and, in the long run it will contribute substantially to enhance the capacity of BRRRI to incorporate marker assisted breeding in our current breeding programmes using QTLs of agronomic importance. This current capacity building support grant therefore aims to equip the BRRRI laboratory for DNA marker technology and to support scientific exchanges between BRRRI, DU and IRRI for further training and technical backstopping to support and complement the ongoing GCP funded projects.

107.02 G4008.39.02: Capacity-building à la carte 2008–Improving capacity for phenotyping for abiotic and biotic stress in Burkina Faso  
*April 2008–April 2009*

**Team Leader**

Issa Drabo, INERA, Burkina-Faso

**Collaborating institutions and scientists**

- UC–R: Jeffrey Ehlers, Timothy Close, Philip Roberts
- IITA, Kenya: Din-Jong Kim
- IITA, Nigeria: Satoru Muranaka, Ousmane Boukar

Cowpea is a major grain and fodder crop in Burkina Faso and one of the few crops adapted to the poor soils, low rainfall and high temperatures found in most of the country. Despite its rusticity, productivity is decreasing due to drought spells and the pressure from pests. With the recent funding of a large GCP project targeting development of improved genomic resources in tropical legumes including cowpea with emphasis on drought tolerance, it is important that capacity exists to properly phenotype germplasm and genetic populations for drought tolerance. Therefore to meet the goals of the TL1 project and better characterise the 500 genotypes for their responses to abiotic stresses (drought and heat) and biotic ones (thrips, nematodes, Fusarium wilt, and bacterial blight), facilities for phenotyping need to be improved. The background of drought research in Burkina Faso is based on multi-location trials and breeding for agronomical traits. Therefore capabilities need to be strengthened. Precise and accurate phenotyping will be needed to take advantage of molecular markers being identified under the TL-1 project. Equipment to precisely link the plants physiological and agronomical responses to water available in the soil is needed.

**Objectives:**

1. Strengthen capacity for drought phenotyping
2. Strengthen capacity for pest control

**107.03 G4008.39.03: Capacity-building à la carte 2008–Improving capacity for phenotyping for abiotic and biotic stress in Senegal**

**April 2008–April 2009**

**Principal Investigator**

Ndiaga Cisse, ISRA; ncisse@refer.sn

Institut Sénégalais de Recherches Agricoles (ISRA–CNRA) BP 53, Bambey,  
Senegal; (221) 76 580 92 37 ; Fax : (221) 33 973 63 48

**Collaborating institutions and scientists**

- UC–R: Jeff Ehlers

Cowpea is a major grain and fodder crop in Senegal and one of the few crops adapted to the poor soils, low rainfall and high temperatures found in most of the country. Despite its rustiticity, productivity is decreasing due to drought spells and the pressure from pests. With the recent funding of a large GCP project targeting development of improved genomic resources in tropical legumes including cowpea with emphasis on drought tolerance, it is important that capacity exists to properly phenotype germplasm and genetic populations for drought tolerance. Therefore to meet the goals of the TL1 project and better characterise the 500 genotypes for their responses to abiotic stresses (drought and heat) and biotic ones (thrips, nematodes, *Fusarium* wilt, and bacterial blight), facilities for phenotyping need to be improved or repaired. The background of drought research in Senegal is based on multilocation trials and breeding for agronomical traits eventhough some collaborative work was initiated between ISRA-CNRA and CERAAS (Regional Centre of Studies for Drought Adaptation) to integrate physiological and molecular parameters with the breeding programme. These capabilities need to be strengthened. Precise and accurate phenotyping will be needed to take advantage of molecular markers being identified under the TL-1 project. Equipment to precisely monitor soil moisture profile is needed to link the plants physiological and agronomical responses to water available in the soil.

**Objectives :**

1. Streighten capacity for drought phenotyping
2. Streighten capacity for pest control

**108. G4008.40: Workshop on “Reference sets of food crop germplasm for international collaboration”**

**January 2008–December 2008**

**Principal Investigator**

Jean-Christophe Glaszmann, Agropolis-CIRAD, [jean-christophe.glaszmann@cirad.fr](mailto:jean-christophe.glaszmann@cirad.fr)  
Agropolis-CIRAD, TA A-DIR PS3, Boulevard de la Lironde, 34398 Montpellier  
Cedex 5, France

M Carmen de Vicente, Generation Challenge Programme, [c.devicente@cgiar.org](mailto:c.devicente@cgiar.org)  
Carretera México-Veracruz, Km 45, El Batán, Texcoco, México.

**Collaborating institutions**

- System wide Genetic Resource Programme (SGRP), Rome, Italy
- Global Crop Diversity Trust (GCDT), Rome, Italy
- Global Partnership Initiative for Plant Breeding Capacity Building (GIPB), Rome, Italy



Access to genetic diversity available in large crop germplasm collections requires identification of representative samples with smaller size to make them amenable to a suite of purposes: screening of traits, evaluation of phenotypic diversity, evaluation of combining ability, assessment of molecular diversity, etc. Moreover, integrating diverse types of characterisation on the same materials makes it possible to assess correlations among traits and investigate gene effects such as epistasis and pleiotropy. Passport data enable selecting based on eco-geographic information; molecular markers offer means to further refine assessment of relatedness and to reduce sample size. Use of standardised methods yields data that can be compared across materials, laboratories and time, providing a durable momentum to enrich global understanding and representativeness.

The first phase of the GCP has yielded massive data sets featuring SSR diversity (12 to 50 loci) of large germplasm samples (300 to 3000 accessions). This resulted in the identification of reference samples of 50 to 500 accessions to be handled as genetic stocks, for which data have been ascertained for a subset of high quality SSR markers.

Altogether this shaped into a major GCP product: germplasm reference samples with validated data of reference markers, accessible as a global public good in a robust form. It is hoped that these samples will be widely shared and used, so that new data can be integrated in order to derive biological understanding useful for germplasm diversity management and use.

In this workshop, all these steps and aspects will be described and discussed, as well as the perspectives and the mode of organisation necessary for taking full advantage of this product. It will be an opportunity for coordination among various players engaged in germplasm management in international programmes.

It will be held in Montpellier, France, on November 13-17, 2008.

---

## FOCUS PROJECTS

### **109. GCP/Rockefeller project G4005.69.01 (CB19a/RF-FS022): Developing and disseminating resilient and productive rice varieties for drought-prone environments in India**

*March 2005–February 2008; no-cost extension to February 2009*

#### **Principal Investigator**

Arvind Kumar, IRRI; akumar@cgiar.org

DAPO BOX 7777, Metro Manila, Philippines; (63) 2-580-5600; Fax: (63) 2-580-5699

#### **Collaborating institutions and scientists**

- IRRI: R. Serraj; T. Paris; S. Haefele; R. Anitha; G. Atlin\*
- IGAU: S.B. Verulkar; P. R. Dongre
- CRRI: O.N. Singh; P. Swain; L. Bose
- CRRI Research Station, Hazaribag: PK Sinha; NP Mandal
- NDUAT: JL Dwivedi
- UAS: S Hittalmani
- TNAU: R. Chandrababu; S. Robin
- BAU: B.N. Singh; R.L. Mahato
- Barwale Foundation, Hyderabad: H.E. Shashidhar; Abhinav Jain

\* Present address: CIMMYT, Apdo Postal 6-641, 06600, Mexico

#### **Research activities and progress at IRRI and collaborating institutes**

The present project funded by the Rockefeller Foundation-Generation Challenge programme supports the activities of IRRI-India drought breeding network that links IRRI with eight institutions in India, the region with the largest rainfed drought prone area (Huke and Huke, 1997). In 2007, on-farm participatory varietal selection trials with promising lines identified in previous years and on-station evaluation of 352 advanced breeding lines were conducted under control and drought stress conditions at all the partners' sites.

##### **1.1 On-farm participatory varietal selection (PVS) trials**

On-farm participatory varietal selection, using mother trials with lines found promising in on-station trials in the previous seasons, were conducted in wet season of 2007 at Raipur, Cuttack, Hazaribag, and Faizabad. At maturity, at each location, several men and women farmers as well as researchers were invited to vote and select the breeding lines of their choice. A preference rating was generated for each variety by expressing the number of effective (“liked” minus “disliked” votes) votes cast for that variety as a proportion of the total number of votes.

At Raipur, women voted for IR55419-04 as the best, followed by IR64 × Mahamaya cross line and Danteswari × Deshi Lal Dhan line, whereas men preferred Mahamaya × CT 9993, followed by IR55419- 04 and SL-100. The combined analysis showed that

Mahamaya × CT 9993 cross was the best preference, followed by IR55419-04. During the focused group discussions, farmers indicated that they preferred Mahamaya × CT 9993 mainly for its high yield attributing characters. However, they would prefer to grow IR55419-04 for home consumption because of its better grain quality. The choice of researchers matched the selection by farmers, as Mahamaya × CT 9993 line was also chosen to be the best, followed by IR55419-04 by researchers. At Cuttack, a largest participation of the farmers was witnessed. Women voted for IR78875-131-B-1-4, followed by IR55419-04 and IR78875-53-2-2-2, and men voted for the latter two as the best varieties, followed by IR80312-6-B-3-2-B. The combined analysis showed that IR55419-04 and IR78875-53-2-2-2 were the most preferred varieties. At Hazaribag, both women and men voted for IR74371-70-1-1 and IR74371-54-1-1 as their most preferred variety. At Faizabad, both men and women voted for RR143-2-2 and IR74371-70-1-1 as equally the best among the 14 tested entries.

### **Advanced yield trials**

A total of 136 entries contributed by both Indian centers and IRRI were screened in advanced yield trials (AYT) in the wet season of 2007. Entries were grouped into three duration groups: less than 100 days (LT100), 100–120 days (100–120), and greater than 120 days (GT120). The AYT was conducted at three stress levels: control, reproductive stress, and rainfed stress. There was one observation yield trial (OYT) that included 216 entries of all durations coming from nominations by collaborators and IRRI. Among the breeding lines of less than 100 days duration, RR 345-2 and RR 433-2 were found promising. In the 100-120 days duration trial, IR72667-16-1-B-B-3, IR55419-04, IR74371-46-1, and Tripuradhan seemed highly promising. They produced about 5 t ha<sup>-1</sup> under control conditions and 2 t ha<sup>-1</sup> under severe stress conditions across trials. Other lines that were high yielding across all treatments were IR70215-70-CPA-3-4-1-3, ARB 8, IR79899-B-179-2-3, DGI75, IR7899-B-179-2-3, and IR78875-23-B-2-B. In GT120 trial, IR78877-181-B-1-2, IR78875-207-B-3-B, and Swarna/IR42253-64-113 performed very well at all levels of stress.

Overall, three years of collaborative testing of breeding lines under drought and control situations has identified breeding lines that have yield potential as high as that of presently grown mega-varieties IR64 and Swarna and yielded 1.0 t ha<sup>-1</sup> higher than them under severe drought stress. In 2008, IR74371-70-1-1, a breeding line from IRRI, has been identified for release as a drought tolerant variety by the Central Variety Release Committee of India for the states of Orissa, Tamil Nadu, and Madhya Pradesh. Promising drought-tolerant breeding lines IR55419-04, IR74371-54-1-1, IR74371-46-1-1, and IR72667-16-1-B-B-3 have been entered into national testing under the All India Coordinated Rice Improvement Programme (AICRIP).

### **References**

Huke RE and Huke EH (1997). Rice area by type of culture: south, southeast, and East Asia. IRRI, Los Baños, Philippines.

**110. GCP/Rockefeller project G5005.69.02 (CB19b/RF-FS029):  
Pathway dissection and candidate gene identification of  
drought tolerance in rice by a forward genetics approach**

*March 2006–March 2007; no-cost extension to March 2008*

**Principal Investigator**

Zhi-Kang Li, Institute of Crop Sciences, CAAS, z.li@caas.net.cn

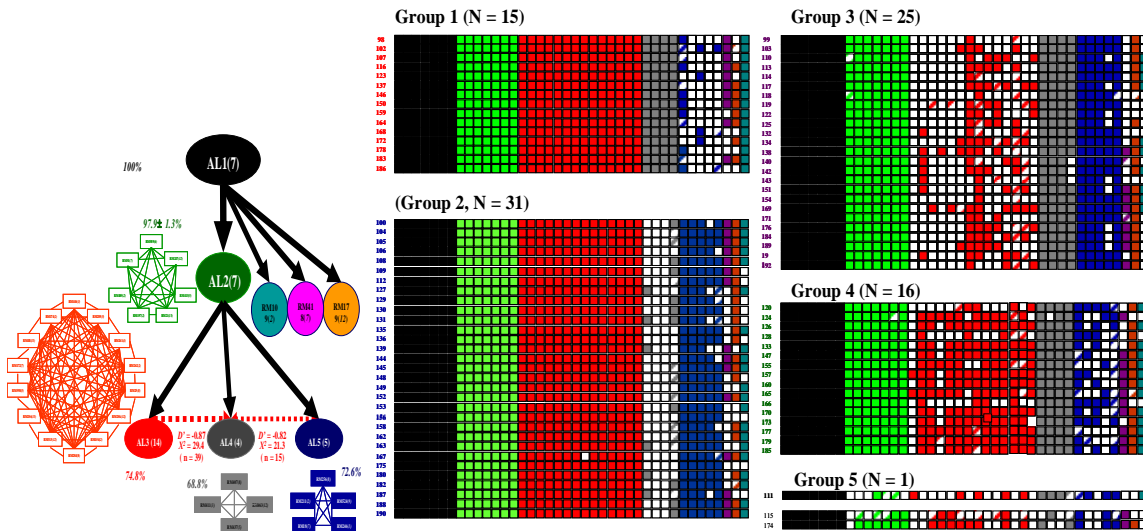
Institute of Crop Sciences, Chinese Academy of Agricultural Sciences (CAAS),  
12 South Zhong-Guan-Cun St., Beijing 100081, China; (86) 10-62136040; Fax:  
(86) 10-68918559

**Collaborating institutions and scientists**

- Peking University and Yale University: Xing-Wang Deng
- IRRI: Racid Saraj (replacing R Lafitte)

Drought tolerance (DT) in rice is a complex trait involving large numbers of loci and complex genetic and physiological/morphological mechanisms. In this project, we took a forward genetics approach and performed a series of genetic, phenotypic and genomic experiments to get insights into the genetic and molecular basis of DT in rice.

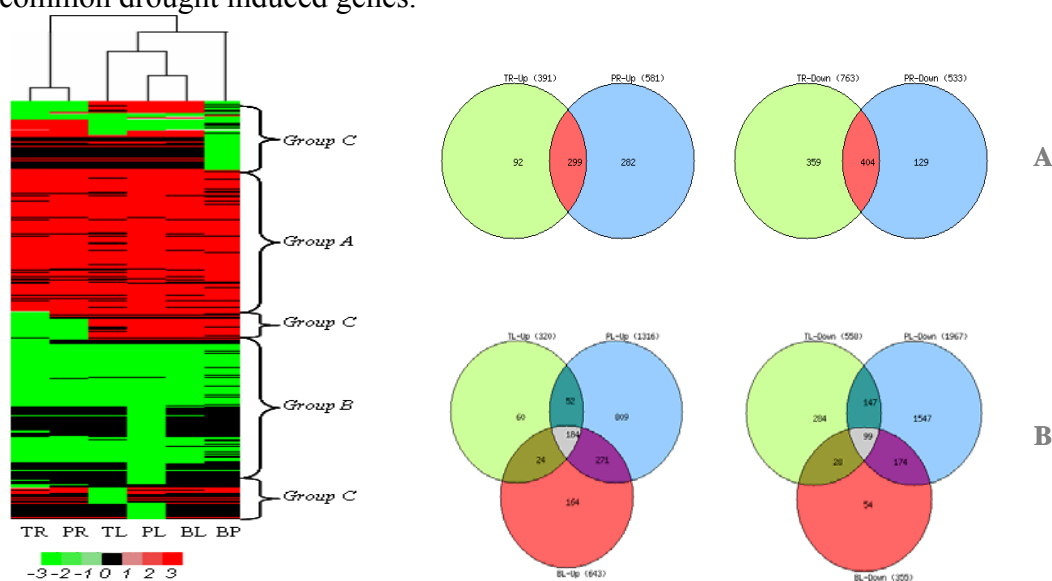
Ninety DT F<sub>2</sub> progeny selected from a cross between 2 DT IR64 ILs segregating for 40 DT QTL regions were genetically analyzed with DNA markers. One dramatic result was that all segregating DT loci in the 90 DT F<sub>2</sub> progeny appeared to be under strong epigenetic control with 14 loci fixed or nearly fixed at one of the alleles, the remaining 26 loci fixed at either of the parental alleles, and the heterozygotes at all segregating loci virtually eliminated. Linkage disequilibrium analysis revealed a complex genetic network and four major group genotypes (GG) (Fig. 1a and 1b)



**Fig. 1. a: The genetic network containing 40 DT QTLs detected in 90 DT F<sub>2</sub> progeny; b: the graphical genotypes of the 4 major group genotypes.**

Replicated phenotyping experiments in 3 consecutive seasons under both drought stress and normal conditions indicated that all 4 GGs had significantly improved DT than IR64, but GGs 1, 3 and 4 had significant yield penalty under normal irrigated conditions, whereas GG2 yield 20-50% more than IR64 even under normal conditions.

The global gene expression patterns of 4 GGs under drought and normal conditions were analyzed using the Affymetrix rice genome array containing 48,564 *japonica* and 1,260 *indica* sequences, revealing a total of 5,284 genes that were differentially expressed under drought stress, including 261 transcription factor genes. Bioinformatic analyses and pairwise comparisons between different GGs and between GGS and IR64 (the genetic background of the GGs) are being performed to explore important cis-elements and putative gene networks related to DT in rice. Preliminary results have uncovered a *cis*-element containing special CGCG box that were over-presented in the upstream of 55 common drought induced genes.



**Figure 1. A. Hierarchical cluster analysis of six tissue types (columns) and all DEGs under drought stress (rows). TL, PL and BL indicate leaves at the tillering, panicle elongation and booting stages, respectively; TR and PR indicate roots at the tillering stage and panicle elongation stage; BP is young panicles at the booting stage. Group A, B and C indicate different sets genes with specific expression patterns. B: Venn diagram of up- and down-regulated genes under drought stress at different developmental stages.**

The 4 GGs and IR64 were also phenotyped for a wide range of physiological and morphological traits, including their responses to different plant hormones in multiple environments. The data analyses are in the progress to link information from all experiments to identify important candidate genes and pathways controlling DT in rice.

---

**111. GCP/Rockefeller project G4005.69.03 (CB19c/RF-028):  
Innovative and integrated approaches to improve the tolerance  
of maize to water-limited environments**

*April 2005–March 2007; no-cost extension to June 2008*

**Principal Investigator**

Yunbi Xu, Genetic Resources and Enhancement Unit, CIMMYT, [y.xu@cgiar.org](mailto:y.xu@cgiar.org)  
Apdo. Postal 6-641, 06600 Mexico, D.F., Mexico. Tel: +52(55) 5804-2004  
Fax: +52(55) 5804-7558

**Collaborating institutions and scientists**

- Genetic Resources and Enhancement Unit, CIMMYT: Carlos Martinez, Debra J. Skinner, Alan F. Krivanek and Jonathan H. Crouch
- Maize Research Institute, SAU: Shibin Gao
- Institute of Crop Science, CAAS: Zhuanfang Hao, Shihuang Zhang, Jiankang Wang

Towards the development a full functional marker-assisted selection (MAS) system for an innovative and integrated approach to improvement of maize tolerance to water-limited environments, we have made progress in the following areas: (1) identifying the bottlenecks and constraints associated with public sector MAS programmes, (2) establishing two strategies (seed DNA-based MAS and selective genotyping) to address the primary bottlenecks, (3) developing a reversed breeding-to-genetics strategy to discover and pyramid favorable alleles for drought tolerance, and (4) bridging the gap between conventional and MAS-based breeding programmes by carrying out an interdisciplinary molecular breeding capacity building workshop as the basis for creating a molecular breeding community of practice.

Converting promising publications into practical applications in MAS requires the resolution of many logistical and genetical constraints. A high proportion of published markers fail at one or more of the translation steps from research arena to application domain. The frequency of success in this translation process is increasing due to developments in the following important fields: sequence-based assays; rapid and efficient large-scale QTL mapping; genotype-by-environment interaction and epistatic analysis; the emergence of high throughput low unit cost genotyping technologies, effective sampling and data acquisition systems, integration of fingerprinting, mapping, and MAS technologies, and developments in breeding informatics and decision-support tools (Xu and Crouch, *Crop Science* 48: 391-407, 2008).

An optimised genotyping method using endosperm DNA sampled from single maize seeds was developed, which can be used to replace leaf DNA-based genotyping for both genetic studies and breeding applications. A similar approach is likely to be suitable for all plants with relatively large seeds. Part of the endosperm was excised from imbibed maize seeds and DNA extracted in 96-tube plates using individuals from eight F<sub>2</sub> populations and seven inbreds. The quality of the resultant DNA was functionally comparable to DNA extracted from leaf tissue. Extraction from 30 mg of endosperm yields 3-10 µg DNA, which is sufficient for analysis of 200-400 agarose-gel PCR-based

markers, with the potential for several million chip-based SNP marker analyses. A substantial advantage of this approach is that it can be used to select desirable genotypes before planting and provides an opportunity for dramatic improvements in the efficiency and selective gain of breeding systems (Gao et al., 2008, *Molecular Breeding*, available online).

Past applications of selective genotyping and pooled DNA analysis have been confounded by the use of small total and tail population sizes and insufficient marker density, which results in a high probability of false positive marker associations. Our simulation studies indicate that when these issues are resolved selective genotyping and pooled DNA analysis can be effectively used for genetic mapping of quantitative trait loci (QTL) with relatively small effects, as well as linked and interacting QTL. Using phenotypic extremes from diverse germplasm, it is theoretically possible that one 384-well plate could be designed to cover almost all major gene/QTL controlled agronomic traits of importance in a crop species. This “all-in-one plate” approach is feasible in all species where high density marker coverage is available. In CIMMYT, over 1600 maize genotypes have been collected from drought tolerance breeding programmes worldwide. A key set of the collection has been evaluated for drought tolerance at both vegetative stage (by measuring biomass changes after drought treatment) and reproductive stage (the final yield harvested from the drought plot). They have been genotyped using a maize 1536 SNP chip, allowing us for the first time to test the feasibility of a one-step simultaneous marker-trait association analysis (Xu et al., *Proceedings of 5<sup>th</sup> International Crop Science Congress*, April 13-18, 2008, Jeju, Korea). A reverse breeding-to-genetics approach has been proposed through a modified marker-assisted recurrent selection scheme to speed up discovering and pyramiding favorable alleles of different sources.

A total of 20 participants attended the Molecular Breeding Capacity Building Workshop, which was held in Nairobi, Kenya, June 8-14, 2008. This workshop was jointly funded by GCP and drought tolerance maize for Africa (DTMA). Participants include seven CIMMYT DTMA scientists, eight national DTMA scientists, and five GCP selected sorghum breeders in Africa. The major objective was to lay the foundations for establishing a Maize Drought Tolerance Molecular Breeding Community of Practice (CoP) in Africa, including a group maize breeders carrying out marker-assisted germplasm evaluation, genetic mapping and MAS. The workshop achieved the following objectives: (1) narrowed the gaps between molecular biologists and field breeders associated with MAS by training participants in basic theories, molecular breeding practice in the private sector and take-home examples; (2) established multidirectional communications among molecular biologists, molecular breeders, field breeders, trait specialists and molecular breeding support group; (3) established molecular breeding working groups for different regions of Africa for maize and sorghum; and (4) initiated proof-of-concept molecular breeding projects for maize and sorghum in Africa through finalizing priority target traits, identifying appropriate donor parent genotypes, and designing suitable breeding schemes.

---

**112. GCP/Rockefeller project G4005.70.01 (CB20a/RF-FS091):  
Tapping crop biodiversity for the resource poor in East and  
Central Africa**

*July 2005–June 2008*

**Principal Investigator**

Dave Hoisington, DDG Research, ICRISAT, d.hoisington@cgiar.org  
Patancheru 502324, Andhra Pradesh, India; Tel 91 40 3071 3071; Fax: 91 40  
3071 3074

**Collaborating institutions and scientists**

- Serere Agricultural and Animal Production Research Institute, Uganda:  
Johnnie Ebiyau
- National Agricultural Research Institute, Eritrea: Amanuel Mahdere
- Ethiopian Agricultural Research Organisation, Melkassa Research Centre,  
Ethiopia: Tesfaye Teferra Tesso
- KARI: Ben M. Kanyenji
- ARI: Amboseni Mbwaga
- Agricultural Research and Technology Corporation, Sudan: Dr. Abdalla  
Mohammed
- Institut des Sciences Agronomiques du Rwanda : Mr. Theophile Ndayayisenga
- Institut des Sciences Agronomiques du Burundi, Burundi: Espérance Habindavyi
- ICRISAT, Kenya: Dan Kiambi
- ICRISAT, India: Tom Hash

**Context**

The knowledge of the extent and structure of genetic diversity in germplasm accessions through characterisation is essential in understanding the evolutionary trends and also the development of strategies for the conservation of germplasm and its efficient utilisation in crop improvement programmes. Analysis of data generated through morphological and molecular characterisation provides information on responses to biotic and abiotic stresses and also on farmer- and market-preferred traits. This project was developed by the Generation Challenge Programme (GCP) and the Rockefeller Foundation in collaboration with Biosciences eastern and central Africa (BecA) in order to harness diversity and use it for crop improvement programmes. It focuses on two crops, cassava (activities coordinated by IITA) and sorghum (activities coordinated by ICRISAT). The project brings together BecA, regional and international initiatives, such as ASARECA and its affiliated networks and the GCP, in a synergistic way that will increase impact through quality research within Africa by releasing the value of African crop germplasm through systematic characterisation of sorghum germplasm using both morphological characters and molecular markers.

**Progress**

*National germplasm collections and phenotypic data*

To date, inventories of germplasm held by the 8 collaborating NARS have been compiled and subsets of between 164 to 298 accessions selected by the countries for phenotyping



and genotyping. Training of the NARS partners has been conducted through the phenotyping workshop and attachment in the BecA lab for trainees to learn genotyping and molecular techniques as part of the postgraduate programmes. Phenotyping of the germplasm is being carried out using 27 descriptors while genotyping is done using 39 SSR markers that have already been optimised and used to generate the first data sets from Tanzania, Rwanda, Burundi, Kenya and Uganda. A data analysis workshop is planned at a later stage when all the phenotyping and genotyping activities are completed.

#### *Genotypic using SSR markers*

A total of 39 SSR markers were optimised and five countries (Kenya, Uganda, Tanzania, Rwanda and Burundi) have finished genotyping their germplasm sets. A total of 29750 SSR data points have been generated and the data is currently being sorted out. Two countries (Sudan, Ethiopia) are currently genotyping their germplasm sets, with Eritrea planned for later in July 2008.

#### *Training and capacity building in molecular techniques and genotyping*

One PhD student from Sudan was registered at the University of Free State in South Africa. Five Visiting Fellows from Kenya, Uganda, Tanzania, Rwanda and Burundi have been trained in molecular techniques and genotyping. Two Visiting Fellows from Sudan and Ethiopia are currently being trained at BecA, Nairobi, and plans to have another one from Eritrea in August have been finalised. Three of the Visiting Fellows (Kenya, Tanzania and Ethiopia) are carrying out the genotyping work as part of their MSc studies. The MSc student from Tanzania has completed his studies, submitted the Thesis and graduated. He is currently drafting a paper from his work

#### **Next steps**

All the genotyping data will be completed by the end of September 2008. A data analysis workshop (for both genotypic and molecular data) with representation from all the participating countries is planned for October 2008. The final project report will be written and submitted after the data analysis workshop.

#### **References**

- Cassa AM, Mitchell SE, Hamblin MT, Sun H, Bowers E, Paterson AH, Aquadro CF, Kresovich S. (2005). Diversity and selection in sorghum: simultaneous analyses using simple sequence repeats. *Theor Appl Genet* (2005) 111: 23–30
- Folkertsma RT, Frederick H, Rattunde W, Chandra S, Raju GS, Hash CT. (2005). The pattern of genetic diversity of Guinea-race *Sorghum bicolor* (L.) Moench landraces as revealed with SSR markers. *Theor Appl Genet*. 111: 399–409

---

**113. GCP/Rockefeller project G4005.70.02 (CB20b/RF-FS090):  
Tapping crop biodiversity for the resource-poor in East and  
Central Africa**

**Principal Investigator**

Morag Ferguson, IITA, [m.ferguson@cgiar.org](mailto:m.ferguson@cgiar.org)

ABSTRACT NOT SUBMITTED

**114. G6007: Tropical Legumes I (TLI): Improving tropical legume  
productivity for marginal environments in sub-Saharan Africa**

**Principal Investigators**

- Objective 1: Improve groundnut (*Arachis hypogaea L*) productivity for marginal environments in sub-Saharan Africa : **R Varshney, ICRISAT (effective June 2008), D Hoisington, ICRISAT (May 2007–June 2008)**
- Objective 2: Improve cowpea (*Vigna unguiculata L*) productivity for marginal environments in Africa : **J Ehlers, UC–Riverside**
- Objective 3: Improve common bean (*Phaseolus vulgaris L*) productivity for marginal environments in Africa: **M Blair, CIAT**
- Objective 4: Improve chickpea (*Cicer arietinum L*) productivity for marginal environments in sub-Saharan Africa : **R Varshney, ICRISAT (effective June 2008), D Hoisington, ICRISAT (May 2007–June 2008)**
- Objective 5: Develop cross-species resources for comparative biology in tropical crop legumes: **D Cook, UC–Davis**
- Objective 6: Provide training and capacity-building for SSA scientists: **C de Vicente, GCP**

This proposal focuses on improving the productivity of legume crops of high importance to food security and poverty reduction efforts in sub-Saharan Africa. Modern biotechnologies offer great potential for enhancing the efficiency of plant breeding programmes, but sufficient genomic resources are needed to implement modern breeding. This project will develop the key genomic resources that are currently lacking in legumes (including cross-legume molecular markers for comparative genomics), identify molecular markers for traits of importance to resource-poor farmers (biotic stresses and drought tolerance), and implement breeding capacities in sub-Saharan Africa. The long term objective of this project (10-15 years) is to double grain legume productivity in farmers' fields. Doing so will generate an additional income for farmers of \$160/h in cowpea, \$370/h in groundnuts, and \$220/h in bean per crop cycle in the target countries of the project, where average agricultural population per capita income today is around \$120 per year.

**114.01. G6007.01: TLI Objective 1--Improve groundnut productivity in marginal environments of sub-Saharan Africa**

**May 2007–April 2010**

**Principal Investigator**

Vincent Vadez, v.vadez@cgiar.org, ICRISAT

GT Biotechnology, ICRISAT, Patancheru 502324, Andhra Pradesh, India –(91) 40 30 71 34 63 – (91) 40 30 71 30 74

**Collaborating institutions and scientists**

- ICRISAT, India: Rajeev K Varshney, P RatnaKumar
- ICRISAT, Mali: Bonny N'Tare
- ICRISAT, Malawi: Emmanuel Monyo
- ARI–Naliendele: Omari M'Ponda
- ISRA–CNRA: Ousman N'Doye, Issa Faye
- INRAN: Scientist to be determined
- Chitedze Research Station, Malawi: Tobias Kapewa
- University of Georgia, USA: Andy Paterson
- UCB: David Bertioli
- EMBRAPA: Soraya Bertioli, Patricia Guimares

**Context**

Groundnut (peanut, *Arachis hypogaea* L.) is an important food and cash crop in Africa. In most countries, the crop is grown by smallholder farmers under very low input and rainfed drought-prone conditions. Lack of seed of improved varieties, poor agronomic practices, drought, diseases, and pests are the major factors limiting yield. The goal of the project's objective is to improve the disease and drought tolerance in farmer-preferred groundnut varieties using modern molecular tools, and involving African NARS. For this purpose major efforts in the following areas are needed: (1) screen a large set of representative germplasm to identify the best source of tolerance to drought and resistance to diseases for their use in breeding; (2) the development of the molecular tools, still dramatically lacking, to speed up the introgression of tolerance and resistance characteristics into farmer-preferred varieties; (3) and (4) apply the tools developed in (2) to determine the genes involved in disease resistance and drought tolerance; and (5) introgress resistance/tolerance into farmer varieties.

**Findings and implications**

Analysis of the molecular diversity data from a GCP-funded project with existing phenotypic data allowed the identification of a reference set of approximately 300 accessions that contain the majority of the diversity in the collection. Seed of the original plant that was genotyped has been multiplied in Year 1 for field and trait screening in Year 2.

First map in cultivated groundnut: The first genetic map of cultivated groundnut, from the cross TAG24 x ICGV86031, developed between parents contrasting for transpiration efficiency, has been assembled in the project.

---

Set up of the lysimetric facility at Patancheru: A large facility has been set up at ICRISAT headquarter in India. The system consists of a set of cylinders (1.2m long, 20 cm dia) filled with natural soil at common bulk density and mimicking a soil profile. Each cylinder is equipped with a collar that allows its lifting. A hanging scale allows regular weighing of the cylinders to measure plant transpiration (after limiting soil evaporation by mulching the soil surface). The system allows assessing water extraction by roots under conditions of imposed water deficit. The results from the evaluation of the reference collection of 300 genotypes has revealed a two-fold range of variation for water extraction, with known check cultivars of groundnut extracting less water than varieties known for their better drought adaptation.

#### Markers development and screening

Two hundred new SSR have been developed. Existing SSR markers have been screen in the parental lines of RIL populations involved in the project (Table)

	<b>Nb of markers</b>	<b>Polymorphic markers</b>	<b>Polymorphic loci</b>	<b>Pop size</b>
ICGV86031 x TAG24	1286	144	150	318
ICGS 44 x ICGS 76	1026	57	46	376
ICGS 76 x CSMG 84-1	633	66	40	277

Efforts are also on-going to develop DArT arrays, one with 14 parents of mapping populations, and one with 45 varied groundnut germplasm, including diploid accessions.

#### Large number of crosses developed

Many crosses have been developed in Year 1 between contrasting parents for disease resistance under activity (5). Also a collection of market- and farmer-preferred varieties has been assembled.

#### **Links**

Phenotyping workshop at Patancheru: Organised between 3<sup>rd</sup> and 28<sup>th</sup> of March 2008 to train NARS technicians and scientists to phenotyping techniques needed to carry out the tasks of (1) and (4) (GCP-funded project G4008-27). The workshop has been the occasion to know better the African team, assess the strengths and weaknesses.

#### **Next steps and/or challenges**

Lot of phenotyping is planned for Year 2, within Activities (1), (3) and (4), to explore the germplasm for best sources of resistance/tolerance, and to phenotype existing RIL population for QTL identification. Activity 2 will focus on sequencing the BAC library, expanded to reach about a 10X coverage, linking the genetic and physical map, developing more markers, and finalise the DArT arrays. Populations developed under (5) will be advanced. Farmer preferred varieties will be evaluated to choose the most suitable one for the introgression work.

**114.02 G6007.02: TLI Objective 2--Improve cowpea productivity for marginal environments in sub-Saharan Africa**

**May 2007–April 2010**

**Principal Investigator**

Jeff Ehlers, UC–R, jeff.ehlers@ucr.edu

Dept. of Botany and Plant Sciences, University of California, Riverside, 900 University Ave., Riverside, CA 92521-0124, USA;; 951-323-5918; Fax 951-827-4436.

**Collaborating institutions and scientists**

- UC–R: Timothy Close; Philip Roberts
- UC–D: Ming-Cheng Luo
- ISRA: Ndiaga Cisse
- IRAD: Ousmane Boukar
- INERA: Issa Drabo
- IITA, Nigeria: Satoru Muranaka; Ousmane Boukar; Christian Fatokun; Dong-Jin Kim

***Progress by Objective 2 Activity and Milestone (Activity leadership in parenthesis)***

**Activity 1. Characterise diversity and develop germplasm for genetic studies. Milestone 1. Seed of 4 RIL populations multiplied for Activities 3 and 4 (UCR).** Seed of 10 RIL populations were multiplied in greenhouses at UCR during the first project year, far exceeding the stated goal of multiplying 4 RIL sets. **Milestone 2. 500 genotypes characterised for drought tolerance, drought germplasm identified.** Rainy season drought phenotyping evaluations were conducted at two drought-prone sites in Burkina Faso and one in Senegal, and in irrigated ‘off-season’ experiments in California, Senegal, Burkina Faso and IITA-Kano. Six genotypes were identified that showed drought tolerance over multiple environments.

**Activity 2: Generate genomic resources for genetic studies and breeding. Milestone 3. 14X coverage BAC library produced, 70,000 clones (UCR, UCD).** Two BAC libraries were produced (*HindIII* and *MboI*) totaling 17x genome coverage (73,728 clones, ~140 kb average insert size). **Milestone 4. 6 cDNA libraries produced, 200,000 ESTs (UCR).** 9 cDNA libraries were produced for sequencing at the Joint Genome Institute (JGI). An initial assessment of SNP yield conducted on 384 clones from each library showed that five libraries were most relevant to African breeding programmes and TLI objectives. EST sequencing was completed on these five libraries. IITA contributed 42,000 cowpea ESTs from four additional African genotypes. More than 10,000 high confidence SNPs were discovered and a subset of 1536 SNPs were chosen for the Illumina GoldenGate assay **Milestone 5. 1000 SNP markers on map, 1440 genotypes scored (UCR).** Design of the Illumina GoldenGate assay for 1536 cowpea SNPs was completed in June 2008 and arrival in August 2008 is expected. We increased the target number of genotypes from 1440 to 1632 DNA to accommodate additional RILs and accessions. DNA extractions have been completed on ~1400 cowpea leaf tissue samples (~900 RILs and ~500 accessions). **Milestone 6. BAC fingerprinting and assembly of physical map (UCR, UCD).** The planned 60,000 BAC fingerprints of IT97K-499-35 have been

---

completed, accomplishing a final depth of coverage of ~10x. **Milestone 7. BAC-end sequencing and linkage to reference genomes and maps (UCD).** The planned 30,000 BAC-end sequencing reactions from the IT97K-499-35 library have been completed; among them 4167 contain at least one SSR. Overall, our vision of nearby marker development through a complete physical map and BAC-end sequences (including SSRs) anchored to the physical and SNP-based genetic map is on track for fruition in latter 2008.

**Activity 3: Identify molecular markers and genes for biotic stress resistance. Milestone 8. 2 RILs phenotyped for resistance to flower thrips (IITA, IRAD, ISRA).** Thrips resistance phenotyping was conducted in Nigeria, Cameroon, and Senegal, with a total of five RIL sets between the three locations. In Senegal, significant differences were observed between the parent lines (Yacine and 58-77) and resistant check line TVx3236, but the frequency of RILs with scores as low as the resistant parent was low and there were no apparent transgressive resistant lines. In Cameroon, differences among RILs and between RIL parental lines were not significant. **Milestone 9. 2 RILs phenotyped for resistance to pathogens (UCR, IRAD, ISRA).** Two RIL sets, CB46/IT93K-503-1 and CB46/IT97K-499-39 were phenotyped for resistance to bacterial blight, viruses, flowering and maturity dates in Cameroon. Two RIL sets, CB46/IT93K-503-1 and CB27/24-125B were phenotyped for Fusarium wilt resistance in California. These two RIL sets are also being phenotyped for root-knot nematode resistance during spring-summer 2008 in replicated plots on infested field sites in California. **Milestone 10. Molecular markers associated with resistance to biotic stresses (UCR).** No progress was anticipated until Year 2.

**Activity 4: Identify molecular markers and genes for drought tolerance. Milestone 11. RILs Phenotyped for drought tolerance (UCR, IITA, INERA, ISRA).** Phenotyping of RIL sets for drought tolerance were conducted in Burkina Faso, California and Senegal. In California and Burkina Faso, significant yield differences were observed among the RILs regardless of maturity classes and transgressive RILs were identified. In Senegal RIL Mouride/Bambey 21 and RIL CB46/IT93K-503-1 were phenotyped for drought tolerance, and transgressive drought tolerant RILs identified. **Milestone 12. Molecular markers associated with drought tolerance (UCR).** No progress was anticipated until Year 2.

**Activity 5: Enhance locally adapted germplasm with target traits. Milestone 13. Targeted SNPs converted to easy-to-use markers (IITA).** Methods to convert known SNPs into easy-to-use markers are being developed, including exploring neighboring intergenic regions which may be easier to convert. **Milestone 14. Breeding populations developed with thrips resistance and drought tolerance (UCR).** Fifteen crosses were made at UCR between several cowpea lines likely to possess drought tolerance and that have good agronomic performance. The F<sub>1</sub>'s of these crosses were planted and F<sub>2</sub> seed produced by July 2008. **Milestone 15. Easy-to-use markers validated for targeted traits (IITA).** No progress was anticipated until Year 2.

**114.03 G6007.03: TLI Objective 3--Improving common bean productivity for marginal environments in sub-Saharan Africa**

**May 2007–April 2010**

**Principal Investigator**

Matthew Blair, CIAT, m.blair@cgiar.org

International Center for Tropical Agriculture (CIAT), AA6713, Cali, Colombia  
Tel: 57-2-4450-000; VIA USA 650-833-6625; Fax: 57-2-4450-073; VIA USA  
650-833-6626

**Collaborating institutions and scientists**

- CIAT: Idupulapati Rao; Manabu Ishitani; Steve Beebe
- ECABREN/Univ. of Nairobi, Kenya: Paul Kimani
- SABRN/CIAT, Malawi: Rowland Chirwa
- SARI, Ethiopia: Asrat Asfaw Amele
- DR4D, Zimbabwe: Godwill Makunde
- DARTS, Malawi: Virginia Chisale
- RIKEN, Japan: Motoaki Seki

Common beans are a major crop for most of the countries of Eastern and Southern Africa and are grown by small-scale farmers because of their high marketability and ready acceptance in the diet. Typical bean yields, however, are only 20 to 30% of the genetic potential of improved varieties due to major production constraints such as drought and several diseases and insects. In this project, we are contributing to improved drought tolerance by breeding common beans for drought tolerance using genetic knowledge, new germplasm, and improved breeding methods, and we are also addressing disease and insect resistance using molecular marker technology.

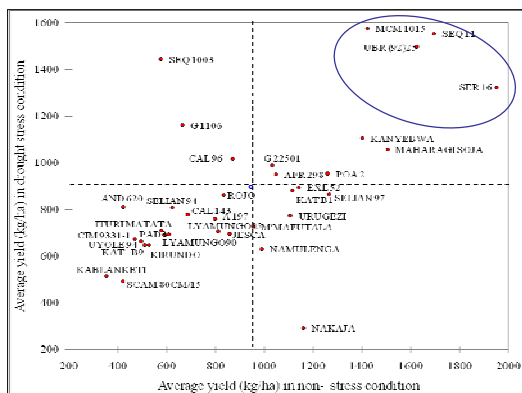
In the first year of the project a wide germplasm base has been established and tested, including the GCP reference collection, CIAT core collection, regional and local varieties from the ECABREN and SABRN bean networks and additional landraces from Ethiopia and Kenya. Drought tolerance experiments have used balanced lattice design experiments with stratification by genotype sources within the networks or across Andean and Mesoamerican genepool boundaries and plantings have been undertaken in CIAT headquarters during the dry season as well as in DART/CIAT-Malawi (Chitedze station), SARI (Awassa, Ethiopia) and the University of Nairobi (Kabete and Thika, Kenya). Figure 1 shows examples of two lattices for ECABREN and SABRN genotypes, out of a total of seven 6 x 6 lattices that were tested across drought (non-irrigated) versus irrigated treatments in CIAT. Notable results were the identification of drought tolerant sources from the Mesoamerican genepool (SER16 for example) which as expected generally outyielded the Andean drought tolerance sources (with SEQ11, SEQ1003, Kanyebwa, among the best and AFR298, DRK149, G17722, KATB1, RAA20, RAA21, SEQ1031 and Selian 97 as intermediately tolerant genotypes).

Varieties needing improvement were identified including CAL143, CIM9331-1, DRK57, Kalima and Sapelekedwa (Malawi), Selian 94, Lyamungu 85 and 90 (Tanzania), and Red Canadian Wonder (Zimbabwe), for example.

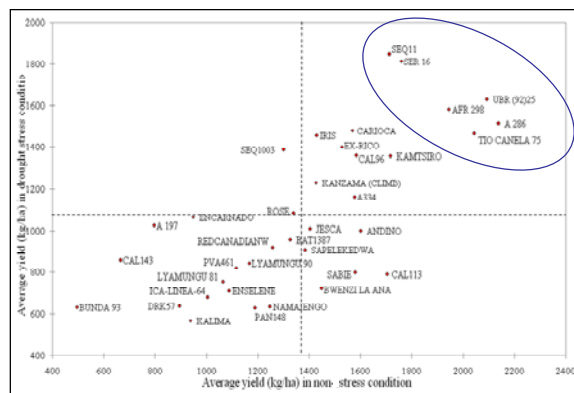
In terms of generation of genomic resources, molecular tools are being developed for the efficient pyramiding of drought tolerance with other needed traits. Marker development has prioritised candidate genes for transpiration/water use efficiency, osmotic adjustment, and root development along with the creation of subtractive libraries between drought tolerant and susceptible genotypes. Expression analysis is being conducted to validate candidate genes as they are converted into molecular markers. A set of new genetic SSR and SNP markers is also being developed cross-legume COS markers have been tested.

For genetic analysis of drought tolerance, two RIL populations have been tested in field and greenhouse trials to determine drought tolerance mechanisms including translocation efficiency and deep rooting. Meanwhile, six new RIL populations are being developed to analyze less well-studied drought tolerance sources such as those in untapped Durango accessions. Finally three innovative crop improvement methods are being used to develop genotypes with consumer acceptance combined with drought tolerance: 1) Advanced backcrossing of drought tolerant sources into grain types popular for Eastern or Southern Africa (red or cream mottled) 2) Intergene pool crosses based on North Carolina Design II mating between elite drought tolerant genotypes and commercial varieties and 3) Marker assisted recurrent selection for drought tolerance in a rapid cycling breeding programme of seven generations over two years.

**Lattice 1 – ECABREN genotypes**



**Lattice 5 – SABRN genotypes**



**Figure 1. Bi-plot distributions for yield of ECABREN (lattice 1) and SABRN (lattice 5) genotypes under drought and non-stress conditions in the CIAT 2007 dry season. Circled genotypes indicate drought / abiotic stress tolerant genotypes.**



**114.04 G6007.04: TLI Objective 4--Improved chickpea productivity for marginal environments in sub-Saharan Africa**

**May 2007–April 2010**

**Principal Investigator**

Dave Hoisington (May 2007- May 2008); ICRISAT, d.hoisington@cgiar.org

Rajeev Varshney (June 2008- April 2010); ICRISAT, r.k.varshney@cgiar.org

International Crops Research Institute for the Semi-Arid Tropics (ICRISAT),  
Patancheru-502324, AP, India

Tel: 0091 40 3071 3305, Fax: 0091 40 3071 3074

**Collaborating institutes and scientists:**

- ICRISAT, India: Rajeev Varshney (Leader, Activity 2), Hari Sharma (Leader, Activity 3), Junichi Kashiwagi (Leader, Activity 4), Pooran Gaur (Leader, Activity 5)
- ICRISAT, Kenya: Said Silim/ NVPR Gangarao (Leader, Activity 1)
- EIAR: Million Eshete
- LZARDI: Robert Kileo
- IIPR: Masood Ali
- University of Frankfurt, Frankfurt am Main, Germany: Peter Winter, Guenter Kahl
- UC-D: Doug Cook
- DArT P/L: Andrzej Killian

**Research highlights**

Chickpea (*Cicer arietinum L.*) is an important grain legume in South Asia and SSA, especially in eastern and southern Africa. Drought, globally the number one constraint to chickpea production, occurs during the terminal growth stages as the crop is largely grown rainfed during the post-rainy season on residual soil moisture. Pod borer (*Helicoverpa armigera* Hubner) is a highly devastating insect pest of chickpea worldwide. It feeds on various plant parts such as leaves, tender shoots, flower buds, and immature seeds. The major goal of this project is to improve incomes and livelihoods by establishing the capacity to improve the insect (pod borer) resistance and drought tolerance in farmer-preferred chickpea varieties using modern molecular tools. This goal will be achieved by: (1) screening a large collection of novel chickpea germplasm to identify important new sources of resistance to the above insect pests and drought component traits; (2) developing a greatly enhanced set of molecular tools for use in chickpea; (3) determining the genes involved in insect pest resistance; (4) determining the genes involved in specific root traits contributing to drought; and (5) introgressing resistance/tolerance into farmer varieties.

Reference collection, developed based on earlier GCP-SP1 sponsored chickpea genotyping project, has been phenotyped for root traits, insect resistance and harvest index (HI) in Patancheru. A higher level of phenotypic variation for all three traits has been observed in the reference collection. These data will be used to develop new genetic populations for root traits and insect resistance for Africa and Asia. Towards developing

---

genomic resources in chickpea, a total of 1,655 novel SSRs have been isolated from a SSR-enriched library (311) and BAC-end sequences (1,344) in collaboration with University of Frankfurt and University of California-Davis, respectively (Nayak et al. 2008). These SSRs are being used to genotype one interspecific reference population (*C. arietinum* ICC4958 x *C. reticulatum* PI 489777) with an objective to integrate newly developed markers to the reference genetic map (Varshney et al. 2008). In order to enhance the density of chickpea genetic map, it is planned to develop either DArT or SNP markers.

The reference population segregates for insect resistance and therefore it has been phenotyped under natural field infestation conditions as well as green house conditions using detached leaf assays. A quantitative variation for different parameters conferring insect resistance was observed. Phenotyping of the population for insect resistance for the second year is in progress. Two intraspecific (*C. arietinum*) mapping populations (ICC 4958 x ICC 1882, ICC 8261 x ICC 283), developed before the start of the TLI project, segregate for root traits and harvest index. These populations have been phenotyped for root traits using PVC Cylinder approach (Kashiwagi et al. 2006) and also for harvest index. For both, root traits as well as harvest index, normal distribution has observed in the mapping populations. While phenotyping of these populations is in progress for the second year, marker genotyping data are being generated on these populations by using newly developed as well as already existing SSR markers in public domain. Marker genotyping data together with the phenotyping data will provide QTLs for insect resistance and root traits for their deployment in chickpea breeding.

As a major root trait QTL was identified before starting TLI project, efforts are being made to introgress this QTL in some drought sensitive chickpea lines of Africa and Asia through marker-assisted back crossing (MABC). In this direction, BC1F1s have already been generated for three crosses. In parallel, marker-assisted recurrent selection (MARS) approach has been initiated and three sets of crosses using good x good genotypes from Africa and Asia have been developed.

## References

- Kashiwagi J, Krishnamurthy L, Crouch J H, Serraj R. (2006) Variability of root length density and its contributions to seed yield in chickpea (*Cicer arietinum* L.) under terminal drought stress. *Field Crops Research*, 95: 171-181.
- Nayak S, Mamo BE, Varghese N, Varma Penmetsa R, Farmer AD, Woodward J, Gao J, Hoisington D, May GD, Bruening G, Varshney RK, Cook DR (2008). BAC-end associated microsatellites in chickpea (*Cicer arietinum*) and cowpea (*Vigna unguiculata*): A resource for genetic and physical mapping. Poster presented at Plant and Animal Genome (PAG) conference XVI, San Diego, USA, 12-16 January, 2008.
- Varshney RK, Penmetsa RV, Varghese N, Farmer A, Reddy P S, Sarma B, Nayak Spurthi, Carrasquilla Garcia N, Lekha P, Gao J, Jayashree B, Steiner S, Gaur PM, Srinivasan R, Hoisington D, Winter P, Bruening G, May G, Cook DR (2008). A genomics platform for molecular breeding and comparative genomics in chickpea (*Cicer arietinum* L.). Oral presentation at Cool season legume workshops at Plant and Animal Genome (PAG) conference XVI, San Diego, USA, 12-16 January, 2008.

**114.05 G6007.05: TLI Objective 5--Development of cross-species resources for comparative genomics in tropical crop legumes**

**May 2007–April 2010**

**Principal Investigator**

Douglas R. Cook; UC–D; [drcook@ucdavis.edu](mailto:drcook@ucdavis.edu)

University of California, Department of Plant Pathology, One Shields Ave, Davis, CA, USA, 95616

**Background**

This project aims to define the ancestral genetic framework of chickpea, cowpea, common bean, and groundnut. If breeders and biotechnologists know how chromosome segments in one species relate to corresponding chromosome segments in other related species, then they will gain insight into the molecular basis of agronomic traits between crop species, and between crops and well-characterised genomes of reference legumes.

During the past year, we have focused on a set of 1440 low-copy genes for allele re-sequencing, leading to the description of thousand of single nucleotide polymorphisms (SNP).

Our current efforts focus on (1) collecting DNA and/or seed from project partners for genetic mapping experiments, (2) continued allele re-sequencing for additional genotypes nominated for common bean and groundnut, and (3) curation of SNP calls.

Although not part of the original project, we also hope to incorporate candidate genes for disease resistance into our SNP data set to develop markers that can be directly applied by breeding programmes in sub-Saharan Africa.

**Activities**

- Last year, we developed bioinformatics tools to identify, compare, and then prioritise a set of low-copy genes present in the public genome sequences of three reference legumes (*Medicago truncatula*, *Lotus japonicus*, and soybean). From a set of 11,250 sequence alignments, we focused on a subset of 1545 genes that were highly conserved and contained features (eg, introns) desirable for discovering genetic variation. We designed and tested 5760 oligonucleotide primers in appropriate pairwise combinations, using DNA from chickpea, common bean, cowpea, and groundnut.
- Based on success in PCR amplification, we obtained a set of about 1440 genes for allele re-sequencing in each targeted legume species. Each of these gene markers has corresponding genome locations in one or more of the reference legume genomes.
- To date, we have conducted allele re-sequencing for all genes in two parents of each of common bean and chickpea, and four parents of cowpea. For groundnut, 10 diploid mapping parent accessions were first surveyed, and *A. duranensis* 2 and 25 were selected as suitably polymorphic for full analysis. As of the writing of this abstract, we have validated 11,545 single nucleotide polymorphisms in these four species, representing an average of 376 orthologous genes in each species.

---

Within the next 9 months, we will continue allele re-sequencing in additional target genotypes, and we will work to complete the design of an oligonucleotide pooled all (OPA) platform for genotype analysis, using the Illumina® GoldenGate assay

**114.06 G6007.06: TLI Objective 6--Provide training and capacity-building for sub-Saharan African scientists**

***May 2007–April 2010***

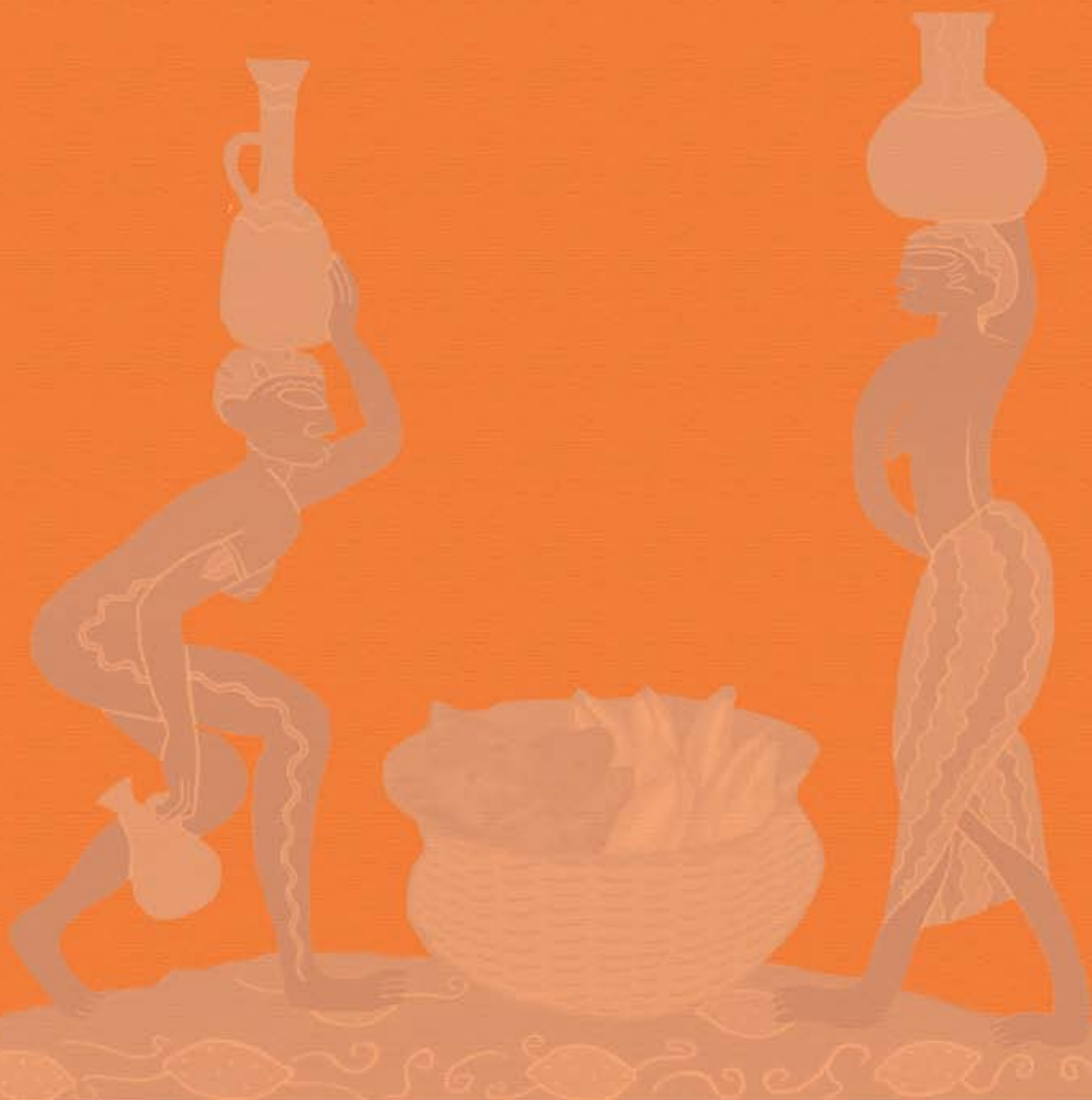
**Principal Investigator**

Carmen de Vicente, GCP, [c.devicente@cgiar.org](mailto:c.devicente@cgiar.org)

Apdo. Postal 6-641, 06600 Mexico, D.F., Mexico, Phone: +52 55 5804 2004

**Collaborating institution and scientist**

- GCP: Nosisa Mayaba (Associate Scientist)
- The project's launch meeting (18–22 September 2007) took place successfully at the Orion Safari Lodge, Rustenburg Kloof, South Africa. The agenda focused mainly on planning Year1 of work and included the development of delivery plans in collaboration with national programme partners present in the meeting.
- One goal of the meeting was to adequately assess the needs of NARS project partners in terms of bioinformatics, laboratory, and field research equipment. To enable sub-Saharan African researchers to start or continue their work in their respective institutions, a list of equipment needs was compiled in the DPKits for each Objective. Priority was given to informatics and field equipment, and purchases are under way.
- Phenotyping was regarded as the most urgent activity for which capacity was needed in the TLI team, and a course was commissioned from ICRISAT on 'Phenotyping for droughtrelated traits across tropical legumes: concepts and practices' (3–28 March 2008). Twentyfour participants attended the course, which focused on Africa.
- In order to increase the prospects for adoption of TLI results and thus to enhance impact, exchanges with the TLII project team have taken place to identify students that could incorporate TLI research components in their degree programmes



# Generation Challenge Programme

CULTIVATING PLANT DIVERSITY FOR THE RESOURCE-POOR

Hosted by CIMMYT

(Centro Internacional de Mejoramiento de Maíz y Trigo;  
the International Maize and Wheat Improvement Center)

**Mailing address:**

Apdo Postal 6-641  
06600 Mexico, DF Mexico

**Physical address:**

Km 45 Carretera México-Veracruz  
El Batán, Texcoco, México, CP 56130

Tel: +52 55 5804 2004 • Fax: +52 55 5804 7558

Email: [generationcp@cgiar.org](mailto:generationcp@cgiar.org) or [info@generationcp.org](mailto:info@generationcp.org)  
[www.generationcp.org](http://www.generationcp.org)