Generation Challenge Programme

CULTIVATING PLANT DIVERSITY FOR THE RESOURCE-POOR

2009 Project updates

July 2009

Generation Challenge Programme

2009 Project updates

July 2009

Generation Challenge Programme (GCP)
Hosted by CIMMYT
(Centro Internacional de Mejoramiento de Maíz y Trigo;
the International Maize and Wheat Improvement Center)

Mailing address:
c/o CIMMYT, Apdo Postal 6–641
06600 Mexico, DF Mexico

Physical address:
c/o CIMMYT 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 or info@generationcp.org
www.generationcp.org

Correct Citation: CGIAR Generation Challenge Programme, 2009. 2009 Project updates. Texcoco, Mexico: Generation Challenge Programme.

Editorial note: Project updates are published as presented by the authors, except for updates that extended beyond the set word limit where text beyond this point has been omitted, indicated by continuation dots/suspension points (...)

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

Table of contents

Acronyms and abbreviations

AARC	Awassa Agricultural Research Center, Ethiopia
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
ACPFG	Australian Centre for Plant Functional Genomics, Pty Ltd
ADOC	allelic diversity for orthologous candidate genes
AGRA	Alliance for a Green Revolution in Africa
AICPMIP	All-India Coordinated Pearl Millet Improvement Project
Al	aluminium
Alt_{SB}	marker diagnostic for aluminium tolerance
APSIM	Agricultural Production Systems Simulator
ARC–Sudan	Agricultural Research Corporation, Sudan
ARI	Agharkar Research Institute, India
ARI(s)	advanced research institute(s)
ARI–HAS	Agricultural Research Institute of the Hungarian Academy of Sciences, Hungary
ARI–Naliendele	Agricultural Research Institute–Naliendele Research Station, Tanzania
ARM	Annual Research Meeting
ARS–Durgapura	Agricultural Research Station, Durgapura, Jaipur, Rajasthan, India
ASTI	CGIAR Agricultural Science and Technology Indicators, Italy
BAC	bacterial artificial chromosome
BAU	Birsa Agricultural University, Ranchi India
BCMV	bean common mosaic virus
BINA	Bangladesh Institute of Nuclear Agriculture
BIOTEC	National Center for Genetic Engineering and Biotechnology, Thailand
Bioversity	Bioversity International
BLB	bacterial leaf blight
BMGF	Bill & Melinda Gates Foundation
BRRD	Bureau of Rice Research and Development, Rice Department, Thailand
BRRI	Bangladesh Rice Research Institute
CAAS	Chinese Academy of Agricultural Sciences
CAPS	cleaved amplified polymorphic sequence (markers)
CARDI	Cambodia Agricultural Research and Development Institute
CAZRI	Central Arid Zone Research Institute, India
CB	conventional breeding
CBI	Crop Breeding Institute, Department of Research for Development, Zimbabwe
cDNA	complementary DNA
CERAAS	Centre d'étude régional pour l'amélioration de l'adaptation à la sécheresse, Senegal
CGIAR	Consultative Group on International Agricultural Research
CGN–WUR	Centre for Genetic Resources–Wageningen University and Research Centre, The Netherlands
CHPRRU	Corn Host Plant Resistance Research Unit, USDA–ARS
CIAT	Centro Internacional de Agricultura Tropical (International Center for Tropical Agriculture)
CIMMYT	Centro Internacional de Mejoramiento de Maíz y Trigo (the International Maize and Wheat Improvement Center)
CIMS	Centro de Inteligencia sobre Mercados Sostenibles of INCAE
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
CLDRI	Cuu Long Delta Rice Research Institute, Vietnam
CMTV	Comparative Map and Trait Viewer
CNG	Centre National de Génotypage, Commissariat à l'Energie Atomique, Evry, France
CNRA	Centre National de Recherches Agronomiques, ISRA
CoP	community of practice
Cornell	Cornell University
COS	conserved orthologous sequence
CP	Challenge Programme (of the CGIAR)
CRI–Ghana	Crops Research Institute, Ghana
CRIL	Crop Research Informatics Laboratory (CIMMYT and IRRI)
CRI–Sri Lanka	Coconut Research Institute, Sri Lanka
CRRI	Central Rice Research Institute, India
CRS	Chitedze Research Station, Malawi
CRURRS	Central Rainfed Upland Rice Research Station, India
CSIRO	Commonwealth Scientific and Industrial Research Organisation, Australia
CSSL	chromosome segment substitution line
CSU	Colorado State University, USA
CStuU	Charles Sturt University, Australia
$\Delta^{13}C$	carbon isotope discrimination
DAR	Department of Agricultural Research, Myanmar
DARS	Department of Agriculture Research Services, Malawi
DArT	diversity arrays technology
DArT P/L	Diversity Arrays Technology Pty, Ltd
DMR	Directorate of Maize Research, India
DNA	Deoxyribonucleic acid
DOA–Thailand	Department of Agriculture, Thailand
DPKit	Delivery Plan Kit
DPSPP–EKC	Department of Plant Sciences and Plant Physiology, Eszterházy Károly College, Eger, Hungary
DREB	drought-responsive element binding protein (gene)
DWR	Directorate of Wheat Research, India
DZARC	Debre Zeit Agricultural Research Centre, Ethiopia
ECABREN	Eastern and Central Africa Bean Research Network
EgU	Egerton University, Kenya
EIAR	Ethiopian Institute of Agricultural Research
EMBRAPA	Empresa Brasileira de Pesquisa Agropecuária (Brazilian Agricultural Research Corporation)
EMU	Eduardo Mondlane University, Mozambique
ERECTA	a leucine rich repeat receptor-like kinase (gene)
EST	expressed sequence tag
ESU	Ebonyi State University, Nigeria
ETH	Eidgenössische Technische Hochschule, (Swiss Federal Institute of Technology), Zürich
F_1 etc	first filial generation etc
FABI	Forestry and Agricultural Biotechnology Institute, University of Pretoria, South Africa
Fedearroz	Federación Nacional de Arroceros, Colombia
FOFIFA–DRA	Foibem-Pirenena Mombra ny Fikarohana Ampiharina Amin'ny Fampandrosoana ny eny Ambanivohitra (National Centre. for Applied Research on Rurual Development) Département de la Recherche Agronomique, Madagascar

GCP	Generation Challenge Programme of the CGIAR
GIS	geographic information system(s)
GISH	genomic *in situ* hybridisation
GOST	GreenPhyl Ortholog Search Tool
GRSS	Genetic Resources Support Service
GSS	Genotyping Support Service
GxE	genotype by environment interaction
HAAS	Hebei Academy of Agricultural Sciences, Institute of Dry Farming, China
HAKI	Research Institute for Fisheries, Aquaculture and Irrigation, Hungary
HPC	high-performance computing
HZAU	Huazhong Agricultural University, China
IAMZ	Instituto Agronómico Mediterráneo de Zaragoza, Spain
IARI	Indian Agricultural Research Institute
IA–Tápiószele	Institute for Agrobotany, Tápiószele, Hungary
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
ICASEPS	Indonesian Center for Agro Socio-Economics and Policy Studies, Indonesia
ICERI	Indonesian Cereals Research Institute
ICFORD	Indonesian Center for Food Crops Research and Development
ICIS	International Crop Information System
ICL	Imperial College London, UK
ICRISAT	International Crops Research Institute for the Semi-Arid Tropics
ICS–CAAS	Institute of Crop Science, Chinese Academy of Agricultural Sciences
IER	Institut d'économie rurale, Mali
IFPRI	International Food Policy Research Institute
IGD	Institute for Genomic Diversity, Cornell University, USA
IGKV	Indira Gandhi Krishi Vishwa Vidyalaya (Indira Gandhi Agricultural University), India
i-GOST	iterative version of GOST
IIAM	Instituto de Investigação Agrária de Moçambique (Institute for Agricultural Research, Mozambique)
IIPR	Indian Institute of Pulses Research
IITA	International Institute of Tropical Agriculture
ILRI	International Livestock Research Institute
iMAS	Integrated Marker-Assisted Selection System
INCA	Instituto Nacional de Ciencias Agricolas, Cuba
INERA–Burkina Faso	Institut de l'environnement et de recherches agricoles, Burkina Faso
INERA–DRC	Institut national pour l'etude et la recherche agronomiques, democratic Republic of the Congo
INIA–Chile	Instituto de Investigaciónes Agropecuarias, Chile
INIA–Uruguay	Instituto Nacional de Investigación Agropecuaria, Uruguay
INIFAP	Instituto Nacional de Investigaciones Forestales, Agricolas y Pecuarias, Mexico
INRA	Institut national de la recherche agronomique, France
INRA–Morocco	Institut national de la recherche agronomique, Morocco
INRAN	Institut national de la recherche agronomique du Niger
INTA–Nicaragua	Instituto Nacional de Tecnologia Agropecuaria, Nicaragua
IP	intellectual property
IPB–The Philippines	Institute of Plant Breeding, The Philippines
IPK	Leibniz Institute of Plant Genetics and Crop Plant Research, Germany
IPM CRSP–VPI	Integrated Pest Management Collaborative Research Support Program–Virginia Polytechnic Institute and State University, USA

IPT	isopentenyltransferase (gene)
IRAD	Institut de la recherche agronomique pour le développement, Cameroon
IRC	Interactive Resource Centre
IRD	Institut de recherche pour le développement, France
IRRI	International Rice Research Institute
ISABU	Institut des sciences agronomiques du Burundi
ISAR	Institut des sciences agronomiques du Rwanda
ISRA	Institut sénégalais de recherches agricoles, Senegal
JCVI	James Craig Venter Institute, USA
JIC	John Innes Centre, UK
JIRCAS	Japan International Research Center for Agricultural Sciences
JLNKV	Jawahar Lal Nehru Krishi Vishwavidyalaya, Jabalpur, India:
KARI	Kenya Agricultural Research Institute
kb	Kilobase
KU	Kasetsart University, Thailand
KUL	Katholieke Universiteit Leuven, Belgium
LAAS	Luoyang Academy of Agricultural Sciences, China
LD	linkage disequilibrium
LIMS	Laboratory Information Management System
LUMC	Leiden University Medical Center, The Netherlands
MAB	marker-assisted breeding
MABC	marker-assisted backcrossing
MAGIC	multiparent advanced generation inter-cross
MahU	Mahidol University, Thailand
MARS	marker-assisted recurrent selection
MAS	marker-assisted selection
MAU	Marathwada Agricultural University, India
Mb	megabase
MPIDB	Max Planck Institute for Developmental Biology, Germany
MSV	maize streak virus
MU	Moi University, Kenya
N/A	not applicable
NAARI	Namulonge Agricultural and Animal Research Institute, Uganda
NaCRRI	National Crop Resources Research Institute, Uganda
NAFRI	National Agricultural and Forestry Research Institute, Laos
NAM	nested association mapping
NARI	National Agricultural Research Institute, Eritrea
NARS	national agricultural research system(s)
NaU	Nagoya University, Japan
NAU	Nanjing Agricultural University, China
NCE	no-cost extension
NCGR	National Center for Genome Resources, USA
NCSRC	National Corn and Sorghum Research Center, Thailand
NCSU	North Carolina State University, USA
NDUAT	Narendra Deva University of Agriculture and Technology, India
NERICA	new rice for Africa
NGO	non-governmental organisation
NIAB	National Institute of Agricultural Botany, UK
NIAS	National Institute of Agrobiological Sciences, Japan
NIL	near-isogenic line
NKLCGGE	National Key Lab of Crop Genetics and Germplasm Enhancement, China
NMRI	National Maize Research Institute, Vietnam
No.	number
NPGRC	National Plant Genetic Resources Centre, Tanzania
NRCPB	National Research Centre on Plant Biotechnology, India

NRCRI	National Root and Tuber Crops Research Institute, Nigeria
NRCS	National Research Centre on Sorghum, India
NSFCRC	Nakhon Sawan Field Crops Research Center, Thailand
NU	Ningxia University, China
NWSUAF	Northwest Sci-tech University of Agriculture and Forestry, China
ORE	Organisation for the Rehabilitation of the Environment, Haiti
OSU	Oregon State University, USA
PAU	Punjab Agricultural University, India
PBI–University of Sydney	Plant Breeding Institute–University of Sydney, Australia
PGRCU	Plant Germplasm Resources Conservation Unit, USDA–ARS
PhilRice	Philippine Rice Research Institute
PI	Principal Investigator
Pioneer	Pioneer Hi-Bred International, Inc
POC	Plant Ontology Consortium
PROINPA	Promoción e Investigación de Productos Andinos, Bolivia
PSU	Pennsylvania State University, USA
PU	Purdue University, USA
Pup1	marker diagnostic for phosphorus uptake
QPIF	Queensland Primary Industries and Fisheries, Australia
QTL	quantitative trait locus
QTLxE	QTL by environment interaction
R&D	research and development
RARS	Regional Agricultural Research Station, Nandyal, India
RAU	Rajasthan Agricultural University, India
RCB–IPB	Research Center for Biotechnology, Bogor Agricultural University, Indonesia
RF	The Rockefeller Foundation
RFLP	restriction fragment length polymorphism
RGDU	Rice Gene Discovery Unit, Thailand
RIKEN	Rikagaku Kenkyūsho (Institute of Physical and Chemical Research), Japan
RIL	recombinant inbred lines
RNA	ribonucleic acid
RYMV	rice yellow mottle virus
SAARI	Serere Agricultural and Animal Production Research Institute, Uganda
SAAS	Shanxi Academy of Agricultural Sciences, China
SABRN	Southern Africa Bean Research Network
Saltol	marker diagnostic for salt tolerance
SARI–Ghana	Savannah Agricultural Research Institute, Ghana
SARK	senescence associated receptor protein kinase
SAU	Sichuan Agricultural University, China
SCRI	Scottish Crop Research Institute, UK
SIRDC	Scientific and Industrial Research and Development Centre, Zimbabwe
SNP	single nucleotide polymorphism
SP	Subprogramme
SP1, SP2 etc	Subprogramme 1, Subprogramme 2 etc.
SPL	Subprogramme Leader
SPS	sucrose phosphate synthase (gene)
SPVD	sweet potato virus disease
SSA	Sub-Saharan Africa
SSR	simple sequence repeat
SUoAg	Sokoine University of Agriculture, Tanzania
TAMU	Texas A&M University
TBD	to be determined
TF	task force
TLI	Tropical Legumes I Project

TLII	Tropical Legumes II Project
TNAU	Tamil Nadu Agricultural University, India
TPE	target population of environments
TSL	The Sainsbury Laboratory, UK
TU	Tishreen University, Syria
UAS	University of Agricultural Sciences, India
UBU	Ubon Ratchatani University, Thailand
UCB	Universidade Católica de Brasilia, Brazil
UCG	Universidade Católica de Goiás, Brazil
UdR	Universidad de la Republica, Uruguay
UdB	Università di Bologna, Italy
UdU	Università di Udine, Italy
UGA	University of Georgia, USA
UKZN	University of KwaZulu–Natal, South Africa
UoA	University of Arizona, USA
UoAa	University of Aarhus, Denmark
UoAb	University of Aberdeen, Scotland
UoAl	University of Alberta, Canada
UoC	University of California, USA
UdAC	Universidad Autónoma Chapingo, México
UoD	University of Dhaka, Bangladesh
UoF	University of Frankfurt, Germany
UoGh	University of Ghana
UoH	University of Hohenheim, Germany
UoMi	University of Missouri, USA
UoN	University of Nairobi, Kenya
UoP	University of Pretoria, South Africa
UoQ	University of Queensland Australia
UoT	The University of Tehran, Iran
URGV	Unité de Recherche en Génomique Végétale, France
USDA–ARS	United States Department of Agriculture–Agricultural Research Service, USA
USDA–ARS PGRU	USDA–ARS, Plant Genetic Resources Unit
USP	Universidade de São Paulo, Brazil
UoV	University of Virginia, USA
VBI	Virginia Bioinformatics Institute, VPI
Virginia Tech	see VPI
VPI	Virginia Polytechnic Institute and State University, USA
WACCI	West Africa Centre for Crop Improvement, University of Ghana
WARDA	Africa Rice Center
WMS	Workflow Management System
WUR	Wageningen University and Research Centre, The Netherlands
YAAS	Yunnan Academy of Agricultural Sciences, China
ZU	Zhejiang University, China

COMPETITIVE PROJECTS

Subprogramme 1: Crop genetic diversity

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

January 2005–December 2008; no-cost extension to September 2009

Principal Investigator and Lead Institutes
Mathias Lorieux, Agropolis–IRD/CIAT; mathias.lorieux@ird.fr, m.lorieux@cgiar.org
Joe Tohme, CIAT; j.tohme@cgiar.org
CIAT Biotechnology Unit; AA 6713, Cali, Colombia; Tel.: +57 2 445 31 43

Collaborating institutes and scientists
- CIAT: César P Martinez
- WARDA: Marie Noelle Ndjiondjop; Baboucarr Manneh
- EMBRAPA: Claudio Brondani
- Cornell: Susan R McCouch
- Fedearroz: Miguel Diago Ramirez

Context
Wild relatives represent a valuable source of under-utilised genetic variation that is available to plant breeders and constitute an invaluable source of genetic information for modern genomics research initiatives. A systematic approach is required to identify and characterise genes from wild species that can be used to enhance crop productivity in a range of environments and under diverse cultural conditions. Using rice as a model, this project aims to develop six libraries of interspecific lines called Chromosome Segment Substitution Lines (CSSLs), targeting chromosomal introgressions from different rice relatives, and to analyse a set of advanced CSSLs generated from Asian x African rice crosses for their phenotypic response to drought stress.

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 Paddy Map Database was created and is available online (http://mapdisto.free.fr/PaddyMap/).

Development of CSSL populations
O. glaberrima (Caiapo x MG12)
59 lines were chosen and backcrossed to the *O. sativa* parent Caiapo and selfed to obtain 59 BC_4F_2 families. 4200 individuals were planted in the field with the aim to identify plants bearing the target chromosomal fragment. The DNAs were bulked and were evaluated with SSRs. The BC3DH population is available to the scientific community upon request and has been already distributed to eleven partners. The BC4F2/3 lines will be available by December 2009.

O. meridionalis and *O. rufipogon*
Double haploid lines were developed at CIAT from the candidate BC3F1 lines. Foreground selection of the BC3DH lines led to the selection of 180 lines that were subsequently backcrossed. These lines will be analysed for their genomic content at Cornell University during summer 2009 using the Illumina

Golden Gate SNP platform. The two populations will then be increased for seed and we expect to have them available for distribution by December 2009.

O. barthii
BC3F1 were obtained at WARDA from the BC2F1s developed at CIAT. Due to germination problems, it was decided to solve this issue before sowing the remaining material. We expect to obtain BC3F2 lines before December 2009.

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

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.

Implications
Generating such resources and knowledge will contribute to the objectives of Subprograms 1 and 3 by utilising *natural genetic diversity* to develop whole-genome libraries of CSSLs as a permanent genetic resource for both breeding and genomics-based research.

Next steps and/or challenges
Depending on the genetic material, BC3DH, BC3F2/3, or BC4F2 seeds will be prepared and made available to the scientific community. 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 “wild” 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 and Lead Institute
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 institutes and scientists
- Cornell: Tim Setter; Edward Buckler
- CIMMYT: Yunbi Xu; Jonathan Crouch
- NSFCRC: Pichet Grudloyma
- KARI: James Gethi
- SIRDC: Ester Khosa
- SAU: Wanchen Li

Major research activates 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).

Table 1. summary of indentified 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	lcye(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 centre of origin

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

Principal Investigator and Lead Institute

Marilyn Warburton m.warburton@cgiar.org

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

Collaborating institutes and scientists

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

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 colonisation. 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 programmes. 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 generalised 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 organisation 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, characterised 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, Maniccaci 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 and Lead Institute
Michael Baum, ICARDA; m.baum@cgiar.org
ICARDA, P.O.Box 5466 Aleppo, Syrian Arab Republic, +96 3 2122 2551

Collaborating institutes and scientists
- NIAB: W Powell; K. Stamati
- UdU: M Morgante
- ICARDA: S Ceccarelli; S Grando; SM Udupa; M Von Korff
- TU: 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. 2009) and from literature. Results of a previous microarray experiment (Guo et al. 2009), 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.

Furthermore, we have also started to investigated cis- and trans effects in two crosses (Stamati at al. 2009). Resukts on cis- and trans effects will be reported soon.

References

Guo P, M. Baum, S. Grando, S. Ceccarelli, G. Bai, R. Li, Maria von Korff, R.K. Varshney, A. Graner, J. Valkoun (2009) Differentially expressed genes between drought-tolerant and drought-sensitive barley genotypes in response to drought stress during the reproductive stage. *Journal of Experimental Botany*, 1 -14, doi:10.1093/jxb/erp194

von Korff M, Radovic S, Choumane W, Stamati K, Udupa SM, Grando S, Ceccarelli S, Mackay I, Powell W, Baum M, Morgante M (2009). Asymmetric allele specific expression in relation to developmental variation and drought stress in barley hybrids. *Plant J.*, doi: 10.1111/j.1365-313X.2009.03848.x.

Lababidi S., N. Mejlhede, S.K. Rasmussen, G. Backes, W. Al-Said, M. Baum, A. Jahoor (2009) Identification of Mutants at the Dhn Loci through TILLING in Barley. Plant Breeding, accepted.

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*, 2008 Sep; 117(5):653-69. Epub 2008 Jul 11.

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.

Stamati K, I. Mackay, S. Radovic, M. von Korff, M. Baum, M. Morgante, W. Powell (2009) Cis- and trans- regulatory variation underlines barley evolution. In prep.

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; no-cost extension to December 2010

Principal Investigator and Lead Institute

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 institutes and scientists

- CIAT: J Tohme; C Martinez; M Lorieux (IRD out-posted)
- Fedearroz: M Diago

- IER: F Cissé
- INERA–Burkina Faso: H Drissa
- LGDP, Agropolis–IRD/CNRS/Perpignan University, Montpellier and Perpignan, France: O Panaud; R Guyot
- PhilRice: A Alfonso
- PAU: JS Sidhu
- UoA: D W 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. The phase of material production consisting in the development of large BC1 populations in using twenty-five consensus accessions of *O. glaberrima* has been achieved mainly at CIAT and WARDA. MAS on the S^1 locus with RM190 and RM 204 markers was also completed (table 1). The transmission of the *O. sativa* allele in back cross (S^1_s) is around 5% regardless of the *O. sativa* recurrent parent. Nevertheless, some variation is observed according to the *O. glaberrima* parent and interestingly some accessions confirmed to derive from natural introgressions with *O. sativa* can give a much more higher transmission of S^1_s.

Table 1: summary of BC1 plants produced by MAS at WARDA

Nb of combinations	*O. sativa* parent	Nb of plants selected by MAS (RM190)	S1s transmission %	
			mean	range
13	WAB165	772	5.44 %	0 - 8.1%
12	Curinga	815	4.29 %	0 - 16%
14	IR64	928	5.50 %	0 - 19%
4	Admixtures	680	33%	16 - 46%

1.2 WP2: Characterisation of interspecific introgressions

As BES of CGI4 were not sufficient to derive new specific markers, Strategy is now focused on the genomic sequence in progress at AGI (IOMAP initiative) and in collaboration with LGDP. The available sequence of Chr. 3 is used to determine if SFP or SNP will be the more efficient technology to monitor the interspecific introgressions.

1.3 WP3: Physical mapping of S^1 locus

Physical mapping and sequencing of the S1 locus have been completed in using 8 BACs clones of the CG14 library and representing a 813 kb contig. Four BAC clones of our home TOG5681 library have been also sequenced for comparison. Two genes in 30 kb represent the best candidates for S1. Complete annotation of the region showed high conservation with Nipponbare sequence but the boundary region evidenced a large 40 kb inversion and multiple insertions of transposable elements. Insertion polymorphism generated by transposable elements is presently used to determine PCR markers along this region (13 *O. glaberrima* specific insertion markers and 7 O. sativa specific insertion markers) in order to verify the status of accessions supposed to derive from natural interspecific introgressions (see 1.1).

Outputs

Complete sequence resolution of S1 locus in 2 *O. glaberrima* accessions (CG14 & Tog5681) and determination of interspecific markers suitable for introgression monitoring and diversity analysis in African rice.

References

A.Garavito et al., Genetic model for interspecific female sterility barrier between O. sativa and *O. glaberrima* mediated by the S1 locus. (submitted to *New Phytologist*)

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

August 2007–July 2009

Principal Investigator and Lead Institute

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

Collaborating institutes and scientists

- ICARDA: Michael Baum; Stefania Grando; Salvatore Ceccarelli
- OSU: Patrick M Hayes
- SCRI: Joanne Russell; David Marshall
- INIA–Chile: Ivan Matus; Alejandro del Pozo
- UoC–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.

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

ICARDA and Chile are phenotyping the RCSLs for components of drought stress. Where possible assaying the same traits in both locations. The RCSLs have been grown for the first season (2008/2009) at two sites in Syria (ICARDA) and three sites in Chile (INIA). Physiological and agronomic traits have

been measured across these locations and sites and will over the next few months be analysed in Activity 4. For the season 2009/2010 the same experiments will be planted and traits evaluated with the addition of carbon discrimination and stem carbohydrates measurements in Chile.

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.*
Three hundred and seventeen of the SJLC accessions have been genotyped with the Barley OPA1. From the 1536 SNPs, 990 SNPs were polymorphic among this germplasm and grouped the 317 accessions into four main geographically based clusters (Moragues et al., 2009). By combining the SNP and previous phenotypic data (correcting for population structure) we have performed whole genome association analysis. Forty three loci were identified which were significantly associated with 4 key traits, including days to heading, grain yield, plant height and thousand kernel weight (Russell et al., 2009). By aligning these mapped barley loci with rice we have identified several key candidate regions which have been back-translated into CAPs markers. These primers are being tested on the complete set of 480 SJLC to assess their usefulness as markers for breeding programmes.

4. *Identification of introgression segments influencing drought response*
This activity is dependent on the availability of the phenotypic data which is now in the process of being gathered and distributed for association analysis. Using the information described in Inostroza et al. (2009), we have identified QTL in these RCSLs using genotype and phenotype data generated prior to the current GCP project. These regions were referenced to drought tolerance QTL reported in the literature and will serve as reference points using the complete genotype data, and more extensive phenotype data, generated by the GCP project.

References

Moragues, M, Comadran, J, Waugh, R, Milne, I, Flavell, AJ & Russell, JR. Effects of ascertainment bias and marker number on estimations of barley diversity from high throughput SNP genotyping data. 2009 Theoretical Applied Genetics (submitted).

Russell, JR, Comadran, J, Ramsay, L, Booth, A, Macaulay, M, Thomas, WTB, Marshall, DM, & Waugh, R. 2009. Exploiting barley diversity using the tools of contemporary genomics. Linnean Society of London Symposium on 'The Future of Plant Genetic Resources'.

Inostroza, L, del Pozo, A, Matus, I, Castillo, D, Hayes, P, Machado, S & Corey, A. 2009. Association mapping of plant height, yield, and yield stability in recombinant chromosome substitution lines (RCSLs) using Hordeum vulgare subsp. spontaneum as a source of donor alleles in a Hordeum vulgare subsp. vulgare background. *Mol. Breeding*. 23:365-376.

7. G3008.01: Generating new wheat germplasm with enhanced drought/heat tolerance using AB genomes genetic diversity

October 2008–September 2011

Principal Investigator and Lead Institute

Satish Misra, ARI; satishmisra@yahoo.com; scmisra@gmail.com
Agharkar Reasearch Institute Genetics Department, Agharkar Research Institute
G.G. Agharkar Road, Pune 411 004, India Tel : +91-20-25654357 (Ext 267)
Fax: +91-20-25651542

Collaborating institutes and scientists

- ARI: S Tetali; S Tamhankar
- CIMMYT: M Zaharieva; S Dreisigacker; J Crossa; T Payne
- PBI–University of Sydney, Australia: R Trethowan; P Sharp
- UAS–Dharwad: RR Hanchinal; A Shreenivas Desai; I Kalappanaar; KK Math; B Nirmala Yenagi

Main achievements

1. Research activities and progresses at CIMMYT, Mexico
The main objective of the project was to establish a world collection of emmer wheat (*Triticum turgidum* ssp. *dicoccon*), analyse its genetic diversity and develop a reference set to be used for creating new diversity by direct crossing with hexaploid wheat and/or the development of new synthetic hexaploid wheat (SHW).For this purpose, a well-documented collection of around 300 accessions of emmer wheat has been constituted at CIMMYT. It included accessions originating from 38 countries, provided by CIMMYT and ICARDA gene banks. The collection was multiplied under greenhouse conditions and morphological descriptors, agronomical and physiological traits related to drought and heat tolerance were measured.

One plant per accession was used as founder for generating a stock of seeds available for further research activities. Leaf tissue of this plant was collected, DNA extracted and diversity analysis of the collection is being conducted at CIMMYT using a set of around 50 wheat A- and B-genome microsatellite (SSR) markers. Samples will be sent soon to Australia for DArT analysis. Around 100 good performing emmer wheat accessions from geographically diverse origins will be crossed to three *Ae. tauschii* accessions. Plants are growing and crosses will be carried out in field conditions at El Batan, Mexico.

2. Research activities and progresses at Plant Breeding Institute, University of Sydney
Emmer and durum based primary synthetics developed and sent by CIMMYT have been passed by AQIS (Australian Quarantine Service) and are now growing in quarantine at PBI. After getting a specific import permit, the seed material of *Aegilops tauschii* were also sent by CIMMYT and are now undergoing vernalisation in quarantine.

Emmer wheat accessions maintained at University of Sydney were crossed to Indian elite bread wheats, derived synthethic hexaploids and primary synthetic hexaploids. F1 seed from 30 successful combinations is being grown out for further evaluation, selection and crossing. Seed of the parental materials used will be sent to CIMMYT for genotyping. Another 70 cross combinations based on the emmer wheat accessions sent by CIMMYT are currently in development in the greenhouse.

The development of synthetic wheats has been slightly delayed by problems in the importation of *tauschii* seeds from CIMMYT. These accessions are now under-going vernalisation and will be available for crossing in September this year.

Emmer and durum based synthetics are growing in quarantine and crosses among these materials are underway.

3. Research activities and progresses at Agharkar Research Institute, Pune, India
Seventy (37 emmer based and 33 durum based) primary synthetic wheats sent by CIMMYT were sown along with 9 popular bread and durum wheat varieties at Agharkar Research Institute farm situated at Hol as well at Wellington, where off season nursery of Indian council of Agricultural Research is situated. Intercrossing between durum based and emmer based synthetics was done at Wellington in March 2009. Seeds were harvested and F1 was sown in June at Wellington. F2 generation seed will be grown at Hol in October 2009.

4. Research activities and progresses at University of Agriculture Sciences, Dharwad, India
Synthetic lines received from CIMMYT and being multiplied at ARI, Pune will be shared with UAS Dharwad for their exploitation in the crossing programme.

5. Research activities and progresses at Pakistan Agricultural Research Council, Pakistan
Due to some administrative problems, request has been made to GCP for making separate contract.

Links to previous work
Emmer wheat accessions from the GCP SP1 project "2005-01f: Genotyping of composite germplasm set, tier 1, wheat" were included in the collection.

The diversity of the emmer wheat collection is being analysed using a set of around 50 micro satellite (SSR) markers of the A- and B- genomes established in the frame of the same GCP SP1 project.

Next steps and challenges

1) Variation in morphological traits measured at CIMMYT will complement and expand our knowledge of emmer wheat diversity, and spike and grain characteristics will provide information about yield and its components.
2) Genetic diversity analysis among existing synthetic wheats have been initiated at ARI and CIMMYT in close collaboration
3) Crosses between emmer wheat and *tauschii* accessions will be carried out at El Batan, Mexico; This will allow the creation of three families of SHW with high intra and inter genetic diversity.

8. G3008.02: Improving grain yield on acid soils by the identification of genetic factors underlying drought and aluminium tolerance in maize and sorghum

October 2008–September 2011

Principal Investigator and Lead Institute

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 institutes and scientists

- EMBRAPA Maize and Sorghum: Jurandir Magalhaes; Claudia Guimarães; Robert Schaffert; Reinaldo Gomide; Vera Alves; Flavio Tardin as collaborator; Lauro Guimarães as collaborator; Sidney Parentoni as collaborator
- IGD–Cornell: Stephen Kresovich, Sharon Mitchell, Martha Hamblin
- Dept. of Botany, MU: Sam Gudu
- Robert W. Holley Center for Agriculture and Health , USDA–ARS: Jiping Liu, Lyza Maron, Miguel Pineros, Owen Hoekenga

Our research consortium has recently cloned a major aluminium tolerance gene in sorghum, identified several major maize Al tolerance QTL, and are close to identifying the genes responsible for these QTL. Hence we are poised to significantly advance our programme on improving the acid soil tolerance of maize and sorghum. We will also expand our programme to investigate the molecular determinants for drought tolerance, in order to improve agronomic performance of crops whose root systems are severely restricted by Al toxicity and hence are more susceptible to yield reductions from drought stress on acid soils. Thus, in this proposal, we will use near-isogenic lines, biparental mapping, and association analysis together with physiological and molecular/genomic investigations to:

1) Develop a SNP genotyping platform for high density genotyping in sorghum; **2)** Identify polymorphisms associated with maize Al and drought tolerance; **3)** Use recently developed computer-based image analysis tools to identify root architecture traits in maize and sorghum associated with drought tolerance; **4)** Determine the genetic architecture of very Al tolerant maize and improve maize Al tolerance by introgressing Al tolerance QTLs and/or genes into maize tropical breeding lines; and, **5)** Assess the yield advantage of Al tolerant maize and sorghum in Kenyan environments and begin to investigate the contribution of Al tolerance to drought tolerance.

The expected outcomes are: **1)** Identification of novel sorghum Al tolerance genes that will be ready by the end of the grant to be introduced into our existing collaborative sorghum breeding programme in Africa; **2)** Identification of regions of the sorghum genome associated with drought tolerance, including root architecture traits; **3)** Identification of novel maize Al tolerance genes; **4)** Identification of candidate maize drought tolerance genes via association mapping; **5)** A SNP genotyping array for sorghum that will be an important community resource for many agronomic traits; and **6)** Field testing in Africa of maize and sorghum lines improved for Al tolerance and the role of Al tolerance in drought tolerance.

Subprogramme 2: Genomics towards gene discovery

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

January 2005–June 2009

Principal Investigator and Lead Institute

John Bennett (Jan 2005–Dec 2008); Rachid Serraj (effective Jan 2009), IRRI
International Rice Research Institute; DAPO Box 7777, Metro Manila, Philippines
(63) 2-5805600; r.serraj@cgiar.org

Collaborating institutions and scientists

- CSIRO: Rudy Dolferus
- NIAS: Shoshi Kikuchi; Kouji Satoh
- IRRI: Ken McNally; Richard Bruskiewich; Ramil Mauleon
- TNAU: R Chandra Babu
- NAU: Zhengqiang Ma

Research activities

This project targets the dissection of the physiological and molecular bases of drought sensitivity of reproductive-stage processes, by developing novel drought-phenotyping approaches and integrating the various genomics and functional genomics tools. We used rice and wheat genotypes with contrasting behaviour under stress to identify candidate genes that underlie differences in drought tolerance, using microarray analysis. QTLs were mapped for reproductive-stage processes, candidate genes were identified, and allele mining was used to identify novel alleles contributing to improved physiological traits under reproductive-stage drought stress.

1. Rice genotypes differ in drought sensitivity at the reproductive stage

A set of diverse rice genotypes was used to analyse the response of reproductive-stage processes to a field-managed drought stress. Effects of drought were measured on various morphological attributes, showing that spikelet fertility was reduced by 64 % on average, followed by peduncle elongation rate (PER), and total peduncle length. These effects were correlated with a reduction of plant water status as measured by leaf and peduncle water potentials. Grain yield reduction under water deficit condition was mainly due to an increase in spikelet sterility, which was found to be strongly associated with yield and PER. One of the major causes of spikelet sterility under pre-anthesis drought stress seems to be associated with the inhibition of peduncle elongation, which in turn reduces panicle exsertion, causing sterility in the spikelets left inside the flag leaf. These traits showed significant variation under stress in three mapping populations phenotyped under field-managed drought, thus several QTL could be associated with these traits.

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

A microarray 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 antagonising 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
Two Australian wheat cultivars were compared 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, 3 NCED and 3 ABA 8'-hydroxylases. We confirmed using RT-PCR that these genes are expressed in wheat anthers and ovules.

4. Allele mining
Six candidate genes related to drought-induced growth arrest were analysed, including a gene related to ABA response element binding factor, cellulose synthase, and cytochrome P450. Primers were designed for the six genes involved in ABA-induced growth arrest as response to drought stress. Out of 28 primer pairs, six pairs (representing four candidate genes) that showed a single PCR product of the expected amplicon size were used to screen for variation in 1,536 *O. sativa* accessions. Samples were contrasted against IR64 and Nipponbare. Haplotypes were generated based on specific mismatch combinations from the two contrasts. Haplotypes derived for the four genes showed a tendency to be differentiated following varietal groupings derived from SSR analysis.

Outputs delivered
1. Bioinformatic analysis conducted on rice microarray data to identify key cis elements in drought-responsive genes.
2. Five genes analysed by TILLING
3. Four candidate genes linked to drought-induced arrest of peduncle growth identified

References
Ji XM, Raveendran M, Oane R, Ismail A, Lafitte R, Bruskiewich R, Cheng SH and Bennett J (2005). *Plant Molecular Biology* 59:945-964.
Liu JX, Liao, DQ, Oane R, Estenor L, Yang XE, Li ZC and Bennett J (2006). *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). *New Phytologist* 178:68-79.

10. G3005.02: Revitalising 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 and Lead Institute
Abdelbagi M Ismail, IRRI; abdelbagi.ismail@cgiar.org
Crop and Environmental Sciences Division, IRRI, DAPO Box 7777, Metro Manila, The Philippines; +63 (2) 580-5600 ext. 2843; Fax: +63 (2) 580-5699

Collaborating institutes and scientists
- IRRI: Michael J Thomson; Sigrid Heuer; David Mackill; Glenn Gregorio; RK Singh
- JIRCAS: Matthias Wissuwa
- UoC–Davis: Eduardo Blumwald
- UoD: Zeba I Seraj
- ICABIOGRAD: Masdiar Bustamam
- UoC–Riverside: Timothy J Close

- ABRII: 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; however, development of adapted germplasm is affordable to farmers and is 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
Gene-based markers were developed for *Saltol*, and additional SSR markers were identified and genotyped across the region for fine-mapping, delimiting the QTL to a 1.2 Mb region between 11.4-12.5Mb on chromosome 1. Several different Pokkali alleles at the *Saltol* locus were identified, which complicated the analyses. 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. NILs containing at least two different Pokkali alleles (one in FL478 and another in other RILs 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, suggesting the need for a pyramiding strategy for higher salinity tolerance. A precision MABC 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 MABC 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 at the *Saltol* region, where 17 genes were short-listed based on expression data and physiological role. Nine of these candidates were subsequently cloned and constructed in overexpression or silencing binary vectors for further analysis. 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. The project is also strengthening the capacity of NARES through degree and non-degree training.

Fine-mapping delimited the *Pup1* locus to about 250 kb region on chromosome 12, and gene discovery at the QTL is ongoing, with few candidate genes identified from the DNA sequence of the tolerant Kasalath genotype. Analysis of *Pup1* in Kasalath, Nipponbare, and 93-11 identified major differences in specific gene models, and Kasalath *Pup1* candidates were further investigated through expression and sequence analysis (Heuer et al., 2009). Functional validation of few candidates through RNAi and overexpression is ongoing. Phenotypic analysis showed that *Pup1* may be more beneficial under drought, and this is being further investigated. *Pup1* markers were characterised across diverse rice accessions, and a MABC system was developed and is being used to transfer *Pup1* into popular varieties through the subsequent GCP-SP3 project G4008.41.

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 stable yields even under multiple stresses. Further efforts are also needed to clone the genes underlying *Saltol* and *Pup1*, to develop more precise markers. While this project has laid the groundwork for optimising a MABC system for *Pup1* and few salinity tolerance alleles, more research is needed to add additional tolerance QTLs to meet the challenges of breeding for more stable rice production in marginal environments and to improve farmers' livelihoods.

References

Heuer S, Lu X, Chin J-H, Tanaka JP, Kanamori H, Matsumoto T, De Leon T, Ulat VJ, Ismail AM, Yano M, Wissuwa M. 2009. Comparative sequence analyses of the major QTL Phosphate uptake 1 (*Pup1*) reveal a complex genetic structure. *Plant Biotechnology Journal* 7: 456–47

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

11. 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 and Lead Institute

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

Collaborating institutes 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: Margaret Smith
- RCB–IPB: Utut Suharsono

1. Linkage and association mapping of QTL for major diseases of rice and maize

Based on multiple years of data, QTL for several diseases of maize and rice were more finely mapped. For example, many loci conditioning resistance to northern leaf blight (turcicum), gray leaf spot and southern (maydis) leaf blight of maize have been mapped in the 5,000-line nested association mapping population developed by Buckler et al. We are in the process of cloning several QTL. [Nelson and Balint-Kurti labs]

The maize diversity panel, comprising of approximately 300 diverse inbred lines (Flint-Garcia et al., 2005), was screened for resistance to northern leaf blight (NLB) across four environments. Association analysis of the maize diversity panel for resistance to NLB using ~ 4000 SNP markers (Buckler et al.) resulted in candidate SNPs associated with NLB resistance. Two of the top five SNPs are located in chromosomal regions that had been identified as NLB QTL. A third SNP is located in tasselseed2 (*ts2*), a gene that is the final step of a jasmonic acid pathway that determines floral differentiation, with a role in cell death. An SSR located at *ts2* was found to be under selection between C0 and C4 of a population developed through recurrent selection (RS) for resistance to NLB and was confirmed in two F_2 populations developed from the last cycle of the RS population. We are currently sequencing the *ts2* locus to further determine if this gene is implicated in defense response. [J. Kolkman, Nelson lab]

2. Near-isogenic lines carrying QTL for major diseases of rice and maize.

In rice, a series of NILs were advanced. The recipients, IR64Sub1, SwarnaSub1 and Hopyeongbyeo were grown for checking blast phenotype, morpho-agronomic traits and are being increased for seed for crossing with donors carrying the target gene(s). The donors, IR65482-4-136-2-2, SHZ 2, IR83260-2-10-5-2-1-B and Moroberekan derivative lines, were selected for corresponding defense response (DR) gene and *Pi40* for transfer into the different recipients are in Table 1. The purity of the seeds has been verified and presence of the genes confirmed in some these lines (others are in progress). [Vera Cruz lab]

Two NLB QTL were characterised in detail. Data from repeated greenhouse and field trials revealed that one is effective only against fungal penetration, while another is effective for inducing the accumulation of callose and phenolics surrounding infection sites, as well as inhibiting hyphal growth into the vascular bundle, and the subsequent necrotrophic colonisation in the leaves. Evaluating the NILs with a number of important diseases suggested that in addition to NLB, one conditions resistance to Stewart's wilt and common rust, and the other confers resistance to Stewart's wilt. The non-specific resistance may be attributed to pleiotropy or linkage. [C. Chung, Nelson lab]

A set of SLB QTL was identified and characterised in B73 NILs. A set of 253 NILs carrying various combinations of these loci were tested in replicated field trials assessing resistance to a number of other foliar and ear-rot diseases. Several of the loci appear to confer resistance to more than one disease. One introgression appeared to confer significant levels of resistance to all three foliar diseases. [A. Belcher, Balint-Kurti lab]

Recent publications

Bruce, M., A. Hess, J. Bai, R. Mauleon, M.G. Diaz, N. Sugiyama, A. Bordeos, G.L. Wang, H. Leung and J.E. Leach. 2009. Detection of genomic deletions in rice using oligonucleotide microarrays. *BMC Genomics*, 10:129 http://www.biomedcentral.com/1471-2164/10/129

Carrillo MGC, PH Goodwin, JE Leach, H Leung, and CM Vera Cruz. 2009. Phylogenomic relationships of rice oxalate oxidases to the cupin superfamily and their association with disease resistance QTL. *Rice.* 2:67-79. DOI: 10.1007/s12284-009-9024-0

Carrillo, MGC. 2009. *Genomic analyses of rice oxalate oxidases and their association with quantitative resistance to rice blast*. Ph.D. Thesis. Univ. of the Philippines, Los Banos. 255 p.

Variar, M., C.M. Vera Cruz, M.G. Carrillo, J.C. Bhatt and R.B.S. Sangar. 2009. Rice blast in India and strategies to develop durably resistant cultivars. In *Advances in Genetics, Genomics and Control of Rice Blast Disease*, G.-L. Wang, B. Valent (eds.) DOI 10.1007/978-1-4020-9500-9 35, Springer Science+Business Media B.V. pp. 359-373.

Liu, B., Xiao Yuan Zhu, Shaohong Zhang, Jianli Wu, Seong-Sook Han, Young-Chan Cho, Jae-Hwan Roh, Jan Leach, Yan Liu, Suzette Madamba, Alice Bordeos, Marietta Baraoidan, Isabelita Ona, Casiana Vera Cruz, Hei Leung. 2009. What it takes to achieve durable resistance to rice blast? In *Advances in Genetics, Genomics and Control of Rice Blast Disease*, G.-L. Wang, B. Valent (eds.) DOI 10.1007/978-1-4020-9500-9 35, Springer Science+Business Media B.V. pp. 385-402.

Poland, J. A., Balint-Kurti, P. J., Wisser, R. J., Pratt, R. C., and Nelson, R. J. 2009. Shades of gray: the world of quantitative disease resistance. *Trends in Plant Science* 14 (1):21-29.

Wisser, R.J., Murray, S.C., J.M. Kolkman, H. Ceballos and R.J. Nelson. 2008. Selection mapping of loci for quantitative disease resistance in a diverse maize population. Genetics 180: 583-599.

12. 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 and Lead Institute
Jean-Benoit Morel, Agropolis−INRA; jbmorel@cirad.fr

PROJECT UPDATE NOT SUBMITTED

13. 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 and Lead Institute
F Tardieu, Agropolis–INRA/LEPSE, Montpellier; tardieu@supagro.inra.fr

Collaborating institutes 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

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. Rice is less drought-susceptible than expected for shoot traits *(INRA, IRRI)*
Rice is considered as sensitive to water deficit, with a greater sensitivity of lowland-adapted genotypes. We have analysed the stomatal control and the responses of leaf elongation rate to soil water status and evaporative demand in 7 genotypes of different subspecies, and are upland or lowland-adapted. Stomatal conductance was controlled in such a way that day-time leaf water potential was nearly maintained in well-watered, droughted or flooded conditions. This resulted in a low sensitivity of leaf elongation rate to evaporative demand, much lower than that observed in maize, with small differences between studied genotypes. The sensitivity of leaf elongation rate to soil water deficit was similar to that of maize. Although observed differences between genotypes were small, a tendency was observed towards a lower sensitivity of upland genotypes than of lowland genotypes to both evaporative demand and soil water deficit. Hence, gas exchanges and leaf growth of rice are not particularly sensitive to water deficit, so detailed omics studies comparing genotypes on these traits may well not be successful. The main origin of rice sensitivity is its root system.

2. Common QTLs for leaf growth between phenotyping platforms and field experiments in maize *(INRA, CIMMYT, KARI, IARI)***.**
We have analysed traits linked either to maize yield or to leaf area in a network comprising 6 experiments with different watering conditions and 210 recombinant inbred lines (RILs).

- Common QTLs were identified for grain yield, grain number, leaf length and total leaf area. The favorable allele came from the drought tolerant parent. There was a good relationship between grain yield an total leaf area, which applied within individual experiments for the different RILs, and between experiments with different water regimes.
- QTLs detected from the field experiments were consistent with QTLs previously detected in the phenotyping platform. (i) 3 colocalisations of QTL of leaf area in the field with QTLs of leaf growth rate measured in the phenotyping platform in favourable conditions. (ii) 3 colocalisation of QTLs of sensitivity of leaf growth to water deficit, evaluated either in the field via contrasts between experiments, or in the phenotyping platform via the responses to evaporative demand and soil water status.

- Hence, the genetic variability in traits determined during the vegetative phase in good part determines that of yield as measured in field experiments. Phenotyping platforms provide genetic information which partly predicts that observed in networks of field experiments.

3. 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).

- Roots. QTLs for the elongation rates of axile roots responded to water stress, more than those of root length or of axile root lengt. Two major QTLs were detected, among which one co-located with a major QTL for the anthesis-silking interval (ASI) in field experiments.
- Leaves. Six 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. Some of them colocate in platform and field experiments (see §2)
- Silks. Half of QTLs of leaf growth maintenance colocate with QTLs of silk growth (ASI) in dry fields in Mexico (CIMMYT).

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.

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

- *Where and when sample?* Leaf growth is restricted to specific zones and times. It is essential to take these spatial and temporal patterns for omic sampling. Rules proposed in 2007 have been essentially confirmed and are now published.
- *Dealing with fluctuating temperatures.* Field experiments requires the calculation of temperature-compensated time. Unfortunately, the classical thermal time method does not apply to rice. We have proposed a new method currently under publication.

14. G3005.16: Isolation and characterisation of aluminium tolerance genes in the cereals: an integrated functional genomic, molecular genetic and physiological analysis

January 2005–December 2007; no-cost extension to December 31, 2008

Principal Investigator and Lead Institute

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 institutes 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; Euclydes Minella; Eduardo Caierão; Alfredo Nacimento Junior
- EMBRAPA Rice and Bean: Flavio Breseghello; Pericles Neves
- MU: Sam Gudu
- Robert W. Holley Center for Agriculture and Health, USDA–ARS, USA; Lyza Maron; Owen Hoekenga; Jiping Liu; Ed Buckler

Research activities and progress

Sorghum and maize Al tolerance. *Sorghum.* Genetic, biochemical and physiological characterisation is ongoing for the major sorghum Al tolerance gene we cloned earlier in this project, Alt_{SB}, 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_{SB} alleles in the same genetic background to identify the best Alt_{SB} 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_{SB} to facilitate maximal Alt_{SB} expression and tolerance. We are currently producing highly Al tolerant lines useful for breeding, in work on our more recently funded GCP grant on advancing sorghum Al tolerance. Finally, at the biochemical level, we have identified a novel Al-binding protein that binds to the Alt_{SB} protein very tightly. It is possible this protein plays a role in the Al activation of citrate transport by Alt_{SB}

Maize. Al tolerance in maize is a typically complex, quantitative trait. We have conducted QTL mapping for maize Al tolerance using a n RIL population generated from Al tolerant and sensitive Brazilian maize inbred lines. 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 also had previously in this project conducted a detailed analysis of root gene expression under Al stress using maize microarrays (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_{SB}, 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. The expression patterns from the microarray studies were also confirmed with quantitative real time PCR analysis.

The MATE gene family is large and complex in plants, and especially 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_{SB}. However, it is interesting to note that neither *ZmMATE1* nor *ZmMATE2* are close homologs of Alt_{SB}, and therefore would not have been identified without the integration of genetic and gene expression profiling approaches. We now have strong functional evidence that *ZmMATE1* is the maize root citrate transporter that plays an important role in maize Al tolerance. We have developed transgenic maize lines overexpressing *ZmMATE1* and once we have lines homozygous for the transgene, we will test these lines for Al tolerance and root citrate exudation.

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 Sega 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 bred and highly selected for superior performance in acid soils. In order to improve acid soil adaptation, we are undertaking to create inbred lines from Brazilian single crosses, 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 in acid soils. We are following these lines both in sorghum and maize for production of inbreds, hybrid seeds and synthetic varieties. In this proposed project we are seeking funds to allow us to produce high yielding hybrid and synthetic varieties suited to acid soils of Kenya.

Tangible outputs. 1) Identification of highly aluminium tolerant alleles of our sorghum Al tolerance gene, Alt_{SB}, that are being used to develop breeding lines. 2) Have verified that additional sorghum Al tolerance genes exist, which we are currently working o identify. 3) Identification of strong candidates for the first identified maize Al tolerance genes, *ZmMATE1* and *ZmMATE2*. 4) Developing maize breeding lines improved for Al tolerance based on QTL and *ZmMATE* information. 5) Developed recombinant maize lines between crosses between Brazilian and Kenyan materials for ongoing field testing on acid soils in Africa.

15. G3007.03: Development of genomics resources for molecular breeding of drought tolerance in cassava

September 2007–February 2010

Principal Investigator and Lead Institute

Pablo Rabinowicz, Institute for Genome Sciences, and Department of Biochemistry and Molecular Biology, University of Maryland School of Medicine prabinowicz@som.umaryland.edu; 801 W. Baltimore St., Baltimore, MD 21201, USA. Tel: (1) (410) 706-6714

Collaborating institutes and scientists

- ACGT–UoP: Jane Morris
- ACGT–FABI, UoP: Alexander Myburg
- ACGT–School of Molecular and Cell Biology, University of the Witwatersrand, South Africa: Chris Rey
- UoC–Davis: Ming-Cheng Luo

Context: Cassava is a major staple crop in developing countries, especially in sub-Saharan Africa where environmental conditions are often extreme due to severe drought. This project aims to develop an extensive set of genome-wide single nucleotide polymorphism (SNP) markers in cassava, taking advantage of a fingerprint-based physical map generated during the first phase of the project. A SNP-based genetic map will be built using a cassava mapping population of contrasting phenotypes for drought tolerance.

Findings and implications: A bacterial artificial clone (BAC) library constructed from CIAT's cassava inbred line AM560-2 has been fingerprinted. As a result, using the programme FingerPrint Contig (FPC), 58,244 BAC clones could be assembled into 2,104 contigs and 5,054 singletons. The longest contig spans ~4.8 Mbp and includes 357 BAC clones. The average number of BAC clones per contig is 25 and the average estimated length is 0.41 Mbp. A set of minimally overlapping BAC clones (minimum tiling path or MTP) has been selected from the assembled contigs. This MTP includes 6,868 BAC clones and its total span is 710 Mbp, which is consistent with the estimated size of the cassava genome.

In order to identify SNP markers as evenly distributed throughout the genome as possible, we sequenced the ends of all 6,868 BAC clones in the MTP and also of an additional 2,000 BAC clones that remained as singletons after assembly.

All FPC and sequence data is available at a public project website (http://cassava.igs.umaryland.edu) in the form of genome browser where users can search and visualise all the BAC contigs and access the end-sequences when available (Figure 1). In addition, the website includes a BLAST server where users can compare their own sequences to our cassava BAC-end sequences, the available cassava EST sequences, an assembly of all cassava ESTs, and the genomic sequences generated by the U.S. D.O.E. Joint Genomics Institute that cover approximately 70% of the cassava genome.

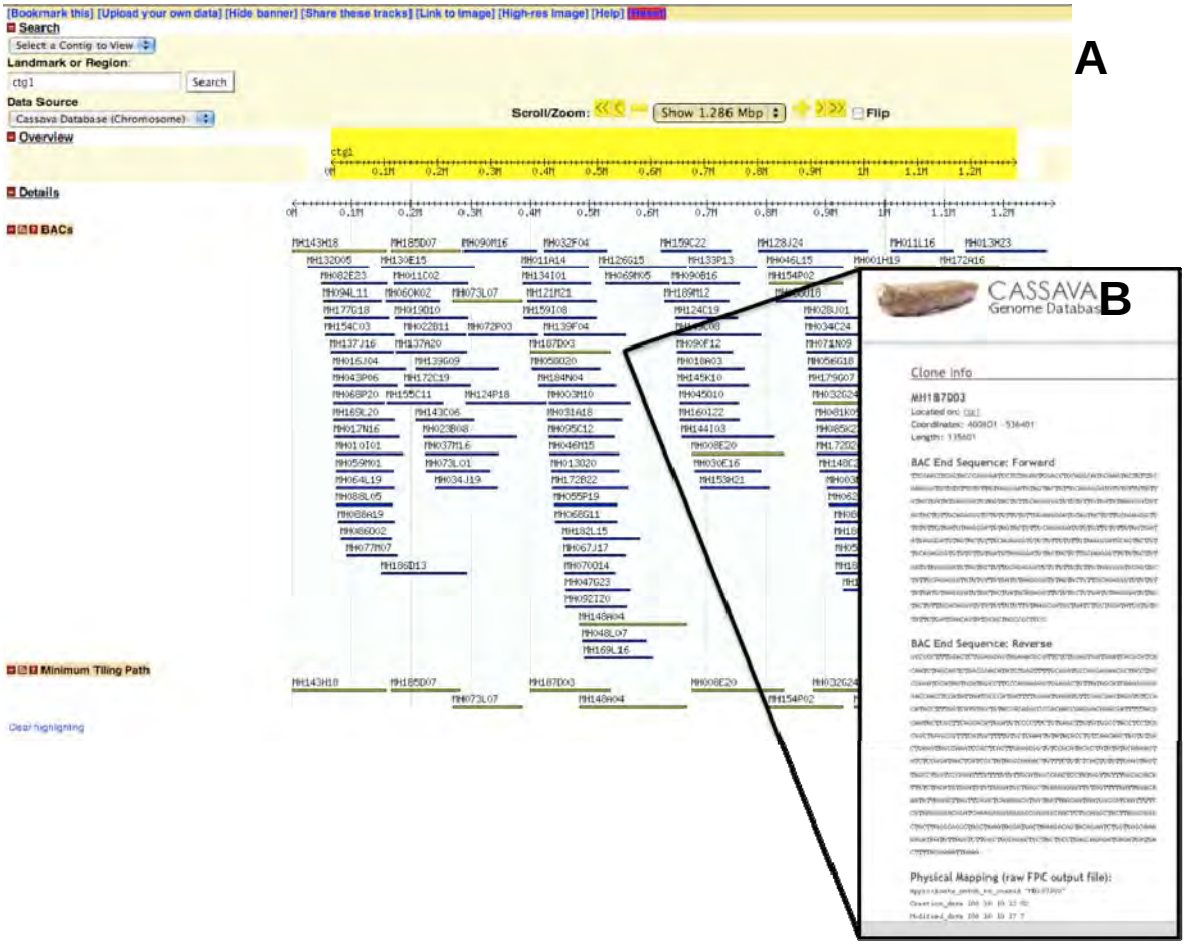


Figure 1. Screen capture of the cassava genome browser displaying a 1.3 Mbp fingerprint BAC contig (A). Each bar represents a BAC clone and those selected for the MTP are shown in yellow. An additional track displays only the MTP of BAC clones (bottom). Clicking in each BAC clone brings sequence and fingerprint information on that particular clone (B).

Next steps

Using the BAC-end sequences, a set of 1,536 primer pairs has been selected to amplify low-copy sequences separated at least by 200 kbp from each other. These primers have been already synthesised and will be used to amplify the same 1,536 loci in 10 diverse cassava genotypes to identify approximately 1,000 SNPs. All data will be made available in a user-friendly manner in the project website. These SNPs will also be genotyped in the cassava mapping population using the Illumina BeadXpress platform that has been set up at CoPI Myburg's lab at the University of Pretoria, South Africa. This instrument will also be used for a Capacity Building Workshop to train collaborators from National Programmes in high throughput SNP genotyping to be held in November 2009.

16. G3007.06: Genetic dissection of drought adaptive mechanisms in bread and durum wheat through large scale phenotyping methodologies

November 2007–October 2009

Principal Investigator
Matthew Reynolds, CIMMYT; m.reynolds@cgiar.org
AP 6-641, 06600 Mexico D.F., Mexico; Ph: +52 (55) 5804-7659

Collaborating institutes and scientists

- CIMMYT: Daniel Mullan; Jose Crossa; Yann Manes
- ACPFG: Peter Langridge; James Edwards; Dion Bennett
- DWR: Jagadish Rane; B Mishra; Ravish Chatrath; Sindhu Sareen
- ARI: 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*) have been genotyped and phenotyped under multiple environments in Mexico, Australia and India. QTL discovery has allowed 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

Extensive field trials have been conducted by CIMMYT (Cd. Obregon and El Batan, Mexico), ACFPG (Multiple sites in S.A., Australia), ARI (Pune, India) and DWR (Karnal, India) during the past three 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 undergoing component trait analysis and QTL discovery. Research highlights for each of the mapping populations include:

Kukri x RAC875

A major heat QTL has been identified on 3BL. The QTL accounts for 12% of genetic variation for yield and results in a 1.3^{o}C cooler canopy, which translates to a 30% yield advantage. The QTL on 3B and additional QTLs on 6A (yield and grain size) and 7A (yield and grain number) are now the focus of fine mapping activities.

Excalibur x Kukri

Four maturity QTL (2BS, 4AL, 5AL, 7AS) have been identified that have significant effects on all other physiological traits measured. QTLs on chromosomes 1B, 6A and 7A are currently undergoing fine mapping.

Atil x T. dicoccum

Primary map construction for the Atil x *T. dicoccum* population has been completed. The map consists of 894 markers (853 DArTs and 41 SSRs), with an average marker spacing of 2.8 cM. The map is now being used for the identification of novel QTLs.

Table 1. QTLs targeted for fine mapping

Population	Location of QTL	Traits
DH1	3B	Yield under heat stress (12% variation), Canopy temperature
	6A	Flag leaf width, Grain size, Early vigor
	7A	Yield, Grain number
DH2	1B	Yield (13% variation), Grain size, Tipping, Grain filling duration
	6A	Yield, Grain size, Tipping, HI, Wax, Crown rot
	7A	Yield

Glasshouse experiments have been designed to assess drought adaptive traits in a controlled environment. Two subsets have been selected from the 11th SAWYT and Seri/Babax RILs for determining the most suitable traits for application in high throughput phenotyping. The first cycle of analysis has indicated that grain number, SPAD, leaf rolling and leaf waxiness are all likely targets for high throughput phenotyping, with a further cycle of analysis to be conducted during 2009. Additional activities include the generation of mapping populations that are not confounded by phenology. Six populations have been selected and are currently at F_4 stage of population development, with an 8–12 day range in anthesis.

17. G3008.03: Delayed senescence and drought-tolerance in rice

October 2008–September 2011

Principal Investigator and Lead Institute

Eduardo Blumwald, Department of Plant Sciences, UoC–Davis; eblumwald@ucdavis.edu
University of California, Davis, One Shields Ave, Davis, CA 95616, USA; (01)530-752-4640; Fax: (01)530-752-2278.

Collaborating institutes and scientists

- IRRI: Abdelbagi Ismail; Rachid Serraj

1. Research activities and progresses at UC Davis and IRRI collaborators

We hypothesised that drought induced plant senescence is due to a type of cell death programme naturally activated during drought and that the down-regulation of such programme would enable plants to maintain efficient source-to-sink relationships during stress, thus enabling plants to acquire vigorous acclimation responses to water deficit resulting in enhanced drought tolerance with reduced yield losses. We generated transgenic plants overexpressing an isopentenyltransferase (IPT) gene (a key enzyme mediating the synthesis of cytokinins) under the control of SARK (senescence activated receptor kinase), an inducible maturation- and stress-dependent promoter. Our results demonstrated that the suppression of drought-induced leaf senescence resulted in significantly enhanced drought-tolerance of the plants (1). The transgenic plants expressing *pSARK-IPT* maintained relatively high relative water content, retained photosynthetic activity and survived longer periods without irrigation (1). Moreover, the plants overexpressing *pSARK-IPT* were able to grow under a restrictive water supply with a lower yield penalty as compared to controls and displayed minimal yield losses when watered with only 30% of the amount of water used under control conditions. There were no significant differences in stomatal conductance between leaves from wild-type (WT) and transgenic pSARK-IPT plants grown under optimal or restricted watering. On the other hand, there was a significant reduction in the maximum rate of electron transport as well as the use of triose phosphates only in WT plants during growth under restricted watering, indicating a biochemical control of photosynthesis during the growth under water deficit (2). The transgenic plants displayed an increase in catalase inside peroxisomes, a physical association between chloroplasts, peroxisomes and mitochondria and an increase in the CO_2-

compensation point, indicating the CK-mediated occurrence of photorespiration in the transgenic plants. The transcription of genes encoding enzymes associated with photorespiration was further enhanced in the transgenic plants grown under restricted watering conditions, indicating a cytokinin-induced increase in photorespiration and suggesting the contribution of photorespiration in the protection of photosynthetic processes and its beneficial role during water stress (2).

We have tested the efficiency of this approach in the monocot plants rice (*Oryza sativa* L.). Rice cv. Kitaake was used to produce transformed homozygous lines expressing *pSARK-IPT*. Transgenic, wild-type (wt) and null plants, were tested in control greenhouse conditions under two water-stress treatments at different developmental stages: (***i***) *pre-anthesis* (booting) stage and (***ii***) *post-anthesis* stage (two weeks after panicle initiation), by slowly drying water-logged pots until visual stress symptoms (~12 days) appeared. Pots were then re-watered and plant productivity, morphological, physiological, and phenological parameters were collected from plants after maturation. In general, both stress treatments had a significant (P<0.001) effect on productivity parameters. When compared with the wt plants, the P_{SARK}-*IPT* rice plants displayed delayed leaf senescence under both stress treatments. Under the *pre-anthesis* treatment, the transgenic plants produced up to 80% of grain yield (GY) (Figure 1) and 98% of 1000-grain weight (TGW) (Figure 2) as compared to the wt plants grown under well-watered conditions, whereas the wt plants showed a dramatic reduction in GY (26%) or TGW (85%). Water stress during panicle initiation (*pre-anthesis*) causes severe reduction in flower fertility, as shown in the wt plants, with 60% reduction in grain number, while only a 20% reduction in grain number was seen in the transgenic plants. At the later stage (*post-anthesis*), water stress predominantly affected the translocation of assimilates from vegetative organs to the developing grain, nevertheless the transgenic rice plants showed significantly greater GY than the wt. Our results indicate the possibility of generating transgenic rice plants with increased water use efficiency and increased tolerance to water deficit.

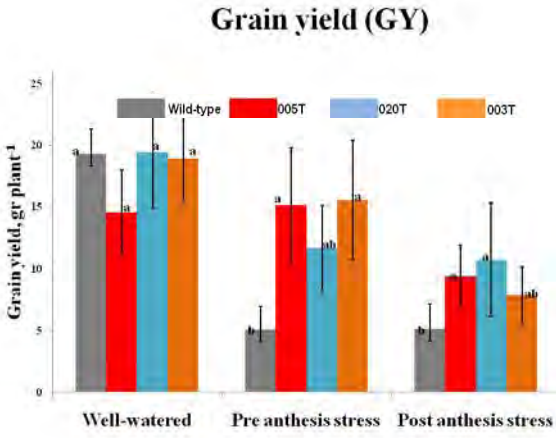


Figure 1 – Grain yield per plant. Plants were water-stressed at pre-anthesis or at post-anthesis and grain was harvested 5 months after germination. Wild-type and three independent transgenic lines (T3) are indicated.

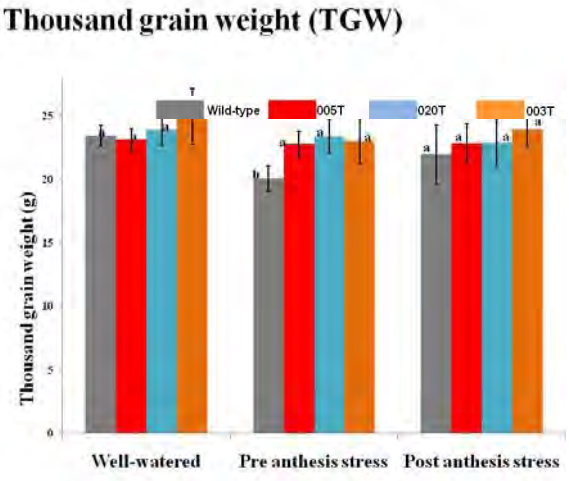


Figure 2 – Thousand grain weight. Grain was harvested 5 months after germinated. Pre- and post-anthesis stress was applied as described in Fig. 1. Wild-type and three independent transgenic lines (T3) are indicated.

References

Rivero RM, Shulaev V, Blumwald, E. (2009) Cytokinin-dependent photorespiration and the protection of photosynthesis during water deficit. *Plant Physiol.* 150:1380-1393.

Rivero, RM., Kojima, M., Gepstein A., Sakakibara, H., Mittler, R., Gepstein, S., & Blumwald, E. (2007). Delayed leaf senescence induces extreme drought tolerance in a flowering plant. *Proc. Natl. Acad. Sci*. USA 104:19631-19636.

18. G3008.04: Drought from different perspective: Improved tolerance through Phosphorus acquisition

October 2008–September 2011

Principal investigator
Sigrid Heuer; IRRI; sheuer@cgiar.org
Plant Breeding, Genetics and Biotechnology Division (PBGB)
DAPO box 7777, Metro Manila, Philippines
phone: 0063 (2) 5805600- 2761

Collaborating institutes and scientists

- JIRCAS: Matthias Wissuwa
- ICABIOGRAD: Masdiar Bustamam; Joko Prasetiyono
- IRRI: Stephan Haefele; Arvind Kumar; Abdelbagi Ismail
- MPIDB: Bernd Mueller-Roeber; Slobodan Ruzicic
- ZU: Ping Wu

Phosphorus (P) deficiency and drought are companion stresses coinciding in rainfed upland and lowland environments. P deficiency can aggravate the negative effect of drought because it reduces root growth and thereby access to water. The major QTL *Pup1* was shown to confer tolerance of P deficiency by maintaining root growth under P stress and converging evidence suggest that *Pup1* is also beneficial under drought (see SP3 project "Application and Validation of the Major QTL *Phosphate Uptake 1*). The *Pup1* genomic region was recently sequenced in the tolerant donor Kasalath and 68 putative genes, including retro-/transposon-related elements (TEs), were predicted (Heuer et al 2009). Many of these putative genes contained fragments of genes joined with TEs and detailed analyses were needed to determine if those genes are functional or represent truncated genes. Based on *in silico* validation of gene models in conjunction with RT-PCR gene expression analyses, the number of candidate genes has now been reduced to six putative genes (fatty acid oxygenase, unknown protein, dirigent, hypothetical protein, protein kinase, Zn-knuckle). In agreement with earlier data, none of these genes is obviously associated with P uptake and *Pup1* might therefore confer tolerance by an unknown mechanism or a regulatory genes. In the centre of attention is currently the protein kinase gene *PupK46* since it might be involved in P signaling and signal transduction by e.g., phosphorylation of high-affinity P transporters. Transgenic plants (T0 overexpression, T2 RNAi plants) are currently growing to analyse the function of this gene in detail. A yeast-2-hybrid screen is in preparation to identify target proteins that are phosphorylated by this kinase.

Allelic sequencing of candidate genes in tolerant and intolerant rice accessions is ongoing and already revealed a tolerant-specific allele for the dirigent gene (*PupK20*). In a first attempt to address putative functions of this gene, the lignin content in roots of drought and P-stressed plants was quantified in contrasting *Pup1* near isogenic lines (NILs). The data showed that lignification of roots is a ubiquitous drought response and that this response is absent in plants grown hydroponically with PEG-induced water deficit (Fig. 1). The quantification of differences between contrasting *Pup1* NILs and between different P treatments is ongoing. Overexpression plants have been generated (T0) and T2 RNAi plants are currently being analysed. Promoter::GUS constructs for *PupK20* and *PupK46* are in preparation. Extensive analyses have been conducted with *Pup1* NILs to assess other root traits associated with *Pup1* and tolerance mechanisms. Root hair length and number, as well as number of adventitious roots were compared in several *Pup1* NILs grown under stress (P and drought), and control conditions (well watered, +P). The data are currently being analysed.

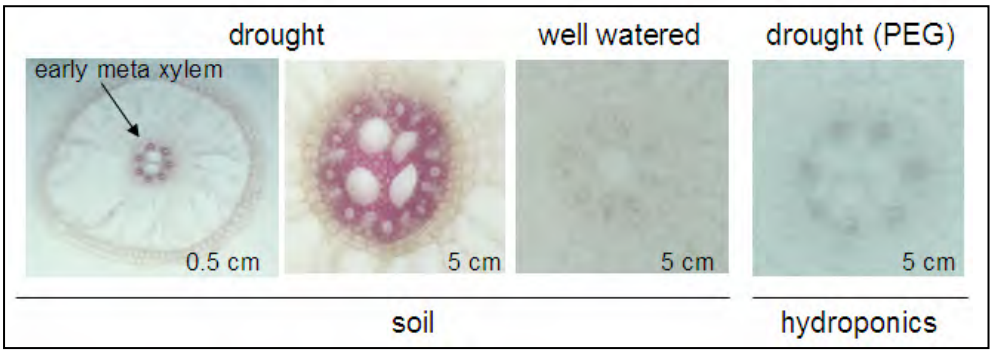


Figure 1. Lignification of drought-stressed roots. Plants were grown either in soil or hydroponics culture solution and drought stress was induced by withholding water or adding polyethylene glycol (PEG), respectively. Lignin was stained with Fluroglucinol in cross sections at 0.5 and 5 cm from the root tip.

References

Heuer S, Lu X, Chin JH, Tanaka JP, Kanamori H, Matsumoto T, De Leon T, Ulat VJ, Ismail AM, Yano M, Wissuwa M (2009). Comparative sequence analyses of the major QTL Phosphate uptake 1 (Pup1) reveal a complex genetic structure. *Journal of Plant Biotechnology Journal*: 456–471.

19. G3008.05: Discovery and development of alleles contributing to sorghum drought tolerance

March 2009–February 2012

Principal Investigator and Lead Institute

Andrew H Paterson, Plant Genome Mapping Laboratory, UGA; paterson@uga.edu
Plant Genome Mapping Laboratory, University of Georgia, 111 Riverbend Road, Rm 228, Athens, GA 30602; Phone 1-706-583-0162; Fax 1-706-583-0160; URL: www.plantgenome.uga.edu

Collaborating institutes and scientists

- SARI–Ghana: IDK Atokple
- ICRISAT: C Thomas Hash
- MAU: SP Mehtre
- NRCS: Nadoor Seetharama

Executive summary

Sorghum is the most drought-tolerant dual-purpose (grain + straw) cereal crop of the semi-arid tropics and subtropics, where development challenges are the greatest and market failure is most acute. As such, it is both a priority for further improvement and a botanical model from which we might glean information about drought tolerance that might be leveraged in improvement of many other cereals by comparative approaches. Sorghum has recently become only the second cereal (after rice), to have its genome fully sequenced, opening new doors to its improvement and enhancing its value for comparative biology.

In a partnership joining African and Asian sorghum improvement researchers with genomic scientists experienced in crop breeding and germplasm enhancement, we will engage the sorghum sequence in a balanced approach to durably increase rates of sorghum improvement. Toward a pathway joining discovery research of increasing scope and sensitivity with application to the needs of resource-poor farmers living in drought-prone environments, early study of a few genes already known to have qualitative effects on drought tolerance will set the stage for identifying a growing pipeline of additional genes/alleles with more subtle effects, engaging several previously GCP-funded resources. Key to both discovery research and product development/delivery will be our focus on breeding populations in which drought tolerance will be combined with other traits that address production constraints in West

and Central Africa, Eastern and Southern Africa, and South Asia. By applying sorghum's fully-sequenced genome to study of these field-proven genetic resources, we will elucidate genotype x environment interactions that render drought tolerance a difficult trait to work with. Improved knowledge of sorghum presents a singularly-promising opportunity to leverage comparative genomics approaches to benefit improvement of many other cereals. NARS scientists are full research partners, and will also benefit from training visits to UGA and/or ICRISAT.

Scientific summary

Fine-scale characterisation of qualitative factors related to sustained photosynthetic output under drought stress, will yield improved understanding of the structure and functions of genomic regions of importance to sorghum improvement, advancing genetic dissection and molecular cloning of genes conferring the phenotypic effects mapped to these regions.

Empirical testing of key combinations of *stay-green* QTLs will reveal the comparative efficacy of various combinations under stress conditions in Ghana and India, also providing for obtaining additional recombinants needed to reduce linkage drag. The sorghum sequence will provide us with the means to design comparative DNA markers that are suitable for utilisation both in sorghum and in many additional cereals.

Transcriptome profiling of a diverse sampling of field-proven germplasm will support development of hypotheses about roles of specific genes and pathways in drought response. We will begin to test these and other hypotheses based on analysis of the sorghum sequence, using breeding populations in which drought tolerance will be combined with other traits addressing production constraints in West and Central Africa, Eastern and Southern Africa, and South Asia.

Subprogramme 3: Trait capture for crop improvement

20. 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 and Lead Institute

Alfredo Alves, EMBRAPA/CNPMF; 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

- EMBRAPA/CNPMF: Miguel Dita; Antonio Souza
- EMBRAPA/CPATSA: Alineaurea Florentino Silva
- CIAT: Martin Fregene; Hernán Ceballos
- Cornell: Tim Setter; Luis Duque
- IITA: Morag Ferguson; Rosemary Mutegi
- ARI–Naliendele: Geoffrey Mkamilo
- SARI–Ghana: Joseph Adjebeng-Danquah
- KARI: Joseph Kamau

1. Introduction

Cassava is one of the most important sources of food for the resource-poor in the tropics, and it is particularly valued in drought-prone areas as a contributor to food security. Given cassava's long breeding cycle and duration to harvest maturity, there is a strong incentive to enhance phenotyping methods and to use molecular markers that will assist the breeding process.

2. Evaluation of contrasting cassava varieties and identification of traits related to drought tolerance

In several trials, carried out in Brazil, Colombia, Tanzania and Ghana, drought tolerance contrasting cassava varieties have been screened and evaluated under drought environments and we have identified physiological traits that are associated with drought tolerance. Among these, are traits related to developmental regulation that permits the most drought tolerant lines to optimally partition resources for storage root initiation and maintenance during drought.

In general, the functional traits that cassava uses to achieve its drought tolerance and yield performance are related to growth parameters, leaf retention, chlorophyll content, leaf conductance, leaf temperature, abscisic acid-ABA and carbohydrates accumulation. At early phase, partitioning ratio to storage root had the best correlation with yield, showing that the phase of storage root initiation is most vulnerable to drought and that phenotyping in this stage indicates genotypic tendency for partitioning growth toward storage roots, which is an indicator of potential for high harvest index. Water status remains high during water deficits due to the sensitivity of stomata conductance, which is positively correlated to yield. But, leaf conductance is a dynamically changing trait. Leaf retention is positively correlated to yield mainly under drought condition. Substantial amounts of starch are stored in stems and petioles which is remobilised during stress. Genotypes with more stem starch might be better able to sustain activities during prolonged stress and tolerate cycles of stress. Better genotypes keep ABA levels low, apparently not due to better stomata or root function, but might be due to growth inhibition by high ABA. Considering that heritability of yield was low, while other drought tolerance traits had higher heritability, the use of a selection index that permits stacking of several tolerance traits is feasible.

3. Development and evaluation of segregating populations for drought tolerance

Three F1 mapping populations were generated and the cross (MCol 1734 x MVen 77) with the largest population of 235 individuals has primarily been selected for use in QTL analysis for drought tolerance. This population and its parents were distributed by CIAT to Brazil, Kenya, Ghana, Tanzania and Nigeria. While genotyping of this population has completed with SSR markers, phenotyping has yet to be done.

Since this drought tolerance segregating population was originated from parents not adapted to Africa, it will be field tested for drought tolerance QTLs only in Brazil and Colombia. However, and because the material has been multiplied and shipped to 4 NARS in Africa (Kenya, Ghana, Tanzania and Nigeria), this population will be grown (as such, or after a cross with material well adapted to local conditions) under field conditions to see potential new sources of alleles for African germplasm.

The 235 individuals of the mapping population MCol 1734 x MVen 77 were genotyped with 187 polymorphic markers identified at IITA-Nairobi and will be phenotyped under drought and control conditions using the traits previously identified as related to drought tolerance. Genotypic and phenotypic data will be analysed to identify QTLs for drought tolerance in cassava.

3. Tangible outputs delivered
Traits highly correlated to yield performance under water deficit conditions identified and phenotyping plans established

Three mapping population segregating for drought tolerance developed and shipped to Brazil, Kenya, Ghana, Tanzania and Nigeria.

Genotypes (374) from three mapping population genotyped with, at least, 180 SSR markers.

21. 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 Recursos Geneticos e Biotecnologia
Parque Estação Ecologica Norte, Final Av W5 Norte, Brasília-DF, 70770-900, Brazil;
valls@cenargen.embrapa.br

Collaborating institutes and scientists

- UCB: David Bertioli
- UCG: Wellington Martins
- CERAAS: Ousmane Ndoye
- ICRISAT: Vincent Vadez; Rajeev Varshney; Aruna Rupakula
- UAS–Bangalore: Udaya Kumar
- Agropolis–CIRAD: Angelique d'Hont
- IBONE: Guillermo Seijo: Germán Robledo
- UoAs: Jens Stougaard; Lene Madsen; Niels Sandal
- TAMU: 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 characterization 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 characterized 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* hybridization (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.

22. G3005.06: Marker development and marker assisted selection for Striga resistance in cowpea

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

Principal Investigator and Lead Institute
Satoru Muranaka, IITA; smuranaka@cgiar.org
IITA Kano, Sabo Bakin Zuwo Road, PMB 3112, Kano, Nigeria

Collaborating institutes and scientists

- IITA: Christian Fatokun; Adebola Raji; Boukar Ousmane; Dong-Jin Kim
- UoV: Michael Timko
- CERAAS: Ndiaga Cisse
- ISRA–CNRA: Moctar Wade

1. Research activities and progresses

The present project seeks to develop molecular markers and to establish MAS method for *Striga* gesnerioides resistance in cowpea. Under this project, four major activities such as Marker development for *Striga* resistance, Identification of genetic resource for *Striga* resistance, development of useful tool kit and capacity building are geared towards achieving effective and efficient breeding for *Striga* resistance not only by IITA, but also by NARS partners in West and Central Africa.

During the project period, two SCAR markers 61R and MahSE2 were identified as reliable markers for *Striga* race 3 with high efficiency (80% and 77% respectively). Three new markers for *Striga* resistance have been developed for several *Striga* races in West and Central Africa. Also, several cowpea populations have been evaluated for their reactions to other *Striga* races in order to identify additional markers. High MAS efficiency of MahSE2 to SG3 resistance was confirmed using populations derived from IT06K-43 series at F_3, F_4 and F_5 generations.

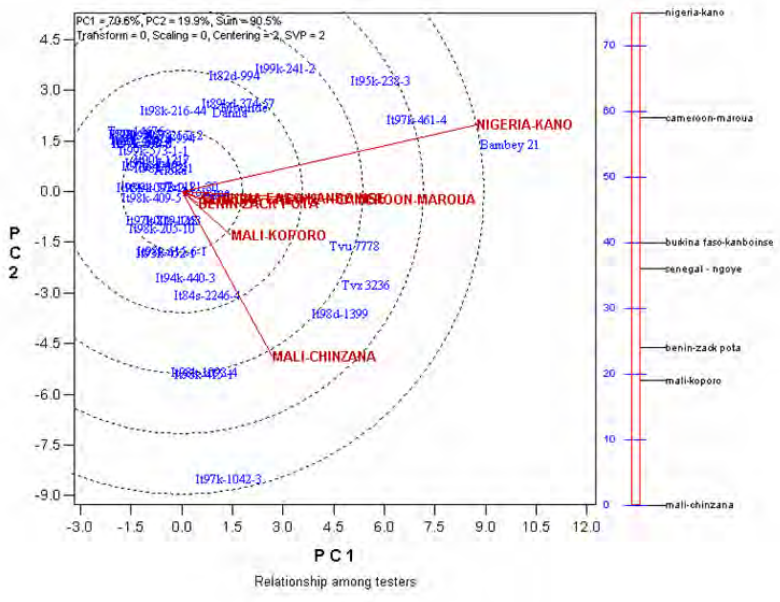


Figure 1. GGE biplot analysis of emerged *Striga* number in 7 *Striga* hot-spots

The *Striga* hot spot trial clearly showed that the ability of cowpea genotypes to resist *Striga* parasitism depends on geographic origin of the parasite. From a three year trial, three germplasm lines were confirmed to exhibit resistance to multiple *Striga* races while several breeding lines resistant to SG4z were identified. However, results of diversity analysis in Senegal showed clear differences in parasitism among the *Striga* collected from within same district thus suggesting a possibility of existence of new *Striga* races.

Three RIL populations were phenotyped for drought tolerance. Significant differences were found among each RIL population for drought tolerance. The phenotypic data will be utilised for QTL analysis.

Table 1. Performance of RIL population and parents under drought condition

Characters	Comparison between parents		RILs
	Bambey 21	Mouride	Means and S.D
50% flowering	41.7*	54	45.4 ± 3.0
95% maturity	57.33	75.3**	66.5 ± 8.8
Pod lenght	13.0*	11.3	10.0 ± 5.3
grain weight/plant	7.6	17.5*	
Biomass yield	1358.8	2375.4**	1420.9 ± 780.5
grain weight/plant,	36	74	33.9 ± 36
grain weight/pod	4.8	14.2	5.138 ± 7
Grain yield	270.7	583.3**	155.3 ± 2

Under this project, many scientists and students from different regions of Africa received specific hands on instructions in general aspects of plant molecular biology, physiology and breeding.

References

Timko, M.P., Ehlers, J.D., and Roberts P.A. 2007. Cowpea. In: *Genome Mapping and Molecular Breeding in Plants*, Volume 3, Pulses, Sugar and Tuber Crops, C. Kole (Ed.), Springer Verlag, Berlin Heidelberg. pp. 49-67.

Timko, MP, Gowda, BS, Ouedraogo, J. and Ousmane, B. 2007. Molecular markers for analysis of resistance to Striga gesnerioides in cowpea. In. *Integrating New Technologies for Striga Control: Towards Ending the Witch-hunt*. G. Ejeta and J. Gressell (Eds.) World Scientific Publishing Co. Pte Ltd, Singapore, pp. 115-128

Timko, M.P. and B.B. Singh. 2008. Cowpea, a multifunctional legume. In: *Genomics of Tropical Crop Plants*, Paul H. Moore and Ray Ming (eds) Springer-Press. pp. pp. 227-258.

Tonessia, D. C., Wade, M., Cissé, N, Timko, M.P and N'Diaye, A. 2007 : *Genetic Structure and Analysis of Striga gesnerioides (Witchweed) from Senegal* (Poster presented in May 2007 by the student Charlotte at the Kenyan meeting on biotechnology)

E. M. Agbicodo, C. Fatokun, S. Muranaka, G. Linden and E. Visser, 2008, *Phenotyping cowpea recombinant inbred lines (RILs) for drought tolerance* (Poster presentation in Sept 2008 by r. Linden at international Gordon Conference on Water and Salt stress in plants, Big sky, MT, USA)

E. M. Agbicodo, C. Fatokun, S. Muranaka, E. Visser and C. G. Linden, 2009, Breeding drought tolerant cowpea" constraints, accomplishments, and future prospects. Euphitica 167: 353-370

23. 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 and Lead Institute

Anthony Bellotti, CIAT; A.Bellotti@cgiar.org
CIAT, A.A.6713, Cali, Colombia

Collaborating institutes and scientists

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

The primary and secondary gene pools of cassava, made up of wild *Manihot* species, are a rich source of genes for pest and disease resistance as well as important root quality traits such as delayed post harvest physiological deterioration (PPD). This project was initiated to accelerate the process of introgressing useful genes from wild relatives into cassava. This project has resulted in the development of invaluable germplasm, with resistance to cassava mosaic diseases (CMD) andassava green mites (CGM), and delayed post harvest physiological deterioration (PPD), and the transfer of these materials to partners. It has also led to development of new molecular tools to increase the cost-effectiveness of breeding cassava for stable yields and novel traits.

1. Low cost technologies
To enable NARs conduct their own MAS projects, markers linked to CMD2 was converted to Sequence Characterised Amplified Region (SCAR) marker SCAR markers that are easily assayed. To simplify assays, RME1, the polymorphic RAPD fragment for CMD resistance, was converted to SCAR marker, a low cost and easy to use marker.

Testing of the FTA paper for low cost MAS of CMD resistance in breeding populations was successfully achieved for both RAPD and PCR markers. This result indicates that FTA paper leaf squashes could replace cumbersome DNA isolation step. However, a two-step DNA isolation method was found to have lower costs than FTA paper technique.

2: Genetic mapping for traits
Multi-locational evaluation for delayed PPD of CW 429-1 and 8 other elite genotypes indicate that results of mean PPD values at 10 days after harvest (DAH), ranged from 0% in CW429-1 to 58% in CM 523-7. The same trend was observed 14 DAH. Three linkage groups with eight putative markers for delayed PPD were identified accounting for 6.2 to 12.8% of phenotypic variance.

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.

New collections of wild *Manihot* species for identification of useful genes
New collections of *Manihot* species in the centre of diversity of the genus, Brazil, were also conducted by EMBRAPA, a collaborating institution. Existing and new collections of wild *Manihot* species were also evaluated for pest and disease resistance to increase yield stability of cassava across many regions. Some of the accessions are listed in Table 1.

Development of advanced lines and transfer of germplasm
Over 2000 genotypes selected with CMD and/or CGM markers were distributed by CIAT to NARS in this project. Useful traits (e.g. delayed PPD) from wild Manihot species were also introgressed into the distributed germplasm via advanced backcross. In 2009, delayed PPD second backcross population, selected with markers for CMD resistance were shipped to NARS partners.

Farmer participatory plant breeding has resulted in the testing of varieties developed in this project for possible release in Nigeria and Ghana. Four advanced lines are under test by famers in Nigeria and a similar number of varieties have be identified and selected by farmers in Ghana. Efforts were also initiated to develop improved gene pools with introduced germplasm through crosses with local farmer preferred varieties in Africa. The F1 genotypes integrated into NARS breeding programme.

Table 1. The number of developed accessions for each Manihot species

Species	Accessions
anomala	152
caerulescens	35
cecropiaefolia	10
compostifolia	1
dimantinensis	1
dichotoma	104
flabellifolia	215
glaziovii	24
irwinnii	11
jacobinensis	1
maracasensis	1
peruviana	260
tomentosa	31
tripartita	1
violacea	2
manicoba	67
pornuncia	3
sete anos	1

Capacity building
The project has contributed to 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. The technical know-how on MAS for breeding CMD resistance was also transferred to National Programmes in Africa with the setting of simple molecular breeding laboratories in Nigeria, Ghana, and Uganda. 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).

Ex impact analysis
Ex impact analysis of MAS bred varieties estimated the change in economic surplus generated by introducing cassava varieties with tolerance to cassava mosaic disease, green mites, whiteflies, and delayed post-harvest deterioration. Results indicate that MAS was significantly better than conventional breeding. The difference was mostly due to the faster timing of release for the varieties developed with MAS and the higher probability of success.

24. 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 aluminium tolerance gene in sorghum (ALTSORGHUM)

August 2007–December 2009; no-cost extension to December 2010

Principal Investigator and Lead Institute
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 institutes and scientists
- Robert W. Holley Center for Agriculture and Health, USDA–ARS: Leon Kochian; Owen Hoekenga; Jiping Liu
- IGD–Cornell: 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 Alt_{SB}, a major gene in sorghum conferring tolerance to aluminium (Al) toxicity, which seriously constrains agriculture in several regions of acid soils throughout the world. Association analysis is being used to identify Alt_{SB} haplotypes that confer superior tolerance to Al toxicity in sorghum, if possible to dissect the molecular nature of Alt_{SB}, develop haplotype specific markers, and by molecular breeding develop near-isogenic lines (NILs) carrying elite haplotypes. Three different association panels are being used: (1) 45 inbred lines current used in breeding programmes in the US and Brazil, (2) 210 accessions assembled to maximize genetic diversity based on racial classification, latitude of origin, and photoperiod response (this is part of the SP1 sorghum reference set, LR-CIRAD) and 300 landraces that were converted to photoperiod insensitivity, early maturity, and short stature phenotypes (i.e. adapted for breeding programmes in temperate regions, LR-Cornell). We are genetically combining these three panels using Illumina GoldenGate technology by assaying 384 SNPs distributed throughout the sorghum genome (see 2008 annual report). This will allow for conducting association analysis with correction for population structure and relatedness in the combined panel. We are also developing at least two 1536 SNP chips that will be designed based on genome coverture and will contain candidate genes for Al tolerance. The combined association panel will be genotyped with these SNP loci in search for new Al tolerance genes in sorghum.

2. Partial Results

2.1 Germplasm exchange: Linking to the commissioned project ALTFIELD led by Bob Schaffert, one of our goals is to conduct allele mining in sorghum breeding panels belonging to INRAN (Soumana Souley) in Niger and ICRISAT – Bamako (Eva Weltzien and Fred Rattunde) in Mali. As noted in the 2008 annual report, the germplasm exchange processes concerning these panels are delayed. The INRAN panel has finally entered quarantine procedures in Brazil and is expected to be made available in approximately 2 months time. After signing the required MTA and SMTA, the Mali collection was sent to Brazil but returned to Mali due to bureaucratic problems with the Brazilian system. The material has been sent again to Brazil about a month ago.

2.2 Phenotyping for aluminium tolerance: In the 2008 annual report we reported on the completion of the phenotyping activities in hydroponics for the Embrapa and LR-CIRAD panels. Phenotyping of the LR-Cornell is now half-way from completion and the combined panel will be genotyped with at 2x1536 SNP chips to look for novel Al tolerance genes in sorghum.

2.3 Association analysis of Alt_{SB}: We have tested several different population structure (Q) and relatedness (K) corrections for association analysis in the Embrapa/LR-CIRAD panel. Upon applying the linear mixed model with Q(6) + K, 5% of the datapoints showed p-values ≤ 0.05 with different phenotypic traits. This optimised model with proper type-I error control enabled us to update our estimates for SNP effects.

2.4 Development of gene specific markers: we have successfully developed agarose-based assays for 5 loci associated with Al tolerance in the LR-CIRAD/Embrapa panels in additional to a SNP specific to one of the strongest alleles we found so far (Table 1).

2.5 Development of NILs carrying elite Alt_{SB} haplotypes: In the 2008 report we included a table indicating the current status of this activity, which is progressing in very good pace. Some of these exotic alleles coming from the association panel are apparently stronger than those of our previous Al tolerant standards. However, Al tolerance as conferred by them behaves much more like a recessive trait, with the ratio of dominance (d) to additive (a) effects falling between 0 (additive gene action) and -1 (recessive gene action) for the majority. This means that for hybrid production, both the male (R line) and the female (A line) parents will need to harbor the tolerant alleles.

3. **Tangible outputs delivered:** In terms of the three key products listed in our proposal: 2 NILs are ready (i.e. fixed and confirmed BC3F2s), 7 BC3F2 are being genotyped, 14 are in BC3F1 and 3 in earlier generations. We have successfully developed 6 tag markers that are being used for MAS for NIL development. Those will be used to introgress Alt_{SB} alleles into 2 selected accessions from INRAN as soon as we receive the germplasm, which will also be used for allele mining.

Table 1: Summary of the association results and tag markers (Al tolerant alleles are in bold)

Polymorphism	Position (bp)	Region (see 2008 report)		p-value	r^2	Gene-specific markers		
						Marker	Primers	Fragments
SNP M (A/C)	6083	A2	second intron	2.10E-009	8.3	four primer ARMS (co-dominant)	HB23+HB24 +HB30+HB37	gene specific 700 pb **A allele**: 240 pb C allele: 484 pb
SNP R (A/G)	5985	A2	second intron	0.0042	-	three primer ARMS (two allele specific markers)	HB23+HB24+HB27 HB23+HB24+HB26	gene specific: 700 pb **A specific**: 361 pb G specific: 361 pb
SNP S (C/G)	6094	A2	second intron	2.53E-007	6.3	three primer ARMS (two allele specific markers)	HB23+HB24+HB33 HB23+HB24+HB32	gene specific: 700 pb **C allele**: 254 pb G allele: 254 pb
Indel 19 pb	12487	A3	3` to Alt_{SB}	1.04E-008	7.7	two primers (dominant)	HB1+HB2	Insertion specific: 685 pb (Al sensitive)
SNP W (A/T)	5519	-	first exon	-	-	three primer ARMS (two allele specific markers)	HB48+HB49+HB51 HB48+HB49+HB50	gene specific: 709 pb **T specific**: 230 pb A specific: 230 pb
MITE indel (1912bp)	1850	MITE	5` to Alt_{SB}	1.04E-004	4.0	two primers (co-dominant)	MJ23+MJ26	**1912**
MITE indel (1184 allele)	1850	MITE	5` to Alt_{SB}	9.92E-004	2.9	two primers (co-dominant)	MJ23+MJ26	**1184**

25. 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; no-cost extension to July 2010

Principal Investigator and Lead Institute
Arvind Kumar; IRRI; akumar@cgiar.org
DAPO BOX 7777, Metro Manila, Philippines; (63) 2-580-5600; Fax: (63) 2-580-5699

Collaborating institutes and scientists

- IRRI: R Serraj; David J Mackill
- IGKV: SB Verulkar
- CRRI: ON Singh; P Swain
- CRURRS: PK Sinha; NP Mandal
- NDUAT: JL Dwivedi
- UAS–Bangalore, India: S Hittalmani; Venkatesh Gandhi
- TNAU: R Chandrababu; A Senthil; S Robin
- BAU: BN Singh; RL Mahato
- JLNKV: P Perraju
- Barwale Foundation, India: HE Shashidhar; Abhinav Jain
- YAAS: D Tao
- UoAl: Dean Spaner

1.1 Fine-mapping loci affecting yield under drought stress

A large effect QTL (qtl12.1) for grain yield under upland reproductive stage drought was identified in Vandana/Way Rarem population around a 10 cM interval between RM 28048-RM 511 on chromosome 12 (Bernier et al. 2007). Additional markers were added in the region on the RIL and $BC_2F_{2:3}$ population. Results indicated presence of at least two independent regions (RM1261-28166, RM28048-

RM511) affecting grain yield under drought. RM1261-RM28166 region showed additive effect of 208 kg which is around 50% of the trial mean (417 kg/ha) under stress. The BC_2F_3 population is being screened in India in 2009 WS and will be screened at IRRI in 2010 DS to validate the results. In BC_3F_2 derived IR55419-04/Way Rarem population, additional markers have been added in the region to fine map the RM17403-RM17503 region on chromosome 4 earlier identified (qtl4.1) to have significant affect on grain yield under upland reproductive stage drought. The population screened under drought in 2009 DS season at IRRI could not be exposed to severe drought stress due to frequent rains and will be phenotyped again in 2010 DS to fine map this region.

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. Additional markers have been added in the region to precisely locate the QTL and data is being analysed. In populations derived from crossing the drought tolerant donors N22 with susceptible cultivars Swarna, MTU1010 and IR 64, bulk segregant analysis (BSA) of high and low tails genotypes for grain yield under drought has identified region around RM 431on chromosome 1 to show significant differences in all the three populations. Further genotyping to precisely identify the region is under progress.

1.3 Physiological effects of drought yield QTLs

Experiments have indicated that qtl12.1 enhanced the grain yield under reproductive stage drought through increased 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 (Bernier et al. 2009). 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.

QTLs effect validation in India and China

Effect of identified regions governing grain yield under upland aerobic situation in drought tolerant and aerobic adapted cultivar "Apo" was validated in Apo/Minguihi $BC_2F_{3:4}$ population for improving the aerobic adaptation of Minguihi. Under surface irrigation system, RM302-RM212 region on chromosome 1 showed significant affect on grain yield with additive affect of 870 kg ha^{-1} at the trial mean level of 2228 kg ha^{-1}. At IRRI, the relative effect of the qtl12.1 on grain yield increased 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. Effect of qtl12.1 has been validated in diverse environments in India (Bernier et al. 2009a) and effect of qtl4.1 was validated in China.

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 J, Serraj R, Kumar A, Venuprasad R, Impa S, Gowda V, Oane R, Spaner D and Atlin G (2009). The large-effect drought-resistance QTL qtl12.1 increases water uptake in upland rice. *Field Crops Res.* 110:139-146.

Bernier J, Kumar A, Spaner D, Verulkar S , Mandal NP, Sinha PK, Perraju P, Dongre, PR, Mahato RN and Atlin G (2009a). Characterisation of the effect of rice drought tolerance qtl12.1 over a range of environments in the Philippines and eastern India. *Euphytica*: 166:207-217.

26. G3008.06: Targeting drought-avoidance root traits to enhance rice productivity under water-limited environments

October 2008–September 2011

Principal Investigator and Lead Institute

Rachid Serraj, IRRI; r.serraj@cgiar.org

Crop & Environnemental Sciences Division, IRRI, DAPO Box 7777, Metro Manila, Philippines. Tel: +63(2)580-5600, ext. 2730

Collaborating institutes and scientists

- IRRI: K McNally (co-PI); A Kumar; N Kobayashi
- WARDA–Nigeria: S Mande (co-PI); T Hiroshi
- TNAU: R Chandra Babu (co-PI)
- Barwale Foundation, India: HE Shashidhar (co-PI)
- UoAb: A Price (co-PI)
- CStuU: LJ Wade (co-PI)
- Collaborators: A Yamauchi (NaU); RE Sharp and HT Nguyen (UoMI); A Kijoji (SUoAg); Drought Breeding Network, India).

Research activities and progresses

Activity1.1. Development of high-throughput, low-cost root phenotyping methodologies.

Field drought phenotyping experiments were carried out at IRRI in dry-season 2009, using the *OryzaSNP* set of genotypes, NILs of IR64 and from the Adaysel-IR64 population with variable performance under drought. The relationships between root architecture and the dynamics of plant water uptake were analysed during progressive soil drying. Interactions between root growth parameters and water uptake were used to analyse the genotypic differences in dehydration avoidance and yield under stress. The dynamics of soil water profiles, canopy development and plant water status using non-destructive measurements (NDVI, IR thermal imaging) were measured as well as destructive sampling of roots and shoots. Large differences in soil water uptake were observed among genotypes, contrasting in their sensitivity to drought. Results also suggest that variation for other root characteristics such as root hydraulic conductance, in addition to rooting depth may be important for root water uptake by rice plants growing in drying soils (Henry et al., 2009).

A root phenotyping platform (900 lysimeters) was established in the greenhouse at IRRI, for nondestructive measurement of plant water use under drought. Root growth at depth in the lysimeters, varied substantially among genotypes under drought, and the growth patterns were correlated with plant water uptake at depth. These initial results will be used to demonstrate the importance of genetic variation for root growth at depth for conferring drought avoidance through improved uptake of soil water under water deficit.

Activity 1.2. Characterisation of root soil environment. Three sites used in multilocation drought screening in India (Raipur, Hazaribagh and Ranchi) were characterised during dry-season 2009 in collaboration with the NARES partners in these sites. Analysis focused on electric conductivity using EM-38 soil mapping. Additional characteristics include soil texture, pH, soil moisture, etc. This will allow analysis of G x E interactions of drought responses and role of root traits for in different target rice environments.

Activity 1.3. Root phenotyping network. A comparative root phenotyping study involving all project partners was initiated, for testing all major root-screening methods that have been previously used in rice root phenotyping. Seed of the reference set of genotypes that includes the OryzaSNP was dispatched to all project partners and collaborators, to be used in a multisystem and multilocation experiment. This will generate G x E data sets for the analysis of root trait expression and comparison of the validity of the various root phenotyping systems.

Activity 2.1. Root growth kinematics under drought. Preliminary root growth kinematic experiments were carried out at IRRI, under controlled soil hydraulic and physical conditions, using contrasting rice genotypes and precisely imposed water deficit. Initial results showed large variation among rice genotypes in the response of root elongation to soil moisture, as measured by the fraction of transpirable soil water.

Activity 2.2. Root architecture and phenotypic plasticity. Studies were initiated (IRRI and Nagoya Univ.) for the characterisation of root morphological traits and the relative contributions of nodal and lateral roots, root volume, root density, and hydraulic properties to soil water uptake under different water regimes.

Activity 3.1. Association analysis. The accessions from the *aus* isozyme in the collection of *Oryza sativa* from the GCP composite collection was evaluated under field conditions (link with project *G4008.05*) and seed was multiplied using SSD for further phenotypic analysis of root traits (McNally et al.).

Activity 4.1. Analysis of root traits in advanced breeding lines and NILs. A large set of advanced breeding lines from the India Drought Breeding Network has been phenotyped for root water uptake and field crop performance during Dry-season 2009 in collaboration with ICRISAT and DRR-Hyderabad.

Tangible outputs delivered–As per initial project timeline

- A standardised high-throughput drought phenotyping protocol (field rainout shelters)
- A reproducible and low-cost root water uptake phenotyping system (900 lysimeters)
- Initial rice root growth kinematics analysed under drought in relation to plant performance

References

Serraj R, Gowda V, Henry A, Mabesa RC, Kano M, Kato Y, Kobayashi N, Yamauchi A (2009). Drought phenotyping and dehydration avoidance root traits for rice improvement under water-limited conditions. *Japanese Journal of Crop Science* 78 (1) – Special issue, 400-401.

Bernier J, Serraj R, Kumar A, Venuprasad R, Impa S, Gowda V, Owane R, Spaner D, Atlin G (2009). Increased water uptake explains the effect of qtl12.1 a large-effect drought-resistance QTL in upland rice. *Field Crop Res*. 110: 139-146.

Henry A, Torres R, Serraj R (2009). Dissecting the role of root architecture in drought avoidance mechanisms in rice. In: *20th Conference of the Federation of Crop Science Societies of the Philippines* (FCSSP), May18-22, Dumaguete, Philippines.

27. G3008.07: Basal root architecture and drought tolerance in common bean

October 2008–September 2011

Principal Investigator and Lead Institute

Jonathan Lynch, Dept. Horticulture, PSU
Pennsylvania State University, University Park, PA 16803, USA;
JPL4@psu.edu, tel 814-863-6139, fax 814-863-6139

Collaborating institutes and scientists

- PSU: Jonathan Lynch; Kathleen Brown
- CIAT: Steven Beebe; Matthew Blair; Idupulapati Rao
- IIAM: Magalhaes Miguel; Celestina Jochua
- SABRN–Malawi: Rowland Chirwa

1. Research activities at CIAT: Screening of Recombinant Inbred Lines (RILs) for Basal Root Whorl Number (BRWN)

As part of objective 1 the reference collection (a set of diverse lines with multiple genepool sources) was planted at IIAM in Sussundenga, Mozambique but suffered from excessive drought stress. Therefore, additional seed is being prepared for this high priority evaluation. This collection will be evaluated in South Africa. Final seed multiplication for the reference collection is underway for objective 1.

As part of the same objective, progeny of 5 inter-specific scarlet runner bean x common bean crosses were evaluated for BRWN. Common bean parents in two crosses included a large seeded Andean common bean, ICA Quimbaya, and another three crosses involved SER 16, a small red seeded drought resistant Mesoamerican line. Average BRWN in the first evaluation of SER 16 was 1.75, while its progenies ranged from 1.16 to 3.30 whorls. Quimbaya averaged 2.90 whorls, and its progeny presented from 2.00 to 4.40 whorls. These could be useful genotypes to test the impact of whorl number on agronomic performance.

As part of objective 2, RILs for the populations DOR364 x G19833 and G2333 x G19839 were multiplied and sent to PSU and IIAM for evaluation. Segregating populations from G19833 have been generated and will be used for QTL analysis. These crosses combine commercial types for Southern Africa, such as sugar lines or large red mottled beans, with genes from G19833 in simple and double crosses that are now in the F2:3 or F1:2 generations. PVA773, a Mozambican release is included in many of these crosses and therefore they will be of high priority for the project. Other crosses combining G19833 with BCMV-resistant parents were shipped to PSU and are awaiting analysis for BRGA and BRWN.

2. Research activities at Penn State: Utility of Basal Root Whorl Number (BRWN) and Basal Root Growth Angle (BRGA) plasticity for drought tolerance

Field experiments have been planted to evaluate utility of BRWN and BRGA plasticity for P acquisition and drought tolerance using RILs from the common bean crosses DOR364 x G19833, G2333 x G19839, and TLP19 x B98311. Ongoing greenhouse experiments evaluate the utility of these traits and pouches are being used to phenotype these populations.

Pouch phenotyping of DOR364 x G19833 has been completed for BRWN and is ongoing for BRGA plasticity. Pouch phenotyping for BRGA and plasticity of BRGA has been completed for the TLP19 x B98311 cross and will begin in July 2009 for the SER16 x *P. coccineus* population. This rapid, high throughput phenotyping of multiple RIL populations will supply our partners in CIAT with enough material to precisely identify QTLs.

A greenhouse pot study evaluated the utility of BRGA plasticity challenged selected genotypes of the TLP19 x B98311 population with phosphorus, water and combined phosphorus and water stress. This pot study will be repeated and results are expected to confirm the outcome of summer 2009 field studies.

Finally, work at a field research site in South Africa has been arranged, which will allow us to perform field experiments in both PA and South Africa every year. The sandy, nutrient poor soil in an area only suitable for irrigated agriculture is expected to induce stronger and more consistent water stress than our current site in Pennsylvania. Similar results in contrasting soils will validate our conclusions.

3. Research activities at IIAM: Field evaluations of the utility of BRWN for P efficiency and drought tolerance

Genotypes contrasting for BRWN have been evaluated in the field for P acquisition efficiency and drought tolerance. P efficiency studies were conducted at Sussundenga Research Station, while the drought studies will be conducted in Chokwe to examine the interactive effects of these two abiotic stresses. The field experiment at Sussundenga evaluated the utility of BRWN for P acquisition

efficiency. It was found that RILs with 3 whorls had a higher shoot dry weight than those RILs with 2 whorls.

Field evaluations of basal root traits for drought (Chokwe) and P efficiency (Sussundenga) of the materials supplied by CIAT will be carried out in the next planting season starting January 2010.

Two sets of RIL populations of common bean (DOR364 x G19833 and G2333 x G19839) were screened for BRWN in May 2009 at Penn State by a Mozambican student.

Evaluating the utility of BRWN and BRGA plasticity for water and nutrient acquisition response and identification of QTLs involved, will enable rapid and precise breeding of common bean varieties that maintain productivity in water and nutrient limited environments. Increased productivity of the common bean in water and phosphorus limited environments will contribute to better human nutrition, slow soil erosion and increase nitrogen available to other crops in the cropping system. As 60% of small-holder and subsistence bean growers in Latin America and 75% Eastern and Southern Africa grow beans in phosphorus limited soils, and 70% of both these regions suffer from drought, this research will have broad impacts.

28. G3008.08: Breeder-friendly high-throughput phenotyping tools to select for adaptive traits in drought environments

January 2009–December 2011

Principal Investigator and Co-PI
Francis C Ogbonnaya, ICARDA
P.O. Box 5466, Aleppo, Syria, Ph: +963-21 2213433; FAX: +963-21 2213490; E-mail: F.Ogbonnaya@cgiar.org
M Fernanda Dreccer, CSIRO Plant Industry
CSIRO Plant Industry Cooper Laboratory, PO Box 863, University of Queensland, Warrego Highway, Gatton, QLD 4343 AUSTRALIA, Ph: 61 7 5466 3010, FAX: 61 7 5462 3797, Mobile: 0408885137; E-mail: fernanda.dreccer@csiro.au

Collaborating institutes and scientists

- ICARDA: Osman Abdalla; Mohammed Karrou
- CSIRO Plant Industry, Australia: Tony Condon
- CIMMYT: Matthew Reynolds; David Bonnett
- INRA–CRRA, Centre Aridoculture, Settat, Morocco: Hassan Ouabbou
- ICARDA–INRA Cooperative Research Programme, Rabat Institute, Morocco: Sripada M Udupa
- INRA–CRRA/EIAR/Kulumsa Agricultural Research Centre, Ethiopia: Solomon Gelacha

1. Research activities and progresses at ICARDA, CSIRO and other collaborators

Two hundred and fifty wheat lines including ICARDA purported drought elite germplasm and 47 possessing tillering, early vigour, transpiration efficiency, low and high stem water soluble carbohydrate and stay green were assembled and seed multiplied at Tel Hadya and Breda, between December, 2008 to July 2009 in Syria. Notes were taken on key agronomic traits and are being genotyped with markers linked to *Ppd*, *Vrn*, *Rht* genes. Sufficient seeds from the just concluded seed multiplication would be used for establishments of trials in Ethiopia, Mexico, Morocco and Syria in 2010 for screen development, trait comparison and characterisation.

2. Research activities at CSIRO, Australia

Sixty-one ICARDA lines were out of quarantine in late May with 5-20g of seed per line made available and used in establishing ongoing trials from May 2009 to December 2009. Trials have been sown to cover objectives 1 (screen development), 2 (trait comparison) and 3 (ICARDA lines characterisation) as shown in Table below. Several trials for trait comparison are on the way that will be kept rainfed sown to a range of initial soil moistures. There is also an irrigated trial and two late sowing dates under irrigation for comparison under warmer conditions/heat stress.

Seed multiplication has been sown in Gatton and Leeton, Australia for lines with enough seed quantities, lines with 5g or less of seed were only sown at Gatton only.

Trial type	Location	Sowing date and water level	Design
Trait comparison	Gatton, Queensland, Australia (27.55 S lat, 152.33 E long)	Sown 4-June-09 Irrigated	Row-Column, 2 reps
Trait comparison	Gatton, Queensland, Australia (27.55 S lat, 152.33 E long)	Sown 6-June-09 Rainfed onto a saturated profile	Row-Column, 3 reps
Trait comparison	Gatton, Queensland, Australia (27.55 S lat, 152.33 E long)	To be sown late July or mid August Irrigated	Row-Column, 2 reps
Trait comparison	Temora, New South Wales, Australia (34.41 S lat; 147.52 E long)	Sown May 28, 2009 Rainfed onto a dry subsoil	Row-column, 2 reps
Trait comparison	Leeton, New South Wales, Australia (34.60 S lat; 146.40 E long)	Sown June 23, 2009 Rainfed onto a dry subsoil to be wetted up by irrigation August (if needed) then maintained Rainfed	Row-column, 2 reps
Trait comparison	Leeton, New South Wales, Australia (34.60 S lat; 146.40 E long)	To be sown July 7, 2009 Rainfed onto a dry subsoil to be wetted up by irrigation August then maintained Irrigated	Row-column, 2 reps
ICARDA line multiplication	Gatton, Queensland, Australia (27.55 S lat, 152.33 E long)	Sown 5-June-09	Paired rows
ICARDA line multiplication	Leeton, New South Wales, Australia (34.60 S lat; 146.40 E long)	Sown 16-June-09	Paired rows

3. Research activities at INRA-CRRA, Centre Aridoculture, Settat, Morocco
Because of limited quantities of seeds, 5 gms each of the 250 lines were sown as observation nursery at Sidi El Aydi, Morocco between December 2008 to July 2009 with data collected on key agronomic and phenology traits. Data are being collated from the preliminary trials from the observation nursery in Morocco and seed multiplication in Syria. These would be analysed and subset germplasm chosen for establishments of trials in Ethiopia, Mexico, Morocco and Syria in 2010 for screen development, trait comparison and characterisation.

4. Research activities at, **Ethiopia Institute for Agricultural Research (EIAR**), **Kulumsa Agricultural Research Centre (KAR).**
An ongoing trial from July 2009 to December 2009 was sown at Dhera and Melkassa, Ethiopia with 200 lines. Trials have been sown to cover objectives 2 (trait comparison) and 3 (ICARDA lines characterisation). Trials were sown as alpha-lattice with three replications.

Data collected from this trial in addition to the preliminary data from Syria would be analysed and subset germplasm chosen for establishments of trials in Australia, Ethiopia, Mexico, Morocco and Syria in 2010 for thorough screen development and validation, trait comparison and characterisation.

Subprogramme 4: Bioinformatics and crop information systems

29. 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 and Lead Institute
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 institutes and scientists
- CIMMYT: Jean-Marcel Ribaut; Matthew Reynolds; José Crossa
- CSIRO: Scott Chapman
- UdAC: 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 QTLxE 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 GxE (QTLxE), 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).

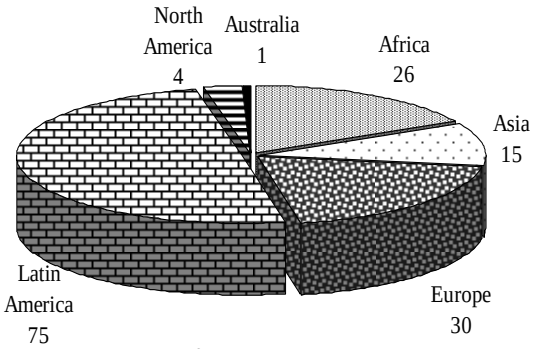


Fig 1. Number of course participants per region

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 QTLxE 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; 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* 117: 1077-1091

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). Springler, 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.

30. G3008.09: Breeding drought tolerance for rainfed lowland rice in the Mekong region

November 2008–October 2011

Principal investigator

Boonrat Jongdee, BRRD; boonrat@brrd.in.th
Ubon Ratchathani Rice Research Center, Ubon Ratchathani 34000, Thailand; Tel: 0066-45-344-103; Fax: 0066-45-344-090

Collaborating institutes and scientists

- BRRD: Grienggrai Pantuwan
- UoQ: Shu Fukai
- NAFRI: Phoumi Inthapanya
- CARDI: Ouk Makara

1. Project summary

Drought is considered to be a major production constraint for rainfed lowland rice in the Mekong region. We propose to develop new effective breeding strategies for selection of adapted genotypes using our populations, which we had developed in Thailand, Laos and Cambodia.

We will use 3 approaches; conventional, , modeling and farmer participatory. In the conventional approach, there will be phenotyping under controlled field screening for tolerance to different types of drought and also for general adaptation. In addition to multi-location trials, evaluation of selected genotypes will use farmer participatory variety selection with the Mother and Baby design. In order to identify target areas of different drought types particularly early and late season drought, maps of drought development patterns will be produced using simulation modeling and GIS techniques.

Traits involved in adaptation to different growing conditions and new drought traits will be investigated. In drought-prone rainfed lowland conditions, rice plants often experience both anaerobic and aerobic conditions. Thus, our hypothesis is that acceptable genotypes need to be well adapted to both growing conditions, and also possess drought tolerant traits.

2. Activities by BRRD, Thailand

A total of 300 lines included 100 DH lines derived from CT9993-5-10-1-M x IR62266-42-6-2 and Surin1 backcross are being screened under well-water conditions, aerobic and anaerobic and drought stress at flowering stage and raised bed system in dry season 2009. There is no result yet.

3. Activity by NAFRI (Laos), CARDI (Cambodia) and BRRD (Thailand)

Seed was multiplied and generation advanced during the dry season of 2008/2009 for major experiments in wet season in 2009 in all 3 countries.

4. Activity at The University of Queensland

Computer hardware and software has now become operational and the mapping activity has commenced. Meteorological data (maximum and minimum temperature, rainfall, sunshine, relative humidity and evaporation) has been collated for 26 locations in Thailand and is currently being manipulated into the required form for input into the rice model. A model has been run for Ubon Ratchathani to generate potential evaotranspiration data which will then be used as input to run the rice model.

Meteorological and soils data from a number of sites in Cambodia, and for Vientiane and Champasak provinces in Laos are currently being collated. Preliminary work has begun utilising the mapping software which in the first instance will be utilised to generate rainfall maps based on the data from 26 locations in Thailand. It is anticipated that more rainfall data will be required for each province.

COMMISSIONED PROJECTS
Subprogramme 1: Crop genetic diversity

31. G4005.01.03: Completing genotyping of composite germplasm set of sorghum

January 2005–December 2006, NCE through September 2008

Principal Investigator and Lead Institute
C Tom Hash, ICRISAT; c.hash@cgiar.org
Patancheru, Hyderabad 502 324, Andhra Pradesh, India; Phone: +91 40-30713322; Fax: +91-40 30713074

Collaborating institutes 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

1. Introduction/Background
Sorghum is the 5th most important cereal crop globally and its germplasm exhibits substantial phenotypic and genotypic diversity. 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 propose 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 were used in 2008 to validate the population structure of the proposed sorghum reference germplasm set. Several of the accessions proposed for inclusion in the reference set are not readily available for exchange, so some modification of its composition is required.

2. Ongoing/completed activities
2.1 Selection of composite collection entries and markers, and SSR data generation
A total of 3372 sorghum accessions, including the CIRAD minicore collection and the ICRISAT core collection, a number of breeding lines and elite varieties of interest to sorghum improvement programmes in Latin America, Africa and South Asia, as well as 250 sorghum accessions of Chinese origin contributed by CAAS, were genotyped with 48 SSR markers detecting loci distributed across all 10 sorghum linkage groups. Reasonably complete data sets were obtained for 3365 accessions and 41 SSRs. A manuscript describing the SSRs used for this analysis has been prepared.

2.2 Analysis of population structure and selection of a reference germplasm set
Genetic diversity analysis based on allelic variation across the 41 SSR markers indicates that population structure in sorghum is explained in large part by botanical race (five basic races and ten hybrid races) within geographic origin, with the West African margaritiferum sub-group within the guinea race forming a distinct cluster that appears to be more closely related to wild and weedy sorghums than to most other cultivated sorghums, suggesting that this group represents an independent domestication event. Race kafir (largely from Southern Africa) was distinct. Accessions of the durra, caudatum and guinea races each formed distinct geographic subgroups. Race bicolor showed limited structure, 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), but was otherwise scattered across the dendogram. Based on this, a sorghum reference set of 384 accession was proposed, which includes a substantial portion of the CIRAD minicore collection for which substantial phenotyping information already exists. Population structure of this proposed reference set was confirmed using a set of 40 additional EST-SSRs.

These results will benefit sorghum research programmes globally, and ultimately sorghum producers and consumers around the world. The implications of this study are that the proposed sorghum reference germplasm set is sufficiently diverse to 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.

A draft manuscript on the SSR-based diversity analysis of the GCP sorghum composite germplasm collection and development of the proposed reference set has been prepared and is under revision. Similarly, a second manuscript on the EST-SSR-based validation of population structure of the proposed sorghum reference set has been prepared.

Unfortunately, at present there is no seed available for exchange for 4 of these accessions originally contributed by CAAS and 30 of the accessions originally contributed by CIRAD, so modification of the reference set is required—replacing these accessions with the most similar accessions for which seed is available for exchange. Analysis required for this revision of the sorghum reference set has not been completed.

3. Future activities
Complete data analysis required for revision of the proposed sorghum reference set and for publication of the analysis of population structure of the sorghum composite germplasm set. Correct submitted data set. Finalise and submit final project report.

4. Expected outputs
Sorghum composite germplasm collection representing diversity of landrace and wild/weedy germplasm available globally for sorghum improvement. A set of well-characterised sorghum genomic and genic SSR markers, detecting single-copy loci that are well-distributed across all 10 sorghum chromosome pairs, that are suitable for use in diversity assessment. Improved information on the population structure of global sorghum genetic resources that will facilitate exploitation of these for sorghum improvement. A reference germplasm set comprised of accessions representative of global diversity of cultivated and wild/weedy, with seed and DNA samples available for exchange.

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

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

Principal Investigator
Ken L McNally, TT Chang Genetic Resources Centre, IRRI; k.mcnally@cgiar.org
IRRI, DAPO Box 7777, Metro Manila, Philippines

Collaborating institutes and scientists

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

Research activities for EcoTILLING in *Oryza* and *Sorghum* germplasm
EcoTILLING, a tool that detects polymorphism in the form of SNPs or indels in natural populations, was employed to survey variation in 3 panels of *Oryza* germplasm: a mini-core collection of 1536 *O. sativa* accessions, 190 accessions of *O. glaberrima*, and 95 accessions from six wild AA genome *Oryza* species. The *Oryza* genomic DNAs were pooled against a japonica (Nipponbare) contrast (all panels) and either an indica (IR 64) contrast (*O. sativa* wild samples) or IRGC 96717 (CG-14) for the *O. glaberrima* samples. Amplification products designed from the Nipponbare genome worked in both cultivated and wild AA genome *Oryza* species for drought candidate genes TPP, ERF3, DREB2, ADF2a, ADF2b, MAPk2, BZIP, SUC, and 14-3-3. Agarose-based ecotilling was developed as a simplified tool for visualising digestion products generated from the endonuclease action of the CEL 1 enzyme (Raghavan et al, 2007, Mol Breeding. 19:87-101). Haplotypes were scored based on mismatch

cleavage paterns and verified by confirmatory sequencing of representative accessions. Table 1 presents some results for the *O. sativa* mini-core. Among the wild AA genome *Oryza* species, species-specific SNPs were identified, proving the utility of EcoTILLING in taxonomic authentication.

For sorghum, the same strategy was deployed: a microcore set of 48 *S. bicolor* samples was nominated to represent maximum diversity in the collection and served as reference lines while 192 samples server to test broader genetic diversity. Although amplifications were clear on agarose, profiles on LiCOR genotyper were unclear (even with PCR, digestion, purification and migration tests). Sequencing of haplotypes revealed clear differences related to the known genetic diversity.

The population structure of the *O. sativa* panel was determined by model-based analysis of SSR genotyping data, allowing stratification into variety groups that were partitioned into haplotypes. The mini-core accessions were phenotyped under upland (vegetative stress) and lowland (reproductive stage stress) for performance under drought. Large phenotypic variation was observed in all variety groups. Several candidate gene haplotypes were associated with tolerance to drought stress. The magnitude of significant differences between haplotypes was highest for the AP2 domain transcription factors, consistent with their pleiotropic effects on multiple drought traits. Only rare haplotypes were observed for TPP with the main haplotype shared across varietal groups, suggesting this locus is under intense selection.

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 work plan

Sequence analysis of representative mismatches (putative SNPs) at candidate genes in rice required longer than expected to convert them to SNP identitites. These results are now being analysed, and these results compared to haplotype mismatch patterns. Further candidate genes for reproductive-stage stress tolerance are being screened. For *Sorghum* and *Musa*, the deviation from the work plan was more critical. Clear and interpretable patterns of polymorphism were difficult to obtain in Sorghum, and were related to primer design. Since the Sorghum genome sequence is now available, this issue can be solved. This technical difficulty is particularly critical when heterozygotes are present and for polyploidy species. Quantification of intensity pattern is not possible, limiting the use of the technique. For those species, other technologies have been tested, including low scale resequencing methods (Sequencher) with low success. A new technology, High Resolution Melting Curve Analysis, seems more promising especially for polyploids.

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 pheonytpes and the utility of EcoTILLING as a biosystematics tool for the AA genome species of rice.

New publication

McNally KL, Naredo MEB, and Cairns J. (2009) SNP discovery at candidate genes for drought responsiveness in rice. In *Drought Frontiers in Rice—Crop Improvement for Increased Rainfed Production* (R. Serraj, J. Bennett, and B. Hardy, Eds.), World Scientific Publishing, Singapore. 14 pp.

33. G4005.06: Supporting emergence or reference drought tolerance phenotyping centres - drought phenotyping network

April 2005–December 2007; no-cost extension to September 2008

Principal Investigator and Lead Institute

Reinaldo L Gomide, EMBRAPA Maize and Sorghum; gomide@cnpms.embrapa.br
Rodovia MG 424 km 45, Sete Lagoas, MG 35701-970, Brazil. Phone: 55 (31) 3027-1328.

Collaborating institutes and scientists:

- EMBRAPA Units: Rice and Beans, Mid-North, Wheat, Savannah and Semi-Arid; Co-PIs: Cleber M Guimarães, Edson A Bastos, Walter Q Ribeiro Júnior and Luiz B Morgado respectively, and others.

Research activities and progresses: The project established a service network for drought tolerance phenotyping in Brazil and developed phenotypic evaluation protocols for cereals (maize, sorghum, rice, wheat) and legume crops (common bean, cowpea), and installed phenotyping site specific experimental areas (2 of excellence and 5 of reference) for drought tolerance (DT) investigation accessing pre-required specific conditions of climatic, soil physical and chemical properties, together 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. ***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. The water depths applied in the irrigations were measured in collectors or catch cans in each genotype field plot. High uniformity of the water distribution in the irrigated plot was assured by setting the Christiansen coefficient ≥ 95 %. 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 in each environment target field, 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 site, configured to register automatically the main microclimatic surface parameters locally. ***Irrigation water management*** was carried out by means of reference evapotranspiration (ETo) and crop evapotranspiration (ETc) computation, using both class A pan and modified Penman-Monteith equation methods, with the crop (kc) and pan (kp) coefficients. Irrigation management strategy and irrigation timing criteria were performed based on spread sheet (Excell) for ETo and ETc 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, with non water stressed condition. Afterwards, the water stress treatments were induced, generating different application of water depths in the plots, and consequently different SWA, at pre-defined crop growth phases, defined for each genotype. ***Soil water content***, in different soil layers, was monitored by gravimetric method and other equipments and sensors (gypsum block, Diviner, tensiometer, neutron probe). ***Remarks & Recommendation*:** Procedures and criteria of controlling and monitoring water stress in contrasting environment for DT phenotyping must be established. Local micro-climatic condition, irrigation water application, and soil and plants water status should be registered with some kind of sensors and

equipments. Irrigation management strategy and timing criteria should be performed. Identifying and characterising what % yield reduction in genotypes due to water stress to better understand the effects of plants genetic and environmental (GxE) interactions for grain yield. Only a response to water regime differentiation! Escape, Resistance or Tolerance? What water stress level intensity to use (lower, intermediate, severe, very severe) and timing of application according to pre-defined crop growth phases for each genotype. It is necessary to search for indexes based on genotypes, such as crop water stress index or on yield reduction basis and soil water depletion.

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*** 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/environment site. ***Genotypes tolerant and sensible***: a reasonable number of genotypes were phenotyped and the main contrasting genotypes (tolerant and sensible) to drought were identified and selected. All the project information and data were transferred into database (Morpho).

Future research
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.

34. G4005.07: Whole Plant Physiology Modelling (WPM)

May 2005–May 2008

Principal inverstigator and Lead Institute
Luquet Delphine, Agropolis–CIRAD; luquet@cirad.fr
Av. Agropolis Lavalette 34398 Montpellier cedex 5 (France); 00 33 4 67 61 55 71 –

Collaborating institutes and scientists

- CSIRO: Scott Chapman
- UoQ: 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 characterising 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 2nd 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). P*artners: 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.

4. References

Dingkuhn, M., Kouressy, M., Vaksmann, M., Clerget, B., Chantereau, J. (2008). Applying to sorghum photoperiodism the concept of threshold-lowering during prolonged appetence. *European Journal of Agronomy* 28, 74-89.

Kouressy M, Dingkuhn M, Vaksmann 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.*

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

January 2006–July 2008

Principal Investigator and Lead Institute

NR Sackville Hamilton, IRRI; r.hamilton@cgiar.org

PROJECT UPDATE NOT SUBMITTED

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

January 2006–December 2007; no-cost extension to December 2009

Principal Investigator and Lead Institute

Dominique This, Agropolis (Montpellier SupAgro); 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 institutes 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

1. Context

The ADOC project aims to characterize 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 the Os01g69030 orthology group was 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 undertaken.

2. Research activities and progresses by crop and gene family

Specific primers were designed and tested for each crop and each member of the target 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. Sequences with good quality scores were obtained from 4 (bean) to 18 (rice) different genes for each crop, and from 5 (DREB) to 20 (SuSy) genes across species in each gene family studied, representing a total of 10 Mbp (see Table 1). SNP identification was done, integrated into INRA GnpSNP database, and is currently the GCP repository.

Ortholabs are analysing sequence data by crop cluster and generating information on candidate gene diversity with crop partners (to characterise population structure of reference collections and map candidate genes when possible) and gene specialists (for establishing orthology relationships and functional inferences).

Table 1. Final point on July 7th 2009 of the sequencing work for the ADOC project

	ASR	VIN	ER / ERL	DREB2A /2B	SuSy	SPS	total/crop	Length	Depth	Total bp
rice	5	1	3	1	7	1	18	22 091	131 to 260	4 317 772
barley	3	3	2	1	1	1	11	8 414	143 to 274	1 833 097
sorghum	5	1	3	0	4	1	14	12 095	127 to 347	3 422 803
cassava	1	0	3	0	4	1	9	6 531	153 to 261	1 405 923
potato	3	0	1	0	2	1	7	5 744	64 to 269	1 004 424
chickpea	1	0	1	1	0	1	4	4 079	79 to 235	622 990
bean	2	0	1	2	2	0	7	6 631	60 to 241	1 182 143
total/gene	20	5	14	5	20	6	70	65 585	60 to 347	10 Mbp

3. Findings and implications

In cereals, population structure influences partially haplotype patterns. Different patterns and intensity of sequence diversity have been found within gene families. For some genes like *OsAsr3*, 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 number of non-synonymous mutations was higher than synonymous mutation sites. The haplotype diversity (HD) for candidate genes in chickpea ranged from 0.00 (DREB2A) to 0.522 (ASR) and in common bean HD ranged from 0.00 (ASR, SuSy) to 0.677 (ERECTA). The SNP diversity in DREB genes was higher in common bean than in chickpea. Among all the genes, ERECTA showed highest number of SNPs across two legumes.

In cassava, heterozygosity induced difficulties in analysing data. However sequences for 10 genes were obtained for SNP analysis. The number of non-synonymous changes in these genes was, as larger than that observed in cereals. 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. Nevertheless, for potato, 7 genes could be analysed for SNPs. Negative Tajima's D value of ASR1 of potato points towards purifying selection for this gene. The number of non-synonymous mutations was higher in potato than in all the other crops. Particularly high variation in amino acid sequence was obtained for sucrose phosphate synthase.

4. Next steps

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 will be used when available to analyse potential effect of agronomical constraints on gene evolution during domestication and breeding history. The SNP database produced by this project will serve as resource for designing markers for association mapping

with drought tolerance traits, tagging favorable candidate gene alleles, and permitting use of relevant SNPs in selection programmes.

37. 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 and Lead Institute
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 institute 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 analysed 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.

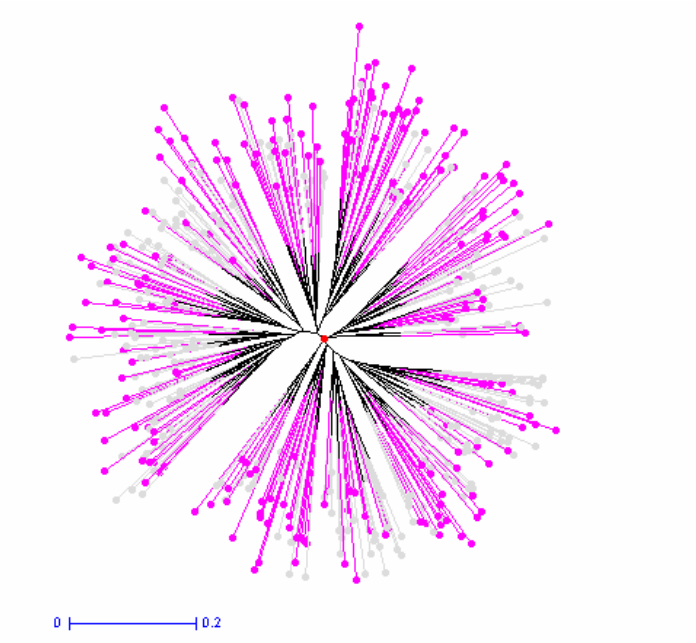


Figure. Un-weighted neighbour-joining tree based on the simple matching dissimilarity matrix of 19 SSR markers across the 452 accessions of foxtail millet composite collection (Grey colour) with proposed reference set (200 accessions) in pink

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 (Steria italica). Genet Resour *Crop Evol* 54:233-236.

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

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

Principal Investigator and Lead Institute

Hari Upadhyaya, ICRISAT; h.upadhyaya@cgiar.org
Patancheru, 502324, India; Phone: 91-40-30713333

Collaborating Institutes 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 germpasm 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 analysed 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 minimising 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.

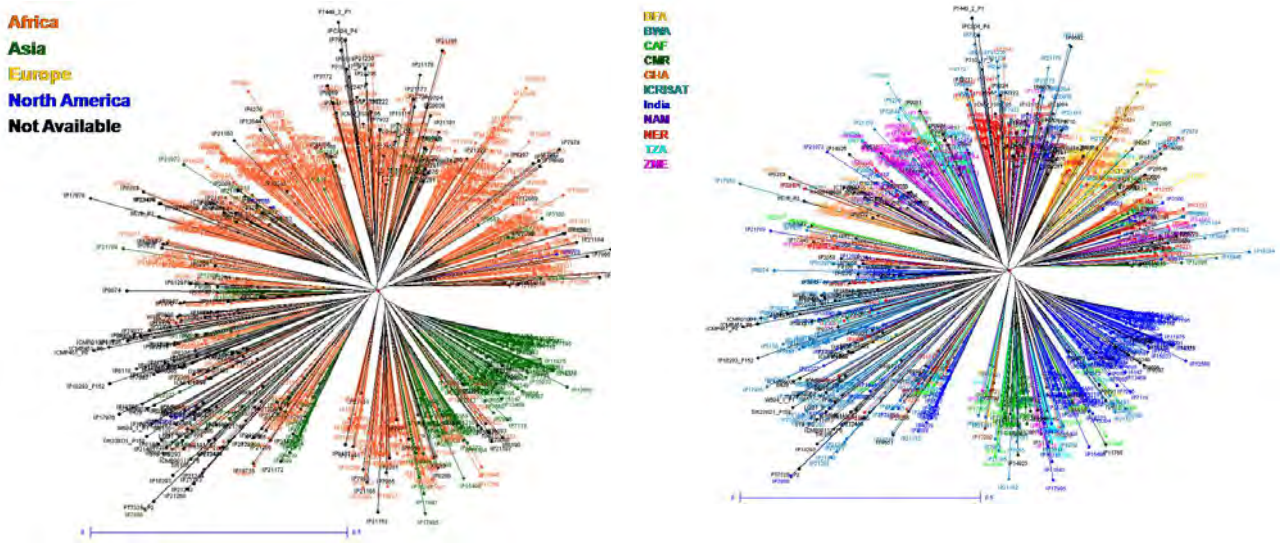


Figure. Neighbour joining tree of composite collection based on Euclidean distance

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.

39. G4007.01: Genotyping validation of the GCP reference sets

January 2008–December 2008; no-cost extension to December 2009

Principal Investigator and Lead Institute
Jean-Francois RAMI, 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

Collaborating institutes and scientists

- ICRISAT: Hari D Upadhyaya
- IRRI: Ken McNally
- CIP: Marc Ghislain
- CIMMYT: Maria Zaharieva; Susanne Dreisigacker
- CIAT: Martin Fregene; Matthew Blair
- ICARDA: Michael Baum
- Bioversity: Nicolas Roux
- Agropolis–CIRAD: Luc Baudouin
- IITA: Robert Asiedu
- ICRISAT: Tom Hash
- CIP: Wolfgang Gruneberg
- UoC–Riverside: Jeff Ehlers

A major output of the global genetic characterisation undertaken under the SP1 umbrella was to identify, from the material analysed, a reference sample, as a subset of the initial collection, meeting several criteria of representativeness and minimised structure. Altogether, these reference samples constitute a key product of the GCP and a public good, which will be widely distributed. The genotypic characterisation of this sample will constitute a core dataset that needs to be validated.

The validation consist in re-genotyping the reference set of samples with a subset of top quality and most discriminant markers (about 20) by a single non-consortium member lab (service provider). This validation process has to be adapted to the past and current management of the genetic material constituting the reference samples. It may possibly become part of it when this management needs to be modified. We want to have this new genotyping based on a stabilised genetic stock that will be further handled as such by gene bank curators and will serve for international distribution

Research activities and progress at CIRAD

Collect of DNA from the different collaborators has been pursued. DNA has been received and quality controlled for 12 of the 15 already defined reference sets among the 21 species involved in the project (Table1). Genotyping with a set of discriminant markers has been achieved for 7 species and is ongoing or scheduled for 4 additional species.

A collection of R scripts has been written for automatic datasets comparison.
The analysis conducted so far on a limited number of species pointed out the needs for such a validation. The general conclusions that can be drawn from this exercise are:

- Format heterogeneity of the datasets available on the GCP central registry. While all of them were based on the GCP template defined for SSR genotyping there were variations on the way it was used. There is a need for dataset curation at this level.
- Comparison between original and validation dataset pointed out different sources of variation:
    - Marker specific:
        - Non-linear shift between independent experiments due to difficulties in SSR binning.
    - Accession specific
        - Completely different results that can be due to any error during seed management or DNA extraction (inversion, mislabeling, etc.)

- Completely different results that can be due to pollen contamination or DNA mixing.
- Low rate of differences caused by intra accession drift that occurred during reference set single plant derived genetic stock constitution.

In the coming months, genotyping will be pursued on the remaining species and data analysis will be finalised on all the datasets following the protocol that has been setup during the first analyses. When complete and curated, validated datasets will be posted on the central registry following a uniform GCP template based format.

Table 1: Status of validation genotyping for the 21 species involved in the project

	Reference set			DNA		Genotyping	
Species	Size	Genetic Stock established	Progress	Received	DNA Quality control	Nb of markers	Status
Barley	300	Y	-	Y	Passed	15	Complete
Coconut	359	Y	-	Y	Passed	20	Complete
Finger Millet	300	Y	-	Y	Passed	20	Complete
Groundnut	300	Y	-	Y	Passed	20	Complete
Maize	234	Y	-	Y	Passed	20	Complete
Pigeon Pea	300	Y	-	Y	Passed	20	Complete
Sorghum	345	Y	-	Y	Passed	20	Complete
Chickpea	300	Y	-	Y	Passed	20	Ongoing
Common Bean	192	Y	-	Y	Passed	20	Ongoing
Wheat	372	Y	-	Y	Passed	20	Ongoing
Musa	48	Y	-	Y	Passed	20	Scheduled
Yam	342	Y	-	Y	Failed		-
Lentil	137	Y	-	N	-		-
Cowpea	345	Y	-	N	-		-
Cassava	250	Y	-	N	-		-
Potato	0	N	Ref. set construction ongoing	-	-		-
Rice	0	N	Ref. set construction ongoing	-	-		-
Sweet Potato	0	N	Ref. set construction ongoing	-	-		-
Fababean	0	N	Ref. set construction ongoing	-	-		-
Foxtail Millet	0	N	Ref. set construction ongoing	-	-		-
Pearl Millet	0	N	Ref. set construction ongoing	-	-		-

40. G4008.01: Population development through Multiparent Advanced Generation Inter-crosses (MAGIC) among diverse genotypes to facilitate gene discovery for various traits in rice

December 2008–December 2009

Principal Investigator and Lead Institute
Hei Leung, IRRI; h.leung@cgiar.org
Plant Breeding, Genetics and Biotechnology Division, International Rice Research Institute, Manila, Philippines

Collaborating institutes and scientists
IRRI: E Redona; RK Singh

Background: Multi-parent advanced generation inter-crosses (MAGIC) is an experimental method that increases the precision with which genetic markers are linked to quantitative trait loci (QTL). Sixteen diverse founder lines involving eight each from the *indica* and *japonica* eco-geographic races of *Oryza sativa* L. were selected to develop two MAGIC populations. These lines were 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. The overall goal is to generate permanent mapping populations suitable for localising multiple QTLs for multiple traits to regions of 3 cM or less. Since the founder lines are elite breeding materials, some of the recombinant inbred lines can be used directly for multiple breeding objectives.

Population development and characterisation

Genotyping and phenotyping of founder parents: Genetic relationship of the 16 founder lines was determined using 50 SSR markers to ensure that the lines are genetically diverse. In addition, the DNA of founder parents was submitted for SNP genotyping at Cornell University for 1536 SNP markers. Phenotyping of MAGIC founder lines was started in the Philippines in 2009 dry season at one rainfed lowland and one upland drought trial site along with control. This will be continued in the coming wet season under non-stress and stress conditions (drought, salinity, and iron toxicity). S1 or S2 progenies derived from the MAGIC populations will be available for phenotyping for highly heritable traits that do not require extensive replicated trials and that are not influenced greatly by intra-plot heterogeneity.

Status of crosses: In the first crossing cycle, 28 single intercrosses in all possible combinations for each population were made using a half-diallel mating scheme. In the 2008 wet season, each single intercross was intermated to generate 70 four-way crosses out of all possible 210 crosses. We will make 35 out of the 105 possible eight-way crosses, with each founder line represented in each eight-way cross. For the MAGIC-indica population, planting of the four-way lines began in the 2009 dry season and 35 eight-way crosses will be completed by the end of July 2009. Due to problems with synchrony of flowering time, the 70 four-way lines of MAGIC-japonica were completed only at the end of the 2009 dry season. Planting of these lines is currently underway to generate the 35 eight-way lines. Line purity and hybridity of the different mating cycles are monitored phenotypically and genotypically with SSR markers.

Next steps: Each population will be advanced to produce at least 1000 RILs by single-seed descent. The MAGIC-indica population is ready for producing RILs that are expected to be harvested by December 2009. However, the RILs for MAGIC-japonica will be available only by April 2010 due to problem of non-synchronous flowering in early crossing cycles. In addition, we plan to subject these populations to two generations of intra-population mating to ensure rapid and uniform decay of linkage disequilibrium across the genome.

Samples from each population and bulks of selfed progenies will be genotyped with a set of genome-wide markers to validate their population structure. We plan to use a panel of SNP markers to generate graphical genotypes anchored to the rice physical map.

These populations will become permanent resources for breeding and genetics research. They will be used for fine mapping of multiple QTLs for specific traits. The novel allele rearrangements and enhanced genetic diversity in these MAGIC populations will facilitate the discovery of genotypic combinations. Beyond trait mapping, the highly recombined MAGIC populations may be used directly as source materials for the development of breeding lines and varieties adapted to different environments in Asia and Africa.

Acknowledgment: We thank Ian Mackay (NIAB) and Collin Cavanagh (CSIRO) for their discussion and active collaboration in this project.

41. G4008.02: Phenotyping sorghum reference set for drought tolerance

January 2008–December 2010

Principal Investigator and Lead Institute

Hari Upadhyaya, Principal Scientist, ICRISAT; h.upadhyaya@cgiar.org

ICRISAT-Patancheru, PO, 502324, AP, India; Telephone: 91-040-30712333 (office); 9866986262 (cell)

Collaborating institutes and scientists

- ICRISAT–India: SL Dwivedi; V Vadez; CT Hash; L Krishnamurthy
- ICRISAT–Mali: B Clerget; E Weltzein-Rattunde
- ICRISAT–Kenya: MA Mgonja
- UAS: PM Salimath
- KARI: CK Karari
- NPGRC: W Ntundu
- IER: M Diourte
- ISRA–CNRA: N Cisse

Drought is one of the most important yields reducing abiotic constraint worldwide. It is proposed to characterise sorghum reference set accessions for morphological diversity, assess drought x seed micronutrient interaction, screen reference set for post-flowering drought tolerance, and evaluate selected postflowering drought tolerant reference set accessions and stay-green QTL introgression lines under farmers conditions in Africa and Asia. The results of 2008/09 crop season across locations have been discussed below.

Pattern of diversity: Sorghum reference set accessions (375) were characterised for morpho-agronomic traits at four locations (Patancheru and Dharwad, India; Samanko, Mali; Kiboko, Kenya), and data is being computerised for statistical analysis.

Variation for seed micronutrient content: Experiment on genotype by drought interaction for Fe and Zn has been conducted under varying moisture conditions (drought and unstressed control) during the 2008/09 postrainy season at Patancheru, India, and replicated seed samples are being analysed for Fe and Zn contents.

Variation for stay-green and chlorophyll content: The accessions differed significantly for stay-green and chlorophyll content, which identified 28 accessions greener at flowering, 76 accessions greener at 4^{rth} week after flowering, and 40 accessions greener at 5^{th} week after flowering. IS#29691, 18876, 8882, 2730, 4285, 7861, 14414, 27164, 22287, 41724 and 8685 and 62(73)509 were found stay-green at 4^{th} and 5^{th} weeks after flowering. Most of these accessions represent to four basic and two intermediate races.

Transpiration efficiency (TE): The 375 accessions along with three controls were screened for TE under water stress (WS) and well-watered (WW) conditions. Plants were grown in large pots. WS was imposed six-weeks after sowing to measure plant biomass before imposing drought. WS plants were exposed to progressive water stress by letting plant loose no more than 150 g per day during the first 4 days after imposing stress; no more than 100 g per day in the following 4 days; and no more than 75 g per days during the rest of the experiment. Since pots were weighed every 4 days only, this corresponded to a maximum water loss of 600 g over 4 days, 400 g in the subsequent 4 days, and then 300 g for any 4 day interval in the rest of the experiment. WW plants were maintained well-watered by re-adjusting pot weight close to field capacity on those days when the pots were weighed (every 4 days), and by adding water two days after weighing to bring back pots close to field capacity, based on transpiration data (from previous weighing intervals). All plants were harvested when the transpiration of the drought set fell below 10% of the transpiration of the well-watered set, and work is in progress to analyse data.

Water uptake: A set of 210 entries was assessed for water uptake under stressed conditions in lysimeters, mimicking roughly the soil volume that sorghum plants would have at usual field planting densities. Two treatments imposed were water stress (WS) and well watered (WW). Preliminary analysis revealed large range variation in water extraction (10.2 kg plant^{-1} to 15.3 kg plant^{-1}) between the top and bottom extractors under WS conditions. The total water used in WW conditions varied from 10.5 kg to 42.3 kg of water per plants from 4 weeks after sowing until maturity. The ten highest water extractors under WS conditions were IS# 2367, 5720, 23988, 20709, 929, 20351, 1127, 20387, 14259 and 3971 while those that extracted lowest water under similar conditions were IS# 13848, 13452, 32234, 4821 18868, 33173, 28645, 30352, 30451 and SSM 547. In contrast, the accessions that extracted highest amount of water under WW conditions were IS# 20842, 14529, 22506, 30443, 18922, 2205, 10978, 23903, 5720 and 5622, while those that extracted lowest water under similar conditions were IS# 24009, 13452, 28645, 13848, 26731, 2398 and 12447 and 393 (411) 659, 452 (484) 510 and 651 (902) 656. Most of these accessions belong to four races and intermediate races. Further, analysis of grain weight and its relation to water extraction is in progress.

The PVC experiment gave also the opportunity to assess TE across a fairly large range of genotypes, over a long period of time (from 4 weeks after sowing until maturity). TE value varied between 2.44 g kg^{-1} water transpired to 6.09 g kg^{-1} water transpired. A detailed analysis is needed to investigate how differences in TE relate to the presence or absence of certain stay-green QTL. In short, in S35 background, few stay-green QTL introgression lines had TE above S35. Introgression lines with *Stg 3*, *Stg4* or *Stg B* had all fairly high TE equal or superior to S35. In R16 background, most *Stg B*, and *Stg3* and few *Stg4* introgression lines had high TE, equal or superior to R16. Some of the sorghum reference set accessions had higher TE than the highest TE of the stay-green trial entries.

42. 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
Km 45 Carretera Mexico-Veracruz, CP. 56130 Texcoco, Edo. de Mexico, Mexico

Collaborating institutes and scientists

- CIMMYT–Mexico: M Reynolds; Y
- Manes; T Payne; M Zaharieva; J Crossa; H Braun
- CIMMYT–Iran: J Kamali
- Seed and Plant Improvement Division, Agriculture and Natural Resources Research Center of Fars Province (SPII), Iran: AR Nikzad; S Tahmasbi; S Sarikhani; S Sarikhani
- INRA–Morocco: R. Dahan; N Nasrolha; H Ouabbou

The purpose of this project is to i) provide the reference samples for durum, winter and spring bread wheat, ii) characterise the spring bread wheat reference sample in multi-location trails for relevant agronomic traits, iii) perform association analyses based on revealed phenotypic and genotypic data.

1) **Seed multiplication and purification of the reference samples for durum, winter and spring bread wheat**
   To make the spring bread wheat reference sample available, 374 accessions selected form the wheat composite set were collated form the genetic resource centres at CIMMYT and INRA, France. Seed was multiplied during the winter cycle 2007/2008 at Obregon and Mexicali, Mexico, during the summer cycle 2008 at CIMMYT Headquarters in Mexico, Batan. Not all accessions could be multiplied. Forty-two accessions did not germinate, flowered too late, or did not enter into flowering stage. From the residual 332 accessions single plant seed was collected to purify the accessions and was planted in Batan in March 2009 for a second round of multiplication.

For the durum and winter wheat reference samples, 96 accessions could be collated from the genetic resource centres at CIMMYT, ICARDA and INRA. The durum wheat accessions were planted in Batan in March 2009. The winter wheat accessions will be planted in Mexicali in November 2009 due to necessary vernalisation requirements.

2) Characterise the spring wheat reference sample in multi-location trails for relevant agronomic traits

During seed multiplication of the spring wheat reference sample in Obregon, heading date of each accession was estimated to be able to classify the entries regarding their penology. The days from planting to heading varied strongly and ranged form 72 to 94. Based on this information a subset of 180 accessions was selected for drought phenotyping in multi-environment conditions. Fifteen drought tolerant CIMMYT lines, two bread wheat, and one durum wheat check were additionally included. Entries were divided into two phenology groups and organised in a modified split-splot design for distribution to collaborators. This subset was planted at three locations (two planting dates in Mexico, Morocco, Iran) and various physiology and agronomic measures were collected (see Table 1). Data are currently processed by each of the collaborators.

Tangible outputs, challenges, and next steps

By the end of 2009, purified accessions of the spring wheat reference sample and accessions of the durum wheat reference sample will be stored in the genetic resource centre at CIMMYT and be available for distribution. A subset of the spring wheat reference sample has already been requested for the GCP project G7009.01 and the challenge initiative: Improving drought tolerance in wheat for Asia.

Also splitting accessions in two phenology groups for phenotyping, genetic variation within each group is still high. Analyses based on the available data have to show the level of variation and if it will be possible to correct for the effect of phenology.

Fourteen accessions form the original durum wheat reference sample could not be collated form the genetic resource centre at INRA in 2008, but are available in July 2009 and will be requested.

Leave samples have been collected form all accessions, which will be used for DNA extraction and further high-throughput genotyping.

Table 1. Summary of physiology and agronomic traits collected on a subset of 200 accessions of the spring wheat reference sample at the three different locations within the project

Institution	CIMMYT	INRA	SPII
Site	Obregon, Mexico	Settat, Morocco	Darab, Iran
Measurements			
Planting	X	X	X
Emergence	X	X	X
% Ground cover	X	X	X
Leaf rolling	X		
Anthesis	X		
Canopy Temperature	X	X	X
Leaf chlorophyll (SPAD)	X	X	
Heading	X	X	X
Canopy Temperature	X	X	X
Wax	X	X	X
Leaf pubescence	X		
Stem sampling for CHO	ongoing	ongoing	ongoing
Plant height	X	X	X
Maturity	X	X	X
Yield	X	X	X

43. G4008.05: Connecting performance under drought with genotypes through phenotype associations

January 2008–December 2010

Principal Investigator and Lead Institute

Arvind Kumar, Plant Breeding, Genetics and Biotechnology Division, International Rice Research Institute (IRRI); a.kumar@cgiar.org
IRRI, DAPO Box 7777, Metro Manila, Philippines; Telephone: +63 (2) 580-5621; Fax: +63 (2)580-5699

Collaborating institutes and scientists

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

Phenological characterisation of germplasm collection

Association studies are dependent on large phenotypic variation, which complicates the ability to synchronise flowering for precision drought phenotyping (McNally et al. 2009). All accessions (227 *aus*, 499 *indica* and 329 *tropical japonica*) for association analysis were initially characterised for phenology at IRRI in 2008. Days to flowering ranged from 63 to 94 days, from 64 to 104 days, and from 67 to 123 days within *aus*, *indica*, and *tropical japonica* accessions, respectively. Within isozyme groups, accessions were separated into four (*aus* and *indica*) or five (*tropical japonica*) maturity groups for subsequent drought phenotyping (Figure 1). Phenological characterisation of all accessions is currently underway at partner sites.

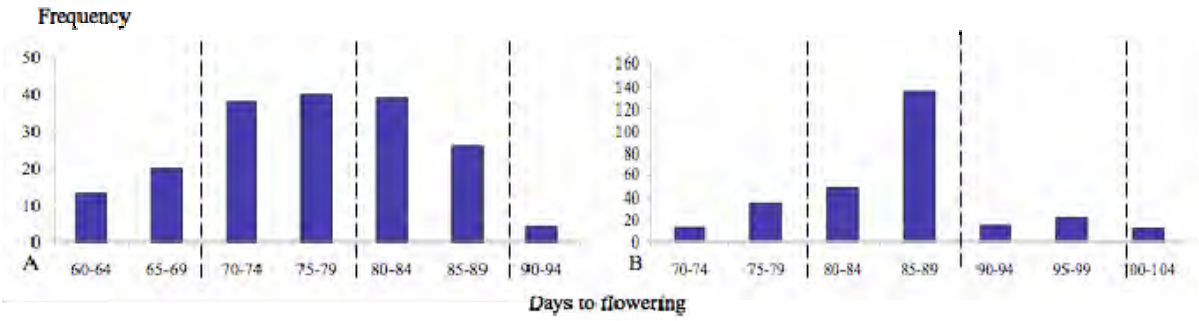


Figure 1. Days to flowering with (A) *aus* and (B) *indica* accessions. Broken lines indicate maturity groups used for subsequent drought phenotyping experiments.

Field phenotyping for agronomic and physiological traits under control and reproductive-stage drought conditions

Aus and *indica* accessions were screened under lowland conditions in the dry season 2009 at IRRI with two treatments: a well-watered control and drought stress imposed at the reproductive-stage. Accessions were separated into maturity groups determined the previous year, when seed was available accessions were sown in two maturity groups to increase the success of synchronising flowering. Groups were sown over a period of 32 days and the majority of accessions reached flowering at the same time with drought stress imposed at 10 days before flowering. Traits measured included leaf drying, days to flowering, biomass production, plant height, tillering, grain yield and yield components (Table 1).

Table 1. Range and means of relative reduction in grain yield and flowering delay under drought stress imposed at flowering for *aus* and *indica* accessions.

Isozyme group		Relative reduction in grain yield			Flowering delay		
	N	Average	Maximum	Minimum	Average	Maximum	Minimum
Aus	*227*	53.9	100	4.37	19	41	0
Indica	*449*	53.0	100	2.09	19	34	0

Tangible outputs delivered

Phenotyping of this diverse collection of accessions for yield under drought stress and yield potential has identified several donors which have now entered drought breeding programmes. A comprehensive dataset on phenology of these diverse accession is now available which will be valuable for future phenotyping of these accessions at these locations.

References

McNally KL, Naredo MEB and Cairns J (2009). SNP discovery at candidate genes for drought responsiveness in rice. In: *Drought Frontiers in Rice-Crop Improvement for Increased Rainfed Production* (Serraj R, Bennett J and Hardy B, eds.). World Scientific Publishing, Singapore, 14 pp.

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

May 2008SFebruary 2011

Principal Investigator and Lead Institute

James G Gethi, KARI; jgethi@wananchi.com
P.O. Box 340-90100 Machakos, Kenya +254 721831166 Fax +254 2 2338320

Collaborating institutes and scientists

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

1. Research activities at KARI

A total of 180 inbred lines from a diverse background grouped into three maturity groups were phenotyped under stress and non stress conditions in the dry season of 2008-9 at Kiboko. Data of up to four inbred lines (early) and six (late) did not have good data or did not survive to harvest.

Table 1. Phenotyping of 180 inbred lines under stress and non stress conditions at Kiboko for selected traits

Maturity Group	No. of lines	Stress level	DTF	DTS	Grain Wt Plot^{-1} (g)	100 kernel Wt. (g)	% of Well Watered
Early	58	Anthesis	63.6	65.6	758.1	23.8	115
	56	Vegetative	65.0	67.7	443.9	22.8	67
	56	WW	65.8	68.7	659.1	21.6	
Interm	60	Anthesis	66.9	68.8	688.2	23.5	75
	60	Vegetative	67.8	70.5	342.3	22.8	37
	60	WW	68.2	70.2	916.0	23.1	
Late	57	Anthesis	70.9	73.4	434.6	23.9	50
	54	Vegetative	70.4	74.0	343.9	23.0	40
	58	WW	70.7	72.5	867.5	24.6	

2. Research activities and progress at INRA-LEPSE

We investigated on phenotyping platform a subset of the panel covering a wide range of genetic origin and 'a priori' adaptation habilities (ie lines from different cycle of selection for drought tolerance, typical lines with good GCA for grain yield, lines selected for other agronomic targets from different breeding pools). Characterisation of leaf growth rate on Phenodyn revealed a large variation for maximum growth and its sensitivity to water deficit. These results presume large variation for biomass accumulation under water deficit within the entire panel that would be characterised next year.

3. Research activities and progress at ETH

Two phenotyping experiments were conducted at ETH in 2008/2009. The first experiment comprised 224 lines from the reference set the in growth pouches. The second experiment comprised of 33 genotypes from the GCP-INRA panel conducted in growth columns. Both experiments aimed at describing basic parameters determining the root distribution in soil. The measurements aimed to describe the potential the variation in rooting depth and root distribution in soil which is important for an efficient water and nutrient acquisition.

Growth pouches

Seven independent growth chamber replications were conducted in 2009 to obtain best linear unbiased predictors for a number of 22 traits. The data is available for QTL mapping (once marker data are available) and for characterisation of the genotypes.

Growth columns

Three independent greenhouse experiments, each comprising three harvest dates per genotype (V2, V4 and V6) were conducted. The data were used to characterise the genotypes for their heterotrophic development (harvest at V2) and during the early autotrophic growth (development V2-V6). These data are available to correlate traits across platforms and environments (pouch platform, phenopsis platform, field experiment).

4. Research activities and progress at CIMMYT

4.1. Genotyping and documentation of the maize reference set

The whole maize reference set of 240 inbred lines was genotyped at CIMMYT with 45 SSRs markers. Relevant information (pedigree, source of origin, climatic adaptation) was collected to be related to genotyping and phenotyping variation. Diversity and structure of the panel are being analysed.

4.2. Seed increase of maize reference set inbred lines and hybrids development

225 inbred lines were seed increased and test crosses with CML 312 developed and 220 hybrids produced. Seeds were provided for phenotyping trials at Tlaltizapan (Mexico) and ETH (Switzerland). During the winter (2009A) cycle lines were multiplied and hybrids produced for phenotyping at Kiboko (Kenya) and INRA (France). Seeds from the founder plants (used for genotyping activities) were increased separately, and will be stored at CIMMYT GeneBank as genetic stock.

4.3. Drought phenotyping of the maize reference set inbred lines and hybrids

Phenotyping of the reference set (225 inbred lines) and the set of 220 hybrids, divided in 3 maturity groups was carried out under drought and well-watered conditions during the 2009 dry season in Tlaltizapan (Mexico). Days to anthesis, silking and maturity, senescence, leaf rolling, panicle weight, plant and ear aspect, yield and yield components data were recorded. Crop establishment and biomass development were evaluated measuring NDVI with a portable spectro-radiometer (Green Seeker).

Tangible outputs

Grieder C (2008) *Depth and vertical distribution of roots: Variation within a diverse panel of tropical maize inbred lines*. Master Thesis, ETH Zurich.

Grossard F (2008) *Drought tolerance in maize: genetic variability for growth and its sensitivity to water deficit within the crop species*. Master Thesis, INRA Montpellier, Univ Blaise Pascal, ENITA Clermont, France

45. G4008.42: Developing DArT markers for several crops in the GCP

January 2008–December 2008

Principal Investigator and Lead Institute
Jean-Christophe Glaszmann, Agropolis−CIRAD/GCP; jean-christophe.glaszmann@cirad.fr

PROJECT UPDATE NOT SUBMITTED

46. G4008.45: A Nested Association Mapping (NAM) population of rice. Laying the bases for highly efficient QTL characterisation

August 2008–July 2010

Principal Investigator and Lead Institute
Mathias Lorieux, Agropolis–IRD/CIAT, Colombia; mathias.lorieux@ird.fr; m.lorieux@cgiar.org
CIAT Biotechnology Unit, AA 6713, Cali, Colombia. Tel.: +57 2 445 31

Collaborating institutes and scientists

- CIAT: César P Martinez; Edgar Torres
- WARDA: Marie Noelle Ndjiondjop

Context

Modern breeding strategies often fail to include precise genetic information. Marker-Aided Selection (MAS) strategies have proven to be more efficient than conventional selection in several cases, but still suffers of (1) lack of precision in the localisation of the genes of agronomical importance (the so-called QTLs, for Quantitative Trait Loci) and (2) are often limited to the alleles available in the crossing scheme used for QTL detection, i.e., the genetic information obtained from a particular cross between two genotypes (or lines) will not be useful when working with other genotypes. We propose to develop a new genetic resource, called a Nested Association Mapping (NAM) population, that would (1) help in linking the genomic tools available for rice, (2) give access to a much higher allelic diversity at the important QTLs than "conventional" mapping approaches do, (3) allow fine mapping of QTLs (i.e., localise them with high precision on the rice genome), thus increasing significantly the efficiency of MAS strategies, and (4) provide interesting and promising genetic materials (advanced lines) for direct introduction in breeding schemes.

Findings and implications

Selection of parental lines

Activities at CIAT: A diversity survey of the 48 candidate lines as parents for the metapopulation was done using a set of 24 SSR markers. The data were analysed using the Darwin 5.0 and NTSYS programmes and the SAS statistical package (Multiple Correspondence Analysis), in order to identify a final subset of fifteen tropical *japonica* lines that maximize the genetic diversity (Table 1).

Production of F1 hybrids

F_1 hybrids have been produced by crossing the *indica* IR64 accession as female with all 48 candidate lines. Ten F1 seeds were sown per retained combination, and were checked for heterozygosity using 3 SSRs. They were then brought to the field in order to be selfed

and to produce the F2 populations that will represent the starting point of the Single Seed Descent (SSD) process.

Activities at WARDA: A similar approach has been followed at WARDA, in order to choose the other candidate accessions.

Table 1: List of selected varieties crossed with acc. IR64 at CIAT

Accession	*Pedigree*
IRAT 122	MAKALIOKA / CHIANAN 8
IRAT 146	IRAT 13 / DOURADO PRECOCE
ITA 164	LAC 23 (RED)/MULTIPLE PARENT 25
TOX 1011-4-1	IRAT 13/DP689//TOX490-1
CT6241-2-2-1-3	NGOVIE/TAIPEI 309//COL 1 X M312A-74-2-8-8
CT6241-19-2-1-3-1P	NGOVIE/TAIPEI 309//COL 1 X M312A-74-2-8-8
CT10011-5-4-M-M	CT6424-12-1-4-1-2//CT6515-18-1-3-1-6/CT8088-14-16
CT10006-7-2-M-2	CT6241-2-2-1-3//CT6516-23-10-1-2-2/CT8071-13-1
CT10048-6-3-M-2	CT6240-12-3-3-5//CT6516-23-10-1-2-2/CT8076-15-2
CT8556-37-1-3-1-M	TOX 1859-102-5M-7/COL 1 X M312A-74-2-8-8//TOX 1837-103-1-4
CT10035-26-4-2-M	CT6258-5-2-6-5//CT6129-17-2-1P-2/CT6196-33-10-2-1
CT10037-56-6-M-M	CT6424-12-1-4-1-2//CT6129-17-2-1P-2/CT6196-33-10-2-1
CT10045-5-5-M-1	CT6258-5-2-5-3-3P//CT6516-23-10-1-2-2/CT8060-2-1
Oryzica Turipana	P 4971/P 5004
Liderança	Not available

Implications
We think that the NAM population will provide the rice research community with a highly efficient and powerful genetic resource that would allow to fully taking advantage of the numerous genomic tools that are now available for this species. Accurate, powerful, multi-allelic QTL detection and fine/ultrafine mapping of QTLs for many important agronomic traits are expected from future studies based on this resource.

We expect that GCP and external partners, based on the resource created in this project, will develop phenotyping projects on drought and other traits.

Next steps and/or challenges
Of each F2 population, 400 seeds will be sowed in order to make sure that we obtain final F7 population sizes of 200 individuals through the SSD process.

Sequencing of IR64 will certainly increase very much the potential of this NAM metapopulation, in allowing us to infer the ancient recombination events based on complete genomic information instead of partial (SNP) coverage.

47. G4008.46: Sorghum MAGIC: Multiparent Advanced Generation Inter-Cross development for gene discovery and allele validation

September 2008–August 2009

Principal Investigator and Lead Institute
C Tom Hash, ICRISAT; c.hash@cgiar.org
Patancheru, Hyderabad 502 324, Andhra Pradesh, India; Phone: +91 40-30713322; Fax: +91-40 30713074

Collaborating institutes and scientists

- ICRISAT–Kenya: Mary A Mgonja
- ICRISAT–Mali: H Fred W Rattunde
- ICRISAT–India: S Senthilvel; SP Deshpande
- NIAB: Ian Mackay
- NRCS: R Madhusudhana

1. Introduction/Background

Sorghum is the 5th most important cereal crop globally and its germplasm exhibits substantial phenotypic and genotypic diversity. Further, an aligned genome sequence, DArT platform, and large numbers of SSR markers are available, and SNP platforms are in pipeline for this species, so many of the tools required for applied marker-assisted breeding are in place or in advanced stages of development. However, although a GCP reference collection that is representative of sorghum germplasm cultivated in developing countries has been constructed, and phenotypic assessment of this for phenology and drought tolerance-related traits has begun to generate data sets required for association mapping, few conventional sorghum bi-parental mapping populations have been produced based on germplasm relevant to developing country needs so there are relatively few relevant targets for applied marker-assisted selection.

A new mapping population type, referred to as the multi-parent advanced generation intercross (MAGIC) population, has been proposed as a complement to conventional bi-parental mapping populations and association mapping using diverse germplasm or breeding materials. This extension of the advanced intercross can be used for fine-mapping multiple QTLs for multiple traits and has the advantage of sampling a larger set of parental combinations than would normally be practical using bi-parental crosses. Further, MAGIC-like populations based on genetically diverse but elite parental founder genotypes can be expected to generate lines with potential for direct use as improved cultivars or hybrid parental lines. Thus MAGIC-like populations could offer a lower cost way to generate coarse-mapping and subsequent fine-mapping information for traits of interest to applied breeding programmes in developing countries. Phenotyping of MAGIC populations would be subject to the same limitations as bi-parental populations or association panels of similar size—and in sorghum would largely be constrained by phenology. This project explores the potential for developing MAGIC-like populations for sorghum targeting four distinct agro-ecologies in sub-Saharan Africa and South Asia.

2. Ongoing/completed activities

2.1 Selection of founder genotypes for four populations

Parental founder lines proposed by regional sorghum breeders in South Asia were in hand at ICRISAT-Patancheru at the start of this project, but several founder lines proposed by sorghum breeders in Africa were still in transit. Missing lines were introduced to India and most passed successfully through post-entry quarantine seed multiplication during the 2008/09 postrainy season. However, three highly photoperiod-sensitive long-duration lines from West Africa failed to flower. As including these parents will further delay initiation of the crossing programme for creating the West Africa-targeted MAGIC-like population, we propose to proceed with the founder materials now in hand (Table 1).

Table 1. Sorghum founder lines proposed for MAGIC-like populations

Rainy season South Asia	Postrainy season South Asia	Western and Central Africa	Eastern and Southern Africa
296B (IS 41245)	B35 (IS 40653)	Grinkan	AF 28
IS 41397	E 36-1 (IS 30469)	IS 15401	E 36-1
BTx623 (IS 40583)	IS 2205	Soumba	IS 9303
ICSV 745 (IS 36524)	IS 18551	Lata 3	IS 23520 (Sima)
ICSV 93046	M 35-1 (IS 1054)	Fambe B	KARI-Mtama 2
NTJ 2 (IS 30468)	N 13 (IS 18331)	Kapelga	Macia
PB 15520	Parbhani Moti (GD 31-4-2-3)	02 SB-F5DT-12B	MB30
SSG 59-3	R 16 (IS 18482)	CSM 63E	Ochuti (IS 21124)

2.2 Develop and assess methods for design and maintenance of MAGIC populations
Genetic models for development of MAGIC populations based on 8 and 16 founder parents were developed by NIAB and communicated to ICRISAT in early 2009, revealing two critical project design flaws. First, numbers of crosses required for the 2nd and 3rd generations of intercrossing to produce "true" MAGIC populations are prohibitive for practical application in sorghum given the relatively high rate of failure of manual emasculation and crossing in this species. Second, a 3rd generation of intermating is essential prior to initiation of inbreeding to develop the 1st sets of inbred lines for each population; but this was not initially planned for in the present project (or the earlier proposal on which it was based). Thus practical sorghum MAGIC-like populations (requiring manageable numbers of crosses) will have less predictable linkage disequilibrium decay than anticipated, and take longer to develop than planned.

2.3 Initiate crossing to develop MAGIC populations
Half-diallel sets of crosses of the 8 founder parents of two proposed MAGIC populations for South Asia were produced, along with DNA samples and selfed seed from each of the parental plants involved. However, this single generation required 6 months, not the planned 4 months, due to problems matching parental flowering times for crossing.

3. Future activities
It is recommended to reduce the scope of this project, initially advancing only one or two sorghum MAGIC-like populations to better assess their practical difficulties in this crop.

48. G4008.51: Use of molecular marker and physiological tests to characterise Hungarian rice diversity and new sources of blast resistance

April 2008–October 2008

Principal Investigator and Lead Institute
Ibolya Simon-Kiss, HAKI; jancsom@haki.hu
Institute for Fisheries, Aquaculture and Irrigation, Anna-liget 8, Szarvas H-5540, Hungary; +36-66-515-300; Fax: +36-66-312-142

Collaborating institutes and scientists
- HAKI: Mihály Jancsó (PhD student)
- Agropolis–CIRAD: Jean Christophe Glaszmann; Claire Billot; Didier Tharreau

1. The Context of the project - Research activities at HAKI
Hungary is the northernmost border of rice (*Oryza sativa* L.) cultivation in Europe (47-48°N) thus the special climatic conditions (cold spring, short season, long daylight) makes it possible to breed only special varieties. The need of scientific plant improvement is very high to maintain the long term sustainable rice cultivation and preserve the genetic material achieved in the last century.

The Hungarian Ministry of Agriculture and Rural Development has signed cooperation with the CGIAR-GCP in 2007. Several new projects was started and one of these projects was the G4008.51 focusing on the genetic diversity and on the blast resistance of the varieties in the frame of a scientific training in 2008 at CIRAD, Montpellier.

2. The research activities at the CIRAD
2.1 Discover the genetic diversity in the Hungarian germplasm collection
This part of the research was hosted by the Genotyping platform (head Dr. C. Billot). The genetic diversity of the germplasm collection selected and maintained at the HAKI (74 accessions) were investigated by SSR molecular markers using the standard MATAB DNA isolation method and LI-COR DNA Sequencers. The results were checked and validated by SAGA Generation 2 and analysed by DarWin5 software.

26 different loci were analysed and the allele sizes were determined using five different standard variety mixtures (Pt-1 to Pt-5) with known allele sizes. Based on our results we can determine that genetic background of the Hungarian germplasm is very narrow but contain also very good tolerance to cold (cv. HSC-55), salt (cv. Dáma) and drought (cv. Sandora) stresses as well as extremely short duration (cv. Ábel). High genetic difference was only found when we compared the traditional Hungarian varieties and the recently introduced varieties like CL-62 from Brazil and IRH-1, IRH-2 and IRH-3 from Iran.

2.2 Blast resistance of Hungarian rice varieties – new sources of blast resistance

Blast is a major constrain of rice cultivation in Hungary as well as worldwide. Blast resistance of 72 genotypes was investigated in a glasshouse experiment by eight different, previously characterised European *Magnaporthe* isolates (from France: FR13, FR94, FR137, Portugal: PR9, PR14, Hungary: HN1, Spain: SP6 and Italy IT10. Infection type of the genotypes was scored using 6 categories to determinate the symptoms. Six standard varieties were used to check the virulence of the isolates and to compare the sensibility pattern of the Hungarian varieties (Table 1).

In our project, six genotypes (CL-62, Dáma-A, GB, Kirara, p140b and IRH3) were found resistant to all *M. grisea* isolates. The states approved varieties were divided into sensitive (cv. M-60, cv. M-488), medium tolerant (cv. Ábel, cv. Janka, cv. Sandora, cv. Risabell) and resistant (cv. Dáma) groups. This fact also highlights the importance of the breeding for blast resistance.

Table 1. The results of the interactions (S-susceptible, R-resistant and 9-non defined) by the isolated *Magnaporthe* strains and some state approved rice variety from Hungary and the control lines (Montpellier, 2008)

Nr.	Name	FR13	FR94	FR137	HN1	IT10	PR9	PR14	SP6
1.	M-488	S	S	S	S	S	S	S	S
2.	M-60	S	S	S	S	S	S	S	S
6.	Ábel	R	S	R	S	R	S	S	R
4.	Bioryza H	S	S	R	S	S	9	R	S
5.	Dáma	R	R	R	R	R	R	R	R
6.	Janka	R	S	R	S	R	S	S	R
7.	Ringola	R	S	R	S	R	S	S	S
8.	Risabell	R	S	R	S	R	S	S	R
9.	Sandora	R	S	9	S	R	S	S	R
C1	Maratelli	S	S	S	S	S	S	S	S
C2	Shin2	S	S	R	S	S	9	R	S
C3	Fujisaka5	R	S	R	S	R	S	S	R
C4	K59	R	R	R	R	R	S	R	R
C5	Aichi asahi	R	S	R	S	R	S	9	9
C6	K1	R	R	S	R	R	R	R	R

3. Next challenges and the first steps further

The results of this project were directly integrated into the Hungarian rice research and breeding activities. We have continued the cooperation by sending blast samples from Hungary to the CIRAD to maintain the good contact (GCP projects are preferred). Based on the efficient and friendly cooperation, new project plans are formed in the field of abiotic and biotic stresses.

References

Jancsó M, Simon-Kiss I, Árvai Zs, Billot C and Tharreau D (2009). Studies on Blast (*Magnaporthe grisea* (Herbert) Barr) Resistance of Hungarian Rice Varieties In: *Hagyomány és haladás a növénynemesítésben* (Veisz O, ed). XV. Növénynemesítési Tudományos Napok, Budapest, pp. 218-222. (ISBN: 978-963-508-575-0)

49. G4008.52: Genetic Resources Support Services (GRSS)—Implementation feasibility work programme

September 2008–December 2008

Principal Investigator and Lead Institute

Jean Christophe Glaszmann, GCP

1. Background

The Generation Challenge Programme (GCP) is at the heart of a research and capacity-building network that uses plant genetic diversity, advanced genomic science and comparative biology to develop tools and technologies that help plant breeders in the developing world produce better crop varieties for resource-poor farmers. All GCP activities aim to generate usable products in the form of knowledge and tools for plant breeders. A core mission is to promote and disseminate these products through provision of cost-efficient services worldwide. A suite of plant breeding support services (PBSS) are being developed or are envisaged. These include:

1. Genetic resources support service
2. Genotyping support service
3. Phenotyping support service
4. Analysis and policy support service

Beneficiaries of the PBSS will be partners in GCP projects, but also breeders outside GCP working on crop improvement in marginal environments.

The work programme contained in this document focuses on preparation of options, feasibility and planning for creation and implementation of the Genetic Resources Support Service (GRSS).

Subprogramme 2: Genomics towards gene discovery

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

August 2007–July 2009

Principal Investigator and Lead Institute

Andy Pereira, VPI; pereiraa@vbi.vt.edu

Washington Street, Blacksburg, VA 24061, USA; 001-540-231-3784; 001-540-231-2606

Collaborating institutes and scientists

- IRRI: Hei Leung; Rachid Serraj; Jill Cairns
- HZAU: Lizhong Xiong

Research within the GCP and other ongoing research on abiotic stress biology, has provided researchers a number of candidate genes with a potential role in drought response and resistance. These genes have been identified in a number of crops, in response to a variety of environmental stresses and by data derived from breeding, genetics, physiology and genomics, but their exact role in drought response/resistance is unknown. The analysis of mutants is one of the most reliable and time-proven ways of correlating the genotype to a phenotype. The international research community has generated significant mutant resources in the two sequenced plants Arabidopsis and rice (Krishnan et al., 2009). In this project the aim has been to provide drought response phenotypes for an extensive list of candidate orthologous genes in the two plants selected for their potential role in drought responses and resistance mechanisms. The comparative analysis of gene functions between the dicot and monocot plants will be applicable across a wide number of crop plants.

A list of candidate Stress-Associated Genes (SAG) was assembled in rice and Arabidopsis for which there were available knockout (KO) mutants. In rice the mutant lines were evaluated for important drought related physiological parameters at vegetative and reproductive drought stages in field or controlled environmental conditions. The field testing at HZAU was done with about 20 plants/line in the moveable rainout shelter and drought stress given at the time of booting to test the drought sensitivity at anthesis stage. Plant phenotypes scored were leaf rolling, wilting, maturity, and grain yield per plant measured under drought stress. From around 100 T-DNA tagged mutant lines (http://rmd.ncpgr.cn/), 12 mutant families showed segregation for drought sensitivity. Segregation analysis was performed by PCR methods to test the co-segregation of phenotype and genotype of T-DNA insertions. Among the 12 families, only 2 families showed 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. In addition over-expression lines were generated for three SAGs whose mutant lines showed drought sensitivity phenotypes identified before (project 2005-9). One of these overexpression lines (Fig. 1), with its gene encoding a bZIP transcription factor (TF), showed improved drought and salt tolerance and elevated sensitivity to ABA.

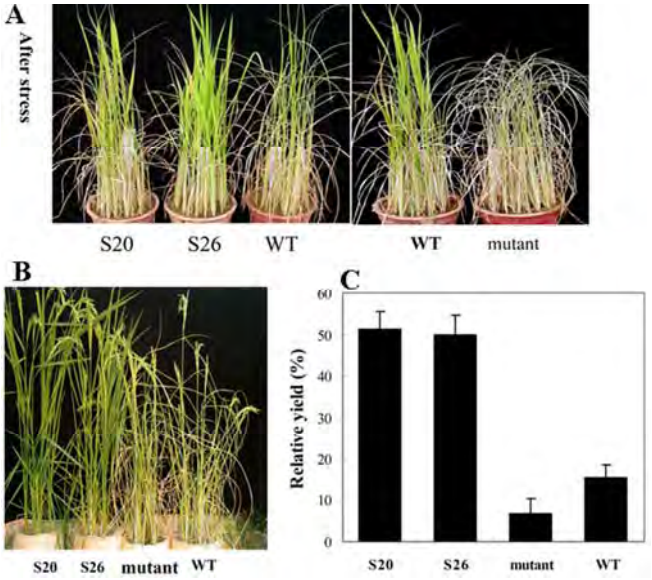


Fig 1: Drought resistance phenotype of rice lines (S20, S26) expressing a bZIP transcription factor and KO mutant. A: Plant phenotype after medium drought stress treatment (water withheld for 5 days followed by recovery).
B: Phenotype under severe drought stress at flowering stage. C: Relative yield per plant after being stressed by drought at flowering stage.

In *Arabidopsis*, drought gene expression studies identified significant differentially regulated genes (Krishnan & Pereira, 2008) orthologous to rice. Out of 200 mutant lines screened in replications for medium drought response phenotypes, measured in the reduction of biomass accumulation under constant soil water deficit, 10% lines showed an altered drought response phenotype. These lines were analysed on microarrays and display a network of drought response genes that have changed in the mutants (Fig 2).

The outputs of the project are: a) Drought response phenotypes of around 100 SAG mutants, revealing altered drought response phenotypes for 2 genes in rice. b) Drought response phenotypes of around 200 Arabidopsis regulatory SAG mutants, revealing altered drought phenotypes for 20 regulatory genes with rice orthologs. These genes are candidate genes for drought resistance in other plants.

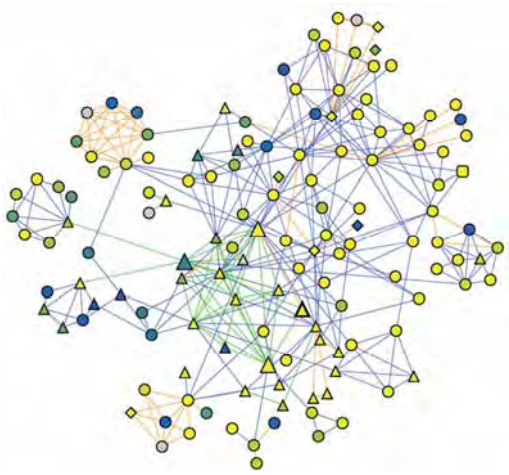


Fig 2: Gene network around a TF gene (large triangle) that when knocked out gives drought resistance. The network of 138 genes whose expression is altered in the KO mutant comprises drought regulated genes.
The gene (nodes) functions in the network are given by form: triangles (TFs), diamonds (kinases).

References

Krishnan A, Guiderdoni E, An G, Hsing YC, Han C-d, Lee MC, Yu S-M, Upadhyaya N, Ramachandran S, Zhang Q, Sundaresan V, Hirochika H, Leung H, Pereira A (2009) Mutant resources in rice for functional genomics of the grasses. *Plant Physiol* 149: 165-70.

Krishnan A, Pereira A (2008) Integrative approaches for mining transcriptional regulatory programmes in *Arabidopsis. Brief Funct Genomic Proteomic.* 7: 264-74.

51. G4008.06: Single nucleotide polymorphism discovery, validation, and mapping in groundnut

January 2008−December 2009

Principal Investigator and Lead Institute
Steven Knapp, UGA; sjknapp@uga.edu

PROJECT UPDATE NOT SUBMITTED

52. G4008.07: Improving molecular tools for pearl millet

January 2008–December 2009

Principal Investigator and Lead Institute
C Tom Hash, ICRISAT; c.hash@cgiar.org
Patancheru, Hyderabad 502 324, Andhra Pradesh, India; Phone: +91 40-30713322; Fax: +91-40 30713074

Collaborating institutes and scientists

- ICRISAT: V Vadez; RK Varshney; T Nepolean; S Senthilvel
- AICPMIP: IS Khairwal
- CAZRI: OP Yadav
- ILRI: Michael Blümmel

1. Introduction/Background

Pearl millet is the 6th most important cereal crop globally, but has limited genomic tools available for its improvement. This project is intended to strengthen genomic resources for pearl millet, producing advanced generation inbred and testcross seed of available partially inbred mapping populations, generating large EST libraries and mining these for markers, and producing a skeleton map and phenotype data for a new RIL population.

2. Ongoing/completed activities

2.1 Development of EST libraries from tissues of two pearl millet genotypes

In the first year of this project, elite pearl millet inbred lines 841B-P3 (drought sensitive) and 863B-P2 (drought tolerant), were subjected to drought stress using a standard dry-down protocol) at boot leaf stage under controlled conditions. Total RNA was extracted separately from stressed root and leaf tissues of each inbred. cDNA was synthesised for the four RNA samples, and the cDNA samples then sent to JCVI for 454 sequencing.

2.2 Bioinformatics analysis of EST sequences produced by 454 sequencing

This year large numbers of short EST sequences generated from each cDNA sample were cleaned of rRNA sequences, vector, ligator and poor quality sequences at JCVI (Table 1), and 112,864 unigenes (34,270 contigs and 78,594 singletons) were identified from the cleaned EST sequences using the PLANTTA pipeline assembly programme. At ICRISAT these unigenes were mined for Class I SSRs. The cleaned EST sequences were also analysed separately using the Cap3 assembly programme. Finally, 236 primer pairs were designed for putative SSR markers using the Primer3 programme, after removing additional redundant sequences.

Table 1. Pearl millet EST sequences (raw and clean) generated by 454 sequencing

	841B-P3 leaf ESTs	841B-P3 root ESTs	863B-P2 leaf ESTs	863B-P2 root ESTs
Total raw sequences (average length ~205bp)	183,657	188,823	174,513	190,534
Total clean sequences (average length ~240bp)	58,445	139,913	71,535	127,705

2.3 PCR optimisation, polymorphism and mapping newly developed markers
PCR conditions for the 236 newly developed EST-SSR primer pairs currently are being optimised and PCR products screened for polymorphism detection against parents of three F7 RIL mapping populations (based on crosses 841B-P3×863B-P2, H 77/833-2× PRLT 2/89-33 and 81B-P6×ICMP 451-P8). Earlier generations of these populations were used to map phenology, grain and stover yield, drought tolerance and disease resistance.

2.4 Generation advance and multiplication of pearl millet RIL mapping populations
During the 2009 summer season (Jan-May), progenies of several RIL populations were advanced by one generation of modified single-seed descent with enforced selfing. These RIL populations included two sets of F5 progenies (F6 seed harvested) [192 based cross ICMB 89111-P5× ICMB 90111-P5 and 174 based on cross ICMP 451-P6×H 77/833-2-P5(NT)], and one large set of 499 F3 progenies (F4 seed harvested) based on cross ICMB 01222-P1×ICMB 95333-P1. In addition, selfed seed of two F7 RIL progeny sets [based on crosses ICMB 89111-P2×ICMB 90111-P2 (191 RILs) and ICMB 89111-P6×ICMB 90111-P6 (204 RILs)] was multiplied for medium-term storage and future distribution.

2.5 Drought tolerance phenotyping of new pearl millet RIL mapping population
Testcross hybrids of 230 lines of a new pearl millet RIL population based on a cross of RIB 335/74 and Jakhrana S8-35 were evaluated, with parental testcrosses and commercial hybrid controls, under fully-irrigated non-stress and terminal drought stress conditions in the summer off-season drought nursery (Jan-May 2009) at ICRISAT. Testcross yield reductions for grain (37%), stover (18%) and biomass (27%) in the stress treatment, compared to the irrigated control, indicate timing of stress onset was optimal. Repeatabilities were >0.5 for most observed non-stress grain and stover yield component traits, while those of stress grain, panicle, and biomass yields were marginally below this.

3. Future activities
Complete mapping of newly developed polymorphic EST-SSR markers on three RIL populations, and develop improved pearl millet consensus linkage map. Develop primers for additional EST-based markers (SNPs and indels) and map some of these. Refine comparative maps of rice, sorghum and pearl millet based on comparisons of pearl millet ESTs with the aligned rice and sorghum genome sequences. Complete data analysis of summer 2009 drought nursery field trial RIL population testcross data set. Generate stover quality data from ground stover samples from this field trial. Skeleton-map new RIL population for which phenotype data is now available. Transfer seed samples of finished F6/F7 RIL populations to ICRISAT Genebank. Advance remaining RIL populations by one additional selfing generation.

4. Expected outputs
Additional pearl millet gene-based PCR-compatible DNA markers mapped. Improved consensus linkage map of pearl millet with enhanced marker density. Enhanced knowledge of syntenic relationships between pearl millet, sorghum and rice. One new skeleton-mapped pearl millet RIL population with phenotype data set suitable for mapping novel QTLs for grain and stover yield, stover quality and drought tolerance. Seed of 10 nearly homozygous pearl millet RIL populations available for future use.

53. 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 and Lead Institute
Shoshi Kikuchi, NIAS; skikuchi@nias.affrc.go.jp
National Institute of Agrobiological Sciences, Kannon-dai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan; (81) 29-838-7007, Fax (81) 29-838-7007

Collaborating institutes and scientists
IRRI: Hei Leung (genetics); Venuprasad Ramaiah, Arvind Kumar (breeding); Rachid Serraj (drought physiology); Ramil Mauleon (bioinformatics) and Violeta Bartolome (statistics)

Context: Two recent advances in gene expression analysis and drought-QTL mapping were applied to test the hypothesis that gene expression patterns in a chromosomal context are causally correlated with manifestation of drought tolerance as detected in near-isogenic lines. Application of a new comprehensive 44K oligoarray platform enabled us 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. NILs were originally established through the backcrossing of Aday-Sel (drought-tolerant) to IR64. Two series of NILs such as IR77298-5-6 and IR77298-14-1-2, and their drought tolerant progeny IR77298-5-6-18 and IR77298-14-1-2-10 and drought susceptible progeny IR77298-5-6-11 and IR77298-14-1-2-13 were used as materials. Parallel to transcriptome analyses, fine-scale genotyping of the NILs was performed to determine whether expression signatures co-segregate with specific regions of the genome. Results from this series of studies revealed 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 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.

Findings and implications:

1) Physiological and phenotypic analysis of NILs
   Each pair of NILs showed difference in biomass under water stressed conditions in the phytotron and the trend of biomass reduction was similar to that observed under field conditions. Compared with IR77298-5-6-18, IR77298-14-1-2-10 showed much higher tolerance to the drought stress treatment (0.5FTSW: Fraction of Transpirable Soil Water)). Ratio of water uptake was measured among IR64 and NILs. Water uptake between IR77298-14-1-2-13 and IR77298-14-1-2-10 was not large, while IR77298-5-6-18 showed larger increase of water uptake relative to IR IR77298-5-6-11.
2) Gene expression profiles of NILs in well-watered and stressed conditions
   Gene expression analysis in leaves and panicles of NILs based on the 44K oligoarray system was performed in NIAS. Differential responses of gene expression to the drought stress were observed between IR77298-14-1-2-10 (highly tolerant) and IR77298-5-6-18 (moderately tolerant). That is, IR77298-5-6-18 (moderately tolerant) showed down-regulation of the drought responsive genes (DRG) even in the well-watered condition; those genes were down-regulated in the stress treated IR77298-5-6-11, whereas IR77298-14-1-2-10 showed much drastic down-regulation of DRG than IR77298-14-1-2-13. These profiles of gene expression to drought stress, suggest that IR77298-5-6-18 seems to be well-prepared for the drought stress, while IR77298-14-1-2-10 can make quick response against the drought stress treatment. Thus, it appears that different strategies to drought stress are operating in the two pairs of NILs.
3) Combination of transcriptome data and genotyping data of NILs
   Genotyping data produced in IRRI suggested that there was no common introgressed DNA region from Aday-Sel to IR77298-5-6 and IR77298-14-1-2 lines. By comparison between IR77298-14-1-2-10 (highly tolerant) and IR77298-14-1-2-13 (susceptible), five regions on three chromosomes showed the difference of DNA introgression. They are, Chr2: 6-11Mb, 18-22Mb, Chr8: 3-19Mb,

Chr11: 3-4Mb, 16-18Mb regions. In the case of IR77298-5-6-18 (moderately tolerant) and IR77298-5-6-11 (susceptible), only one region on Chr9: 14-16Mb showed the difference of DNA introgression. Currently the size of introgressed DNA region is still too large to speculate which gene may play the key role. However, we have dramatically reduced the number of candidate genes for further investigation. For example, among the 596 predicted genes underlying the Chr02:18~22Mb region, enes with functional categories can be identified: encoding Transcription factor: 31, encoding protein kinase: 16, encoding protein phosphatase: 1, and Ubiquitin-related proteins: 5.
Our results suggest a) considerable potential to identify genes with large contribution to drought tolerance, and b) combination of genotyping and gene expression analysis is a good approach to narrow down specific gene(s) related to the phenotypes. This methodology could be applied to other traits.

Next steps and challenges:
1) We will generate transcriptome data in roots. The NILs showed differnces in water uptake ratio. We also have root samples of the NILs under non- and stressed conditions.
2) We will further backcross the NILs to narrow down the regions of introgresed DNA. Next generation of NILs will be used to validate the predicted results
3) To gain additional evidence that the genes in the target regions are important for drought response, we have produced F1 plants from two sets of NILs for conducting allelic imbalance (AI) assay. This work is done in collaboration with National Institute for Agricultural Botany (NIAB) in Cambridge. Significantly imbalanced expression would suggest the presence of cis-regulatory elements that may play a role in response to drought stress.

54. G4008.09: Development of genetic and genomic resources for breeding improved sweetpotato

January 2008–December 2009

Principal Investigator and Lead Institute

Roland Schafleitner, CIP; r.schafleitner@cgiar.org
Avenida La Molina 1895, Apartado 1558, La Molina, Lima 12, Peru
Tel.: +51 (0) 1 349 6017, Fax: +51 (0) 1 317 5326

Collaborating institutes and scientists

- CIP: Luz Tincopa; Reinhard Simon; David Tay; Genoveva Rossel; Wolfgang Gruneberg; Merideth Bonierbale
- EMU: Ivone Martins Muocha
- NAARI: Robert Mwanga, Gorretti Ssemakula
- INIA–Uruguay: Francisco Vilaro
- EMBRAPA: Andre Dusi
- DArT P/L: Andrzej Kilian

Context

Sweetpotato production 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. Improved and well adapted sweetpotato varieties with increased tolerance to biotic and abiotic stresses can significantly contribute to increasing productivity and therefore 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.

In the present project we develop a well defined Composite Genotype Set and obtain information on sweetpotato gene sequences, new markers and a reference map to mobilize allelic diversity for enhancing sweetpotato breeding programmes in CG Centers and at NARS.

Research activities and progresses

A Composite Genotype Set for sweetpotato has been established. It consists of 480 accessions representing the genetic and the trait diversity of this crop. The majority of the Composite Genotype Set has been cleared from viruses and is ready to be distributed to users.

At project start, the sequence information for sweetpotato genes was limited to about 500 genes and ~ 22.000 EST sequences that were assembled to 2476 contigs and 6598 singletons (Childs et al., 2007). To augment the information on sweetpotato gene sequences, we have submitted two cDNA libraries, one derived from leaves and one from stems, to 454 sequencing. The obtained 500.000 reads have been assembled together with about 22.000 publicly available ESTs to 31.166 contigs and 29.080 singletons. The new sequences were annotated via sequence comparisons with known plant genes and are made available in form of a Sweetpotato Gene Index, which represents a large part of the sweetpotato transcriptome. This index can be used e.g., for functional genomics, primer and oligonucleotid design for expression studies and SNP search.

The newly obtained sequence information was used to identify microsatellite (SSR) marker sequences. Up to now, PCR-primers have been designed for 94 new SSR loci and 72% of them yielded amplification products with sweetpotato germplasm and mapping parents. A set of ~130 markers remains to be tested.

A DArT microarray has been developed based on 96 sweetpotato accessions including the most important mapping parents and clones of the Composite Genotype Set. The chip performs well and yields close to 2,000 markers for the tested accessions.

We are progressing towards a diploid reference map for *Ipomoea trifida*, a diploid relative of cultivated sweetpotato. An *I. trifida* population was produced from a cross between the accessions M09 and M19. 130 plants of this population have been established in vitro up to now. A diploid map for this population is under construction using SSR, DArT and COS markers. This reference map will help to synthesize genetic information already available from independent hexaploid populations and enables comparative genomics among sweetpotato and other crops.

Collaboration with National Programme partners in Sub-Saharan Africa and Latin America allowed incorporating germplasm from these regions into our genotype set and is stimulating the use of the genetic and genomic resources in ongoing sweetpotato breeding programmes.

Tangible outputs and products (status June 2009)

- Composite Genotype Set
- New sweetpotato ESTs: We planned to obtain 20.000 new ESTs. Due to progress in sequencing technologies, we were able to increase the number of new sequence tags to about 500.000 at lower costs than projected. The gene index resulting from the assembly of these sequence tags together with the publicly available ESTs yielded in total 60.246 sequences including contigs and singletons. The mean length of the contigs and singletons is 803 and 321 bp respectively.
- New markers: A DArT chip with ~ 2000 polymorphic markers is available. Up to now 68 new functional SSR have been identified. This number will be increased to 200 until the end of the project.
- Diploid mapping population: up to now parents and 130 progenies of a cross with two *I. trifida* clones have been established in vitro.
- Diploid reference map: under construction.
- Characterisation of clone collection: planned for Oct. 2009

References

Childs KL, Hamilton JP, Zhu W, Ly E, Cheung F, Wu H, Rabinowicz PD, Town CD, Buell R, Chan AP (2007). The TIGR Plant Transcript Assemblies database. *Nucleic Acids Research* 35(Database issue): D846–D851.

55. G4008.47: Developing genomic resources for pigeonpea using next generation sequencing technologies

August 2008–July 2010

Principal Investigator and Lead Institute
Gregory D May, NCGR; gdm@ncgr.org

PROJECT UPDATE NOT SUBMITTED

56. G7009.01: Natural variation in the transcriptional regulation of drought responses in wheat

January 2009–December 2011

Principal Investigator and Lead Institute
Peter Langridge, ACPFG; peter.langridge@acpfg.com.au
Plant Genomics Centre, Hartley Grove, Urrbrae SA 5064, Australia; (61) 8 83037368; Fax: (61) 8 83037102

Collaborating institutes and scientists

- ICS/CAAS: Jizeng Jia (Co-PI); Xiuying Kong; Guangyao Zhao; Lifeng Gao
- CIMMYT: Matthew Reynolds
- ACPFG: Sergiy Lopato; Serik Eliby

Drought stress can affect plants in many ways and plants have evolved complex response pathways that involve the activation or silencing of many genes and many interactions between regulatory proteins or compounds. Despite this complexity, our knowledge of the regulatory pathways is developing rapidly. Key to the drought response is the activity of transcription factors and associated proteins that lead to the activation of multiple pathways. Many of the regulatory sequences that these transcription factors bind to have been described and additional components, such as phosphorylation of the transcription factors are also known. When the expression level of the genes encoding these regulatory proteins is altered, for example in mutants or in transgenic plants, enhanced, or reduced, drought tolerance can be seen in the plants. This project will build on a well established programme to isolate and evaluate these regulatory proteins to screen for natural variation in expression of regulatory genes shown to moderate the drought tolerance response in wheat. Several genes are already available for screening and more will be identified over the life of this project. A wheat germplasm collection assembled to encompass a wide section of variation in cultivated, land race and wild wheat will form the base for the screen. Tissues collected from field grown plants under both well-watered and drought stress conditions will provide the RNA for evaluation. The screen will give preliminary correlation of expression with drought tolerance. These results will be confirmed using introgression lines and other genetic populations. Where expression correlation is validated the germplasm plus diagnostic marker will be made available to breeders for introgression.

Progress

There are three major areas of activity in this project

1. Isolation and characterisation of drought related transcription factors (TFs) at ACPFG
2. Screening of germplasm collections for variation in expression levels of the TFs
3. Correlation of expression levels of TFs with drought tolerance

1. Research activities and progress at ACPFG

A set of primer sequences based on existing and validated TFs were sent to the colleagues in China to commenced expression screening

Nine transcription factors were isolated from flowering part and early grain of drought tolerant wheat cultivar *T. aestivum* cv.RAC875 subjected to drought stress using *cis*-elements from drought inducible promoters (Lopato et al., 2006). Constructs with constitutive (maize ubiquitin) and drought/cold/salt

inducible (maize Rab17; barley DHN8; wheat (*T. durum*) cor410H; rice Jc189 and Jc285) promoters were generated and used for transformation of either of four cultivars of *T. aestivum*: Bobwhite, Gladius, Drysdale, or Frame by biolistic bombardment (Kovalchuk et al., 2009). The transgenic wheat plants are currently at different stages of regeneration and development in growth cabinets and glasshouse. Several transgenic lines, which were generated before the commencement of this project, are currently being tested for drought tolerance. At the moment four transcription factors (TaDREB2, TaDREB3, TabZipI, and TaERF1) are selected for the work in this project and information about their sequences will be transferred to collaborators.

2. Research activities and progress at ICS/CAAS

Isogenic line (IL) evaluation

100 IL generated from Laizhou953, a released variety, and Am3, a synthetic wheat, were evaluated for their drought tolerance in droughted and irrigated trials. Eight drought tolerance lines were selected from the population. These lines will be used for evaluating the genes for drought tolerance.

RNA extraction

Each of the ten varieties being evaluated for drought tolerance and sensitivity were previously grown under drought stress or irrigation. Flag leaves and developing grains from five plants of each variety at ten days after anthesis were collected and their RNA isolated. The cDNA were synthesised by reverse transcription of RNA and they will be used for Q-PCR in next step.

The first set of primer sequences have been received from Australia and these are currently being used to establish the screening procedures.

References

Lopato S, Bazanova N, Morran S, Milligan AS, Shirley N, Langridge P. Isolation of plant transcription factors using a modified yeast one-hybrid system. *Plant Methods*. 2006 22; 2:3-17

Kovalchuk N, Smith J, Pallotta M, Singh R, Ismagul A, Eliby S, Bazanova N, Milligan AS, Hrmova M, Langridge P, Lopato S. Characterization of the wheat endosperm transfer cell-specific protein TaPR60. *Plant Mol Biol*. 2009 [in press]

57. G7009.02: Mapping and validation of QTLs associated with drought tolerance traits in chickpea

January 2009–December 2011

Principal Investigator and Lead Institute

Pooran M Gaur, ICRISAT; p.gaur@cgiar.org

ICRISAT, Patancheru, Hyderabad 502324, AP, India; Tel: (91) 40-30713356 (O), (91)-9866080915 (Mobile); Fax: (91) 40-30713074/30713075

Collaborating institutes and scientists

- ICRISAT: Rajeev Varshney; L Krishnamurthy; Vincent Vadez, Shailesh Tripathi
- UAS–Bangalore: KP Viswanatha; MS Sheshashaye
- RARS: Veera Jayalakshmi
- ARS–Durgapura: SJ Singh
- RAK–CoAg: Md Yasin

Drought is the most serious constraint to chickpea (*Cicer arietinum* L.) production globally and together with heat stress accounts for over 40% yield losses annually. Over 90% of chickpea crop is grown rainfed on residual soil moisture stored during the previous rainy season and the crop often experiences drought at the critical stage of pod filling and seed development.

The grain yield under drought environments is the product of Transpiration (T), Transpiration Efficiency (TE) and Harvest index (HI). The root system that can extract water from deeper soils can increase T and contributes to improving the total biomass productivity and also the HI. A measure of carbon isotope discrimination (δ13C) gives a good estimation of TE as these are positively correlated.

This project builds on Tropical Legume I project, where efforts are being made to map QTLs for root traits. In this project, we propose to map and validate QTLs affecting all three components, T, TE and HI, of the grain yield under drought environments. The root traits will be used for T, carbon discrimination factor for TE and biological and grain yield for HI.

The major objectives of this project are as follows:

- Mapping QTLs for drought tolerance traits (root characteristics, δ13C and HI) in chickpea
- Validating of molecular markers for major QTLs controlling root traits, δ13C and HI in chickpea
- Identifying parents and initiating marker-assisted breeding for drought tolerance in chickpea

The project officially started in January 2009, but we started work early (Oct 2008) to make use of the chickpea crop season (Oct-March). A brief account of the research progress made so far is given below:

Phenotyping of ICC 4958 x ICC 1882 RILs for yield, biomass and HI: The RILs (n=264) from the cross ICC 4958 x ICC 1882 along with parents and two checks were evaluated for phenology, seed yield and biomass at four locations (ICRISAT-Patancheru, RARS-Nandyal, RAKCA-Sehore and ARS-Durgapura) during the crop season 2008/09. The ranges and means obtained for different traits are presented in Table-1. The data is being analysed for GxE interactions.

Table1. Phenotypic variation in ICC 4958 x ICC 1882 RILs for various traits

	Days to 50% flowering				Days to maturity				Harvest Index (%)			
Location	P*	N	S	D	P	N	S	D	P	N	S	D
Range	46-58	34-63	62-69	55-65	99-111	75-98	113-119	124-135	35.1-60.8	35.9-51.4	34.9-71.3	33-49.1
Mean	52	51	65	60	104	86	116	129	47.1	43.4	54.9	41.5
SEM±	1.24	1.28	0.54	-	1.48	1.16	0.48	-	3.07	1.32	4.06	3.74
CV%	3.41	3.5	1.18	-	2.01	1.9	0.58	-	9.22	4.3	10.5	12.7
LSD (5%)	3.5	3.6	1.5	-	4.1	3.2	1.3	-	8.5	3.7	11.3	10.4

*P=Patancheru, N=Nandyal, S=Sehore, D=Durgapura

Phenotyping of ICC 4958 x ICC 1882 RILs for δ13C: Leaf samples of 264 RILs of ICC 4958 x ICC 1882 and the parents were collected from three locations (Durgapura, Sehore and Patancheru) and sent to University of Agricultural Sciences, Bangalore, for analysis of carbon isotope discrimination (δ13C). The data on δ13C will be available by October 2009.

Genotyping of ICC 4958 x ICC 1882 RILs: The RILs of ICCV 4958 × ICC 1882 are being genotyped using SSR markers and phenotyped for root traits under TL-I project. These will be used for detection of QTLs for root traits, carbon isotope discrimination and harvest index.

58. G7009.04: Development and evaluation of drought-adapted sorghum germplasm for Africa and Australia

January 2009–December 2011

Principal Investigators and Lead Institutes

David Jordan, Hermitage Research Station, QPIF; david.r.jordan@deedi.qld.gov.au
604 Yangan Road, Warwick, Queensland 4370, Australia; (61) 7 46603622; Fax: (61) 7 46603600

Andrew Borrell, Hermitage Research Stn, QPIF; andrew.borrell@deedi.qld.gov.au
604 Yangan Road, Warwick, Queensland 4370, Australia; (61) 7 46603640; Fax: (61) 7 46603600

Collaborating institutes and scientists

- IER: Sidi Bekaye Coulibaly; Niaba Teme; Mamoutou Kouressy
- Agropolis–CIRAD, Mali: Michel Vaksmann

Context: The aim of this project is to improve drought adaptation and productivity in Malian sorghum by integrating three complementary activities. 1. Evaluating the stay-green drought resistance mechanism in plant architectures and genetic backgrounds appropriate to Mali. 2. Developing sorghum germplasm populations enriched for stay-green genes that also carry genes for adaptation to cropping environments in Mali. 3. Carrying out training activities for African sorghum researchers in drought physiology and selection for drought adaptation in sorghum.

Findings and implications

Activity 1. **Evaluating stay-green (SG)**

The SG trait has been shown to enhance grain yield under post-anthesis drought in elite three dwarf sorghum germplasm. The potential trade-offs involved in deploying SG in the 1- or 2-dwarf photoperiod sensitive sorghums commonly grown in sub-Saharan Africa are unknown. In order to provide evidence that enrichment of Malian germplasm with SG from elite QPIF germplasm will be valuable, a pilot experiment was conducted in Australia to determine the impact of variation in height on the expression of the SG trait. A factorial experiment was conducted consisting of two height isoline pairs (2 vs 3 dwarf) of two elite QPIF lines contrasting in SG (R931945-2-2, SG; R955343-1, senescent) grown under two irrigation regimes (irrigated vs rainout shelter) at two population densities (high and low).

The isogenic pairs clearly differed in their SG response (Fig. 1a), as expected. Leaf greenness, as measured by SPAD, was equivalent between pairs under irrigation for FL-3 (3rd leaf down from the flag). However under stress, the senescent pair (R955343-1) exhibited significantly lower SPAD for FL-3 than the SG pair (R931945-2-2). Differences in SG response were also observed between tall and short plants. Under low stress (irrigated or low density), short plants had greater SPAD than tall plants, probably due to the dilution effect of N in tall plants. Under stress, no differences were observed between tall and short in the SG pair, but SPAD was higher in the tall than short in the senescent pair. Overall, tall plants displayed an equivalent or better SG phenotype than short plants under stress conditions. Tall plants of the SG pair yielded better than short plants of the same pair and, under stress, the SG pair generally yielded better than the senescent pair (Fig. 1b).

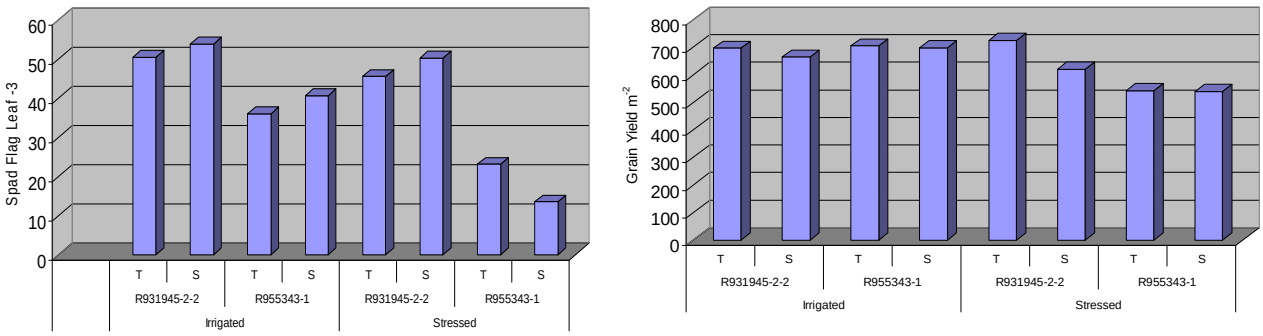


Fig. 1a (left). Leaf greenness (SPAD) of the 3rd leaf down from the flag (FL-3) in tall (T) and short (S) isolines of SG (R931945-2-2) and senescent (R955343-1) sorghum lines grown under irrigated and stressed conditions. Fig. 1b (right). Grain yield (g/m^2) in tall (T) and short (S) isolines of SG (R931945-2-2) and senescent (R955343-1) sorghum lines grown under irrigated and stressed conditions.

Activity 2. **Developing sorghum germplasm populations enriched for stay-green genes that also carry genes for adaptation to cropping environments in Mali**

Develop germplasm adapted to Mali and enriched for stay-green (SG)

Four F2 populations previously produced by QPIF were grown in the field in Australia under long days. The populations were the result of crosses between a genetic male sterile version of R931945-2-2 (SG) and four different lines from Mali. A number of photoperiod sensitive plants were selected from each population, transferred to the glasshouse, and screened with markers linked to five stay-green QTL. Stay-green donor plants were identified and crossed back to the respective Malian parent lines. These crosses were made with the objective of implementing marker assisted selection for the stay-green QTL in the BC1F2 generation in 2010.

F1 hybrids were made between two Malian lines and elite QPIF cytoplasmic male sterile lines with contrasting levels of stay-green. These hybrids will be used in experiments to follow on from the experiment described in Activity 1 (see above) to further evaluate the impact of height on the expression of stay-green in tall photoperiod sensitive sorghum typical of Sub-Saharan Africa.

Develop photoperiod sensitive tall mapping populations for stay-green & grain yield

Approximately 2000 plants from each of 9 F2 populations were grown at Ayr in tropical Queensland. The F2 populations were produced by crossing a genetic male sterile version of R931945-2-2 with nine unique accessions from Mali (PI585749, PI585750, PI585805, PI585818, PI609084, PI609108, PI609114, PI609278 & PI609321). Approximately 500 lines from each population with similar height to the Malian parent were self pollinated. These families will be evaluated for photoperiod sensitivity in the spring of 2010.

Subprogramme 3: Trait capture for crop improvement

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

July 2005–June 2007; no-cost extension to June 2008

Principal Investigator and Lead Institute
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 institutes and scientists
- CIMMYT–CAAS: Jiankang Wang
- UoQ: Mark Dieters; Graeme Hammer
- CSIRO: David Bonnett; Greg Rebetzke
- Agropolis–INRA: Francois Tardieu; Claude Welcker

Background
The aim of this project was 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 finished in June 2008, with 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 2007/08 was vast improvement in the 2nd 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/. 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

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 (see Chenu et al 2009)
- 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.

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 and Bonnett DG (2008). Designing crossing and selection strategies to combine diagnostic markers and quantitative traits. Int. Wheat Genetics Symp. Brisbane Aug 22-25 2008.

Chenu K, Chapman SC, Tardieu F, McLean G, Welcker C and Hammer GL (2009) Revealing the Yield Impacts of Organ-Level Quantitative Trait Loci Associated with Drought Response in Maize - A Gene-to-Phenotype Modeling Approach. *Genetics* (accepted with revision)

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.

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

August 2007–July 2009; no-cost extension to July 2010

Principal Investigator and Lead Institutes
Chalermpol Phumichai, BIOTEC
National Centre for Genetic Engineering and Biotechnology (BIOTEC), Department of Agronomy Faculty of Agriculture Kasetsart University; chalermpol.ph@gmail.com, agrcpp@ku.ac.th

Julapark Chunwongse, Kasetsart University, Thailand
Department of Horticulture, Kasetsart University, Kamphaengsaen campus, Nakhon Pathom 73140; julapark.c@ku.ac.th

Collaborating institutes and scientists

- NCSRC: Sansern Jampatong
- NSFCRC: Pichet Grudloyma

Context: A few lines to explain the context of the project

Association studies is an advantages over linkage mapping in traditional biparental populations which individual genes may be responsible for the expression of the target trait and have the potential to evaluate and characterise a large number of alleles per locus (Buckler and Thornsberry, 2002; Flint-Garcia et al, 2003). Therefore, association mapping based on LD may allow identification of the actual genes that are underlying QTLs. Informative polymorphism identified in gene haplotypes and associated with a contrasting phenotype can be converted into DNA markers for use in MAS. The application of association mapping to plant breeding appears to be a promising approach to overcome the limitations of conventional linkage mapping.

The objective of this research was investigate genetic diversity within germplasm of maize inbred lines in Thailand and investigate positions of markers that show high association with trait values that were compared with those of loci known to confer resistance to downy mildew.

Findings and implications

Analysis of variance mean squares values for DMR score of 60 maize inbred lines grown at two locations. The analysis of variance show significant variation among entries (Table 1). There was a significant location effect as well as a significant entry by location effect. The analysis of variance results were used to measure broad sense heritability. The value was 0.50.

Table 1 Analysis of variance mean squares values for percentage of downy mildew at two locations

Source of Variation	df	MS
Location	1	11878.2**
Location*Rep	4	154.899*
Entry	59	7045.59***
Location*Entry	59	227.203***
Residual	236	51.9696

*$P \leq 0.05$, **$P \leq 0.01$, ***$P \leq 0.001$

Sixty maize inbred lines were genotyped for 48 SSR markers on all maize chromosomes. All of 48 markers produced a total of 489 alleles among 60 entries. The alleles varied from 2 to 29 with an average of 10 alleles per locus. A dendrogram was generating using the UPGMA algorithm with GD matrix that all of the entries could be group in to three clusters.

Population structure was estimated using the model based approach as implemented in the software programme STRUCTURE (Pritchard and Wen 2007). The number of subpopulation (K) was identified likelihood value. However, it was difficult to determine the optimal number of subgroups, since the posterior or probabilities for the number of clusters increased steadily.

Association analysis identified marker trait association (P-adj < 0.05). Three significant SSR/trait associations were detected with the Q GLM model over 2 locations. These marker loci could explain 38.25–70.93% of the total variation.

References

Buckler ES and Thornsberry JM (2002) Plant molecular diversity and applications to genomics. *Curr. Opin. Plant Biol.* 5, 107–111.

Flint-Garcia SA, Thornsberry JM and Buckler ES (2003) Structure of linkage disequilibrium in plants. Ann. Rev. *Plant Biol.* 54, 357–374.

Jumpatong C, Jampatong S, Balla C, Grudloyma P, Jompuk C and 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.

Pritchard JK, Wen X and Falush D (2007) Documentation for structure software: Version 2.2. Department of Human Genetics, University of Chicago; Department of Statistics, University of Oxford. Available at http://pritch.bsd.uchicago.edu/software

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

January 2007–December 2009

Principal Investigator and Lead Institute

Abbad Andaloussi Fouad, INRA–Morocco; abbad@inra.org.ma; abbadandaloussi@yahoo.fr
Institut National de la Recherche Agronomique, Division Scientifique, Département Protection des Plantes, Postal address 415 RB, Rabat, Morocco. Phone: 212 537 72 72 74, Fax: 212 537 70 64 99.

Collaborating institutes and scientists

- INRA–Morocco: Nsarellah Nasserlehaq; Jlibene Mohammed; Lhaloui Saadia; Labhilili Mustapha; Saidi Seddik
- ICARDA: Sripada M Udupa
- UdB: Roberto Tuberosa
- CIMMYT: Susanne Dreisigacker
- Cornell: Mark E Sorrells
- UoMi: 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

The germplasm screening (during 2008) 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). The reference barley set (99 lines) was screened (during 2009) for diseases (net bloch and powdery mildew, under controlled conditions) and pests (Barley Stem Gall Midge under field and Hessian fly under controlled conditions). The screening resulted in identification of 27, 76, 24 and 47 accessions were resistant to net bloch, powdery meldew, barley stem gal midge and the Hessian fly respectively. The Mediterranean durum wheat collection (96 lines) was screened (during 2009) for variation in days to collapse under salinity (saline water, 18g of NaCl/l). The preliminary result revealed that variation in days to collapse under salinity ranged from 4 to 28 days, with more than twenty entries showed number of days that was greater than 20. The useful genetic diversity observed during current reporting period will incorporated into the crossing programme during coming cropping season.

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 various quality traits, high protein content, resistance to leaf rust, Hessian fly, cooking quality and tolerance to drought. These populations are in their F3 and F4 for durum wheat, and in F2 stage for 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 nurseries and trials of the bread and durum wheat breeding programmes were evaluated in several experiment stations and important entries were selected for genotyping work and marker validation. The crossing programmes for MAS are at various stages, aiming mainly on pyramiding genes of agronomic interest. Doubled haploid technique is being integrated into these MAS schemes, to speed up the varietal developmental process. The ongoing MAS scheme intended to incorporate high grain protein content and yellow rust resistance (GPC/Yr) dwarfing genes and introduce rye chromosome segment 1R (the short arm of rye chromosome 1R, because it provides resistance to insects, diseases and reported improvements in yield potential and water-use efficiency) into a cultivar. 'Yecora Rojo" (GPC/Yr) was crossed with 'Tilila' (1RS) during spring 2008. The F1s were grown during summer 2008 and were backcrossed with 'Tilila'. During 2009, around 200 BC1F1 were screened using markers for GPC/Yr and 33 F1BC1 plants were selected to produce haploids. Upon production of haploids, approximately 100 haploids were screened using three markers (1RS, GPC/Yr and Glu 3D 22) and selected several haploid plants for doubled haploid production. The rest of the selected BC1F1s were used for producing BC2F1, by backcrossing with 'Tilila'. Seventy three BC2F1 were screened using markers for GCP/Yr and selected 17 plants for further backcrossing and haploid/DH production. Several other crosses were performed and carried forward for marker validation and MAS for drought, cereal cyst nematode resistance and leaf rust.

4. Generation of new markers including candidate gene/allele-based molecular markers (functional markers) for enhancing the breeding strategies
Field evaluation of the Simeto/levante cross has been completed for the second season in two stations in Morocco with both agro-morphological ad physiological data collected, (including radiation and SPAD) while profiling will be done in Bologna Italy, during later part of this year.

5. Mining the novel genes and alleles conferring tolerance to stresses and improving end use quality from the germplasm collection of Morocco Primers were designed to amplify the genes (Srg6, EDS1 and Coi1 genes) and genetic polymorphisms (genic-microsatellites, SNPs, and STS marker approach) in cereals using PCR techniques. A set of genic markers (GPC, glutine protein, waxy and grain texture) were tested for polymorphism among the bread and durum varieties of Morocco, in order to study allelic variation. The result revealed narrow genetic diversity among the varieties.

6. Empower Moroccan researchers, young scientists and students to implement new tools of genomics in their breeding programmes
So for, five undergraduate students completed their internship training, and four master degree students from Moroccan Universities completed their thesis.

62. 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 and Lead Institute
Rui-Lian Jing, Institute of Crop Sciences, CAAS; jingrl@caas.net.cn
No 12 Zhongguancun South Street, Haidian District, Beijing 100081,China; Tel & Fax: ++86 (0)10 62186706

Collaborating institutes and scientists
- NU: Xing Xu
- NWSUAF: Hui-Min Xie
- Institute of Crop Sciences, SAAS: Mei-Rong Sun

- LAAS: Can-Jun Zhang
- HAAS: Xiu-Min Chen

1. Selection of drought tolerant individuals from DT introgression lines

A Total of 156 accessions with improved drought tolerance (DT) have been identified from BC_3F_2 to BC_3F_5 introgression lines (ILs) by phenotyping in multi-location under rainfed and well watered conditions in Beijing, Shanxi, Shaanxi, Henan and Hebei Provinces, respectively. Among them, 61 lines are identified from 11 crosses of BC_3F_4 or BC_3F_5 ILs, 95 individuals selected from BC_3F_2 to BC_3F_4 ILs. The recurrent parents (female parent) Jinmai 47, Jinmai 54, Lumai 14 and Yumai 18 are the elite Chinese cultivars. The donor parents of ILs, Xiaoyan 54, Zhongyou 9507, Gaoyou 504, Shanyou 225, Dali 1, Jingnong 79-15, Yunhan 2028 etc. are Chinese cultivars with drought tolerance and favorable agronomic traits, Salgemma and Pandas are European common wheat cultivars with drought tolerance and good quality.

Table 1 Drought tolerant individuals selected from DT introgression lines with the elite Chinese wheat genetic backgrounds

Recipient	Donor parent	IL generation	Individual selected
Jinmai 47	Chinese cultivar	BC_3F_2 ~ BC_3F_5	52
	European cultivar	BC_3F_4	14
Jinmai 54	Chinese cultivar	BC_3F_2 ~ BC_3F_5	35
	European cultivar	BC_3F_4	12
Lumai 14	Chinese cultivar	BC_3F_3 ~ BC_3F_5	23
Yumai 18	Chinese cultivar	BC_3F_2 ~ BC_3F_4	20

2. Marker assisted selection for DT introgression lines

Seven introgression line populations derived from recurrent parent Jinmai 47 were genotyped with 50 ~ 137 DT QTL candidate SSR markers. Total of 23 DT ILs have been selected based on the genotype and phenotype evaluations. A high yield DT line, JL28 is recommended to be pretested in Shaanxi provincial yield trial this year.

Table 2 Marker assisted selection for DT introgression lines derived from recurrent parent Jinmai 47

Donor parent	IL generation	Introgression line	No. of polymorphic SSR marker
Lumai 14	BC_3F_4	148	137
Salgemma	BC_3F_4	150	116
Xiaoyan 54	BC_3F_4	25 selected lines	52
Zhongyou 9507	BC_3F_4	25 selected lines	55
Gaoyou 504	BC_3F_4	25 selected lines	57
Shaanyou 225	BC_3F_4	25 selected lines	50
Pandas	BC_3F_4	25 selected lines	54

The challenge for MAS is that majority of QTL involved significant additive and epistatic effects with interactions of QTL and environments, QTL and genetic backgrounds, which badly impact the MAS effect. In the next year, we are going to select ~60 DT ILs by MAS combining DT phenotype in the fields of multi-location, continuously exchange and utilise the information and techniques among breeders and geneticists for improving the capacity of wheat modern breeding in China.

References
Yang DL, Jing RL, Chang XP and W Li (2007). Identification of quantitative trait loci and environmental interactions for accumulation and remobilization of water- soluble carbohydrates in wheat (*Triticum aestivum* L.) stems. *Genetics*. 176: 571-584. (http://www.genetics.org/cgi/content/abstract/176/1/571)
Yang DL, Jing RL, Chang XP and W Li (2007). QTL Mapping for chlorophyll fluorescence and associated traits in wheat (Triticum aestivum L.). *Journal of Integrative Plant Biology*. 49: 646-654. (http://www3.interscience.wiley.com/cgi-bin/fulltext/118492091/PDFSTART)
Zhang KJ, Tian L, Zhao and S Wang (2008). Mapping QTLs with epistatic effects and QTL × environment interactions for plant height using a doubled haploid population in cultivated wheat. *J Gen Genomics*. 35: 119-127. (http://www.jgenetgenomics.org/qikan/epaper/zhaiyao.asp?bsid=13385)

63. G4007.07: Marker−assisted selection for Sweetpotato Virus Disease (SPVD) resistance in sweetpotato germplasm and breeding populations

August 2007−July 2009

Principal Investigator and Lead Institute
WJ Grüneberg, CIP; w.gruneberg@cgiar.org

PROJECT UPDATE NOT SUBMITTED

64. G4007.08 Integration of genomic tools with conventional screening for developing NERICA rice cultivars for West Africa

September 2007–September 2009

Principal Investigator and Lead Institute
Marie Noelle Ndjiondjop, WARDA; m.ndjiondjop@cgiar.org
01 BP 2031 Cotonou, Benin; Tel: 00 299 21 35 01 88; Fax: 00 229 21 35 05 56

Collaborating Institutes and scientists

- WARDA: Manneh baboucarr; Drame K. Nani; Sere Yacouba
- Agropolis–IRD: A Ghesquiere; Mathias Lorieux; V Verdier
- IER: Fousseyni Cisse

Background of the project
The NERICAs (New Rice for Africa) are interspecific hybrids that bridge two cultivated rice species (*Oryza 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. Morpho-agronomic and molecular characterisation of* Oryza glaberrima *germplasm from Mali
In order to exploit the important agronomic traits that make *Oryza glaberrima* Steud suitable for harsh growing conditions, 79 *Riz Africain du Mali* (RAM) including 4 checks of known cultivated rice species (*O. sativa* and *O. glaberrima*) were used for both field evaluation in 2007 and 2008 and assessment of molecular diversity. The molecular characterisation showed all the 37 SSR markers used amplified and were polymorphic. The average heterozygosity of the accessions were relatively higher (0.18) than in

some earlier studies. The population from Mopti and Gao had a higher genetic diversity than other populations. Clustering of individuals showed no clear pattern of grouping based on either location or growing condition; this is due to the higher gene flow (Nm= 7.83) that may occur between different growing condition within one location. Moreover, populations from different locations were highly differentiated (fst= 0.45) as compared to populations from different growing conditions (fst= 0.03). The dendrogram based on populations from different growing conditions revealed more similarity among populations with relatively similar growing conditions. Three accessions (RAM116, RAM100 and RAM122), performed better under drought with a yield advantage of 40 %, 16 % and 6.7 %, respectively over the best performing check.

2. Progress on research activities at IRD

2a. Screening of resistance to RYMV and identification of a second resistance gene.

Screening of *O. glaberrima* for RYMV accessions (including parents of *iBridges* project) was focused to reconfirm high resistance pattern in using the more aggressive strain S2 BF1. 29 accessions were reconfirmed and the MIF4G domain of RYMV1 resistance gene was sequenced in all those accessions. A third resistance *rymv1-5* allele was identified and the presence of the susceptibility allele in 15 resistant accessions suggested a new type of resistance. Allellism tests evidenced a second major recessive resistance gene, *RYMV2* in one of those accessions, which mapped on chromosome 1 and exactly in the interval of a resistance QTL previously identified in the IR64 x Azucena RI population. Markers specific of the different *RYMV1* resistance alleles were developed for marker-assisted selection of resistant genotypes for disease management.

2b. Mapping of resistance genes/QTLs to* Xanthomonas Oryzae pv. Oryzae *(Xoo)

The (IR64 x Azucena) RI population was used to map 4 strains of *Xoo* representative of the diversity of the strains in West Africa, namely, BAI3 and NAI8 (race 1) from Burkina-Faso and Niger respectively, BAI4 (race 2) from Burkina-Faso, MAI1 (race 3) from Mali. two Philippine strains PXO61 (race 1) and PXO86 (race 2) were used as control since African and Asian strains of *Xoo* are very different. A QTL common to the 4 African races was detected on chromosome 11 where several major resistance genes have been previously evidenced in using Philippines strains. Interestingly an original QTL specific to African strains and explaining 36% of the phenotypic variance has been found on chr. 7. Allelism test have been developed to confirm if Xa4 and xa5 can be efficient against African strains of *Xoo*.

3. Progress on research activities at IER

A set of 100 promising drought tolerance *O. glaberrima* selected last year is being evaluated 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 analysed using SAS (version 9.1).

Outputs

- 29 *O. glaberrima* accessions confirmed to have high resistance to Rymv
- Specific markers to reveal the 3 different *RYMV1* resistance alleles in *O. glaberrima*
- Identification of a new resistance gene *RYMV2* with initial mapping and markers on chromosome 1
- Identification of a new major resistance QTL to the African strains of *Xoo*`
- 3 *O. glaberrima* (RAM) performed better under drought with a yield advantage of 40 %, 16 % and 6.7 %, respectively over the best performing check (Moroberekan).

Perspective

(1) Data collected from drought and blast screening of *O. glaberrima* and interspecific lines in Benin and Mali will be compile and analyse
(2) Promising interspecific lines evaluated and the best interspecific lines made available to breeders and geneticists for validation as NERICA
(3) Publications on the outcome of this project.

References
D. Thiémélé et al., A wealth of resistant phenotypes to Rice yellow mottle virus (RYMV) is observed in the African cultivated rice *Oryza*. glaberrima Steud. and involves a second major recessive gene, RYMV2. (Submitted to *Theor. Appl. Genet*).

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

December 2007–November 2010

Principal Investigator and Lead Institute
Márta Molnár-Láng, ARI–HAS; molnarm@mail.mgki.hu
2462 Martonvásár, Brunszvik u. 2., Hungary; Tel: 36-22-569-505; Fax: 36-22-569-576

Collaborating institutes and scientists
- CIMMYT: Maria Zaharieva
- Institute of Crop Sciences, CAAS: Ruilian Jing
- DPSPP–EKC: Sándor Dulai
- ARI–HAS: Éva Darkó

1. Research activities and progresses at ARI HAS
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, 7H Mv9kr1/Igri and 4H, 6H, 7H Asakaze komugi/Manas (Akom/Manas) disomic addition lines, the 4H(4D) substitution line as well as the 3HS.3BL, 6BS.6BL-4H, 2DS.2DL-1HS and the 7DL.7DS-5HS translocation lines were sown in the field of Martonvásár in 2008. Harvesting of seeds will take place in July 2009.

Activity 2. The presence of the barley chromosome pair in wheat background was demonstrated using genomic *in situ* hybridisation (GISH) and sequential fluorescence *in situ* hybridisation (FISH) with the help of repetitive DNA probes pTa71, Afa family, HvT01 and GAA sequences in the 2H, 4H, 6H, 7H Akom/Manas disomic addition lines. The cytogenetic identification was confirmed with molecular markers. Four 3H disomic addition plants have been selected from among the progenies of the 3H Akom/Manas monosomic plant.

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

Activity 1. A wheat-barley centric fusion containing the long arm of 4H was produced from the 4H(4D) substitution line after crossing it with the CS ph mutant. Dry seeds of the 4H(4D) substitution line were irradiated with 60Co -rays at dosage of 100 Gy. Several wheat-barley translocations (centric fusions, terminal translocations, reciprocal translocations) were detected by GISH in the irradiated seeds.

Activity 2. DNA samples isolated from the 4H, 6H, 7H Akom/Manas wheat/barley addition lines, the parental lines and the 4H(4D) substitution line were sent to CIMMYT, Mexico beside the DNA samples sent earlier.

Physical mapping of two wheat/barley translocation lines were carried out in Martonvásár. In the 7DL.7DS-5HS Mv9kr1/Igri translocation the elimination of the 7DS terminal region was proved by three (Xbarc184, Xwmc506, Xgdm130) of the twenty-four tested SSR markers. The breakpoint of the

5HS.7DS translocation was considered to be closer to the telomere than the breakpoint of known deletion lines, which provides a new physical landmark for future deletion mapping studies.

A 4BS.7HL translocation was detected in the progenies of the Akom × Manas hybrids by sequential GISH, FISH and SSR markers. The presence of 4BS and the absence of 4BL was proved by six wheat SSR markers. The 7HL chromosome arm was identified with six barley SSR markers. The absence of two 7HL-specific markers from this translocation line (EBmac0785 and Ebmac0827), revealed the absence of the 7HL centromeric region. The translocation breakpoint could be placed between markers Ebmac0827 and GBM1102

Objective 3. Drought study of the wheat/barley addition, substitution and translocation lines
Activity 1. 2H, 3H, 4H Mv9kr1/Igri, 4H Akom/Manas, 2H Akomugi/Betzes, 6H Mv9kr1/Betzes wheat/barley addition lines; 4H(4D) Mv9 kr1 wheat/barley substitution line; 3HS.3BL, 2DS.2DL-1HS, 6BS.6BL-4HL Mv9kr1/Betzes, 7DL.7DS-5HS Mv9kr1/Igri wheat/barley translocation lines and the parental lines were sown in the field under the rainshelter and in irrigated conditions as control in three replications in Martonvásár in 2008. The experiment was harvested in the first week of July in 2009. Evaluation of heading time, plant height, tillering, fertility, thousand grain weight etc. will be executed in the near future. The same lines, except two, were sent to the Chinese collaborator in 2008, and these lines were sown under a rain shelter in China. The experiment was harvested in the end of June and evaluation of the results will be carried out this summer in 2009.

Four wheat/barley translocation, four addition, one substitution and the parental lines were sown in the field experiment at Pannon University, Georgikon Faculty in Keszthely in 2008. Each genotype was sown in a 15 m long row. The half length of the rows was covered with a plastic folia in April to protect plants from rain. Plant heigth, root/shoot ratio, leaf water potential, ear length, thousand grain weight, number of kernels and grain yield will be determined.

A new set of wheat/barley disomic addition lines (2H, 3H, 4H, 6H, 7H) was produced with the six-rowed winter barley cultivar Manas, which showed good Al-tolerance in the experiment carried out in the previous year. It is planned to study the Al-tolerance of the Akom/Manas addition lines when the required amount of seeds will be available from the five lines.

The drought tolerance test under rain shelter will be repeated in the third year again in Martonvasar and in China simultaneously. Conclusions about the drought tolerance of the different lines can be made after several years observations.

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

October 2008–June 2009

Principal Investigator and Lead Institute

Professor Ervin Balazs, ARI–HAS; balazs@mail.mgki.hu
Department of Applied Genomics, ARI–HAS, 2462 Martonvasar, Brunszvik str 2.; Hungary; Tel.: +3622569500 ; Fax.: +3622569514

Collaborating scientists and institutes

- ARI–HAS: Vilmos Soos; Angela Juhasz
- Research Centre for Plant Growth and Development, 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 (butenolide) 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

To determine the smoke- and butenolide-responsive genes, we treated kernels of Mv255 maize strain with smoke and butenolide solutions for 1.5, 3, 6, 9, 12 and 24 h. The changes in the total transcriptome of the embryos were recorded using maize oligonucleotide microarray slides. Initial experiments were carried out to determine the effect of smoke-water and butenolide on the growth parameters of 5-day-old seedlings. Smoke-water yielded significantly longer shoots and roots compared to controls, although the germination rate was not affected. Throughout the course of the experiment, or a part of the whole experiment only a narrow subset of genes were affected by smoke treatment. A sulforedoxin-like protein gene and a leucine rich repeat (LRR) gene, were upregulated, while the transcript abundance of the ABA signalling negative regulator calcineurin 9 like gene and a tetratricopeptide repeat containing gene were sharply declined. In smoke-treated seeds, the most obvious changes were observed in the expression of the ubiquitin activating enzyme 1 (UBE1) which was upregulated at all time points. Known as the E1 enzyme, UBE1 catalyses the first step in the ubiquitination reaction that targets proteins for degradation via the proteasome. To prove that smoke exposure has an effect on the ubiquitination process, we blotted the protein samples extracted from maize embryos after 3, 4.5, 6, and 7.5 h smoke-treatment onto PVDF membrane and treated it with polyubiquitin antibody. Comparing these samples with controls, and samples treated similarly with butenolide, it was apparent that smoke-treatment, and not butenolide, enhanced the ubiquitination of the proteins dramatically after 6 h suggesting that the smoke treatment resulted in accelerated ubiquitination.

Treatment of maize kernels with the 10^{-7} M solution of butenolide resulted in a very similar frequency distribution of shoot/root size as observed in smoke-treated kernels. However, butenolide treatment yielded a completely different gene expression pattern in comparison to smoke-treated samples. The overlap between the two gene set was only 2%. The butenolide-responsive gene list is restricted to only few genes. A senescence-associated protein-related gene of unknown function was upregulated except at 9h, when a sharp decline in the expression was observed. The most notable gene, which was upregulated ubiquitously during the course of the experiment is an aquaporin. We applied silver nitrate, a potent aquaporin inhibitor, on maize seedlings which resulted in a reduction of the growth parameters of the seedlings. Treatment of the seedlings with a combination of butenolide and silver nitrate showed an alleviation of the adverse effect of the silver nitrate treatment, whereas simultaneous treatment with both smoke-water and silver nitrate show no such reduction in the effect of silver nitrate inhibition. This effect of the butenolide in combination with silver nitrate was demonstrated by the frequency distribution of the seedling shoot/root size which was not significantly different from the butenolide-treated control plants. Analysis of the microarray data obtained from comparison of the butenolide- and smoke-treated samples showed that the master genes aquaporin and SAM-dependent methyltransferase, and the senescence-associated protein-related gene proved to be butenolide responsive, were downregulated in smoke treated plants. In addition to the butenolide used in this study, three other active butenolides are known to be present in the smoke. Surprisingly, the activity of these compounds seems to be different in terms of the relative expression level on selected hormone related genes. It was previously reported that smoke-water has a “dual regulatory” effect on germination, since high concentrations of smoke-water were shown to inhibit germination, whereas lower concentrations had a promontory effect. These findings, together with our present results, clearly indicate that the array of compounds present in the smoke results in distinctly different effects on the gene expression in germinating maize kernel in comparison to that observed with the treatment of butenolide alone.

Next steps, expected results

The microarray data will soon be placed in the ZEAMAGE and GEO database. Transgenic lines of fire prone tobacco species harboring the candidate genes and their promoters will be applied for functionalisation tests and promoter-based reporter assays.

Phytotron and field tests on smoke-water, butenolide and dry smoke-treated kernels of maize will be launched soon in order to facilitate the transfer of knowledge on smoke to the agricultural practice.

67. G4008.10: Assessment of the breeding value of superior haplotypes for *Alt$_{SB}$*, a major *Al* tolerance gene in sorghum: linking upstream genomics to acid soil breeding in Niger and Mali (ALTFIELD)

January 2008–December 2010 (Contract effective February 2009)

Principal Investigator and Lead Institute

Robert Schaffert. EMBRAPA Maize and Sorghum; schaffer@cnpms.embrapa.br; reschaffert@hotmail.com
Caixa Postal 151, 357001-970 Sete Lagoas, MG, Brazil; E-mail:; tel: +55-31-3027-1176, cel: +55-31-9986-1698, fax:+55-31-3027-1188)

Collaborating institutes 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; Flavia C Santos, João Herbert M Viana

Project Scope: Aluminium (Al) toxicity is a major agricultural constraint on acid soils, which comprise over 50% of the world's potentially arable lands, particularly jeopardising food security in the poorest regions of the globe. A major sorghum Al tolerance gene, *Alt$_{SB}$*, 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. The ALTSORGHUM project (competitive call) is using association genetics to identify superior haplotypes and develop haplotype-specific markers and identify new Al tolerance genes in sorghum. This project (ALTFIELD) aims at establishing a connection between the outputs of ALTSORGHUM and sorghum breeding programmes in Niger and Mali, ensuring that selected cultivars will be validated in specifically developed phenotyping sites and effectively used to attain higher and more stable yields in farmer's field on African soils where aluminium toxicity is a crop production limiting factor.

The ALTSORGHUM project has identified several haplotypes of *Alt$_{SB}$* from several different sources with varying degrees of tolerance to Al toxicity. Our strategy is to screen a few hundred landraces from Mali and Niger for tolerance to aluminium toxicity. Those encountered with tolerance to Al toxicity can be recommended and used immediately by the farmer. Using Marker Assisted Backcrossing (MAB) protocols we plan to introgress the best aluminium tolerance haplotypes into the best landraces and improved breeding improvement genotypes. In this process, we will establish aluminium 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 increase the availability of this nutrient.

Project activities: This project was to be initiated in January 2008, but funding was only transferred to Embrapa in the first trimester of 2009. The Embrapa Legal Department has several requirements that are different the GCP standard contracts. This delay has limited Embrapa´s onsite participation at INRAN and ICRISAT/Mali and ICRISAT/ Niger. However, these problems have been resolved and the project will proceed as planned. With the exception of site visits by Embrapa, the other components began in early 2008. Germplasm exchange has been slow initially, but all exchanges proposed are in the process of being executed. Germplasm lists of landraces from both INRAN and ICRISAT Mali have been prepared and germplasm exchange initiated. The INRAN germplasm is currently in plant

quarantine in Brasilia with expectation for release to our project within the next 30 to 60 days. The seed from ICRISAT Mali with appropriate import permit was sent by courier directly to Brasilia as required, but was returned to the sender (ICRISAT in Mali) for reasons unknown. The seed has been redispatched (June 2009) to plant quarantine in Brazil by airmail. Genetic resources from Embrapa, including NILs, elite haplotypes, and sources of elite cytoplasmic male-sterility and elite fertility restorer lines with Alt_{SB} *have* been increased for exchange, all official documents have been prepared and signed and the seed has been being sent by courier service during the month of July 2009.

Project results: Both ICRISAT and INRAN have phenotyped the material from Niger at low pH and corrected pH. The trial at INRAN was lost due to drought, but the same material was phenotyped in the field with supplemental irrigation at ICRISAT Niamey. A blank trial was conducted with genetic resources from Mali but differences were not observed. A new site with higher levels of Al saturation has been selected at the ICRISAT station in Mali. When the genetic resources from Mali and Niger arrive at EMBRAPA seedlings will be phenotyped the for tolerance to Al toxicity at Embrapa and genetic introgression of the best haplotypes for tolerance to Al toxicity in the best landraces and best breeding materials from both INRAN and ICRISAT Mali will be initiated using MAB. Trials are being repeated in 2009. Technical visits by scientists from Embrapa to Mali and Niger are planned for September 2009. During these visits we will discuss on how to implement MAB in Niger and Mali.

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

January 2008–December 2010; no-cost extension to July 2011

Principal Investigator and Lead Institute

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

INIFAP, 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 institutes and scientists

- CIAT: Steve Beebe; Matthew Blair
- INTA–Nicaragua: Aurelio del Llano; Julio Molina
- INCA: Humberto Rios Labrada; Orlando Chaveco
- INIFAP: Ernesto Lopez Salinas; Raul Rodriguez Guerra; Victor Montero

1. Research activities and progresses at INIFAP

Activities conducted in the first semester 2009 include: five trials under full irrigation and terminal drought during the dry season at Celaya and two trials at the lowlands in Veracruz (data is being analysed).

1.1. Drought trial. Main trial included 108 lines from CIAT (black and red seeded) and 13 local genotypes. In the stress treatment some lines were hit by *Macrophomina phaseolina* and the whole set will be inoculated with this pathogen under controlled conditions. Outstanding genotypes in the trial are shown in Table 1 (output 3).

1.2. Drought and gene expression trial. This trial included 39 recombinant lines plus parents, Pinto Saltillo (resistant) and Pinto Durango (susceptible). Samples were taken to analyse the expression of eight genes related to abiotic stress. Relative water content, biomass and seed yield were recorded.

Specific primers were designed to amplify 22 gene fragments reported as associated to drought stress resistance in dry beans. Eight out of 22 genes allowed for differentiation of susceptible from drought resistant parental cultivars: Pinto Saltillo (resistant), Pinto Durango (intermediate) and Bayo Madero (susceptible). These primers include: a ribosomal gene, a protein cinasa receptor, LEA 18, serin treonin protein phosphatase, a clon from Medicago truncatula, 9-cisepoxicarotenoid dioxigenase, casein cinasa,

Lox 2, and Histidin cinasa. These will be used to screen lines derived from Pinto Saltillo/Pinto Durango; meanwhile more sets of lines are being developed. Same primers will be tested on outstanding lines from CIAT (Table 1).

1.3. New populations. Thirty population were developed, focusing on combining outstanding drought resistant materials developed at CIAT (Central America and Caribbean basin adapted) with bred lines/cultivars from Mexico: Mesoamerica (lowland adapted), Durango and Jalisco races (Highland adapted).

Research activities at CIAT

2.1. Germplasm characterisation. Mexican core and variety collections are being genotyped as part of the project to compare genetic diversity with world-wide collection evaluated at CIAT. A group of 200 genotypes representative of INIFAP holdings (>10,000 accessions) was analysed with 30 fluorescent microsatellites based on the CIAT marker kit for bean germplasm evaluation and data analysis showed a large amount of diversity with 359 alleles detected so far across the two collections analysed. An additional 16 SSR markers will be evaluated on the collection of released varieties to see if population structure is consistent across the initial part of the study and additional marker evaluations.

2.2. Drought and low P trials. Lines selected under drought pressure were planted in a low fertility plot in Darién, Colombia, site in which low P is normally the only limiting factor. In this case significant rainfall suspended between day 13 and day 50 after planting, resulting in significant drought stress combined with low fertility. In one trial the elite line SXB 412 yielded 1257 kg/ha, versus 730 kg/ha for low P tolerant Carioca, 709 kg/ha for check Tio Canela and 653 kg/ha for a cowpea 'Mouride'. F3.4 families were also planted in this environment and individual plant selections taken for subsequent increase of which 78 lines were planted under drought stress. These would be candidates for testing for root rot resistance and low fertility tolerance under output 1.

2.3. New populations. Forty seven populations were created and have been planted for drought screening (output 3). F3.5 families from 18 populations have advanced to the drought nursery (output 4). More than 6000 F1 plants from four-parent crosses were evaluated for the bgm-1 gene (output 7).

Table 1. Outstanding genotypes in seed yield in the main stress trial conducted under irrigation during the dry season at Celaya, Mexico, 2009.

Genotype	Ave. from two trials Kg/ha	Genotype	Terminal water stress, one trial kg/ha
ALB 213	2322	SER 89	2304
SER 83	2260	MIB 451	2103
SER 109	2588	SCR 4	2028
SER 113	2142	SCR 17	2590
SCR 3	2091	SCR 18	2055
SCR 11	2093	SMR 20	2302
SEN 56	2027	SEN 10	2113
SEN 70	2017	SEN 18	2784
B98311 (MSU)	2042	SEN 44	2447
		NCB 229	2261
BAT 477 (Check)	1477	BAT 477	1097
Negro Durango (Check)	1804	Negro Durango	1719

3. Research activities in Nicaragua. A drought trial conducted in nine environments during 2008 identified four superior genotypes, INTA Rojo, INTA Sequía, 243 SS 19-4 and IBC 302-29. 243 SS 19-4 has an outstanding seed color and will be released as new cultivar in Nicaragua, whereas IBC 302-29 will be released in Honduras and Nicaragua throughout a related SICTA project.

Tangible outputs delivered
Genotype 243 SS 19-4 has been identified by collaborators as outstanding under drought stress in Nicaragua and will be released as a new cultivar.

References
Beebe E. S., I. M. Rao, C. Cajiao, and M. Grajales. 2008. Selection for drought resistance in common bean also improves yield in phosphorus limited and favorable environments. *Crop Sci.* 48:582-592

69. 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; no-cost extension to September 2010

Principal Investigator and Lead Institute
L Krishnamurthy, ICRISAT; L.Krishnamurthy@cgiar.org
Patancheru, 502 324, Andhra Pradesh, India

Collaborating institutes and scientists
- ICRISAT: Rajeev Varshney; T Mahender; Hari Upadhyaya; David Hoisington
- JIRCAS: Satoshi Tobita; Osamu Ito
- UAS–Bangalore: MS Sheshshayee

I. Introduction/Background
Chickpea (*Cicer arietinum L.*) is the third most important grain legume crop and is an important source of protein for the millions of poor in Asia and Africa. About 90% of world's chickpea is grown rainfed and therefore is prone to terminal drought stress. Thus, it becomes important to improve the adaptability of this crop to terminal drought stress. The global demand for chickpea in 2010 is projected to be 11.1 million tonnes from the current demand of 8.6 million tonnes. A combination of modern biotechnological tools and crop improvement, integrated crop management and expansion of areas to new niches can help achieve the targeted crop demand. The effectiveness of agronomical and physiological traits in improving drought tolerance would vary depending on opportunities provided by growth environments, e.g., precipitation pattern and amount, soil type, moisture status of the soil, soil compaction, etc. There is a need to characterise the drought environments so that better understanding would be obtained for suitable mechanism(s) and traits of drought tolerance for each specific environment.

High G x E interaction of chickpea yields and a low heritability of yield, particularly under drought, emphasises the need for a trait based breeding approach for achieving greater yield stability. Growth rates and productivity under water limited conditions can be improved only when relevant drought tolerance traits are brought under a single genetic background. It has been well demonstrated that better water mining abilities of the root system, superior water use efficiency or transpiration efficiency (TE) and good water conservation strategies are the most relevant physiological traits that deserve exploitation. Existence of significant genetic variability in water mining or root system extent has already been demonstrated.

Traits $\Delta^{13}C$, SPAD, and SLA, provide a convenient option for accurate phenotyping of the reference collection developed under the framework of SP1. As chickpea is a self pollinated crop species with a low level of linkage disequilibrium (LD), genome scanning approach was considered to be more appropriate for association mapping. Diversity Array Technology (DArT) markers were chosen appropriate as it offers low cost and genome wide marker profiling system for this crop species.

ICRISAT and UAS Activities
The reference collection accessions of the chickpea germplasm had been grown in a Vertisol field during the postrainy season in 2008-09 successfully. The data on 50% flowering, maturity, yield and its components were recorded. For estimating the $\Delta^{13}C$ leaf samples were collected and JIRCAS is

currently performing the analysis. A similar field trial conducted at UAS, Bangalore had exhibited unacceptably poor growth, partly due to late sowing, needing to conduct the experiment again and extend the project period till September 2010. The assessment of genetic variability in TE based on $\Delta^{13}C$ values is being planned to be taken up during Oct 2009-Mar 2010.

Progress in genotyping of the reference collection with DArT markers: As existing DArT arrays developed at DArT Pty Ltd. (Australia) turned to be monomorphic in cultivated chickpea germplasm tested, an expanded DArT array of about 15,360 genomic clones from 96 diverse chickpea genotypes (which include parental genotypes of mapping populations and diverse accessions from reference collection) have been developed in collaboration with DArT Pty Ltd. This expanded DArT array includes 7,680 new clones and 7,680 earlier clones. The DNA of chickpea reference set has already been sent to DArT Pty Ltd (Australia) for genotyping. The results are expected soon. These results would also provide for association of DArT markers with other important component traits like root system.

JIRCAS activities
With a technician trained for the routine analyses of stable carbon and nitrogen isotopes using IRMS and the capacity increased to 300 samples (with 2 replicated measurements) in a week, the analysis of about half of the samples from ICRISAT is completed. The experience gained already with chickpea leaf sample analysis for $\Delta^{13}C$ in a previous joint effort in 2005-06 on the mini core chickpea germplasm accessions (n = 211) turned out to be useful for this rapid progress.

Further plans
As the root system advantage under receding soil moisture had been demonstrated and progress made towards identification of diverse germplasm for larger root system and harvest index through SP1 of GCP, the identification of sources for better TE becomes important. With identification and validation of sources of good TE, introgression of these component traits into well accepted backgrounds through marker-assisted breeding would be considered. Currently available chickpea genotypes are narrow in adaptation and any specific genotypes being suitable only for a limited agro-ecological region. It is necessary to improve the crop for a wider adaptation to face the challenges of global warming.

Further, the changes that seem to occur with the global warming can be expected to change the agro-ecological zones and it becomes necessary to take a fresh look at the agro-ecological zones. These initial efforts towards characterisation of drought in agro-ecological zones through this project need to be expanded further keeping chickpea or cool season food legumes in view.

70. G4008.13: Improving drought tolerance phenotyping in cowpea

January 2008–December 2010

Principal Investigator and Lead Institute
Jeff Ehlers, Dept. of Botany and Plant Sciences, UoC–Riverside; jeff.ehlers@ucr.edu
900 University Ave., Riverside, CA 92521-0124, USA;; 951-323-5918; Fax 951-827-4436.
Collaborating institutes and scientists

- UoC–Riverside: Timothy Close; Philip Roberts
- ISRA: Ndiaga Cisse
- INERA–Burkina Faso: Issa Drabo
- IITA: Satoru Muranaka
- TAMU: William Payne; Bir B Singh

Our work is focused on developing and refining efficient drought screening protocols applicable to large-scale screening of germplasm and breeding populations. Such methodologies are needed to enable discovery of drought tolerant germplasm and to support breeding efforts focused on developing varieties with improved drought tolerance. Four primary activities of this project that will contribute to the overall objective are: 1) measuring grain yields of 30 early and 30 medium cycle cowpea varieties under drought at multiple environments; 2) investigating the relationship between grain yield under drought and various traits; 3) determining the relationship between drought tolerance and shoot and root

traits, and selecting potential drought tolerant materials with beneficial root characteristics which contribute higher productivity under drought conditions; and 4) measuring canopy thermal images of cowpea genotypes subjected to drought using modern thermal imaging systems. These activities will create the following outputs: 1) identification of 'check' entries for future comprehensive evaluations of germplasm for drought tolerance; 2) establish whether off-season controlled irrigation screening environments provide results that are relevant to terminal droughts that occur in main season African environments; 3) provide estimates of the consistency of drought tolerance evaluations across three West African environments; 4) identify efficient indirect screening tools for drought tolerance screening (such as the seedling drought tolerance assay we developed that is described in Muchero et al 2008); and 5) provide information on the role of root and shoot traits in the expression of drought tolerance. One complete field season has been completed to date evaluating grain yields and physiological and agronomic characteristics of 30 early and 30 medium cycle cowpea varieties under drought in Senegal, Burkina Faso, IITA-Kano and California. Other trials completed included a trial conducted at IITA-Kano to assess the relationship between canopy temperature and transpiration rate under different air temperatures and a preliminary greenhouse pot test evaluating the ability of 5 genotypes to maintain growth under mild and severe water stress. We are also currently conducting 'root cylinder' trials with the 30 early and 30 medium maturity entries in greenhouse. Significant conclusions that can be reached so far include: 1) Air temperature significantly affects the relationship between leaf temperature and transpiration rate and needs to be considered when calculating Canopy Temperature Depression as a tool for drought screening; 2) Severe drought treatments are more effective than mild treatments in discriminating cowpea genotypes in dry-down pot tests; and 3) Breeding line IT98K-555-1 was identified as a genotype with the ability to maintain relatively high growth under severe water stress, compared to other genotypes evaluated, while genotypes IT99K-241-2 and IT98K-208-5 maintained leaf greenness and relatively high transpiration rates under prolonged drought compared to other genotypes evaluated.

Tangible outputs delivered

A trial protocol manual was developed for all collaborators, including a provision for G x E analysis.

At least two complete field seasons will be needed to establish other tangible outputs with a high degree of confidence. That said, breeding line IT98K-555-1 was identified as a genotype with the ability to maintain relatively high growth under severe water stress, compared to other genotypes evaluated, while genotypes IT99K-241-2 and IT98K-208-5 maintained leaf greenness and relatively high transpiration during prolonged drought compared to other genotypes evaluated.

References

Muchero W, Ehlers JD, Roberts PA (2008). Seedling stage drought-induced phenotypes and drought-responsive genes in diverse cowpea genotypes. *Crop Sci.* 48:541-552.

71. G4008.14: Breeding for drought tolerance with known gene information

January 2008–December 2009

Principal Investigator and Lead Institute

Jiankang Wang, CAAS and CIMMYT–China; wangjk@caas.net.cn or jkwang@cgiar.org
No 12 Zhongguancun South Street, Beijing 100081, China; (86) 10-8210 5846; Fax: (86) 10-8210 5819

Collaborating institutes and scientists

- CSIRO: Scott Chapman (co-PI); Nick Hansen
- UoQ: Mark Dieters
- CAAS: Xianchun Xia; Huihui Li Changbin Yin
- GCP: Jean-Marcel Ribaut
- CIMMYT: Matthew Reynolds; Yunbi Xu
- Agropolis–INRA: Francois Tardieu; Claude Welcker
- ICRISAT: Dave Hoisington; Shyam Nigam; Vincent Vadez

1. Design of strategic breeding tools
The QuGeneUI has been extended to allow it to write and read data in the form of Excel sheets. The current interface is able to capture information from their molecular map, their QTL map and their list of additive (or epistatic) effects associated with QTL. In addition to the connection to iMAS, the QuGeneUI is able to write dataset to JoinMap formats and to GGT (Graphical Genotype) formats so that the QTL and marker map models for simulated datasets can be viewed in existing iMAS viewers (GGT). A series of other tools (collectively known as QuGeneView) have been developed to compare and review breeding strategies. A software library has been developed to allow the simple implementation of QU-GENE simulations into a virtual computer cluster of PCs running the Condor open-source software.

2. QTL mapping in F2 populations
F2 populations are commonly used in genetic studies of animals and plants. We have demonstrated that dominance can cause the interactions between markers, and proposed an inclusive linear model that includes marker variables and marker interactions so as to completely control both additive and dominance effects of QTL. The proposed linear model is the theoretical basis for inclusive composite interval QTL mapping (ICIM) for F2 populations, which consists of two steps: first, the best regression model is selected by stepwise regression, which approximately identifies markers and marker interactions explaining both additive and dominance variations; second, the interval mapping approach is applied to the phenotypic values adjusted by the regression model selected in the first step. Due to the limited mapping population size, large number of variables, and multicollinearity between variables, coefficients in the inclusive linear model cannot be accurately determined in the first step. Interval mapping is necessary in the second step to fine-tune the QTL to their true positions. The efficiency of including marker interactions in mapping additive and dominance QTL was demonstrated by extensive simulations using three QTL distribution models with two population sizes, and data from an actual rice F2 population.

3. Mapping QTL for drought and drought-related traits in maize
Following a new approach based on inclusive composite interval mapping additive and di-genic epistatic effects were estimated in a recombinant maize line population evaluated under different years, locations and water regimes, for a total of 10 environments. For each of the trials large di-genic epistasis effects were identified for male and female flowering traits, respectively up to 48 and 51% in a single trial and to a less extend for plant height. The segregation of plant height was regulated mainly by a single QTL identified under 9 environments and expressing an average of 19% of the phenotypic variance (bin 1.06). Several loci with significant additive affects were identified across trials and in general a low correlation was observed between loci expressing additive and epistatic effects. Stable epistatic effects were identified across experiments, as the result of same di-genic interaction between two loci across trials (plant height), or di-genic interaction between few regulatory genes, but with different loci depending to the environments (flowering traits). Major regulatory genes for flowering epitasis were identified on bins 2.08 and 10.05 for female flowering, bins 1.06, 8.04 and 10.04 for male flowering. Those genes might play a key role in plant adaptation, interacting with different genes involved in physiological pathways depending on environmental changes. These results may help to achieve genetic gain and improve plant adaptability in maize.

4. Simultaneous selection of major genes and QTL: a case study using coleoptile length in wheat
Loci targeted in selection composed six major genes affecting plant height, disease resistances, and grain quality, and six known and 11 “unidentified” QTL affecting coleoptile length (CL). Of the two parent lines, HM14BS contributed the target allele at two of the major gene loci, and Sunstate four. Both have a similar plant height, but HM14BS has a longer CL, a desirable attribute in many rainfed environments. Given the breeding objective of developing the maximum number of inbred lines combining the six desired major genes and long CL QTL, simulation indicated that a single biparental cross F1 produced the highest frequency of target genotypes compared with backcross populations. On average, 2.4 individuals with target genotype were present in unselected F1-derived DH or RIL populations of size 200. A selection scheme for the six major genes increased the number to 19.1, and additional marker-assisted selection (MAS) for CL increased the number of target individuals to 23.0.

Phenotypic selection (PS) of CL outperformed MAS in this study due to the high heritability of CL, incompletely linked markers for known QTL, and the existence of unidentified QTL.

Tangible outputs delivered

1. QTL IciMapping software package v2.2, available from http://www.isbreeding.net.
2. Improved QU-GENE modules, available from http://www.uq.edu.au/lcafs/qugene/.
3. Li, H, et al. 2008. *Theor. Appl. Genet*. 116: 243-260.
4. Zhang, L., et al. 2008. *Genetics* 180: 1177-1190.
5. Chapman, S. C., et al. 2008. IWGS 2008, Brisbane, Australia.
6. Wang, J., et al. 2009. *Theor. Appl. Genet.*119: 65-74.
7. Wang, J., et al. 2009. *Theor. Appl. Genet.* 118: 683-694.

72. G4008.15: Developing potato cultivars adapted to Southern Africa countries

February 2008−February 2010

Principal Investigator and Lead Institute
F Vilaró, INIA−Uruguay, fvilaro@lb.inia.org.uy

PROJECT UPDATE NOT SUBMITTED

73. 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 and Lead Institute
Abdelbagi M. Ismail, IRRI; abdelbagi.ismail@cgiar.org
International Rice Research Institute DAPO Box 7777, Metro Manila, Philippines; Phone: +63 (2) 580-5600; 845-0563; 844-3351 to 53 (ext 2843); Fax: +63 (2) 580-5699; 845-0606

Collaborating institutes and scientists

- IRRI: Michael J Thomson; David J Mackill; Thelma R Paris
- BRRI: MA Salam
- UoD: Zeba I Seraj
- BINA: Mirza M Islam

Salt-affected areas cover 1 million ha across the southern parts of Bangladesh, and pose a serious problem for resource-poor farmers who depend on rice production for their livelihoods. Developing salt-tolerant high-yielding rice varieties adapted to these areas provides enormous opportunities for improving the lives of farmers living off these marginal lands. This project aims to take advantage of modern breeding tools, such as marker-assisted backcrossing (MABC), to develop high-yielding salt-tolerant rice varieties adapted to the conditions in southern Bangladesh.

Objective 1. Introgression of salinity-tolerance QTLs into popular varieties

IRRI: Development of BR28-*Saltol* using FL478: MABC was used to transfer *Saltol*, the salt tolerance QTL, from the donor parent IR66946-3R-178-1-1 (FL478) into the popular, but salt-sensitive, Bangladeshi popular variety BRRI dhan28 (BR28). A BC_3F_2 homozygous individual with a 1.4 Mb Pokkali introgression at the *Saltol* region (10.8-12.2 Mb) and 99% recurrent parent content across the rest of the genome was identified using marker selection. When tested under salt stress of 12 dS/m, 58% of BR28-*Saltol* seedlings survived whereas all of BR28 and the sensitive check IR29 did not.

Progress in developing BR11-*Saltol* using FL478*:* A BC_2F_2 homozygous individual was identified using marker selection, with the donor introgression at the *Saltol* region (10.8-12.2 Mb) and 96% recurrent parent content across the rest of the genome. Genotyping of a BC_3F_1 population is in progress with additional background primers to further convert the background to the BR11 genome.

BRRI and Dhaka University: Development of BR11-*Saltol* using FL378: For the BR11/ FL378 population, 434 BC_3F_1 progeny were genotyped for foreground selection (RM1287, RM3412 and RM493) and background selection (77 markers) to confirm the *Saltol* introgression and recovery of BR11. Two lines NIL-52 (BR8509-73-149-52) and NIL-58 (BR8509-73-149-58) were identified with *Saltol* (10.7 to 12.9Mb) with almost no background introgression. The selected two lines will further be genotyped to fix *Saltol* and BC_3F_3 seeds will be field tested by December 2009.

Progress in developing BR28-*Saltol* using FL378: A total of 799 BC_1F_1 progeny were grown and genotyped by foreground (RM1287, RM3412 and RM493) and 49 background markers to select progenies having *Saltol* locus with maximum background from BRRI dhan28. After foreground selection, 16 progenies were selected for background check. Recovery obtained was from 32.65 to 69.39%. The selected progenies were backcrossed again with BR28 and these lines will be advanced in 2009-10.

IRRI: Mapping QTLs from BR29/Capsule: An $F_{2:3}$ mapping population from the salt tolerant Bangladeshi landrace Capsule crossed with BR29 was evaluated for 9 physiological and agronomic traits pertaining to seedling stage salt stress tolerance. Using 104 SSR markers, QTLs were identified on the long arm of chromosome 1 (different from *Saltol* region) for Na^+ uptake, Na/K ratio and survival, on chromosome 3 for Na^+ uptake, survival and SES, and chromosome 5 for K uptake and SES.

Dhaka University: Progress in mapping non-*Saltol* QTLs: Four QTLs were identified using a BC_2F_2 population of Boilam/BRRI dhan27 at seedling stage in 2008 and the progenies were advanced to BC_2F_4 for mapping QTLs of reproductive stage tolerance. Screening at reproductive stage will be completed by September 2009.

Objective 2. Seed production and evaluation of Saltol in farmers' fields. Seed multiplication of the BR28-*Saltol* line (IR89573-84) has been carried out and 15 kg pure seeds are ready to distribute to NARES partners for field trials in Bangladesh.

Objective 4. Capacity building of BRRI and other national research institutions
A marker-assisted selection (MAS) laboratory in the Plant Breeding Division of BRRI, Gazipur, was developed with funding support from the GCP. Newly installed equipment includes PCR machine, pipette sets, vertical PAGE rigs, micro-centrifuge, gel documentation system, pH meter, EC meter and a computer. A 10-day in-country training workshop on MAS was accomplished at BRRI from November 18 to 27, 2008. The objective was to strengthen the capacity of Bangladeshi research institutes in the use of molecular markers for rice improvement by providing hands-on training to apply MAS in rice breeding programmes in Bangladesh. Sixteen scientists including 10 from BRRI, 3 from BINA and 3 from DU were trained on MAS technology by resource persons from IRRI, BRRI and DU.

Next steps and/or challenges
We plan to complete the development of BR11-*Saltol*, produce sufficient seeds of both BR28-*Saltol* and BR11-*Saltol*, and extend their testing to farmer's fields starting in the dry season of 2009-2010. *Saltol* needs to be incorporated into additional varieties and further work is needed to complete the development of marker systems for newly identified QTLs and to pyramid them with *Saltol* for higher level of tolerance. More efforts also need to be devoted to field evaluation to facilitate release of these MABC varieties through the Bangladesh national seed release system.

74. G4008.17: Application of marker-assisted selection for *Striga* resistance in cowpea

January 2008–December 2009

Principal Investigator and Lead Institute
Jean Baptiste Tignegre, INERA–Burkina Faso; racinetignegre@yahoo.com or jb.tignegre@messrs.gov.bf
CREAF de Kamboinse 01 BP 476 Ouagadougou 01 Burkina Faso ; Tel (Off.): +226 50 31 31 92 02/08 Tel (Hom): +226 50 38 27 70; Cell.: +226 76 02 49 50; Cell.: +226 70 04 92 72

Collaborating institutions and scientists

- INERA–Burkina Faso: Jeremy T Ouedraogo; Issa Drabo
- IITA: Satoru Muranaka; Boukar Ousmane
- IER: Mamadou Toure

Research activities and progresses at INERA, IITA and stakeholders

1. Validation of markers for *Striga gesnerioides* resistance in cowpea

This activity aims at providing marker(s) that are effective in identifying *Striga gesnerioides* resistance genes race 1 of Burkina Faso and race 3 in Niger, both being Striga-prone area. Three AFLP markers developed during previous works were transformed into SCARs to ease their use in cowpea MAS breeding. Among this set of primers, AFLP-SCAR marker Mah-SE2 proves relatively effective in identifying Rsg 3 in the current multiple Strike race resistant lines such as IT93K-693-2, B301 and all F_1s and derived Back Cross. Then such a resistance was confirmed to be so, by using *Striga*-infested pot screenings. However, some lines resistant to *S. gesnerioides* race 1 prevailing in Burkina Faso and their subsequent crosses with susceptible lines were monomorphic. The presence of such Striga resistance marker seems to be specific of the genetic of certain combinations of crosses only.

The *Striga* seed collected in Maradi, Toumnia and Tessawa weer tested in pot wtih 60 lines of cowpea in INRAN Maradi. The 60 lines showed similar response to *Striga* seed collected in Maradi and Touminia. And, since almost all cowpea *Striga*-resistant lines (except 1 line) to the *Striga* race (SG 3) collected in Kano also showed resistance to Maradi and Toumnia seeds, the *Striga* race in these two locations may be same as the race in Kano, Nigeria and it is expected that the existing SCAR marker is effective for the MAS/MAB activities in Niger. However, the response of 60 cowpea lines to the Striga seed collected from Tessawa was different from Maradi and Toumnia. Currently, pot experiment to confirm this issue is on going in INRAN Maradi station, Niger. As well as Burkina Faso, a few *Striga* susceptible lines selected in Niger also showed presence of both Mah-SE2 and 61R and could not have good polymorphyism to conduct MAB.

2. Development of a simple MAS protocol for *Striga gesnerioides* resistance in cowpea

A simple MAS protocol based on the use of FTA cards for sampling candidate cowpea lines or segregating population for Striga resistance were adapted successfully to implement the MAS selection in cowpea. This protocol is available and works perfectly.

3. Allelic relationship study of *Striga* gesnerioides resistance in cowpea

A set of resistant lines identified in 2008 as sources of resistance to S. gesnerioides were crossed to know whether identified Striga resistance genes in new genotypes are allecic with the current Striga resistance genes. The genotypes were tested in infested plots to know their status vis avis du Strike.

4. Farmers' participatory breeding for *Striga gesnerioides* resistance

Photo 1: Farmers selecting their preferred cowpea plants (BC4F2 types in a Striga-hot spot in Est Burkina Faso; 2008.

Photo 2: A lab technician sampling and harvesting farmers' selected plants for a MAS screening purpose in Est Burkina Faso.

A participatory breeding implemented in 2008 at farmers' field of Burkina Faso showed that farmer's preferences for cowpea varieties were mostly oriented towards cowpea grain quality. Some isolated farmers still consider *S. gesnerioides* as part of cowpea rooting system and therefore as being harmless to cowpea. One hundred and eight cowpea genotypes were screened in infested pots for *Striga* resistance. This has resulted in a set of new resistant lines, but none matched farmer's ideotypes of cowpea.

The data of farmers' participatory selection activities in 4 villages of Niger conducted in 2008 in Niger, showed that grain yield and fodder yield were important criteria of adoption. Most farmers recognise *Striga gesnerioides* as a major constraint sffecting cowpea productivity.

5. Breeding line development
Four farmers' preferable cowpea lines were involved in developing back-cross populations using MAS, and currently F1BC1 populations are under development at IITA. Four back-crosses and pedigree breeding F3 lines crossed for resistance are currently being achieved. Similarly, in Burkina Faso, four BC2F1 to BC4F4 populations and four pedigree breeding lines are being assessed using participatory approach in farmer's Striga-infested fields this 2009 rainy season.

Aknowledgements
Dr Mike P Timko, Dr. Ouedraogo, Dr. T. Jeremy, GCP, Kirkhouse Trust foundation, IITA

75. G4008.19: Incorporation of an MSV resistance gene in Mozambican maize varieties, mediated by use of MAS

January 2008−Decemeber 2010

Principal Investigator and Lead Institute

Mark Laing, UKZN; laing@ukzn.ac.za
University of Kwa Zulu Natal, PBX01, Scottsville, Pietermaritzburg, South Africa

Collaborating institutes and scientists

- UKZN: Jedidah W Danson; John Derera
- IIAM: Pedro Fato; David Mariote

Research activities and progresses

A total of 9 lines adapted to low and mid altitude in Mozambique that are good for mildew, drought tolerance and yiled advantage were collected and planted at the University of Kwa Zulu Natal (UKZN). A total of 81 crosses were made randomly depending on the pollination and silking of the donor and recurrent materials. Another second set involved 60 crosses. The materials were planted in the UKZN screen houses February are due for harvest in July 2009.

1. Objectives: To access and introgress the CIMMYT CML lines with *MSV-1* resistance gene into important Mozambican maize lines.

Germplasm from Mozambique:

1. Zm 621- 4 lines
2. Zm 521-1 lines
3. Zm 421- 2 lines
4. P501-1 line
5. LP19
6. LP21
7. LP23
8. LP37D
9. LP37F

CIMMYT MSV donor lines

1. Osu23i- immune (Zim) (5)
2. CML 202- R (Ken)
3. CML 204-R (Ken)
4. CML 206-R (Ken)
5. CML 440 (Ken)
6. CML 445

Checks:

1. Hybrid 1: Olipa (QS7707)
2. Hybrid 2: Hluvukani
3. OPV1: Changalana
4. OPV2: Susuma

SSR markers for MSV-1 introgression (Chr 1.04-0.5)

locus	Bin No	Repeat motif	Forward primer	Reverse primer
Umc1144	1.04	(CT)8	ATGGCCCACTCATCATATCTCTGT	TGTGTTGATTAGCAGCGGATAAAA
Umc1243	1.04	(CAT)4	AACTGCAGAGTCGCCTGATCC	AAGCAGACTATGCTATGCTACGCC
Umc1770	1.04	(GGC)4	GAGGGATCATGGCTCTCTTCC	GTCCATCATCAGCCTGTCACC
Umc1917	1.04	(CTG)6	ACTTCCACTTCACCAGCCTTTTC	GGAAAGAAGAGCCGCTTGGT
Umc2112	1.04	(GA)12	AGCTCTACCAAACACGAGCTTCAT	CAAATGCAGAAAGATAACGCGAAT
Umc2228	1.04	(AGC)4	ACCATACCTCTCTGAACATGAGCC	GTGAGGTGAAAATGAAGCTGGAAC
Umc2229	1.04	(CGC)4	CGAAGAGCACGATGTTGACG	GAGAAGGGCGGGAGGAATAAC
Umc2390	1.04	(TAC)4	GAAATGGCAGGGAAACTTGTTTAT	AAGAGGCAAGCAAGTGTACAGTGA
Bnlg1811	1.04	AG(16)	ACACAAGCCGACCAAAAAAC	GTAGTAGGAACGGGCGATGA
Bnlg2086	1.04	AG(18)	CGGAACCTGCTGCAGTTAAT	GAGATGCAGGAATGGGAAAA
Bnlg2295	1.04	AG(30)	CGGAGGAGTGGTTCTTGAAA	GGTTAGTGAAAGGGTTGCCA
Bnlg1832	1.05	AG(15)	GCGCCCACAACAAGTAAATT	CCTCATTGTAAGGGGCAGAA
Bnlg1886	1.05	AG(11)	TCTCTCTCACATGCACGCC	TTTGATTTGGGGAACCAGAG

Tangible outputs expected

The F1s will be planted at Makatini breeding station in South Africa. The F2s will be sampled and analysed with SSR markers using the Genotying Supporting Services by the GCP. The materials

selected will be advanced to F3 at the research station. Then the F3 seed harvested will be sent back to Mozambique for selection of other traits such as mildew, drought tolerance and resistance to fungal diseases.

76. G4008.30: Development of a GCP phenotyping network

February 2008–January 2009

Principal Investigator and Lead Institute

Abraham Blum (consultant); ablum@plantstress.com
Plantstress.com, PO Box 16246, Tel Aviv Israel.

Participating consultants

- Greg Edmeades, New Zealand
- John C O'Toole, OK, USA

This project aimed to establish a strategic network of field drought phenotyping sites for GCP target crops and projects. In year 1 the project is to identify a total of 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 relied on analysis of geo-referenced climate data, water balance and target crop distribution. These were combined with site visits and previous experience of requirements to conduct uniform managed stress field trials.

The first step was to prepare a list of potential candidate drought phenotyping sites. The information base for listing candidate sites was various reports, publications, personal information, personal contacts, listed GCP projects, GIS data, weather databases, etc. We sought sites according to the following criteria:

> Locality; expertise in growing and researching GCP target crops; weather and precipitation amount and distribution; soils; expected infrastructure and staffing; motivation; experience in drought research; expected ease of access; security; etc.

A questionnaire was then designed and sent to a leading scientist or director of each potential candidate site. The questionnaire requested information such as: address, contact(s), access and transportation, weather and soil data, crops grown and their seasons, drought phenotyping practices if any, crop growing constraints (biotic, environmental, technical), infra-structure, etc'. A final list of sites was selected for visits based on responses to the questionnaire.

Site visits were divided among the three consultants: Dr. Greg Edmeades for Africa, Dr. John O'Toole for South America and Dr. Abraham Blum for Asia. A standard form for data recording was prepared and used so that all visits by all consultants would acquire a minimum standard essential dataset. A handheld GPS was used to record site/field coordinates and altitude.

Information collected on visits, together with data provided by the questionnaires was very extensive. It was summarised into a concise form made available on an excel sheet with accompanying linked figures, documents and photographs. The pros and cons for each site was summarised as final recommendation. The complete data sets and the recommendations were submitted to GCP Management Team as a confidential report. All final decisions towards phenotyping sites as based on this report or other information resources are made by GCP management team.

77. G4008.34 Environmental assessment for phenotyping network

March 2008–March 2010

Principal Investigator and Lead Institute

Glenn Hyman, CIAT; g.hyman@cgiar.org
AA 6713, Cali, Colombia; (1) 650-833-6625 Fax: (1) 650-833-6626

Collaborating institutes and scientists

- KUL: Dirk Raes; Sam Geerts; Nirman Shrestha
- EMBRAPA: Reinaldo Lucio Gomide
- Waen Associates, Wales: Peter Jones

1. Project summary

The objective of the project is to develop data, information, tools and knowledge to support site selection and germplasm deployment for the Generation Challenge Programme Phenotyping Network.

The project delivers these information, tools and knowledge of environment and agroecology in the form of databases, tables, graphics, maps and models. Tables, graphs and maps of phenotyping network test sites and associated environments are being used to support the selection of Field Phenotyping Platform (FPP) sites. The FPP sites are characterised to measure how representative they are of larger environments, and to provide information for future field trial planning. Project researchers are developing an analysis of the relationships between testing sites and environments. They are creating maps and models of soil-water relationships.

2. Summary of recent progress

Drought stress profiling

Analysis of climate data at potential trial sites can be useful for assessing the utility of a site for a given cultivar trial. The analyses can also be used to plan experiments. This project has developed a protocol for analysing weather station data to develop a drought stress profile for a site. The method draws on climate and crop simulation software freely available in the public domain. 1 - ETa/ETc was used as indicator of water stress, where ETa is the actual evapotranspiration and ETc is the crop evapotranspiration under the same conditions but for non-limiting water conditions. Water stress profiles were calculated for 9 different crops and 17 sites in Latin America, Africa and Asia. Figure 1 shows model output as a graph of water stress throughout the growing season. The data can be used to decide when to start an experiment and for selecting sites where effects of drought-timing (early, mid-season, late) needs to be evaluated.

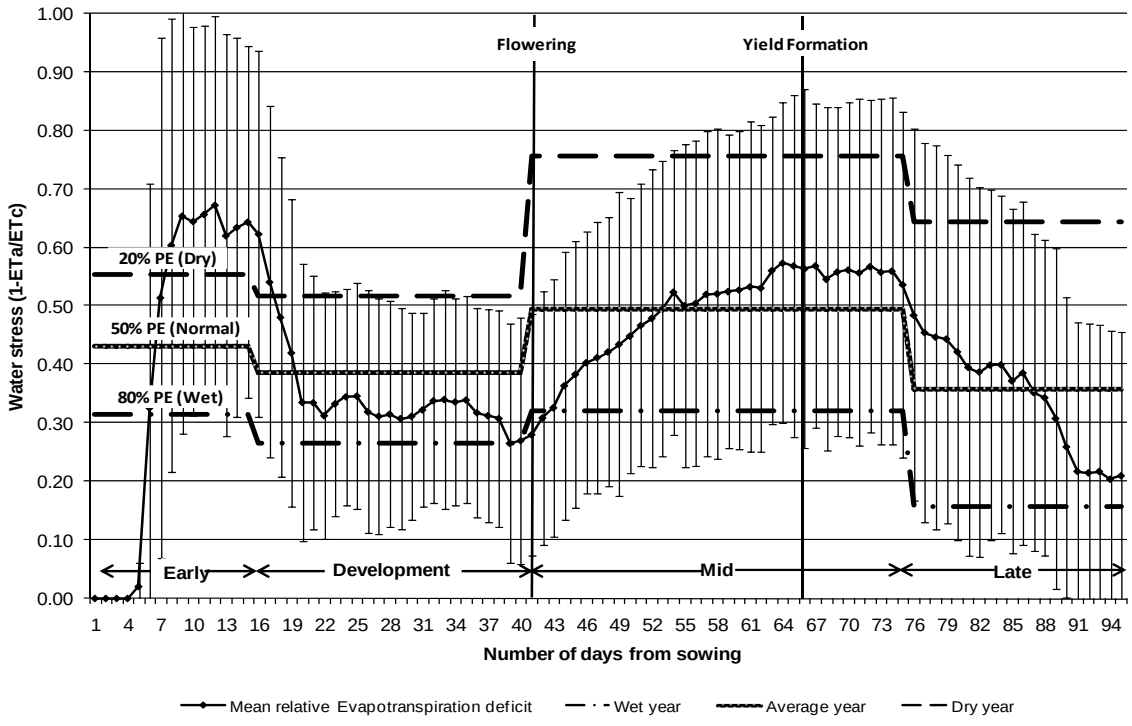


Figure 1. Example of drought stress profile model output

Link with African Trial Sites Catalogue
This GCP project motivated the development of a new collaboration among four CGIAR centres to develop a trial site catalogue for Africa (Figure 2). The project builds on the phenotyping network concept. Trial sites are characterised and information is made available to the crop improvement research community. The catalogue includes climatic data, soil profile data, cultivar performance information and contact information for the site. Trials sites are grouped according to environmental similarity. The project is scheduled to be completed at the end of 2009. This linked project provides a way to extend the phenotyping network in Africa.

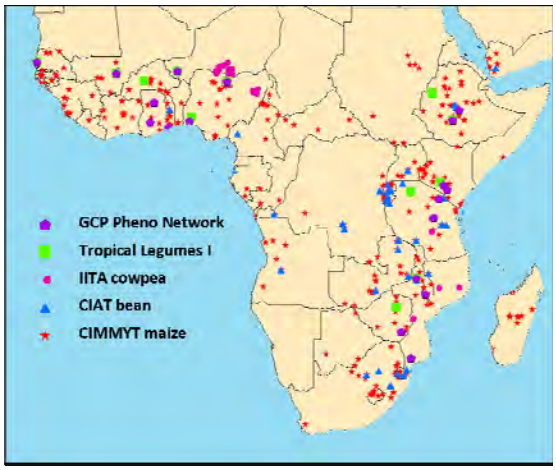


Figure 2. Trial sites of the Africa Trial Sites Network.

78. 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 (PBGB)
DAPO box 7777, Metro Manila, Philippines
phone: 0063 (2) 5805600- 2761

Collaborating institutes and scientists

- JIRCAS: Matthias Wissuwa
- ICABIOGRAD: Masdiar Bustamam; Joko Prasetiyono
- IRRI: Abdelbagi Ismail

The *Pup1* major QTL confers tolerance of P deficiency by a yet unidentified mechanism. Sequencing of the *Pup1* genomic region in the tolerant donor parent Kasalath revealed a complex genetic structure and short listed *Pup1* candidate genes are now being evaluated to identify the *Pup1* major gene(s) within the parallel SP2 GCP project "Drought from Different Perspective: Improved Tolerance through Phosphorus Acquisition".

Based on the *Pup1* sequence information, molecular markers were designed that target genes that are partially conserved in the Nipponbare reference genome (co-dominant markers) as well as Kasalath-specific genes that are located in a large insertion-deletion (INDEL) region that is absent from the Nipponbare reference genome (dominant markers). These markers were tested in more than 150 diverse rice accessions and the dominant markers showed highly diagnostic for *Pup1* (Fig. 1a, b).

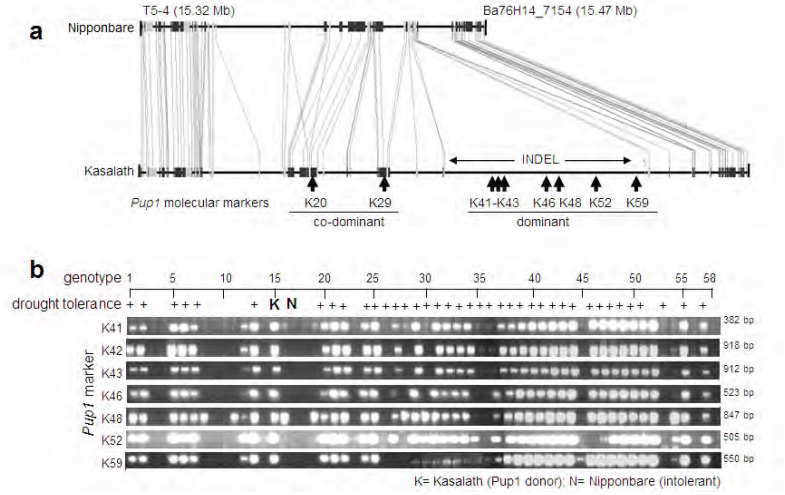


Figure 1. Alignment of the *Pup1* genomic sequence and gene-specific markers. (a) Co-dominant molecular markers were developed for two genes (*K20*, *K29*) that are at least partially conserved in Nippponbare (N, top) and the *Pup1* donor Kasalath (K, bottom). The dominant markers target six genes (*K41-K59*) located in an INDEL region that is specific to Kasalath. These markers were tested in diverse rice genotypes (b) showing that *Pup1* is present in most drought-tolerant accessions (+).

The co-dominant markers were less diagnostic suggesting that the targeted polymorphisms are not functionally associated with *Pup1* (Chin et al 2009). Allelic sequencing of these genes in tolerant and intolerant genotypes is now ongoing to identify functional polymorphisms. The marker survey revealed that *Pup1* is conserved in rice accessions that are grown in unfavorable rainfed environments whereas it is absent from modern, irrigated varieties (Fig. 2). Furthermore, *Pup1* was detected in more than 80% of the analysed drought tolerant genotypes. This finding is well in agreement with the observation that *Pup1* co-localises with a major QTL for yield under drought (Bernier et al 2007) and that *Pup1* improves yield under aerobic/drought conditions. The *Pup1* haplotype is now being analysed in a mapping population segregating for the drought QTL qtl12.1 and data will be available soon.

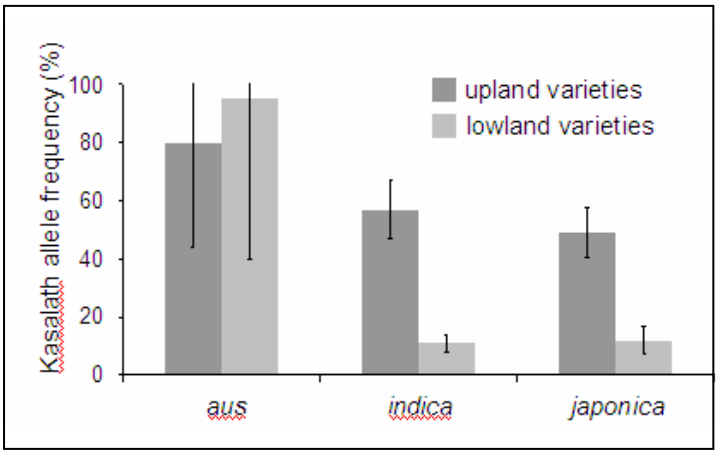


Figure 2. Representation of the *Pup1* locus in diverse rice genotypes. A *Pup1* marker survey showed that the *Pup1* QTL is conserved in most *aus*-type varieties regardless of the cropping system. In *indica* and *japonica* genotypes, *Pup1* is conserved in about 50% of accessions developed for rainfed environments, in contrast to only in about 10% of accessions grown in irrigated/lowland environments.

Several *Pup1* introgression lines have been developed by marker assisted backcrossing (MABC). The most advanced *Pup1* introgression lines of three Indonesian upland varieties (Situ Bangenit, Batur, Dodokan) have been evaluated at two sites in Indonesia (Jasinga, West Java; Lampung, Sumatera) last season and the data are currently being analysed. Selected lines were additionally grown in the greenhouse in both, P-deficient soil and Yoshida culture solution. Some of the lines, mainly those derived from crosses with the *Pup1* NILC443, performed well and data are currently being analysed. Seeds of the lines have been sent to IRRI and JIRCAS for further phenotyping and genotyping. Two *Pup1* NILs in the background of irrigated varieties (IR64 and IR74) are at an advanced stage (BC_2F_1) and will be tested in 2010 after seed increase. A first attempt to introgress *Pup1* into the upland variety Ashoka 228 failed since no F1 seeds were obtained which was likely due to hybrid sterility.

References

Bernier J, Kumar A, Venuprasad R, Spaner D, Atlin G (2007). A Large-Effect QTL for grain yield under reproductive-stage drought stress in upland rice. *Crop Sci* 47:507–516.

Chin JH, Lu X, Haefele SM, Gamuyao R, Ismail AM, Wissuwa M, Heuer S (2009) Development and application of gene-based markers for the major rice QTL Phosphate uptake 1 (Pup1). Submitted.

79. G4008.48: Improving sorghum productivity in semi-arid environments of Mali through integrated MARS

July 2008–December 2012

Principal Investigator and Lead Institute
Jean-Francois Rami, 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
Collaborating institutes and scientists
- IER: Niaba Témé; Michel Vaksmann; Sidy Bekaye Coulibaly
- Syngenta: Denis Lespinasse; Michel Ragot

The present project proposes to associate recent approaches on sorghum breeding that have been developed at IER and methodologies for marker assisted recurrent selection (MARS) that have proven to provide significant improvement of breeding efficiency for complex traits, especially in the case of maize.

Two populations dedicated to two different environments of sorghum crop in Mali will be developed from the cross of local well-characterised advanced breeding cultivars exhibiting complementary traits for the target environment. A multi-location evaluation of the progenies as F4 families, together with genotyping will provide accurate QTL detection for as many traits that have to be considered for breeding. This QTL information will be used in several consecutive cycles of recurrent selection aiming at monitoring recombination and pyramiding favourable alleles for selected QTLs. All along the recurrent process material will be released for evaluation and selfing to develop new varieties.

This project will illustrate through a private-public partnership the value of the MARS approach for sorghum breeding in Mali.

Research activities and progress at IER
Population development is ongoing from a set of selected elite and/or farmer preferred varieties. Sixteen different crosses have been realised during the 2008 rainy season in Sotuba research station, Mali (Table 1).

Three populations (V33 x Tiandougou, Tiandougou x V248, Tiandougou x Lata3) have followed an accelerated scheme comprising 2 off-season generations between Nov. 2008 and June 2009. A hiding system was specially designed to allow large-scale flowering induction of photosensitive material during the warm off-season. About 600 F2 progenies have been harvested for each population.

Seven remaining populations are being advanced from F2 to F3 during the 2009 rainy season. All the populations will be observed either as F3 or F2 during the 2009 rainy season in Sotuba. The final choice of the two populations used for MARS will be finalised based on those results in October 2009.

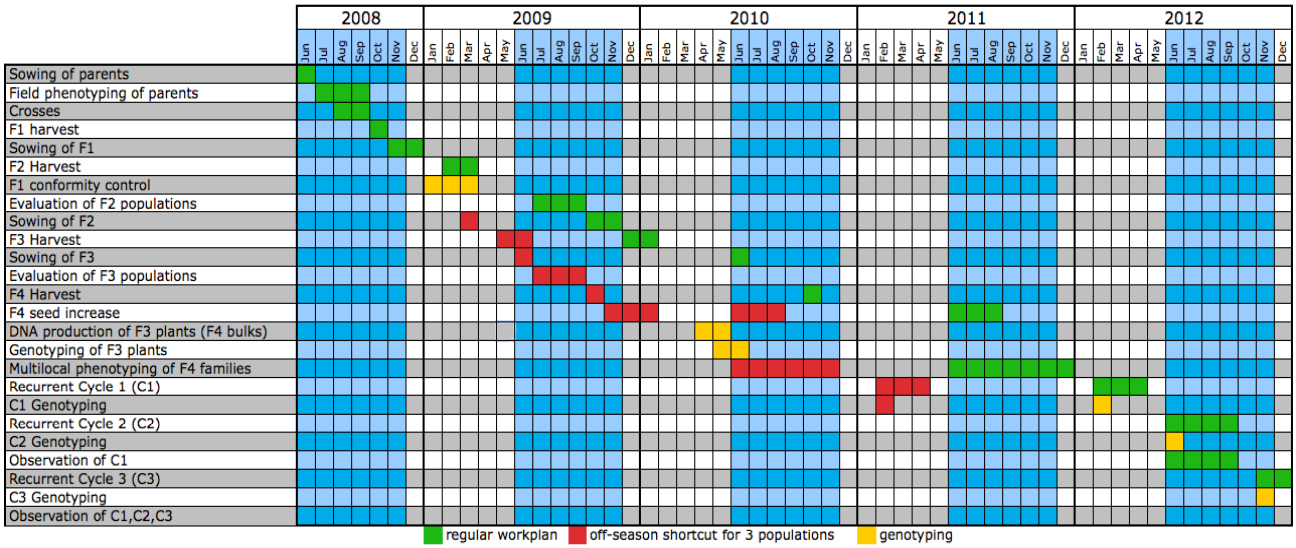
Research activities and progress at CIRAD
All the F1 plants and all the corresponding parental plants used for F1 production have been controlled for conformity and homogeneity using the sorghum SSR kit (http://sat.cirad.fr/sat/sorghum_SSR_kit/). Genotyping was performed on the Montpellier LR genotyping platform, using DNA produced from silica gel dehydrated leaves samples collected in the field in Sotuba. These analyses allowed discarding some crosses that showed unexpected level of heterozygoty for one parental plant or non-conform F1 plants.

The whole project is progressing as expected (Table 2). Next steps include production of F4 seeds for the 3 populations following the accelerated workplan and advance of the other populations up to the F3 generation. DNA for F3 individuals will be produced from a bulk of F4 plantelets grown in CIRAD and genotyping will be performed mid 2010. The first year of multi-location evaluation of F4 families for QTL detection will occur during the 2010 rainy season in Mali.

Table 1: Crosses and corresponding F1 plants realised and controlled

N°	Cross	♀	♂	F1 plants											
115	V33 x V248	10	18	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10		
116	Tiandougou x V33	9	24	A11	A12	A13	A14	A15	A16	A17	A18	A19			
117	Lata3 x V248	7	18	A20	A21	A22	A23	A24	A25	A26	A27	A28	A29		
118	Tiandougou x Lata3	9	6	A30	A31	A32	A33	A34	A35	A36	A37	A38	A39	A40	A41
119	V33 x Tiandougou	14	9	A42	A43	A44									
120	V248 x V33	22	25	A45	A46	A47	A48	A49							
121	Tiandougou x V33	5	4	A50	A51	A52	A53								
114	Tiandougou x V248	9	27	A54	A55										
279	Foulatieba x Tiandougou	13	15	A56	A57	A58	A59	A60	A61	A62	A63				
280	Foulatieba x V33	13	3	B1	B13	B14	B15	B16	B17						
281	Tiandougou x Manganié	1	19	B2											
284	V248 x Lata3	27	6	B3	B18	B19	B20	B21	B22	B23	B24	B25	B26	B27	
294	V820 x V33	11	4	B4											
295	Tiandougou x CSM63	8	26	B5	B28	B29									
296	Tiandougou x Foulatieba	8	20	B6	B30	B31	B32	B33	B34	B35	B36				
297	V33 x Foulatieba	23	13	B7	B37	B38	B39	B40	B41	B42	B43	B44	B45	B46	B47

Table 2: Workplan of the project including accelerated scheme



80. G4008.49: Enhancing groundnut (*Arachis hypogaea L.*) genetic diversity and speeding its utilisation in breeding for improving drought tolerance

August 2008–August 2009

Principal Investigator and Lead Institute

Ousmane Ndoye, ISRA; ousndoye@refer.sn

ISRA–CNRA B.P. 53 Bambey, Senegal

Collaborating Institutes and Scientists

- Agropolis–CIRAD: Jean-François Rami
- UCB: David Bertioli
- EMBRAPA: Soraya Leal Bertioli
- ICRISAT: Vincent Vadez; Hari Upadhayaya
- PROINPA: Antonio Gandarillas; Jorge Rojas; Rene Adolfo Maita
- PGRCU, USDA–ARS**:** Roy Pittman

Groundnut (*Arachis hypogaea* L.) is the most largely cultivated legume in Africa, with most of the production originating from drought-prone areas. Drought considerably reduces yield and production. Cultivated groundnut has a narrow genetic basis and the first step for improving drought tolerance in

this crop is by enhancing genetic diversity. This can be done either by accessing more effectively the genetic diversity present in the cultivated species or by tapping the genetic diversity from wild related species.

1. Research activities and progresses at PROINPA

A Bolivian core collection is being established to facilitate the evaluation of agronomical characteristics (mainly drought tolerance); in fact, 375 accessions have been repatriated from USDA collection and they were cultivated in Yacuiba (Tarija) and Sucre and new accessions of Bolivian cultivated and wild groundnuts have been collected from Southern Bolivia (Chaco boliviano). DNA of 54 Bolivian wild groundnuts (section Rhizomatosae) was obtained from the "Instituto de Botánica del Nordeste" (IBONE) - Argentina.

The data bases of the groundnut collections held by USDA, ICRISAT and Argentina (INTA – Mafredini) are being homologated in order to identify unique accessions in each collection.

DNA extraction are been performed successfully. Leaves from 200 accessions from USDA and Bolivian new collection (Chaco boliviano) has been grinded and DNA was extracted with a protocol provided by CIRAD (France). All samples were quantified and their quality was evaluated. Each accession was extracted in duplicate and their concentrations range from 30 to 40 ng/µl. The DNA extraction protocol has been adapted; hence, fifty percent (50%) of total accessions to be evaluated were collected. Thus, DNA was successfully extracted from 254 accessions

2. Research activities and progresses at CIRAD

Important progresses have been achieved in the way toward enlarging the genetic diversity of the cultivated peanut. Two marker-assited pre-breeding populations have been developed from the cross between the amphidiploid AiAd (*A. ipaensis* x *A. duranesis*) and the Fleur 11 cultivated variety. For the development of the AB-QTL populations about 150 BC_2F_2 and 300 BC_3F_1 individuals were selfed to produce the BC_2F_3 and the BC_3F_2 populations, respectively. These populations are now available and will be characterised in early September in Bambey (Senegal) for drought related traits. From this experiment, we expect to identify wild alleles that favorably contributed to drought tolerance. The CSSL(s) population is still under-development. In the initial project work-frame, it was planned to self the BC_3 selected individuals to produce the CSSL(s). However, in the BC_3 generation the number of wild segregating fragments in each selected lines was not optimal (more than 1 fragment per line) (Figure 1). A fourth generation of back-cross has been performed for achieving the best possible return to the background of the cultivated parent. The BC_4 seeds have been harvested at the end of June (2009) and 800 BC_4 seeds were sown at the beginning of July. The leaves of the BC_4 lines will be harvested in early August, dried and sent to CIRAD for DNA extraction and genotyping. The BC_4 lines carrying the desirable wild fragment in an optimal recurrent background will be selfed to produce BC_4F_2 population. The CSSL(s) population will be available by the end of this year.

In the timeframe of the project, a research paper concerning the genetic mapping of the wild introgression in the BC_1 and the BC_2 populations has been submitted for publication and a solid expertise in the construction of CSSL library has been developed at ISRA and CIRAD.

3. Research activities and progresses at ISRA

One hundred varieties randomly selected among the collection made available by ICRISAT in were involved in that study. These varieties were compared to three local varieties: 55-437, most drought resistant and 73-30, moderately drought resistant and Fleur 11 yielding variety but more susceptible to drought. The experiment was conducted to assess the TE on the accessions in well-watered and water-stressed conditions.

Under WS, the general mean of TE was 2.125 g kg^{-1} while it was 2.23 g kg^{-1} under WW. ICGV 928 showed the highest TE mean (3.383) under WS and ICGV 6394 had the highest mean (3.356 g kg^{-1}) under WW, which is higher than for 55-437 (1.700). In addition to ICGV 6394, four other genotypes had higher TE than Fleur 11 under WW.

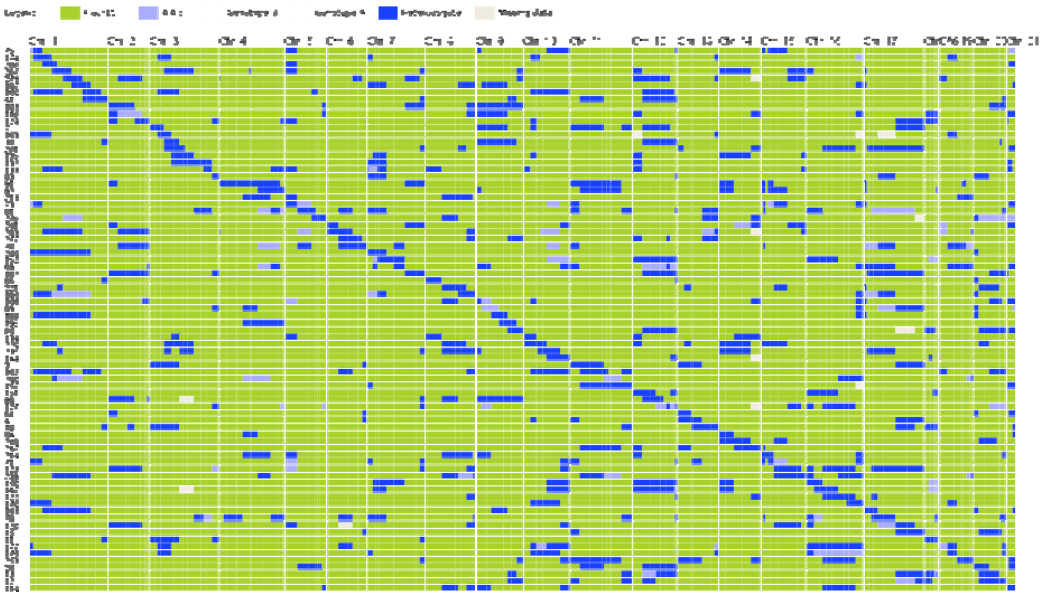


Figure1: Graphical genotypes of the 80 BC_3 lines selected for developing the BC_4 generation

81. G4008.56: Asian maize drought tolerance (AMDROUT project)

November 2008–October 2013

Principal Investigator and Lead Institute

Bindiganavile S Vivek, CIMMYT; bvivek@cgiar.org

Senior Scientist, Asian Regional Maize Programme, C/o ICRISAT, Patancheru-502324, Greater Hyderabad, (A.P.), INDIA, Phone: +91-40-30713780, Fax:+91-40-30713074/75

Collaborating institutes and scientists

- Krishidhan Seeds Ltd., India: IS Singh
- Syngenta India Limited: Raghvendra Pratap Singh
- ICFORD: Muhammad Azrai
- NMRI: Phan Xuan Hao
- IPB–The Philippines: Eureka Teresa M Ocampo
- DOA–Thailand: Pichet Grudloyma
- YAAS: Fan Xingming

Maize area in South and South-East Asia has been expanding by 2.2% annually from 16.5 (2001) to 18.0 (2006) million hectares. Over 80% of the maize is grown under rainfed conditions and prone to drought. Addressing the problem of drought has been estimated to provide the highest technical returns to rainfed maize R&D investments in Asia. Based on substantial breeding progress made for drought tolerance in maize in other regions (Central America and eastern and southern Africa), this project proposes to apply marker-assisted selection within pedigree breeding or backcrosses made between drought tolerant source inbreds and a minimum of four elite Asian adapted inbreds, and more through execution of additional self-funded and donor-funded MARS projects by public and private partners. Inbred lines will be extracted from improved populations, using either selfing or doubled haploids, and new drought tolerant Asia-adapted hybrids tested. GCP support will result in a minimum of four Asian adapted drought tolerant inbreds and hybrids, molecular marker information associated with drought tolerance, and NARS and private sector scientists with experience in integrating MARS in applied breeding programmes. The project intends to integrate self-funded public and private sector partners for a larger number of MARS breeding projects (for drought tolerance and other traits) and wider capacity building. This project is expected to become the impetus for significant levels of drought tolerance being introduced into highly relevant Asian maize germplasm with resulting impact in diverse environments and by diverse suppliers, and for a molecular community of practice being established among the Asian maize breeding community.

Subprogramme 4: Bioinformatics and crop information systems

82. G4005.22: Development of GenerationCP domain models and ontology

January 2005–December 2008; no-cost extension to June 2009

Principal Investigator and Lead Institute
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 institutes 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

The design of Generation CP information systems, tools and data exchange protocols must be specified by domain models that ensure semantic compatibility across the Consortium and that create robust global public goods from crop information. The design of such domain models, derived from scientific use cases of the GenerationCP, is forming the basis for "model driven architecture" generation of templates, web services and software driven by scientific use cases underlying the project. This project officially started in 2005, formalising preliminary efforts made in 2004. The GCP domain model and ontology form the backbone of semantic integration standards for GCP data and tools in the GCP platform and network.

Previous years of effort in this task established a mature scientific domain model documented on the Pantheon project web site (http://pantheon.generationcp.org) and used in GCP platform and network systems. Since 2007, the focus of the task shifted to the development of ontology to parameterise the model for specific semantics, in particular, ontology for plant and trait characteristics across GCP mandated crops. Existing third party standards and tools such as OBO-Edit (www.oboedit.org) have been used for this work, which is also documented on the Pantheon web site and also, for end users, on the GCP McClintock web site (http://mcclintock.generationcp.org). The GCP ontology database is deployed with a web user interface for browsing ontology at http://koios.generationcp.org:8081/ontology-lookup/.

The original project, G4005.22, led by Richard Bruskiewich of IRRI, was merged in January 2009 with a new consolidated project, G4009.03, combining GCP ontology, data templates and the GCP central registry - "*Development of data standards and community of practice enabling the capture of and access to GCP quality data sets*" -being led by Elizabeth Arnaud of Bioversity.

83. G4005.23: Implementation of web services technology in the Generation Challenge Programme Consortium

January 2008–December 2008

Principal Investigator and Lead Institute
Milko A Skofic, Bioversity; m.skofic@cgiar.org
Via dei Tre Denari, 472/a, 00057 Maccarese (RM) – ITALY; Tel: +39 06 6118286

Collaborating institutes and scientists

- IRRI: Martin Senger
- CIMMYT: Guy Davenport
- Bioversity: 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 Programme.

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:

- **getAccessionFeatureOntologyTerms:** This service returns the feature codes and descriptions that the data source makes available. It can be considered the data dictionary of the data provider, allowing clients to assess whether the provider has the required data types. The result of this service can be used as the input to getAccessionsByFeatures and countAccessionsByFeatures.
- **countAccessionsByFeatures**: This service accepts a series of features coupled with an operator and value that constitute a query to the data provider. The result is the number of accessions that match the provided query. This web service is useful to plan the next steps and have an idea of the amount of data that is available.

- **getAccessionsByFeatures:** This service accepts the same arguments as countAccessionsByFeatures, but it returns the list of accession identifiers that match the provided query, these identifiers can then be used to request the full feature set of these accessions.

- **getFeaturesByAccessionId:** This service accepts a list of accession identifiers, and for each element it returns the passport features – either the full set or only the requested features.

These web services cover the passport data of all the accessions available in the CGIAR, including the reference sets used by the GCP. In addition, they provide a view of the germplasm stored in European collections.

84. G4005.24: Application and development of web services technology

January 2007–December 2007 (NCE 2008)

Principal Investigator and Lead Institute

Mathieu Rouard, Bioversity; m.rouard@cgiar.org
Parc Scientifique Agropolis II 34397 Montpellier - Cedex 5 - France;
Tel: (+33) 4-67612908; Fax: (+33) 4-67613034

Collaborating institutes and scientists

- IRRI: Richard Bruskiewich; Martin Senger; Mylah Anacleto
- Agropolis–CIRAD: Manuel Ruiz; Pierre Larmande; Alexis Dereeper; Matthieu Conte
- NIAS: Shoshi Kikuchi; Masaru Takeya; Koji Doi

Research activities

An important goal of the Generation Challenge Programme has been to achieve biological data integration through an information platform shared between all GCP partners, allowing data sharing and analyses of data sets produced by GCP research projects. Web services have been a key technology for the effective integration of a distributed network of tools and data sources. Therefore, it appeared obvious that this technology could address this problem, particularly in an international context, and technical solutions such as 'BioMoby' have been adopted.

Development of a MOBY Tool Kit

Functionalities of BioMoby technology were extended through the implementation of specific tools that were necessary to facilitate the work of GCP developers:
MOby SErvices Support ("MOSES"), software development kit for web service provider and client development
MOBY Dashboard, an integrated MOBY developer's graphical user interface application called the that streamlines the development, registration and testing of MOBY applications in general.
MOSES and Dashboard code have been deposited in the public domain of the BioMOBY project (in the JMoby branch in the CVS at http://biomoby.org). This toolkit was adopted by members of the MOBY community and some of them participated in further developments.

Supply data to the GCP platform

Data access via Web Services was developed for several databases, including TropgeneDB (http://tropgenedb.cirad.fr), OrygenesDB (http://orygenesdb.cirad.fr), Oryza Tag Line (http://urgi.versailles.inra.fr/OryzaTagLine), IRIS (www.iris.irri.org) and the NIAS microarray database. Data sets accessible via Web Services for several crops studied at Agropolis–CIRAD, rice genome annotations (BACs, genes, ESTs, orthologs/paralogs, etc.) and the rice mutant collection (e.g. IRRI, Agropolis–CIRAD). Status and documentation can be found at http://moby.generationcp.org/.

Rice Moby network

Using the web services described in the previous section, the IRRI team implemented a mutant browser which is the first component of the Rice Moby network. The main purpose of this Browser is to find all germplasm records that have the given phenotype, or combination of phenotypes, indexed and queried by various levels of description: plant anatomy or developmental, trait or full phenotypes (trait and trait value assignments). The browser assumes the use of specific ontology (Plant Ontology, Trait Ontology and a "rice mutant phenotype" ["IRFGC"] ontology). Starting from phenotypes, other web services will be progressively integrated to link with genomics and gene expression databases.

GOST: ortholog prediction tool available through web services

With an increasing amount of data provided by GCP projects on full or partial genome sequencing, there is an urgent need to transfer the information from model species to newly sequenced ones. GreenPhyl Orthologs Search Tool called GOST(figure 1) rapidly integrates a new sequence into a pre-computed phylogenic tree developed on O. sativa and A. thaliana to infer orthologs relationships. GOST is also accessible via the GreenPhylDB website (http://greenphyl.cirad.fr/).

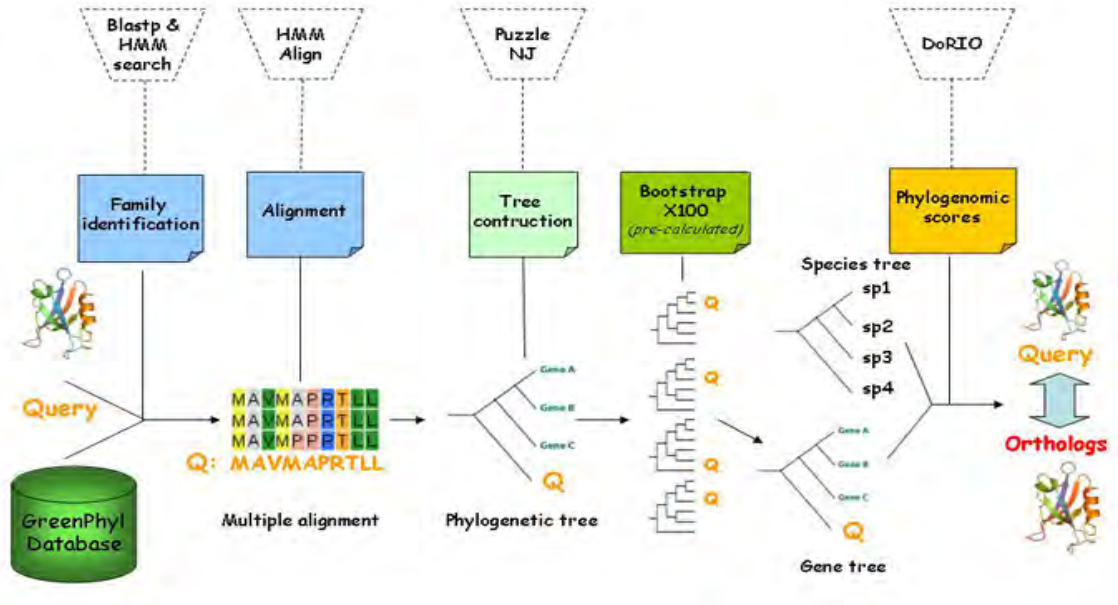


Figure 1: GOST pipeline

References

R. Bruskiewich, G. Davenport, T. Hazekamp, T. Metz, M. Ruiz, R. Simon, M. Takeya, J. Lee, M. Senger, G. McLaren, and T. Van Hintum. 2006. The Generation Challenge Programme (GCP) *Standards for Crop Data OMICS* 10(2):215-219.

R. Bruskiewich, M. Senger, G. Davenport, M. Ruiz, M. Rouard, T. Hazekamp, M. Takeya, K. Doi, K. Satoh, M. Costa, R. Simon, J. Balaji, A. Akintunde, R. Mauleon, S. Wanchana, T. Shah, M. Anacleto, A. Portugal, V. J. Ulat, S. Thongjuea, K. Braak, S. Ritter, A. Dereeper, M. Skofic, E. Rojas, N. Martins, G. Pappas, R. Alamban, R. Almodiel, L. H. Barboza, J. Detras, K. Manansala, M. J. Mendoza, J. Morales, B. Peralta, R. Valerio, Y. Zhang, S. Gregorio, J. Hermocilla, M. Echavez, J. M. Yap, A. Farmer, G. Schiltz, J. Lee, T. Casstevens, P. Jaiswal, A. Meintjes, M. Wilkinson, B. Good, J. Wagner, J. Morris, D. Marshall, S. Kikuchi, T. Metz, G. McLaren and T. van Hintum. 2008. The Generation Challenge Programme (GCP) Platform: Semantic Standards and Workbench for Crop Science. *International Journal of Plant Genomics 2008.*

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

85. G4005.27: High performance computing facilities for the GenerationCP

Principal Investigator and Lead Institute
Anthony Collins, CIP; a.collins@cgiar.org
1895 Avenida La Universidad, La Molina, Lima 12, Peru; (51) 1-3496017; Fax: (51) 1-3175268

Collaborating scientists and institutes

- CIP: Reinhard Simon
- ICRISAT: Jayashree Balaji, via Mike Butterfield
- IRRI: Guy Davenport, via Ramil Mauleon

Activities and Outputs, their status
CIP: Porting of R and Structure to alternative sustainable future Grid platforms, available to collaborators via web services, ongoing from 2008
CIP: Development of a programmatic access to the programme ***Structure***, running as a Grid resource
ICRISAT: Ongoing support for collaborators performing data analysis on SP1, 2 and 3.
IRRI: Ongoing support for collaborators performing data mining or analysis on SP1, 2 and 3.

Individual technical sub-reports by the collaborators
CIP (Anthony Collins)
Development, deployment, and support of HPC/Grid applications.

GCP HPC application support for R statistical and related bioinformatics
HPC computations is ongoing from 2008:

- Sustainable HPC options beyond the existing GCP systems identified as the BOINC and EELA grids.
- R and Structure are now being ported to both alternative platforms
- Upgraded HPC website is in construction to
  - Improve documentation and use case support for the GCP community
  - Offer processing options beyond the GCP systems: BOINC and EELA grids.
  - Improve HPC task performance

Development of a programmatic access to the programme ***Structure***, running as a Grid resource, integrated and complying with existing GCP Platform APIs:

- Startup delayed by receipt of the project budget in May 2009
- Java programmer now in training on the GCP API, and advances will be reported at the 2009 ARM

ICRISAT (Mike Butterfield)
The standalone Information system for MAB (ISMAB) is being developed to manage information generated by the marker assisted breeding programme, and for integration of information relevant to the breeding programme from local and/or external databases, with four basic components:

- Molecular Breeding Design Tool (MBDT) to assist breeders in selecting parental germplasm based on phenotypic data, and managing genotype data processing.
- Interfaces with the LIMS, providing the list of germplasm and markers to be screened which carries out the sample tracking for germplasm.
- Molecular Selection Tool (MOSEL) to facilitate the selection of the most promising lines in terms of closeness to the target genotype.
- Loading of data into the central database, for future use and dissemination.

IRRI (Ramil Mauleon)
Objectives and Outputs for the HPC:
Ongoing support for collaborators performing data mining or analysis on SP1, 2 and 3.

HPC usage report:
Primary use of the HPC is providing the computing power for:
GCP Project 2006-08 (Data analysis support for existing projects in SP2 with emphasis on integrating results across gene expression and QTL mapping experiments)

Principal Investigator and Lead Institute
Guy Davenport (CIMMYT)

Collaborators & collaborating institutes
Richard Bruskiewich, Hei Leung (IRRI)
Shoshi Kikuchi, K. Satoh, Masaru Takeya (NIAS)
Andreas Magusin (JIC)
Jose Crossa, Yunbi Xu (CIMMYT)
Reinhard Simon (CIP

Objective 2 (*which requires HPC support*) : Support to SP2 projects generating and utilising gene expression and mapping data
The GCP pipeline and analysis tools (MAANOVA, Regions of Correlated Expression, and Differentially Expressed Genes aggregation analyses, which has R code that run only on R/MPI in the HPC) were used to generate high throughput microarray data:

GCP SP2 analysis includes:
- Significance analysis of the GCP SP2-15 project (Tardieu, Serraj, et al)
- Transcriptome analysis for response to rice blast and drought response of various NILS (Kikuchi, Serraj, Leung), with Dr Kikuchi giving a plenary talk at ARM2008
- Integrated transcriptome and QTL data analysis from rice drought response by J. Bennett, S. Kikuchi, H. Leung et al, as presented by Dr Serraj at ARM 2008.
- Previous GCP-generated and public data for salinity tolerance using Affy-rice (A. Ismail, et al)

Non-GCP data analysis
- Significance analysis of disease response gene expression profiling for *Xoo* and *M. grisea* using NSF 20k platform (Wang, Ronalds, Leung et al)
- IRRI-NIAS BRAIN rice tungro virus response transcriptome project data analysis (Choi et al)
- Chilling response transcriptome analysis using NSF45K platform (Dr delos Reyes – University of Maine).
- Aggregation analysis of SNPs data for the OryzaSNP project (K. McNally - IRRI).
- Aggregation analysis of Affymetrix-rice mutant genomic deletion data in collaboration (Dr J. Leach et al. at Colorado State University).

Tangible outputs delivered
Scientific publications from the analysis outputs utilising the HPC were made (three under revision namely: Choi et al on the tungro rice virus transcriptome analysis submitted to The Plant Journal in 2008, delos Reyes et al on the chilling response transcriptome submitted to BMC Genomics in 2009, and the OryzaSNP paper by McNally et al, submitted to PNAS in 2009. One was published in 2009 at BMC Genomics from J. Leach et al , submitted in 2008 (PubMed ID 19320995).

The analysis pipeline component scripts and documentation are deposited at the cropforge website (http://cropforge.org/frs/?group_id=73) in the gcpmicroarray/GCP_developed and gcpmicroarray/R_MAANOVA_GCP releases.

86. G4006.08 Data analysis support for existing projects in SP2 with emphasis on integrating results across gene expression and QTL mapping experiments

August 2006–December 2008; no-cost extension to July 2009

Principal Investigator and Lead Institute
Guy Davenport, CIMMYT; g.davenport@cgiar.org
Apdo. Postal 6-641, 06600 México, D.F., Mexico

Collaborating institutes and scientists

- CIMMYT: Jose Crossa; Yunbi Xu; Trushar Shah
- IRRI: Richard Bruskiewich; Hei Leung; Ramil Mauleon
- NIAS: Shoshi Kikuchi; Kouji Satoh; Koji Doi; Masaru Takeya
- JIC: Andreas Magusin
- CIP: Reinhard Simon

Objective 1: Development and integration of tools for the management and analysis of gene expression and QTL data

The 3 main activities for this objective were (1) modifying existing public 3rd party software and database systems to enable storage of rice & maize – specific microarray data (maxd) and QTL information (GMOD CMap, Gbrowse), (2) extension of public 3rd party analysis software for use in rice & maize genome-wide expression data, and (3) develop/implement published genome-wide microarray analysis algorithms.

For (1), the maxd system (maxdLoad2, maxdView) source code was modified to integrate of the MIAME-Plant extensions, enabling the storage of crop-specific metadata. The 3 array platforms used in maize & rice (University of Arizona maize 46k oligoarray chip, Agilent 22k rice oligochip, Affymetrix-rice chip) were created and loaded in maxdload2, enabling the storage of microarray data generated from these platforms.

For (2), biological themes analyses tools EASE (Hosack et al) and MAPMAN (Thimm et al.) were extended to enable the analysis of rice & maize microarray results & to include rice/maize-specific metabolic pathways. The Blast2GO programme was used to extend the Mapman software for the maize oligonucleotide array using both gene ontologies as well as KEGG mappings.

For (3), case studies for significance analysis of several common microarray experimental designs are published in the cropforge website under the gcpmicroarray release. The code and documentation for the genome-wide analysis for Regions of Correlated Gene Expression (RCE) and Aggregation of Differentially Expressed Genes (ADEG) are also available in the same website.

Generic Model Organism (GMOD, http://gmod.org) tools Gbrowse and CMap were installed and adapted to enable loading of data analysis results and QTL data. These were converted to GFF & CMap table specifications and loaded into the respective tools.

Details on the availability of the software solutions are listed in Table 1 in the Tangible Outputs section.

Objective 2: Support to SP2 projects generating and utilising gene expression and mapping data

For the GCP community, a two day workshop for Candidate Gene Discovery was held in Bangkok as a pre-ARM2008 activity, with 20 participants. The workshop utilised the tools and methods developed by the project.

The analyses pipeline and software tools implementing the pipeline from Objective 1 were used on the following projects (both GCP & non-GCP) that generate high throughput microarray data (mostly run under R/MPI in the HPC):

- Significance analysis of the GCP SP2-15 project (Tardieu, Serraj, et al), primarily for gene expression data from the Affymetrix-rice microarray experiments with ~38,000 genes being analysed. Integrated results from maize-rice gene expression and QTL analyses were presented in a poster (Cairns et al) and a plenary talk (Claude Welcker) at ARM2008.
- Transcriptome analysis for response to rice blast and drought response of various NILS using Agilent 44k platform, representing ~44,000 genes (Kikuchi, Serraj, Leung), with Dr Kikuchi giving a plenary talk at ARM2008
- Integrated transcriptome and QTL data analysis from rice drought response studies using Agilent 22k platform (representing ~18,000 genes) done by J. Bennett, S. Kikuchi, H. Leung et al, as presented by Dr Serraj at ARM 2008.
- Previous GCP-generated and public data for salinity tolerance using Affy-rice (A. Ismail, et al), with ~38,000 genes represented.
- An automated annotation of the Cassava (22903 unigenes) and Cowpea (15964 unigenes) EST sequences was carried out using the Blast2GO programme. The results have been summarised into spreadsheets as well as mapped onto KEGG metabolic pathway maps.
- Analysis of the Illumina maize 1536 SNP data for diversity studies as well as association mapping, in-silico mapping of the SNP data onto the maize genome sequence associated golden path (AGP) then used for LD analysis. Coreset selection was also performed using powermarker
- Maize – rice synteny analysis using CMap. The synteny blocks can be visualised on comparative maps using CMap.

Objective 3: Common factors for abiotic and biotic stress responses across species through comparative analysis
The analysis pipeline for integrated analysis of gene expression and QTL across rice & maize genomes was presented at ARM 2008 in Bangkok and documented in a poster (Poster 4-12). This pipeline was used to identify common drought-resistance responsive regions in rice and maize for SP2 project 15 (Tardieu, Serraj, et al), and for common QTL & expression regions associated to disease resistance in rice and maize (Targeted discovery of superior disease QTL, alleles in the maize and rice genomes - Rebecca Nelson, et al).

The comparative CMap resource that show maize and rice syntenic genome regions and QTLs for various traits are available in a CIMMYT-hosted public site and in an internal IRRI website (to be migrated to a public resource this year).

87. G4006.16: Development of an integrated GCP Informatics Platform

Principal Investigator and Lead Institute
Martin Senger, IRRI; m.senger@cgiar.org
Collaborating institututes and scientists

- Bioversity: Milko Skofic
- CIMMYT: Guy Davenport
- CIP: Reinhard Simon; Anthony Collins
- Agropolis–CIRAD: Manuel Ruiz
- ICRISAT: Jayashree Balaji
- IRRI: Richard Bruskiewich

Activities and outputs, their status

- The GCP web query and display application (“Zeus”) reached the level of testing with end-users.
- GCP Ontology browser was developed and inrtoduced to the GCP Ontologists.

- Germplasm and Accession pedigree viewer was developed and introduced to the end-users.
- Development of standalone molecular breeding components MBDT and MOSEL started.
- GCP Central Registry germplasm/accession passport and metadata (indexing) started. The GCP platform development tool kit was enhanced (component inventory and testing/validation facilities, reworked GCP Data Consumers and Transformers components

Individual technical sub-reports by the collaborators

Agropolis–CIRAD (Manuel Ruiz)

The work started around March-April 2009 due to the delay in receiving the GCP 2009 funds. The achieved activities include:

- We hosted the GCP Developers Workshop 8-10th March 2009 at CIRAD, at Montpellier.
- We added the Structurama format (http://fisher.berkeley.edu/structurama/manual.html) as output of GenDiversity. Structurama is a programme for inferring population structure from genetic data.
- We improved the HaploPhyle pipeline for connection with GenDiversity.
- We are working closely with Bioversity (Max Ruas) in order to develop a MGIS GCP DataSource for musa passport data, which should be enough generic for using it for cocoa and coconut passport data. This DataSource will be connected to GenDiversity.

CIMMYT (Guy Davenport)

- Modelling work on Demeter 2.0, which is at RC1.
- The three genotyping formalise use cases, now available and documented on Pantheon web site and SVN.
- Considerable work was done with GenoMedium, paving the way for MoSel tools.

Bioversity (Milko Skovic)

- Installed *Ontology Browser* at Bioversity. The ontology browser will be used as a local ontology term repository that will work in conjunction with the web services to provide and use the terms as a data dictionary.
- Created ***GCP_SimpleIdentifierWithTypeAndOperator*** Moby data type which adds to the inherited *GCP_SimpleIdentifier* Moby data type.
- Started compilation of the Generation Challenge Programme Central Registry meta-data data dictionary for use by web services covering the Central Registry uploaded datasets.
- Started development of the Germplasm Passport Excel template parser.
- Set up Central Registry passport datasets database.
- Started development of passport web services.

ICRISAT (Mike Butterfield)

The Information system for MAB (ISMAB) is a standalone system being developed for the management of information generated during the course of a marker assisted breeding programme. The system allows the integration of information relevant to the breeding programme that may come from local and/or external databases. Based on user requirements and developer discussions the system has been designed to consist of four basic components. These include a Molecular Breeding Design Tool (MBDT) that will assist breeders in selecting parental germplasm based on phenotypic data, check availability of genotyping data for potential recipients and donors, display graphical genotypes, allow user to choose markers and design crosses and the target genotype. The second component interfaces with the LIMS, providing the LIMS with the list of germplasm and markers for which they need to be screened which carries out the sample tracking for germplasm. The third component is the Molecular Selection Tool (MOSEL) that facilitates the selection of the most promising lines in terms of closeness to the target genotype, filter sort and scroll graphical genotypes for proximity to target and parental lines, choose lines and crossing schemes for further development. The fourth component is the loading of data into the central database, for future use and dissemination.

CIP (Anthony Collins)

Development, deployment, and support of HPC/Grid applications

i. GCP HPC application support for R statistical and related bioinformatics

HPC computations is ongoing from 2008:

- Promising sustainable HPC options beyond the existing GCP systems were identified as the BOINC and EELA grids.
- R and Structure are now being ported to both alternative platforms
- An upgraded HPC website is in construction with objectives of
    - Improving documentation and use case support for all installed programmes at the service of the GCP community
    - Offering HPC processing options as the CIP Paracel system. and the BOINC and EELA grids.
- Improved HPC task performance will be demonstrated for the 2009 ARM

ii. Develop a programmatic access to the programme ***Structure***, running as a Grid resource, that can be integrated in the GCP platform by complying with the existing GCP Platform APIs (CIP)

- Startup delayed by receipt of the project budget in May 2009
- Java programmer has been recruited and is training on the GCP API

88. G4006.17: GCP quality management and data quality improvement

January 2006–December 2008; no-cost extension to July 2009

Principal Investigator and Lead Institute

Thomas Metz, IRRI; tmetz@cgiar.org

DAPO Box 7777, Metro Manila, Philippines; Tel: +63 (2) 580-5600 ext. 2895; Fax: +63 (2) 580-5699

Collaborating institutes and scientists

- CGN–WUR: 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 incorporated the previous project GCP Software Engineering and Collaboration Platforms as an objective. The project addresses issues that have strong implications on data quality and/or quality management in the GenerationCP, in particular, a Laboratory Information Management System, data quality indicators and best practices, collaboration systems for software engineering and support, and on requirements for GCP projects producing primary data.

Findings and implications

LIMS (ICRISAT): A dedicated developer provides support with evaluation, customisation and adaptation of the LIMS to interested users. During the course of the year 2008, the LIMS code underwent some rewriting to accommodate necessary changes in the workflow and machines that the LIMS interfaces with. Job scheduling was an added function incorporated into the LIMS. Three NARES partners evaluated the LIMS and one of them adopted the system for their use (NRCG – National Res. Centre for Groundnut). Amongst international institutes, evaluation support was provided to the Lab des Interactions Plantes Micro-organismes at CNRS-INRA and evaluation and implementation support to the University of Washington. Users and/or programmers from Centres/Institutes adopting the LIMS were invited to attend a 2.5 day LIMS workshop held from the 30th March -1st April 2009. During the workshop, the most recent version of the LIMS was demonstrated; an interactive meeting with ICRISAT scientists/users was arranged to allow LIMS developers to understand user perspectives and approaches to applications such as the LIMS. User participants were trained on the process of installation, deployment and use of LIMS, while developer participants held discussions on the code structure, about the reusable style sheets, common classes created to implement common functionality,

followed by a demonstration of adding a new module to existing code structure. Finally, participants were trained on the use of SVN for code versioning and the need to keep code in the public domain under the CropForge LIMS project was emphasised. The technical support from ICRISAT-Nairobi has been trained and entrusted with the task of enabling users at the BecA facility (ICRISAT, IITA, CIMMYT, ILRI) to use the current version (2.9) of the LIMS.

Data quality indicators and best practices (CGN, CIP, CIRAD): Analysis of data quality of GCP SSR data sets were presented by Theo Van Hintum (CGN) in a poster during ARM 2008 and in a presentation during a GCP Workshop in Montpellier. The reports of the tests were made available to the GCP Management Team, and the results were also presented and discussed within the Management Team. Major findings include:

- the conformity of the files to the GCP SSR template varies strongly and needs to be improved to allow proper interpretation and automatic processing many errors in the files can be identified by visual inspection or simple analysis
- the enormous range observed in both DR and M50 cannot be explained by lack of population structure; low DR combined with high M50 values are an indication of poor data quality
- many data sets have a poor to very poor data quality
- the fact that allowing for binning-errors improved nearly all DR values indicates that this is a major weak spot in the genotyping protocols using SSRs in the GCP

Based on the GCP template structure, a data uploading and decentralised quality control system has been developed at CIP (Reinhard Simon). The system is based on a set of freely available tools plus custom scripts in R and has been tested with real-world examples. The report creating infrastructure is a one-click batch driven approach based on R and tools around R. The system enables de-centralised quality control by users before uploading of datasets as well as automated runs on the repository server.

Collaboration systems for software engineering (IRRI): The systems were maintained and updated and user support was provided. The CropForge system now hosts 84 projects and has become the standard software engineering platform for SP4 projects. Average availability for the CropForge server (http://cropforge.org) was 96% and for the CropWiki server (http://cropwiki.irri.org/gcp) 99.7%. The availability is constantly monitored and reported here: http://www.pingdom.com/reports/ojapqnsplqwg/).
User manuals for using the CropForge server for collaborative software development projects are available here: http://cropforge.org/projects/cforgeinfo.

White paper on requirements for GCP projects producing primary data (IRRI): The first ever meeting of staff responsible for research data management in the different CG Centers was hosted by Bioversity International in Rome during 9-11 June 2008. A prior questionnaire had collected information on research data management problems, solutions, and the current status. Subsequent to the meeting a paper was written by the workshop participants and presented to the CG-ADE during the AGM 2008 in Maputo, Mozambique, December 1-5, 2008. The paper analyses common research data management problems and makes specific recommendations for strategy and actions. A copy of the paper is available from here: http://cropwiki.irri.org/gcp/images/4/40/ResearchDataManagement_CG-ADE.pdf

89. G4006.35: Statistical support for the design and data analysis of GCP projects

January 2009–December 2009

Principal Investigator and Lead Institute

Marcos Malosetti, WUR Biometris; marcos.malosetti@wur.nl
P.O. Box 100, 6700AC, Wageningen, The Netherlands; (31)317482627; Fax: (31)317483554

Collaborating institutes and scientists

- WUR–Biometris: Hans Jansen, Fred van Eeuwijk, Marco Bink
- UNAM (presently UoC–Davis): Joost van Heerwaarden

Context

This project aims to provide statistical support to GCP scientists on issues that ranges from experimental design to different types of data analyses (field and lab data analysis, diversity analysis, QTL mapping, association mapping, etc). The statistical support is channelled through three major activities:

1. Development of training materials encompassing data quality, general statistical theory, experimental design, analysis of single and multiple trials, GxE analysis, QTL and association mapping, diversity analysis, selection response theory
2. Training courses on selected relevant subjects
3. A Statistical Helpdesk providing one-to-one support with specific statistical questions coming from researchers.

Outputs (period January-July)

Two training courses/workshops have been delivered in the first half of the year. A first workshop was held at CIAT headquarters in Cali, Colombia, from January 26-30, 2009. This workshop was organised in collaboration with the Genotyping Support Services with the objective of assisting beneficiaries of the GSS programme in the analysis of the data received from the lab shortly before the workshop started. Researchers brought in the data consisting of panels of genotypes assessed by DArT or SSRs markers in combination with phenotypic characterisation and/or passport data in order to answer specific research questions. Most of the research issues related to the assessement of the genetic diversity in germplasm collections in relation to adaptation to abiotic (drought) or biotic stresses (resistance). The participants were from Bolivia, Ethipia, Ghana (3), and Kenya (2), working in potato, cassava, maize, rice, ensete, and yam. The dynamic of the workshop consisted of: a) presentations/lectures where key concepts useful to perform the different types of analyses were introduced, and b) a one-to-one interaction assistance on data analysis. The issues addressed in the different presentations during the course were: data quality control, experimental design, population genetics and diversity analysis, introduction to clustering, and association mapping. In addition, and due to the fact that many of the research questions related to drought stress trials, extra attention was dedicated to phenotyping procedures under drought stress (by Sam Geerts, Leuven University, Belgium).

A second workshop was organised by SP5 to provide statistical support to researchers working within the framework of the project *Improving tropical legume productivity for marginal environments in Africa*. The venue of this event was the IAMZ (Mediterranean Agronomic Institute of Zaragoza), Spain, from June 29th to July 3rd, 2009. This hands-on training activity was attended by 21 researchers mainly from Africa (Tanzania, Senegal, Malawi, Niger, Burkina Faso, Mozambique, Ethiopia, Kenia, Zimbabwe, India, and Iran) working in the following species: groundnut, cowpea, common beans, and chickpea. The activity consisted of a combination of oral presentations and a one-to-one interaction with trainers to discuss and delineate a strategy for data analysis. The presentations addressed aspects including experimental design, field data analysis, molecular maps construction, QTL mapping, and association mapping. These two workshops constitute the planned outputs for activity 2.

As a result of the two workshops, an updated set of training material has been compiled, including aspects on data quality control, experimental design diversity analysis, molecular map construction, and theory on QTL mapping either using biparental crosses (conventional QTL mapping), or diverse populations (association mapping). As such, this material is a first output for activity 1 as it has been compiled and made available to participants for distribution in the form of either a CD, or by uploading it to internet from where it can be downloaded. In the second part of the year, the efforts will concentrate on the standardisation of the material so as to make it available on a website and additionally in the form of written material, thereby fulfilling the planning with respect to outputs in this respect.

The third type of output from this project corresponds to consultations from researchers. During the first half of the year, we have received some consultations, mainly involving questions related to QTL and association mapping. Those were re-directed to us via the Bioinformatics Portal.

90. G4007.09: Design and analysis of marker trait associations studies with special attention for genetically challenging crops

January 2009–December 2009 (No-cost extension)

Principal Investigator and Lead Institute
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 institutes and scientists

- UoH: Hans Peter Piepho; Albrecht Melchinger
- ICL: David Balding
- NIAB: Ian Mackay; Wayne Powell
- SCRI/BIOSS: Christine Hackett; Dave Marshall
- LUMC: Hans van Houwelingen; Jeanine Houwing-Duistermaat
- WUR: Marcos Malosetti; João Paulo; Marco Bink; Hans Jansen

Context
The project proposal included six major objectives: 1) Describe the current status of LD studies; 2) Design a data quality protocol for phenotypic and genotypic data to be used in LD studies; 3) Make an inventory of the major factors causing LD and how to account for them in LD studies, 4) Establish guidelines about the design of LD studies (number of markers, number of genotypes, which marker systems, etc); 5) Develop statistical methods for LD analysis accounting for factors affecting LD; and 6) Deliver training on statistical methodology for LD analysis. Last year it was discussed with the GCP SP4 leader that developing the objectives exclusively for challenging crops would be too hard within the financial constraints of the project and it was decided to develop general protocols for design and analysis of LD studies, paying attention to genetically challenging crops like potato and sugar cane. Also, as GxE and QTLxE are of particular interest to the GCP, attention will be given to methodology for the detection and description of QTLxE in association mapping contexts.

Outputs as of July 2009
Advances have been achieved on all fronts (objectives 1 to 6). Work on QTLxE and association mapping was carried out by three students; two post-docs and a PhD, namely Dindo Tabanao, working on association mapping and QTLxE in barley, Caroline Castro working on the same topics in potato, and Maria Marta Pastina working in sugar cane. The QTLxE work was supervised by Fred van Eeuwijk and Marcos Malosetti. On another topic, Hans Jansen led work on data quality requirements and design questions. A further PhD, Thomas Odong, and post doc, Joost van Heerwaarden, worked on identifying population structure in association panels. This work was supervised again by Fred van Eeuwijk, Hans Jansen and Theo van Hintum.
Objectives 1, 2 and 6 were well underway last year, but new material/documentation has been added (or will soon be available):

- Objective 1: a review on LD studies will submitted for publication in Trend in Plant Science in the second part of this year.
- Objective 2: a paper treating different aspects on data quality control has been submitted to Theoretical and Applied Genetics; "*How many markers can be put on a dense genetic map?"* by H. Jansen, L. Bardaji, M.P. Boer, P. Stam. A guidelines document/ web pages will be prepared.
- Objective 6: LD website updated, further updates are necessary and will follow before December 2009.

With respect to objectives 3, 4 and 5 the outputs include research results and new publications:

- Objectives 3: different methodologies to account for population structure in the type of populations used in LD mapping have been investigated. Those include both, the evaluation of different methods to estimate coancestry between genotypes based on molecular marker information, and the evaluation of alternative methods to define the number of subgroups within a population. The work has been done using as example crops barley and potato and the results will be published in the second part of the year (see below).
  In addition to this, and in relation to the same topic, work was done on the identification of population substructure. Part of this work appears in a draft entitled "Determination of genetic structure of germplasm collections: Are traditional hierarchical clustering methods appropriate for genetic marker data?" by TL Odong, J van Heerwaarden, J Hans Jansen, TJL van Hintum & FA van Eeuwijk has been submitted.
- Objective 4: a document entitled "The power for testing association between a marker and a QTL" has been produced by Hans Jansen that will be made available through the Biometris LD website.
- Objective 5: A number of publications on LD methodology has appeared in which issues of LD mapping in relation to GxE and QTLxE:
  - Pswaray, A; van Eeuwijk FA, Ceccarelli S, Grando S, Comadran J, Russel JR, Pecchioni N, Tondelli A, Akar T, Al-Yassin A, Benbelkacem A, Oudabbou H, Thomas WTB, Romagosa I (2008) Changes in allele frequencies in landraces, old and modern barley cultivars of marker loci close to QTL for grain yield under high and low input conditions. *Euphytica* 163:435-447.
  - Romagosa I, FA van Eeuwijk, WTB Thomas (2009) Statistical Analyses of Genotype by Environment Data. In: *Cereals* (Carena MJ, ed). Springler, New York USA, pp 291-331.
  - Ignacio Romagosa, Jordi Voltas, Marcos Malosetti, Fred A. Van Eeuwijk (2008) INTERACCIÓN GENOTIPO POR AMBIENTE. La adaptación al ambiente y los estreses abióticos en la mejora vegetal. CAPÍTULO 5, p. 107-136. JUNTA DE ANDALUCÍA. Instituto de Investigación y Formación Agraria y Pesquera, Consejería de Agricultura y Pesca Publica: Dirección General de Planificación y Análisis de Mercados. Servicio de Publicaciones y Divulgación. Editores Científicos: Carmen María Ávila Gómez, Sergio Gustavo Atienza Peñas María Teresa Moreno Yangüela, José Ignacio Cubero Salmerón

  Two more publications are to be submitted in the second part of the year presenting the major results from the work of the two postdocs at Biometris where different methodologies to account for population structure (objective 3), and models to perform marker-trait association (objective 5) will be discussed.

91. G4007.10: Support to GCP scientists regarding issues related to bioinformatics and data handling

January 2008–December 2009

Principal Investigator and Lead Institute

Theo van Hintum, CGN–WUR; theo.vanhintum@wur.nl
Centre for Genetic Resources, The Netherlands (CGN), P.O. Box 16, 6700 AA Wageningen, The Netherlands; +31 317 480884
Co-PI: Elisabeth van Strien, CGN–WUR; elisabeth.vanstrien@wur.nl

Collaborating institutes and scientists

N/A

Context

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. 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 was not very accessible. Therefore an easy access to the available facilities, expertise and products needed to be created.

To address these issues the 'SP4 helpdesk on Bioinformatics and Biometrics' (in short: SP4 helpdesk) was established, a one-stop-shop to support GCP scientists regarding issues related to bioinformatics and data handling.

Findings and implications
In the year 2008 the GCP-SP4 helpdesk became operational. The GCP-SP4 helpdesk is the entry point for any GCP scientist who has questions regarding handling, storing, or analysing GCP data. The SP4 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.

The GCP-SP4 helpdesk pro-actively advised on the improvement of GCP web-sites, created an expert network and acted as a point of reference for GCP scientists. The expert network accompanying the GCP-SP4 helpdesk was updated in 2009.

In the course of the first year it also became apparent that many of the created products in SP4 do exist, but were hidden on local hard disks, often not in a condition to be shared with colleague scientists. The GCP-SP4 helpdesk tried to identify which products need to be accessible, approached the PI and supported him/her in making these products available for use to others.

This involved reformatting of reports and presentations, adding explanatory information and exporting information into formats that could be easily be accessed and interpreted by others.

2009, the first half of the second year of the helpdesk project, showed a slight increase of the number of requests for support, all of which were handled as described in the protocol established in 2008.

The subjects of the requests varied from questions regarding biometrical methodology, to assistance on template choice, guidance and assistance on uploading to the Central Registry and assistance to other helpdesks. In order to increase the awareness of the services that the SP4 Helpdesk can offer to breeders, scientists and software developers, handouts were produced and dispatched to GCP crop researchers and breeders at two recent GCP organised courses.

In order to make SP4 products more accessible to researchers and breeders, as well as software developers both from within and outside the GCP, a new bioinformatics portal has been developed, allowing users to walk through the various aspects of SP4, and guiding them to information about and access to products and expertise produced in the framework or with the support of SP4. This includes access to the material used at SP4 meetings, explanation on how to get access to the HPC facilities, but also who to contact if a user is interested in the GCP activities in fields such as ontology development or webservices. In this context four SP4 researchers were personally contacted and interviewed on the progress of their activities, and the products their groups developed (mainly in the area of programming software and webservices).

The number of visits and the origin of the visitors of the GCP SP4 bioinformatics pages were regularly monitored. The SP4 old portal pages were steadily visited in 2009. The 828 visitors originated from 61 countries, which resulted in 1163 pageviews.

Next steps
The bioinformatics portal on the GCP website will continue to be improved, expanded and updated, and in this context the attempts to capture the SP4 products (identify, describe, and make them accessible) will continue. These products will be promoted via the SP4 bioinformatics portal.

The SP4 helpdesk will continue to complement and to be complemented by other GCP helpdesks. As such, it will 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. The increased promotion of the helpdesk will be continued by dispatching handouts at courses and meetings, and by actively contacting PIs, mainly via the data hunting route for project G4009.03.

92. G4007.11: Further development and support for use of iMAS by NARS and the other user communities

January 2008–December 2008; no cost extension to December 2009

Principal Investigator and Lead Institute

Jayashree Balaji (until October 2008); Abhishek Rathore (effective October 2008); Scientist, Biometrics, ICRISAT; a.rathore@cgiar.org

Patancheru 502324, Andhra Pradesh, India; Phone: +91 (40) 3071 3413, Fax: +91 (40) 3071 3074

Collaborating institutes and scientists

- ICRISAT: Mike Butterfield; Tom Hash; Jayashree B
- IRRI: Richard Bruskiewich
- CIMMYT: Guy Davenport

Context

Over the years molecular marker technology has become an integral part of any breading programme due to its advantages over conventional breeding programs including study of multiple genes, late expression of trait, seasonal and geographical considerations. However, the bottleneck in application of molecular markers is to understand and use complex algorithms (many are available in form of computer programmes) to construct appropriate linkage map and detection of quantitative trait locus (QTLs).

The iMAS system is designed to provide a single unified computing and decision support platform to facilitate marker-aided selection and breeding which involves the use of a number of different computing tools. Each of these tools has its own input data file requirements, which require manual preparation of the required input data files. This is both time-consuming and error-prone. Also, most of these tools lack the availability of simple-to-use guidelines for their correct and appropriate use. The current version of iMAS (iMAS 1.8) has mitigated these difficulties by providing a single unified computing and decision support platform. iMAS 1.8 allows researchers to carry out appropriate biometrical analysis, construct linkage map for single and multiple populations, multiple environment QTL analysis, graphical display of genome of progenies and also planning of marker aided backcrossing.

System consists of six modules Data Validation, Phenotyping, Linkage Map Building, QTL Analysis, Genome/QTL Display and Marker Assisted Breeding.

Progress

During 2009 several GUI modifications within Phenotyping, Linkage Mapping and QTL analysis modules were carried out to make several analyses easier. Based on feedbacks and suggestions received phenotyping module was redesigned to enable user to summarise biometric analysis in a much easier fashion. In addition to this Linkage Map building module was extended to use multiple populations to build consensus map.

QTL mapping module has been extended to work with multi-environment data and correspondingly input files (*.qdt and *.qin) for linked programme PlabQTL has also been modified. Integration of simulation system QUGene with iMAS is also under progress, in which the creation of input file for simulation system is under progress. However based on discussions with Martin Senger, it was decided that iMAS integration with GCP information platform will not be feasible. Integration with CMTV has been implemented successfully and accordingly, online decision guidelines and user manual has been updated.

iMAS setup file has also been modified for easier installation with less dependencies and the project with source code has been updated to Crop Forge Repository.

iMAS has been used in several training programmes to NARS and other partners and their feedbacks has been recorded. In addition to this a final debugging workshop has been scheduled from 24-26 August, 2009, which will be followed by a final release workshop in December.

93. G4007.12: Development of tools and technology to increase the functionality of the GCP Information Platform

February 2008−December 2008

Principal Investigator and Lead Institute
Martin Senger, IRRI–CRIL; m.senger@cgiar.org

Collaborating institutes and scientists

- IRRI−CRIL: Graham McLaren; Richard Bruskiewich
- Bioversity: 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 biologicaly-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.

94. G4008.21: Large scale phylogenomic analyses to gene function prediction for GCP crops

January 2009–December 2009

Principal Investigator and Lead Institute

Mathieu Rouard, Bioversity; m.rouard@cgiar.org

Bioversity International, Parc Scientifique Agropolis II 34397 Montpellier - Cedex 5 - France; Tel: (+33) 4-67612908; Fax: (+33) 4-67613034

Collaborating institutes and scientists

- Bioversity: Matthieu Conte
- Agropolis–CIRAD: Christophe Périn

1. Research activities and progresses

1.1. Genome Clustering and update procedure

Following the 2008 activities, we continued with the insertion of new plants' genomes to enrich the previously-computed clustering on the 12 plant genomes currently featured in GreenPhylDB. In order to facilitate this time consuming task, we developed a semi-automatic procedure to handle insertions of new genome releases. Thus, we added 4 recently sequenced genomes (maize, papaya, castor bean, brachipodium) and updated genomes already integrated in the database (e.g. v7 to v9 for arabidopisis t. and v5 to v6 for Oryza s.). Updates are a critical step to keep the database attractive for end users.

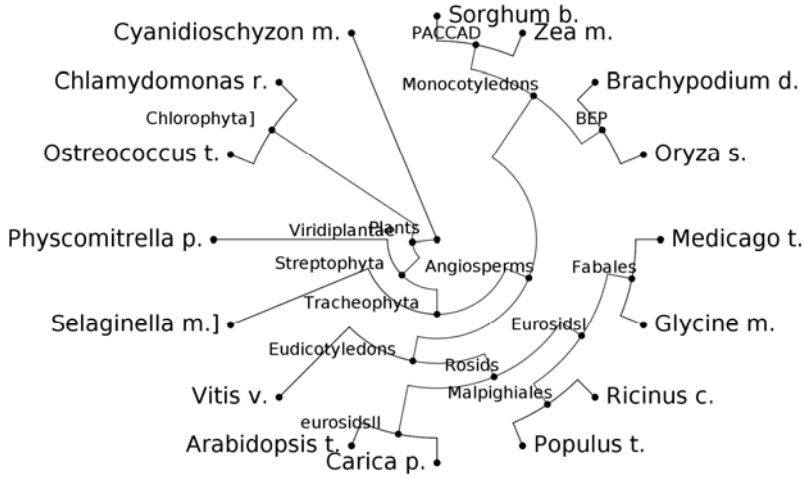


Figure 1: full genomes clustered in GreenPhylDB

1.2. Annotation of gene families

Although,GreenPhylDB stores results of an automatic clustering, overall consistency of the clusters is manually checked. A web interface displaying relevant information (Statistics, Domain shuffling, cross-references, publications etc.) and guidelines were produced to help with manual annotation of gene families. Annotation has been started and almost 2000 genes families were annotated with emphasis on gene families related to agronomic traits and evolutionary processes. Following the example of gene annotation, we believe that gene family annotation is an essential step before any phylogenetic analyses. For instance, using Gene Ontology terms (GO), we can now identify lists of gene families involved in plant stress response.

1.3. Pipeline of analyses

In order to cope with the increase of sequences contained in the clusters, a more efficient pipeline for phylogenomic analyses has been designed and is being set up on the supercomputer centre (http://www.cines.fr). Annotated gene families will be processed shortly and ortholog predictions will be made available on line.

1.4. Website

Version 2 of the the website is available on line but still in beta version (http://greenphyl.cirad.fr/v2/cgi-bin/index.cgi). Although the number of data has increased drastically, we have also optmised the database structure and the script code to speed up loading of pages. A tool to localise genes belonging to same gene family along the chromosomes was added to the website.

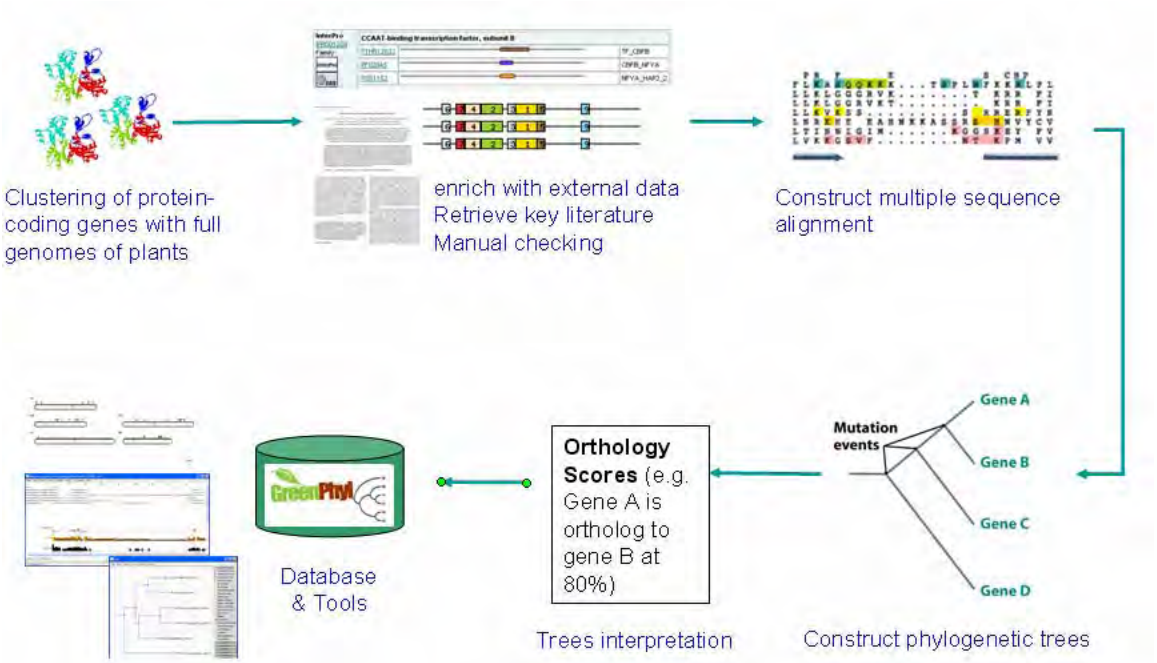


Figure 2: main steps of the phylogenomic analyses in GreenPhylDB

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

Conte, Matthieu, Laporte, Marie-Angélique, Périn, Christophe, and Rouard, Mathieu . GreenPhylDB: A Gene Family Database for plant functional Genomics. Available from *Nature Precedings* <http://dx.doi.org/10.1038/npre.2009.3136.1> (2009)

95. 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, no-cost extension in 2009

Principal Investigators and Lead Institute

Claire Billot and Manuel Ruiz, Agropolis–CIRAD
UMR DAP, Avenue d'Agropolis, 34398 Montpellier Cedex 5, France, Claire.billot@cirad.fr, manuel.ruiz@cirad.fr, 33 4 67 61 44 65, Fax: 33 4 67 61 56 05

Collaborating institutes and scientists

- Agropolis–CIRAD: Xavier Perrier; 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, analysing 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.

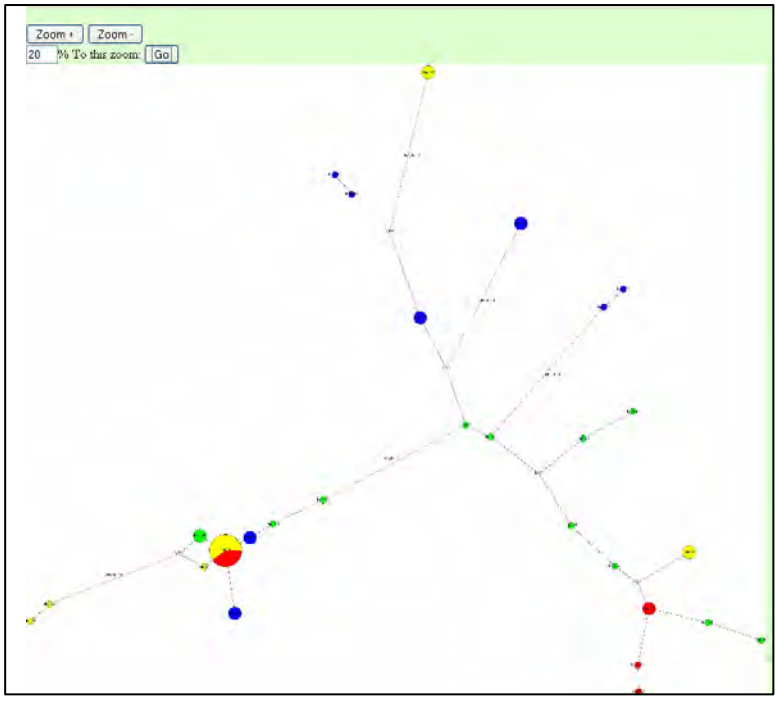
This project provides the community with new software for analysis of genotyping data, HaploPhyle (version 1.0 is downloadable at haplophyle.cirad.fr). This Web based pipeline includes haplotype definition, haplotype network analysis and connexion with external data, such as geographic origin, evolutionary history or genetic group assessment. It differs from existing projects in the fact that different methods in haplotype definition and haplotype network are available for users, with tuneable choice criteria, as well as sub-optimal networks. It is developed and integrated by two research groups: one group at Agropolis-Cirad takes care of the pipeline including haplotyping, haplotype network construction and its illustration by external data, as well as some methodological aspects of network construction. CIP is more involved into connexion with DIVA-GIS, an already existing tool which manages geographic information, in order to integrate geographic information and enable pertinent modes of graphical representation.

After an extensive study of haplotyping methods and software associated, with the major constraints of being used with Linux and being free of public use, the Java code of Gevalt was slightly modified for integration into our pipeline. Visualisation at once of the haplotypes ordered according to a tree can be performed with a R function.

Haplotype diversity results from an evolutionary process that can be represented as a tree graph connecting the different haplotypes weighted by their frequency in the population. However the process is sometimes more complex than a tree structure and the level of diversity between haplotypes is often low. Consequently a large number of optimal or sub-optimal trees can be equally likely solutions. So it is generally proposed to represent the haplotype diversity as a network combining these trees. We implemented new Java modules for haplotype network construction. Two methods were implemented and tested: median joining network and minimum spanning network.

We developed a module, using GraphViz, for providing graphical network representations. Nodes sizes are proportional to the numbers of germplasms sharing a haplotype and edges length is proportional to the difference between haplotypes. Nodes of the network are represented by pie-charts taking into consideration the repartition of external information (set by the user) among germplasms owning this haplotype.

Once a network defined, the user may want to illustrate it with any kind of external information, such as geographic origin, racial characterisation, specific traits, previous genetic information, etc... With HaploPhyle, users can integrate any 'layers' of geographic information, or any other illustrative information.



Graphical output of the haplotype network

96. 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 and Lead Institute

Robert Koebner, CropGen International; mockbeggars@gmail.com

Mockbeggars, Townhouse Rd, Old Costessey, Norwich, NR8 5BX, UK, +44 (0) 1603 745 036; **www.cropgeninternational.com**

Collaborating institutes and scientists

- CIMMYT: Guy Davenport
- CRIL/IRRI:Warren Vincent E Constantino

The major goal of the project was to create a wizard-driven template able to capture phenotypic observations and all associated data to make them interpretable, whilst assuring compatibility with the GCP domain models. In June 2009, the PI and the programmer worked together for two days at Wageningen to finalise the template for its delivery in time for the 2009 ARM. Current activity consists of the programmer working remotely with the PI to debug the wizard using anonymised sample data sets provided by other GCP PIs. As the PI will not be able to attend the ARM in person, our intention is to present the wizard in the form of an animated video, to be fronted by Theresa Fulton. We will also prepare a poster to illustrate the features of the wizard. The video will also represent an appropriate user manual for the wizard.

97. G4008.32: Promotion of data quality management in genotyping laboratories, development of crop ontologies, comparative genomics for drought tolerance and evaluation and promotion of selection indices

January 2009–December 2009

Principal Investigator and Lead Institute
Guy Davenport, CIMMYT; g.davenport@cgiar.org
Apdo. Postal 6-641, 06600 México, D.F., Mexico

Collaborating institutes, Co-Principal Investigators and scientists

- CIMMYT: Jose Crossa; Trushar Shah; Rosemary Shrestha
- Central Science Laboratory: David Galsworthy

This project is a collection of related objectives that are are being carry out under the following objectives. A report for each objective is given below or is referenced as part of another existing project:

Objective 1: Development and implementation of a Data Quality Management at BecA
The immediate objective is the 'Promotion of a Quality Management System in BecA Laboratories'. However, under the wider objectives of the Generation Challenge Programme it aims to serve as a pilot for the promulgation of quality management systems amongst the GCP laboratories. It is thus expected that the outcomes of the project will have a major impact on the quality of the data produced in BecA laboratories and ultimately on the wider delivery of the GCP.

The following objectives have been identified:

a. To assess the present status of quality systems operating within BecA laboratories, benchmarked against ISO17025
b. To increase amongst BecA scientists awareness of the Principals of quality management in a laboratory environment, and
c. To develop the framework for a quality management system in terms of identification of 'process critical points', documentation required (eg SOPs) and the production of a plan for implementation.

Objective 2: A validated Crop Agronomic Traits Ontology for the GCP priority crops
Progress reported in G4009.03: Development of data standards and community of practice enabling the capture of and access to Generation Challenge Programme quality data sets

Objective 3 - Analysis of maize and other recently sequenced genomes in order to find candidate genes for drought tolerance
Progress reported in *G4006.08: Data analysis support for existing projects in SP2 with emphasis on integrating results across gene expression and QTL mapping experiments*

Objective 4: Biometrics and statistical support for selection indices and simulation of breeding programmes

The following selection indices (SI) are in development:
Phenotypic
(1) Smith
(2) Kempthorne and Nordskog
(3) ESIM (Eigen value Selection Index methods)
(4) RESIM (Restrictive Eigen value Selection Index methods)

Molecular marker
(5) Lande and Thompson – marker assisted selection SI
(6) Lange and Whittaker – genome wide SI
(7) MESIM. – marker assisted selection and genome wide SI (Molecular Eigen value Selection Index methods)

The SAS codes for the 7 SIs are finished and the programmes are up and running. When Molecular Markers are included the Lange and Whittaker and MESIM genome wide SIs were tried with up to 1000 molecular marker. The Penrose generalised inverse is used when matrices do not have standard inverse. The indices can be constructed based on variance-covariance matrices or correlation matrices. The outputs include the % of selected individuals with their trait values, the expected and realised genetic gains, the variance of the SI and the correlation between the SI and the breeding value. Each SI can be used in traits or in traits-environment combinations.

A complete user manual is currently being preparing in Spanish and English and will be ready by the end of October. Later a publication will be prepare for Crop Science with a brief theoretical back ground on the SI and with several examples showing how the SAS codes are used and the outputs are produced for each of the 7 SI. The SAS codes will be converted to R scripts for use with the user interface that is in development in objective 5.

Objective 5 - Development of user friendly tools for selection indices
The aim of this objective is to develop a user friendly user interface for the SI software developed in objective 4. Initial work has concentrated on the development of an R plugin for GenoMedium (http://www.genomedium.org). When the SI are availble in R, a user interface will be developed to run these R scripts using data queryed from the GCP informatics platform and combine the results of the R scripts with the input data in order to rank lines by selection index.

98. G4008.54: Implementation of a Molecular Breeding Platform

September 2008–June 2009

Principal Investigator and Lead Institute
Graham McLaren, Subrprogramme 4 Leader, Generation Challenge Programme, g.mclaren@cgiar.org
Apdo. Postal 6-641, 06600 Mexico, D.F., Mexico**,** Phone: +52 55 5804 2004

Collaborating institutes and scientists
N/A

This project aims to develop a plan, strategy and resources for a functional and sustainable molecular breeding platform providing breeders access to molecular breeding services, an information system and a toolbox of analysis and decision support applications. Such a platform will enable breeding programmes in the public and the private sectors to access well-characterised germplasm, modern marker technology and germplasm information in a simple and reliable way, to accelerate variety development in South Asia and sub-Saharan Africa, among other regions.

The project to implement this plan will build on existing GCP products and technology and work with existing MAB projects supported by the Bill & Melinda Gates Foundation (BMGF) and GCP, targeting improved crops for developing countries as initial use cases. The platform will address the needs of these projects in the first instance. Success with these examples will demonstrate the feasibility and effectiveness of the platform, creating incentive for other breeding projects of the same nature to use the platform.

A comprehensive project was developed and approved in July 2009. This 20 million dollar project funded by BMGF and core GCP donors will provide a one-stop shop for breeders wanting to implement molecular breeding projects and integrate them into their mainstream breeding programmes.

The preparation of the proposal was a consultative process over the past year. Consultations were held with individual user cases and private companies since summer 2008 and a one-day workshop was appended to the GCP Annual Research meeting in Bangkok in September 2008.

A planning workshop was organised by GCP in Montpellier in March 2009. The overall objective of the workshop was to provide a forum to learn about the latest achievements and strategies in molecular

breeding from a panel of world experts, and provide an opportunity for the platform user community to exchange views with colleagues from other projects, and with the scientists who will develop the different elements of the platform (http://www.generationcp.org/sp5/?da=09144305). This objective was fully met and the output of the discussions surpassed the expectations of the GCP Management Team. Major outputs were:

- A series of notes and recommendations made by the workshop participants on critical issues related to MB and the potential role and *modus operandi* of the platform.
- A 'contract' that describes the rules of engagement of users and the commitments of the platform management to work together.
- A detailed list of unmet needs formulated by the different use cases for data management, informatics, analysis and decision support tools and support services.
- Clear guidelines for the platform managers to revise the project proposal taking into account the outputs of the workshop.
- A vision of success for the MB platform: how will we know we have achieved platform objectives as given below.

At the end of the workshop the participants were asked to describe their vision of success for the Molecular Breeding Platform. A broad set of diverse response was provided that can be summarised in the following different criteria:

- The platform is adopted and the number of users is increasing over time because its facilities promote faster genetic gain at low cost.
- The platform attracts users from small and medium sized enterprises
- The number of molecular breeding projects in the public sector is increasing and the number of new varieties developed using facilities of the MBP is increasing
- User willing to release data to the platform database and share germplasm to add value to their own programmes.
- User willing to provide feedback and continue to be the development of different elements of the platform.
- Stakeholder communities are taking on the promotion of the platform in their respective crop/region communities and with donors

99. G4009.03: Development of data standards and community of practice enabling the capture of and access to Generation Challenge Programme quality data sets

January 2009–December 2009

Principal Investigator and Lead Institute

Elizabeth Arnaud, Bioversity; e.arnaud@cgiar.org

Bioversity International, via dei Tre Dinari, 472/a, 00057 Maccarese, Rome, Italy, tel: (39) 066118323

Collaborating institutes and scientists

- Bioversity: Adriana Alercia; Elizabeth Arnaud; Stephanie Channelière; Milko Skofic
- CIMMYT: Rosemary Shrestha; Guy Davenport
- CIP: Reinhard Simon
- Cornell: Chich-Wei Tung
- ICRISAT: Jayashree Balaji (until April 2009); Mike Butterfield (effective May 2009)
- IRRI: Thomas Metz; Ramil Mauleon; Warren Constantino
- OSU, Plant Ontology: Pankaj Jawal
- Robert Koebner (consultant)
- WUR: Theo van Hintum; Elizabeth van Strien

1. Research activities and progresses

Component 1: A validated Crop Agronomic Traits Ontology for the GCP priority crops The GCP ontology is renamed 'Crop Ontology' and will be registered on the OBO-Foundry. It can be browsed using the new GCP ontology Browser. A workshop was held in Bioversity, with the participation of Margherita Sini (FAO) and Chih Wei (Cornel). Collaboration was identified with AGROVOC, Thai Rice Ontology, Plant ontology consortium, SGN genomic network (Potato trait ontology), and Maize GDB. Wheat, Chickpea, and Sorghum validated terms were submitted with their definitions to the Gramene Trait Ontology (TO) and to Plant Ontology (PO). *Musa* terms on plant structure and definitions were validated by a botanist. For the ontology on rice mutant phenotype descriptor, 55 internal nodes were added for a total of 147 ontology terms and deposited in the GCP ontology SVN repository. 57 new mutant phenotypes, as observed in IRRI are now under evaluation for inclusion into the current ontology. 63 terms were collected for drought-related ontology from IRRI researchers. The controlled vocabulary terms describe experimental designs & environment, temporal traits, plant traits, and yield/biomass. Data curation and annotation using the Ontology, was initiated on wheat and maize in the International Wheat Information System (IWIS) and International Maize Information System (IMIS). Crop ontology will include assay/methods and scale. Two tools were tested for Ontology mining: MSWord-2007 Ontology add-in and Terminizer (http://terminizer.org/). A poster was presented at the Plant and Animal Genomics (PAG) XVII meeting in San Diego (January 10-14, 2009). The group presented a poster in the 3rd Biocuration Conference in Berlin that is posted in 'Nature precedings': http://precedings.nature.com/documents/3087/version/1

Component 2: GCP data templates for quality data capture. A version 0.5 of Genomedium validator that creates reports on the data format problems, is in development. The phenotyping wizard version 1.1 has been debugged on the basis of data sets sent by PIs.

Component 3: A Central Registry for the GCP data sets. The file upload function was made more obvious. The attribution of restricted access is now more flexible and can be allocated by the Principal Investigator (PI) only to its project's partners.

Component 4: A collection of comprehensive quality data sets. The analytical software applied in the Data Resolution Method was redeveloped and applied to SSR data sets quality checking. The work includes an extensive set of checks for the template and a quantification of the quality of the data. In most cases, the dataset had no value, unless improvement is provided because of non-corresponding sample-identifiers and germplasm-identifiers in the different sheets or lack of passport data. PIs of the projects or collaborators were contacted to get additional data or corrections but it was moderately successful. In the case of lentils, a new dataset has been prepared for uploading. Other data sets will be loaded into ICIS, enabling regeneration of the dataset from the database with unique germplasm identifiers. Datasets provided by different sources on the same crop, or different material with the same markers, were combined. The next step is to define a process of interaction with the Central Registry and how to interact with the PIs.

Component 5 Helpdesk

A standard procedure was established to handle support requests. The newly created mail address to collect the reports on automatic validation is: gcpcr.validation@generationcp.org. 33 Principal Investigators of the five Sub-programmes were contacted in the process of data chasing. Researchers were assisted to select the correct data templates, and received personalised guidance through the data set upload procedure on the Central Registry. This support led to successful uploading of their datasets. An overview of the validation reports was created.

Tangible outputs delivered

A new look-up service, called the GCP Ontology Browser is now online: http://ontology.grinfo.net/ontology-lookup/. OBO files for the Ontology are posted on Cropforge. The version 1.1 of the phenotyping data-entry wizard is debugged. The attribution of restricted access to project partners is available on the Central registry for testing. 14 SSR datasets for 6 crops have been checked for quality: barley, chickpea, finger-millet, foxtail-millet, lentil and maize. The quality check results are available in Excel files. Central registry homepage shows the number of files registered and

uploaded. In 9 months, PIs have registered additional 33 data files and posted 23 new data sets which is the result of the active data chasing by the helpdesk and subsequent assistance for the upload procedure.

References

Shrestha, Rosemary; Mauleon, Ramil; Simon, Reinhard; Balaji, Jayashree; Channelière, Stephanie; Alercia, Adriana; Senger, Martin; Manansala, Kevin; Metz, Thomas; Davenport, Guy; Bruskiewich, Richard; McLaren, Graham; and Arnaud, Elizabeth. Development of GCP Ontology for Sharing Crop Information. 3rd International Biocuration Conference, 17 April 2009 - Available from *Nature Precedings* <http://dx.doi.org/10.1038/npre.2009.3087.1> (2009)

100. G4009.04: Data analysis support for existing projects in SP2 with emphasis on analysis of next generation sequencing data

January 2009–December 2009

Principal Investigator and Lead Institute

Rajeev Varshney, ICRISAT; r.k.varshney@cgiar.org
International Crops Research Institute for the Semi Arid Tropics, Patancheru, Andhra Pradesh, 502324, India; (91) 40-30713305; Fax: (91) 40-30713074

Collaborating institutes and scientists

- ICRISAT: Vivek Thakur
- NCGR: Greg May; Andrew Farmer
- TSL: David Studholme; Jonathan Jones
- IRRI: Richard Bruskiewich

Background

SP2 projects using 454 FLX and Solexa sequencing technologies are beginning to generate EST and SNP marker resources in pigeonpea and chickpea that will help overcome a serious bottleneck in the development of these crops – namely shortage of markers and absence of genetic maps. NGS methods however generate a deluge of data; the shorter read lengths require considerable bioinformatics effort in assembly. Since the sequencing itself is being carried out at NCGR and uses the NCGRs computational pipeline; bioinformatics efforts at ICRISAT are related to putting together an open access, open-source alternative to the NCGR proprietary pipeline (consisting of *Alpheus* pipeline for Solexa data and *XGI* pipeline for EST data).

Efforts through 2009 will involve developing a protocol for the analysis of NGS data with the express objective of evaluating accuracy of tools in predicting SNPs, differential gene expression in a pair of genotypes. Further, the project also aims for completion of pipeline for variety of NGS data analysis applications, validation methods for markers generated using Illumina GoldenGate assay on identified mapping populations, development of genetic maps and computational prediction of miRNAs in the transcriptome data.

Findings and Implications

Based on simulations and other analysis, we had selected MAQ (1) and NOVOCRAFT (2) to be superior in several features (speed, flexibility in mapping/assembly, accuracy, etc.) for mapping/assembly of Solexa datasets generated for two genotypes i.e. ICC 4958 and ICC 1882. For mapping these Solexa transcript datasets, the transcript assembly developed based on ca. 400,000 454/FLX ESTs generated from a normalised pool of cDNA coming from >20 tissues was used.

For the SNP identification in these two genotypes, we evaluated MaqSNP filter, a SNP identification programme, for false discovery rate estimation. The analysis showed high rate of false positives identified from MaqSNP filter. On other hand, we found *de novo* assembly approaches (using Velvet) to be limiting in identification of SNPs from (the assemblies of) two or more genotypes. Given these limitations, we developed ad-hoc approaches for identification of *high confidence* SNPs, which involves comparison of assembly of one genotype, either against reference or against assembly of other genotype. The SNPs identified from above two tools and also from Alpheus (3) however showed poor degree of overlap across the tools, such that SNPs unique to each set is as high as ~50% of the total. Now experimental validation of SNPs generated from above three approaches is under progress.

The pipeline for NGS data analysis with the above open source tools is underway. We have a plan to develop the pipeline to allow the users to get the mapping done by Maq/Novo for at least two genotypes with variety of Solexa input file formats.

It is also important to mention that under the framework of this project, an International Workshop is being organised on **Next Generation Sequence (NGS) Data Analysis** during 21-23 July 2009 at ICRISAT, Patancheru. More than 30 participants including 12 from overseas (3 from UK, 2 from Korea, 2 from France, 1 from US, 1 from Australia, 1 from Japan, 1 from CIMMYT and 1 from IRRI), 11 from India and 10 from ICRISAT will participate in this workshop. The main objective of the workshop is to discuss on developing the analytical tools, platforms and infrastructure so that the genomics community may overcome the data analysis bottleneck of NGS.

Next steps and/or challenges
The next steps involve completion of some of the ongoing development work/analysis. The completion of experimental validation of SNPs identified from three tools/approaches will provide a basis for preferring a mapping/assembly tool over another. The NGS data analysis pipeline requires to include components for SNP identification, conversion of SNPs to CAPS by using SNP2CAPS tool as well as assessing the quality of SNPs for designing the OPA assays for developing the Illumina SNP genotyping platform. For visualisation of assembly the GBrowse configuration needs to be completed so that it can be used for displaying reference, mapped NGS reads along with score, no. of mismatches, mapping quality and SNPs.

References
Li H, Ruan J and Durbin R (2008). Mapping short DNA sequencing reads and calling variants using mapping quality scores. *Genome Res*. 18(11):1851-8.
Hercus C (2008). Manual of NovoCraft short read alignment package. URL: http://www.novocraft.com/downloads/downloadpage.php
Neil AM, Andrew F, Stephen FK, Raymond JL, Faye DS, et al. (2008). Management of high-throughput DNA sequencing projects: Alpheus. *J. Comp. Sci. Syst. Biol.* 1: 132-148.

101. G7009.03: Rice Challenge Initiative start-up project
June 2009–August 2009

Principal Investigator and Lead Institution
Graham McLaren, GCP; g.mclaren@cgiar.org
Apdo. Postal 6-641, 06600 Mexico, D.F., Mexico**,** Phone: +52 55 5804 2004

Collaborating scientists and institutions
WARDA: Marie Noelle Ndjiondjop

Executive summary
The objective of this project is to establish the Rice Challenge Initiative by convening a project development workshop and commissioning small activities to start the research process in 2009. The main proposal(s) will start in 2010.

Objectives and outputs
Develop proposals for the Rice Challenge Initiative and commission small activities necessary to ensure the start of the main projects is not delayed

Rationale
The Rice CI team is new and needs an opportunity to come to a group understanding of the project and to develop coherent and effective set of proposals to achieve overall goals of the Challenge Initiatives.

Activities
Project development meeting was held at WARDA in June 2009. Other activities as may be deemed essential to start the Initiative.

Subprogramme 5: Capacity-building and enabling delivery

102. G4005.53: The use of molecular markers in efficient crop improvement: Marker-Assisted Breeding Learning Module

August 2007–July 2008

Principal Investigator and Lead Institute
Theresa Fulton, IGD–Cornell; 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

A training module on the topic of Marker-Assisted Breeding has been completed. This module was developed as a complement to other GCP modules completed or in progress, including those on molecular markers in plant diversity, phenotyping, genomics, and bioinformatics, creating a "bookshelf" of material available for scientists world-wide. The targeted audience is plant breeders in developing countries; it is assumed the reader has a basic background in genetics and plant breeding, although key background points are reviewed. The module has been developed in such a way as to be useful either as a self-tutorial, or as the basis of a training course. It will be freely available both online and as a CD-ROM. The module was reviewed by 3 reviewers (2 selected by GCP), revisions made, is now being re-reviewed, and final revisions will be made upon receiving last comments.

The module (165 total slides) includes the following sections:

1: Introduction (18 slides)
2: Selection of markers for MAB (27 slides)
3: Genetic diversity and germplasm selection (22 slides)
4: Tips for phenotyping in MAB (20 slides)
5: Genetic linkage mapping (32 slides)
6: Quantitative trait analysis (30 slides)
7: Applications and future directions (12 slides)
8: Resources (4 slides)
9: References
10: Glossary
11: About this module

103. G4005.63: Interactive Resource Centre & Helpdesk

August 2007–July 2009

Principal Investigator and Lead Institute
Theresa Fulton, IGD–Cornell; 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 Centre & Helpdesk 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, learning modules, literature and other general resources, such as information on writing proposals. Freely available data is also available for download. Also posted are key links, including funding opportunities, 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 (specialising in molecular markers, population genetics, plant breeding, genetic diversity, etc.). Recently we have been coordinating with the new GCP Bioinformatics helpdesk to better direct users' questions to appropriate assistance, as well as working with the new Toolbox. The number of users continues to grow each year, with over 4000 unique visitors to the site so far this year.

Recently an associated Facebook group has been started, to better facilitate discussion and postings. This group is open to anyone; you must first register with Facebook (which is free); then join the group "Resource Centre for global plant scientists". Please join us!

104. G4006.13: Targeting and impact analysis of Generation Challenge Programme Technologies

January 2007–December 2008

Principal Investigator and Lead Institute
Glenn Hyman, CIAT; g.hyman@cgiar.org
AA 6713, Cali, Colombia (57-2) 445-0000; Fax: (57-2) 445-0073

Collaborating institutes and scientists

- CIAT: Sam Fujisaka; Peter Jones
- CIMMYT: John Dixon
- IFPRI: Stanley Wood

1. Project summary

The GCP commissioned this research to support their priority-setting and strategy development activities. The research team developed global data sets and tools to help the GCP decide which crops and regions to focus on. The project assesses poverty, drought and farming systems in the context of GCP goals and objectives. A database was developed for further use by the GCP community.

2. Activities and progress

Poverty profiles were developed for 15 GCP farming systems. These profiles draw on global maps of per capita gross domestic product, stunted children, underweight children and infant mortality. The profiles also include detailed country data where available.

The project produced a global drought probability map, based on the "failed seasons" model. A second objective was to conduct crop-specific drought mapping. Our initial effort utilised a crop modeling approach, which turned out to be cumbersome and unreliable. As an alternative, the Seasonal Drought Index (SDI) was developed to assess drought during 20-day periods after planting (Figure 1). For 18 km pixels, the maps show the proportion of days where actual evapotranspiration (Ea) divided by potential evapotranspiration (Et) is less than a threshold value that corresponds to crop water requirements. Ten 20-day periods after planting date were mapped for water stress thresholds (Ea/Et) of .3, .35, .4 and .45, producing a total of 40 global maps. We are presently asking GCP scientists to evaluate this model in the context of the crops and regions where they focus their efforts.

A climate change assessment was carried out to evaluate potential changes in temperature and precipitation in GCP priority farming systems. Algorithms calculated averages, standard deviations and model congruence for 18 Intergovernmental Panel on Climate.

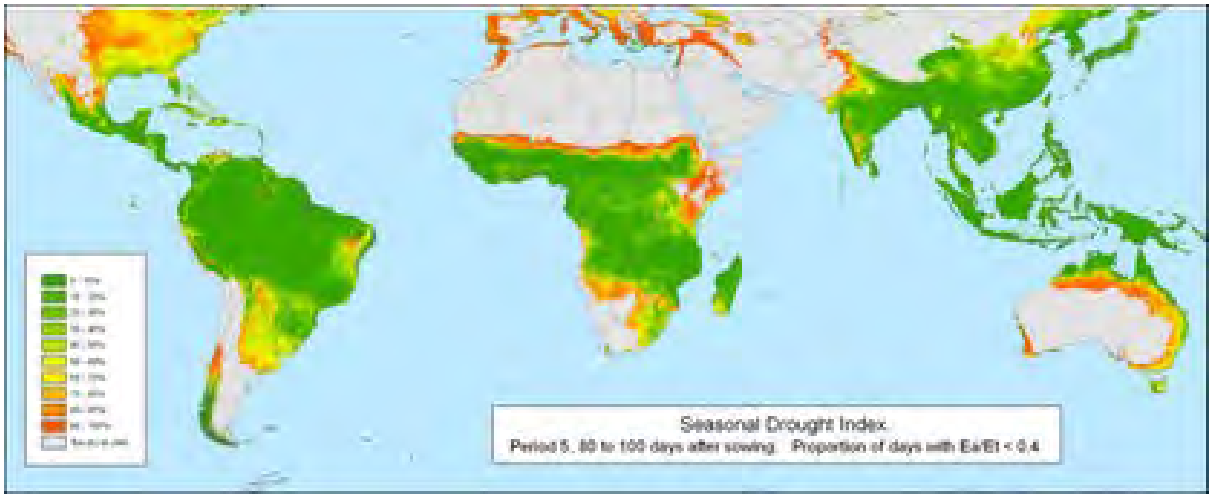


Figure 1. Seasonal drought index: Period 5, 80 to 100 day after sowing, proportion of day with Ea/Et < 0.4

Change (IPCC) models at the pixel level and for entire farming systems. Table 1 below shows current and future average temperatures and expected change. The analysis agrees with some other climate change studies arguing the importance of heat tolerance in crops. Highland and temperate farming systems show the greatest temperature changes. Temperatures in the Sahel systems *agro-pastoral millet/sorghum* and *cereal-root crop mixed* are expected to increase substantially.

Table 1. Temperature and changes by farming system for key GCP priority areas.

Farming system	REGION	Current Temperature	2050 Temperature	Temperature Change
Temperate mixed	EAP	6	9	2.8
Highland mixed	SA	14	17	2.7
Agro-pastoral millet/sorghum	SSA	26	28	2.5
Highland temperate mixed	SSA	18	20	2.4
Rice-wheat	SA	25	28	2.4
Cereal-root crop mixed	SSA	26	28	2.3
Upland intensive mixed	EAP	15	17	2.3
Maize-beans (Mesoamerica)	LAC	21	23	2.2
Root crop	SSA	24	26	2.2
Maize mixed	SSA	22	24	2.2
Rainfed mixed	SA	26	28	2.2
Dry rainfed	SA	26	29	2.2
Lowland rice	EAP	21	23	2.1
Highland extensive mixed	EAP	17	19	2.1
Rice	SA	27	29	1.9

Data Repository and Query

The data sets developed for this project have been made available as a Web site and as an ACCESS database. The Web site utilises Google Groups. More than 50 tables are available on crop production, micronutrient content of staple crops, population, childhood stunting, infant mortality, market accessibility, length of the growing period, soil constraints and climate change. GCP scientists interested in these data should contact Glenn Hyman at CIAT.

105. G4006.14: Ex ante impact analysis of marker-assisted selection technologies supported by the Generation Challenge Programme

December 2006–December 2008; no-cost extension to April 2009

Principal Investigator and Lead Institute

George W Norton, VPI; gnorton@vt.edu

Virginia Polytechnic Institute and State University, Department of Agricultural and Applied Economics, 205B Hutcheson Hall, Blacksburg, VA, USA (1) 231 7731; Fax: (1) 540 231 7417

Collaborating institutes and scientists

- VPI: Jeffrey Alwang
- IRRI: Abdelbagi Ismail
- CIAT: Martin Fregene

1. Context

Ex ante impact analysis was used to estimate benefits of GCP investments and to validate an approach to impact assessment. Two GCP projects: "Revitalising 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 used as case studies for the impact assessment. Two students completed Masters theses in 2008 (Alpuerto, 2008 and Rudi, 2008) and a third student (Pricilla Marimo) completed her Masters thesis in June 2009. She developed and applied 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. 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. These results were reported on last year and are summarised in one paper with the rice results which has been submitted with second revisions to a refereed journal (Alpuerto, et al, 2009) and in a second paper with the cassava results which is under revision for second refereed journal.

A survey of 200 cassava producing households was completed in Nigeria and a Probit analysis was conducted to assess potential impacts by gender of improved cassava varieties developed on the GCP. Specifically, the issue of labor use and the effects of increased income on household decision making were examined. It was found that households that adopt cassava varieties with improved insect and disease resistance allocate more female labor to cassava production, processing, and marketing than do non adopting households. One reason appears to be the concurrent expansion of female labor for planting, fertiliser application, weeding, harvesting, processing and marketing of cassava that occurs with the now more profitable crop. One implication is that improved varieties may increase pressures for mechanisation of some of the tasks completed by women in cassava production, harvesting, and processing. There is little change in male labor use. However, women in adopting households relinquish some control over decision making to men with respect to input purchases, labor allocation, and borrowing when improved varieties are adopted. It is not clear why this change occurs.

3. Next steps

Explore avenues for reducing constraints to rapid adoption of improved cassava varieties and for evaluating impacts of nutritionally enhanced cassava. We are revising the cassava paper based on reviewer comments from the African Journal of Agricultural Economics and are working on a journal article manuscript out of the gender analysis.

4. References

Alpuerto, V.L (2008). *Economic Analysis of Marker Assisted Breeding in Rice*. MS thesis, Virginia Tech, June.

Alpuerto, V.L., G.W. Norton, J. Alwang, and A. Ismail (2009). Economic Impact Analysis of Marker-Assisted Breeding for Tolerance to Salinity and Phosphorous Deficiency in Rice. paper submitted with second revision to *Applied Economic Perspectives and Policy.*

Marimo, P. (2009). *Gender Impacts of Molecular Assisted Breeding: The Case of Insect and Disease Resistant Cassava in Nigeria,* 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, August.

Rudi, N., G.W. Norton, J. Alwang, and G. Asumugha (2009). Economic Impact Analysis of Marker-Assisted Breeding for Resistance to Pests and Post Harvest Deterioration in Cassava, paper submitted to *African Journal of Agricultural Economics*.

106. G4006.36: Capacity-building and research project

October 2006–December 2009

Principal Investigator and Lead Institute

Mark Laing, ACCI–UKZN; laing@ukzn.ac.za

African Centre for Crop Improvement, University of Kwa Zulu Natal, PBX01, Scottsville, Pietermaritzburg 3209 South Africa

Collaborating institutes and scientists

- ACCI–UKZN: Jedidah W. Danson
- CIMMYT–Kenya: Dan Makumbi
- ACCI students
- NARs breeders in 14 countries

Research activities and progresses

The main objective was to provide a teaching position for a biotechnologist to teach ACCI students at the University of Kwa Zulu Natal (UKZN). This was achieved in November of 2008. The scientist would introduce biotechnology tools to breeders from 14 different countries in Africa that are trained at the ACCI. Emphasis to be on marker assisted breeding. This would involve developing and optimisation of a cheap, simple and reliable system for breeders to sample DNA and package it for analysis to the GSS centres. Secondly, carry out small scale marker validation for different traits of breeder's interest.

Assist in developing genotyping centres

The aim was to collect materials from CIMMYT breeding program and sharfe with the ACCI breeders after MAS analysis for QPM and MSV traits with markers. To-date a total of 4500 samples from S1 and S3 populations collected from MSV, QPM and drought tolerant CIMMYT nurseries in Kenya are undergoing analysis. The handi-cap as mentioned above is the FTA optimisation using the Rotor-gene. The primers aligned for use for MSV and QPM had been optimised on the ABI 3100 sequencer. The PCR product was visualised on an agarose gel. But was found to be labor intensive for a large number of samples. Therefore the need to optimise the FTA with the Rotor-gene remains the best option.

Assist in institutional support of biotechnology for 6 African Universities

Focus on SADC countries. The capacity in Kenya, Uganda and Tanzania is relatively good. The institutions also have close proximity to BecA (soon may be operational). Other countries found to be relatively capable of extracting DNA are Zambia and Zimbabwe. The labs should be capable of DNA extraction and gel resolution to test on quality and quantification. The approach is to combine with objective 6 and take inventory of the capacities in these universities plus possible MAS applications and interest by breeders.

Assist in developing a dynamic inventory of biotechnology capacity in southern and eastern Africa

There is no question breeders are interested in MAS especially in genotyping, diversity studies breeders rights etc. However, the availability of individual researchers to conduct this analysis is limiting. Therefore, objective 5 forms a critical part in how objective 6 is implemented.

Tangible outputs delivered

The 1st and 2nd PhD have already received lessons in Biotechnology tools. However, no practicals have been conducted yet due to equipment and space limitations. Being a relatively new subject, supervision of students is yet to be established. However, genotyping of students germplasm is set to start. All the phenotypic data for the 4500 maize samples has been completed. However, due to cost and volume, the samples have been reduced to 2000 by selecting only those that scored for resistance to MSV phenotypically. The presence of the region conferring resistance to maize streak virus disease will be confirmed by markers. This is on going. The best lines identified with markers and phenotypic data (already collected) will be used to improve germplasm for Mozambique (low-land) and the mid-highland areas. The FTA paper technology optimisation has been finalised. It is now possible to extract DNA from the FTA and run the samples on a Rotor-gene which initially showed interference from the FTA paper disks. The 1400 rice samples are being prepared for shipping to a GSS centre for MAS. The rice materials are being improved for the low shattering trait using markers developed by Prof Susan McCouch of Cornell University.

107. G4007.03: Community of Practice 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; no-cost extension to April 2009

Principal Investigator and Lead Institute

Theerayut Toojinda, BIOTEC; theerayut@dna.kps.ku.ac.th
113 Thailand Science Park, Phaholyothin Road, Klong 1, Klong Luang Pathumthani 12120 Thailand; (66-2) 5646700; Fax: (66-2) 5646705

Collaborating institutes and scientists

- RGDU: Jonaliza L Siangliw
- DAR: Toe Aung
- NAFRI: Monthathip Chanphengxay
- CARDI: Men Sarom; Ouk Makara
- UBU: Sureeporn Katengam

Line conversion of popular rice varieties in Mekong Region: MAS, trait validation and preliminary field trials

Application of biotechnology like marker-assisted selection (MAS) in rice breeding had been proven effective in Thailand. This technique has been used in line conversion and gene pyramiding in Khao Dawk Mali 105 (KDML105) and RD6 backgrounds. The traits incorporated include resistance to leaf and neck blast, bacterial leaf blight, brown plant hopper, submergence, salinity and drought. Rice Gene Discovery Unit (RGDU), BIOTEC had the opportunity to convey knowledge to its neighbouring countries through trainings thus molecular breeding of rice in Mekong Region begun when long term training on MAS was initiated in 2004 as sponsored by Rockefeller Foundation, BIOTEC and Kasetsart University. This training aims to develop popular rice varieties like CAR3, TDK1, Manawthukha, IR53936 and IR57514 from Cambodia, Laos, Myanmar and Thailand, respectively which lack traits that may improve quality and adaptation in the local area. In the previous training, plant materials that were developed have not reached the final target. Under the Generation Challenge Programme, the lines initially developed had the chance to reach field trial testing and further trials will be initiated in the future to hasten the release of improved varieties.

Table 1 shows the products of line conversion through MAS. After the line conversion, these materials were self pollinated to produce seeds for trait validations and field trials (Table 2). Improved varieties from Myanmar, Laos, Cambodia and Thailand were produced. A paper on line conversion of Manawthukha had been published in Field Crops Research 113 (2009) 178-186. All materials will be tested further in the field particularly in target locations where these popular rice varieties are planted in each country. Farmer's participation in selection will be applied as initiated by each institute and after farmer's field trials, the best line will be selected for release.

Table 1. Popular rice varieties in Mekong Region introgressed with QTL/gene using marker-assisted selection.

Institute/Trait	Recipient	Donor	Selected Lines	Markers
DAR/Aroma and intermediate AC	Manawthukha	Basmati	12 BC4F2	BADH, Waxy, RM204
DAR/salinity tolerance	IR53936	Pokkali	56 BC3F4	SalT, RM472, RM3412, RM10720, RM1287, RM10772
NAFRI/aroma and low GT	TDK1	Homnangnuane	21 BC3F2	BADH, RM21, RM5349
CARDI/aroma and soft GC	CAR3	Pkha Rumdoul	46 BC3F2	BADH, RM587, RM589
UBN/aroma, low AC, low GT and Xa21	IR57514	KD571-77	380 BC3F5	BADH, Waxy, SNP2341, PB7/PB8

Table 2. Trait validations and preliminary field trials of BILs of popular rice varieties in Mekong Region.

Institute/Trait	Traits	BIL Selected
DAR/Aroma and intermediate AC	Aroma – Sensory and 2AP concentration OYT – 4 locations in Myanmar and 1 in Thailand; grain yield and flowering	12 BC4F3
DAR/salinity tolerance	Salinity – Yoshida and pond screening OYT – 3 locations in Myanmar; grain yield and salt injury score	14 BC3F5
NAFRI/aroma and low GT	Aroma – Sensory and 2AP concentration Agronomic characterisation in NAFRI, Laos	21 BC3F3
CARDI/aroma and soft GC	Aroma – Sensory and 2AP concentration OYT – Irrigated and drought stress conditions in CARDI, Cambodia; grain yield and spikelet sterility	15 BC3F3
UBN/aroma, low AC, low GT and Xa21	Aroma – Sensory and 2AP concentration Submergence screening in RGDU, KU; percent plant survival Drought screening in Chumphae, Khon Khaen, Thailand; grain yield and spikelet sterility OYT – RGDU, KU; grain yield	10 BC3F5

OYT – Observation Yield Trial

108. G4007.13: Capacity-building *à la carte* 2007

Principal Investigator and Lead Institute
Carmen de Vicente, GCP; 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 Capactity-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 Institute: KARI)
- *Marker-aided development of Nutritionally-enhanced cassava for Nigeria* (Lead Institute: NRCRI)
- *Application of molecular tools for controlled wild introgression into cultivated germplasm in Senegal* (Lead Institute: ISRA)
- *Characterisation of maize germplasm found in Ghana, using the bulking technique* (Lead Institute: CSIR, Ghana)
- *An integrated proteomics and genomics approach to discover salt tolerance genes* (Lead Institute: ABRII)
- *Enhancing capacity of ICABIOGRAD in phenotyping and molecular analysis to develop elite rice lines suitable to Indonesian uplands* (Lead Institute: 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.

108.01 G4007.13.01: Capacity-building à la carte 2007−Capacity building for characterising maize for water stress tolerance at KARI−Katumani

May 2008–February 2009

Principal Investigator and Lead Institute
James G Gethi, KARI; jgethi@wananchi.com
P.O. Box 340-90100 Machakos,Kenya +254 721831166 Fax +254 2 2338320

Collaborating institute and scientists
Agropolis–INRA: Francois Tardieu; Claude Welcker

1. Research activities at KARI
Inbred lines from a diverse background phenotyped in the first year of the project were test crossed to two testers, CML 312 and CML 204. In addition to these materials, other inbred lines were test crossed. The resulting hybrids were very early and seed set was not very good due to problems with synchronised flowering. The inbred lines that produced seed were evaluated under well watered conditions and under random drought. The hybrids under random drought dried due to severe drought. The hybrids from the selected lines and CML 312 under well watered conditions in the dry season of 2008-9 at Kiboko.

Table 1. Evaluation of testcross hybrids developed from a subset of 60 inbred lines using CML 312 under stress and non stress conditions at Kiboko 2008-2009

Entry	DTF	DTS	ASI	Plant Ht	Ear Ht	Stand harv	Grain Yield
	Days	days	Days	cm	cm	No.	Tons/ha
RF-291-10-5-2-5/CML312	51.3	52.3	1.0	127.7	70.3	39.0	5.0
RF-295-2-1-1-4/CML313	59.3	62.0	2.7	202.0	89.3	39.3	5.5
RF-62-5-3-1-1/CML314*	57.0	58.0	1.0	208.0	98.0	11.5	6.5
RF-291-8-3-4-9/CML315	59.7	62.0	2.3	202.7	81.0	37.0	7.0
RF-291-3-10-11-1/CML316	59.3	62.0	2.7	191.0	84.3	37.0	7.0
RF-195-4-6-1-7/CML317	57.3	59.3	2.0	174.7	73.7	36.0	7.0
RF-295-2-1-1-2/CML318	59.3	61.3	2.0	191.0	80.0	38.3	7.7
RF-5-3-4-1-5/CML319	58.0	59.7	1.7	179.7	76.7	39.3	7.9
RF-291-3-9-6-2/CML320	58.7	62.3	3.7	169.7	72.3	37.3	8.3
RF-291-3-7-5-8/CML321	60.7	63.3	2.7	211.3	96.3	38.3	8.5
RF-291-3-10-15-2/CML322	59.0	62.0	3.0	188.7	81.7	39.3	8.6
RF-195-4-6-1-5/CML323	54.7	56.7	2.0	172.7	73.3	39.3	8.2
RF-295-2-1-1-5/CML324	58.3	60.7	2.3	188.9	85.0	39.3	8.9
RF-291-10-7-1-4/CML325	57.7	59.3	1.7	203.7	89.0	38.7	9.2
RF-5-3-4-1-4/CML326	57.7	60.0	2.3	171.7	74.3	39.0	9.5
RF-292-3-6-4-2/CML327	58.3	59.7	1.3	222.3	94.7	38.0	9.1
RF-291-10-5-3-3/CML328	55.7	57.7	2.0	192.7	87.3	38.0	9.8
RF-292-3-6-4-6/CML329*	57.0	58.5	1.5	192.0	80.0	12.0	11.0
DH02	52.9	54.8	1.9	169.9	78.5	28.7	4.9
DH04	61.7	63.4	1.8	200.9	105.0	30.3	7.0
DK8031	60.9	62.8	1.9	205.1	99.8	27.6	8.1
PH3253	62.5	65.2	2.7	202.0	102.5	29.8	6.4
Means of checks	59.5	61.6	2.1	194.5	96.4	29.1	6.6
Overall mean	57.5	59.5	2.0	187.8	84.6	29.3	8.1
CV (%)	2.0	2.1	34.0	6.1	9.2	8.6	17.6

*Single row plot

The same lines were test crossed to CML 204 and the resulting hybrids are being evaluated in the field under well watered conditions.

2. Training

16 technical staff in the maize programme were trained in drought trials management, data collection equipment that included chlorophyll meters (SPAD), infra red thermometers and GreenSeekers. This will enable the staff to be more efficient in data collection and handling in the trials this season. Data analysis scheduled for INRA in the second year of the project will take place as soon as possible during the off-peak period.

108.02 G4007.13.02: Capacity-building à la carte *2007–Marker-aided development of nutritionally-enhanced cassava for Nigeria*

July 2007–July 2009

Principal Investigator and Lead Institute

Chiedozie Egesi; NRCRI; cegesi@yahoo.com
National Root Crops Research Institute, Umudike, Nigeria; Email; Tel: +2347034962100

Collaborating institutes and scientists

- NRCRI: Emmanuel Okogbenin; Ada Mbanaso; Nnamdi Eke-Okoro; Khaya Shuaibu; Oluwakemi Ogundapo; Samuel Baiyeri
- CIAT: Martin Fregene

1. Research activities and progress
One NRCRI research staff visited CIAT's Cassava Genetics Laboratory for training in marker-assisted selection and post-flask management of *in vitro* plants. Ms Princess Onyeyirichi Onyegbule spent six weeks learning relevant techniques for molecular breeding with different scientists at CIAT.

This is brings to 3 the number of scientists trained in marker-assisted selection in cassava at CIAT through the project. The urgent need for enhancement of manpower for molecular breeding in relevant breeding programmes in NRCRI is very critical for the application of modern tools in agriculture for food security, poverty alleviation and income generation in Nigeria.

The two NRCRI research fellows 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 nearing completion of their programmes. 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 in the process of data analyses and writing-up of their thesis.

2. Field establishment and multiplication of the introduced germplasm
The 138 genotypes of cassava mosaic disease (CMD) resistant Latin American cassava introduced from CIAT were hardened and transferred to the field for initial assessment and multiplication for a replicated yield trial. These plants are growing in the field and shall be harvested in due time for evaluation of field performance and determination of beta carotene and protein contents.

Of the 138 genotypes of cassava combining high beta carotene and protein contents, 60 were sub-cultured and shipped to NARs partners in Ghana, Tanzania and Uganda as part of germplasm exchange efforts in a new network of African cassava breeders.

Tangible outputs delivered
138 CMD resistant, high beta carotene and protein Latin American cassava genotypes growing in the field in Nigeria for initial assessment and multiplication. A subset of these genotypes were micro-propagated and shipped to Ghana, Tanzania and Uganda for evaluations by partner NARs. One NRCRI research staff was trained on-hand in molecular breeding in an advanced laboratory. Also two other NRCRI research fellows are continued their MSc - degree training in molecular breeding in 2 Nigerian universities.

Next steps/challenges
The genotypes selected after an initial assessment for resistance to CMD shall be evaluated for the traits of interest (beta carotene and protein content) as well as other important agronomic characters in a replicated yield trial in 2 agro-ecologies in Nigeria. Genotyping for resistance CMD, and marker assisted selection for beta carotene and protein contents shall be done. Crosses with local elite germplasm shall be done with selected germplasm to derive the best of local and introduced gene combinations. Participatory evaluation of selections by farmers shall be conducted.

108.03 G4007.13.03: Capacity-building à la carte *2007–Application of molecular tools for controlled wild introgression into cultivated germplasm in Senegal*

July 2007–July 2009

Principal Investigator and Lead Institute
Ousmane Ndoye, ISRA; ousndoye@refer.sn
ISRA–CNRA B.P. 53 Bambey, Senegal

Collaborating institutes and scientists

- Agropolis–CIRAD: Jean-François Rami
- UCB: David Bertioli
- EMBRAPA: Marcio Moretzsohn

Groundnut is an important crop of the Sahel Zone of Africa. It is a cash crop as well as a major source of dietary proteins and oil, and also a source for animal feeding. Groundnut cultivation in this area 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.

During a previous project supported by the GCP, two amphidiploid varieties (*A.ipaensis* x *A. duranensis* and TxAG6) have been transferred to ISRA / CERAAS and each of them have been crossed to four different *A. hypogaea* cultivars from Senegal to produce backcross populations. Populations derived from crosses of this type segregate strongly for many traits. However, considering the nature of the parental cultivars and breeder's priorities in Senegal, investigation of components of drought tolerance, resistance to leaf spot and seed dormancy have been given top priority.

The main objective of this project is to allow the best use of the molecular tools developed earlier in order to optimise the development of breeding material for these priority traits, from the populations. Introgression lines were then developed from the available material using MAS. This requires the use of integrated genotyping at each step of the breeding process. To achieve this goal, we propose to build on the ISRA/CIRAD/EMBRAPA collaboration to ensure capacity building to PhD students and scientists involved in peanut breeding at ISRA and provide technical backstopping at the key steps of the breeding process for all activities related to MAS.

The BC1 population developed with the variety 73-30 as seed dormant cultivated parent 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 between Fleur11 and 73-30, using controlled seed lots of the parents. During the F1 to F2 generation, individual F1 plants were controlled to select plants to be self-pollinated, based on their conformity to the parental lines.

A 3-month stay of the PhD student allowed to genotype the F2 population with SSR markers.

Percentage of true F1 hybrids identified with SSR markers

Of seventy eight (78) putative F1 individuals tested, thirty eight (38) were true hybrids, corresponding to a percentage of 48%. In the view of the percentage of true F1 hybrids, it seems to be reliable to use molecular markers to discard inadvertent selfed seeds when managing cross progenies from intrasubspecific crosses in peanut. That average percentage of true F1 hybrids found in this study was lower than the percentage (60-70%) observed by Gomez et al. (2008). However, percentage of true hybrids in self-pollinated crops depends on the female parent used and climatic conditions during hybridisation.

Inheritance of fresh seed dormancy

All the seeds of the non-dormant parent (Fleur 11) germinated within 11 days following the test start while no seed did germinate for the dormant parent (73-30).

The chi square test performed on F2 population derived from true F1 hybrids, assuming 3:1 dormant: non-dormant ratio was not significant (P=0.08), indicating that the trait is controlled by a single dominant gene. Furthermore, the frequency distribution curve of fresh seed dormancy was bimodal, indicating that the trait can be treated in a qualitative fashion. These findings agreed with those previously reported by Upadhyaya and Nigam (1999) and recently by Asibuo et al. (2008).

Polymorphism level
From a previous work 558 SSR markers were screened for polymorphism between parents (Fleur 11 and 73-30). Among them 116 markers were polymorphic between the parents. All these markers were mapped (Fonceka et al. 2009, submitted). The percentage of parental polymorphism was 4.8% in this cross. This low polymorphism detected in this work may be due to the fact that both parents belong to the Spanish type. This is in agreement with previous findings on low level of polymorphism among cultivated groundnut (Halward et al. 1991; Stalker et al. 1994; Hopkins et al. 1999; Mace et al. 2006).

Bulk Segregation analysis
The 116 markers that were polymorphic between the parents are being surveyed on bulked DNA from extremes phenotypes of fresh seed dormancy along with ten individual DNA from dormant and non-dormant individuals. The work is ongoing at CIRAD to find molecular markers linked to the trait under study.

108.04 G4007.13.04: Capacity-building à la carte *2007–Characterisation of maize germplasm found in Ghana using the bulking technique*

July 2008–July 2009

Principal Investigator and Lead Institute
Allen Oppong, CSIR–Crops Research Institute, Kumasi, Ghana
Email: alnopp@yahoo.cio.uk; a.oppong@cropsresearch.org; Tel. +23324483`3895, +2335160391

Collaborating institutes and scientists
- CHPRRU, USDA–ARS: Marilyn Warburton
- CIMMYT: Yunbi Xu; Claudia Bedoya
- UdR: Jorge D Franco

Supporting scientists
- CSIR/CRI–Ghana: Manfred Ewool; Maxwell Darko Asante; Ruth Thompson

Maize streak virus disease and drought are major constraints to maize production in Ghana as well as other parts in Africa. Over the years research efforts have been made to improve the yield as well as the nutritional quality of the crop in Ghana with the CSIR-Crops Research Institute playing a dominant role in these efforts. These efforts led to the development of several improved varieties with high yield potential, improved nutritional attributes and better resistance to local diseases particularly maize streak virus disease. These improved varieties were adopted by many farmers. Over time, however, farmers started to complain about the storage qualities of these improved varieties. This led to a situation where some farmers went back to grow their local landraces which they consider to store and taste better. Some of these landraces are highly susceptible to the MSV disease. Also, most of the landraces have long maturity periods and are, therefore, exposed to higher disease infection rates and late season droughts. Other accessions are early maturing and with some good resistance to diseases. In 2007 the CSIR-Crops Research Institute with the support of CIMMYT and GCP initiated this project with the aim of identifying and characterising as much as possible the maize germplasm/landraces found in Ghana with a very efficient but cost effective technique-bulk fingerprinting (Dubreuil et al., 2006) with the objective to identifying sources of MSV resistance from the local landraces. In the process, the institutional capacity of the CSIR-Crops Research Institute will be boosted with the purchase of some basic equipment to improve its equipment needs of the Biotechnology laboratory as well as train its human resource capacity in advanced laboratories outside Ghana.

Towards the realisation of the objectives of this project, 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 savannah in the South-East, High Forest in the South-West and the middle zone, Forest transition in mid-north and the Guinea savannah in the northern parts of Ghana with some few collections from the Republic of Togo.

Field phenotyping of the Ghanaian landraces started in June 2008, which coincided with minor season rain. About 580 landraces that were collected from various parts of Ghana were planted in about 1.5 hectare land at Fumesua near Kumasi, within the forest ecozone. Not all the germplasm earlier collected germinated at the time of planting as a result of pest infestation and damage. A second planting this time during the major raining season in Ghana is currently in progress.

Phenotypic data with reference to the Bioversity International maize descriptors (1991) were collected. Phenotypic data for the first planting is ready and will be deposited at the GCP central registry and then analysed.

Genotyping using the bulk fingerprinting technique is currently on-going at CIMMYT and as at the time of reporting 12 SSR markers have been analysed using the bulk fingerprinting technique involving 24 populations and data will again be deposited at the GCP central registry and again analysed for the creation of core subset.

Marker assisted selection using 12 MSV resistant markers have just began at the Biotech laboratory of the CSIR-Crops Research Institute to complement what is currently being done at CIMMYT. This activity will also build local capacity for the challenges ahead.

Various equipment to be procured as part of this project started last year and is still on-going. A scanner with a printer, and a desk top computer were purchased. Processes have been initiated for the acquisition of a 17 Hp irrigation pumping machine and other laboratory equipment.

References

Dubreuil P., M. Warburton, M. Chastanet, D. Hoisington, and A. Charcosset. 2006. More on the introduction of temperate maize into Europe: large-scale bulk SSR genotyping and new historical elements. *Maydica* 51:281-291

Descriptors for Maize, 1991. *International Board for Plant Genetic Resources (IBPGR); International Maize and Wheat Improvement Center* (CIMMYT)

108.05 G4007.13.05: Capacity-building à la carte *2007–An integrated proteomics and genomics approach to discover salt tolerance genes*

July 2007−July 2008

Lead Institute

Ghasem Hosseini Salekdeh, ABRII

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 institutes and scientists

- IRRI: Abdelbagi M. Ismail
- IPK: Mohammad Reza Hajirezaei

We performed proteome and metabolome analyses on samples collected from FL478 and IR29. Proteomics analysis was performed at Agricultural Biotechnology Research Institute of Iran and proteome expression pattern was acquired using two dimensional gel electrophoresis. The expression pattern was analysed using Melanie 4 software and more than 150 salt responsive proteins were identified using MALDI TOF-TOF mass spectrometry. This led to identification of about 100 proteins involved in several cellular processes involved in plant adaptation to salt stress including oxidative stress defense, signal transduction, metabolisms, translation as well as photosysnthesis. Metabolome analysis of similar samples was performed at IPK, Germany. Several metabolites were measure and their statistical analysis and establishing a correlation between metabolome and proteome is in progress.

108.06 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
Jalan Tentara Pelajan 3A Bogor 16111, Republic of Indonesia, 62251338820

Collaborating institutes 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 2008, ICABIOGRAD conducted four solutions, i.e. six months hands-on research on marker-assisted selection and phenotyping at IRRI, partial support for ongoing long-degree programme, minigrants for one refrigerated centrifuge apparatus and minigrants for building blast nursery. By the end of this July 2009, we have finished two solutions and the project of building blast nursery that are now still under constructions. To ovoid the delay immigration administration for seed transfer in Indonesia and Philippine disturbed the condition of seeds and the schedule of project, as for six months hand-on research at IRRI, we did not carry out the activity; in spite of that we are performing the activity in Indonesia. We have obtained the BC3F1 seeds for two lines at IRRI. We conducted phenotypic evaluation by artificial inoculation with dominant blast races and genotypic evaluation using SSR primers at ICABIOGRAD. The selected plant will be then backcrossed to Way Rarem to get BC4F1. For partial support of the ongoing long-degree programme, we have used it to support the PhD programme of Wening Enggarini in the second year. This includes three semesters tuition fee, living expenses and travel for one month. Mini-grant for one refrigerated centrifuge apparatus, have used for purchasing one set Thermo Fisher Scientific Refrigerated Centrifuge 17R. Mini-grants for building blast nursery, we have used it for the purpose as mentioned in the proposal. We are now in the process of constructing blast nursery and preparation room. This project was still ongoing. The picture of the apparatus and building blast nursery in the bottom will send later.

Figure 1. Refrigerated centrifuge apparatus

Figure 2. A Building head-house of blast nursery, and B. Nursery installation

109. G4007.20: Managing the Generation Challenge Programme in a post-International Treaty world

August 2007−July 2008 (No-cost extenstion requested until 31 December 2008)

Principal Investigator and Lead Institute
Michael Halewood, Senior Scientist, Head, Policy Research and Support Unit, Bioversity International, m.halewood@cgiar.org

PROJECT UPDATE NOT SUBMITTED

110. G4007.21: Genotyping Support Service

Principal Investigator and Lead Institute
Humberto Gómez Paniagua, GCP; h.gomez@cgiar.org
Apdo. Postal 6-641, 06600 Mexico, D.F., Mexico**,** Phone: +52 55 5804 2004, Fax: + 52 55 5804 7558

Collaborating institutes 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 second call, GSS received 78 applications, out of which 32 were selected and are being processed:

Crop	No Proposals
Barley	1
Coconuts	1
Sorghum	1
Sweet potatoes	2
Yams	1
Rice	2
Cassava	3
Musa	3
Phaseolus	3
Potatoes	4
Maize	5
Cowpeas	6

These applications were received from the following countries:

Country	No Proposals
Bolivia	2
Brazil	1
Bulgaria	2
Chile	3
Ecuador	1
Ethiopia	2
Ghana	5
India	4
Kenya	2
Malaysia	1
Mexico	1
Mozambique	1
Nigeria	1
Peru	1
Philippines	2
Sri Lanka	1
Thailand	1
Uruguay	1

Twenty two of these applications aim to conduct molecular characterisation of germplasm and ten address plant breeding issues such as marker assisted selection or gene tagging.

The GSS is considering the possibility to announce a new call for proposals

111. G4007.22: GCP Workflow and Repository System (WMS)

July 2007–March 2008 (Phase I); April–August 2008 (Phase II)

Principal Investigator and Lead Institute
Norbert Niederhauser, Cropster GmbH; norbert.niederhauser@cropster.org
Tellenweg 10, 6850 Dornbirn, Austria, 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.

112. G4008.23: Statistical rules for defining characteristic genotype and marker sets

January 2008–December 2008; no-cost extension to September 2009

Principal Investigator and Lead Institute
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 institutes and scientists
- WUR: Thomas Odong; Hans Jansen; Jac Thissen; Theo van Hintum; Fred van Eeuwijk
- UNAM (presently UoC–Davis): Joost van Heerwaarden

Context
Large sets of genotypic score tables for molecular markers (mainly micro satellites) have been constructed for a wide series of crops within various activities within GCP-SP1. This marker information can be used as a starting point for diversity and association studies. 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: micro-satellite kits. 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 general philosophy underlying the activities in this project was that micro satellite scores define a multi-dimensional genetic space from which genotypes need to be sampled. A sampling strategy for genotypes can then consist in choosing genotypes such that the whole of the original genetic space is covered as good as possible by a defined number of genotypes. This can be achieved by cutting up the total genetic space in a number of subspaces equal to the number of genotypes in the final micro satellite kit and then choosing one genotype from each subspace. For the markers, next, that subset of

markers can be chosen that best approximates the full genetic space, or, more or less equivalently, that best preserves the genetic distances between the genotypes chosen in the micro satellite kit. If clear population substructure exists in the initial genotype collection, it would be useful to identify that structure and take it as a basis for choosing genotypes, where the number of genotypes to be chosen from individual subpopulations can be proportional or log-proportional to the size of the subpopulations.

Besides statistical principles, also molecular genetic requirements should be taken into account, 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 consists in 1) general guidelines for constructing micro satellite kits 2) defined reference sets for a number of crops (chosen were coco nut, potato, common bean, rice, chickpea).

Progress and outputs

Identifying subpopulations of genotypes

Common hierarchical cluster techniques (UPGMA, Ward's clustering, single and complete linkage), as well as neighbour joining and the Bayesian model based clustering algorithm implemented in the package STRUCTURE were applied to SP1 data sets and simulated population genetic data sets containing subpopulations exhibiting increasing genetic separation (as measured by F_{st} statistics). The idea was to investigate to which extent different cluster techniques 'identify' common and different types of subpopulations in real SP1 data. To help interpreting the results of the various cluster analyses on the SP1 data, comparisons were made with applications of the same clustering techniques to simulated data with known separation/differentiation. The results of this work is documented in a document called 'Determination of genetic structure of germplasm collections: Are traditional hierarchical clustering methods appropriate for genetic marker data?' by TL Odong, J van Heerwaarden, J Hans Jansen, TJL van Hintum & FA van Eeuwijk. Parallel to the study on the performance of cluster algorithms to identify subpopulations, a similar kind of study was undertaken looking at techniques derived from principal components analysis to identify subpopulations. The results of this work are written down in a document called 'Determination of genetic structure by adaptations of PCA' by J van Heerwaarden, TL Odong, , J Hans Jansen, TJL van Hintum & FA van Eeuwijk.

Identifying subsets of informative markers

Two main approaches were investigated: 1) choosing markers based on their squared loadings in PCA type of analyses as described in van Heerwaarden et al. (see above); 2) a stepwise approach in which markers are selected on the basis of their potential to distinguish subpopulations of genotypes (if present) or individual genotypes (in the absence of subpopulations). Results are described for the first approach in Van Heerwaarden et al. (see before), and for the second approach in Jansen et al. (see below).

Sampling strategies for genotypes

Depending on the degree of population structure one or more sampling strategies for genotypes can be chosen to create a micro satellite kit. Genotype sampling strategies are described in a document 'Sampling genotypes and markers for the construction of reference kits' by J Hans Jansen, Jac Thissen, TJL van Hintum & FA van Eeuwijk.

113. G4008.24: From attractiveness to feasibility: A strategic assessment of the capacity to develop and adopt GCP technologies

January 2008–December 2009

Principal Investigator and Lead Institute

Stanley Wood, IFPRI; s.wood@cgiar.org

International Food Policy Research Institute (IFPRI), 2033 K Street, Washington D.C. 20006-1002, USA; (1) 202-862-5600; Fax: (1) 202-467-4439

Collaborating institutes and scientists:

- CIAT: Glenn Hyman
- Consultant/Analyst, Canada: Carlos Tovar
- IPM CRSP–VPI: George Norton
- ASTI: Nienke Beintema
- Consultant, Tanzania: Sindi Kasambala
- ICASEPS: Erna Lokollo; Muchjidin Rachmat
- INERA–Burkina Faso: Mathieu Ouedraogo
- IER: Lamissa Diakite
- NRCRI: Godwin Asumugha

Summary of progress

A preliminary analysis of adoption capacity at country level was initiated using the Agricultural Science and Technology Indicators (ASTI) database. Several different indicators were evaluated. One of the key ones – agricultural research intensity – measures the agricultural R&D investment as a percentage of the agricultural gross domestic product (Alston, Chalfant, and Pardey 1993). Others assess human resource capacities in agricultural research and development. The results of our analysis suggested that countries have different levels of capacity to adopt and develop GCP technologies. These countries can be ranked according to their capacities and compared to various combinations of farming systems ranked by crop area, poverty, and drought index. Project researchers used these rankings, along with a need to align such selection with the seven GCP Challenge Initiatives, to select country case studies for the next phase of our research. Consequently, 5 case study countries and associated commodities were chosen: Burkina Faso (cowpea and rice), Mali (Rice and Sorghum), Nigeria (Cassava and Rice), Tanzania (Cassava), and Indonesia (Rice).

The next step was to select suitable case study collaborators for each country. A total of 20 applications were screened and the best suited candidates were interviewed by the project team and final five of them were selected based on their background, experience, and availability. Employment contracts were prepared and signed, and four of the five country case study collaborators participated in a survey planning and analytical design workshop that took place in Toronto, Canada in June 8 – 10, 2009. The central themes of the workshop focused around building awareness and understanding of the project goals, validating the conceptual approach proposed in the project design, and establishing a standardised national survey instrument and protocol. To provide a framework for this process, we developed a capacity evaluation matrix that addresses the two basic dimensions of the project (adaptation and adoption of GCP technologies). By collecting and analysing information that addresses the themes and questions set out in the matrix, the study generates strategically important information on the probability of successful broad-scale local adaptation and adoption of GCP technologies.

The capacity evaluation matrix was used during the meeting to come up with a list of 50 indicators for which data collection will now take place. The indicators were initially grouped in 5 categories:

1. **Crop improvement**: goals of national breeding programs, germplasm collections, research investments, personnel, infrastructure, training and capacity building, technology platform, use of farmer/community based processes of crop breeding, and varietal release.
2. **Seed systems**: seed production and producers, dissemination, traits of interest, and public financial support.
3. **Enabling environment**: extension (structure, investments, and awareness system), infrastructure (roads, irrigation systems, geographical distance to market), access to inputs and processing technologies, value chains, and market information.
4. **Household characteristics**: natural resource base, capital base, labour availability, literacy, varietal preferences, and membership to farmer associations.
5. **Adoption**: time from release to adoption, and variety adoption (area, share of farmers, and varieties with GCP traits).

The indicators have been incorporated into a survey form and regrouped into three sections: Crop Improvement; Seed Systems and Extension; and Household Characteristics, Enabling Environment and Adoption.

The survey will be administered to several target groups (NARS, seed companies, ministries of agriculture, universities, NGOs, and other specialised agencies) by means of documentary research and focus groups, comprising approximately 8 people (experts/key informants) from the said agencies. The elicitation is expected to take place between mid July and mid August and preliminary findings will be documented and presented at the forthcoming Annual Research Meeting of the GCP.

In addition, the project team has been gathering a broad range of relevant background material for the countries, commodities and national development capacities of interest. This includes identifying sources of secondary data that will be correlated with the findings of the study-specific surveys. Combined with this secondary data, the rapid survey and its analysis will help provide new information methods and tools for subsequent capacity assessments for other countries beyond those of the case studies.

Reference

Alston, Julian M. & Chalfant, James A. & Pardey, Philip G., 1993. "Structural Adjustment In OECD Agriculture: Government Policies And Technical Change," Working Papers 14473, University of Minnesota, Center for International Food and Agricultural Policy.

114. 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 Investigator and Lead Institute

Emmanuel Okogbenin, Cassava Programme, NRCRI/CIAT; eokogbenin@yahoo.com
NRCRI, Umudike, PMB 7006, Umuahia, Abia State, Nigeria/CIAT, Cassava Genetics Unit, A.A. 6713, Cali, Colombia.

Collaborating institutes and scientists

- CRI–Ghana: E Parkes
- NRCRI: C Egesi
- NaCRRI: Yona Baguma
- National Root and Tuber Program, ARI–Naliendele: Geoffrey Mkamilo
- IITA: Gedil Melaku
- CIAT: Martin Fregene

Cassava suffers from several pest and diseases that reduce yield by nearly 48 million tons yearly in Africa, about 50% of its current production, valued at US$1.4billion (FAO 2003). The cassava breeding community of practice (CoP) in Africa project is aimed at improving the capacity of NARs in modern breeding in field based strategies and marker assisted selection (MAS) in the development of cassava varieties resistant to pests and diseases. The CoP is therefore a platform for building and strengthening a network of African cassava breeders, in the task of creating a new generation of varieties that meet the needs of small farmers in Africa.

Population development and field evaluation of selected cassava clones

Development of appropriate breeding populations for specific traits is one of the goals of the CoP. NARs members carried out crosses for breeding for resistance to cassava mosaic disease (CMD), high dry matter content and yield. Crosses for resistance breeding to CBSD were also carried out in East Africa. In the CoP, $F_{1,}$ populations and selfed populations (S_1 to $S_{3)}$ have been generated so far. The project also initiated crosses for the pyramiding of two sources of CMD genes (CMD2 and CMD3).

Diallel crosses of elite clones were made to estimate GCA and SCA in these clones and to access their breeding values and determine best hybrid combinations.

In Nigeria, three MAS bred clones which are undergoing pre-release trial have been disseminated to farmers. NRCRI selected additional five clones from the GCP developed materials and these are to be tested at national level with partners (institutions and universities). One of the clones (CR41-10) has been identified for its very good architecture and canopy as well as high resistance to CMD. In Ghana farmer participatory breeding trial in three agro-ecological zones identified four varieties for adoption and it includes CR52A-25, AR 14-10, CR52A-4 and CR52A-31. These genotypes were selected for their good yields, architecture, and superior cooking characteristics. In Tanzania, CIAT derived lines improved for CBSD resistance are also at advanced stages of evaluation in breeding cycle and about ten of the clones will be nominated for release at the end of the evaluation cycle in 2010.

Genotyping and MAS

Genotyping and MAS activities for CMD resistance were conducted in backcross populations developed for delayed post harvest physiological deterioration (PPD) and increased protein/ beta carotene rich populations introduced from CIAT into Africa. The introgression of CMD resistance in these populations is important for its effective utilisation in Africa and for use in breeding programmes for the development of delayed PPD genotypes in NARs. About 500 clones of the delayed PPD population were selected for resistance to CMD with molecular markers. Beta carotene/protein cassava germplasm was also selected for CMD resistance with the selection of 138 genotypes which were then transferred to Africa. Field performance indicates that most of the genotypes are showing good resistance to the disease.

Through GSS platform of the GCP, F_1 progenies from CRI, Ghana developed using farmer preferred local varieties and CIAT elite lines were genotyped in routine MAS to select for CMD resistance and to identify new sources of CMD resistance in the CRI germplasm were also made. Genotyping results revealed a popular local variety in Ghana, “Dabudabu” as a possible source of CMD resistance gene (s) and will be further tested in further crosses to to validate preliminary results.

Sharing experiences and information

A web-based database, http://190.144.167.140/cbcopa, was created to enhance strong information exchange among breeders and to strengthen linkages with end users. This activity was handled at the International Center of Tropical Agriculture (CIAT). Activities for the design and implementation of website for CoP were anchored by the data services unit of CIAT with support of an external provider. Other basic features of the site include web-links, contacts, customised manual for the site, and gallery of documents. In addition a breeder –to-breeder visit among national agricultural research systems (NARS) breeders were undertaken to facilitate knowledge sharing and to improve interaction among breeders in 2008.

Capacity building

Capacity building in MAS and field based breeding strategies is a key output of this project. Two workshops were organised at the International Institute of Tropical Agriculture (IITA), Ibadan with resource persons drawn from IITA and outside Africa. Resource persons for the workshop were drawn from IITA, International Rice research Institute (IRRI), Mosanto, West African Biotechnology Workshops (WABWS), and the University of Ibadan. The two workshops were conducted between October 20 and 31, 2008. The first workshop was organised for conventional breeding (50%), quantitative genetics/statistics (30%), and molecular marker tools (20%). The second workshop was on molecular breeding covering a wide array of topics.

Reference

FAO (2003). Cassava production statistics, 2002. http//www.fao.org

115. G4008.28 Characterisation of maize diversity in Central Europe

March 2008–June 2008

Principal Investigator and Lead Institute

Peter P Papp, ABC; ppapp@abc.hu

Agricultural Biotechnology Center, Gödöllő, No 4 Szent-Györgyi A., Gödöllő, H-2100, Hungary; +36 (28) 526-100; Fax: +36 (28) 526-145

Collaborating institutes and scientists

- ABC: Tünde Golenar; György Botond Kiss
- IA–Tápiószele: Laszlo Holly
- CIMMYT: Marilyn L Warburton

The Hungarian maize collection is comprised of over 4000 accessions. Locally adapted varieties and ecotypes were collected from all over Hungary in the last five decades and they are maintained by the Institute for Agrobotany at Tapioszele. Collaborations between the Hungarian Ministry of Agriculture (affiliated institutes: Agricultural Biotechnology Center and Institute for Agrobotany) and GCP (affiliated institute: CIMMYT) aimed to assess Hungarian maize collections by characterising a subset of accessions and comparing their allelic diversity with the diversity present in GCP's reference sets. The project also supported the training of a Hungarian scientist on molecular characterisation who worked at CIMMYT to complete the characterisation during the training period.

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.

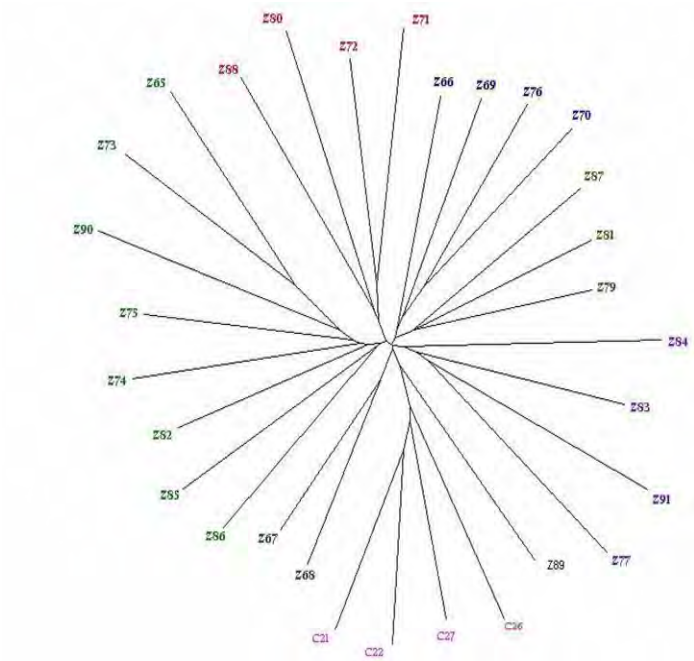


Fig. 1. Relationship between maize genotypes of 26 Hungarian landraces (Z65-Z77 and Z79-Z91) and 4 GCP control populations (C21-C22 and C26-C27) based on Euclidean distance analysis of data from 45 SSR markers generated by PowerMarker. Having several branches on the tree that begin near the centre (origin) indicates the presence of substantial diversity in Hungarian landraces.

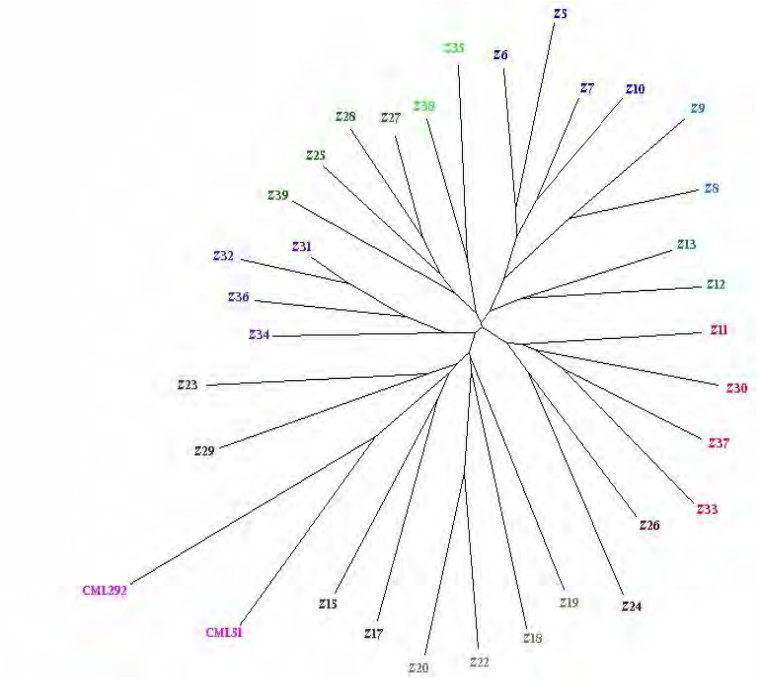


Fig. 2. Relationship between maize genotypes of 32 Hungarian inbred lines (Z5-Z13, Z15, Z17-Z20 and Z22-Z39) and 2 GCP control lines (CML51 and CML292) based on Euclidean distance analysis of data from 27 SSR markers generated by PowerMarker.

A common set of GCP and Hungarian markers for populations was generated and analysed using appropriate softwares (R, Powermarker, Darwin, NTSYS).

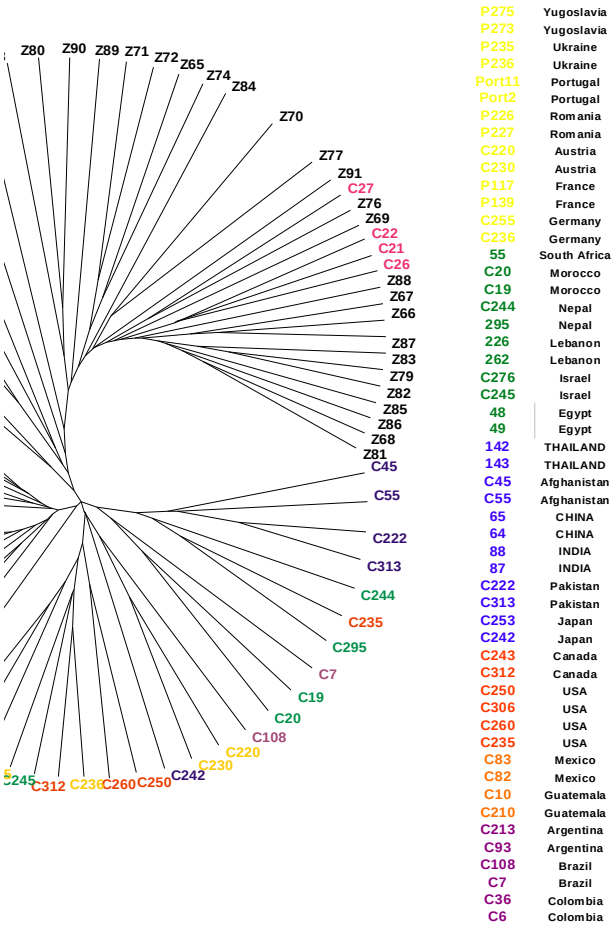


Fig. 3. Radial dendrogram presenting the comparison of genotype data of 26 landraces (Z65-Z77 and 79-Z91) derived from the Hungarian maize collection and four controls from CIMMYT (C21, C22, C26 and C27) with 53 landraces of the GCP reference set, based on 21 SSR markers. The reference set contained genotypes from Europe, Africa, Asia, North-, Central- and South- America.

Genotypes of the Hungarian maize landraces were compared to the GCP reference set with genotypes from different continents. Twenty-one SSR markers were found to be suitable for merging the two sets. The strange feature seen in the result is that Hungarian genotypes clustered away from the other genotypes. Even European genotypes, obtained from neighboring countries, did not clustered together with the Hungarian ones, contrary the four controls from CIMMYT, used in the genotyping experiments. Although we can't rule out that the result is real, since only two genotypes were selected from each represented country, it is also possible that the phenomenon might reflect some problems with merging the two sets. Repeating the genotype determination step with some of the accessions selected from both the Hungarian and GCP reference set could clarify the situation. Within-set analysis of the Hungarian lines that were all run at the same time, however, should be correct and show real relationships and diversity between the populations. Six alleles found in the Hungarian populations with 5 markers were not represented in the reference set. PIC values for the landraces were quite high, as compared to many previous studies (Warburton et al., 2002), showing that these accessions are a good source of diversity.

The analyses of our marker data revealed that the Hungarian maize germplasm contains substantial diversity, which could support the national programme's breeding objectives as well as allow participation in international programmes aiming at maize improvement and conservation.

Our aim is to create a genotyped core collection from the Hungarian maize germplasm. Samples have already been collected from 180 inbred lines and 180 landraces for obtaining DNA for genotyping.

116. G4008.29 Characterisation of bean diversity in Central Europe

January 2008–December 2008

Principal Investigator and Lead Institute

Peter P Papp, ABC; ppapp@abc.hu
Agricultural Biotechnology Center, Gödöllő, No 4 Szent-Györgyi A, Gödöllő, H-2100, Hungary; +36 (28) 526-100; Fax: +36 (28) 526-145

Collaborating institutes and scientists

- ABC: György Botond Kiss
- IA–Tápiószele: Laszlo Holly
- CIAT: Matthew W Blair

The Hungarian bean collection is comprised of over 3000 accessions, collected all over Hungary in the last five decades and managed by the Institute for Agrobotany at Tapioszele. Characterisation of the Hungarian bean germplasm has been lunched 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) were initiated 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.

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 (32 genomic and 20 gene-based), 50 were polymorphic and they identified 6 alleles in an average, spanning from 2 to 35 alleles/marker.

Out of the 100 Hungarian bean accessions 41 genotypes were clustered with the Mesoamerican controls (DOR364 and ICA Pijao) and 56 genotypes with the Andean controls (G19833 and

CALIMA), while 3 samples represented hybrid genotypes (INT) and they could be found between the two clusters. Within the Andean gene pool two major subcluster (A1 and A2) were identified (Fig.1). A minor excess of Andean gene pool (56% versus 41%) might related to the preference of the Hungarian costumers for large seed size. STRUCTURE analysis confirmed the clustering of genotypes shown by DARwin analysis. Further subgroups can be seen within both gene pools but they do not correlate with known races characteristic for the given gene pool. Although the controls clearly marked the two kinds of gene pools, they stood out from the Hungarian genotypes within both clusters, and similar phenomenon could be recognised in the comparisons with GCP reference sets. This may indicate diverged allelic compositions in the Hungarian genotypes caused by adaptation to the local environment, selections by farmers or market preferences.

The characterised set of the Hungarian germplasm and GCP's reference set were compared to each other. Accessions belonged to the two gene pools were compared separately. Comparisons were based on genotype data obtained by 35 fluorescent SSR markers.

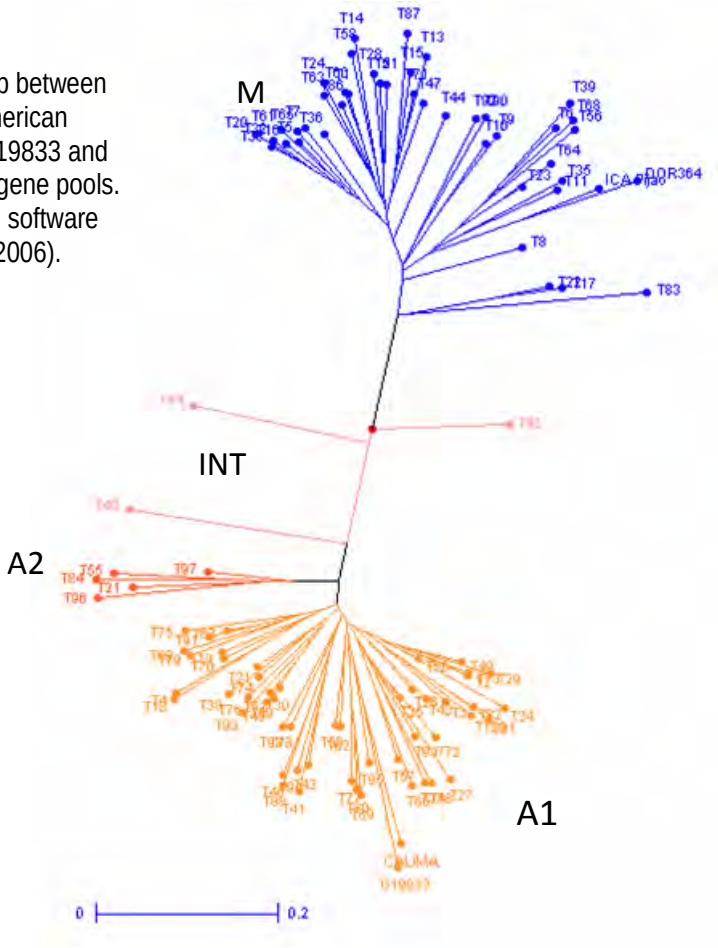


Fig. 1. Radial representation of relationship between 100 Hungarian bean genotypes. 2 Mesoamerican (DOR364 and ICA Pijao) and 2 Andean (G19833 and CALIMA) controls represent the two major gene pools. Dendrogram was created by using DARwin software package (Perrier and Jacquemoud-Collet, 2006).

The closest relationship could be seen with the Durango-Jalisco race when Hungarian bean genotypes were compared with GCP references within the Mesoamerican gene pool.
It is interesting that Hungarian bean genotypes, compared with GCP references within the Andean gene pool by correspondence analyses, clustered closely with G19833, representing Peru race. UPGMA dendrogram indicated some associations of Hungarian Andeans with race Nueva Granada.

The analyses of the selected accessions 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. We believe that our study will lead to the recognition of the value of the Hungarian bean germplasm and genotyping of additional accessions will result in justifying that Hungary is an important secondary centre of diversity for common bean.

Another 360 accessions have been selected to continue genotyping of the Hungarian bean germplasm, collection of plant samples for obtaining DNA is under progress.

117. G4008.35: Toolbox of available molecular markers useful for marker assisted selection in GCP crops

March 2008–March 2010

Principal Investigator and Lead Institute
Veerle Van Damme, consultant; veerle.vandamme@gmail.com
Cali, Colombia; (57) 2-5552650

Collaborating institute and scientists

- GCP: M Carmen de Vicente; Humberto Gómez Paniagua

This project deals with the development of a toolbox that provides free and easy access to succinct information of all effectively used and publicly available molecular markers (SSR, SCAR, SNP, STS) for marker-assisted selection (MAS) 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.

This toolbox is mainly aimed to agricultural researchers and plant breeders, in particular in developing countries, who face difficulties to access the scattered and constantly evolving scientific information on useful molecular markers. The information generated for each marker includes general information (trait identified by this marker, marker code, marker type, ….), the lab protocol, the validation process, the most relevant references, and the name of corresponding experts.

A first step in collecting the information for the toolbox consisted of a literature study on the available markers useful for MAS for each of the 19 crops. In a second step the result of this literature study, a preliminary list of markers, was presented to several crop experts who checked if the found markers were effectively used in marker-assisted selection. As the breeder community possesses the most up to date information and breeders have hands on experience on the use of markers their collaboration was considered vital. After this revision by experts the updated information was entered in the toolbox.

The current information available in the GCP Toolbox can be consulted at http://www.generationcp.org/sp5/?da=09148937 and consists of supporting texts and the toolbox itself (See “Enter the GCP Toolbox”). The supporting texts were considered essential to introduce the toolbox and to explain its contents and use.

For *Musa* spp., coconut, lentil, millet, pigeonpea, sweet potato and yam MAS is not applied yet or the used markers do not comply with the conditions to be included in the toolbox. For these crops the provided information is limited to the actual status of use of molecular markers in these crops. An overview of the available markers useful for MAS in the remaining 17 crops is shown in table 1.

As the development of markers is a very dynamic and rapidly evolving field, updating the toolbox is an ongoing process whereby scientists and experts are invited to comment and collaborate with their latest findings and experiences in markers, especially those related to the effective uptake of it in breeding programmemes.

The toolbox is complementary to the Genotyping Support Service (GSS) of the Generation Challenge Programmeme – subprogramme 5, launched in 2006, which aims at facilitating the access of national agricultural research systems (NARS) to genotyping technologies. It will be disseminated by CDs or USB sticks for those scientists and breeders who have difficulties accessing computers or the Internet.

In a near future the toolbox will be extended with other types of markers e.g. CAPS, MITE and allele-specific markers. Markers that are used for tracing genes used in genetic modified crops will be added as well. Any other extension can be included as this can comply with the requests of the plant breeders as informed through the feed-back button of the toolbox or the comments received by the realised Survey Monkey.

By sharing the latest advances in molecular plant breeding, the toolbox will be an important contribution to support modern agricultural science for the benefit of the poor in developing countries.

Table 1. Summary of the markers presented in toolbox for those crops for which marker-assisted selection

Crop	No of traits for which markers are available		No of markers according to trait type		No of markers according to the type of marker				Total No of markers
	biotic	abiotic	biotic	abiotic	SSR	STS	SNP	SCAR	
Barley	8	7	23	14	20	16	0	1	37
Bean	8	0	13	0	0	0	0	13	13
Cassava	2	0	5	0	4	0	0	1	5
Chickpea	1	0	2	0	2	0	0	0	2
Cowpea	2	0	2	0	0	0	0	2	2
Faba bean	1	0	0	2	0	0	0	2	2
Groundnut	1	0	1	0	0	0	0	1	1
Maize	2	2	2	4	6	0	0	0	6
Potato	2	0	2	0	1	0	0	1	2
Rice	4	4	15	5	4	9	5	2	20
Sorghum	0	1	3	0	3	0	0	0	3
Wheat	10	10	30	24	15	34	0	5	54
Total	41	24	98	49	55	59	5	28	147

118. G4008.36: Getting the focus right: Food crops and smallholder constraints

April 2008–February 2009; no-cost extension to April 2009

Principal Investigator and Lead Institute

John Dixon, CIMMYT Impacts Targeting and Assessment Unit; j.dixon@cgiar.org
Apdo. Postal 6-641, 06600 México, D.F., México; (52) 595 952 1900

Collaborating institutes and scientists

- CIMMYT: Stephen Waddington; Xiaoyun Li
- CIAT: Glenn Hyman
- Food crop breeders, crop management, socio-economics and GIS specialists in CG centers including IRRI, CIMMYT, ICRISAT, CIAT, WARDA and IITA
- Numerous NARS and NGO institutions and staff in South and East Asia, Sub-Saharan Africa, Latin America

Context

To generate information to support existing projects and identify areas for future investment, the GCP commissioned a study of production constraints and opportunities for important food crops in priority farming systems with high degrees of poverty following an approach successfully piloted for maize in an earlier GCP study. The study focused on wheat, rice, sorghum, cassava, cowpea and chickpea, grown in 15 broad farming systems with large numbers of stunted children, drought incidence and large production areas of food crops (Hyman et al 2008). These systems were located in South and East Asia, sub Saharan Africa and Latin America. Applying a modified Delphi method (Dalkey 1969), three rounds of surveys were conducted April 2008 – March 2009 with over 670 panelists familiar

with the crops and systems. They included plant breeders, agronomists, socio-economists and other researchers; extension and training specialists, input suppliers and farmer organisations representing government institutions, international organisations, NGOs and the private sector. Panelists identified important abiotic, biotic, management and socio-economic constraints that contribute to the smallholder farm yield gap (defined as *Highest achieved yield on farm – Average yield on farm)* and estimated yield losses associated with the constraints (Evenson et al 1996) which were adjusted for frequency of occurrence. Constraints were ranked by size of yield loss. Panelists also proposed solutions to the most serious constraints.

Findings and implications

Large smallholder farm yield gaps were reported for most crops in most farming systems, implying significant scope for improvement of farm yields if the most serious constraints can be identified and alleviated. Generally yield gaps were smallest (averaging 60% of current farm grain yields) for rice, mid size for wheat and cassava, and usually larger (sometimes double current farm yields) for sorghum and the two legumes. They were larger in the marginal, dryer farming systems, particularly in Sub-Saharan Africa, and smallest in the high input and yield systems of East Asia. Abiotic, biotic, management and socio-economic constraints were all important contributors to yield gaps. Abiotic and management constraints were more important for wheat, socio-economic and management issues for rice and cassava, and abiotic constraints for sorghum. Biotic constraints dominated the legumes.

Many serious specific constraints were reported for the crops in the systems. Most serious constraints were considered to be getting worse. Overall, the combined yield losses from the ten most serious constraints identified contributed between 45% (cassava) and 56% (chickpea) of the average yield gap. The most serious constraints for wheat involved the management, high cost and deficiency of N fertilizer, and problems associated with grain filling and mid season drought stress and irrigation management. Those for rice included the deficiency, high cost and poor management of N fertilizer, soil fertility depletion, various leaf, stem and head pests and diseases, weed competition and issues related to water management. *Striga* and weed competition, constraints with the soil resource and soil fertility management, and drought were the most serious for sorghum. Pod, leaf, stem and flower insect pest problems and the high costs of their control dominated for cowpea. *Helicoverpa* pod borer, *Botrytis* grey mould and the high cost of control were the most serious for chickpea. Soil fertility depletion and fertilizer input and management problems were also widespread and serious with the legumes. Marketing and finance issues were concerns for cassava along with weed competition, African cassava mosaic virus and poor varieties/planting materials.

Respondents proposed many variety/germplasm interventions to address the most important constraints, and suggested policy/socio-economic and crop management solutions. Proposed solutions to biotic and abiotic constraints often involved the development of improved germplasm with tolerance or resistance to various pest, disease, nutrient and water stresses. Some addressed management issues such as the need for earlier maturity types to fit into intensifying cropping systems. Effective deployment of seed of new varieties with the appropriate traits and farmer awareness were emphasised.

Next steps

A comprehensive report of findings has been developed for the GCP and is being circulated more widely. A journal article of results from the study is in preparation. These findings on constraints and opportunities for a range of important food crops will be useful to guide GCP priority setting. The GCP should find many of the suggestions on serious constraints and on varietal, germplasm or genetic solutions helpful to support their current investments in particular thrusts with the crops surveyed and to guide deployment of their resources into the future.

References

Dalkey NC (1969). The Delphi method: An experimental study of group opinion. The Rand Corporation, Santa Monica CA, USA. 79 p.

Evenson RE, Herdt RW and Hossain M (1996). Priorities for rice research: Introduction. In: Rice Research in Asia: Progress and Priorities. Evenson RE, RW Herdt and M Hossain (eds.). CAB International and IRRI. Wallingford, UK. pp. 3-15.

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.

119. G4008.37: Training plant breeders at the West Africa Centre for Crop Improvement

January 2009–December 2012

Principal Investigator and Lead Institute

Eric Danquah, WACCI; edanquah@wacci.edu.gh
University of Ghana, PMB 30, Legon, Ghana; (233) 21-520605-8; Fax (233) 21-520604

Collaborating institutes and scientists

- Biotechnology Centre, UoGh: Kwame Offei
- College of Agriculture and Life Sciences, Cornell; Vernon Gracen

The West Africa Centre for Crop Improvement (WACCI) was established in June 2007 with initial funding from the Alliance for a Green Revolution in Africa (AGRA) to equip the next generation of plant breeders with the necessary knowledge and skills to develop improved varieties of the indigenous crops that feed the peoples of the West and Central Africa regions. WACCI, a collaboration between the University of Ghana and Cornell University, is located on the main campus of the University of Ghana. The PhD at WACCI is a 5-year programme. Students undertake two years of coursework at the University of Ghana and three years of field research at the students' home institution.

The AGRA funding allows five cohorts of eight students to enroll each year over a 5-year period. The first cohort of eight students enrolled in February, 2008. In March, 2008 WACCI received a grant from the Generation Challenge Programme (GCP) to train four additional students at the PhD level. The first batch of two students enrolled with the WACCI second cohort in January 2009. Currently, there are 18 doctoral candidates from Mali, Niger, Ghana, Nigeria and Burkina Faso. The two GCP students, Ruth Thompson, an assistant research scientist at the Biotechnology Division of the Council for Scientific and Industrial Research - Crop Research Institute (CSIR-CRI), Ghana and Dramane Sako, from the Kaye's Regional Centre for Agronomic Research, Mali have undergone the first semester coursework at WACCI. They took courses in Biometry and Research Methods, Integrated Pest Management (IPM), Plant-microbial Interactions and the Control of Plant Diseases, Plant Genetics and Genetic Improvement of Crop Plants. In the second semester which begins in August, 2009 the students will take courses in plant tissue culture, molecular genetics and biotechnology in plant breeding, quantitative genetics and other related subjects.

Ruth Thompson and Dramane Sako will work on improving cassava and sorghum respectively. Their project proposals will be developed in the second year which is dedicated to some advanced modular course work, critical writing and execution of mini-project.

WACCI is committed to training the next generation of plant breeders with the capacity to lead the conversion of genetic and molecular discoveries into innovative solutions that result in improved varieties. To achieve this, WACCI is exploring win-win linkages to assure quality in both teaching and research at the Centre and the student's home institution. Currently we seek two full-time faculty in molecular genetics and plant breeding to complement staff at the Centre.

120. G4008.38: Fellowships and travel grants 2008

Principal Investigator and Lead Institute
Carmen de Vicente, GCP; c.devicente@cgiar.org
Apdo. Postal 6-641, 06600 Mexico, D.F., Mexico**,** Phone: +52 55 5804 2004

The following travel grants and fellowships were awarded in 2008:

Travel grants awarded for training

	Name	Institute	Country	Topic	Host
1.	Sarkar Ramani Kumar	CRRI	India	Molecular markers	IRRI
2.	Afiukwa Celestine Azubuike	ESU	Nigeria	Molecular markers	FABI
3.	Triwitayakorn Kanokporn	MahU	Thailand	Phenotyping for drought	Cornell
4.	Geoffrey Mkamilo	ARI–Naliendele	Tanzania	Marker-assisted selection	IRRI

Travel grants awarded to attend the 2008 GCP Annual Research Meeting

	Name	Institute	Country
1.	Shailaja Hittalmani	UAS–Bangalore	India
2.	Kongpanh Kanyavong	NAFRI	Laos
3.	Issa Drabo	INERA	Burkina Faso
4.	MA Salam	BRRI	Bangladesh
5.	Nguyen Thi Lang	CLDRI	Vietnam
6.	Ndiaga Cissé	ISRA	Senegal
7.	Zeba Islam Serraj	UoD	Bangladesh
8.	Chiedozie Egesi	NRCRI	Nigeria
9.	Khin Than Nwe	DAR	Myanmar
10.	Sabariappan Robin	TNAU	India
11.	Habibul Bari Shozib	UoD	Bangladesh
12.	Lalith A Perera	CRI–Sri Lanka	Sri Lanka

Travel grants awarded to attend the groundnut workshop

	Name	Institute	Country
1.	René Maita	PROINPA	Bolivia
2.	Jorge Rojas	PROINPA	Bolivia

Fellowships awarded in 2008

	Name	Institute	Country	Topic	Host
1.	Zhengzhi Zhang	NKLCGGE, NAU	China	Translation of SNPs underlying drought tolerance in wild barley into high-throughput low-technology markers	SCRI, UK
2.	Priscilla Sabadin	USP	Brazil	Multi-trait multi-environment QTL analysis for agronomical performance of sorghum on acid soils	EMBRAPA Maize and Sorghum, Brazil and WUR, The Netherlands
3.	Caroline Marques Castro	EMBRAPA Clima Temperado	Brazil	The use of relatedness information in linkage disequilibrium mapping: pedigree information versus molecular marker information	WUR, The Netherlands
4.	Mateo Vargas Hernández	CIMMYT	Mexico	Data analysis of a network of field trials involving a population of recombinant inbred lines: dissecting the genotype x environment interaction	Agropolis–INRA
5.	Dindo A Tabanao	PhilRice	Philippines	The use of relatedness information in linkage disequilibrium mapping: pedigree information versus molecular marker information	WUR, The Netherlands

121. G4008.39: Capacity-building *à la carte* 2008

Principal Investigator and Lead Institute
Carmen de Vicente, GCP, 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 1st December 2007 to January 31st 2008 with capability to accommodate 10 grants. In the end three teams were selected, and continue to carry out activities in 2008:

- *Enhancing MAS capacity for salt-stress rice breeding in Bangladesh* (Lead institute: BRRI)
- *Improving capacity for phenotyping for abiotic and biotic stress in Burkina Faso* (Lead institute: INERA–Burkina Faso)
- *Improving capacity for phenotyping for abiotic and biotic stress in Senegal* (Lead institute: ISRA)

121.01 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

PROJECT UPDATE NOT SUBMITTED

121. 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

PROJECT UPDATE NOT SUBMITTED

121.03 G4008.39.03: Capacity-building à la carte 2008–Improving capacity for phenotyping and biotic stress in Senegal

April 2008–April 2010

Principal Investigator and Lead Institute
Ndiaga Cisse, ISRA; ncisse@refer.sn
ISRA/CNRA BP 53, Bambey, Senegal

Collaborating institute and scientists
UoC–Riverside: N Ndack Diop; Philip Roberts; Jeff Ehlers

1. Research activities and progresses
Two drought phenotyping trials including each 30 genotypes (early and medium maturing) were conducted. The Trial seeks to (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 two water regimes; (2) study the relationship between grain yield under drought and various traits.

Early maturity group

Growth cycle

Survival rates were poor for Bambey 21, UCR-P-24 , SH-50 and UC-CB27. It was high for IT84S-2246, IT85S-3139, IT95K-181-9, IT95M-190, IT97K-499-39, IT98K-1111-1, Mouride and Ifé Brown. About 1/3 of the genotypes had 50% of its plants producing flowers in the stressed trial and half produced few small immature pods. These pods did not produce seeds.

Physiological traits

For about 30% of the genotypes, relative chlorophyll content remain practically the same after 6 weeks of stress. Some of the lines are: IT95M-190, IT98K1111-1, IT97K-499-39, Ifé Brown, IT95K-1491 and Mouride. Little variations were observed for: IT84S-2049, KVX61-1 and Prima. Significant differences were observed between the two periods of measurement for this parameter for IT85S-867-5, Bambey 21, Sh-50, IT82E-18, UC-524B, Yacine, UCR-CB27etc

Differences in canopy temperatures were significantly important for most the genotypes. It is smallest for IT99K-1111-1, IT97K-499-39 followed by IT95M-190, Mouride and Prima. Genotypes with the highest canopy temperatures and relative differences are: Bambey 21, IT85F-867-5, IT82E-18, SH-50, Yacine and the four UC lines.

Performance

Significant differences were observed between well-watered and stressed plots for biomass yield and per plant. Eight lines have high biomass production in stress conditions; These include IT84S-2049, IT85F-3139, IT95M-190, IT98K-1111-1, IT97K-499-39, KVX-61-1, Mélakh, Mouride and Ifé Brown. These genotypes have also with the highest grain yield in the watered conditions.

Conclusions

The drought stress conditions seemed particularly severe and no grain production was observed. However some genotypes appears to have potentials to adapt to drought conditions. These include: IT95M-190, IT98K-1111-1, IT97K-499-39 and KVx-61-1. Whereas the lines Bambey 21, Sh-50, IT82E-18, UC-524B, Yacine, UCR-CB27 appears to show susceptibility to water stress.

Medium maturity group

Growth cycle

For all genotypes excepted Petite-n-grn, survival rates were very good to excellent in the stressed plots. It was 100% for 1/3 of the genotypes such as Mougne, Kvx-525, IT89KD288 etc. All genotypes excepted 58-53, 58-57 and Petite-n-grn had attained 1 flower stage. However, only half of them had 50% of plants with at least one flower. About 20 genotypes produced flowers which did not attain maturity in the stressed trial. Mougne and Kvx-525 obtained the highest number of pods. These pods did not produce seeds.

Physiological traits

Relative chlorophyll content expressed by the Spad values remained high and practically the same after 6 weeks of stress for 8 genotypes. The line IT89KD-288 had the highest Spad value after 6 weeks of drought stress. This value was high also for Mougne, Petite-n grn etc. The variation in Spad values between the evaluation times were high for about 1/3 of the genotypes including 58-57, 58-53, Suvita 2 and IT97K-556-6. Significant differences were observed between the two periods of evaluation for this parameter.

Differences in canopy temperatures were significantly important for most of the genotypes. It varied the least for IT00K-901-1, IT83D-442, IT98K-317-2, IT98K-428-3 and Mougne. Genotypes with the highest canopy temperatures and relative differences include: 58-53, 58-57, IT97K-556-6, IT97K-819-132 and Suvita 2.

Performance
Significant differences were observed between well-watered and stressed plots and per plant biomass yield. The genotype mougne had the highest biomass production in stressed plots, followed by Suvita 2, Kvx-525 and IT00K-901-6. The genotype 58-57 had below mean production.

Conclusions
The drought stress conditions were also very severe for the medium maturity genotypes and no grain production was observed. Mougne appears to be the most adapted to the stress condition of the trial. The genotypes IT89D288 and Kvx-525 show some attributes which confer adaptation to water stress. Genotypes such as 58-57, Suvita 2 and 53-53 which are reputed as drought resistance, are showing susceptibility to the drought conditions of this trial. This result may indicated that genotypes have different responses to drought during this period and the rainy season.

122. G4008.40: Workshop on 'Reference sets of food crop germplasm for international collaboration'

January 2008–December 2008

Principal Investigator
Jean Christophe Glaszmann, Agropolis–CIRAD/GCP; jean-christophe.glaszmann@cirad.fr
Agropolis–CIRAD, TA A-DIR PS3, Boulevard de la Lironde, 34398 Montpellier Cedex 5, France
M Carmen de Vicente, GCP; 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 sise 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 sise. 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 were described and discussed, as well as the perspectives and the mode of organisation necessary for taking full advantage of this product. It was an opportunity for coordination among various players engaged in germplasm management in international programmes.

The workshop was held in Montpellier, France, on November 13-17, 2008.

Outcomes of the workshop can be seen on the GCP Capacity-building website
http://www.generationcp.org/sp5/?da=08137824

123. G4008.43: Improve cowpea productivity for marginal environments in Mozambique

July 2008–June 2010

Principal Investigator and Lead Institute
Rogério Chiulele, EMU; rchiulele@uem.mz
Eduardo Mondlane University, Maputo, Mozambique; P.O. Box 1715, Phone: +25821492177/8; Fax: +25821492176

Collaborating institutes and scientists

- UoC–Riverside: Jeff Ehlers; Timothy Close; Philip Roberts
- PSU: Jonathan Lynch

This project has three objectives. The objective 1 offers opportunity to build capacity in drought tolerance screening through conducting drought trials and interacting with other groups doing the same type of trials. This also offers training in analysing data for genotype by environment interaction and presentation of results. Apart from the capacity building this objective will provide baseline information on drought tolerance for early and medium cycle cowpea varieties and assess the importance of genotype x environment interactions for grain yield under drought in Mozambique. The objective 2 provides experience in larger-scale germplasm screening for drought tolerance by assessing the genetic variability for drought tolerance of 300 Mozambican cowpea landrace accessions. The objective 3 provides experience on how to design and implement an MAS-based programme, in close collaboration with mentors at UC Riverside. This objective will also enable to develop breeding populations suitable for application of marker-assisted selection (MAS) using SNP-based markers. In Objective 1, thirty early maturing and thirty medium maturing cowpea varieties will be compared for grain yield under terminal drought conditions using late plantings at two drought-prone sites in Mozambique during the main growing season and in one trial under irrigation during the off-season. This will provide baseline drought tolerance information for a wide range of cowpea genotypes in Mozambique and will allow identification of drought tolerant and susceptible 'checks' for future drought studies. By comparing results from identical trials being conducted in West Africa by an associated GCP project "Improving Drought Tolerance Phenotyping in Cowpea" of the SP3, it will be possible to estimate genotype x environment interactions for grain yield under drought across a wide range on conditions, including the degree of correlation between the results of off-season controlled environment screening and results from main-season African growing environments. The effectiveness of a new root screening protocol developed for evaluating drought tolerance and rooting characteristics in common bean (Lynch, 2007) will be evaluated in the Objective 1 trials to determine associations between root ratings and grain yield under drought in cowpea. In Objective 2, 300 landrace accessions from Mozambique will be assessed for tolerance to drought using screening protocols developed for the GCP-TL I project and the rapid root screening assay (Lynch, 2007). This will complement assessments of other sets of cowpea germplasm being assessed in the GCP-TL I and ICRISAT Tropical Legume II projects by including unique germplasm not being evaluated in these other projects. In Objective 3, ten breeding populations appropriate for Mozambique and for marker-assisted recurrent selection using SNP-based and SSR markers developed under the GCP-TL I project. Overall, the funding offers opportunities for capacity building on phenotyping for drought tolerance, design and implementation of MAS-based programme.

This project aims to build capacities in cowpea screening for drought tolerance and on data analysis and interpretation. Apart from the capacity building this project aims to provide baseline information on drought tolerance screening of early and medium cycle cowpea varieties, to assess the importance of genotype x environment interactions for grain yield under drought in Mozambique and to determine the genetic variability of drought tolerance on Mozambican germplasm. In addition, the project was

also aimed to develop breeding population suitable for the country and for the use of marker-assisted selection.

- As a result four (4) genotypes consisting of two drought tolerant and two drought susceptible were identified as checks and are being used in drought tolerance screening trials in the country.
- The relevance of off-season drought evaluation to drought performance in the main season is still being determined. However, preliminary results suggest to be useful screening criteria for drought tolerance in Mozambique.
- Three hundred (300) genotypes were screed for drought tolerance during the off-season to assess the genetic variability of drought tolerance on local germplasm. From the results of this trial, a set of 84 best performing genotypes was further evaluated for drought tolerance under field conditions during the main season to determine the usefulness of off-season screening on main season drought screening and for identifying local checks for future studies.
- Ten breeding population were developed from the crosses between 2 drought tolerant sources and 5 local landraces. These populations are being advanced to F2 generation.
- Because of the results of this project a cowpea breeding programme for drought tolerance is being established in the country that will contribute for improved livelihood of the small-scale farmers growing cowpea in the drought prone environments of the country.
- The assessment of the genotypes to determine the usefulness of off-season drought screening will still continue to generate consistent results. The set of 84 best performing genotypes are being evaluated in two contrasting environments and seasons to determine the GxE.
- The 10 breeding populations developed will be advanced further to generate the F3 population necessary for application of markers-assisted selection.

124. G4008.44: GCP Learning Materials

January 2008—Dec 2009

Lead institution
GCP (Carmen de Vicente)

GCP learning materials are derived from workshops and/or commissioned at Programme level by GCP's Subprogramme 5, in close collaboration with the technical Subprogrammes. The learning materials cover topics of contemporary relevance to GCP's work and mission, and are offered for free public use. We would however appreciate acknowledgement of the CGIAR Generation Challenge Programme whenever you use or adapt these materials, and we would also appreciate hearing from you on how you have used the material (please send an email to c.devicente@cgiar.org)

You are warmly invited to freely use and/or print any of these materials for educational or other non-commercial purposes without prior permission, provided due credit is attributed.

The materials for the three courses below are accessible via GCP's Capacity-building corner website (http://www.generationcp.org/sp5/) and are also available in CD format. If you would like to receive a copy of the CDs by mail, please contact the GCP Communications Assistant (info@generationcp.org). Another four sets of learning materials are under preparation.

1. Genetic resource policies course (http://www.generationcp.org/distantpolicies/)
A distant learning module for scientists, covering genetic resource policies and implications on freedom to operate. This course material was developed in collaboration with Wageningen University and Research Centre (WUR).

2. The McClintock crop bioinformatics course (http://mcclintock.generationcp.org/)
This self-study introductory online course targets scientists with a reasonable background in germplasm, biology and genetics, and their application in plant and agricultural sciences. Named in

honour of Nobel prize winning crop geneticist, Barbara McClintock, this course is a joint project between IRRI and the CGIAR Generation Challenge Programme, designed to demonstrate how basic bioinformatics tools, techniques and resources can help molecular biologists, geneticists and other scientists to effectively manage sequencing projects. This course is a 'living resource' and will be continually updated to attune it with the latest developments in the dynamic field of bioninformatics.

3. Genomics and comparative genomics (http://www.generationcp.org/genomics/)
For use by scientists and advanced students with a strong background in biology and genetics as basic class material or self-tutorial. The principal audience includes plant breeders, molecular biologists and other plant scientists on the fringe of—but not fully engaged in—genomics research. A modular approach accommodates the different backgrounds and needs of users of this material, developed jointly by Cornell University's Institute for Genomic Diversity and the CGIAR Generation Challenge Programme

***The following learning materials are under preparation*:**

1. Genetic diversity
2. Marker-assisted selection
3. Phenotyping
4. Association genetics

125. G4008.50: Delivery plan remote learning modules

January 2009–June 2010

Principal Investigator and Lead Institute
Lawrence Pratt, CIMS; lawrence.pratt@incae.edu
Apdo. 960-4050, Alajuela, Costa Rica; (+506) 2437-2297; Fax: (+506) 2443-0506

Collaborating institutes and scientists

- Cornell: Stefan Einarson
- GCP grantees, to be determined
- Related NARS, to be determined

1. Research activities and progress
A critical component of the Generation Challenge Programme's (GCP's) mission is to demonstrate how its efforts lead to direct positive impacts for resource-poor farmers. GCP has discovered that agricultural research scientists are clear on the expected benefits, but they are often very unclear about the process and mechanisms by which their innovations actually get to these farmers. Perhaps most importantly, they are often unaccustomed to considering the various steps between research and field implementation, nor identifying the actors and institutes in the "critical path" between bench and field.

GCP developed a high-quality tool, called the DPKit (for Delivery Plan Kit), to help GCP grantees articulate the expected outputs of their research, and present the logical framework by which these products will be passed on to organisations that in turn support farmers. This project is producing interactive tools to assist scientists involved in GCP programmes to develop high quality "Delivery Plans" based on the current DPKits. The resulting "DP Remote Learning Modules" will allow GCP grantees to successfully complete their required delivery plans to the satisfaction of GCP (and its various stakeholders and funders) without the need of dedicated, in-person, technical assistance seminars.

GCP grantees will be able to follow, from their home country offices, a pre-established sequence of tasks that result in a completed plan that:

- clearly identifies and articulates expected programme outputs
- explains how the programme outputs will translate into improved conditions for resource poor farmers in the target areas and crops of the GCP

- defines the path by which their innovations will travel through "downstream" organisations to reach the organisations that provide direct support to farmers (normally NARS)
- provides detailed commitments and milestones (timing, description, and type of product) by the research team to ensure synchronised efforts and clear expectations.
- can serve as a reference point for GCP management in follow-up and tracking, and to communicate goals and achievements to its various stakeholders.

The development of a remote learning strategy will permit better plans because it will allow more time and opportunity to internalise the concept and its objectives, and allow for pacing of the DPKit development based on the specific conditions of the grantee teams. It will also be considerably more cost-effective than the principal alternatives (seminars, or traveling training teams).

The structure of the DP Remote Learning Modules and DPKit is web-based and interactive (see Figure 1).

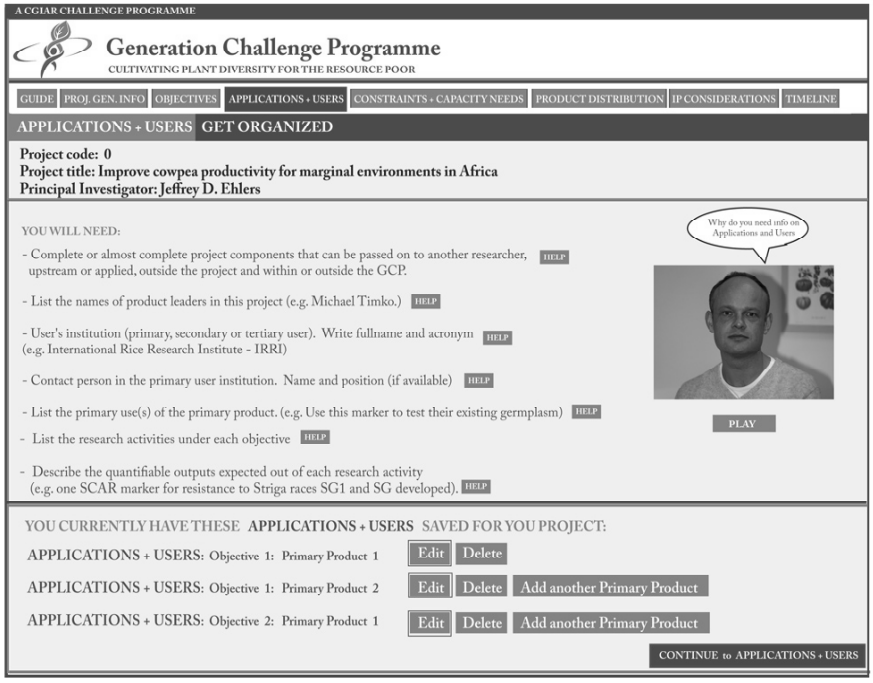


Figure 1: sample page of DPKit and remote learning modules

The structure of new DPkit will be as a relational database, while the remote learning modules will be web-based interactive pages for filling in the DPKit. Each step will be accompanied by explanations and a series of short videos to explain the importance of each step. These will include (a) a scientist explaining why it is necessary to think through the downstream consequences of his or her work; (b) an overview of the information that needs to be provided and the reasoning behind it, with graphics and animation; (c) an interactive component that will assist the users in thinking through the entire "value chain" in which they are contributing their upstream research. It will follow the logic of "look forward, reason backward." It will permit the user to map out the chain backwards from (crop, attribute, farmer, extension, breeder, NARS, bench). (d) This component will address the concept of "what can go wrong?" Using real examples, ideally with humor, of several possible scenarios of "what we are trying to avoid"

It is anticipated that the preliminary online version will be ready by September 2009.

126. G4009.02: Fellowships and travel grants 2009–Study of Burkina Faso rice landraces diversity and breeding for resistance to *Rice Yellow Mottle Virus* (RYMV)

March 2009–March 2010

Principal Investigator

Honoré KAM, INERA–Burkina Faso; kamhonore@yahoo.fr

Institut National de l'Environnement et de la Recherche Agronomique, INERA; 01 BP 910 Bobo-Dioulasso 01, Burkina Faso; tel: (+226) 70 23 11 39; Fax: (+226) 20 97 01 59

Collaborating institutes and scientists

- Agropolis–CIRAD: Nour Ahmadi
- Agropolis–IRD: Alain Ghesquière
- WARDA–Benin: Marie Noelle Ndjiondjop
- UKZN: Mark D Laing

Rice is the staple food in many countries of Africa and constitutes a major part of the diet in many others. A series of abiotic and biotic stresses continue to limit rice productivity. Rice yellow mottle virus (RYMV) is one of the most important rice pathogens in most rice-growing countries of Africa and Madagascar, but not elsewhere. Two types of natural resistance to RYMV have been reported in rice: a partial-resistance in *Oryza sativa* cultivar Azucena and a high-resistance on cultivars Gigante and Tog5681, which represent *Oryza sativa* and *O. glaberrima*, repectively. The high and partial resistances are controlled by a single recessive gene (*rymv*) and several genes, respectively. IRD in collaboration with The Africa Rice Center (WARDA) developed a fine genetic map and the cloning of the high level of resistance and the SNP gene markers tight are used to facilitate the screening of germplasm for their resistance to RYMV. Recently, however, the partial resistance in Azucena has been completely broken down, and high level of resistance in both Gigante and Tog5681 has been overcome by several resistant-breaking-isolates from five countries of the west and central African Sudano-savannah zone. Therefore, there is an urgent need for searching other rice genotypes with high and durable resistance to RYMV in Africa. This project seeks to: (i) genotype 335 accessions collected recently in Burkina Faso with 26 SSR diversity markers used by The Generation Challenge Programme, and (ii) conduct extensive search for durable RYMV resistance among traditional rice accessions from Burkina Faso.

127. G4009.05: Training workshop on Marker-Assisted Breeding

April 2009–September 2009

Principal Investigator and Lead Institute

Carmen de Vicente, GCP; c.devicente@cgiar.org

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

Collaborating institutes and scientists

- IAMZ
- WUR

The ultimate aim in plant breeding is to produce improved varieties by crossing superior parental genotypes. A crucial step is to accurately assess the genotypes from where to select. This task is largely complicated by the genetic complexity behind most of the traits under selection, with the extra difficulty that genotypic responses are usually dependent on the environment (a phenomenon known as genotype by environment interaction, GEI). As a consequence of that, breeding programmes assess the response of genotypes in their dependence on the environment in multi-environment trials (METs).

Advanced statistical procedures together with modern molecular marker techniques can assist improving the accuracy with which breeders can assess genotypic performances. The design of the individual trials within METs is often complicated by the large number of genotypes to be evaluated making difficult to assure homogeneity of field conditions (required for example in standard randomised complete block designs). Special types of designs have been developed, which combined with powerful statistical methodologies improve the accuracy of genotypic evaluations. Powerful models can also be used in analysing METs to describe GEI, among which well-known exploratory models are Finlay-Wilkinson and AMMI models. Alternatively, factorial regression models allow the incorporation of explicit genotypic and environmental covariables to describe GEI. Factorial regression within a mixed model context provides a suitable framework for the mapping of quantitative trait loci (QTLs), including extensions to QTL by environment interaction (QEI). An asset of this modelling framework is that can equally be applied to the analysis of segregating populations (traditional QTL analysis) and to unstructured populations in linkage disequilibrium (LD) analysis.

The objective of the course is to introduce the participants to the analysis of single and MET trials using mixed models, and demonstrate how these models can be extended to detect QTL and QEI. The methods will be illustrated both in the context of conventional QTL mapping (i.e. using designed segregating populations) and in the context of LD mapping (i.e. using diverse populations). The course will also touch on the crucial steps of molecular map construction as well as on the analysis of population structure in diverse populations.

The following were participants in the course:

No.	Name	Home institute	Country	Crop
1.	Mashamba Phillipo	ARI–Naliendele	Tanzania	Groundnut
2.	Issa Faye	ISRA–UCAD-ISRA/CNRA Peanut Breeding Program	Senegal	Groundnut
3.	Tobias Kapewa	DARS	Malawi	Groundnut
4.	Issa Karimoune	INRAN	Niger	Groundnut
5.	Abdou-Razakou Bio Yérima	INRAN	Niger	Groundnut
6.	Ndeye-Ndack Diop	ISRA	Senegal	Cowpea
7.	Jean-Baptiste Tignegre	INERA–Burkina Faso	Burkina Faso	Cowpe
8.	Rogerio Chiulele	EMU	Mozambique	Cowpe
9.	Eugene Agbicodo	IITA	Benin	Cowpea
10.	Wellington Muchero	UoC–Riverside	USA	Cowpea
11.	Asrat Asfaw Amele	AARC	Ethiopia	Bean
12.	Godwill Makunde	CBI	Zimbabwe	Bean
13.	Lizzie Kalalokesya	SABRN	Malawi	Bean
14.	C Sivakumar	ICRISAT	India	Chickpea
15.	Richard Mulwa	EgU	Kenya	Chickpea
16.	S K Chaturvedi	IIPR	India	Chickpea
17.	Paul Kimurto	EgU	Kenya	Chickpea
18.	Mohammad Reza Naghavi	Agricultural College, UoT	Iran	GSS (Germplasm characterisation)
19.	Joyce Malinga	KARI	Kenya	GSS (Germplasm characterisation)

FOCUS PROJECTS

128. G6007: Tropical Legumes I (TLI): Improving tropical legume productivity for marginal environments in sub-Saharan Africa

May 2007–April 2010

Principal Investigator and Lead Institutes

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**

In Year 2, TLI's workplan advanced at full speed, with remarkable progress being made overall.

Basically all milestones were completed, with the remaining being slower because of their dependency on the season for field experiments or a need for further validation, rather than delays per se. All Objectives exceeded at least some of their expected milestones.

Below are a few of the Year 2 research highlights:

Diversity studies

The Objectives evaluated the respective reference collections, identified promising germplasm for traits of interest to partner countries, and began developing populations for mapping or breeding.

Genomic resources

- 3,200 microsatellite markers, with at least 142 associated with disease resistance and the first cultivated groundnut genetic map, are now available to the groundnut community.- 1,536 high-confidence cowpea SNPs were selected for an Illumina® GoldenGate Assay and the first full physical cowpea map assembled.
- 1,532 genomic and gene-based microsatellites were created from cDNA, small insert libraries and BAC end sequences for common bean.
- An expanded chickpea DArT array with 15,360 clones is now available and the first chickpea Illumina® GoldenGate Assay (768 SNPs) made.
- 16,000 SNPs were identified, with an average of 616 genes per species.
- A database was created to contain all DNA sequences and SNP data related to comparative markers for the four legumes.

Marker discovery and validation

- The first QTLs for disease resistance in groundnut were identified.
- In cowpea, two QTL each was identified for resistance to flower thrips and for resistance to the fungus *M. phaseolina*, and one QTL for root-knot nematode.

- Several bean populations were screened with a QTL based marker for common bacterial blight resistance.
- Two putative QTLs were identified for resistance to the *Helicoverpa* pest in chickpea.
- Minor QTLs for drought-related traits were identified in three groundnut populations.
- QTLs for drought are currently being validated in cowpea, already evaluated in bean populations, and identified in chickpea.

Pre-breeding

- Nine groundnut elite varieties were identified as good candidates for MABC.
- Elite × elite crosses for MARS have been advanced in cowpea (15) and chickpea (7).
- 2,524 BC2F1-derived segregants were obtained for five advanced backcrosses between drought-tolerant bean parents and three commercial Andean cultivars from Africa.

Capacity building

- Infrastructure, particularly for field phenotyping, was secured for all African partner institutions with support from Objective 6.
- Particular Objective under-expenditures for Year 1 were used to help build up infrastructure at partner institutions in Africa and conduct training activities for students working with TLI.
- Principal Investigators visited African institutions more often and monitored their field activities more closely. This led to positive results in terms of engagement and team building. The involvement of African partners in data analysis also increased.
- GCP has leveraged funds for training and for topping off some infrastructure needs.

128.01 G6007.01: TLI Objective 1–Improving groundnut productivity in marginal environments of sub-Saharan Africa

May 2007–April 2010

Principal Investigator

Vincent Vadez, ICRISAT; v.vadez@cgiar.org
GT Biotechnology, Patancheru 502324, Andhra Pradesh, India –(91) 40 30 71 34 63; (91) 40 30 71 30 74

Collaborating institutes 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
- CRS: Tobias Kapewa
- UGA: Andy Paterson
- UCB: David Bertioli
- EMBRAPA: Soraya Bertioli; Patricia Guimares

Context

Groundnut (peanut, *Arachis hypogaea* L.), an important food and cash crop in Africa, is often grown by smallholder farmers under very low input and rainfed drought-prone conditions. The project's objective is to improve disease resistance and drought tolerance of farmer-preferred groundnut varieties using modern molecular tools, and involving African NARS. The project involves the screening of representative germplasm for disease resistance and drought tolerance, developing the molecular tools needed for faster/more efficient introgression of beneficial traits in cultivated groundnut, using these tools for identifying the genome regions involved in resistance/tolerance, and eventually introgress disease resistance and drought tolerance into farmer-preferred varieties.

Findings and implications

Diversity assessment – A large and representative set of groundnut germplasm was tested in India and Africa for drought and disease. Large range of variation was found for transpiration efficiency in India (1.93-2.65 g DW kg^{-1} water), with low TE in popular JL24 and TMV2. A yield test under intermittent drought in Niger showed a 0.5 – 1.7 t/ha range of variation, with 30 new sources above popular variety 55-437 and Fleur11 (1.44, and 1.51 kg/ha). Thirty lines were identified with disease scores lower than 5 (from 1, resistant to 9, sensitive), compare to a score of 9 of polular varieties JL24 and Fleur 11.

Genomic tools – Physical mapping is in progress. Pilot hybridisation has been completed with RLFP markers. Hybridisation is on the way with comparative markers. Fingerprinting is on-going. New BAC clones will supplement the original library.

Sequencing of a long SSR enriched library has been completed and 139 functional primer pairs were developed, of which 83 are polymorphic for a panel of 22 diverse cultivated peanut. About 3200 SSR markers are now available to the groundnut community.

Disease mapping - 34 sequence-confirmed candidate disease resistance genes and five QTLs were mapped onto the wild Arachis AA-genome. Groupings of candidate genes and QTLs for late leaf spot resistance were apparent on the upper region of linkage group 4 and the lower region of linkage group 2, indicating that these regions are likely to control disease resistance. F2/3 populations derived from disease resistant lines and farmer preferred varieties being advanced to RILs for disease resistance mapping.

Drought trait mapping - RIL ICGV86031xTAG24 was re-phenotyped for TE under lower VPD, and large variation was obtained for TE (2.8-5.2 g DW kg^{-1} water). Seeds of that population were sent to Senegal, Malawi and Tanzania for field and TE assessment there. RIL ICGS76 x CSMG84-1 phenotyped for TE (range 1.0-3.0 g DW kg^{-1} water). Field assessment of RIL ICGS44 x ICGS76 is on-going.

202 polymorphic markers screened in RIL of ICGV86031xTAG24 with 165 loci mapped. 83 polymorphic markers screened in RIL ICGS44 x ICGS76 and 63 loci mapped. 119 polymorphic markers screened in RIL ICGS76 x CSMG84-1 and 84 loci mapped. 59 QTL, most with small effect, identified for several drought related traits, including TE.

Development of breeding populations - 28 populations between farmer-preferred parents and disease resistance sources are being developed: (i) 12 for rosette; (ii) 8 for ELS, (iii) 8 for rust. Nineteen of these populations are being phenotyped at F2 and all of them advanced to BC1F1. These populations are also advanced to F3 (F2/3 populations for disease resistance QTL mapping and advancement to RILs should be completed within about 2 years.

Lessons learnt

Transfer of material, imports permit have been a constraints in certain cases. Enhancing physical and human capacity is time consuming (especially for phenotyping issues), like setting up / refurbishing new irrigation facilities, but efforts appeared to be rewarding. Capacity to use markers for breeding (eg to follow backcrossing work) is a limitation. Time limited to fully analyse phenotyping data

Next steps and/or challenges

Third year will repeat germplasm assessment to confirm Year 2 data. The physical mapping should be completed. Mapping of drought related trait will continue actively while disease resistance populations will be further advanced towards RILs. Challenge here will be to find resources for genotyping of the large number of populations being developed. The breeding activities consisting in disease resistance and drought tolerance introgression should take more and more importance and link up with the Tropical Legume II project.

128.02 G6007.02: TLI Objective 2–Improve cowpea productivity for marginal environments in Sub-Saharan Africa

May 2007–April 2010

Principal Investigator and Lead Institute
Jeff Ehlers, Dept. of Botany and Plant Sciences, UoC–Riverside; jeff.ehlers@ucr.edu
900 University Ave., Riverside, CA 92521-0124, USA; 951-323-5918; Fax 951-827-4436.

Collaborating institutes and scientists

- UoC–Riverside: Timothy Close; Philip Roberts
- UoC–Davis: Ming-Cheng Luo
- ISRA: Ndiaga Cisse
- IRAD: Ousmane Boukar
- INERA–Burkina Faso: Issa Drabo
- IITA: Satoru Muranaka; Ousmane Boukar; Christian Fatokun; Dong-Jin Kim

Our project seeks to enable efficient application of modern breeding of cowpea in sub-Saharan Africa. Progress during months 13 through 24 of this 36 month project are reviewed below for each of the 5 project activities:

Activity 1. Characterise diversity and develop germplasm for genetic studies. Drought phenotyping of 500 genotypes was completed early in year 2 and a subset of the most promising 200 genotypes selected and evaluated for grain yield and drought tolerance traits trials in Senegal, Burkina Faso, Nigeria and California. Twenty drought tolerant genotypes were identified; association genetic analysis combining phenotypic and marker data (Activity 2, Milestone 5) has been initiated to identify drought tolerance QTL.

Activity 2: Generate genomic resources for genetic studies and breeding. An assembly of 183,118 ESTs yielded ~10,000 high confidence SNPs, 1536 of which were selected for an Illumina GoldenGate Assay. The 1536-plex GoldenGate Assay was applied to 1632 DNA samples including 981 RIL from eight populations and 640 germplasm accessions. 1375 SNPs (90%) were technically successful, among which 991 SNPs from seven RIL populations were placed on a consensus genetic map spanning 680 cM (0.69 cM average marker distance) (Muchero et al 2009b). The seven individual maps ranged from 600 to 665 cM. A 17x genome coverage BAC library produced in year 1 from African genotype IT97K-499-35 was used to produce an online physical map (http://phymap.ucdavis.edu/cowpea) containing 43,717 BACs (11X coverage) with an average of 52 BACs per contig and only 5.8% singletons (A minimal tiling path of the physical map was anchored to the genetic map using the 1536-plex GoldenGate assay. The final ~4,000 BAC end sequences (BES) were completed, taking the total to 30,611 BES (674 base average length, 20.6 Mb, 3.3% of the 620 Mbp genome). 39% of BES had a high BLAST hit to non-transposon plant sequences.

Activity 3: Identify molecular markers and genes for biotic stress resistance. Field-based thrips resistance phenotyping was conducted in Cameroon and Senegal with two RIL populations. Results in Cameroon were inconclusive due to excessive numbers of thrips/flower (>30) overwhelming the resistance. Data from Senegal provided sufficient discrimination among parents and RILs for QTL analysis. Two RIL populations were phenotyped for resistance to bacterial blight and a virus complex (mainly CAbMV and CSMV) at Maroua in Cameroon. Two other RIL populations were phenotyped for Fusarium wilt in greenhouse tests and root-knot resistance in field and growth chamber pouch evaluations in California. A RIL population was phenotyped for resistance to *Macrophomina phaseolina* (ashy stem blight) in an infested field in California.. These phenotype data were used to identify the following QTL: two QTL for resistance to flower thrips, QTL for resistance to foliar thrips (Muchero et al 2009a), three major and six smaller QTL for root-knot nematode resistance, and three QTL for resistance to *Macrophomina*.

Activity 4: Identify molecular markers and genes for drought tolerance. Three RIL populations were phenotyped for drought tolerance in California and Senegal for QTL analysis. A drought trial at Saria

in Burkina Faso was conducted but compromised by prolonged and excessive rainfall. The phenotype data were used to identify ten QTL for drought response explaining 4.7 to 24.2% of the trait variation.

Activity 5: Enhance locally adapted germplasm with target traits. Sequence information around ten SNPs linked to three traits (response to CpMV, CpSMV and chilling at emergence) were used to design primers for allele-specific PCR for the purpose of testing marker conversion methods. Fifteen elite x elite **c**rosses for future MARS breeding were advanced 2 generations.

Tangible outputs

1) A high-throughput SNP genotyping platform was developed. This represents a major advance in tools available for marker-assisted breeding.

2) A consensus genetic linkage map with 991 SNPs from seven RIL populations was produced, from which legume synteny is readily apparent.

3) A physical map of cowpea, anchored to the genetic map, provides a link to an average of seven BAC-end sequences per cM for additional marker development.

4) SNP markers linked to determinants of several abiotic and biotic stress tolerance and resistance traits relevant to Africa were identified.

References

Muchero W, Ehlers JD, and Roberts PA (2009a). QTL analysis for resistance to foliar damage caused by *Thrips tabaci* and *Frankliniella schultzei* (*Thysanoptera: Thripidae*) feeding in cowpea [*Vigna unguiculata (L.) Walp*.]. *Mol. Breed*. DOI 10.1007/s11032-009-9307-6 (Published online 7 July 2009).

Muchero W, Diop NN, Bhat PR, Fenton RD, Pottorff M, Hearne S, Cisse N, Fatokun C, Ehlers JD, Roberts PA, and Close TJ (2009b). A consensus genetic map of cowpea [*Vigna unguiculata* (L) *Walp*.] based on EST-derived SNP markers and six RIL populations, and synteny with reference genomes. (submitted May 2009).

128.03 G6007.03: TLI Objective 3–Improving common bean productivity for marginal environments in sub-Saharan Africa

May2007–April 2010

Principal Investigator and Lead Institute

Matthew Blair, CIAT; m.blair@cgiar.org
International Center for Tropical Agriculture, AA6713, Cali, Colombia

Collaborating institutes and scientists

- CIAT: Steve Beebe; Manabu Ishitani; Idupulapati Rao
- ECABREN/UoN: Paul Kimani; Felix Waweru
- DR4D: Godwill Makunde
- RIKEN: Motoaki Seki; Kazuo Shinozaki
- SABRN: Rowland Chirwa; Lizzie Kalolokesya
- SARI: Asrat Asfaw Amele

Background: Common bean are a major food legume crop for drought affected regions of Eastern and Southern Africa where marketable yields are also influenced by viral and bacterial diseases as well as storage insects. Typical bean yields are only 20 to 30% of the genetic potential of improved varieties due to these major production constraints. Drought is especially important in Sub-Saharan Africa as its frequency and severity are on the increase. Participants in the Tropical Legumes I project for common bean have worked on five activities of yield improvement for the crop: germplasm enhancement for drought, genomic tool development, biotic constraint breeding, analysis of drought tolerance mechanisms and advanced breeding for drought tolerance.

Findings and implications: In the first activity, germplasm and genetic resources, trials have been established for a reference collection of common bean diversity in CIAT and at ECABREN/SABRN sites. The full reference collection consists of 288 genotypes but was reduced to a total of 212 bush beans that have been adapted to mid elevation sites. As part of this activity, six new recombinant inbred line populations have been developed to the F5 generation. Regional varieties from ECABREN (87) and SABRN (121) networks were evaluated in Kenya and Malawi and distributed to the same sites as for the reference collection.

In the second activity, for the development of genomic resources, candidate gene markers and BAC end microsatellites have been tested and mapped. A set of AT-rich, small insert based microsatellite markers have been well-characterised. Two subtractive root and shoot tissue libraries (drought versus control) were developed for the genotypes DOR364 and BAT477 with an additional library in preparation for tepary bean versus common bean drought tolerant lines. Full-length cDNA library construction proceeded as planned with RNA prepared from drought and control treatments for genotypes BAT477 and G19833 and pilot sequencing has begun.

In the third activity, genetic crosses have been made for marker assisted selection of bruchid, common bacterial blight (CBB) and bean common mosaic necrosis viruses (BCMNV) resistance. A total of 236 F1-derived families were screened for arcelin presence indicating bruchid resistance in combination with CBB or BCMNV resistance genes. For combining with drought, a total of 495 selections were screened for the arcelin marker (256 positive, 181 fixed) in small white beans for Ethiopia.

In the fourth activity two QTL mapping populations were phenotyped across 3 countries and 6 different sites under terminal or intermittent droughts. DOR364 x BAT477 population was tested in Awassa (early and late rainfed plantings) and Amaro (short rains), Ethiopia and Chitedze, Malawi for drought tolerance. BAT881 x G21212 population was tested in Kiboko (off-season), Mwea and Thika (short rains) in Kenya for photosynthate mobilisation under irrigated and rainfed treatments. The primary plant traits measured included grain yield and yield components at harvest (number of pods per plant, number of seeds per pod, and 100 seed weight) at all sites. Canopy dry weight per plant at mid podfill and dry matter distribution (vegetative versus pods) were measured at the three Kenyan sites. Harvest index, pod partitioning index and pod harvest index was also determined. In addition, data were analysed for the rooting depth evaluation in the greenhouse with the DOR364 x BAT477 population under two levels of water supply: 80% field capacity (well-watered) and withholding of watering (to simulate terminal drought stress conditions).

In the fifth activity, 22 Andean x Andean (AA) and 17 Andean x Mesoamerican (AM) cross populations based on North Carolina Design II crosses between elite drought tolerant genotypes and commercial varieties advanced to the F7 generation under drought stress with 1500 selection made across CIAT and Zimbabwe environments and 347 genotypes now in Colombia (DAB lines) or Zimbabwe (CBIB lines). For the development of advanced backcross populations, 2524 BC2F1 seed were obtained for 5 crosses between two drought tolerant donor parents: SER16 and SER48 (both good combiners from the Mesoamerican gene pool) and three commercial Andean cultivars from Africa used as recurrent parents: CAL 96, CAL 143, and CIM 9314-36.

Tangible output delivered: Advanced lines from recombinant inbred line populations or inter- and intra-genepool drought crosses were the first tangible outputs of the project along with constitution of the reference and regional collections which were widely tested. Benefits have been to national research programmes in Ethiopia, Kenya, Malawi and Zimbabwe with spillover to Tanzania both in terms of research methodologies and germplasm that will be useful for breeding programmes.

Next steps and/or challenges: Genomic tool development has proceeded well and we will be analysing the results of full-length cDNA sequencing. Germplasm development challenges are managing the large number of populations for marker aided selection which will require the further development of user-friendly markers. Incorporation of drought tolerance traits from various sources will emphasise pyramiding of translocation efficiency and deep rooting traits with physiological measurements and QTL selection.

128.04 G6007.04: TLI Objective 4–Improved chickpea productivity for marginal environments in sub-Saharan Africa

May 2007−April 2010

Principal Investigator and Lead Institute
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 30713305; Fax: 0091 40 3071 3074/ 3075

Collaborating institutes and scientists
- ICRISAT–India: Pooran Gaur; Hari Sharma; Junichi Kashiwagi; L Krishnamurthy; C Sivakumar; T Mahender
- ICRISAT–Kenya: Said Silim; NVPR Gangarao
- EIAR–Ethiopia: Million Eshete
- LZARI–Tanzania: Robert Kileo
- IIPR: Masood Ali; Aditya Kumar
- UoF: Peter Winter; Guenter Kahl
- UoC–Davis: Doug Cook
- Diversity Array Pty Ltd: Andrzej Killian

Summary
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. The major goal of this project is to improve incomes and livelihoods by establishing the capacity to improve the drought tolerance and insect (pod borer) resistance in farmer-preferred chickpea varieties using modern molecular tools.

The research progress made under five activities during 2008- 2009 is presented here.

Germplasm for genetic studies and modern breeding : A total of 291 accessions of the reference set were evaluated for root traits, harvest index (HI), and pod borer resistance in India (Patancheru). High and significant levels of genetic variability were observed for all these traits such as overall insect resistance (scores ranging from 5.0 to 9.0) and combined data analysis of root traits for 2007 and 2008. A set of 289 accessions of the reference collection were also evaluated at Egerton University (Kenya), EIAR (Ethiopia), and IIPR (India) for HI, seed mass, and maturity. As a result promising genotypes were identified for use in breeding programmes in different countries. For instance, ICC 7272, ICCV 92311, ICCV 10, and ICC 14595 genotypes for India; ICC 8200, ICC 1510, ICC 15248, ICC 3325, and ICC 10393 accessions for Kenya; and ICC 11198, ICC 4495, ICC 7668, ICC 4918, and ICC 3325 accessions in Ethiopia should be proven useful. Moreover, phenotyping of 80 genotypes (coming from the mapping populations as well as the reference set) was completed for HI, seed mass, and maturity at Egerton University (Kenya), EIAR (Ethiopia), and IIPR (India). The promising genotypes identified were, ICC 1052, ICC 4958, ICC 15333, ICCRIL04-0239, and ICCRIL03-0135 fo Kenya; ICCRIL03-0168, ICCRIL03-0041, ICC 708, and ICC 1882 for India; ICC 4958, ICC 14435, ICC 14199, ICCRIL03-041, and ICCRIL04-0189 for Ethiopia. Multi-location evaluation of the International Chickpea *Helicoverpa* Resistance Screening Nursery in Kenya, Ethiopia, and India resulted in the identification of the following promising lines, EC 583250, EC 583264, EC 583311, ICC 144402, and ICC 16903 (Kenya); ICC 10393, ICC 1356, ICC 506, ICC 14402, and ICCV 10 (India).

Genomic resources for genetic studies and modern breeding: A new DArT array with 7,680 features was developed from 96 diverse genotypes was found monomorphic. Hence, an expanded DArT array with 15,360 clones became available for chickpea. A large SNP dataset (26,082) was also identified, and 134 SNPs were used to develop the first Illumina® GoldenGate Assay (768 SNPs) for chickpea in collaboration with the *Objective 5* team. A reference genetic map with 526 loci was developed for the interspecific mapping population *Cicer arietinum* ICC 4958 × *C. reticulatum* PI 489777. The loci included newly developed SSR markers, gene-based markers (UC–Davis, USA), and public domain markers.

Molecular markers and genes for biotic stress resistance: A total of 128 lines of the mapping population ICC 4958 × PI 489777, along with two checks- ICC 506 (resistant) and ICCC 37 (susceptible) were evaluated for resistance to the pod borer *H. armigera*. The detached-leaf assay and natural infestation were used to assess leaf damage, egg and larval numbers, pod damage, and grain yield. The detached-leaf assay indicated that larval weight gain varied from 1.7 to 9.0 mg per larva in the mapping population. The genetic map developed as mentioned above for 526 loci was used to conduct QTL analysis with insect resistance phenotyping data for 2007 and two putative QTLs were identified. Phenotyping data obtained for 2008 will be analysed with genotyping data to identify the QTLs for 2008 as well as consistent QTLs for 2007 and 2008.

Molecular markers and genes for drought tolerance: We evaluated 281 RILs from the cross ICC 283 × ICC 8261 for root traits (depth, length density, volume, and dry weight) and shoot weight. These RILs were also evaluated for HI, phenology, shoot biomass, and yield in India (Patancheru). SSR marker genotyping of the population ICC 4958 × ICC 1882 was completed with 295 SSR markers and the genetic map was developed for 225 SSR loci. At least six QTLs for root traits were identified in population ICC 4958 × ICC 1882 by using root-trait phenotyping data for years 2005 and 2007, and 225 SSR marker data. Among these QTLs, one QTL flanked with the marker TAA170 was found very promising as this was identified for a number of component root traits and was also stable across two environments.

Improvement of locally adapted germplasm for target traits through modern breeding: For introgression the root trait QTL (marker TAA 170), marker-assisted back crossing (MABC) was undertaken in one desi (ICCV 93954; donor parent ICC 4958) and two kabuli cultivars (ICCV 92318 and ICCV 92311; donor parent ICC 8261)……

128.05 G6007.05: TLI Objective 5–Development of cross-species resources for comparative genomics in Tropical crop legumes

May 2007−April 2010

Principal Investigator and Lead Institute

Douglas R Cook, UoC−D; drcook@ucdavis.edu

University of California−Davis, One Shields Ave, Davis, CA 95616, USA; (office) 530-754-6561, (fax) 530-754-6617.

Collaborating institutes and scientists

- UCB: David Bertioli
- UGA: Andrew Paterson

Activity 1: Comparative Marker Development

We developed a set of 1440 degenerate primer pairs to amplify 1369 tentative orthologous genes (TOGs) from the crop legumes chickpea, cowpea, common bean, and peanut. We completed sequencing and analysis of these 1440 amplicons across the parents of mapping populations in each species and obtained high quality sequences for a majority of the TOGs in all species, ranging from a low of 1043 in peanut to a high of 1251 in chickpea. 950-1200 TOGs provide the basis for pair-wise comparisons between species.

Within-species sequence alignments yielded ~16,000 single nucleotide polymorphisms (SNPs) in ~2500 genes (average of 616 genes per species). A relational database has been developed that contains all DNA sequences and SNP data related to the comparative markers. SNPs with high assay design scores were used to develop Illumina GoldenGate assays for each species. Genotyping and comparative map development are currently in progress.

Project investigators have also worked to integrate bacterial artificial chromosome (BAC) resources and genetic maps in each species. >800 orthologous markers have been linked the diploid peanut physical map based on overgo hybridisation. In diploid peanut, chickpea and cowpea, a combined total of >4,000 BAC end-associated simple sequence repeats (SSRs) are being analysed for suitability as molecular genetic markers; genetic mapping of these SSRs is serving to integrate BAC resources to genetic maps. A similar, but smaller effort is underway in common bean, focusing on resistance gene-associated BAC SSR markers.

Activity 2: Analysis of the Arachis-species complex
Twenty five thousand BAC clones were end sequenced, resulting in 41,856 sequences and ~25 Mbp of data. >4K SSRs were identified and PCR primers were designed for a subset of 1535 SSRs, including 142 resistance gene-associated SSRs. These SSR markers will initially be tested for amplification, followed by re-array and testing for polymorphism across a panel of diploid and cultivated tetraplid genotypes.

Ultra-long SSRs have been mined from SSR-enriched libraries in cultivated peanut, resulting in 147 candidate SSRs of which 83 have been determined to be polymorphic. These markers increase the total number of SSRs available in cultivated peanut by ~25%.

Activity 3: Estimating genome divergence at orthologous loci
To estimate fine scale genome divergence, conserved loci were selected for BAC clone isolation and DNA sequencing in each species. Target loci include one member of the Ara allergenic gene family of peanut; one member of the Dreb2 gene family implicated in drought response; and the dmi2 gene, a key innovation leading to nodulation. Additional BACs for arbitrary RFLP loci are slated for comparison in diploid and tetraploid peanut. Candidate BACs have been identified for AraH1-3 in peanut, and Dreb2 in Phaseolus. Dmi2 screens are in progress. Eight peanut BACs have been sequenced, and the first Phaseolus BAC is in validation, providing sequence to guide selection of maximally overlapping BACs from the other taxa.

128.06 G6007.06: TLI Objective 6–Provide training and capacity-building for Sub-Saharan African scientists

May 2007−April 2010

Principal Investigator and Lead Institute
Carmen de Vicente, GCP, c.devicente@cgiar.org
Apdo. Postal 6-641, 06600 Mexico, D.F., Mexico**,** Phone: +52 55 5804 2004

Collaborating institution and scientists
GCP: Nosisa Mayaba (Associate Scientist – left June 2008); Alejandro Boeta (Administrative Coordinator – left 2009)

Activity 1.
The first annual meeting took place between 30 June and 3 July 2008 in Dakar, Senegal. The meeting was a great opportunity for the entire team to discuss progress achieved during Year 1, revise workplans for Year 2, and identify capacity-building and training needs of NARS partners.

The second annual meeting took place between 16 and 20 April 2009 in Lilongwe, Malawi. In addition to updating the entire team on progress achieved during Year 2 and providing information on the workplans for Year 3, time was allocated for the Objective teams to prepare this annual report to encourage the participation of all team members in this task. Very importantly, it also allowed brainstorming for the preparation of the proposal for the TLI Project's second phase (2010–2014).

A course devoted to the analysis of phenotyping and marker data has been organised and is planned for 29 June–3 July in Zaragoza (Spain). Logistic support will be provided by CIHEAM's Instituto Agronómico Mediterráneo of Zaragoza (IAMZ). Biometris of Wageningen University and Research Centre will be in charge of instruction. We expect 22 TLI-related participants to take the course.

Activity 2
The prioritised list of capacity-building needs for Year 2 was compiled, new contracts with NARS institutions were signed, and funds are now being disbursed. Funds are still to be disbursed to EIAR (Ethiopia) and DARS (Malawi) because the requirement of submitting several quotations has not yet been fulfilled.

Regular contacts have been maintained between the PIs of TLI and TLII, as well as between the PIs of both projects. As a result, research components of TLI have been incorporated into the research of TLII-funded students (Objectives 1, 3, and 4).

Extra funds are being secured through the GCP to cover human resources development for NARS partners. Final decisions still pending, but the expectation is to benefit either degree students or technical staff in Objectives 1 to 4.

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

Collaborating institutes and scientists
- IRRI: R Serraj; T Paris; S Haefele; R Anitha
- IGKV: SB Verulkar; PR Dongre
- CRRI: ON Singh; P Swain; L Bose
- CRURRS: PK Sinha; NP Mandal
- NDUAT: JL Dwivedi
- UAS–Bangalore: S Hittalmani
- TNAU: R Chandrababu; S Robin
- BAU: BN Singh; RL Mahato
- Barwale Foundation, India: HE Shashidhar; Abhinav Jain

Research activities carried
During the four years of the project period, more than 2,000 advanced breeding lines coming from IRRI as well as eight different partners from India in the Drought Breeding Network (DBN) and eight different partners from India in Upland Rainfed Shuttle Breeding Network (URSBN) have been evaluated under severe reproductive-stage drought stress, moderate drought stress, and control irrigated situations in lowlands under DBN and in uplands in URSBN.

Progress in developing varieties for drought prone lowlands

In the DBN, from the evaluation of more than 1,300 advanced breeding lines, several promising breeding lines with high yield potential and good yield under severe drought stress were identified. The identified promising lines provided yield advantage of 0.7 to 1.0 t ha^{-1} over the presently grown varieties IR64 and IR36 under severe drought stress situation while maintaining the high yield potential under normal irrigated situation (Verulkar et al., 2009). In participatory varietal selection (PVS) trials on farmers' fields at Raipur, Hazaribag, Ranchi, Pusa, and Faizabad, IR70215-70-CPA-2-4-1-3, NDR1045-2, IR55419-04, and IR74371-70-1-1 were identified as farmers' preferred early duration breeding lines whereas RRF-23 and RRF-69 as farmers' preferred medium duration breeding lines. Promising drought-tolerant breeding lines were nominated for testing under the All India Coordinated Rice Improvement Program (AICRIP). Breeding line NDR1045-2 has been released as drought-tolerant variety "Shusk Samrat" in the state of Uttar Pradesh province and IR74371-70-1-1 has been identified by AICRIP for release as a central variety "Sahbhagi dhan". New improved and traditional sources of drought tolerance under lowland-IR77298-14-1-2, Dagad Deshi, Laloo 14 Jonga, Kallurundkar, Baranideep, Lalmati, Kalamkati, Sadabahar, Khiradhan, and Mattaikar were identified for use by the different national breeding programmes.

Table 1: Overall mean of 100-120 days duration breeding lines pooled across environments (location × year × stress level) under irrigated control and managed-stress conditions for 2005-2007.

Breeding lines	DTF	Grain yield (t ha^{-1})		
		Control	Moderate	Severe
NDR 1045-2	89	5.333	3.863	1.438
IR79906-B-5-3-3	92	5.77	3.125	1.637
IR70213-10-CPA-4-2-2-2	91	5.508	3.402	1.566
IR70215-70-CPA-3-4-1-3	84	4.843	3.214	1.942
IR74371-54-1-1	86	5.206	3.401	1.543
IR78937-B-4-B-B-B	91	5.531	2.699	1.735
IR74371-70-1-1	84	5.167	3.233	1.574
IR55419-04	84	4.647	3.266	1.802
IR64	87	5.084	2.601	0.872
IR36	87	4.363	2.249	0.716
$SED_{0.05}$	4.1	0.4972	0.4962	0.3361
% reduction compared to control			42	73

Progress in developing varieties for drought prone uplands

Under URSBN, over four years, more than 700 breeding lines were evaluated under upland conditions and breeding lines with at least 0.5 t ha^{-1} yield advantage over the presently grown varieties Anjali, Vandana were identified. RR 347-2, an elite line identified by the network, has been released by the Central Variety Release Committee (CVRC) of India as "Virendra" for the uplands of Gujarat and Orissa. DDR 97 (IET 19258) has been released for cultivation in Gujarat (Mandal et al. 2009). New sources of drought tolerance- VLDT 1, VLDT 2 and Sukhawan for uplands were identified.

References

Verulkar SB, Mandal NP, Dwivedi JL, Singh BN, Sinha PK, Mahato RN, Swain P, Dongre P, Pyasi D, Singh ON, Bose LK, Robin S, Chandrababu R, Senthil S, Jain A, Shashidhar HE, Hittalmani S, Vera Cruz C, Paris T, Hijmans R, Raman A, Haefele S, Serraj R, Atlin G and Kumar A (2009) . *Breeding resilient and productive rice genotypes adapted to drought-prone rainfed ecosystems of India* (under preparation)

Mandal NP, Sinha PK, Variar M, Shukla VD, Perraju P, Mehta A., Pathak AR, Dwivedi JL, Rathi SPS, Bhandarkar S, Singh BN, Singh DN, Panda S, Mishra NC, Singh YV, Pandya R , Singh MK, Sengar RBS, Bhatt JC Sharma RK, Anitha K, Kumar A. and Atlin G (2009). *Implications of genotype X input interactions on breeding superior genotypes for favorable and unfavorable rainfed upland environments* (Under preparation)

130. 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 and Lead Institute

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 institutes 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_2 progeny selected from a cross between 2 DT IR64 ILs segregating for 40 DT QTL regions were genetically analysed with DNA markers. One dramatic result was that all segregating DT loci in the 90 DT F_2 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)

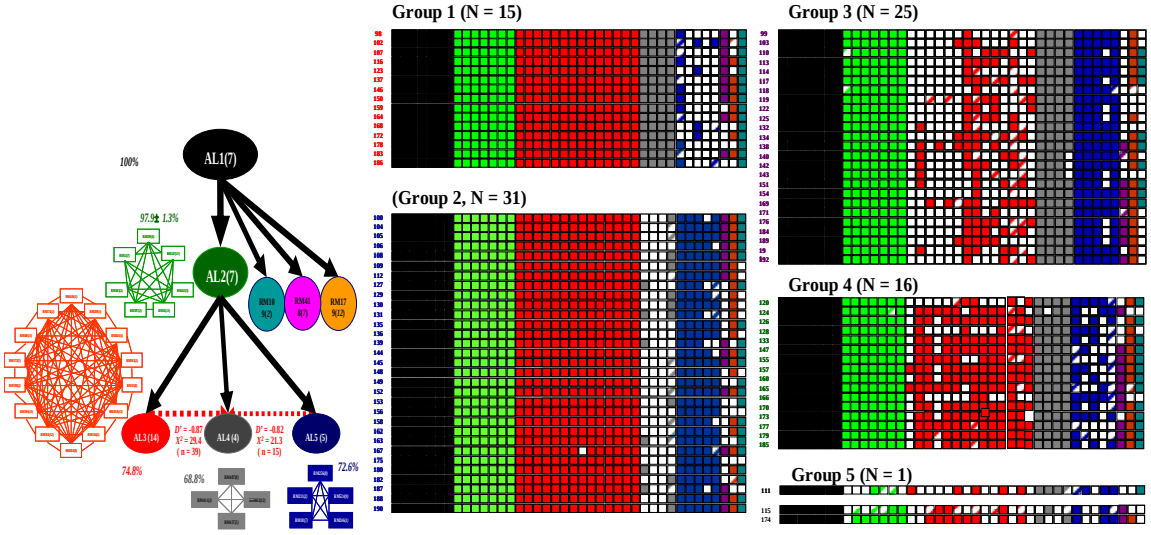


Fig. 1. a: The genetic network containing 40 DT QTLs detected in 90 DT F2 progeny; b: the graphical genotypes of the 4 major group genotypes.

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 analysed 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.

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.

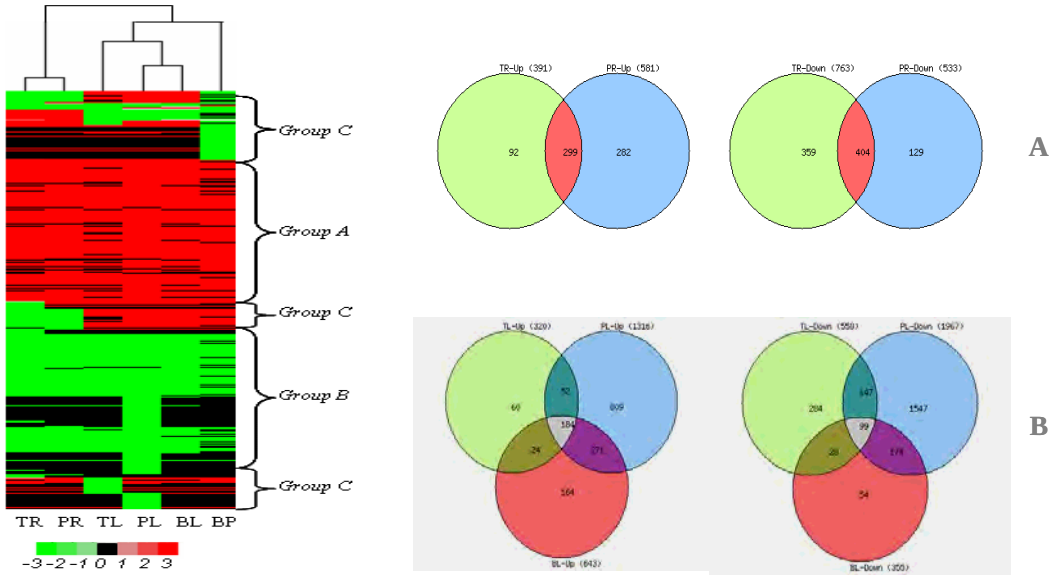


Figure 1. A. Hierarchical cluster analysis of six tissue types (columns) and all DEGs under drought stress (rows). TL, PL and BL indicate leaves at the tillering, panicle elongation and booting stages, respectively; TR and PR indicate roots at the tillering stage and panicle elongation stage; BP is young 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.

131. 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 and Lead Institute
Yunbi Xu, GREU–CIMMYT; y.xu@cgiar.org
Apdo. Postal 6-641, 06600 Mexico, D.F., Mexico. Tel: +52(55) 5804-2004
Fax: +52(55) 5804-7558

Collaborating institutes and scientists

- GREU–CIMMYT: Carlos Martinez; Debra J Skinner; Alan F Krivanek; 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.

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_2 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).

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 (Sun et al 2009. 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. A reverse breeding-to-genetics approached 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). 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.

References

Sun, Yanping, Wang, Jiankang, Crouch, J.H., and Xu, Yunbi. 2009. Efficiency of selective genotyping for complex traits and its innovative use in genetics and plant breeding. *Molecular Breeding* (in press)

Hao, Z., Liu, X., Li, X., Xie, C., Li, M., Zhang, S., and Xu, Y. 2009. Identification of quantitative trait loci for drought tolerance at seedling stage by screening a large number of introgression lines in maize. *Plant Breeding* doi:10.1111/j.1439-0523.2009.01642.x

Xu, Yunbi., Skinner, D.J., Wu, H., Palacios-Rojas, N., Araus, J.L., Yan, J., Gao, S., Warburton, M.L., and Crouch, J.H. 2009. Advances in maize genomics and their value for enhancing genetic gains from breeding. *International Journal of Plant Genomics* (in press)

Gao, Shibin, Carlos Martinez, Debra J. Skinner, Alan F. Krivanek, Jonathan H. Crouch, and Yunbi Xu. 2008. Development of a seed DNA-based genotyping system for marker-assisted selection in maize. *Molecular Breeding* 22:477-494.

Xu, Yunbi and Jonathan H. Crouch. 2008. Marker-assisted selection in plant breeding: from publications to practice. *Crop Science* 48:391-407.

132. GCP/Rockefeller project G4005.70.01 (CB20a/RF−FS091): Tapping crop biodiversity for the resource poor in East and Central Africa (*Sorghum component*)

July 2005–June 2008

Principal Investigator and Lead Institute

David Hoisington, ICRISAT; d.hoisington@cgiar.org
ICRISAT Patancheru 502324, Andhra Pradesh, India; Tel 91 40 3071 3071;
Fax: 91 40 3071 3074

Collaborating institutes and scientists

- SAARI
- NARI
- Melkassa Research Centre–EIAR
- KARI
- ARI–Naliendele
- ARC–Sudan
- ISAR
- ISABU
- ICRISAT–Kenya: Dan Kiambi
- ICRISAT–India: Tom Hash

The knowledge of the extent and structure of genetic diversity in germplasm accessions through characterisation is essential in the development of strategies for 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 in collaboration with the Generation Challenge Programme (GCP), the Rockefeller Foundation and Biosciences eastern and central Africa (BecA) in order to harness diversity and enhance its use in crop improvement. The project brought together BECA, regional and international initiatives, such as ASARECA and its affiliated networks and the GCP, in a synergistic way that would 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. The project was coordinated by ICRISAT and involved NARS from 8 countries namely Burundi, Eritrea, Ethiopia, Kenya, Rwanda, Tanzania, Uganda and Sudan. It was funded by the Generation Challenge Programme (GCP), Rockefeller (RF) and Biosciences eastern and central Africa (BecA).

The project aimed at:

- Designing a database with passport data, farmer-knowledge, pedigrees, phenotyping and genotyping data of accessions used by NARS
- Developing standardised phenotypic characterisation methodologies
- Characterising NARS germplasm using selected phenotypic descriptors
- Characterising selected accessions genotypically using SSR markers
- Analysing diversity at the national and regional levels and comparing it with the GCP global sorghum diversity data set.

A project inception meeting in which all the participating NARS were represented was held from 15-17 August 2006. Both technical and administrative logistics for the project implementation were discussed and agreed upon. Initially the NARS crop-based programmes compiled a computerised inventory of the germplasm available in their gene banks. From this inventory, subsets of between 164 to 298 accessions were selected to make up a regional composite set eventually used for phenotyping and genotyping. A phenotyping workshop was held in April 2006 and a protocol developed through a participatory approach. The protocol was field-tested and later used to score for 27 qualitative and quantitative characters. In phenotyping, subsets of between 100 and 298 accessions per country were used. Phenotypic data on 1387 sorghum accessions was generated and documented. Each of the eight countries was supplied with a computer and relevant software to enhance data documentation and retrieval. Morphological diversity data is currently being analysed through a PhD Thesis.

A total of 1405 accessions were genotyped using 39 SSR markers that are part of the GCP set of high quality microsatellite markers used for the survey of the global composite set of sorghum germplasm. A total of 54,794 data points were generated. The data has been analysed and genetic diversity between and within countries assessed using a wide range of parameters including the total number of alleles detected, the average number of common alleles per locus, rare alleles (occurring at < or = 5%), number and percentage of polymorphic loci, observed heterozygosity (Hobs) and average expected gene diversity corrected for small simple sizes (Hunb). The country specific molecular data has been analysed and comparison of diversity at the regional and global levels is being finalised through a PhD Thesis.

As part capacity building for the NARS project partners, a phenotyping training workshop was held in April 2006 and was attended by 18 participants. Four project partners received postgraduate training in the use of molecular markers for diversity studies. These include three MSc students from Kenya, Tanzania and Ethiopia and one PhD student from Sudan. In addition, five Visiting Scientists from Rwanda, Burundi, Eritrea, Uganda and Sudan were trained in molecular techniques, genotyping and data analysis at the ILRI-BecA hub. In October 2008, a data analysis workshop was organised in order to consolidate all the data generated by the project and equip the NARS collaborators with the relevant data analysis skills. The workshop was attended by 11 sorghum breeders and they were trained in theoretical aspects of molecular and morphological diversity, data cleaning, preparation, formatting and interpretation using *R*, DARwin, PowerMarker, Structure and GenStat. The molecular data is publicly available through ICRIS, an ICRISAT web-based database. All the project data will be stored and made available in form of access databases, excel spreadsheets, word files and tables, JPEG and BITMAP images. It will finally be posted in the GCP Data Registry.

The project has managed to deliver the expected outputs with the exception of submission of one MSc and one PhD Thesis which are due for completion by the end year 2009. Publications emanating from the project research activities are at different stages of preparation and it is hoped that some of them will be ready by the end of the year.

133. GCP/Rockefeller project G4005.70.02 (CB20b/RF−FS090): Tapping crop biodiversity for the resource−poor in East and Central Africa (*Cassava component*)

October 2005–October 2008; no-cost extension to June 2009

Principal Investigator

Morag Ferguson, IITA; mferguson@cgiar.org

International Institute of Tropical Agriculture (IITA), c/o International Livestock Research Institute (ILRI), P.O. Box 30709, Nairobi, 00100, Kenya. Tel +254 20 4223000; Fax +254 20 4223001

Collaborating institutes and scientists

- INERA–DRC: Mpansu Bidiaka
- KARI: James Gethi
- FOFIFA/DRA: Isabelle Ralimanana
- IIAM: Fernando Chitio
- ISAR: Gashaka Gervais
- ARI–Naliendele: Geoffrey Mkamilo
- NaCRRI: Robert Kawuki

1. Phenotypic activities and progress

The germplasm included in this study represented elite and farmer varieties. Twenty nine qualitative traits were scored on cassava germplasm from: Tanzania (130 genotypes), Uganda (317 genotypes), Kenya (97 genotypes), Rwanda (177 genotypes), DRC (182 genotypes) and Madagascar (188 genotypes). In addition, quantitative traits (root dry matter content, harvest index, leaf retention, and root cortex thickness) were measured on the germplasm in one field season.

Table 1 Mean squares for dry matter content, harvest index and leaf retention of cassava germplasm available with selected national cassava breeding programmes[1]

Source of variation	MS LR	MS DMC	MS HI	MS root cortex
Country (C)	84.20*	1405.9*	2.389*	19.795
Group (G)	6.91*	239.6*	0.823*	6.178
C x G	14.13*	141.5*	0.160*	1.802
Residual	0.51	21.5	0.011	0.224

[1] Countries represent the national cassava breeding programmes; Groups represent the elite and local cassava genotypes; *indicates significance at 5%. MS = mean square. LR = leaf retention in DRC, Kenya, Madagascar and Rwanda. DMC = dry matter content assessed in Kenya, Madagascar, Rwanda, Tanzania, Uganda and DRC. HI = harvest index assessed in Kenya, Madagascar, Tanzania, Uganda and DRC.

2. Genotyping activities and progress

Leaf samples were collected from: Tanzania (279), Uganda (270), Kenya (237), Rwanda (192), DRC (192), Madagascar (188) and Mozambique (82). All cassava genotypes were assayed with 26 highly polymorphic SSR markers. Amplicons were subjected to capillary electrophoresis using the ABI 3730 DNA sequencer (Applied Biosystems) and allele calls made using the GENEMAPPER® software version 3.7 (Applied Biosystems). All genotyping was done at the Biosciences Eastern and Central Africa (BecA) hub in Nairobi, Kenya.

Table 2 Trends in allelic richness, gene diversity and gene flow of cassava germplasm available within the NARS[1]

Country	Allelic richness		Gene diversity		F_{ST}	*Nm*
	Local	Elite	Local	Elite		
Tanzania	89.58	82.52	0.60	0.55	0.017	14.14
Uganda	81.94	87.84	0.53	0.57	0.044	5.36
Kenya	82.24	92.35	0.58	0.61	0.038	6.21
Rwanda	88.36	85.71	0.59	0.56	0.040	5.98
DRC	93.51	97.53	0.61	0.63	0.024	9.84
Madagascar	84.47	86.95	0.58	0.60	0.010	24.41
Mozambique	85.29	85.68	0.55	0.55	0.074	3.10
Mean	86.4	88.3	0.57	0.58		

[1] Allelic richness computed using the rarefaction method as descried by El Mousadik and Petit (1996). F_{ST} and gene flow (*Nm*) estimates compare local and elite cassava genotypes for each national breeding programme. *Nm* estimated from $F_{ST} = 0.25(1 - F_{ST})/ F_{ST}$ as described by Nei (1987).

Tangible outputs delivered

Considerable variation in both root dry matter content (16.3-49.6%) and harvest index. (0.04-0.9). For the genetic analysis, most of the genetic variation was distributed within individuals (89.36%), with marginal variation quantified among groups (5.48%; countries) and among populations within groups (5.15%; local and elite genotypes).

Generation Challenge Programme

CULTIVATING PLANT DIVERSITY FOR THE RESOURCE-POOR

CGIAR Generation Challenge Programme

Hosted by CIMMYT

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

Mailing address:
c/o CIMMYT, Apdo Postal 6-641
06600 México, DF México

Physical address:
c/o CIMMYT, 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 *or* info@generationcp.org
www.generationcp.org