2005
Generation Challenge Program Competitive and Commissioned Research Project Summaries

1 September 2005

Correct Citation: Generation Challenge Program, 2005. *2005 Generation Challenge Program Competitive and Commissioned Research Project Summaries.* Mexico D.F.: Generation Challenge Program.

Table of Contents

COMPETITIVE GRANTS

1. Identifying Genes Responsible for Failure of Grain Formation in Rice and Wheat under Drought

Principal Investigator:
H. Renee Lafitte, IRRI
Co-Principal Investigators:
R. Richards, CSIRO, Australia
A.G. Condon, CSIRO, Australia
L. McIntyre, CSIRO, Australia
J. Bennett, IRRI
K. McNally, IRRI
G. Atlin, IRRI
S. Kikuchi, NIAS
R.C. Babu, Tamil Nadu Agricultural University, India
Zhengqiang Ma, Nanjing Agricultural University, China
Collaborating Scientist:
S. Robin, Tamil Nadu Agricultural University, India

EXECUTIVE SUMMARY
Rice and wheat provide approximately 50% of the calories consumed directly by the human population. The projected increase in this population from 6 billion in 2000 to 9 billion in 2050 requires that production of rice and wheat continue to increase as it has over the last 40 years, following the introduction of high-yielding modern varieties. Future increases will come principally from further intensification of production in the limited irrigated areas and from improved yields in the larger rainfed areas. Drought is the main cause of yield loss in rainfed rice and wheat, and losses are most severe when drought occurs at the flowering stage. Water-saving strategies for irrigated areas must also deal with the sensitivity of the flowering stage to water deficit. For these reasons, we focus here on a comparative study of drought tolerance in rice and wheat, exploiting on the one hand the greater drought tolerance of wheat and on the other hand the recent explosion of information on the rice genome. The rice genome is approximately one-twentieth the size of the wheat genome, but these two cereals are comparatively closely related, with highly similar genes controlling growth, reproduction, and protection. Our team combines expertise on drought-stress physiology, gene expression, genome structure, biodiversity, and plant breeding. Years of research have produced detailed knowledge of which rice and wheat varieties and mutants show contrasting responses to drought during key steps of flowering such as panicle/spike emergence and pollination. Progeny derived by crossing these contrasting lines provide highly informative comparisons that help scientists to interpret the large data sets emerging from modern studies of gene expression using such techniques as microarrays and proteomics, and to identify and validate genes crucial to drought tolerance. Superior forms (alleles) of these genes can be identified in traditional varieties and other sources. Such alleles can then be efficiently transferred into popular rice and wheat varieties via DNA-assisted backcrossing to enhance drought tolerance in both cereals.

2. Revitalising Marginal Lands: Discovery of Genes for Tolerance of Saline and Phosphorus Deficient Soils to Enhance and Sustain Productivity

Principal Investigator:
Abdelbagi M. Ismail, IRRI
Co-Principal Investigators:
Matthias Wissuwa, IRRI
Glenn B. Gregorio, IRRI
David J. Mackill, IRRI
Eduardo Blumwald, University of California, Davis,
Emmanuel Delhaize, CSIRO, Australia
Zeba Seraj, Dhaka University, Bangladesh
Masdiar Bustamam, Indonesian Centre for Agricultural Biotechnology and Genetic Resources and Research Development, ICABGRRD
Collaborating Scientists:
Massahiro Yano, NIAS
Timothy J. Close, University of California, Riverside
Ghasem H. Salekdeh, Agricultural Biotechnology Research Institute of Iran

EXECUTIVE SUMMARY

Soils that contain toxic levels of salts and/or are 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. In Asia alone, more than 12 million ha are currently affected by salinity and about 50% of the rice lands are P-deficient. Salt stress often coexists with other abiotic stresses such as drought and P deficiency. Amendments and management options for these soils are too expensive for the resource-poor farmers commonly living in these areas. However, solutions through improved germplasm are affordable to farmers and are becoming more feasible with the developments in modern molecular tools that are becoming available to unravel the genetic basis of tolerance. Combining mechanisms underlying tolerance of complex traits such as salt and P-deficiency, as well as those for multiple stresses, is now feasible once the genetic components or genes for tolerance are tagged to allow them to be traced in the breeding process. We aim to identify and tag the genes for tolerance of salinity and P-deficiency. For both stresses, we have made excellent progress in understanding the biology and in identifying major chromosomal regions that are associated with tolerance. We will further fine-map these regions and use modern molecular approaches to discover the genes that are involved in tolerance using a range of molecular strategies. We will also use biological information and genes discovered from other crops to facilitate the identification of similar genes in rice. Ultimately, we will develop a marker system to allow the efficient incorporation of these genes into popular, yet intolerant, varieties, initiate a marker-assisted breeding system with NARES partners, and provide them with the training needed to carry out these activities.

3. Identifying the Physiological and Genetic Traits that Make Cassava One of the Most Drought Tolerant Crops

Principal Investigator:
Alfredo Alves, EMBRAPA/CNPMF
Co-Principal Investigators:
Hernán Ceballos, CIAT
Martin Fregene, CIAT
Yvonne Lokko, IITA
Tim Setter, Cornell University

EXECUTIVE SUMMARY

Cassava is usually cultivated in areas considered marginal for other crops, with soils of low fertility and long periods of droughts. Cassava's photosynthesis and growth decrease to near zero during episodes of water deficit, and it achieves most of its growth after rainfall resumes. This suggests that a key to cassava's success is its ability to regulate numerous plant processes to rapidly change course as it navigates between episodes of favourable and unfavourable weather. The general objective of the proposed work is to determine the best traits to be used in breeding programmes for drought tolerance by elucidating the mechanisms of cassava's remarkable tolerance to drought and making full use of the expanding body of information on the physiological and molecular bases of drought tolerance in other well-studied crops. Contrasting genotypes for several traits related to drought tolerance will be selected for evaluation and segregating progenies will be developed for genetic studies. The effect of water deficit on traits which are related to the probable mechanism(s) for drought tolerance in cassava will be evaluated and compared with other well-studied crops. The selected contrasting genotypes will be crossed to generate segregating populations. In addition, drought tolerant genotypes will be selfed to provide S1 families to study recessive gene action. Evaluations will be conducted on the parental clones and the segregating progenies in semi-arid environments of Brazil, Colombia, Ghana, and Tanzania, to screen phenotypes. Segregating progenies will be analysed using a set of genome-wide molecular markers and candidate genes to identify quantitative trait loci (QTL) of component traits of drought tolerance. To assess the value of enhanced leaf retention during stress, a transgenic cassava in which a cytokinin synthesis gene is over-expressed will be field evaluated. Expected outputs of this project include an improved understanding of drought tolerance traits and their biological bases, molecular markers for key drought tolerance traits, and cassava genotypes ready to be introduced into breeding programmes.

4. An Eco-physiological – statistical Framework for the Analysis of GxE and QTLxE as Occurring in Abiotic Stress Trials, with Applications to the CIMMYT Drought Stress Programmes in Tropical Maize and Bread Wheat

Principal Investigator:
Fred van Eeuwijk, WUR
Co-Principal Investigators:
Jean-Marcel Ribaut, CIMMYT
Matthew Reynolds, CIMMYT
Scott Chapman, CSIRO, Australia
Collaborating Scientists:
José Crossa, CIMMYT
Mateo Vargas, Universidad Autónoma Chapingo, Mexico.
Sergio Ceretta, INIA, Uruguay
Marco Bink, WUR

EXECUTIVE SUMMARY

When breeders try to develop adapted genotypes for abiotic stress conditions, i.e., plants with on average superior genetic constitution with respect to yield, they are faced with the problem that it is hard to get reliable estimates of genetic superiority under stress conditions. Under stress, the phenotype, that which the breeder can measure and observe, provides little information on the underlying genetics. A traditional solution uses measurements on yield or

other, secondary traits in non-stress conditions to predict performance under stress. The idea is that under non-stress conditions the genetic value can be estimated more precisely, and as long as the genetic basis of the trait observed under non-stress is closely enough related to the genetic basis of yield under stress or, if the genetic correlation is high enough, then selection under non-stress is preferable. Recently, the traditional approach was challenged by an alternative approach originating from CIMMYT researchers that was built on physiological understanding of the stress response and relevant environmental characterisation of selection and stress environment. The alternative approach would facilitate a better choice of secondary traits and selection environments. Molecular marker techniques make this alternative even more attractive, because of the possibility of selection at the genetic level. However, the new approach still does not live up to the expectations and we think that one of the important reasons for this partial failure is the use of a less-than-adequate statistical framework for analysing data from abiotic stress trials.

The present statistical approaches do not incorporate any explicit physiological knowledge on the part of the genotype nor the environment. We propose the development of an integrated eco-physiological statistical framework, modeling yield responses on both the phenotypic and genetic level in direct dependence on physiologically relevant environmental factors. Application of this framework to existing CIMMYT data on drought stress in maize and wheat will significantly add value in the form of deeper insight in the genetic and physiological mechanisms underlying drought stress in those crops. Additional features of our approach include facilities for the analysis of multiple traits and crosses. To make the methodology generally available to students and researchers in developing countries, course material and corresponding software modules will be developed. This teaching material will be presented in one-week courses in Uruguay and Kenya.

5. Unlocking the Genetic Diversity in Peanut's Wild Relatives with Genomic and Genetic Tools

Principal Investigator:
José Valls, EMBRAPA
Co-Principal Investigators:
David Bertioli, Universidade Católica de Brasília, Brazil
Serge Braconnier, Centre d'Etude Régional pour l'Amélioration de l'Adaptation à la Sécheresse, Senegal
Jonathan Crouch, CIMMYT
Pietro Piffanelli, CIRAD
Guillermo Seijo, IBONE
Jens Stougaard, University of Aarhus, Denmark
Vincent Vadez, ICRISAT

EXECUTIVE SUMMARY

Legumes, unlike other crops, fix nitrogen, need little fertilizer, and help maintain the productivity of the soil. Legume seeds are among the most important sources of protein and iron for the poor. Peanut (A.hypogaea) is a legume grown throughout the tropics on about 24.8 million ha (>90% cultivated by small farmers). Peanut is particularly important in Africa, where production greatly exceeds that of any other legume, and in Asia, where production is almost as high as that of soybean. Peanut is sensitive to fungal diseases and drought stress and these factors are important reducers of yield. Improvement of peanut has been limited by an extreme genetic bottleneck at its origin, which occurred via hybridisation

of two wild species followed by a rare spontaneous duplication of chromosomes. The resultant plant had hybrid vigor, but because of the difference in chromosome number, was reproductively isolated from its wild relatives. Therefore, all peanuts are probably derived from one or a few plants. This led to low diversity for important agricultural traits and very limited genetic diversity, which has constrained advances in genetics necessary for modern breeding. In contrast, wild Arachis species are very diverse and have been selected during evolution by a range of environments and diseases, providing a rich source of variation in agronomically important traits.

Recently, partners in this proposal have artificially recreated the events that gave rise to peanut, using a wide range of diploid species. So far, four viable synthetic hybrids have been created, thus bringing for the first time to peanut breeding the genetic diversity of the genomes of eight wild Arachis species. In parallel, major breakthroughs in genetic mapping have been made using a new strategy that will allow plant breeders to work complex hybrids more efficiently. This proposal aims to build on these advances to enable the creation of peanut varieties resistant to disease and drought. In addition, we propose to include peanut in a single genetic system for legumes, allowing peanut research to benefit from the knowledge of modern "genomics."

6. Marker Development and Marker-assisted Selection for Striga Resistance in Cowpea

Principal Investigator:
Festo Massawe, IITA
Co-Principal Investigators:
M.P.Timko, University of Virginia
B.B. Singh, IITA
V. Mahalakshmi, IITA
Ndiaga Cissé, CERAAS, Senegal
N'deye Ndack Diop, CERAAS, Senegal

EXECUTIVE SUMMARY
Cowpea is an important food grain legume grown on 9.8 million hectares of small farms in the dry savannah of tropical Africa. Current estimates place world cowpea production at 3 million tons, with 80% of its production in Africa, principally West and Central Africa where the crop productivity is low due to pests and diseases. The parasitic angiosperm Striga gesnerioides (Willd.) is one of the major limitations to cowpea productivity. Conventional breeding efforts have helped to alleviate some of the Striga problems, but pyramiding resistance to the parasite with other important agronomic and resistance traits is time-consuming and difficult. Modern technologies, such as marker-assisted selection (MAS), in combination with conventional breeding, have been successfully used for genetic enhancement of other crop species. The cooperative work proposed here, involving the International Institute of Tropical Agriculture (IITA), the Centre d'Etude Regional pour l'amelioration de l'Adaptation a la Seccheresse (CERAAS), the Institut d'Environnement et de Recherches Agricoles (INERA) of Burkina Faso, and the University of Virginia (UVA), seeks to develop a MAS strategy for cowpea that will allow the rapid, reliable identification of race-specific Striga resistance genes in breeding lines and integration of MAS for Striga resistance in their breeding programmes. The outcome of this work will be superior-performing, well-adapted cowpea varieties containing pyramided agronomic productivity, and disease and pest resistance traits available to farmers. This project will also contribute to

the development of human and institutional capacity to fully integrate the use of MAS technologies in cowpea breeding. It is expected that farmers will achieve higher yields of better quality cowpea that would impact favourably on their general livelihoods.

7. Measuring Linkage Disequilibrium across Three Genomic Regions in Rice

Principal Investigator:
Susan McCouch, Cornell University
Co-Principal Investigators:
Michael Thomson, ICABIOGRAD, Indonesia
Endang Septiningsih, ICABIOGRAD, Indonesia

EXECUTIVE SUMMARY
Rice is an important staple crop worldwide. Rice also has many advantages for genetic research, most notably the complete genome sequence and a wealth of genetic diversity. Rice therefore presents an excellent opportunity to use linkage disequilibrium (LD) mapping and association studies for allele mining to identify superior alleles that lay hidden in the vast reservoirs of global rice germplasm. Although LD mapping presents a powerful technique, little is known about the actual amount of LD across different genomic regions in rice. The main objective of this start-up grant is to measure the extent of linkage disequilibrium in three genomic regions to test the feasibility of LD mapping as the basis for allele mining in rice, through collaborative research between Cornell University and the Indonesian Centre for Agricultural Biotechnology and Genetic Resources Research and Development (ICABIOGRAD) in Bogor, Indonesia. Currently, 250 Indonesian landraces are being analysed with a set of genome-wide simple sequence repeat (SSR) markers to define the population structure across these accessions. This Generation start-up project will use the same accessions to measure the LD in three genomic regions surrounding the bacterial blight resistance genes Xa7, Xa13, and the cluster of Xa4/Xa22/Xa26. Single nucleotide polymorphism (SNP) markers will be developed across each region by sequencing PCR products from a small subset of landraces. These SNP markers will then be genotyped across a larger subset using a high-throughput SNP assay. The SNP data will be used to measure the LD across each region and define the haplotype block structure that exists in this set of rice germplasm. This project will provide a specific measure of LD across three genomic regions that will be useful in planning the strategy for larger, more complex allele mining experiments using the global collection of rice germplasm.

8. Targeted Discovery of Superior Disease QTL Alleles in the Maize and Rice Genomes

Principal Investigator:
Rebecca Nelson, Cornell University
Co-Principal Investigators:
Casiana Vera Cruz, IRRI
Darshan Brar, IRRI
Hei Leung, IRRI
Margaret Smith, Cornell University
Peter Balint-Kurti, USDA
Jan Leach, Colorado State University
Jane Ininda, KARI. Kenya

Jedidah Danson, KARI, Kenya
James Gethi, KARI, Kenya
Masdiar Bustamam, ICABGRRD, Indonesia
Utut Widiyastuti Suharsono, Bogor Agriculture University, Indonesia

EXECUTIVE SUMMARY

We propose to identify, characterise, and utilise sections of the rice and maize genomes that provide superior disease resistance to cereal diseases of critical and global importance. Durable, broad-spectrum resistance would be valuable to resource-poor farmers. Although much research has been focused on qualitative (complete, race-specific) resistance, the work proposed here will focus on quantitative (incomplete, presumably race non-specific) disease resistance (QDR) because QDR is usually the more durable form or the only form available. At present, the chromosomal regions associated with QDR are defined with very low precision, and germplasm has not been systematically analysed to identify superior alleles at the loci of greatest potential utility. We propose to characterise selected maize and rice germplasm for urgently needed disease resistance. We will initiate development of near-isogenic lines (NILs) capturing useful segments of maize and rice chromosomes in a susceptible background for detailed analysis. We will use a set of complementary strategies in the development of the NILs, including backcrossing of advanced resistant lines derived from rice varieties known for durable resistance; selection of allelic series at loci of outstanding interest based on a summary of all available disease QTL studies in maize; and selection of lines carrying alleles showing increases in frequency under recurrent selection for a maize disease. We will make use of the existing collection of rice mutants to validate the function of candidate QDR genes. The superior chromosomal segments identified in this project will be analysed in detail and utilised in the applied breeding programmes in which improving disease resistance is a high priority.

9. Development of Low-Cost Technologies for Pyramiding Useful Genes from Wild Relatives of Cassava into Elite Progenitors

Principal Investigator:
Anthony Bellotti, CIAT
Co-Principal Investigators:
Martin Fregene, CIAT
Alfredo Alves, EMBRAPA-CNPMF
Collaborating Scientists:
Hernan Ceballos, CIAT
Elizabeth Alvarez, CIAT
Elizabeth Okay, CRI, Ghana
Chiedozie Egesi, NRCRI, Nigeria
Anton Bua, NAARI, Uganda
Titus Alicai, NAARI, Uganda
Yona Baguma, NAARI, Uganda

EXECUTIVE SUMMARY

Cassava (Manihot esculenta Crantz) is increasing in importance in the tropics due to its hardy nature but it suffers from a plethora of anthropod pests and diseases as well as post harvest physiological deterioration (PPD). It has been estimated that cassava farmers, typically resource-poor farmers, lose 48 million tons of fresh root to pests, diseases, and PPD every year. This makes up some 30% of total world production, valued at US$1.4billion. Wild

relatives of cassava are important sources of genes for resistance to pests and diseases and longer shelf life. Dramatically delayed PPD has been identified in inter-specific hybrids from Manihot walkerae. The only source of resistance to the cassava hornworm and a widely deployed source of resistance to the cassava mosaic disease (CMD) were identified in 4th backcross derivatives of M. glaziovii. Moderate to high levels of resistance to cassava green mites (CGM), white flies, and the cassava mealy bug have been found in inter-specific hybrids of M. esculenta sub spp flabellifolia. Furthermore, M.glaziovii, M. catingae, and M. carthaginensis, are adapted to semiarid lands and are potential sources of genes for tolerance to drought. But the heterozygous nature and long reproductive cycle of cassava makes introgression and pyramiding of these genes a long-term effort. For several years molecular marker tools and a modified Advanced Back Cross QTL (ABC-QTL) scheme have been tested for cost-effective pyramiding of useful genes from cultivated and wild gene pool through the elimination of phenotypic evaluations in each breeding cycle. This proposal seeks to make marker-assisted introgression of exotic genes into elite cassava progenitors widely available through the development of low cost approaches, expand the gene tagging effort to other traits, and establish a systematic approach of collection, evaluation, and use of additional wild germplasm.

10. Exploring Natural Genetic Variation: Developing genomic resources and introgression lines for four AA genome rice relatives

Principal Investigators:
Joe Tohme, CIAT
Mathias Lorieux, CIAT/IRD
Co-Principal Investigators:
Susan R. McCouch, Cornell University
Claudio Brondani, CNPAF-EMBRAPA
Howard Gridley, WARDA
César P. Martinez, CIAT
Miguel Diago Ramirez, Fedearroz, Colombia

EXECUTIVE SUMMARY

Cereals provide the majority of calories consumed by humans. Cereal production faces growing challenges due to increasing human population, changing nutritional requirements, and variable environmental conditions that require new approaches to crop production. Wild relatives of modern crop species have survived for millions of years using natural genetic defenses to endure biotic and abiotic aggressions. These wild relatives represent a valuable source of under-utilised genetic variation that is available to plant breeders and 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, we propose to (1) develop four libraries of interspecific lines called Chromosome Segment Substitution Lines (CSSLs) targeting chromosomal introgressions from different rice relatives, (2) develop a set of 140 molecular markers (called SNPs) identified in genes associated with tolerance to abiotic stress (drought, acid soils, mineral deficiencies, or toxicities), (3) validate the utility of the SNPs by using them in the development of the CSSLs in this project and exploring their value in breeding programmes for other cereals, (4) analyse a set of advanced CSSLs generated from Asian x African rice crosses for their phenotypic response to drought stress. Generating such resources and knowledge will contribute to the objectives of Subprogrammes 1 and 3 by (i) utilising natural

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

11. Functional Genomics of Cross-species Resistance to Fungal Diseases in Rice and Wheat (CEREALIMMUNITY)

Principal Investigator:
Pietro Piffanelli, AGROPOLIS
Co-Principal Investigators:
J-L Notteghem, AGROPOLIS
M. E. Ferreira, EMBRAPA
R. Singh, CIMMYT
P. Ronald, University of California-Davis
M. William, CIMMYT
J-B Morel, AGROPOLIS
D. Tharreau, AGROPOLIS
S. Kikuchi, NIAS
F. Dedryver, University of Rennes, France
L. Boyd, JIC
S. Brammer, EMBRAPA
AS. Prabhu, EMBRAPA
M.C. Chaves, EMBRAPA
E. Guiderdoni, AGROPOLIS

EXECUTIVE SUMMARY
Resistance shown by a plant species to the majority of potentially pathogenic microbes is known as non-host resistance. The events leading to non-host resistance in plants represents one of the least understood phenomena and a remaining challenge in the field of plant-microbe interactions. Comparative genomics is a promising method to identify key genes involved in cross-species interactions and to better understand their regulation at the genetic level and their evolution. Non-host resistance also represents one promising defence mechanism in developing durable resistance against plant pathogens, namely due to its effectiveness against a broad range of pathogen species and its durability in nature. The proposed project will strengthen and extend ongoing research in rice and wheat and aims to define the signalling and effector genetic components involved in non-host resistance in cereals to devise novel defence strategies with the potential to yield durable resistance against host pathogens in cereals. This project aims to implement existing breeding programmes for resistance to blast and rust diseases in developing countries, taking advantage of advanced genomic platforms and technologies.

12. Drought Tolerant Rice Cultivars for North China and South/Southeast Asia by Highly Efficient Pyramiding of QTLs from Diverse Origins

Principal Investigator:
Zhi-Kang Li, CAAS

Co-Principal Investigators:
Gary Atlin, IRRI
Jian-Min Wan, CAAS

EXECUTIVE SUMMARY
Rice is the staple food for most Asian and Chinese people, but rice production uses large amounts of water. Drought has become the single largest factor limiting rice production in Northern China and the rainfed areas of South/Southeast Asia. Developing drought tolerant (DT) rice cultivars is the most efficient way to stabilise rice production and alleviate food insecurity and poverty in China and Asia. In this proposed project, we aim to develop high yielding and DT rice cultivars for the Northeast/Northwest China and the rainfed areas of South/Southeast Asia by exploiting the rich genetic diversity in the primary gene pool of rice in a large backcross breeding programme integrated with efficient selection and DNA markers. Using molecular markers, linkage disequilibrium mapping, and two large sets of introgression lines (ILs) in elite Chinese japonica backgrounds having introgressed DT from 67 diverse germplasm accessions and breeding populations derived from 7 well characterised DT IR64 lines, our goal is to discover and characterise important DT QTLs in the process of breeding for high yielding and DT cultivars for the target environments. The expected outcomes from the project will include four major aspects: (1) important DT QTLs and multiple alleles at many QTLs identified, confirmed and characterised in the elite rice backgrounds; (2) development of superior high yielding and DT rice cultivars for the Northeast/Northwest China and the rainfed areas of South/Southeast Asia; (3) knowledge, theory, and strategy generated for genetic improvement of complex phenotypes; and (4) training of 10 young scientists from China and South/Southeast Asia in molecular breeding. More importantly, information and knowledge generated from the proposed project will allow CAAS to establish modern breeding systems that fully integrate the molecular tools with the current breeding programmes for genetic improvement of major crops in China.

13. Development of Informative DNA Markers through Association Mapping in Maize to Improve Drought Tolerance in Cereals

Principal Investigator:
Jean-Marcel Ribaut, CIMMYT
Co-Principal Investigators:
Edward Buckler, Cornell University
Alain Charcosset, INRA
James Gethi, KARI
Grudloyma Pichet, NSFCRC, Thailand
Luke Mehlo, SIRDC
Mark Sawkins, CIMMYT
Tim Setter, Cornell University
Wanchen Li, Sichuan Agriculture University
Collaborating Scientists:
Marianne Bänziger, CIMMYT
Javier Betran, Texas A&M
Jose Crossa, CIMMYT
Luz George, CIMMYT
Philippe Monneveux, CIMMYT

EXECUTIVE SUMMARY

Drought and low soil fertility are the major limiting factors for cereal-crop production in developing countries. The objective of this project is to use the natural variation inherent in the maize genome for the dissection of drought tolerance and for the identification of superior alleles. While maize grows in a wide range of environments and is the most diverse crop in the world, we do not know the genes that are responsible for these adaptations. For phenotypic selection, crops need to be fully evaluated in every environment, which is costly and time consuming, although it does allow genetic progress. Association studies, proposed in this project, are based on correlation between a gene sequence and plant performance for target traits, and represent a powerful approach to evaluate candidate genes regulating plant phenotype. This project will focus on evaluating the genes in two major pathways that are involved in drought tolerance. We will build upon previous mapping approaches that have identified genomic regions containing a few hundred genes, and use high resolution approaches that can evaluate individual genes. This high resolution mapping will require combining rapid molecular approaches with careful evaluation of diverse germplasm for drought tolerance and physiological response. Additionally, by screening several hundred diverse lines, this project maximizes its potential to identify the best alleles in the maize gene pool. The discovery of superior alleles at the gene level will permit the development of molecular markers that can facilitate breeding drought tolerance in a wide range of germplasm. One important benefit of working with the natural variation, it is that any discovery can be rapidly converted to improved breeding materials without the societal and regulatory obstacles of transgenics materials. Because of the genetic and physiological commonalities among cereal crops, this knowledge gathered in maize can be applied to all other cereal crops.

14. Characterisation of Genetic Diversity of Maize Populations: Documenting global maize migration from the centre of origin

Principal Investigator:
Marilyn Warburton, CIMMYT
Collaborating Scientists:
Luz George, CIMMYT
S. Taba, CIMMYT
V. Mahalakshmi, IITA
Abebe Mentir, IITA
Alain Charcosset, INRA
Zachary Muthamia, KARI
S.H. Zhang, CAAS
B. M. Prasanna, Indian Agriculture Research Institute
Sutrisno, Indonesian Department of Agriculture
Pichet Grudloyma, Nakhon Sawon Field Crops Research Centre
Phan Xuan Hao, National Maize Research Insitute,Vietnam
Artemio Salazar, Philippine Department of Agriculture

EXECUTIVE SUMMARY

Although maize hybrids represent the most economically important portion of the species, maize breeding populations, open pollinated varieties (OPVs), landraces, and wild relatives contain the majority of the diversity found in maize, much of which has never been incorporated into improved varieties. Populations introduced into other countries, originally from the centre of origin in Central America but following a complicated pattern of

introductions, have become adapted to many new growing conditions and local stresses, including drought. Past studies of maize population diversity have revealed useful clues as to relationships and patterns of diversity; however, a complete, global picture of maize diversity is lacking because analysis of heterogenous populations has been until recently very expensive and time consuming. Phenotypic characterisation of cultivated maize and wild teosinte populations for traits important to breeders and farmers has been done only in a very limited manner, and at the molecular level, essentially not at all.

Drought tolerance is a trait of extreme importance to farmers who have access to limited resources, but is difficult to phenotype (especially in wild species) and sufficient diversity is lacking in current breeding germplasm, so a great need for new diversity exists. This study aims to complete the global picture of maize diversity and spread by collecting and analysing maize populations from geographic regions that have been underrepresented in previous studies, and representatives of the wild ancestor of maize (teosinte). Structural characterisation will occur at the molecular level using SSR markers. The populations containing the most unique alleles at the SSR loci will then be characterised for markers associated with drought tolerance, as these are the populations most likely to contain new alleles in general and potentially for drought related loci. The genetic characterisation data will provide useful information for utilising these populations in genomic studies and breeding efforts to create drought tolerant maize.

15. Determination of a Common Genetic Basis for Tissue Growth Rate Under Water-limited Conditions across Plant Organs and Genomes

Principal Investigator:
Mark Sawkins, CIMMYT
Co-Principal Investigators:
Jean-Marcel Ribaut, CIMMYT
Francois Tardieu, INRA
Peter Stamp, ETH, Switzerland
Matthew Reynolds, CIMMYT
Peter Langridge, ACPFG, Australia
Renee Lafitte, IRRI
Ravindra Kumar, IGAU, India
Luke Mehlo, SIRDC, Zimbabwe
Collaborating Scientists:
John Bennett, IRRI
Marianne Bänziger, CIMMYT
Claude Welcker, INRA
Yvan Fracheboud, ETH, Switzerland
Richard Trethowan, CIMMYT

EXECUTIVE SUMMARY

The effort to minimise the impact of drought on yield needs new approaches for bridging traditional breeding to molecular genetics. Recent advances in comparative genomics allow information to be moved from one genome into another for identifying key genes controlling drought tolerance. However, comparison between species remains difficult because compared processes, organs, and conditions differ between species in most published studies. We will undertake a multiple-species, multiple-organ study on a key process: growth maintenance under water deficit. The project combines new approaches of phenotyping

(controlled conditions and field), modeling, quantitative genetics, comparative genomics, and first steps towards association genetics. It also combines the strengths of research in "advanced" countries, CGIAR centres, and developing countries. It is applied to three cereals (wheat, maize, and rice) for growth maintenance of leaves and to three organs (leaves, roots, and reproductive organs) in maize. The project will adopt the approach of characterising environmental conditions in all experiments (including those for genomics), and analysing germplasm under controlled environment and field conditions using a modeling approach. Common genomic regions and genes important for growth will be identified through existing and new QTL data across the three cereals. Comparison of gene expression in common tissue across and within species will be used to identify candidates for detailed analysis. Questions to be addressed will include: How do identified genes contribute to growth maintenance in different climates over the world and how does that correlate with yields? And, what combinations of alleles optimise the growth of key tissues in droughted rice, wheat, and maize under different environments? A comparative study of the three species will generate results that feed into modeling work, thereby interpreting and using (for breeding) the genotype x environment interaction of key traits involved in drought tolerance such as early vigor, high light interception, or maintenance of reproductive development.

16. Isolation and Characterization of Aluminum Tolerance Genes in the Cereals: An Integrated Functional Genomic, Molecular Genetic and Physiological Analysis

Principal Investigator:
Leon Kochian, Cornell University
Co-Principal Investigators:
Ed Buckler, Cornell University
Owen Hoekenga, Cornell University
Jurandir Magalhaes, EMBRAPA
Claudia Guimarães, EMBRAPA
Vera Alves, EMBRAPA
Newton Carneiro, EMBRAPA
Robert Schaffert, EMBRAPA
Sandra Brammer, EMBRAPA
Pericles Neves, EMBRAPA
Rosangela Bevitori, EMBRAPA
Samuel Gudu, Moi University, Kenya

EXECUTIVE SUMMARY
One of the most important soil-related factors limiting agriculture in developing countries is acid soil pH (pH < 5). Acid soils occur for both natural and humanity-derived reasons. On acid soils, regardless of their source, toxic levels of aluminum (Al) ions are released into soil solution, where they damage roots and impair their growth and function. This damage results in reduced nutrient and water uptake, with concomitant reductions in crop yield. There is considerable natural variation in Al tolerance both within and between plant species, and we have assembled an interdisciplinary group of scientists to take advantage of this variation to improve crop tolerance to Al toxicity on acid soils. This proposal details an interdisciplinary project that will characterise recently isolated cereal Al tolerance genes as well as identify novel Al tolerance genes and physiological mechanisms in a range of cereal species (sorghum, maize, rice, and the Triticeae). The research group we have assembled has considerable expertise in the genetics, molecular biology, and physiology of aluminum

tolerance in these crops, and has available the necessary genetic resources to ensure the success of this project. We will use information from candidate genes identified in wheat and sorghum, as well as ongoing progress from our genetic mapping and cloning programme in maize, to identify and verify candidate Al tolerance genes in several cereals species. The long-term goals of this research are to generate cereal genotypes expressing improved Al tolerance that ultimately can be distributed to farmers who till acid soils in Africa and other developing regions, thus exploiting a wide range of still hidden genetic variation for Al tolerance. Increasing the Al tolerance of staple crops, such as maize and sorghum, will help increase yields and, thus, food security.

17. Allele Mining Based on Non-Coding Regulatory SNPs in Barley Germplasm

Principal Investigator:
Michael Baum, ICARDA
Collaborating Scientists:
W. Powell, University of Adelaide, Australia
P. Langridge, Australian Centre for Plant Functional Genomics Pty Ltd
Mark Tester, Australian Centre for Plant Functional Genomics Pty Ltd
J. K. Eglinton, University of Adelaide, Australia
M. Morgante, Universita' di Udine Via delle Scienze, Italy
Salvatore Ceccarelli
Stefania Grando, ICARDA
Sripada Udupa, ICARDA
Wafaa Choumane, Tishreen University, Syria.

EXECUTIVE SUMMARY
In recent years analysis of genetic variation has focused on the study of changes in DNA coding for proteins. It is now becoming increasingly clear that this only accounts for one aspect of heritable variation and for many traits, notably tolerance to environment stresses, the level of gene expression is also likely to be of great importance. If changes in gene expression underlie many evolutionary changes in phenotype, then identifying the genetic variants that regulate gene expression is a significant and important endeavour. One of the key problems in genetics is how to identify this type of variation. We propose a robust, quantitative approach to efficiently identify plant genes that harbor such regulatory variants. The approach is novel and particularly amenable to plants since it is based on monitoring gene expression in experimentally created hybrids. A successful outcome will provide a new mechanism to connect genotype to phenotype based on changes in gene expression rather than changes in the structure of an encoded protein. This approach will be used to characterise a series of genes identified and to reveal potential candidates for tolerance to drought, frost, cold, and salinity stresses. The approach is generic and widely applicable. The project will also involve training researchers in Developing Countries and create a high quality collaborative network of researchers delivering new knowledge on genetic diversity and translatable outputs for the Developing World.

SP1 COMMISSIONED GRANTS

1a. Completing Genotyping of Composite Germplasm Set of Barley

Principal Investigator:
Michael Baum, ICARDA
Collaborating Scientists:
J. Valkoun, ICARDA
S. Grando, ICARDA
Zhang Jing, CAAS
Joanne Russell, SCRI, Scotland

EXECUTIVE SUMMARY
The application of association genetics through linkage disequilibrium studies of germplasm adapted to key regions represents an innovative solution to the problem of identifying markers for genomic regions key to specific characteristics (Risch and Merikangas, 1996; Nordborg and Tavaré, 2002). Linkage Disequilibrium (LD) mapping is a relatively new concept in plant breeding, developed in human genetics, to facilitate trait-marker allele association (Risch and Merikangas, 1996; Cardon and Bell, 2001; Boehnke, 2000; Clayton 2000; Nordborg and Tavaré, 2002). The power and discrimination of association studies is largely determined by the underlying patterns of LD in the chosen population and studies in humans have shown the importance of past history of the population, bottlenecks, selection, admixture, and population structure. Other important factors for association studies are aspects of the phenotype, such as heritability and accuracy of measurement. Association genetic studies in plants have been limited to the single gene context; for example, the association of plant height in maize with a known candidate gene bronze (Remington et al., 2001). The application of LD outside the single gene context meets with scepticism. However, such scepticism has mainly an empirical basis, as it is based on little data and is probably due to the use of non-optimal statistical methods. Recently, Kraakman et al., (2004) found significant association of genomic regions with yield and yield stability using a limited number of anonymous markers in barley. The efficient implementation of MAS requires association and QTL detection methods that can be applied in multiple related crosses using populations that are directed more toward the situation in a practical breeding programme.

The project takes advantages of recent developments in barley genomic research and modern marker technology to rapidly generate genotypic data that can be related to the wealth of phenotypic data. The application of a low-resolution screen to a range of genotypes, followed by more detailed study of candidate genes in the region, will enable a selective filtration of the vast accumulation of genomic information to target factors that will be key to adaptation to the specific environments chosen for this project. The resources available in barley are now considerable and include a large number of Expressed Sequence Tags (ESTs), from which gene based functional markers have been developed at the Scottish Crop Research Institute (SCRI).

The importance of the use of gene based markers is that they have been generated from ESTs that have been used to design the Barley Affymetrix chip array, which means that there is a direct relationship between the variation detected using these markers and on-going studies showing the importance of particular genes in the control of biologically and agronomically important traits throughout the world. This change to functional markers more relevant to target traits is therefore expected to result in information that can then be utilised in a robust

MAS scheme for barley improvement in such areas, as well as serving as a demonstration project for the selective improvement of other crops. Certainly, deployment of a strategy implementing MAS to form a pool of elite germplasm fixed for key target QTLs, followed by phenotypic selection to identify the best of these, would improve the efficiency of barley breeding. This would be because expensive yield trialing can be concentrated upon lines that are most likely to meet desired standards and the overall numbers of lines in advanced trials would consequently be reduced without reduction in genetic gain.

1b. Completing Genotyping of Composite Germplasm Set of Wheat

Principal Investigator:
Marilyn Warburton, CIMMYT
Participating Institutions:
CIMMYT
ICARDA

EXECUTIVE SUMMARY
The objective of SP1C2 is the molecular marker analysis of the accessions belonging to the core subsets of up to 3,000 accessions per crop so that the structural genetic diversity within the subset may be known. The selection of priority crops to include in SP1 for Year 1 was done at the Planning Workshop held in Wageningen and included: barley, maize, rice, sorghum, wheat, chickpea, cowpea, common bean, cassava, potato and musa. The criteria to select them included the relevance of the crop for food security worldwide, the degree of genetic/genomic tools already available for the crops, and their potential as good sources of genes for drought tolerance. The definition of the core subsets for each crop will be done in SP1 Cluster 1.

In year one, for wheat, 2,600 genotypes were identified based on earlier phenotyping data analysed with 50 markers. Because the final target had been 3,000 genotypes, this proposal seeks to finalise phenotyping and to fingerprint the remaining 400 drought tolerant lines or populations in order to reach the target number of 3,000 based on proper phenotyping data.

In the first year of the Challenge Programme, due to communications difficulties, only 200 entries from ICARDA were included in the composite genotype set. However, because ICARDA holds an extensive collection of germplasm, especially genotypes that may be adapted to dry growing conditions, ICARDA will again participate in the second phase of fingerprinting by choosing 200 additional entries (thus doubling what is included from ICARDA holdings) and will participate in DNA extraction of the 200 entries. CIMMYT will choose the additional 200 entries, paying particular attention to landraces, related species, and synthetic wheats, as the majority of the diversity in wheat will likely be found in these sources. Georeferencing of the materials, which is being done in a CIMMYT GIS database, but in particular ongoing drought trials on 3,100 landraces and synthetics, will allow the entries to be chosen with particular emphasis on drought. Which entries are included in the composite genotype set is of extreme importance, as we will only find the genes contained in this set when we move towards a more intensive study of the genotypes chosen from this set. If we do not include the right material, we will not find the best allelic diversity for drought resistance.

1c. Completing Genotyping of Composite Germplasm Set of Sorghum

Principal Investigator:
C T Hash, ICRISAT
Collaborating Scientists:
Claire Billot, Agropolis
Monica Deu, Agropolis
Jean-François Rami, Agropolis
Jacques Chantereau, Agropolis
P Ramu, ICRISAT
HD Upadhyaya, ICRISAT
RT Folkertsma, ICRISAT
Yu Li, CAAS

EXECUTIVE SUMMARY
Sorghum is the fifth most important cereal globally following maize, rice, wheat, and barley. It provides staple food grain and stover (used for fodder, fuel, and construction material) in semi-arid and sub-humid tropical, sub-tropical and temperate environments too harsh for rainfed production of maize or rice. Sorghum is well known as a drought-tolerant grain and fodder crop, is closely related to maize, and has a much smaller genome than maize or pearl millet. Substantial genetic variability is available in global collections of wild and cultivated sorghum germplasm, but only a very limited fraction of the more than 40,000 collected accessions have been exploited in applied sorghum breeding programmes. These activities are intended to provide a better understanding of the structure of global sorghum genetic resources so that these can be more effectively harnessed for crop improvement—not only of sorghum in the short term, but potentially for other crops as well in a longer-term horizon.

The activities proposed for 2005 build upon GENERATION CP activities initiated in 2004, including constitution of a composite germplasm set of 3,000 wild and cultivated (including landraces, elite cultivars and hybrid parents, mapping population parents, trait donors and recurrent parents for current and future marker-assisted breeding programmes) sorghum accessions, assessment of a set of 104 sorghum SSR primer pairs for their ability to consistently detect variation (across Agropolis, CAAS, and ICRISAT labs) among a mini-core of 48 sorghum accessions, and selection of an initial 30 of these primer pairs for genotyping approximately 700 accessions that are included in the sorghum composite germplasm set.

1d. Completing Genotyping of Composite Set of Chickpea

Principal Investigator:
HD Upadhyaya, ICRISAT
Collaborating Scientists:
PM Gaur, ICRISAT
CLL Gowda, ICRISAT
S Chandra, ICRISAT
SM Udupa, ICARDA
BJ Furman, ICARDA
M Baum, ICARDA

EXECUTIVE SUMMARY
The application of anonymous markers to germplasm of a given crop has a very high added value. This allows structuring the collections, targeting new perspectives, rationalising

breeding strategies, etc. It also allows determination of representative collections for phenotyping efforts and association studies using candidate genes orthologous to those that are proven on the most advanced crops.

The main objective of this commissioned grant is the marker analysis of a subset of 2,714 chickpea accessions. This will result in defining the genetic structure of the global chickpea composite germplasm collection for functional and comparative genomics studies in the Challenge Programme.

The analysis of genetic diversity will help to elucidate population structures that influence the analysis of the associations between markers and phenotypes, which will be performed later in the GCP. Based on the genotypic data set generated, a subset of about 300 accessions will be identified for detailed study of various traits of economic importance including functional genomics analysis. Subsequently, additional SSR and orthologous markers will be screened across this subset. Genetically broad-based mapping and breeding populations could be developed using information generated in this project.

2. Supporting Distribution of Reference Germplasm

Principal Investigator:
Germplasm bank curators
Participating Institutions:
Any CG centre mandated for germplasm conservation for a particular crop studied in year one

EXECUTIVE SUMMARY
The genotyping activities started in year one end up with the definition of various samples, including a microcore collection of 48 accessions best adapted for allele discovery, and a reference sample of about 500 accessions which will be promoted for use in phenotyping experiments and further molecular characterisation with potential functional markers. Managing, securing, and supplying the corresponding accessions, some of which require controlled selfing, may require a specific effort in the CG germplasm centres in charge. It is proposed that the GCP contributes to this initial effort.

3. Molecular Characterization of Tier 2 (Orphan) Crops

3a. Development of Composite Collection of Tier 2 (Orphan) Crops—Finger Millet

Principal Investigator:
HD Upadhyaya, ICRISAT
Collaborating Scientists:
C.T. Hash, ICRISAT
Subhash Chandra, ICRISAT

EXECUTIVE SUMMARY
The Rajendra S. Prasad gene bank at ICRISAT holds 5,940 accessions of finger millet from 24 countries that were characterised for various morphological and agronomic traits. However, the use of germplasm in finger millet improvement programmes is limited. An important goal of the Generation Challenge Programme is

extensive genetic characterisation using molecular markers of vast genetic resources (including wild relatives, landraces, breeding materials, cultivars, and genetic stocks) held by ICRISAT. Since the entire collection cannot be used for molecular characterisation, it is important to develop a composite collection representing finger millet germplasm using available data on geographical and morphological descriptors as well as characterisation and evaluation data. The composite set then can be used for marker analysis in 2006 to elucidate the genetic diversity within the subset which will help in defining the population structure that influences the analysis of association between markers and traits' phenotype.

3b. Molecular Characterisation of Tier 2 (Orphan) Crops--Pigeonpea

Principal Investigator:
HD Upadhyaya, ICRISAT
Collaborating Scientists:
Subhash Chandra, ICRISAT
KB Saxena, ICRISAT

EXECUTIVE SUMMARY
An important goal of the Generation Challenge Programme is extensive genetic characterisation using molecular markers of vast genetic resources (including wild relatives, landraces, breeding materials, cultivars, and genetic stocks) held by ICRISAT.

Development of a composite subset of 1,000 accessions of cultivated and wild pigeonpea using available phenotypic characterisation and evaluation data will be useful in deciding a representative subset of this collection. Marker analysis of this subset will be done to elucidate the genetic diversity within the subset which will help in defining the population structure that influences the analysis of association between markers and trait's phenotype. Using the data on genetic diversity, the composite collection could be further sampled to a manageable number, about 300, which will be used in further studies in the Challenge Programme.

3c. Application of Molecular Markers for Gene Pool Division and Heterosis Estimation under Drought Stress Conditions in Sweet Potato

Principal Investigator:
Marc Ghislan, CIP
Collaborating Scientists:
Wolfgang Grüneberg, CIP
Jorge Benavides, CIP
Robert Mwanga, NAARI, Uganda

EXECUTIVE SUMMARY
The spontaneous polyploidization of species by combining their genomes (auto and allopolyploidy) has played a prominent role in plant evolution. Most allopolyploids species are propagated by cloning – with the option of sexual reproduction – and have a large genetic diversity (i.e. Ipomoea batatas, Solanum tuberosum, Manihot

esculenta, Dioscorea ssp., Colocasia esculenta). A main reason for the success of polyploids is the favorable interactions between genes on their homologous chromosomes (Heterosis), which can be fixed in vegetative propagated species. Moreover, heterosis is assumed to be significantly larger under stress conditions such as drought or nutrient deficiency compared to under non-stress conditions.

From CIP's germplasm collection and CIP's actual breeding material 500 clones will be investigated for their genetic diversity by 20 validated SSR markers. Additionally, a group of 25 new SSR markers will be tested for polymorphism. The aim is to use at least 30 SSR markers to estimate genetic distances between 500 clones and to group these into gene pools. Moreover, material from within and between gene pool crossings will be developed in order to estimate the amount of exploitable heterosis in the follow up study. Therefore 128 (2 x (8 + 8) = 32 parents) intra-gene pool and 64 (8 + 8 = 16 parents) inter-gene pool crossings will be conducted in a factorial design. This material will be tested in the follow up study under field conditions under drought stress and non-drought stress environments. In total 240 entries (families and parents) will be tested in Peru and Uganda. These field evaluations will start at the end of this project and will be financed by core budget or by other projects from CIP's sweet potato breeding programme.

The heterosis in these drought and non-drought field tests will be determined on the basis of family mean and genetic variation of each cross combination as well as on the basis of comparing the population mean and genetic variation of the inter-gene pool cross population with the corresponding two intra-gene pool cross populations. The combination of this study and the follow up study will make it possible to estimate the correlations between genetic distances and heterosis under drought and non-drought stress conditions for each: (i) intra-gene pool and (ii) inter-gene pool cross population. This will allow us to make conclusions to which extend SSR markers can be used for gene pool division and cross prediction.

3d. Assessment of Genetic Diversity of West African *Dioscorea sp* Collection

Principal Investigator:
R.Asiedu, IITA
Co-Principal Investigators:
V.Mahalakshmi, IITA
H. Chair, Cirad

EXECUTIVE SUMMARY

Yam, a multi-species, polyploidy, and vegetatively propagated tuber crop, is cultivated widely in the tropics and subtropics. Over 90% of world yam production occurs in West and Central Africa where white yam (Dioscorea rotundata Poir.) is the most important cultivated species. It is grown in diverse agroecologies including humid forest and lowland/mid-altitude savannas. Early season drought tolerance in varieties adapted to the savannas allows flexibility in planting periods. Identification of germplasm with tolerance to early season stress would provide the needed flexibility in planting.

IITA holds over 3,000 landrace accessions from 10 African countries of this crop in its ex-situ seed bank. Despite yam's importance in sub Saharan Africa, breeding efforts and dissemination of improved varieties have been limited and farmers continue to grow local landraces that are low in productivity. Genetic diversity of these landraces can be evaluated using molecular techniques as a first step towards identification of suitable diverse parents for use in breeding programmes. In the past diversity of few of the landraces were assessed with isozymes, AFLP, RAPD, or SSRs. Diversity and Genomes of Cultivated Plants" (DGPC), IRD, and Cirad are focusing on SSR markers, which were developed jointly (or in collaboration) for yam diversity analysis. Eleven microsatellite loci have been used to analyse over 500 accessions from Benin. Their current objective is to have 20 SSR loci, which would be ideal for diversity analysis under the Challenge Programme. A systematic approach would be to establish a rational collection of this African collection using morphological traits and then to determine the diversity of this collection using SSRs.

3e. Genotyping a Composite Germplasm Set of Lentil

Principal Investigators:
Bonnie J. Furman, ICARDA
Michael Baum, ICARDA
Collaborating Scientist:
Christian Jung, Universität Kiel, Germany

EXECUTIVE SUMMARY

Lentil (*Lens culinaris* Medik.) is an important cool-season crop in North Africa, West Asia, the Middle East, the Indian Subcontinent, and North America (Erskine 1996). It is an important source of dietary protein (25 percent) in both human and animal diets, second only to soybeans as a source of usable protein (CGIAR). Lentil ranks seventh among grain legumes and is grown on over 3.5 million hectares in over 48 countries with a total production of over 3 million metric tons. The major lentil producing regions are Asia (58 percent of the area) and the West Asia-North Africa region (37 percent of the acreage of developing countries).

The Generation Challenge Programme Subprogramme 1 has the main goal of exploring genetic diversity of global germplasm collections of the Consultative Group of International Agricultural Research (CGIAR). For each crop, a "composite sets" of germplasm, representing the range of diversity of each crop species and its wild relatives, will be identified and characterised with anonymous molecular markers. This molecular characterisation will allow for a study of the diversity across a given genus as well as potentially identify candidate genes involved in resistance to biotic and abiotic stresses, thus providing the base for the research activities of the other 4 subprogrammes.
The International Centre for Agriculture in Dry Areas (ICARDA) has a global mandate for research on lentil improvement. As such, ICARDA houses the world collection of Lens, totaling 10,509 accessions. The ICARDA collection includes 8,789 accessions of cultivated lentil from 70 different countries, 1,146 ICARDA breeding lines, and 574 accessions of 6 wild Lens taxa representing 23 countries. From this collection, a composite germplasm set of approximately 1,000 accessions will be identified and characterised utilising molecular microsatellite markers.

Microsatellite markers in lentil (about 80) have been developed by ICARDA recently and some of them (30) have already been assigned to linkage groups (Hamwieh et al. 2004, Eujayl et al. 1998).
Microsatellite-DNA markers will be used to obtain baseline data on allelic diversity of a composite germplasm set of lentil. These data will then be used to determine allelic frequency distributions for each locus within the collection as a whole and within source regions, as well as the geographical population genetic structure displayed by these loci among source regions. The analysis of genetic diversity will help elucidate population structures that influence the analysis of the associations between markers and phenotypes for important traits. Phenotypic data will be collected for the population.

3f. Molecular Characterisation of Tier 2 (Orphan) Crops

Principal Investigator:
HD Upadhyaya, ICRISAT
Collaborating Scientists:
Subhash Chandra, ICRISAT
JFM Valls, EMBRAPA
MC Moretzsohn, EMBRAPA
S Leal-Bertioli, EMBRAPA
Patricia Guimarães, EMBRAPA
David Bertioli, Universidade Catolica de Brasilia, Brazil

EXECUTIVE SUMMARY
Development of a composite subset of 1,000 accessions of cultivated and wild Arachis using available phenotypic characterisation and evaluation data will be useful in deciding a representative subset of Arachis. Marker analysis of this subset will be done to know the genetic diversity within the subset which will help in defining the population structures that influence the analysis of association between markers and trait's phonotype. Using the data on genetic diversity, this composite collection could be further sampled to a manageable number, about 300 which will be used in further studies in the Challenge Programme.

3g. Molecular Characterisation of Tier 2 Crops: Coconut

Principal Investigator:
Jean Christophe Glaszmann, Cirad-Agropolis

EXECUTIVE SUMMARY
Coconut is a major tropical tree crop and represents a source of livelihood for many resource poor farmers. As an oil crop, it is facing the concurrence of soybean and oil palm on the international market. It retains, however, an important economical role in the tropics as a multi-purpose crop, providing income and various products for domestic consumption. Several lethal diseases, especially lethal yellowing (LY) represent major threats for coconut in different regions of the world. As yet, no cultivar has been found wholly resistant to LY, even though evidence for genetic differences between cultivars does exist. Setting up resistance trials is particularly difficult, because the vector of the disease is unknown and no method for inoculation has been found. In addition, coconut, as a tree crop, is bulky and has a long interval

between generations, making coconut breeding uneasy, especially in the case of multiple breeding objectives (yields, fruit component quality, and disease resistance).

Most coconut cultivars – with the exception of self-pollinating Dwarf coconuts – represent highly heterozygous populations, some of which have been improved in research stations. High yielding F1 hybrids between populations are also available but represent a relatively small proportion of cultivated coconuts.

In the present situation, disease resistance is a major challenge for coconut researchers and breeders. The latter have to concentrate on two objectives: firstly, characterising the few coconut cultivars for which there are evidences of the presence of resistance factors, as well as related cultivars; secondly, identifying genetic factors (e.g. QTLs) involved in total or partial resistance.

4. Validation of Diversity Arrays Technology (DArT) as a Platform for Whole Genome Profiling in Orphan Crops

Principal Investigators:
Andrzej Kilian, DArT P/L, Australia
Carmen de Vicente, IPGRI
Jean Christophe Glaszmann, Cirad-Agropolis
Co-Principal Investigators:
Eric Huttner, DArT P/L, Australia
Peter Wenzl DArT P/L, Australia
Ange-Marie Risterucci, Cirad-Agropolis
Collaborating Scientists:
Ken McNally, IRRI
Claire Billot, Cirad-Agropolis
Michael Baum, ICARDA
M Fregene, CIAT
Nicolas Roux, IPGRI-INIBAP
Patricia Lebrun, Cirad-Agropolis
Everard Jayamanne, Coconut Research Institute, Sri Lanka
Prapit Wongtiem, Rayong Field Research Station, Thailand

EXECUTIVE SUMMARY
There are many constraints to the widespread use of molecular markers for diversity analysis of germplasm and the subsequent identification of associations between traits and genes. Limitations include some of the following: theoretical and practical lack of knowledge of the tools, cost of development, low reproducibility, low data yield (limited throughput) of the experiments, restriction of access to proprietary technologies, and insufficient resources (laboratory facilities, equipment, chemicals, etc).

The project proposes to test the usefulness of Diversity Array Technology (DArT) as an alternative for detecting DNA variation in ways that will result to be more effective and resource-efficient. DArT possibly offers the highest throughput available up to date and allows for whole genome scanning in a speedy manner. In addition, the types of polymorphism detected by DArT (single nucleotide polymorphisms, insertion-deletions, and methylation changes) may expand the potential of traditionally used markers, increasing the power to ascertain the structure of germplasm collections. Lastly, the experimental

procedures to obtain DArT take into account the complexity of genomes and its effect on the extent of diversity shown by a collection of germplasm. Therefore, a set of different cases will be addressed, using the biological diversity of the crops and the issues, focusing the development of new resources on orphan crops that are not likely to gain much attention elsewhere.

Two additional goals of the project are facilitating DArT technology transfer to members of the Generation CP Consortium (CIAT, Colombia, and Agropolis, Europe), and building capacity of NARS scientists through their involvement in technology development at DArT P/L so that they can be a resource for subsequent technology transfer.

5. Assessing EcoTILLING as a Methodology for Targeted Genotyping and SNP Discovery

Principal Investigator:
Kenneth L. McNally, IRRI
Co-Principal Investigator:
Claire Billot, Cirad-Agropolis
Collaborating Scientists:
Luca Comai, University of Washington
Jeff Harford, Li-Cor, Inc., USA
Abdelhafid Bendahmane, Unité de Recherche en Génomique Végétale, INRA, and CNRS, France

EXECUTIVE SUMMARY
TILLING (Targeting Induced Local Lesions IN Genomes) is a new technique that can identify polymorphisms in a target gene by heteroduplex analysis. A variation of this technique (EcoTILLING) represents a means to determine the extent of natural variation in selected genes in crops. EcoTILLING may be a cost-effective approach for haplotyping and SNP discovery.

The objectives of the projects are i) to assess Eco-tilling as a reliable and cost-effective method to detect SNP in a large number of accessions, ii) to test for validity in triploid species, and iii) to establish Eco-tilling transfer technology platforms at IRRI and Agropolis-Cirad. These will be performed through the study of 10 orthologous genes in three related species, two diploid (rice and sorghum) and one presenting different ploidy levels (Musa).

6. Supporting Emergence of Reference Drought Tolerance Phenotyping Centre

Principal Investigator:
Frederico Ozanan Machado Durães, EMBRAPA
Co-Principal Investigators and Collaborating Scientists (EMBRAPA):
Antonio Carlos de Oliveira
Antonio Marcos Coelho
Camilo de Lélis Teixeira Andrade
Elto Eugenio Gomes e Gama
Fredolino Giacomini dos Santos
Paulo Emílio P. de Albuquerque
Manoel Xavier dos Santos

Reinaldo Lúcio Comide
Beatriz da Silveira Pinheiro
Cleber Morais Guimarães
Orlando Peixoto de Morais
Natoniel Franklin de Melo
Luiz Balbino Morgado
Hélio Wilson Lemos de Carvalho
Luciana Marques de Carvalho
Milton José Cardoso
Edson Alves Bastos
Francisco Rodrigues Freire Filho
Maria da Glória Trindade
Walter Quadros Ribeiro Jr.

EXECUTIVE SUMMARY

The development of drought tolerant varieties for crops of economical importance presents a major challenge for the 21st. century, considering that agriculture growth will be limited by world water availability. A first step to be taken in this direction is to select germplasm adapted to water stress conditions through appropriated screening techniques and defined protocols. Thus, the great challenge is the identification and characterisation of drought tolerant genitors to provide material to be used in genetic breeding programmes focused on regions historically known as prone to water deficit during crop growing season. The improvement of drought tolerance relies on the manipulation of the traits that limit yield and their accurate phenotyping under the prevailing field conditions being targeted. This issue is particularly crucial for the breeding programme and identification of QTLs for traits categorised as adaptive as compared to constitutive traits per each species. For this purpose it is necessary to amplify an infrastructure to allow plant exposure to water deficit pressure to be used for the evaluation of genotypes and characterisation of plant physiological responses to these stress conditions.

The objectives of this project are to develop and make useful phenotypic evaluation protocols for cereals (maize, sorghum, rice, and wheat) and legume crops (common bean and cowpea), as well as to establish the amplification of the three Phenotyping Centre of Excellence for Drought Tolerance Studies composed of phenotyping central laboratories, including controlled environment field and greenhouse and a training unit for researchers and research assistants, and six-eight experimental stations located in regions with facilities and well defined dry season periods to assure total soil moisture control during the drought phenotyping field experiments. In fact, the project seeks to establish a scientific and service net, like a model to drought tolerance phenotyping of cereals and legumes, including national and international genotypes.

Embrapa-National Maize and Sorghum Research Centre, as the main coordinating institution, accumulates over 30 years of experience working with the application of phenotyping methodologies, conducting maize and sorghum breeding programmes, and releasing drought tolerant germplasm. Also, all partners have extensive experience with others crops such as rice, bean, wheat, etc. The innovative character of the present proposal consists of having different crop expertise aggregated in one single project stimulating the exchange of personal experiences, providing simultaneous experiment conduction and data integration, establishing news and future partnerships for simulation models, and also promoting knowledge diffusion by planning and organising training courses.

7. Whole Plant Physiology Modeling of Drought Tolerance in Cereals

Principal Investigator:
Marcel de Raïssac, Cirad
Collaborating Scientists:
Delphine Luquet, Cirad-Agropolis
François Tardieu, Cirad-INRA
Renée Lafitte, IRRI
Jean-Marcel Ribaut, CIMMYT
M Dingkuhn, Cirad-Agropolis
M de Raïssac, Cirad-Agropolis
JC Combres, Cirad-Agropolis
C Welker, Cirad-INRA
Scott Chapman, CSIRO. Australia
Graeme Hammer, University of Queensland
B. Bouman, IRRI
M Bänzinger, CIMMYT
M Reynolds, CIMMYT
R Trethowan, CIMMYT
Eva Weltzien, ICRISAT
Frederico Duraes, EMBRAPA
Mark Cooper, Pioneer

EXECTUTIVE SUMMARY
The present project gives continuation to the GCP phenotyping workshop organised in July in Montpellier, where more than 40 breeders and physiologists from inside and outside the consortium met for a week. Conclusions of the workshop (available on GCP website www.generationcp.org) stressed the importance modeling in supporting phenotyping process for drought tolerance by : (i) a quantification of traits and integration of their impact on yield, (ii) a genetic analysis of adaptive traits, (iii) a characterisation of target population of environments.

The need for better interactions between physiologists, modelers, and breeders to develop a comprehensive approach and improve phenotyping methods and outputs was also stressed during the meeting and must be kept in mind as a main issue of this project.

This project is not unique within the GCP to develop modeling approaches and deliver new tools. Consequently, it proposes interactions or complements with other initiatives:
- The project 28 “An ecophysiological-statistical framework for the analysis of G X E and QTL X E,” focusing on more statistical concepts, in which some environment characterisations will be carried out using the same models for wheat and maize in this project.
- The commissioned project on “Simulation on marker-assisted selection strategies.” A tentative will be done here to link the two projects on the basis of component 3 activities.
- The commissioned project on “Reference drought tolerance phenotyping centres.” A first interaction is planned on environment characterisation of these centres by model use.
- Interaction will be sought with the project on “Modeling alternate drought tolerance strategies on globally important crops” lead by IFPRI on the impact of improved drought tolerance characterization.

8. Population Structure, Phenotypic Information, and Association Studies in Long-Generation Crops

Principal Investigators:
M Carmen de Vicente, IPGRI
Martin Fregene, CIAT
Luc Baudoin, CIRAD
Kodjo Tomekpe, CARBAP, Cameroon
Merideth Bonierbale, CIP
Jean-Louis Noyer, CIRAD
Co-Principal Investigators:
Toby Hodgkin, IPGRI
Jean Christophe Glaszmann, CIRAD
Marc Ghislain, CIP
Reinhard Simon, CIP
Vincent Lebot, VARTC, Vanuatu
Collaborating Scientist:
Nicolas Roux, IPGRI-INIBAP

EXECUTIVE SUMMARY
Identification of useful genes or chromosome segments involved in traits of agricultural interest rests on the search for co-occurrence of molecular tags with desired values for the target traits. This is commonly undertaken by segregation analysis in controlled progenies, where co-occurrence will be indicative of linkage on the genome. In such experiments, a particular progeny is generated, planted, and evaluated while it is genotyped in the laboratory for markers covering the genetic map of the species. This process has several constraints:
- it requires time for making the crosses and growing the progeny
- it confronts a limited number of alleles at each locus (as many as there are in the few parents)
- it requires specific phenotyping experiments, which usually represent an additional burden to current breeding programmes.

The results generally suffer from several drawbacks. The progeny often represents types that are far from the breeding standards, exhibiting potential interactions between traits that may confound variation for the target features. Phenotyping is often done with limited numbers of plants, i.e. few repetitions over space and still fewer over time. The use of the materials thus produced and monitored is not easy and they are seldom incorporated in the breeding process.

An alternative option could be to try and make use of materials and evaluation data that are regularly produced in the mainstream of the breeding activities. From the collections of potential parents to the advanced breeding materials going to multilocation trials and to the elite materials close to varietal release, there is a wealth of information produced, which is hardly used for deriving genetic information that can in turn be used to enhance global understanding and mastership. The condition for using these materials is that there be significant linkage disequilibrium (LD) that correlates variation in genetically (recombinationally) linked genes/markers. LD can exist among traditional (modern-pedigree wise unrelated) materials if the germplasm has known significant bottlenecks in the past. For example, LD is actually generated in the breeding materials by the crosses made; in case of a well documented breeding programme, the multi-generation pedigrees can be used to derive

linkage information more efficiently. A nice example of such an application has just been described in barley where standard varietal evaluation trials could be used to detect QTLs for yield and yield stability (Kraakman et al, 2004).

SP2 COMMISSIONED GRANTS

Principal Investigator:
Andy Pereira, WUR
Co-Principal Investigators:
Hirohiko Hirochika, NIAS, Japan
Hei Leung, IRRI
Emmanuel Guiderdoni, Agropolis
Mathias Lorieux, IRD/CIAT
Manabu Ishitani, CIAT
Tiegang Lu, CAAS
Qifa Zhang, HAU, China
Deming Jin, HAU, China
Collaborating Scientists:
Members of the IRFGC
Gyn An Pohang, University of Science and Technology, Korea
Srinivasan Ramachandran, Temasek Lifesciences Laboratory, Singapore
Narayana Upadhyaya, CSIRO, Australia
Venkatesan Sundaresan, UC Davis

EXECUTIVE SUMMARY
The rice genome sequence provides the basic framework for the functional analysis of monocot genomes. International efforts have produced rice mutant resources that are a powerful functional genomics tool to identify the function of the sequenced genes. This consortium proposes to create a platform to identify genes that can contribute to a phenotype of resistance/tolerance to abiotic/biotic stresses. Expression analysis of various plant genomes has revealed genes that respond to environmental stresses. In addition, the ongoing functional analysis of Arabidopsis and other model plants identifies genes and associated mechanisms involved in stress tolerance. This comparative genomics information will be integrated to identify a set of candidate rice genes predicted to be associated with drought and other stress responses.

In a reverse genetic strategy, we will use sequence-indexed knockout mutant resources developed by the international rice community to identify insertions in the target stress-associated genes. These insertions include rice Tos17 transposons, heterologous Ac-Ds transposons, and T-DNA inserts that are identified by available flanking sequences (FSTs), supplemented by PCR-based screens for inserts. In addition TILLING will be used to supplement the mutant coverage, and also provide non-transgenic stocks for field testing. To resolve gene redundancy, gain-of-function overexpression lines will be generated by transformation of appropriate constructs, and available activation tag populations will also be accessed for overexpression mutants.

To facilitate novel gene discovery a forward genetics mutant screen will be carried out with random genotypes that have been sequence indexed, thus aiding their further analysis. The mutant genotypes will be phenotyped for drought and disease stress parameters at appropriate

growth stages in greenhouse and field-based screens. Genotypes with altered stress tolerance phenotypes will be tested for other abiotic/biotic stresses. Whole genome microarray analysis will be conducted to identify the downstream genes and characterise the stress response mechanism involved. The genes can be utilised directly by transformation to provide stress tolerance, or for cereal comparative genomics studies and allele mining for breeding.

10. Collection, Distribution, Phenotyping, and Genotyping Directed towards Utilisation of Existing Wheat Genetic Stocks to Enhance Tolerance/Resistance of Wheat Cultivars to Abiotic and Biotic Stresses with Emphasis on Drought

Principal Investigators:
Maarten van Ginkel, CIMMYT
Hans-Joachim Braun, CIMMYT
Collaborating Scientists:
Peter Langridge, University of Adelaide
Xueyong Zhang, CAAS
Marion Röder, Gatersleben, Germany
Tetsuo Sasakuma, Kihara Institute for Biological Research, Japan
Hitashi Tsujimoto, Tottori University, Japan
Masahiro Kishii, CIMMYT
John Snape, John Innes Centre
Jorge Dubcovsky, University of California Davis
Bikram Gill, Kansas State University
Bernd Friebe, Kansas State University
Perry Gustafson, USDA-ARS
Adam Lukaszewski, University California Riverside
Mark Sorrells, Cornell University

EXECUTIVE SUMMARY
Since 1936, a wealth of genetic stocks has been developed in tetraploid (Triticum turgidum L.) and hexaploid (T. aestivum L.). These genetic stocks include material containing intervarietal and interspecific translocations, chromosome and chromosome arm additions and deletions, chromosome and alien substitution addition lines, mono- and polysomic series, recombinant doubled haploid populations, mapping populations, NILs, point and other mutations, and synthetics. Hexaploid wheat is allopolyploid in origin, and the homoeology existing between its three component genomes allows for a range of aneuploidy (i.e. additions, substitutions, deletions, etc.) to be tolerated.

These genetic stocks were often developed for specific purposes, e.g. as new source of disease resistance genes, but have seldom been systematically screened for other value-added traits of interest. A few genes introduced from other species, e.g. rye (Secale cereale L.), have had a tremendous impact on wheat improvement. It can be assumed that a systematic screening of available genetic stocks will reveal useful genetic variation for many value-added traits of immediate interest to breeders. An important advantage is that any gene characterised in any of the existing or newly created wheat genetic stocks can be transferred to improved wheat cultivars without requiring the utilisation of any biolistic or Agrobacterium transformation techniques. And finally, it is anticipated that the distribution of these genetic stocks worldwide will result in many research projects, in particular in the area of gene identification (including across species/genera), marker development, and

association mapping, which will greatly increase our knowledge of wheat genetics and breeding efficiency in the future.

It is intended that, to the fullest extent feasible, all genetic stocks within this project will be placed within the Multilateral System of Access and Benefit Sharing to be established under the International Treaty of Plant Genetic Resources for Food and Agriculture (ITPGRFA), and that they may be freely distributed under the standard ITPGRFA material transfer agreement (MTA). (FAO has an explanatory page about the Treaty at http://www.fao.org/ag/cgrfa/itpgr.htm, and a video at http://www.fao.org/videocatalogue/index.jsp?lang=EN).

It is also intended that, to the fullest extent feasible, results obtained from using these genetic stocks will be considered to be public goods which can be used without restrictions by others. Due to a number of ambiguities in the ITPGRFA that have not yet been resolved, it is not feasible to anticipate all potential implications of the ITPGRFA at this time. Accordingly, to the extent that any germplasm used in this project is protected by intellectual property rights, and to the extent that placing such germplasm within the Multilateral System would be inconsistent with such rights, the parties will endeavor to devise a system for working with and distributing such germplasm that both respects those intellectual property rights and achieves the public goods aims of this project as outlined above.

The preference of this programme, however, shall be for using germplasm that is free of intellectual property rights. Stocks that are initially requested through the proposed central database-facilitated CIMMYT website or by normal email to CIMMYT representing the consortium of key genetic research groups listed below, may later actually be sent out from the originator's programme. This may be especially the case for cytogenetic stocks requiring the resident expertise to guarantee quality maintenance of the stocks. CIMMYT would function in that case as a clearing house for routing the requests. The latter routing would allow the material to be shared under the originator's MTA.

11. Legume Mutant Resource Development

Principal Investigator:
Matthew Blair, CIAT

EXECUTIVE SUMMARY
Grain legumes have a paucity of mutant resources compared to Arabidopsis and the cereals. Between the two broad branches of grain legumes, the tropical legumes have fewer mutant stocks compared to the temperate legumes. A well-developed mutant stock, particularly in genotypes of common bean (Phaseolus vulgaris), a simple diploid species with a small genome (650 Mb), will serve the broad community involved in tropical legume improvement aiding gene-discovery both in common bean, the most widely consumed grain legume for human consumption and a major protein and mineral source in East Africa and Latin America, as well as in two tropical legume relatives: cowpea (Vigna unguiculata), a food crop important in West African agriculture ,and soybean (Glycine max), a major industrial and feedstock crop around the world. Mutant stocks in common bean will allow researchers to conduct both forward (systematic phenotypic screening) and reverse genetics (TILLING or Targeting Induced Local Lesions In Genomes) experiments aimed at understanding the genes involved in abiotic and biotic stress tolerance as well as those genes involved in biological nitrogen fixation in the tropical legumes. Mutations will be sought in common bean genes

that have been isolated at CIAT and shown to be associated with drought tolerance. The phenotypic effect of these mutations will be analysed as a proof of concept for the value of the mutant stocks generated by this project.

12. A Saturated Potato Mutant Population for Functional Genomics among Solanaceae and Tuber Crops

Principal Investigator:
Meredith Bonierbale, CIP
Marc Ghislain, CIP
Collaborating Scientists:
Glenn Bryan, Scottish Crop Research Institute
Robbie Waugh, Scottish Crop Research Institute
Dani Zamir, The Hebrew University of Jerusalem

EXECUTIVE SUMMARY
Complete sequencing of genomes and transcriptomes provides baseline genetic information, which is increasingly being exploited between related species by comparative genomics. However, the identification of gene function remains one of the most challenging tasks of plant functional genomics. The virtually unlimited number of alleles and allele combinations possible for a single organism makes necessary the development of special tools and genetic stocks designed for functional analyses. Tuber crops, including the potato, generally have limited or poorly accessible genetic stocks. A saturated mutant population of a wild potato species is proposed here as a genetic resource useful for forward and reverse genetics in Solanaceae and other tuber crops. Ethyl methane sulphonate (EMS) mutagenesis will be performed on botanical seed of a diploid self-compatible and near-homozygous accession of the tuber-bearing species Solanum verrucosum to produce approximately 400,000 mutant chromosomes in 20,000 M2 seeds. Validation of the mutant population will be conducted for target genes already characterised in tomato and the data will be presented in the Solanaceae Genome Network (SGN).

13. Crop Gene Expression Profiles and Stress-gene Arrays

Principal Investigator:
Tiegang Lu, CAAS
Co-Principal Investigators:
Guozhen Liu, Beijing Genomics Institute (BGI), China
Shoshi Kikuchi, NIAS
Manabu Ishitani, CIAT

EXECUTIVE SUMMARY
cDNA/oligo microarrays currently provide a robust and accessible platform for genome-wide expression analysis. Whole-genome arrays have been developed for rice, barley, maize, and possibly wheat. The maize and rice chips are being used in Year 1 commissioned research in the Generation Challenge Programme (GCP). The identification of stress-tolerance genes by a combination of genome-wide expression and protein-protein interaction analyses has proven promising (Cooper et al. 2003). While gene expression analyses are being pursued in multiple crops, relatively few attempts have been made to integrate experiments to enable cross-species comparison, which will provide opportunities to study the evolution of biological systems (Zhu et al. 2001; Zhou and Gibson, 2004). The GCP has a unique

opportunity to generate pan-crop gene expression data for comparative analysis and data mining which allows us to draw evolutionary inferences concerning specific trait(s) and to elucidate the global properties of expression networks in crops.

Drought tolerance is a complex trait in plants. Our challenge is to link molecular phenotype to physiological or morphological trait(s) that have been observed in drought tolerant plants/crops. It will be difficult to comprehend gene expression patterns across crops without knowing the physiological phenotype of drought resistance even if existing or developed ortholog arrays are feasible for cross-species comparisons. To address this problem we plan to focus on a common physiological trait found across crops under drought. This will allow us to identify underlying molecular components for a common physiological trait cross crops.

We propose to apply single or multiple microarray platforms to identify candidate genes that are associated with a phenotype of drought resistance across selected crops. The proposed workplan will apply existing rice microarray technology created by NIAS and BGI and test the feasibility of develop ortholog arrays for use in multiple crops. The proposed project will enhance synergy with other projects involved in gene expression studies (maize, millet, wheat) in the GCP toward elucidation of gene function.

14a. Stress Response-enriched EST Resources for Targeted Species—Pearl Millet

Principal Investigator:
Tom Hash, ICRISAT
Co-Principal Investigators:
MK Reddy, International Center for Genetic Engineering and Biotechnology
Arjula Reddy, Central University of Hyderabad, India

EXECUTIVE SUMMARY

ESTs are good tools for producing candidate genes as gene-based markers as well as providing the basis for SSR and SNP marker discovery. The GCP is interested in enhancing EST resources in inherently stress-tolerant species that are unlikely to be supported by others. The proponents have been requested to identify target genotypes of a species for which a modest investment will significantly elevate the potential of applying EST resources in the species. Preference will be given to cases that can leverage existing efforts or collaboration.

For this round of commissioned work, the GCP is particularly interested in pearl millet and cowpea because they represent the “hardy” species of each crop group, and progress has already been made. EST resources developed for both crops will be duplicated/deposited at the BioScience Centre of East and Central Africa (BECA). Depending upon efficiency considerations, the lead institutions may perform the tasks themselves or coordinate the work (i.e., outsource specific aspects of the EST work to one or more appropriate high-throughput labs or institutions).

14b. Stress Response-enriched EST Resources for Targeted Species—Cowpea

Principal Investigators and Collaborating Scientists:
Sarah Hearne , IITA
Morag Ferguson, IITA

Chris Town, The Institute For Genomic Research, USA

EXECUTIVE SUMMARY

Cowpea is grown on at least 12.5 M Ha worldwide with two thirds of global production being concentrated in Central and Western Africa (FAO). Some 3.7 million tonnes of seed were produced globally in 2003 (FAO). Cowpea is an important staple for approximately 200 million Africans where it is valued as a protein-rich nutritional supplement to cereal grains and as stover for improved human and livestock nutrition. Cowpea has a genome size only slightly larger than that of the model legume Medicago truncatula, is homozygous, has a short life cycle and well-studied conventional genetics, yet is poorly characterised at the genomic level. Cowpea varieties exhibit wide variability in the degree of drought tolerance. Thus, cowpea is exceptionally well suited to the application of contemporary technologies, including comparative genomics, to advance gene discovery, marker development and marker assisted breeding (MAB) for enhanced drought tolerance, and other desirable traits. However, there is a lack of molecular tools (SSRs, SNPs, etc.) available to cowpea researchers to enable the efficient exploitation of these contemporary technologies. Thus, the aim of this proposal is to develop a range of ESTs that can be utilised to develop other molecular tools such as SSRs, SNPs, and COS markers, oligonucletides for microarray chips, and probes for quantification of expression using real time PCR.

15. Musa Genome Frame-map Construction and Connection with the Rice Sequence

Principal Investigator:
Takuji Sasaki, NIAS
Co-Principal Investigators:
Nicolas Roux, INIBAP-IPGRI
Isabelle Hippolyte, Agropolis-Cirad
Manoel Souza, EMBRAPA
Collaborating Scientists:
Pat Heslop-Harrison, University of Leicester
Jaroslav Dolezel, IEB

EXECUTIVE SUMMARY

Developing basic genomic tools to assist germplasm exploitation is important for Musa (banana), especially in the context of the use of Musa genomic diversity. Currently, BAC libraries and germplasm resource collections are available, but EST collections are not publicly available and there are no characterised mapping populations available. The development and assembly of publicly accessible EST collections and framework maps anchored with genetic markers are top priorities for this species, along with initiation of genomic sequencing to identify genes and regulatory genome regions in this important staple crop of developing countries that involves two different genomes and many triploid cultivars.

The GCP is interested in promoting the development of such genomic resources and integrating them with ongoing trait-based research in Musa. The GCP also awaits assessment of the utility of rice-based information for application in Musa and other mandate crops The GCP is thus requested to support the development of genomic resources for anchoring of chromosomal regions harbouring valuable traits as well as targeted comparative genomic sequencing relating the A and B Musa genomes and the rice genome. The approach suggested will utilise existing resources of Musa germplasm, maps and genomic resources,

and the outputs will be publicly available resources and markers for exploiting breeding-relevant genetic variation within Musa.

16. Validation of Conserved Orthologous Markers

Principal Investigators and Collaborating Scientists:
Jizeng Jia, CAAS
Lifeng Gao, CAAS
Qingming Sun, CAAS
Nicolas Roux, INIBAP
Pat Heslop-Harrison, University of Leicester
Mathieu Rouard, INIBAP
Merideth Bonierbale, CIP
Roland Schafleitner, CIP
Reinhard Simon: CIP
Michael Baum, ICARDA
Sripada M. Udupa, ICARDA

EXECUTIVE SUMMARY
Generation Challenge Programme support is requested for consolidated experimental work on conserved markers to bring Year 1's commissioned research in SP2C2 to a meaningful conclusion (Year 1 proposal; SP2 Progress report and Brisbane reports of Jia and Bonierbale).

Gene-based PCR markers, also called functional markers, are a third generation molecular marker, more powerful than RFLP markers (first generation) and other PCR types (second generation) for marker-assisted selection and allele discovery. However, the shortage of functional markers has become a bottleneck in a number of marker-based genetic and breeding applications. Orthology is an important aspect of functional markers, particularly for applications in comparative genomics. Efforts toward developing sets of sequenced COS (conserved orthologous set) markers within and across crop families have been undertaken by labs at Cornell University and U.C. Davis with encouraging results. In silico research first identified orthologues across dicots by comparing gene sequences from distant genomes ranging from tomato to Arabidopsis. Research at UC Davis has assembled a bioinformatics pipeline for COS identification. The availability of a set, or sets, of sequenced COS markers would be a great advance for the study and understanding of genetic diversity.

17. Comparative QTL Mapping for Drought Tolerance

Principal Investigator:
Mathias Lorieux, IRD/CIAT

EXECUTIVE SUMMARY
The first sub-project (led by CIAT) aims at 1) pursuing the identification of QTLs for drought that were mapped on the common bean genome. The target trait will be the capacity of common beans to maintain the transport of photosynthetically-fixed carbon into grain under drought stress. 2) Using candidate genes to define optimal genetic markers for the transfer of these QTLs. 3) Aligning drought QTLs between legume species and determining the most important regions for saturated mapping.

The second sub-project (led by Agropolis) will focus on 1) refining rice QTLs for root development, 2) providing a rice genetic framework for comparative mapping of drought related QTLs and candidate genes, 3) providing QTL data for in silico determination of candidate genes.

The genetically determined QTL from these sets of experiments conducted in different crop species will contribute to the convergent evidence for pinpointing common or unique candidate genes important for drought tolerance.

SP3 COMMISSIONED GRANTS

18. Development of Low Cost Gene Based Trait Assay Technologies in Cereals

Principal Investigator:
Casiana M. Vera Cruz, IRRI
Co-Principal Investigator:
Manilal Williams, CIMMYT

EXECUTIVE SUMMARY
The Generation Challenge Programme (GCP) has a primary focus on the development of gene-based markers for drought tolerance using comparative genomics and comparative biology. These efforts aim to drive rapid progress across three main crop groups (cereals, legumes, and clonal crops) through comparative analysis with model systems. The development of effective gene-based molecular breeding systems for drought tolerance is clearly a long-term goal in most crops. However, it is fundamentally important that the GCP works with end-users from an early stage to foster the integration of successful, low cost applications of gene-based MAS technologies. Most importantly, we must build champions for these new technologies among breeding programmes of national agricultural research systems (NARS) and small-to-medium-sized enterprise (SME) breeding companies across the major regions of the developing world. Clearly, gene-based technologies for abiotic stress tolerances are only just emerging. Thus, other target traits must be focused upon for this initial pilot project concerning the proof-of-concept for technology product delivery pathways associated with gene-based marker-assisted selection tools. This also offers the added advantage for this preliminary project of being able to target traits with a somewhat simpler genetic basis than drought tolerance. In this context, we envisage two very different models for helping NARS and SME breeders establish their own molecular breeding success stories:
(i) low cost (both in terms of capital investment and assay unit costs), low tech (to save time and offer scalable options) assay technologies that can be used by anyone anywhere;
(ii) low cost (most critically with respect to assay unit costs), high throughput (in terms of millions of samples per year) assay technologies that can be used in shuttle genotyping/MAS service labs.

19. Evaluation and Deployment of Transgenic Drought-Tolerant Varieties

Principal Investigator:
John Bennett, IRRI

Collaborating Scientists:
Renee Lafitte, IRRI
Matthew Reynolds, CIMMYT
Vincent Vadez, ICRISAT
Enrique Chujoy, CIP
Kazuko Yamaguchi Shinozaki, JIRCAS
Kazuo Watanabe, University of Tsukuba

EXECUTIVE SUMMARY
Drought is an important limitation on productivity for all of the mandated crops of the CGIAR system. While land and water resource management improves water harvesting and water use efficiency at the field, community, and regional levels, progress in breeding for enhanced drought tolerance is essential for achieving improved crop water productivity and greater food security for hundreds of millions of the rural poor who depend on rainfed agriculture (Bennett, 2003). Water savings in irrigated agroecosystems will also require breeding for tolerance of water deficit.

Although lines of several mandated crops are now available with shorter duration (drought escape) and DNA markers are being identified for deeper and more penetrating roots (drought avoidance), there is also a considerable potential for deploying drought tolerance genes that enable plants to survive and recover from unavoidable periods of low plant water status, especially at the sensitive reproductive stage.

20. Optimising Marker-assisted Breeding Systems for Drought Tolerance in Cereals through Linkage of Physiological and Genetic Models

Principal Investigator:
Scott Chapman, CSIRO
Collaborating Scientists:
Mark Dieters, University of Queensland
Graeme Hammer, Agricultural Production Systems Research Unit (APSRU)
Jiankang Wang, CIMMYT
Maarten van Ginkel, CIMMYT
Richard Trethowan, CIMMYT
Eva Weltzien, ICRISAT
Tom Hash, ICRISAT
Gary Atlin , IRRI
Marianne Banziger, CIMMYT
Mark Cooper, Pioneer

EXECUTIVE SUMMARY
The dynamic linkage of crop modeling and genetic/breeding simulation allows us to simulate such things as the introgression or marker assisted selection of traits as affected by population genetic structures, selection criteria (e.g. direct or indirect selection for yield), and trait by environment interactions. The aim of this project is to build up the gene to trait information using data from QTL and physiological experiments and to further improve the ability of our crop simulation models to capture the effects of traits and their integration to yield. It will aim to combine these ‘gene-to-phenotype’ physiological models with existing genetic models for other traits such as disease and quality. Simulating molecular breeding programmes will

enable optimisation of MAS strategies and provide a platform for integrating a range of knowledge-based outputs from GCP into breeding programmes.

21. Planning for Effective Product Development, Delivery, and Use

Principal Investigator:
Victoria Henson-Appollonio, IPGRI
Collaborating Scientists:
Silvia Salazar, Consultant (Lawyer), Costa Rica
Maria Ines Mendosa, Consultant (Lawyer), Columbia
Shawn Sullivan, Consultant (Lawyer), USA
Zenete Franca, IFPRI-ISNAR
Rosemary Wolson, University of Capetown
Jocelyn Webster, AfricaBio

EXECUTIVE SUMMARY
Through competitive and commissioned research projects, the Generation Challenge Programme (GCP) is building an extensive number and range of research outputs targeted at the identification of useful germplasm, traits, genes, and alleles for use in enhancing drought tolerance in cereal, legume, and clonal crops. The GCP is also generating knowledge and tools to help plant breeders manipulate those genetic factors. It is well accepted that these resources, tools, and knowledge can be used to improve the food, nutritional, and economic security of small-holder farmers and their families in drought-prone areas. However, uptake of research outputs and seed-based technologies has been patchy and often disappointing. It is, therefore, essential for biotechnologists to consider the entire, innovation-to-impact, pathway whilst designing and carrying out their research programmes. This type of holistic approach is a highly complex multidisciplinary and multisector endeavour. Yet this is the way to increase the rate of uptake of GCP research outputs by ensuring the development of products that provide end-users with practical, user-friendly, appropriate technology packages that meet their needs, capabilities, and capacities.

Emphasis on using an explicit end-user orientated mapping and planning process to improve product development and delivery is a new approach for many of the scientists involved in GCP research projects. Current strategic biotechnology research projects often define indicators for their outputs in terms such as “publication of the results in …”. However, the GCP Subprogramme 3 is committed to populating the intellectual space between conventional academic research outputs and the deployment of research-based products that can have tangible impacts on primary agricultural development parameters. Historically, this is a hugely neglected research domain in the public sector and is surely a major reason for the sub-optimum uptake of research outputs. For this reason, the GCP wishes to move far beyond traditional technology hand-over and training models. Instead, tangible product development and deployment plans must include information on how outputs will serve the end-users, why those end-users would favour transfer/adoption of these new technologies, where the critical linkages for product testing, refinement, and delivery are embedded in the plans, and who or what may be the rate limiting steps for product development and deployment. It is fundamentally important that these issues are considered at the project design phase by adequate representation and decision-making influence by stakeholders from across the entire innovation-to-impact pathway.

Adoption of a more holistic and product-driven approach requires substantial changes in institutional policies and processes, in the perspectives of individual scientists. Many consortium member institutions are already advancing towards these general goals. To complement this, the GCP is committed to fostering targeted progress by creating resources, tools, and case studies for assisting current and future projects. It is envisaged that ultimately the GCP will establish a framework of criteria and guidelines in support of mandatory inclusion of product development and deployment activities in all competitive and commissioned projects. Initially we propose to pilot this approach in a facilitated mode during 2006 and autonomously from 2007 (with assistance from GCP Helpdesk support services). Thus, this project proposes to collaborate with project scientists to develop a number of diverse case studies from fundamentally important areas of the current competitive and commissioned programmes. These case studies will be used to develop specific product development and delivery pathways, generic process templates, resources and guidelines for pathway analysis, and defining policy statements. In this way, not only will the most immediate concerns be dealt with but resources and patterns of behaviour will be established for use in the development of future project proposals.

SP4 COMMISSIONED GRANTS

22. Development of GenerationCP Domain (Data) Models

Principal Investigator:
Richard Bruskiewich, IRRI
Collaborating Scientists:
Reinhard Simon, CIP
Manuel Ruiz, CIRAD
Tom Hazekamp, IPGRI
Masaru Takeya, NIAS

EXECUTIVE SUMMARY
The design of GenerationCP information systems, tools, and data exchange protocols must be specified by domain ("data") 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, will form the basis for "model driven architecture" generation of templates, web services and software driven by scientific use cases underlying the project.
Task objectives (by end of 2005):
• Assist in the compilation of scientific use cases, to elucidate domain models pertinent to the GenerationCP.
• Extend domain models initiated in year 1 to meet the requirements for priority scientific informatics use cases, including capture of priority SP* data sets.
• Commission a community editorial process for domain model development.
• Collaborate with "Creation and maintenance of templates for GenerationCP data storage in repositories" to translate domain models into data templates.
• Commission software tools to translate the domain models into data type and service type ontology specifications for web services implementation.
• Commission software tools to translate the domain models into components of the reference platform for "Improvement of quality of existing databases."

23. Implementation of Web Services Technology in the GCP Consortium

Principal Investigator:
Samy Gaiji,, IPGRI/SGRP
Participating Institutions:
All Generation CP Consortium members
Other CGIAR Centres (CIFOR, WorldFish, World Agroforestry Centre, ILRI, IFPRI), SGRP1, MOBY2, GBIF3, FAO4, EURISCO5, USDA-GRIN6, ICT-KM7, CSI8

EXECUTIVE SUMMARY
The Generation CP will generate a wide spectrum of information from conventional phenotypic data to molecular data. The primary objective of Subprogramme 4 is to establish an informatics network that will allow the sharing and analysis of such information in support of the other Subprogrammes. To meet the challenge of accessing dispersed information sources, the Generation CP has decided to adopt Web Services as its core technology for information sharing. In 2004, the technology was extensively tested by ICIS9, the Global Musa Genomics Consortium10 and SINGER11/EURISCO. The results of these studies demonstrated the scalability of the technology as well as its potential for the Generation CP. At its Annual Research Meeting in Brisbane, Australia, in September 2004, the Generation CP adopted the technology and decided to move rapidly to its deployment by all Consortium members.

The purpose of this project is the practical implementation of Web Services at Generation CP sites that would contribute to distributed access to information generated by scientists. This project will build on the outputs of other SP4 commissioned projects, namely: “GCP Model Driven Platform and Network Architecture,” “Creation and maintenance of templates for Generation CP data storage in repositories,” and “Creation and maintenance of Generation CP Repository.” The objective of this project is to deploy and sustain the management of Web Services at Generation CP sites to support programme activities requiring rapid access to data generated by scientists.

24. Application of MOBY for GenerationCP Consortium

Principal Investigator:
Richard Bruskiewich, IRRI
Collaborating Scientists:
Natalia Martins, EMBRAPA
Mathieu Rouard, INIBAP
Shoshi Kikuchi, NIAS
Masaru Takeya, NIAS
Koji Doi, NIAS

EXECUTIVE SUMMARY
Web services are a key technology for the effective integration of a distributed network of GenerationCP tools and data sources. Year 1 of the GenerationCP Subprogramme 4 activities saw the successful pilot testing of web services technologies by three teams: IRRI, INIBAP/CIRAD, and IPGRI/EURISCO. Year 2 will extend GenerationCP web service integration. Parallel to the Subprogramme 4 commissioned task in year 2 to deploy basic MOBY-S software in GenerationCP institutions, there is a need for state-of-the-art

development and application of web services technology. This task will more firmly establish the role of both the technology and the GenerationCP in the bioinformatics community. Furthermore, it will support our role in the improvement of the technology itself.

Therefore, currently specified and constructed GenerationCP tools and datasets need to be made accessible using web services and included in appropriate registries. Based on gained experience, in this and other activities, GenerationCP should play a role in the further development of the pertinent web service technology. This task is geared to meet these objectives:
• Web services fully embedded in the GenerationCP Java reference platform1.
• Perl web services provider tool kit developed for partner sites using Perl.
• Web services API2 commissioned for a significant set of research applications on the GenerationCP High Performance Computing (HPC) facilities.
• Web services client embedded into EMBRAPA Genoma workbench.
• Web services provider API deployed for comparative gene catalog and expression data.

25. Creation and Maintenance of Templates for GCP Data Storage in Repositories

Principal Investigator:
Guy Davenport, CIMMYT
Participating Institutions:
Agropolis, CIMMYT, IITA, IPGRI, IRRI, and SCRI

EXECUTIVE SUMMARY
The Generation Challenge Programme (GCP) will generate a wide spectrum of data. Part of these data will go into public databases, part will be included in GCP or institutional databases that will be available as web services. However, there will always be data sets that do not or not yet fit in one of the available structures. Therefore a facility has to be maintained that allows access to these data sets in downloadable files in a consistent, but flexible format. The main objective of this task is to provide simple templates for the temporary storing or distributing of the different data sets that are being produced within SP1, SP2, and SP3, for which there is no current provision in public or institutional databases.

These templates must contain concise, but sufficient explanatory notes on how they should be used. The completed data sets should contain the necessary information to be stand-alone and should be simple enough to be understood. For example, enough description of material and methods used, and no use of acronyms or a coding system that only the data provider can understand. Where possible the templates should be consistent with the data models being produced under task ‘SP4-1 Development of GCP data models’ and will be used to store data in the central repository under task ‘SP4-5 Creation and maintenance of GCP Repository’.

26. Creation and Maintenance of Generation CP Repository

Principal Investigator:
Raj Sood, IPGRI/SGRP
Participating Institutions:
All Generation CP Consortium members
Agropolis, CIMMYT, IRRI, IITA, IPGRI/SGRP, ICRISAT, CIP

EXECUTIVE SUMMARY

An enormous amount of data is being generated by the partners in the Generation Challenge Programme. These data are stored and maintained at different locations, using different methods and standards. Most importantly, these data are not available for all partners to use. Acquiring, organising and publishing the information on the Web through a Generation CP repository is critical for the successful completion of the tasks that require data from various sources. The repository will also play a crucial role in supporting the Generation CP network by deploying supporting infrastructure for a CP-specific MOBY1 registry. Furthermore, the repository will store and provide a collaborative environment for the development of templates, standards, data models, XML schemas, and evaluation reports of the analysis of software packages and tools. The objectives of this project are to create a Generation CP repository for different datasets, data models, indexes, standards, XML schemas, evaluation reports from analysis of software packages, and to establish an online collaborative environment for the development of these components and their availability on the Web.

27. Integration of the High Performance Computing (HPC)-facilities in the Generation CP toolbox

Principal Investigator:
Anthony Collins, CIP
Participating Institutions:
CIP, IRRI, ICRISAT, IITA, CIRAD, ICARDA

EXECUTIVE SUMMARY

In 2004 High Performance Computing (HPC)-facilities have become available to the Generation CP. These must now be made fully available to the consortium members and possibly others. Therefore some technical steps will need to be taken, but also promotion of the possibilities, including documentation and training material, and a policy for access by the community needs to be developed. The objective is to consolidate use of HPC facilities on 3 sites involving all consortium members to the maximum level.

28. Improvement of Quality of Existing GCP Databases

Principal Investigator:
Graham McLaren, IRRI
Collaborating Scientists:
Guy Davenport, CIMMYT
Edwin Rojas, CIP
Fernando Rojas, CIAT
Chandra Subas, ICRISAT
Akinnola Akintunde, ICARDA
Visvanathan Mahalakshmi, IITA
Elizabeth Arnaud, INIBAP
Manuel Ruiz, CIRAD

EXECUTIVE SUMMARY

The GCP decision to integrate data across crops using Web Services allows partners to manage data locally provided they meet standards which conform to agreed data models and which guarantee the completeness and quality of GCP data. Hence the responsibility for

LIMS and platform development rests with the institutions collecting the data. Nevertheless the GCP recognises the need to support the development and implementation of standards, LIMS, databases,and curation tools to ensure timely and secure management of quality data from GCP activities.

This commissioned project will support the development of LIMS, the implementation of agreed data models in local databases, the definition of quality standards and quality assurance protocols, and the development of data curation tools. The emphasis of the GCP contribution will be on generic solutions and collaborative development so that outputs have wide application across the CP partners. Generic solutions are those that can readily be employed for different databases and crops.

29a. Creation of Institutional Bioinformatics Capacity (CIAT)

Principal Investigator:
Joe Tohme, CIAT
Co-Principal Investigator:
Mathias Lorieux, CIAT

EXECUTIVE SUMMARY
CIAT is part of a CGIAR consortium that aims to develop bioinformatics tools for genetic mapping, QTL analysis, and germplasm evaluation. A special focus has been given to the generation of databases to manage genetic mapping and genotyping, and to the storage of sequences. Molecular marker data are continuously updated for databases of bean, cassava, and the Genbank database at the NCBI.

A laboratory information management system (LIMS) is being developed (advanced phase) to manipulate the data of all the studies in the agrobiodiversity and biotechnology project. Currently, the available information generated by CIAT is scattered and not standardised (different operating systems, different hardware platforms, and different programming languages). Furthermore, the on-line bioinformatics databases (Galperin, 2004) are heterogeneous and not completely known (Gramene, TAIR, AcDB, NCBI), and the platforms to work at a high throughput level are not well implemented. Finally, the necessary tools for the storage, manipulation, analysis and sharing of the information is scarce or non-existent. For these reasons the development and application of bioinformatics resources is necessary to make an efficient and in-depth analysis of the information generated by CIAT.

Though we are in the process of developing tools for high-throughput analysis in the form of pipelines and data storage reservoirs (Galindo et al., 2004; Rojas et al., 2004), we are in need of trained personal in the area of bioinformatics to develop the implementation processes in a more efficient manner in order to transfer the knowledge and the developed tools to the users inside CIAT. Likewise, the generated capacities can be transferred to parallel projects created by the Challenge Programme.

29b. Creation of Institutional Bioinformatics Capacity (CIMMYT)

Principal Investigator:
Guy Davenport, CIMMYT

EXECUTIVE SUMMARY

Datasets produced by genomics projects at CIMMYT are not currently available in a fully queryable and integrated form. The datasets are either stored in stand-a-lone MS access databases, with limited querying facilities and no integration with other data, or in most cases are only available as text or excel files on each investigator's own computer. The major reason for this state is due to the lack of experienced bioinformatics staff within CIMMYT in the past. This has now been remedied by the appointment of a bioinformatics specialist in September 2004, which was made possible by the '2004 CGP task 28 – Identify informatics staff at germplasm characterisation sites'. The following objectives are therefore also a continuation of the work under this task:

1. Develop in consultation with investigators producing genomics and phenotyping datasets use cases and user requirements for data management, integration, and analysis each of the following types of data:
    i. Genotyping data
    ii. Genetic maps
    iii. QTL (including phenotypic data relating to QTL analysis)
    iv. Functional genomics, including transcription profiling using micro arrays and RT-PCR
    v. Laboratory information (as part of LIMS deployment)
    vi. Any addition genomic data
2. Develop in collaboration with the Software Development Department (SDD) and biometrics unit at CIMMYT the plans for design, installation/implementation, testing and curation of these systems (from objective 1).
3. Oversee the installation/implementation and testing of these systems described in objective 1. Where possible the installed/implemented systems at CIMMYT will be compatible with those developed within CGP.
4. In collaboration with other staff at CIMMYT develop user requirements and plans for integration of genomic data with other systems at CIMMYT, including but not exclusively
i. Genebank (passport data)
ii. Crop improvement (pedigree, phenotying)
iii. Environmental data
iv. GIS systems
5. Coordinate and lead CIMMYT activities in CGP SP4 and advise CIMMYT staff on bioinformatics requirements within the other GCP subprogrammes.
6. Advise the newly formed Crop-Related Information Systems Committee (CISCO) at CIMMYT on Bioinformatics and integration of genomic data.
7. Advise CIMMYT scientists on the uses and requirements for bioinformatics within their work, including commenting on any bioinformatics component with grant proposals from CIMMYT scientists.
8. Keep up to date with the latest developments in Bioinformatics and promote CIMMYT bioinformatics outside of CIMMYT.
9. Train CIMMYT, CGIAR and NARS staff on the use of bioinformatics systems installed/implemented at CIMMYT.

29c. Creation of Institutional Bioinformatics Capacity (CIP)

Principal Investigator:
Reinhard Simon, CIP

EXECUTIVE SUMMARY
The main output of this task at CIP will be the upgrading and expanding of existing bioinformatics skills. Upgrading of capacity will be realised through upgrading existing applications of interest to sub-programmes 1 to 3 and addition of new applications. Expansion of skills will be realised through training local staff and trainees.

There are three specific aspects:
- The implementation of use cases as specified by CIP scientists under the GCP and/or as identified and sub-contracted under specific activities with other commissioned research.
- the integration of the HPC in bioinformatics applications and workflows.
- the exploration of a modular architecture for desktop applications as a complementary tool for collaborative software development.

29d. Creation of Bio-informatics Capacity for Central and West Asia and North Africa

Principal Investigator:
Murari Singh, ICARDA

EXECUTIVE SUMMARY
ICARDA currently has limited manpower resources in this area, distributed in various programmes/units which undertake specific components of bioinformatics. Unlike many single-discipline oriented projects, genomic projects, due to their complex and integrated nature, need experts from a spectrum of disciplines to complete the projects. In coming years, we plan to explore ways to develop its manpower resources using training, conference participation, exposure to the new science tools, and recruiting a Bioinformatician to teams of biotechnologists, breeders and other plant scientists, biometrician and software developers, and information system specialists.

ICARDA has been aiming to develop Bioinformatics capacity at ICARDA and to promote the development and use of state-of-the-art bioinformatic tools for all crop improvement and biodiversity conservation programmes to support its Integrated Gene Management megaproject research activities as well as developing capacity building in the NARSs of CWANA

29e. Creation of Institutional Bioinformatics Capacity at ICRISAT

Principal Investigator:
Subhash Chandra, ICRISAT

EXECUTIVE SUMMARY
Using sound experimental protocols and appropriate methods and systems for analysing resulting and available genomics data lies at the heart of quality genomics research. To make this effectively happen, it is critically important that (a) genomics researchers are trained in properly using appropriate bioinformatic tools, (b) advice and support are readily available to them in choosing and using appropriate bioinformatic tools, and (c) appropriate bioinformatic systems and tools are developed, maintained, and regularly updated to facilitate proper management, analysis, and use of genomics data.

Effectively achieving these three goals demands building and maintaining a professional bioinformatics team. ICRISAT built this team in 2004 with funding from the GCP. With reduced funding available in 2005 and 2006 for this purpose, the size of this team needs to be optimally adjusted following an approach that could allow, as effectively as possible, meeting the bioinformatics needs of GCP genomics researchers at ICRISAT. A basic element of this approach is to place relatively more emphasis on training in 2005. This will allow genomics researchers to undertake at least their routine bioinformatic tasks themselves in coming years, maintaining in 2006 and in later years only the critical bioinformatic staff, and recruit as necessary additional staff from other sources to meet project-specific needs.

29f. Creation of Institutional Bioinformatics Capacity

Principal Investigators:
Guy Davenport, CIMMYT
Graham McLaren, IRRI
Joe Tohme, CIAT
Murari Singh, ICARDA
Simon Reinhard, CIP
Samy Gaiji, IPGRI
Subhash Chandra, ICRISAT
Visvanathan Mahalakshmi, IITA (BECA)
S.Hearne, IITA (BECA)

EXECUTIVE SUMMARY
The SP4 work plan of 2004 included a budget line supporting each of the eight listed institutions for the buildup of bioinformatics capacity, without specifying the outputs expected. This is not a desired mechanism for the GenerationCP. Since the GenerationCP funds many bioinformatics related activities, it can be expected that this capacity can be used in regular GenerationCP activities. Since abruptly stopping this funding might cause problems in terms of continuity, it is decided to reduce the amount made available by one third every year. This budget can be made available, provided that the institution can show that the funds will be used for building up bioinformatics capacity in the organisation.

Since it is already decided that in 2006 there will be a budget of 16.5 k$ available for each organisation, the proposal should also cover this period. The activities should be appropriate to achieve the objective and should include:

- Hiring additional staff, or funding existing staff, for bioinformatics activities related to the GenerationCP and
- Training staff in the field of bioinformatics.

29g. Creation of Institutional Bioinformatics Capacity (IPGRI)

Principal Investigator:
Samy Gaiji, IPGRI/SGRP

EXECUTIVE SUMMARY
The Generation CP Subprogramme 4 plan of work for 2004 included a budget allocation (US$50,000) to support each of the eight CGIAR Centres involved in the CP in building up bioinformatics expertise. The Generation CP has agreed to renew this support at a level of US$33,000 in 2005 and US$16,500 in 2006. The support is to target activities involved in

implementing the Subprogramme 4 workplan over the 2 years. Eligible activities include the recruitment of additional staff or support for additional staff time for bioinformatics activities as well as training of existing staff in the field of bioinformatics. IPGRI proposes to allocate the funds to support recruitment of a bioinformatics staff member at IPGRI-HQ to work on the Generation CP research activities led by IPGRI.

29h. Creation of Institutional Bioinformatics Capacity

Principal Investigator:
Morag Ferguson, IITA

EXECUTIVE SUMMARY
IITA has established a modest bio informatics facility to support the Central Biotechnology Laboratory in Ibadan and at Biosciences in East and Central Africa (BECA) in Nairobi the bio informatics facilities are accessible to the GCP staff. BECA has hired a full time bio informatican who can be consulted for the needs over there. They also have plans to hold a bio informatics training course in BECA similar to the MAS breeding workshop currently scheduled to held in Nov 2004. BECA also has good links with The Institute for Genomics Research (TIGR) both formally through projects and informally through contacts. We expect to benefit from this collaboration with BECA. However, the need to exchange data and information between the two sites is critical to ensure seamless integration of data coming from the two facilities. Genebank is located at Ibadan while genotyping and phenotyping are carried out in both locations. The need to establish standards and interoperability including ontology issues were dealt with in IITA in the first year of the GCP. Working in two locations needs both understanding the flow of information and standardisation of the data collection and curation. The budget this year would be to cover part of the salary of the bioinformatics staff and their training needs.

Ouputs:
- Quality and complete data on germplasm and genotyping from SP1 checked and shared with CP partners
- Support storage of raw data and help in linking to web services
- Analysis of genotyping, mapping and QTL data from SP1 and SP3
- EST data collation and analysis in SP2
- Support the curation of genotyping data at both Ibadan and Nairobi
- Link between SP4 and other SPs at IITA to feed back on special needs and also quality and design of templates as determined by other tasks in SP4
- Uploading GCP data on central repository from IITA
- Establishing workflow in the highthroughput genotyping
- Training technicians and other staff on data curation

30. Development of Decision Support Systems for Sampling Germplasm

Principal Investigator:
Xavier Perrier, Agropolis-Cirad
Participating Institutions:
CIRAD-Agropolis, IPGRI, WUR

EXECUTIVE SUMMARY

The general background for this task is clearly described in conclusions of the Data Analysis Workshop organised by SP 1 (Genetic diversity of global genetic resources) in June, 2004 (Zaragoza, Spain). Schematically, the demand could be summarised in the following steps:

- we will use molecular markers as tags for genes involved in agronomic traits (QTL detection, association mapping),
- thus, to detect relations between markers and genes, we will use germplasm collections where we expect that allelic diversity is maximised,
- however, as phenotyping assessment takes time and is expensive, we wish to concentrate our efforts only on a sub-sample but in minimising the detection power loss.

The proposed concepts and sampling strategies will be tested and validated on real data in collaboration with SP1 scientists. Population structures depend strongly on the considered species and a large range of cases will be used considering the dominant mating system (outcrossing, selfing, vegetative), recent or old domestication, narrow or large genetic base, recent drastic bottleneck, etc. We will use as privileged working data a collection of 205 sorghum accessions which presents several appealing features. A set of RFLP markers is already available (i) 69 markers scattered over the genome, (ii) 12 markers localised in a contig of 4 cM (600Kb). Microsatellite markers are or will be soon available (CPGeneration-SP1 project): (i) 30 scattered markers, (ii) 6 markers in the 4 CM contig. Scattered markers will be used to detect 'structural' LD and efficiency of sampling strategies will be tested on markers in the 4 CM contig. Moreover a current GABI-Génoplante project will produce data on SNP polymorphism in 8 to 10 genes of interest (grain quality and photoperiodism) which could be used for validation. The high-density region may also be used to validate the applicability of the LD fine-mapping approach that was described by Bink and Meuwissen (2004).
Other data are available at WUR and Cirad and for example, we plan to work on rice (diploid, homozygote, strong structure) and on cocoa (diploid, heterozygote, medium structure). Requests in IPGRI databases on biological characteristics and available data will provide some other cases of particular interest for validation and testing. IPGRI will facilitate collaborations in case of restricted access to these data sets.

Since several data sets are accessible from the participating institutions, it does not seem necessary to plan additional observations in the project.

31. Development of Ortholog-Function Display Tools

Principal Investigator:
Richard Bruskiewich, IRRI
Collaborating Scientists:
Kimmen Sjölander, University of California at Berkeley
Brigitte Courtois, CIRAD
Manuel Ruiz, CIRAD
Christophe Perin, CIRAD
Mathieu Conte, CIRAD
Masaru Takeya, NIAS
B. Jayashree, ICRISAT
Natalia Martins, EMBRAPA

EXECUTIVE SUMMARY

Comparative biology across multiple crop species is a key strategy in the GCP for the identification of stress-responsive gene loci and their corresponding alleles of high agronomic value for application in breeding for stress tolerance. Critical to the task of comparative biology is the elucidation of evolutionarily conserved gene orthology relationships across species and related paralogy relationships within a gene family. Such orthologous and paralogous gene loci almost invariably share common gene functions, thus important inferences of comparative biology may be possible once such relationships are clearly defined.

This GCP Subprogramme 4 commissioned research project therefore focuses on the task of compiling a catalog of orthologous and paralogous plant genes for GCP target crops and related model plant species. This task also includes the development of useful (graphical) user interfaces to the catalog and provisions for (web services) integration with other GCP project data and tools, such as comparative gene expression experiments.

32. Development of Crop gene Expression Database and Data Mining Tools

Principal Investigator:
Shoshi Kikuchi, NIAS
Collaborating Scientists:
Richard Bruskiewich, IRRI
Hei Leung, IRRI

EXECUTIVE SUMMARY

During the completion of genome sequence analysis in Arabidopsis and rice, many kinds of technologies for comprehensive analysis of gene expression have been developed. Those are microarray, SAGE (Serial Analysis of Gene Expression), MPSS (Massively Parallel Signature Sequencing), etc. Gene expression data using such technologies have been accumulated and many databases have been opened. The problem is that the databases have been independently constructed and, for the researchers, current databases are not easy to access and to obtain the data they want.

Shoshi Kikuchi's team in NIAS has been contributing to the establishment of rice microarray systems since 1999. Starting with 1265 cDNA-based microarray system, 8987 cDNA-based microarray system was finally established in the Rice Microarray Project in Japan (1999~2003). About sixty research groups have joined in the Rice Microarray Project and almost all microarray experiments were performed in the Rice Microarray Centre established in NIAS at that time. All gene expression data using 8987 microarray system were deposited and accumulated in the Rice Expression Database. Probed cDNA clones originated from the large scale EST collection of Rice Genome Project (1991~1997). The total gene number of rice has been estimated at 40,000~50,000 and the unique set of EST collection corresponds to 11,000. This means the coverage is only a quarter of the whole expressed genes.

The goals of this project include:

- Establishment of the user-friendly gene expression database of the crops
- Connecting data among crops by the linkage of orthologous genes
- Making accessible all the data are through WEB service

33. Development of an Integrated Decision Support System for Marker-assisted Plant Breeding

Principal Investigator:
Subhash Chandra, ICRISAT
Collaborating Scientists:
J-M Ribaut, CIMMYT
CG McLaren, IRRI
AE Melchinger, University of Hohenheim, Germany
FA van Eeuwijk, Wageningen University
H Mohanty, University of Hyderabad, India
JH Crouch, CIMMYT

EXECUTIVE SUMMARY
Marker-assisted selection (MAS) and marker-accelerated breeding (MAB) require reliable identification and application of simply inherited markers that are in close proximity to genetic factors affecting simple, oligogenic, and polygenic traits of interest to stakeholders. This can proceed through a mapping-population-based or an association mapping approach, the latter typically based on a sample of unrelated individuals. The mapping-population-based approaches, with or without a linkage map, remain the most popular in plants. As a result, these approaches are relatively well-developed and practical nuances of their application are well understood. The use of association mapping in plants, on the other hand, has only recently started – for example in maize and rice. Association mapping is likely to provide a reduced level of redundancy during the transition from marker identification to application in diverse breeding populations and germplasm. It may be a particularly useful route to follow where the creation of a large enough mapping population and a good linkage map is not feasible due to constraints in the development of these populations or non-availability of sufficient polymorphic markers.

In either approach, the journey from the phenotyping-and-genotyping of individuals to the identification and application of markers in molecular breeding involves the sequential use of a number of different software – each with its own input data format requirement and many alternative choices for data analysis. The use of appropriate experimental design and data analysis is a critical component for successful development and application of marker-assisted breeding systems. Making these choices rightly is a highly specialised function. There is a lack of proper and simple-to-use guidelines for non-specialists, which makes it difficult for them to confidently choose the appropriate design and analysis methods offered by various software. Having a centralised and evolving resource offering biometric inputs required for molecular breeding would be a tremendously valuable asset to the research and breeding community. In this project, we propose to provide simple-to-follow guidelines embedded into the front-end of the system to help users choose the most appropriate experimental design and data analysis methods, and provide them with a regularly updated selection of the currently most appropriate options. Converting the format of datasets to meet the differing requirements of various software is also an extremely time-consuming process fraught with errors. The system developed in this project will provide automatic transition for data flow between all permutations and combinations of software to be used. This integrated molecular breeding analysis platform (analogous to AgroBase) will be designed to facilitate an integrated, error-free and appropriate data analysis from the beginning to end of the molecular breeding pathway.

We believe that such an integrated system, based on freely available quality software, will be of profound value to scientists in marker research laboratories and breeders in national programmes and small breeding companies across the developing world. With these end users in mind, the core of the platform will be based on freely available software with the provision of simple-to-use on-line decision guidelines to help users make the right choices of software, experimental design and data analysis for any given task. It is proposed to call this integrated system iMAS. To ensure delivery of the system in a realistic timeframe within the available resources, focus will be on crops and marker identification and application systems for which adequate subject matter knowledge, methodology, experience and free software are already available. However, iMAS will have a built-in flexibility to incorporate new, relevant, freely available software as they become available. It is envisaged that the availability of this system will foster further necessary developments of freely available tools to be integrated into it in future. Also, it is hoped that other parallel GCP competitive projects will support the development of solutions to any fundamental gaps in computational support for molecular breeding for which there are no alternative suppliers of freely available software.

34. GenerationCP Use Case and Software Engineering Collaboration and Management

Principal Investigator:
Thomas Metz, IRRI

EXECUTIVE SUMMARY
To facilitate the design of MOBY interfaces to GenerationCP data and tools, and to allow optimal interactions with the 'rest of the world' in terms of access and exchange of SP4 project information, the effective specification and management of informatics "use cases" is critical.

The project software engineering activities need to be coordinated and the resulting software elements need to be integrated. In addition to face-to-face group coordination of software developers, a virtual "groupware" infrastructure specifically suited for collaborative software development and team management is required. This task is designed to meet these needs. The task is strongly SP4 software engineering process (not informatics product) oriented, providing tools and support that allow geographically distributed teams to collaborate efficiently and effectively in the area of content and software development. The functionality of these collaborative tools will complement the GenerationCP virtual workspace and could be integrated as required.

The task goals include: to commission suitable "groupware" infrastructure, provide training and support for its use, and coordinate face-to-face project meetings for efficient GenerationCP SP4 software and use case development and management in 2005.

SP5 COMMISSIONED GRANTS

CB1. Training Programme on Genetic Diversity Analysis of Germplasm

Participating Institutions:
CIRAD

IRD

EXECUTIVE SUMMARY
This project includes a two week course curriculum in the field of genetic diversity analysis. For each topic, we will elaborate a support note in French to accompany the slides in Powerpoint format. Each slide is a succinct summary of a topic used as a support for trainers and a guide for trainees. This work will be done by Cirad and IRD scientists with the help of a young scientist with a computer specialisation and a scientific archivist.

CB2. Development of Training Materials for a Course in Genomics and Comparative Genomics, and Design of Course Curriculum

Principal Investigator:
Theresa M. Fulton, Cornell University

EXECUTIVE SUMMARY
Comparative genomics is a relatively new field of research that has developed due to increasing technological capabilities, particularly in the area of DNA sequencing, and an increased understanding of the conservation of biology. The potential benefits of this new field are extensive, not only on the academic level of an increased understanding of biological organisms, but also including many practical applications. In the field of plant breeding and crop improvement the promise of comparative genomics holds special potency as it has implications on global hunger, nutrition, conservation, and environmental protection. Information that once seemed particular to only one species or limited group of species can now often be useful across many species, genera, or even further. For example, genes involved in disease resistances often employ similar processes and biochemical pathways across many crops, and are often even located in similar chromosomal positions (R. Nelson, pers. comm.).

One result of comparative genomics has been the increasing speed at which new data, information, discoveries and technologies are now developing. This is exciting and promises many benefits; however, it also means that it has become increasingly difficult to keep abreast of current research and technologies. This is particularly true for researchers in developing countries who have limited access to scientific publications or extended academic communities. Thus, the very researchers who could most benefit from comparative genomics are those least able to, due at least in part to the lack of know-how (together with infrastructure insufficiencies, etc.).

This project aims to begin to solve this problem by creating materials to be used in training global scientists in the use of genomics and comparative genomics. A recent collaboration of IPGRI and the Institute of Genomic Diversity at Cornell University resulted in 2 learning modules, “Volume 1: Using Molecular Marker Technology in Studies on Plant Genetic Diversity” and “Volume 2: Genetic Diversity Analysis with Molecular Marker Data.” These learning modules (available as a CD in several languages and downloadable as pdf) were used as the basis of a molecular marker training workshop in Venezuela in 2003, and more than 300 copies have been requested by researchers worldwide. Similar training materials covering genomics and comparative genomics are now in great demand to follow and complement the concepts covered in the previous learning modules.

The goal of this project is to develop training materials and a curriculum for a course in genomics and comparative genomics. These materials will be developed in such a way as to be useful either as a self-tutorial, or as the basis for a course of approximately 2 weeks duration. They will include definitions of terms, illustrations of concepts, photographs, real-life examples, appropriate applications, lists of key references, and other items as appropriate.

CB4. Development of Training Materials for a Course in Bioinformatics and Design of Course Curriculum

Principal Investigator:
Richard Bruskiewich, IRRI
Collaborating Scientists:
Tin Wee Tan, National University of Singapore/APBioNet
Theresa Fulton, Cornell University
Marja Thijssen, Wageningen University
Natalia Martins, EMBRAPA
Francis Moonan, IITA

EXECUTIVE SUMMARY
This project will strive to assemble or develop comprehensive and educational training materials for bioinformatics, intended for scientists with a background in germplasm, biology, and genetics, together with a working knowledge of their application to plant and agricultural sciences. The materials described here will be designed as an educational resource or reference tool that can be used as a self-tutorial or as a course curriculum appropriate for a two-week course incorporating key references, photographs, illustrations of concepts, examples, and suggested applications. The course materials will be meant as a guide to plant genome informatics and bioinformatics useful for modern plant breeding, genomics and comparative genomics, data analysis, and interpretation. The materials will serve as the basis for training and capacity building of scientists attending the courses organised within the Training Programme of the Generation Challenge Programme (GCP). Later on they will be made available for the GCP partners and scientists at large through a repository of training materials so that as many people as possible can benefit from it.

Also, the training materials will be designed to facilitate translation into other languages, in particular, French, Spanish, and Portuguese. Ownership of the IP generated by the project shall be governed by the GCP consortium agreement dated 10 August 2004.

Project objectives include:
• Develop a detailed course curriculum for crop bioinformatics scoped for 2 week delivery and packaged as a web based “distance education” course, structured in a modular fashion with incremental lessons cross-referenced to other publicly available training materials.
• Develop a course operational plan following the distance education model of the S* Alliance (www.s-star.org), specifying the design of various internet resources to be used for such a course including the course website providing access to the online learning environment, a discussion forum, chat facilities, and email integrated for communications between students and course facilitator, and between students.

CB8. Functional Genomics to Improve African Crops

Principal Investigator:
Roeland CHJ van Ham, Wageningen University

EXECUTIVE SUMMARY
Most research centres in Africa have access to the Internet. However, most African scientists have not been exposed to the opportunities offered by plant genomic sequence database tools freely available through the Internet. The visit of Dr.Vorst from a premier plant science institute in the Netherlands to Africa provides a unique opportunity to build capacity in genomic data mining, especially at a time when an annotated version of the rice genome (a model for cereals) is expected to be released publicly.

The visit would fit perfectly into the goals of the GCP as it would help to establish a firm basis for the GCP to expand its activities in this part of the continent. Furthermore, it strengthens the collaborative links between WUR in the Netherlands, ACGT/University of Pretoria in South Africa, CIMMYT in Africa, and CG institutions in Nairobi, Kenya. The University of Pretoria and the CSIR have applied, through the ACGT, to become Consortium Members of the GCP, and if this application is approved, the visit would be well timed to facilitate their future involvement.

CB9. Use of Molecular Markers for Mining Useful Allelic Diversity – A Summary of SP1 Genotyping for Germplasm Scientists

Principal Investigator:
Carmen de Vicente, IPGRI

EXECUTIVE SUMMARY
Subprogramme 1 is charged with the exploration of the genetic diversity of global germplasm collections. This exploration should help identify variants of genes involved in complex traits, and as such will be the basis for the research activities of the other subprogrammes. In year 1, eleven* crop teams built ‘composite sets’ representing the range of diversity of each crop species and its wild relatives. These sets were then genotyped with anonymous microsatellite markers. As a result of this process, it should be possible to document the structure of each crop set and select a "reference sample" suitable for advanced characterisation, which in the end will help identify useful traits in collections and the genes that control them.

This activity proposes a workshop based on genotyping results of the year 1 SP1 work plan including participation of germplasm managers/curators from the CG and NARS, in and outside the Consortium. The detailed content and agenda of the workshop will be defined in concert with Jean Christophe Glaszmann, SP1 leader. It is expected that the workshop will deal with subjects such as practical as sampling of germplasm, standardisation of protocols between different laboratories, analysis and interpretation of results, among others. In general the content will refer to the lessons learned about structural characterisation, but also will tackle issues such as re-sampling for fine genotyping and further functional characterisation and approaches for finding useful genes in well-phenotyped and genotyped samples.

In addition, time will be allocated for constructive discussions between GCP engaged scientists and the rest of germplasm scientists on the broad usefulness (impact) structural characterisation with molecular markers may have for the exploitation of germplasm collections.

The overall objective of the workshop is to convey the knowledge gained through the SP1 activities in year 1 to CG and non-CG scientists involved with germplasm management so that results can benefit a wider community outside the GCP. By inviting this array of participants to the workshop, SP5 also wants to ensure these scientists develop a sense of 'ownership' and become practitioners in their own institutions, expanding the impact of the GCP.
* Barley, Cassava, Common bean, Cowpea, Chickpea, Maize, Musa, Potato, Sorghum, Rice, Wheat.

CB10. GR Policies: A Workshop Session in China Devoted to Genetic Resource Policies

Principal Investigator:
Niels Louwaars, Wageningen UR

EXECUTIVE SUMMARY
The GCP is collaboration among researchers, breeders, and social scientists from a wide variety of institutions, serving an equally broad group of potential users. Much of the success of the programme depends upon having access to materials and technologies under clear agreements. In some cases individual institutes are free to develop guidelines for access (such as in the case of protected inventions) and in some cases national authorities have an important say (as in international transfer of genetic resources). In some cases, regulations are not very clear. Nor is it always clear which Authority can provide access. The latter is in some countries unclear among ministries of agriculture and environment.

In addition, a number of countries now have to include provisions of the International Treaty in their regulatory systems, creating in some cases some form of confusion in the meantime. The proposed GCP Conference that will be organised in China in April 2005 is considered a perfect venue to put emphasis on this important topic and to share experiences among policy makers and managers from different GCP Consortium member countries.

CB11. CIMMYT Plant Genetic Diversity and Molecular Marker Assisted Breeding: A Training Course

Principal Investigator:
Marilyn Warburton, CIMMYT

EXECUTIVE SUMMARY
The GCP Training Programme (Subprogramme 5) will offer a series of regional courses (Africa, Asia, and Latin America) focusing on the main subjects of each of the Subprogrammes. Subprogramme 3, Trait Capture for Crop Improvement, has as its main goal the increased efficiency of crop breeding, particularly using new tools such as biotechnology. The uptake of new tools depends on the ability of national partner scientists to use the new technologies, which may include the need for capacity building in some cases. Therefore, the training courses offered in the regions for Subprogramme 3, entitled "Plant Genetic Diversity and Molecular Marker Assisted Breeding," is geared towards National Programme scientists with the desire and possibilities to utilise markers (via diversity analyses and Marker Assisted Selection) in their breeding programmes. The course proposed here will be offered in conjunction with a National Programme in the region (Latin America) who has a

demonstrated track record in both training and utilisation of biotechnology in crop improvement.

The Instituto Nacional de Investigaciones Agropecuaria (INIA) in Chile is the country's national programme for investigations in agriculture and fisheries, funded by the Ministry of Agriculture. INIA has been in operation for 40 years and has regional offices in all states of the country. The main office is in the capital, Santiago, where they have extensive laboratories in many disciplines, including biotech. INIA's mission is to "generate, adapt, and transfer technologies" to the agriculture sector in Chile, which is completely compatible with what the Generation Challenge Programme is working on worldwide. Located in Santiago, the La Platina Experiment Station is convenient for travel to Chile, and has good accommodation facilities nearby. Laboratories and classrooms are available for the workshop. In addition to INIA staff, several nearby universities could provide resource people for the classes. Resource staff from CIMMYT, CIAT, and CIP (the Latin American CG centres) will be called upon as necessary to complement the resource people from Universities and NARs from Chile and other countries in Latin America. The identification of resource people has begun already, and should be finished by July 1, 2005.

The objective of this proposed project is to organise a training course in Chile on "Plant Genetic Diversity and Molecular Marker Assisted Breeding." The course is tentatively planned for a 3-week period to be determined in October, 2005.

CB11. Cornell University Plant Genetic Diversity and Molecular Marker Assisted Breeding: A Training Course

Principal Investigator:
Theresa M. Fulton, Cornell University

EXECUTIVE SUMMARY

The CGIAR centres have long been hubs of training in molecular marker and marker-assisted breeding techniques for their respective regions. However, in many regions of the world, the nearest CG centre is not easily accessible by travel, rendering it difficult for researchers in a region to go to nearest one for training or collaborations. Thus, institutions in the national agriculture research system (NARS) are becoming increasingly important to help fill this need in many regions. While many NARS are quickly advancing in technologies and infrastructure, as yet most of these institutions have not been commonly used as centres for training.

Since its change to a democratic, non-apartheid government, South Africa has made rapid progress in education and scientific research, specifically in the training of African scientists. The University of Pretoria, in particular, is becoming widely known as centre of excellence of higher learning. The university recently took another step in this direction by developing FABI, the Forestry and Agriculture Biotechnology Institute. FABI is conveniently located near Johannesburg International Airport and has various accommodation facilities nearby. With the nearest CG centre more than 1,700 miles away (in Nairobi, Kenya), FABI has the potential to become an extremely useful location for training and collaboration in agricultural research for the large Southern African region that is otherwise not served. With strong connections to and shared facilities with the University, in addition to well-equipped laboratories, FABI has the potential to become a true "centre of reference" for the entire Southern African region.

Researchers at FABI are already involved in the training of Sub-Saharan African scientists and are financially supported by UP, National Research Foundation (South Africa), the local forest industry, and several international research organisations. FABI presently hosts 16 Research Fellows as well as 41 PhD and 45 Masters students from 10 different African countries. Facilities at FABI include all basic equipment required for the execution of molecular work, including PCR facilities, automatic DNA sequencers, a Microarray and Real-time PCR facility, and a Odessy infra-red Scanner connected to a LI-COR Genotyper. In June 2004, a Bioinformatics facility was established at FABI and in 2005 a plant transformation unit will be also established. Senior scientists from FABI have all been evaluated by the National Research Foundation rating programme, and they extensively collaborate with many Research Institutions worldwide. This includes collaboration with US, European and Asian scientists.

The objective of this proposed project is to organise a training course in S. Africa on "Plant Genetic Diversity and Molecular Marker Assisted Breeding." The course is tentatively planned for a 2-week period of May 22 – June 3, 2005. This coincides with the end of the semester at the University of Pretoria at May 24 (thus freeing up schedules of both possible resource persons and lab and conference rooms); no holidays or conflicting conferences have been noted as of yet.

CB11. ICRISAT Plant Genetic Diversity Analysis and Marker-Assisted Breeding

Principal Investigator:
Subhash Chandra, ICRISAT
Collaborating Scientists:
Morakot Tanticharoen, BIOTEC, Thailand
Theerayut Toojinda, BIOTEC,Thailand
Julapark Chunwongse, BIOTEC,Thailand

EXECUTIVE SUMMARY
The objectives of this course will be to:
• Provide practical hands-on training in relevant state-of-the-art genetic diversity analysis, linkage mapping, QTL analysis, and association mapping to facilitate marker-assisted breeding,
• Establish mechanisms for technical backstopping and trouble-shooting,
• Empower practitioners in the Asia region to use GCP knowledge, services, and products.

The course duration will be two weeks (12 days). It will be held sometime in August 2005 in Thailand in collaboration with BIOTEC and Kasetsart University. The exact dates will be finalised in consultation with Drs Tanticharoen, Theerayut and Julapark.

CB12. Writing Quality Project Proposals that Connect Agricultural Scientists, Stakeholders, and Donors

Principal Investigator:
Karine Malgrand, IPGRI

EXECUTIVE SUMMARY

The aim of this project is to increase the capacity of the Generation Challenge Programme's (GCP) African, Latin American, and Asian partners (NARS) in proposal development that will lead to more effective distribution of research outputs and results and an increased fund raising ability. This will enable them to become stronger partners in alliances with the GCP consortium members in the implementation of our common agenda.

Many donors have sought to infuse more competition into the market for agricultural research for development (ARD) and are moving from unrestricted grants to competitive grants and increasingly relying on evaluations of research and delivery performance to determine future funding levels. National and international public institutions and individuals carrying out ARD are competing for these competitive grants. The GCP realises that universities and partners from the North submit more successful proposals than the South because they have the capacity to submit well-written and well-presented excellent scientific proposals that incorporate stakeholder input and well-defined impact pathways.

Since National Agricultural Research Systems (NARS) have an important role to play in planning and implementing the research agenda of the GCP. One objective of the 'capacity building strategy of the GCP' is to empower NARS from Africa, Latin America, and Asia, as well as less experienced GCP consortium member scientists, in writing and submitting quality research proposals to the GCP and to other donors: In addition to having a quality scientific project design, their project proposals would also be well presented.

Furthermore, as outlined in the impact initiative of the GCP SP5, scientists must consult with farmers and other stakeholders at the planning stage of a proposal to ensure that their research objectives correspond to the needs of their beneficiaries. The proposed training courses on Project Proposal Development will therefore include a session on participatory planning to ensure that project proposals would include an effective communication strategy for an effective delivery chain process.

CB13. The Institute for Genomic Diversity's Interactive Resource Centre

Principal Investigator:
Theresa Fulton, Cornell University

EXECUTIVE SUMMARY

One of the many challenges for scientists in the international community, particularly in developing countries, is getting personalised, interactive support in addressing the needs specific to their own research programmes. The lack of access to trained personnel, key literature, and a support system can be limiting factors in the progress of scientists' research programmes. At the Genetic Resources Challenge Programme's Stakeholders Meeting in Alexandria, January 2003, one of the recommendations of the Working Group on Capacity Building was to create an Interactive Resource Centre (a "helpdesk") to support scientists involved in the Challenge Programme.

The IGD is in a key position to take the initiative in setting up this Resource Centre. Our location at Cornell University gives us access to important resources, including facilities and faculty, that could play an important role in addressing the needs of scientists all over the globe. Nearby facilities include many large libraries (and networked resources), information

technologies such as the Cornell Theory Centre and the new Computational Biology Support Centre. These resources, together with IGD's combined expertise in research and education, make it possible for our Institute to have a significant positive impact, particularly as a capacity-building contribution to the Challenge Programme. Many international scientists have taken advantage of these resources while on training visits to the IGD; establishing a Resource Centre will allow us to extend our resources to scientists that do not have to opportunity to visit in person.

Targeted for, but not limited to, the international CGIAR centres, National Agricultural Research centres and their clients, users of the Resource Centre would include scientists around the world working on plant genetic resources and needing support. Support given could be in the form of answering questions about experimental design, laboratory protocols, data management, laboratory maintenance, funding, training possibilities, software specifics, statistics questions, making literature available, etc. The goal is not necessarily to directly answer each question, but to be able to direct the client to the appropriate person or place that will be able to address the question.

CB14. Regional PGR Courses

Principal Investigator:
Marja Thijssen, Wageningen UR

EXECUTIVE SUMMARY
This project creates a curriculum for regional courses on institutional genetic resource policies in order to create awareness, extend relevant knowledge, and share experiences among scientists and science managers, which will allow them to develop or strengthen institutional policies and tools to handle Freedom to Operate on IPR and ABS in partner institutions of the CP.

The GCP is a collaboration among researchers, breeders, and social scientists from a wide variety of institutions, serving an equally broad group of potential users. Much of the success of the programme depends upon having access to materials and technologies under clear agreements. In some cases individual institutes are free to develop guidelines for access (such as in the case of protected inventions) and in some cases national authorities have an important say (as in international transfer of genetic resources). Provisions of international agreements such as CBD, IT-PGRFA, and TRIPs are translated into national law, which in turn affect freedom to operate in research institutions and in relation to the products developed by these institutions.

Awareness is relatively low in circles of scientists and in some cases science managers as well. The proposed curriculum will be a tool to increase awareness and knowledge in various regions. This initiative is immediately linked with the priority setting workshop in Costa Rica.

CB15. Distant Policies: A Distance Learning Module for Scientists on Genetic Resource Policies and Their Implications for Freedom-to-Operate

Principal Investigator:
Niels Louwaars, Wageningen UR

EXECUTIVE SUMMARY
The purpose of this project is to provide a basic and practical distance-learning module to help GCP scientists understand the importance of limitations to the rights to use plant genetic resources and tools, methods, and products protected by intellectual property rights.

The course introduces scientists from GCP consortium members, partners, and stakeholders on topics concerned with access, benefit sharing, and intellectual property matters at a rather basic level, but in such detail that they are well-prepared to understand the impact of the relevant rights systems on their projects and on the implications of the use of materials and research tools for the availability of the output of their work for target intermediate and end-users. The course will include a test to verify that the course participants have sufficiently understood the issues. It may include a possibility to follow the course at two levels.

The GCP is a collaboration among researchers, breeders, and social scientists from a wide variety of institutions, serving an equally broad group of potential users. Much of the success of the programme depends upon a clear freedom to operate on the materials, tools, and methods. Scientists commonly are not sufficiently concerned with the legal implications of the use in their research of materials and tools that are developed by others. This is particularly true for scientists in institutions that primarily produce public goods.

CB16. IP Matters: An Intellectual property and Access & Benefit-Sharing Helpdesk and On-line Resource for the GCP Community, Partners and Stakeholders

Principal Investigator:
Victoria Henson-Apollonio, IPGRI

EXECUTIVE SUMMARY
The objective of this project is to provide a practical on-line service desk for assistance, clearing-house activity, and feedback to GCP consortium members, partners, and stakeholders on topics concerned with intellectual property matters in the broadest sense. This project will field questions related to GCP agreement requirements, access issues, asset identification, and product development and delivery that will be responded to by a network of IP and ABS professionals. In addition, information that is relevant to IP management, ABS experiences, and IPRs will be provided as original material and in links to other on-line IP sites.

The GCP is a collaboration among researchers, breeders, and social scientists from a wide variety of institutions, serving an equally broad group of potential users. Much of the success of the programme depends upon having agreements, plans, and other tools in place that encourage communication among scientists, provision of the necessary materials (including germplasm), participation of end-users in project design and implementation, as well as effective product development, distribution, and uptake. Negotiating agreements, access to materials and information, obtaining the necessary licenses/permissions to use 3rd party technologies, etc, can be daunting tasks, even for experienced administrators or technology transfer personnel at advanced research institutions. We believe that providing a facility to help GCP participants with Freedom-to-Operate management would be very useful and timely. This service desk would assist GCP institutions, scientists, and staff with fulfilling GCP IP management requirements, complying with national legislation, answering questions

related to IP issues, while also providing additional information on IP and ABS topics that may be of further interest to the GCP community. The “IP Matters” facility would be modeled, in part, on the “IPR-Helpdesk,” available at the URL: http://www.ipr-helpdesk.org/index.htm.1. In addition, this facility would look to the SP5 GCP-sponsored technical (scientific) helpdesk, hosted by Cornell University, located at URL: http://irc.igd.cornell.edu/, for materials and methods.

CB17. Reporting for Product Distribution: An Asset Inventory System for the Generation Challenge Programme

Principal Investigator:
Victoria Henson-Apollonio, IPGRI

EXECUTIVE SUMMARY

The purpose of this project is to develop a practical on-line service that will facilitate the production of a dynamic inventory of 3rd party materials used by GCP scientists and products produced by GCP research. The GCP is a collaboration among researchers, breeders, and social scientists from a wide variety of institutions, serving an equally broad group of potential users. Much of the success of the programme depends upon having GCP assets and products used and taken up by as broad a user-base as possible. This requires that an inventory of GCP products be created and also that 3rd party materials used in the production of these assets be reported as well.

CAS-IP has developed several reporting tools over the past several years that can be used as the basis for developing this GCP reporting system. In addition, a CAS-intern is currently working on a project to make this reporting system more efficient, informative, and easier for the scientist-originators to use.